UNIVERSIDADE DE LISBOA
INSTITUTO SUPERIOR TÉCNICO

LONG-TERM UNCERTAINTY IN ELECTRICITY MARKET MODELLING, POLICY, AND PLANNING

Ian James Scott

Supervisor: Doctor Carlos Augusto Santos Silva

Co-Supervisor: Doctor Pedro Manuel Santos de Carvalho

Thesis approved in public session to obtain the PhD Degree in

Sustainable Energy Systems

Jury final classification:

Pass with Distinction and Honour

2020

UNIVERSIDADE DE LISBOA
INSTITUTO SUPERIOR TÉCNICO

LONG-TERM UNCERTAINTY IN ELECTRICITY MARKET MODELLING, POLICY, AND PLANNING

Ian James Scott

Supervisor: Doctor Carlos Augusto Santos Silva

Co-Supervisor: Doctor Pedro Manuel Santos de Carvalho

Thesis approved in public session to obtain the PhD Degree in

Sustainable Energy Systems

Jury final classification: **Pass with Distinction and Honour**

Jury

Chairperson:

Doctor João Miguel da Costa Sousa, Instituto Superior Técnico, Universidade de Lisboa

Members of the Committee

Doctor Magnus Korpås, Faculty of Information Technology and Electrical Engineering, NTNU-Norwegian University of Science and Technology, Noruega;

Doctor Carlos Alberto Henggeler de Carvalho Antunes, Faculdade de Ciências e Tecnologia, Universidade de Coimbra;

Doctor Manuel António Cerqueira da Costa Matos, Faculdade de Engenharia , Universidade do Porto;

Doctor Luís António Fialho Marcelino Ferreira, Instituto Superior Técnico, Universidade de Lisboa;

Doctor Pedro Manuel Santos de Carvalho, Instituto Superior Técnico, Universidade de Lisboa.

Funding Institutions:

Fundação para a Ciência e a Tecnologia (FCT).

2020

Acknowledgements

I have been lucky with my supervisors and collaborators who I am glad to thank. Firstly, I would like to thank Prof. Pedro Carvalho for the support, ideas, dedication, flexibility, and encouragement that contributed to every part of this work. I would like to also thank Prof. Carlos Silva for the encouragement and flexibility shown with me throughout this journey as well as his continuous support in allowing me to pursue my diverse interests. Next, I would like to thank Dr. Audun Botterud for the encouragement, ideas shared, detailed and diligent input, and hosting me in such a friendly and supportive manner at the MIT, an experience from which I gained a lot. I would like to thank my collaborators Felix Diawuo and Diana Neves for the interesting research topics that kept me distracted and interested and the friendship provided along the way. I would also like to thank Patricia Baptista for the advice and input provided to Felix and myself. I would like to thank my InnoEnergy Lisboa colleagues Sonia, Kaisa, Julia, and Eduardo for the support, distraction, and friendship as well as the many others from the MIT Portugal Program, IN+, and InnoEnergy.

Perhaps most importantly, I wish to thank my partner, Filipa de Almeida, without whom I surely would not have begun, let alone completed, this thesis. Her love and support in this endeavour, as in all things, makes anything possible. I also cannot thank enough my parents for their support in all forms on which I have always been so fortunate to be able to rely on. I am forever grateful for the love, sacrifice, guidance, proof-reading, and upbringing they provided me that has allowed me to chase my interests around the world, any and all achievements are always entirely due to them.

This work is supported by the Fundação para a Ciência e Tecnologia (FCT) through the

MIT-Portugal Program and the FCT scholarship PD/BD/113715/2015 as well as the InnoEnergy PhD school at the European Institute of Innovation Technology (EIT). I am truly grateful for this support which created the opportunity for this thesis.

Abstract

Uncertainty is an increasingly important aspect of decision-making relating to the electricity systems of the future. Over the long-term time horizons required for investment decisions and government policy making, history indicates that forecasts tend to be varied and uncertain. Hence, long-term forecast uncertainty should be an important aspect of any electricity market modelling or planning exercise. However, surprisingly, we do not find this to be the case in the literature.

This thesis contributes to the study of the incorporation of long-term uncertainty into the generation expansion planning class of decision-making models in a number of ways. Firstly, in order to represent a wider range of long-term uncertainties into the generation expansion planning model this thesis first investigates one of the more promising possibilities for reducing model complexity, the representation of time. I examine a number of different approaches to selecting representative days for generation expansion planning models, focusing on ensuring future power system challenges are represented and that the role of energy storage can be accurately assessed. A methodology for adjusting the weighting derived from common representative day clustering algorithms is proposed for use in generation expansion planning models that ensures the targeted level of net demand is captured in the model without altering the underlying net demand shapes that define ramping challenges. The results demonstrate the importance of carefully performing the clustering of representative days with the model selected expansion plans differing greatly in terms of both the total installed capacity and technology choice.

The thesis then investigates the role of uncertainty in the important system planning

question of quantifying decarbonization costs. I focus on the Ghanaian system to provide a benchmark for developing countries and provide insight into the relatively under-studied sub-Saharan region. To do so a generation expansion planning model is modified to incorporate the reality of fuel shortages and fuel switching typical of a developing country's power system. From this modelling, a range of emission reduction costs are generated that provide important benchmarks and I identify drivers of these costs specific to developing countries. The results demonstrate that discount rates, representing Ghana's access to capital, are a particularly important variable for developing countries. Lower discount rates can lead to more investment in capital intensive renewable energy in the long run but can also lock-in additional conventional generation investment in the short term.

The thesis then turns to the investigation of the importance of representing a wide range of economic and physical sources of uncertainty into the modelling of the electricity system and focuses on the method with which uncertainty is incorporated, both for investment decision making and policy analysis. The results of a United Kingdom case study demonstrate the importance of combining uncertainty across different inputs, finding that the difference between a deterministic and stochastic solution increases non-linearly when uncertainty inputs are combined. Further, it is demonstrated that combining uncertainty sources by adding a limited number of scenarios to multiple sources of uncertainty outperforms adding additional scenarios to any individual source of uncertainty. Finally, the representation of uncertainty as individual scenarios is shown to underestimate the range of price outcomes and overestimate the range of potential CO_2 emission outcomes, given uncertainty.

The final study of the thesis compares six different policy options for reducing carbon emissions in the electricity system: a cap on CO_2 emissions (as with a cap and trade scheme), a CO_2 price, a renewable capacity target, a green certificates scheme, a renewable generation subsidy, and a renewable capital grant under different treatments of long-term uncertainty. In a case study of a small power system, the results show that using common modelling approaches that attempt to capture uncertainty as multiple different independent scenarios (such as scenario analysis or Monte-Carlo simulation) perform poorly at representing the reaction of a competitive electricity market as measured by a stochastic optimisation model. A policy maker using a scenario-based approach to make decisions could set a policy 55% more restrictive than required to meet their emission target. Further, a deterministic model that ignores uncertainty can underestimate carbon abatement costs by up to 85%. Incorporating uncertainty as individual scenarios only slightly improves this result and biases the estimated costs between price and quantity-based policy approaches to decarbonizing the system.

Throughout this thesis, a continuous set of results are presented that make the case for long-term uncertainty being a critical consideration for the electricity system modeller. In addition, contributions are made to the difficult problem of how to reasonably represent such uncertainty into a generation expansion planning model while maintaining a level of complexity that means the model can be reliably solved on standard hardware. The research is intended to provide motivation for decision makers to dedicate greater effort into considering this important variable and to stimulate continued academic research into this subject.

Keywords

Uncertainty modelling

Generation expansion planning

Electricity market modelling

Decarbonization policy

Renewable energy support policy

Resumo

A incerteza é um aspecto cada vez mais importante da tomada de decisões relacionadas com os sistemas elétricos do futuro. A história sugere que as previsões de horizontes temporais a longo prazo necessários para decisões de investimento e formulação de políticas públicas tendem a ser variadas e incertas. Portanto, a incerteza da previsão a longo prazo é um aspecto importante de qualquer modelo ou exercício de planeamento do mercado de eletricidade.

Para representar uma gama mais ampla de incertezas de longo prazo no modelo de planeamento de expansão de geração, esta tese começa por investigar uma das possibilidades mais promissoras para reduzir a complexidade: a representação do tempo no modelo. A tese examina várias abordagens diferentes para selecionar dias representativos para modelos de planeamento de expansão de geração, com foco em garantir que desafios futuros do sistema de energia sejam representados e que o papel do armazenamento de energia possa ser avaliado com precisão. Proponho uma metodologia para ajustar a ponderação obtida com os algoritmos de *clustering* para os dias representativos a utilizar nos modelos de planeamento de expansão de geração que garante que o nível alvo de carga líquida é capturado no modelo sem alterar as formas subjacentes de carga líquida que definem os desafios de seguimento das rampas de consumo. Os resultados demonstram a importância de executar cuidadosamente o

clustering de dias representativos, apresentando planos de expansão que diferem significativamente em termos de capacidade total instalada tecnologia escolhida.

Esta tese centra-se de seguida no papel da incerteza no importante planeamento do sistema ou na questão do modelo de mercado para quantificar os custos da descarbonização numa economia em desenvolvimento. Com esse fim, modificou-se um modelo de planeamento de expansão de geração para incorporar a realidade da falta de combustível e da troca de combustível típica do sistema de energia de um país em desenvolvimento. A partir desse modelo, são gerados vários custos de redução de emissões que fornecem importantes parâmetros de referência para a região subsaariana, uma região relativamente pouco estudada, identificando fatores determinantes desses custos específicos para países em desenvolvimento. Os resultados encontrados sugerem que as taxas de desconto, representando o acesso do Gana ao financiamento, são uma variável particularmente importante para os países em desenvolvimento. Os resultados sugerem que taxas de desconto mais reduzidas podem levar a maior investimento em energia renovável intensiva em capital a longo prazo, mas também podem travar investimento adicional em geração convencional a curto prazo.

O estudo do Gana permitiu constatar que diferentes suposições acerca dos principais factores de incerteza, tais como custos da tecnologia e crescimento da carga, resultaram em sistemas de produção e em custos de redução de CO_2 significativamente diferentes. No entanto, o tratamento dado à incerteza, apesar de típico de um exercício de GEP, pode na verdade subestimar o efeito geral que a incerteza tem nos resultados.

Esta tese investiga assim a importância de representar uma ampla gama de fontes económicas e físicas de incerteza no modelo do sistema elétrico, assim como o método com o qual a incerteza é incorporada, tanto para a tomada de decisões de investimento quanto para a análise de políticas.

Os resultados de um estudo de caso do Reino Unido demonstram a importância de combinar incerteza entre diferentes *inputs*, sugerindo que a diferença entre uma solução determinística e estocástica aumenta de forma não linear quando os *inputs* são combinados. Para além disso, mostro que a combinação de fontes de incerteza, através da adição de um número limitado de cenários a múltiplas fontes de incerteza, supera o desempenho de adicionar cenários adicionais a qualquer fonte individual de incerteza. Finalmente, mostro que a representação da incerteza como cenários individuais subestima a gama de resultados de preços e sobrestima a gama de possíveis resultados de emissões de CO_2, dada a incerteza.

O estudo final da tese compara seis opções de políticas diferentes sob diferentes tratamentos de incerteza a longo prazo: a limitação de emissões de CO_2 (como num esquema de limitação e comércio), um preço para o CO_2, uma meta de capacidade renovável, um esquema de certificados verdes, um subsídio à geração renovável, e uma concessão de financiamento renovável.

Para um caso de um pequeno sistema de energia, os resultados mostram que a utilização de abordagens comuns que tentam capturar incerteza como múltiplos cenários independentes (como análise de cenário ou simulação de Monte-Carlo) tem um desempenho mau na representação da reacção dos mercados de eletricidade competitivos, conforme medido por um modelo de otimização estocástica. Os meus resultados sugerem que um formulador de políticas que use essa abordagem para tomar decisões pode definir uma política 55% mais restritiva do que o necessário para atingir sua meta de emissão.

Para além disso, descobri que um modelo determinístico que ignora a incerteza pode subestimar os custos de redução de carbono em até 85%. Incorporar a incerteza como cenários individuais apenas melhora ligeiramente esse resultado e distorce a relação entre as abordagens políticas baseadas em preço e em quantidade para descarbonizar o sistema.

Ao longo desta tese, apresento um conjunto sustentado de resultados que justificam a consideração da incerteza de longo prazo como aspecto crítico na modelização de sistemas de energia eléctrica. Para além disso, esta tese contribui para a solução do difícil problema de como representar razoavelmente essa incerteza num modelo de planeamento de expansão de geração, mantendo, simultaneamente, um nível de complexidade tal que permita tratar o problema com segurança com recurso ao *hardware* padrão. Creio que a investigação que conduzi motivará os decisores a dedicar maior esforço a este aspecto importante do tratamento da incerteza e que contribuirá para estimular um esforço continuado de investigação académica sobre este assunto.

Palavras-chave:

Planeamento da Expansão de geração

Incerteza de Longo Prazo

Modelização de Mercados de Electricidade

Políticas de Decarbonização

Políticas de Suporte às Energia Renováveis

Table of Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

CCGT	Combined cycle gas turbine
CCS	Carbon capture and storage
CFL	Compact fluorescent lighting
CGTEP	Combined generation and transmission expansion planning
CO2	Carbon dioxide
CPU	Central processing unit
D	Deterministic
DC	Duration curve
DR	Discount rate
DSM	Demand side management
DSR	Demand side response
DWS	Demand, wind, and solar series
EIT	European institute of innovation technology
ES	Energy storage
ETS	European union emission trading scheme
EU	European union
EV	Expected value
EVPI	Expected value of perfect information
FCT	Fundação para a ciência e tecnologia
FFS	Fast forward selection
GBM	Geometric Brownian motion
GDP	Gross domestic product
GEP	Generation expansion planning

HFO	Heavy fuel oil
IAM	Integrated assessment model
IS	Independent scenarios
LCO	Light crude oil
LDC	Load duration curve
LEAP	Long-range alternatives planning
LNG	Liquified natural gas
LPG	Liquid petroleum gas
MCP	Mixed complementarity problems
MILP	Mixed integer linear programming
MIP	Mixed integer program
MRU	Max ramp up
NCRE	Non-conventional renewable energy
ND	Net demand
NDHF	Net demand historical and future
NES	National electrification system
NRMSD	Normalised root mean square deviation
O&M	Operation and maintenance
OCGT	Open cycle gas turbine
OPF	Optimal power flow
OU	Ornstein Uhlenbeck
PAM	Partitioning around medoids
PPMP	Power plant mix problem
PV	Photovoltaic
RAM	Random access memory
RC	Renewable cost
RDC	Ramp duration curve
RE	Renewable energy
RES	Renewable energy sources
RPS	Renewable portfolio standard

RSDM	Relative slope decision making
RSMD	Root mean square deviation
SA	Scenario average
SDG	Sustainable development goal
SO	Stochastic optimisation
SRA	Scenario result average
STOR	Short term operating reserve
TEP	Transmission expansion planning
UC	Unit commitment
UK	United Kingdom
US	United States
USD	United States dollar
VOM	Variable operating and maintenance
VOPI	Value of perfect information
VOSS	Value of stochastic solution
VRES	Variable renewable energy sources
VSRAS	Value of scenario result average solution
VSS	Value of the stochastic solution

Nomenclature

Indexes and Sets

$d \in D$	Set of days modelled in year.
$f \in F_p$	Set of fuels for power plant p.
$p \in P$	Set of all generating power plants.
$p \in PB$	Set of build candidate power plants.
$p \in PC$	Set of conventional generating power plants.
$p \in PH$	Set of large-scale hydro generating power plants.
$p \in PR$	Set of variable renewable generating power plants.
$p \in PS$	Set of storage generating power plants.
$s \in S$	Set of scenarios in horizon.
$t \in T$	Set of time periods in day.
$y \in Y$	Set of years in horizon.
$y \in Y^{NA}$	Set of years to apply non-anticipativity constraints.

Parameters

$\overline{B}_p$	Maximum number of units of plant p
B_p^E	Number of existing units of pant p
C_{syp}^B	Build cost for a unit of plant p in year y for scenario s($/MW)
C_{sy}^e	Cost carbon emissions in scenario s, year y ($/tonne)
C_p^F	Annual fixed cost of plant p ($/unit)
C_{syp}^g	Variable generation cost for a unit of plant p in year y for scenario s($/MWh)
C_{syp}^o	Period operating cost for a unit of plant p in year y for scenario s (must include generation cost at minimum stable level) ($/h)

C_p^{rc}	curtailment cost for a renewable unit of plant p (\$/MWh)
C^{rD}	Cost of spin down violation (\$/MWh
C^{rU}	Cost of spin up violation (\$/MWh)
$\mathrm{C_p^S}$	Cost to start a unit of plant p (\$/MW)
C^{UE}	Cost unserved energy (\$/MWh)
CF_{sydtp}	Maximum capacity factor of renewable unit p, year y, day d, time t for scenario s
D_{sydt}	Electricity demand in year y, on day d, at time t, for scenario s (MW)
DF_y	Discount factor applied to year y costs
DF_y^S	Discount factor applied to year y costs for multi- year step
dr	Annual discount rate (%)
EF_p^c	full cycle efficiency of energy storage plant p (%)
EF_p^F	Fuel efficiency of plant p (%)
ER_p^{CO2}	Emission rate for plant p (tonne CO_2 / MWh)
$\overline{ES}$	Maximum energy storage per unit (MWh)
$\overline{G_p}$	maximum capacity of a unit of plant p (MW)
$\underline{G_p}$	minimum stable level a unit of plant p (MW)
$\underline{G_p^H}$	Minimum production at each point in time of hydro unit p (MW)
L_y^f	Annual fuel limit for fuel f in year y (MWh)
L_y^e	Annual carbon emission limit in year y (tonne)
L_p^H	Annual production limit for hydro station p (MWh)
ND_d	Total net demand for day d (MW)
ND^A	Total annual net demand (MWh)
R^{cg}	ratio of maximum charge rate to maximum discharge rate (%)
$\mathrm{R_{sydt}^D}$	Required spinning reserves (down) in year y, on day d, at time

	t, for scenario s (MW)
$\mathrm{R}^{\mathrm{U}}_{\mathrm{sydt}}$	Required spinning reserves (up) in year y, on day d, at time t, for scenario s (MW)
RL^{D}_{p}	negative ramp limit per unit of plant p (MW/h)
RL^{U}_{p}	positive ramp limit per unit of plant p (MW/h)
T^{RC}_{sy}	Annual renewable capacity target (MW)
T^{RGP}_{y}	Target annual demand met by renewable generation (%)
W^{C}_{d}	Clustering based weight of representative day d
W^{D}_{d}	Weight of representative day d
W^{S}_{s}	Weight of representative scenario s
	Variables
b_{syp}	Units built of power plant p in year y for scenario s
BC_{s}	Total costs relating to build of new units in scenario s ($)
c_{sydtp}	Charge of energy storage plant p, year y, day d, time t for scenario s (MW)
c^{B}_{syp}	Build cost of power plant p in year y for scenario s ($)
c^{S}_{sydtp}	Start cost incurred for power plant p, year y, day d, time t for scenario s (check unit)
DF_{y}	Discount Factor for year y
DF^{S}_{y}	Total discount factor for all years in step, starting at model year y
e^{CO2}_{sydt}	Carbon emissions for power plant p, year y, day d, time t for scenario s (tonnes)
es_{sydtp}	Stored energy of energy storage plant p, year y, day d, time t for scenario s
FC_{s}	Total fixed costs of all units in scenario s ($)
g_{sydtp}	Generation over minimum stable level of power plant p at year

	y, day d, time t for scenario s (MW)
g_{ydtp}^{f}	Generation over minimum stable level of power plant p at year y, day d, time t for scenario s using fuel f (MW)
o_{sydtp}	Number of operating units of power plant p at year y, day d, time t for scenario s
o_{ydtp}^{f}	Number of operating units of power plant p at year y, day d, time t for scenario s using fuel f
$\mathrm{r}_{\mathrm{sydt}}^{\mathrm{D}}$	Spin down provision of unit p, year y, day d, time t for scenario s
$\mathrm{r}_{\mathrm{sydt}}^{\mathrm{U}}$	Spin up provision of unit p, year y, day d, time t for scenario s
rc_{sydtp}	Generation curtailed for renewable power plant p, year y, day d, time t for scenario s (MW)
s_d	Scalar to be applied to clustering weight for day d
SC_s	Total system costs in scenario s ($)
v_{sydt}^{rD}	Violation of spinning reserve DOWN year y, day d, time t for scenario s($/MWh)
v_{sydt}^{rU}	Violation of spinning reserve UP year y, day d, time t for scenario s($/MWh)
v_{sydt}^{UE}	Unserved energy year y, day d, time t for scenario s ($/MWh)
VC_s	Total variable costs of all units in scenario s ($)
	Clustering weighting methodology
$\|D\|$	The count of days in the year.
ND_d	The total net demand of day **d.**
ND^A	The total annual net demand for the year.
s_d	The scalar to be applied to the weight for day d.
W_d^C	The weight derived from the clustering for day d (the direct

output of the clustering).

Clustering Settings

DWS	Individual historical demand, wind capacity factor, and solar capacity factor included in clustering.
ND_DC	Historical net demand as duration curve used in clustering
ND_DC_RDC	Historical net demand as duration curve and historical net demand ramp duration curve used in clustering
ND_S	Historical net demand as original chronological series used in clustering
ND_S_DC_RDC	Historical net demand as original series, as duration curve, and as ramp duration curve all included in clustering
NDHF_DC	Historical and future year net demand as duration curve used in clustering (two data series total per day)
NDHF_DC Fut MCRD	Historical and future year net demand duration curve with day with max consecutive ramp down included as medoid
NDHF_DC Fut MCRU	Historical and future year net demand duration curve with day with max consecutive ramp up included as medoid
NDHF_DC GMMD	Historical and future year net demand duration curve as original chronological series with day with max greatest min max distance included as medoid
NDHF_DC GMMD6H	Historical and future year net demand duration curve with day with max greatest min max distance over six hours included as medoid
NDHF_DC MinV	Historical and future year net demand duration curve with day with min value included as medoid
NDHF_DC MRD	Historical and future year net demand duration curve with day with max ramp down included as medoid

NDHF_DC MRU	Historical and future year net demand duration curve with day featuring max ramp up included as medoid
NDHF_DC_RDC	Historical and future year net demand as duration curve and historical net demand ramp duration curve used in clustering (four data series total per day)
NDHF_S	Historical and future year net demand as original chronological series used in clustering (two data series total per day)
NDHF_S Fut MCRD	Historical and future year net demand as original chronological series with day with max consecutive ramp down included as medoid
NDHF_S Fut MCRU	Historical and future year net demand as original chronological series with day with max consecutive ramp up included as medoid
NDHF_S GMMD	Historical and future year net demand as original chronological series with day with max greatest min max distance included as medoid
NDHF_S GMMD6H	Historical and future year net demand as original chronological series with day with max greatest min max distance over six hours included as medoid
NDHF_S MinV	Historical and future year net demand as original chronological series with day with min value included as medoid
NDHF_S MRD	Historical and future year net demand as original chronological series with day with max ramp down included as medoid
NDHF_S MRU	Historical and future year net demand as original chronological series with day featuring max ramp up included as medoid
NDHF_S_DC_RDC	Historical and future year net demand as original series, as duration curve, and as ramp duration curve all included in

clustering (six data series total per day)

Chapter 1

Introduction

This chapter provides a brief overview of the interaction of electricity systems, uncertainty, and modelling, before providing a summary of the contribution of this thesis to the literature.

1.1 Context

The last three decades have seen successive growth in global temperature due to increased greenhouse gas emissions in the process referred to as global warming, which is resulting in climate change. The continuation of this trend is projected to have extraordinary social, economic, political, and environmental costs and the international response to address this crisis is slowly building [1]. Currently, more than 75% of the greenhouse gas emissions are related to the extraction, conversion and use of energy with 25% attributed directly to electricity and heat production [1]. It is clear that a key pillar of our hope to reduce the quantity of greenhouse gases emitted by human activity is a reduction in the quantity of energy we consume as a society and the decarbonization of the sources of this energy [1].

The energy sector is, as a result, in the process of a great transformation towards sustainability. This transformation is particularly acute for the electricity system as the desire for decarbonisation of energy sources leads to the push for greater electrification across sectors, such as transport and heating [2]. The simple business of traditional utilities then has to adapt to a host of new technologies, on both the supply and demand sides, introduced by this drive to achieve sustainability in the sector through the decarbonization of the energy sources.

All this essential change, however, introduces a raft of new uncertainties that complicate the ability of decision makers to achieve their policy goals and for the system to adjust as a whole. In additional to uncertainty around the level of electricity the system needs to deliver and the cost of fossil fuels, there are now key uncertainties in the level and cost of renewable energy sources as well as the form and length of the policies being introduced to support them. Unfortunately, these long-term sources of uncertainty have traditionally been either underestimated [3], oversimplified [4], or even ignored [5] in the modelling tools typically used for decision making. Therefore, addressing the challenge uncertainty poses has been identified as a key challenge for decision making models and the corresponding tools in the energy sector [6].

Quality decision making is then both challenging and of the utmost importance when it comes to the provision of electricity infrastructure. For example, policy makers, regulators, and system operators are entrusted to make the decisions that achieve these public goals of decarbonization and sustainability while maintaining secure energy supplies at affordable prices [6]. At the same time individual utilities must make profitable investment decisions valid over long time horizons in an increasingly unknowable marketplace [7]. The consequences of the decisions made both depend on, and shape, the transition of the sector.

This thesis intends to contribute to the quality of decisions in this challenging environment by exploring the role of long-term uncertainty in the methodologies commonly used for modelling and forecasting that support decision making in the energy sector.

1.2 Energy sector modelling

To make decisions in the energy sector we often turn to the use of models, simplified representations of the system, to forecast and assess different future investment options. However, the future challenges facing the energy sector are predicted to be not just different but more complex than those seen historically [8]. For example, the problem of integrating variable renewable energy sources (VRES), and the demand for flexibility it places on the system, increases with the proportion of VRES in the system and will therefore be significantly higher in the future and require greater adaption of the system [8,9]. To arrive at robust results in studies of the power system, the planner needs to ensure that the models used for decision making contain the detail required to represent these future challenges.

1.2.1 Types of models

There are many different questions and decisions that the decision maker may wish to apply models to, in order to generate insight. The chosen model will partly depend on these questions, in particular the size and scope required. The models used for long term decision making in the energy sector broadly can be partitioned based on their scope and goals [10]:

1. **Integrated assessment model (IAM)** – IAMs cover the broadest scope of the modelling approaches considered here, incorporating societal changes such as demographic transition and the interaction between the macro-economy and the energy system. Both the timeframe (50-150 years) and geographic basis (regional - world) are also large. As a result of this broad scope this class of model typically incorporates the lowest level of specific power system detail.
2. **Energy-economy models** – Energy-economy models focus on the interaction between the energy system and the economy, typically at the national or regional level. This class of model is motivated by the insight that the macroeconomy is a key driver of the demand for energy, and the price and availability of energy has an important influence on macroeconomic outcomes.
3. **Energy system models** – Energy system models reduce the scope of the modelling problem by focusing on modelling the energy system and taking outcomes from the macroeconomy as inputs[1]. These models can then allow for additional detail in the

[1] A simplified representation of the relationship between energy prices, the macroeconomy, and demand, can be incorporated as a demand curve or demand elasticity

representation of the energy system, typically incorporating the entire supply chain for all major energy sources and including all requirements for energy in a given society. In capturing the needs for energy as a whole, trade-offs can then be evaluated between different sources of energy; for example, providing household heating with natural gas, biomass, or electricity.

4. **Long-term power systems models** – Power system models restrict the scope of the model even further to that of the power system itself. The advantage of this approach is that it allows for the most detailed representation of the complex operation of the power system [11]. Traditionally, these models were developed for a central planner (or a vertically integrated utility) who is fundamentally trying to meet the demand for electricity with the least cost combination of technologies[2]. For long-term planning, the class of model depended on whether the planner was considering generation or transmission. However, the importance of combining these decisions is increasingly considered important. Broadly the classes of models are:
    a. **Generation Expansion Planning (GEP) -** The GEP problem is that of "*determining WHAT, WHEN, and WHERE new generation units should be installed over a planning horizon, satisfying the expected energy demand*" [12]. As such, this problem was sometimes called the power plant mix problem (PPMP) [13]. The traditional GEP model either ignores the question of the transmission network or includes an overly simplified representation that excludes the additional complexity of determining optimal power flow (OPF). However, in determining which power plant mix is appropriate, detailed

for energy.

[2] We discuss extension of these models beyond central planning in section 1.3.

modelling of the operational characteristics of different technologies can be incorporated and is increasingly considered essential for representing future systems [14].

b. **Transmission Expansion Planning (TEP)** – The transmission expansion planning problem is similar to the GEP described above. Its goal is typically cost minimisation. However, the decision being made is the selection of investments in the transmission system subject to operational constraints. As such, the transmission system is modelled in significant detail and the generation stations are either modelled as a simplified decision or taken as an input assumption.

c. **Combined Generation and Transmission Expansion Planning (CGTEP)** – The CGTEP recognises that there is often a trade-off between the type and location of generation investments and the increased cost in transmission investment. The CGTEP problem then seeks to optimise these decisions together. However, this approach requires model complexity to be managed in some way, usually by reducing the level of detail in some aspect of the system.

The above classification allows a high-level view of the different models available to decision makers and how each of these relate to each other. In this thesis, I focus on the GEP class of models for their superior ability to represent the challenges the power system faces [11] and its popularity in industry, academia and in policy decision making. While I examine long-term uncertainties in this thesis, it is clear from the literature that the incorporation of detailed short term dynamics remains important for the power sector and I, therefore, select the GEP approach as the best able to represent both the long-term and short-term implications of uncertainty for the power system.

1.2.2 Decision making objectives

GEP models are bottom-up models that optimise the quantity of installed capacity and operation of different electricity generation technologies to meet consumer demand (at the local, national, or international level, depending on scale). Broadly, these models are used for decision making in the energy sector in a number of ways [10]:

- **Normative perspective**: GEP models can be used to find the optimal expansion plan, typically in terms of selecting the lowest cost set of technologies, but possibly subject to achieving certain targets such as decarbonisation levels or system reliability metrics. The GEP class of models was originally developed for the centralised planner seeking to meet consumer demand and reliability requirements by the least cost combination of technologies;
- **Descriptive perspective:** With the transition away from central planning towards competitive electricity markets, the GEP model found new purpose as a means to forecast decentralised market outcomes for policy and investment analysis [10]. In this application, the modeller uses GEP tools to model the expected outcomes in the electricity market under different sets of conditions[3]. In this case, by minimising cost, the GEP model is finding the solution that maximises social welfare, which is also the equilibrium point of a perfectly competitive market [10][15]. As such, a GEP model can be used to provide a descriptive view of future power system outcomes under perfect competition.

[3] Including what the modeller considers the most likely set of future conditions.

In the following section, I explore in more detail the use of the GEP model for descriptive market analysis. I also discuss in detail the application of the GEP model for descriptive policy analysis in Section 1.6.1.

1.3 GEP for market modelling

The GEP and TEP models were developed as a tool for centralized vertically integrated utilities to determine the least costs means to provide the electricity system required to meet our demand for electricity. In such a situation the utility acts as a central planner with complete control of the investment and operational decisions of the entire system. As such, the GEP and TEP are intuitively cost optimisation problems and the interpretation of the resultant decisions is straight forward (the least cost means to achieve the planner's goals of meeting electricity demand). However, with the transition to a decentralised competitive market for electricity, decision makers are now interested in descriptive analysis of what will happen in the market. In a competitive market, there are multiple individual firms acting to maximise their own profits, as opposed to a single entity with complete control acting to minimise the overall system costs.

Probably the first insight into this issue was discussed as early as 1988 in a presentation of the spot market concept [16]. The authors found that at the solution point of the OPF optimisation, the resultant decisions maximise the profit for producers and the surplus for consumers[4] and produces a competitive price. This set of outcomes[5], where a market

---

[4] This result is extended beyond a system with nodal prices in [137], where the authors allow for physical transmission rights, although still exclude uncertainty.

[5] The outcomes for a competitive market can be considered to be simply a price and

system maximises total welfare, is an important one in economics. In normative economics or "welfare economics" the first fundamental theorem of welfare economics tells us that a competitive market will be Pareto efficient and maximise social welfare[6] [17]. This is the original insight of Adam Smith [18]: self-interested individuals maximising their own benefits in a competitive interaction can produce the outcome that maximises the total benefits to society overall. In positive economics the outcome of a competitive market, the "competitive equilibrium", is used to forecast market outcomes, particularly for commodities [19]. However, these results only hold under a reasonably strict set of assumptions which holds in the case of a stylized perfectly competitive electricity spot market without uncertainty [16].

In fact, the finding that the competitive market can maximise benefits to consumers and producers (and not just a central planner) was a key motivation for the change in structure towards competitive markets [15]. In [20] the authors extend this insight, finding that under a range of proposed electricity market structures (with more or less decentralization of decision-making) the market equilibrium is equivalent in most cases, and again equivalent to the centralized optimization case. However, these results relate only to the short-term operation of the electricity market and do not account for

---

allocation of goods [17].

[6] Whether the market result supports weak or strong Pareto efficiency depends on whether there are only non-satiated preferences (whether there are no "fat" indifference curves). Non-satiated preferences are required for strong Pareto efficiency as otherwise one participant may be indifferent between outcomes that reduce the utility of another participant.

uncertainty.

The finding that the competitive market should be expected to achieve the same welfare maximising result of a central planner allows us to extend the application of optimisation models, such as the GEP, to the case of the competitive electricity market [10]. The extension to the case of the generator expansion planning under uncertainty is surprisingly intuitive from this point. As [15] states, *"One can immediately note that the primal-dual conditions of the stochastic capacity expansion model can again be interpreted as describing a stochastic equilibrium in a perfectly competitive market provided one assumes that all agents are risk neutral and share the same probabilities of the different possible developments of the market."*

In other words, it is clear that the result found in short term markets extends to the case of long-term investment decisions even with the introduction of uncertainty. The literature on the application of mixed complementarity problems (MCP) to electricity markets provides a more rigorous exposition of this claim. In the MCP approach to GEP, each investor individually optimises their investment decisions to maximise their individual profits, before competing against each other in a short-term electricity market that minimises the costs of meeting consumer demand[7]. In this way, the decentralised market is modelled directly with investors making individual decisions based on their own profit maximising objective. The MCP framework extends beyond what can be achieved by a GEP optimisation model by allowing for the existence of market power (relaxing the condition of perfect competition). In [21] and in [10], the authors derive the

[7] Many "gross pool" electricity markets work explicitly like this where an optimisation clears all bids and offers. The work of [20] shows this result extends to other electricity pool designs such as "net pool".

result that with the assumption of perfect competition, the MCP formulation is equivalent to a centralized least-cost planning problem. This demonstrates once again, this time for a GEP model, that the solution to the centralized optimization problem can be interpreted as the result of a decentralized market. Further in [21], the authors extend this result to the case where uncertainty is included. The authors show that where investors consider uncertainty in their investment decisions, the MCP formulation has an equivalent stochastic optimisation model representation. That is, a market where individual investors are competing for profit with uncertainty about future market conditions is equivalent to a model where a central planner is optimising cost, assuming that investors share the same views of the future, are risk neutral, and perfect competition holds. In [15], the authors extend this insight and demonstrate that the assumption of risk neutrality can be relaxed. In the model of [15], even under the conditions of heterogeneous risk aversion among investors (different investor preferences for risk), a stochastic optimisation with risk constraints can be used to represent a perfectly competitive market assuming participants have a means to trade risk.

The conditions for which a centralized optimisation can be used to model market outcomes are discussed in [10] and [22]. Broadly, the following conditions are required:

- Perfect competition – It is required that the market is perfectly competitive, and no firm has the ability to influence price strategically (by, for instance, withdrawing capacity). Additionally, the standard assumptions that there are no non-internalised externalities, that no participants hold asymmetric information, and that electricity can be considered a normal good, are required.
- Integratable demand curves – As electricity cannot be easily exchanged between different points in time, technically electricity at each point in time is a separate good. If consumers have multiple different demand functions, with different cross price elasticities, it would be possible that consumer demand could not be integrated

to a single demand function to be used in the optimisation model[8]. In the GEP model it is, however, common to assume "inelastic demand" in the short term, that is that consumers do not respond to short term changes in the price of electricity. This assumption is a simplification used to reduce the complexity of the GEP model and has the additional property of ensuring integratable demand.

- Shared information set – All agents must have the same view of the costs and revenues of participating in the market, this is a relaxation of the assumption that all agents have perfect information as the model can now incorporate uncertainty, but we require that all agents have the same view of this uncertainty.
- Market clearing – The market can be cleared, and participants receive the unique competitive equilibrium price (a pay-as-clear market). Further, this market clearing price cannot influence variables not relating to the clearing of that market. In practice, this means there cannot be an input variable that depends on the market price as this price is a model output (a constraint dual value). For instance, the subsidy to a renewable technology could not be set to be a certain mark-up (%) over energy prices (but the subsidy could be an absolute value in terms of €/MWh).

The work of [10], [15], [21] and [22] demonstrates the ability of a stochastic optimisation GEP to perform descriptive analysis and forecast market outcomes where investors must make investment decisions taking uncertainty into account. This result helps explain the popularity of using the centralised optimisation approach to forecast results in competitive electricity markets and to perform policy analysis [11]. In fact, a report by the European Commission finds that the majority of the pressing electricity market

[8] For instance, this would not be possible if cross-price elasticities between electricity demand at different points in time were not symmetrical [10].

design and policy questions for competitive markets can be represented in optimisation models such as the GEP [23].

1.4 GEP in the literature

While the use of the GEP model to represent electricity markets depends highly on the modellers willingness to accept a restrictive set of assumptions, these models typically offer the greatest level of technical detail of the underlying power system [11]. GEP models are applied both in industry for investment decision making and, in a recent review of the literature, were found to be the most common approach to modelling the electricity system for government policy decision making [11]. In fact, the European Commission recently commissioned a model for market modelling and policy analysis based upon the optimisation approach [24].

However, the level of detail incorporated into the GEP model comes at a price in terms of complexity. The GEP problem remains an active area of research not simply due to its importance but further, due to its notorious size and complexity. The GEP suffers from nonlinearity, non-differentiability, high-dimensionality, and discrete decision variables [7]. The literature on this class of models then is extensive and growing, both for the challenge presented by formulating and solving the GEP problem and the usefulness of the GEP model to a wide range of applications across industry, government and academia. This section examines a number of recent publications that attempt to review this literature, under different lenses, and draws on the open questions they identify.

In a relatively early 2014 review of challenges facing energy system models in the 21st century, [6] identifies four key challenges:

1. **Resolving time and space -** For future energy system models, the level of temporal

and spatial detail must be balanced against the increased model complexity. The increasing introduction of VRES makes this trade-off more important than ever. A detailed representation of time is required to incorporate the effect of VRES on the power system and asses the role of ES and DSR. However, VRES are also more restricted in the locations in which they can be exploited (for example, offshore wind) and so increase the challenge of operating and expanding the transmission system. These requirements mean many of the simplifications traditionally used in energy system modelling, that allow us to reduce model complexity, are no longer appropriate.

2. **Balancing uncertainty and transparency** – Uncertainty is increasingly a challenge faced by decision makers in the energy sector and incorporating its effect on decisions is a key determinant of the usefulness of a model. However, adding this uncertainty into the modelling add complexity to the modelling which can also complicate the interpretation of results. At the same time, transparency of results is an additional important criterion for ensuring decision makers can have faith in their model results, particularly in the case of public policy making. Balancing the reduction in transparency of increasing the uncertainty implemented in these models will remain an increasing challenge.
3. **Addressing complexity and optimisation across scales** – Energy systems appear to be examples of complex systems, that is, systems that cannot be formulated compactly. This complexity grows as energy systems incorporate increasingly diverse and decentralised resources. In addition, the complexity grows as the scope increases with greater interconnection between systems. As such, many existing models may be over-simplified and no longer be representative of future power systems.
4. **Integrating human behaviour and risk** – Finally, the challenges discussed above, and the energy modelling paradigm as a whole, focus on the technical details of the systems. However, outcomes will increasingly depend on human factors such as

political will, public acceptance, and behavioural outcomes that differ from the typical model of rational entities. These factors become increasingly important as energy systems move towards the use of decentralised resources and attempt to bring forward demand side participation in the market. As such, increasingly incorporating human behaviour into technical based models will be an important future challenge.

The above classification appears to be relatively broad and comprehensive, and these broad classes of challenges identified are repeated throughout the review of literature presented here. The authors consider addressing complexity to be the key challenge faced by power system or market optimisation models. However, this challenge inherently relates to the other three challenges presented here in that it is the level of time and space, uncertainty, and human behaviour included in the model that determines the complexity. More specific formulations of these challenges can be found by examining the literature focusing on GEP and TEP models. For example, a 2018 review of energy modelling tools capable of representing power systems with a high share of renewable energy [14] finds a similar set of challenges:

1. **Representation of variability -** The authors find it clear from the literature that the representation of VRES variability and short-term operation are a key challenge due to their importance for accurate modelling of power system outcomes and the complexity added to longer term models.
2. **Consumer participation, electrification, and sector coupling** – Traditionally, energy sector models focus on the supply side of the system. However, the demand side will play an increasingly important role into the future. Demand participation and electrification will determine the level of demand to be met and the ability to integrate renewable resources.
3. **Impacts and links beyond the energy system** – Typically, energy sector modelling is limited in scope to the energy sector. However, the authors argue that the additional environmental impact of different technologies and system

infrastructure should no longer be ignored. It makes little sense to design models to improve environmental outcomes through minimising CO_2 emission while ignoring the role of other pollutants or the lifetime impact of different technologies on the environment.

4. **Validation and transparency of long-term models** – As discussed previously, transparency and validity remain essential to ensuring modelling tools can contribute to decision making. In particular, the sensitivity of results to the uncertainty in our knowledge of the system (and inputs to the model) must be represented.
5. **Representing the future** – The authors classify the representation of the future in the modelling a key challenge due to the inherent uncertainty. The authors specify the level, type, and cost of VRES, and forecasting the effect of climate change on energy demand and available resources, as example of long-term uncertainties to be considered.

The above findings are based upon a review focused on tools that allow for a high share of VRES. However, it is clear that a high share of VRES will be a feature of most systems in the future and as such these challenge repeat in less targeted review studies. For example, in a highly comprehensive 2017 review and classification of the GEP literature, [7] identifies the following as the areas for development:

1. The continued development of combining GEP and TEP in a liberalized environment that incorporates short-term and medium-term operation. The authors emphasise not just that the CGTEP model is an area for future research, and that short-term operational dynamics such as unit commitment should be included, but also that the incorporation of market behaviour is important given the current regime in which many energy systems operate.
2. The impacts of clean coal, CCS, and nuclear hazards and, more generally, incorporating the environmental impact of different technologies outside of their respective levels of CO_2 emissions. Again as previously emphasised in [14], the focus

on CO_2 emissions can potentially be detrimental where other environmental effects are neglected.

3. The role of smart grid technologies and their contribution to integrating distributed and renewable technologies is identified as understudied in comparison to the introduction of VRES technologies.
4. Considering the impact of energy policies including the increasing uncertainty introduced by such policies and their effect on decision making in a market environment.

Similar results are found in a 2018 literature review of the current state-of-the-art of the literature [25], the authors draw the following conclusions for future research:

1. The representation of competition, risk, and uncertainty has significant impact on results and should be carefully considered. In particular, a wider range of uncertainties (economic, political, regulatory, environmental, technical, social) should be, at least, partially included.
2. Due to the increasing share of VRES, short term market dynamics must be incorporated into long term models.
3. The role of consumer demand response and ES in the modelling framework should be considered, particularly in the case of electric vehicles.
4. The introduction of VRES increases the importance of the combination of GEP with TEP. Additionally, combining the modelling of the natural gas system with GEP can add value.

The majority of the above challenges clearly focus around capturing the effect of the increasing share of VRES in the system. However, the importance of the vast array of long-term uncertainties (outside of the typical short-term uncertainty associated with VRES) is also emphasised. In a 2018 review of GEP models for policy analysis, [11] identifies the following as the important directions currently being explored by the literature:

1. The effect of the inclusion of uncertainty into the GEP model.
2. The incorporation of short-term operational constraints into the long-term GEP model.
3. The representation of renewable promotion incentives and policies.
4. The integrated modelling transmission and generation expansion planning.

A number of clear trends emerge from these wide-ranging reviews of the GEP literature. For instance, the role of VRES is clearly identified across all studies, particularly that its role on the short-term operation of the system is included in long-term models. A number of these reviews also draw attention to combining transmission and generation expansion planning, again at least in part due to the geographically differentiated nature of VRES.

Additionally, several authors note the importance of including the changing policy environment as a factor within the GEP model itself. Both current and expected future policies will clearly drive investment decisions in the real world and so must be represented in the tools used for modelling, in particular, the review of [7] highlights this as lacking in the literature. A number of authors note that the focus on CO_2 emissions and related policies should not mean the neglect of additional environmental effects of the energy system.

The other area where all of this literature reaches a clear consensus is in the importance of the treatment of uncertainty, which is often overlooked in GEP models [5]. It is clear that uncertainty is a key feature of the energy system and influences greatly decision making in the market environment (I discuss uncertainty in the electricity system further in the following section). As such, it is not surprising to find uncertainty so clearly and consistently identified as a challenge for the future of GEP modelling, whether for central planning, market modelling, investment decision making, or policy design and evaluation.

This thesis contributes to one aspect of this key challenge: the incorporation of long-term uncertainty into the GEP model. In exploring this uncertainty, this work invariably touches on other key themes identified above, including: managing model complexity, in particular, through the temporal representation to allow for the incorporation of greater uncertainty (Chapter 3); the representation of investor behaviour and the competitive market under uncertainty (Chapter 5); and the representation and assessment of renewable support policy under uncertainty (Chapter 6). The model formulation used throughout (Chapter 2) also incorporates the detailed short-term unit commitment decisions found to be important in capturing the impact of VRES.

In the following section, I explore in detail why the treatment of uncertainty is so consistently identified as an important factor in representing the energy system.

1.5 The role of uncertainty in the energy market

Uncertainty plays a key role in the energy sector and is becoming increasingly important in planning over both the long- and short-term planning horizons [26]. Both the number of areas where uncertainty is considered material, and the degree to which that uncertainty is perceived, are increasing [5]. The policy maker, energy system planner, and utility are all faced with making or influencing long-term investment decisions, typically for assets with thirty year plus lifetimes, where significant uncertainty exists over a number of important dimensions.

For a demonstration of the magnitude of these uncertainties, consider the historical review of long-term energy demand forecasts undertaken in [3]. The authors demonstrate that historical forecasts of energy demand were both systematically high and underestimated the level of uncertainty. The realised level of total energy demand in the US for the year 2000 was less than half the most optimistic forecasts and well below the average. Given the importance of energy demand growth for investment decisions, it seems probable that assets built based on these forecasts could easily have become stranded. For a more recent demonstration of uncertainty errors, consider a review of forecasts for levelized CCGT costs for the period ending in 2040 [27]. The authors find a near consensus in cost projections for forecasts generated before 2005, i.e. in the 20-40£/MWh range. However, forecasts created between 2005 and 2010 range in values between 60-160£/MWh. This large difference in the spread of the cost projections demonstrates that, historically, uncertainty has oftentimes been under-estimated, and the magnitude of the difference was significant enough to have serious implications for investments made under pre-2005 forecasts[9].

These examples illustrate the scale of uncertainties that exist for decision makers. Additionally, the sources of uncertainty are broad and have been increasing (see [5] for a characterisation), particularly on:

- **Climate change goals** are extremely uncertain due to the number of factors on which they hinge [28].These goals and the policies to achieve them can have significant effects on fuel prices, expected uptake of renewable energy, and the price and availability of carbon credits. Ignoring this source of uncertainty can risk over investing in carbon based generation and creating stranded assets or increasing the costs of achieving climate goals in the long run [28].
- **Climate change policy,** as our climate change mitigation goals change so do our means for trying to achieve them. Uncertainty around climate change policy is an often cited concern of private investors and has been demonstrated to effect investors decisions, and in this way the system's ability to achieve climate targets [29]. Even after a policy is implemented, any uncertainty about when it will be removed (and

[9] Although, at the current point in time the pre-2005 estimates appear closer to the mark than those in the 2005-2010 period. However, the uncertainty in these estimates continues to change with the level of oil market volatility.

whether it will be removed retroactively for existing investments) can have a significant influence of the deployment of renewable energy sources [30]. The review of [7] classifies the incorporation of the effect of uncertainty in climate change policy as a key challenge for energy system modellers into the future.

- **Fuel price uncertainty** has long been an important source of uncertainty and motivated significant growth in the practice of long term energy planning in the 1980's [6]. Fuel price volatility has returned over the last two decades with both unexpected highs and lows, invalidating many of the forecasts on which planning was based [5].
- **Load growth uncertainty** is one of the most important problems for GEP modelling and one of the earliest to be examined [31]. Recently, however, the drivers of this uncertainty have increased due to the promotion of energy efficiency, electric vehicles, smart grid technology, and behind the meter generation and storage. These factors greatly increase the uncertainty in the level of load to be met by the system at any point in addition to the still important (and uncertain) factors of population and economic growth.
- **Technology costs** of renewables and storage technologies are developing at a rapid pace and long term forecasts of their cost structures are inherently uncertain both as the technology matures and as demand for the necessary primary materials increases (see for example [32]). However, waiting for certainty and maturity, risks underdevelopment and drastically increases the costs of achieving carbon goals in the long run [33]. Technology costs are identified as a key sensitivity for estimates of system cost produced in long-term power system models [5].
- **Generation and transmission asset reliability** is another traditional area for uncertainty in electric and transmission systems as the availability (and unexpected faults) of generators and transmission lines is uncertain. This source of uncertainty is particularly important in short term modelling [34]. Interestingly [5] finds that while

reliability is a key driver of the reported system cost of a given model, it has a significantly less effect on the actual investment decisions.

- **Production from renewable resources** suffers as the integration of the uncertainty in production from wind and solar resources increases in importance with their ever-growing contribution to the energy mix. The study of generation expansion with high levels of renewable energy is a particularly active area of study (see [8] and [35], for reviews). However, the uncertainty introduced by the water available for generation from hydro stations remains extremely important in many systems.

Given the scope and scale of these uncertainties, it is then of no surprise that their representation is one of the key challenges in energy system modelling [6]. In the following section, the existing representation of uncertainty in the GEP model is discussed.

1.6 Uncertainty in modelling the energy sector

Uncertainty is an important component of our understanding of the future of the energy sector. Uncertainty should then be an important driver of the decisions that are made whether by investors or policy makers [6]. But how is uncertainty treated in GEP models? Given this class of models is widely used for decision making, this is a crucial question.

The inclusion of uncertainty into energy system models was attempted reasonably early in this class of model (early examples include [36–38]). These early attempts were likely hindered by the computational complexity and available computational resources. Overall, uncertainty seems to be often overlooked or excessively simplified in this class of model. A recent review of the literature on GEP models [7] revealed that of the 227 papers cited, only 54 were classified as containing uncertainty on at least one input, and that only 17 included uncertainty in more than one input dimension. Furthermore, as it can be seen in Figure 1, the types or sources of uncertainty dealt with are primarily physical (level of demand, outages, renewable energy sources) as opposed to economic (fuel price, emission price, technology costs).

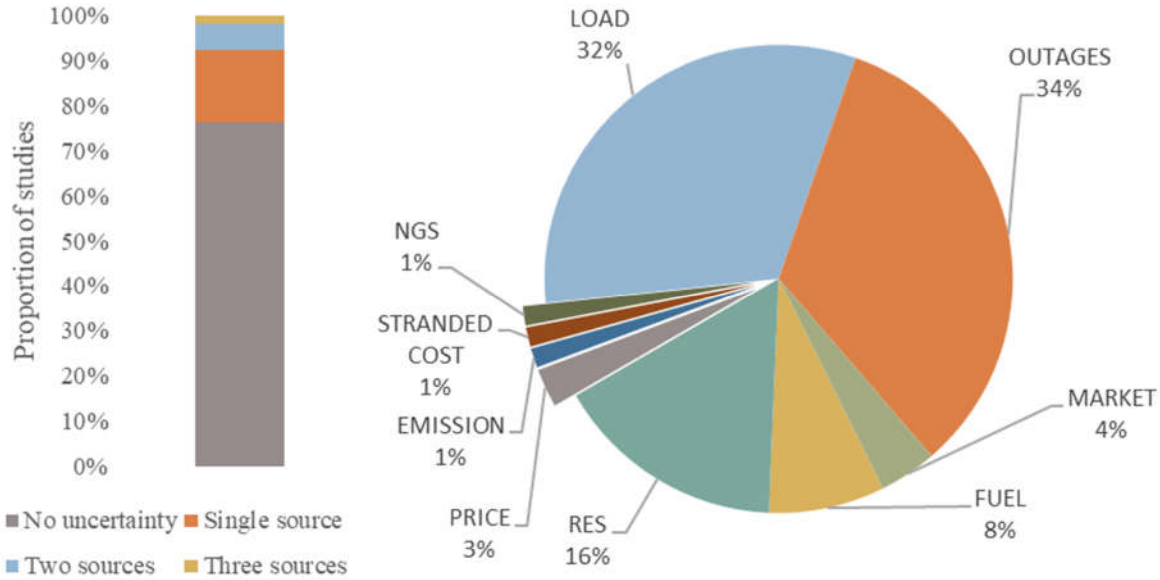


Figure 1 Sources of uncertainty in GEP studies by number of sources and type, adapted from [7].

In addition to the sources of uncertainty, the method used to include uncertainty into the model is also an important consideration. Broadly, uncertainty in energy planning models[10] can be treated in the following ways:

No Uncertainty:

- **Deterministic,** in which the best estimates of inputs are used in the GEP model and uncertainty not otherwise treated.

Uncertainty outside the model:

---

[10] For the case of numeric models, for analytic models see [138]

- **Scenario analysis** in which a number of representative scenarios are selected for key variables and results produced for each scenario (or combination of scenarios). These scenarios can either be chosen explicitly by the decision maker or generated based upon an assumed probability distribution for the underlying input [5].
- **Monte-Carlo analysis**, which involves fitting a statistical process to each key variable and generating possible future scenarios (often thousands). The model is then simulated on each of these scenarios producing thousands of individual results. Averages and ranges are then often taken from this output. Monte-Carlo analysis is one of the most common approaches to generating and including scenarios into energy models [39].
- **Sensitivity analysis** is a common method that allows modellers to test how sensitive results and conclusions are to changes in input parameters. If the modeller has an idea of the uncertainty of a given input, they can combine this with knowledge of the sensitivity of that input to understand at a high level the uncertainty in the modelling results.

Uncertainty inside the model:

- **Stochastic optimisation (SO)** is a modelling approach in which the model is forced to make a single set of decisions given the range of uncertainty or given a number of uncertain scenarios provided. This approach reflects the problem decision makers face with uncertainty in the real world where decisions must be made without knowing the future values of many important factors [40]. While stochastic optimisation offers the ability to incorporate a quantified view of uncertainty into decisions made in the model, it comes at a large computational penalty [40].
- **Chance constrained optimisation**, contrasts with the SO approach of applying all constraints to all futures which can be considered unnecessarily restrictive [41]. Instead, chance constrained optimisation allows the model to look at a less restrictive problem and produce a less conservative solution. In effect chance constrained

programming allows the modeller to choose a specific probability of the constraint violation (or ensure that the constraint would only be violated a certain proportion of the time) [39].

- **Robust optimisation** is another popular alternative to stochastic optimisation which seeks a solution "immunized against uncertainty" [42]. Robust optimisation seeks a solution robust again the worst cases, in effect taking the worst outcomes that are possible into account while ignoring both the best outcomes and the expected outcome overall. In this robust optimization is typically considered a highly conservative approach to modelling uncertainty. Two common approaches to robust optimisation include:
    - **Minmax** robust optimisation which ensures that the cost of the worst-case scenario is minimised. While extremely conservative, or risk averse, this approach in effect means only the worst-case scenario is considered and this case then drives decisions with no reflection of how improbable or costly to improve this scenario may be.
    - **Minimise maximum regret** optimisation changes the minmax approach slightly by considering the costs of the given worst-case scenario. The least regret optimisation minimises the regret, how much it costs us to not address a given scenario. In this way an extremely costly scenario, that cannot be easily mitigated, does not drive results when greater cost improvements can be achieved in other scenarios.

The approaches that treat uncertainty outside the model suffer from a number of issues despite their relative popularity. Firstly, a range of very different results can be produced (one for each scenario) but the decision maker must make a single decision and, therefore, needs a way to aggregate these results (for example the average decision of the results). As I demonstrate in Chapter 5, these approaches can perform poorly.

The second problem is crucial for policy makers attempting to use the GEP model to represent a competitive market or forecast the actions of an integrated utility. In the deterministic scenario analysis, and the scenario analysis or Monte-Carlo simulation methods, the model optimises decisions with perfect knowledge of the future (any uncertainty is effectively treated outside of the model by repetition with changing inputs). Given the real-world decisions that the model is attempting to predict are made taking uncertainty into account this can lead to large differences between all modelled scenario results and real-world outcomes.

As an illustrative example consider a case where there are two possible future levels of electricity demand; one scenario where the transport system has switched to electric vehicles and there is significant GDP growth leading to a high level of demand; and second a scenario where demand remains flat at current levels, potentially due to improved energy efficiency. For simplicity, I assume that these scenarios each have a 50% chance of occurring.

A policy maker can then use a GEP model to forecast the future power system under the two possible futures that exist. In scenario one, the GEP model will optimise for the high load level and predict that new capacity is built to meet demand. In scenario two, the model will predict that the current level of capacity is maintained. The policy maker would conclude that under either scenario there will be a sufficient level of capacity brought forward by the market.

However, consider a profit seeking utility operating in this situation. Would the utility be willing to build new capacity with only a 50% probability of it being needed (this new capacity would be wasted in the low load growth scenario)? This is unlikely to be the case, so the end result is that the utility does not invest in additional capacity. Then in the future, if there is actually high demand growth, the utility has not expanded capacity, meaning possible shortages of energy, and the policy makers decision making model will have completely missed this result (even given that in this illustrative example there is a perfect understanding of the future uncertainty).

This result is illustrative of the problem for decision makers treating uncertainty outside of their models and part of the motivation for this thesis (this effect is discussed in more detail in section 2.2).

When comparing between the approaches that actually represent uncertainty inside the model, the stochastic optimisation offers an important property. As previously discussed, the result of the centralised SO GEP approach has a decentralised market interpretation (broadly what a perfectly competitive market would achieve when participants consider uncertainty in their investment decisions and attempt to maximise their expected profit). In other words, independent investors making decisions over uncertainty in an approach that maximises their individual expected profits can be modelled as a centralised SO cost optimisation. This property does not apply to the other approaches presented here. For example, even if in a market with individual investors minimising their individual worst case scenarios with robust optimisation (maximising profits in the scenario that is the least profitable for them), this market cannot necessarily be modelled as a single robust optimisation problem. In addition, for typical competitive firms the maximisation of expected profits, potentially bounded by risk considerations, would appear a more realistic behaviour than only maximising their profit in a worst case scenario (if this was the case all firms would be extremely risk averse and make minimal investments).

These arguments illustrate the strong case for the stochastic optimisation approach to GEP modelling, however, this is not overly common in the literature. In a review of the related field of storage expansion planning [4], the authors demonstrate that the inclusion of stochasticity in research increased drastically after the year 2000. However, the majority of this research deals with the more simplistic approaches of scenario modelling or Monte-Carlo simulation. For papers published after 2010, still less than 20% feature the more sophisticated stochastic optimisation (SO) approach to handle uncertainty modelling.

The simplified treatment of uncertainty in terms of both limiting the sources included in the model and making use of approaches that treat uncertainty outside of the model is likely a result of the 'curse of dimensionality' where the problem size grows non-linearly as scenarios are added to the SO GEP problem. The inclusion of uncertainty, and balancing the increased model complexity and reduced transparency, is identified as a key challenge energy modellers face [6].

The above discussion covers the inclusion of uncertainty into energy models generally. However, there is one specific application for which uncertainty plays an additional key role, which I discuss in the following section.

1.6.1 Evaluating decarbonization, renewable support policies and the role of uncertainty

One common application for the GEP model is the problem faced by policy makers who are tasked with encouraging the decarbonization of the energy sector to mitigate the effect of CO_2 emissions on the environment. The problem created by the effect of CO_2 emissions, produced in the generation of electricity, is a classic example of an externality. The textbook prescription for an externality is typically a Pigouvian tax [43], due to the simplicity and efficiency of a market based intervention [44]. However, in the reduction of CO_2 emissions the history of debate between quantity measures, such as a cap and trade system, and price measures, such as a tax on CO_2 emissions (often called a carbon price, or CO_2 price), is long. A cap and trade system is argued to have additional benefits in this particular case due to political acceptability and the ability to operate at an international level (where a single worldwide government does not exist to impose a tax) [45].

As it turns out, uncertainty plays a key role in this debate. In a seminal work [46], Weitzman showed that these two approaches are equivalent in terms of social welfare under a set of simple assumptions. However, they differ in the case where carbon abatement costs are unknown. In such a situation the preferred policy depends on the relative slope of the marginal benefit and cost of abatement curves at the equilibrium. Specifically, a price-based approach is preferred if, and only if, the slope of the marginal cost of emissions abatement curve is greater than the slope of the marginal benefit of the abatement curve. Comparing decarbonisation policies on this basis became known as relative slope decision making (RSDM) and the literature has been expanded to relax a number of the more stringent assumptions (see [44], for a review). However, no consensus remains on the preferred policy prescription [44].

The above literature relies on a high-level theoretical analysis, for example, a generic choice is available between supplying to the market with carbon abatement and supplying to the market without carbon abatement. This approach also relies critically on the modelled marginal abatement curves [44]. However, investment decisions in the electricity market are significantly more complex due to the complexity of transmission constraints, the sizing constraints of new entrants, the difficulty of storing electricity inter-temporarily, the requirement to provide ancillary services, the variability and uncertainty of variable renewable generation sources, among other concerns [25]. The interaction of decarbonisation or renewable support policies and the operation of the power system is complex, and policies can have adverse consequences on important power system requirements such as resource adequacy [47].

For the practical purpose of evaluating long-term policy in the electricity sector, GEP models are then the approach favoured by policy makers [11]. In fact, a report by the European Commission finds that the majority of the pressing electricity market design questions for competitive markets can be represented in optimisation models such as the GEP [23] and the European Commission itself has selected this approach for the model it has commissioned to perform policy analysis [48]. In fact, 90% of energy system models available commercially or academically aim to address the questions asked by policy makers [49].

It appears clear then that inclusion of long-term uncertainty into the GEP models will have important implications for policy makers that rely on these models.

1.7 Summary and Motivation

Decision makers in the energy sector face increasing uncertainty across a number of important dimensions. Over the period of interest for long-term decisions, such as the investments in generation assets or the evaluation of government policies, this uncertainty becomes a fundamental characteristic of the system as a whole. However, uncertainty has typically been ignored or considered simplistically in the models used for decision making, including in the popular GEP class of model.

The representation of uncertainty into energy sector models is then considered an important challenge, which must be addressed to ensure decision makers can rely on the results of these models. This thesis contributes to exploring one aspect of this challenge, the incorporation of long-term uncertainty into the GEP class of model across a number of different applications.

1.8 Thesis structure

This thesis explores the incorporation of long-term uncertainty into generation expansion planning models for system planning, market modelling, decarbonization, and renewable energy support policy evaluations. The thesis is organised as follows:

- The first chapter provides the background to the current challenges in the electricity system, with a particular focus on the role of uncertainty in decision making and modelling.
- The second chapter outlines the formulation of the GEP model that will be used

through the remainder of the thesis and describes the approaches to generating a representation of uncertainty and to including it into the model.

- The third chapter addresses one of the key challenges to incorporating a large number of scenarios into the GEP model; it addresses the representation of time and short-term system dynamics in the selection of representative days used in the model.
- The fourth chapter applies the GEP model to examine the costs of decarbonisation in a developing country, finding that these costs vary greatly depending upon the assumed scenario which illustrates the importance of uncertainty in GEP modelling.
- The fifth chapter explores further the interaction between different sources of uncertainty and the means for incorporating uncertainty into the GEP model for electricity market modelling. This chapter demonstrates the importance of both using the stochastic optimisation approach to GEP market modelling and incorporating a wide range of uncertainties simultaneously.
- The sixth chapter applies this insight about the need to incorporate long-term uncertainty in decision making models to an important and common GEP problem, the assessment of different renewable energy support policies.
- Finally, the seventh chapter draws conclusions from the thesis as a whole and discusses the direction of future work.

1.9 Scientific contributions of the thesis

Firstly, in the third chapter, the research extends the previous literature on the selection of representative days for power system expansion models. The main contributions of this work are:

- The development of a new methodology for adjusting the weights calculated by clustering algorithms for representative days before using them in the GEP model. The new approach has the advantage of ensuring that both the forecast annual energy requirements and relationship to peak value are represented in GEP and that ramping implications of increased renewable penetration are also accurately captured.

- The exploration of different approaches to representing historically observed data (i.e. separate time series for load and renewables or a combined time series for net load, the means for determining net demand, and the use of duration curves or ordered series) in the clustering algorithm and the finding of large differences in the quality of GEP modelling results based on this choice.
- The testing of the inclusion of a number of potential 'extreme points' or 'days of interest' to improve the representation of future flexibility requirements and determination of the benefits for accurately modelling the role of ES.

In the fourth chapter, I look to apply the GEP model to the important issue of quantifying decarbonization costs and, in doing so, explore the role of uncertainty. Research into the introduction of VRES and the quantification of the cost for reducing carbon dioxide (CO_2) emissions from electricity generation is typically based on case studies from the developed world, of which the West African region is particularly under served. This creates an important gap in the literature, as opportunities and challenges of contributing to the efforts to reduce carbon emissions may differ in these less developed areas. To address this gap chapter four provides:

- An application of the GEP model to the West African country of Ghana for the purpose of quantifying costs of contributing to decarbonization efforts. This represents an extension to the current literature due to the detailed incorporation of unit commitment and operational constraints. The formulation is also extended to account for the unique challenges of fuel switching, a common occurrence in Ghana where natural gas supply can be limited.
- An estimation of the cost of decarbonisation based upon achieving different CO_2 emission targets, and the role of uncertainty explored through six scenarios on key uncertainty inputs.

- Conclusions on the unique challenges and opportunities faced by developing countries in contrast to developed countries.

The work on Ghana identifies uncertainty as a key driver of results in the application of GEP models, however, the inclusion of this uncertainty is highly simplified in the form of individual scenarios where a single assumption is changed in each. In Chapter 5, I demonstrate the importance of the inclusion and careful treatment of uncertainty across all relevant inputs, in combination, as a stochastic optimisation. In the context of a United Kingdom case study we:

- Provide a simple method for generating uncertainty scenarios for GEP modelling when the modeller has only a single central forecast.
- Demonstrate and quantify the importance of including a wide range of sources of uncertainty in combination. In particular, I find that the often-ignored sources of economic uncertainty can be more important than the traditional uncertainty in respect to the physical quantity of energy required in future years.
- Find that it is important to include multiple sources of uncertainty in combination. I find that assessing each source individually can both underestimate the importance of uncertainty in that source and lead to investment decisions that actually increase costs.
- Demonstrate the benefits of stochastic optimisation for dealing with uncertainty in either normative or descriptive modelling. I find that the stochastic optimal solution differs significantly from the average of scenarios modelled individually.

Having identified the importance of including a wide set of sources of uncertainty with the stochastic optimisation approach, I apply this insight to a typical GEP application. In chapter six, I explore the inclusion and careful treatment of uncertainty across all relevant dimensions when assessing renewable energy policies. More specifically, in the context of a real-world generation expansion planning problem we:

- Compare the effectiveness of six different policy options; a cap on CO_2 emissions (as with a cap and trade scheme), a CO_2 price, a renewable capacity target, a green certificate scheme, a renewable generation subsidy, and a renewable capital grant under different modelling treatments of uncertainty.
- Demonstrate the importance of the inclusion of uncertainty into the GEP model for policy assessment, with many policies designed in a deterministic model not achieving target levels of carbon reduction or significantly underestimating costs.
- Demonstrate that common modelling approaches that attempt to capture uncertainty as multiple different individual scenarios (such as scenario analysis or Monte-Carlo simulation) perform poorly at representing the reaction of a competitive electricity market. These approaches can underestimate the expected carbon abatement response to decarbonisation policies and significantly underestimate costs.
- Demonstrate the importance of incorporating uncertainty into a model with stochastic optimisation for determining the technology expected to be built by a competitive market. I find a scenario analysis approach underestimates the quantity of conventional capacity that the market would construct without renewable support policies and underestimates the benefits of solar capacity compared to wind capacity in an uncertain market environment.

1.10 List of publications

The work developed during the course of this thesis resulted in a number of publications:

- **Scott IJ,** Carvalho PMS, Botterud A, Silva CA (2019). Clustering representative days for power systems generation expansion planning: Capturing the effects of variable renewables and energy storage. Applied Energy 2019; 253:113603. [50]
- Diawuo FA*, **Scott, IJ***, Baptista PC, Silva CA (2020). Assessing the costs of contributing to climate change targets in sub-Saharan Africa: The case of the Ghanaian electricity system. Energy for Sustainable Development, 57, 32-47 ***Joint first authorship**. [51]
- **Scott IJ,** Carvalho PMS, Botterud A, Silva CA (2021) Long-term uncertainties in generation expansion planning: Implications for electricity market modelling and policy. Energy [52]
- **Scott IJ,** Botterud A, Carvalho PMS, Silva CA (2019) Renewable support policy evaluation: The importance of uncertainty. AEAB2019-Applied Energy Symposium: MIT A+B Conference Proceedings. http://www.energy-proceedings.org/wp-content/uploads/2020/01/AEAB2019_paper_121.pdf [53]
- **Scott IJ,** Botterud A, Carvalho PMS, Silva CA (2020). Renewable energy support policy evaluation: The role of long-term uncertainty in market modelling. Applied Energy, 278, 115643. [54]

The model developed in this thesis also contributed to the following publication:

- Neves D, **Scott I**, Silva CA (2020). Peer-to-peer energy trading potential: An assessment for the residential sector under different technology and tariff availabilities. Energy, 118023. [55]

Chapter 2

Generation Expansion Planning Problem Formulation

In this chapter I discuss the model used throughout the thesis, with an emphasis on the means for generating and incorporating uncertainty.

2.1 Introduction

The aim of this thesis is to compare different treatments of uncertainty in a GEP model for planning, market modelling, and policy making. The methodology in this chapter covers the three steps in this process. Firstly, I discuss the means for including uncertainty in this model before providing the full model formulation. Then, I outline the methodology used for quantifying input uncertainty so that in can be represented in the GEP model using a scenario generation and reduction process.

2.2 Model treatment of uncertainty

The model is capable of incorporating three different representations of uncertainty, the distinction between which is discussed below:

Deterministic (D): A single deterministic scenario is optimized, where all inputs (for example demand, fuel prices, and technology construction costs) take a single value. When using the model in such a way the modeller is effectively ignoring uncertainty and using the expected future value for all inputs, including those for inputs for which it is clear that uncertainty exists.

Independent Scenarios (IS) or Scenario Average (SA): One simple and common means to incorporate uncertainty into energy models is by examining different scenarios [56]. In this approach uncertainty is represented as a set of scenarios where the values for the uncertain inputs differ between scenarios. All possible combinations of source of uncertainty (eg. fuel price forecasts, demand forecasts, and construction costs forecasts) are combined to create independent scenarios. In the case that there are 4 scenarios each for fuels prices, demand, and constructions costs, this would mean a total of 64 total scenarios. The model is then used to simulate each scenario independently, so in this case would produce 64 sets of results indicating what is expected to happen in each of the 64 scenarios. Modelling approaches with independent scenarios such as scenario analysis or Monte Carlo analysis are the most common method to incorporate uncertainty [4].

It is important to recognise that the investment decisions made in each scenario reflect the model having perfect foresight (no uncertainty) within that particular scenario. When modelling a liberalized market or a central utility for descriptive analysis or for policy making this implies that market participants do not take risk or uncertainty into account when making investment decisions (as the model is optimized to a known future). Instead here I am representing uncertainty as a range of different outcomes each determined without taking uncertainty into account. This process does generate a range of different results (one for each scenario), which is often used to represent the uncertainty of future outcomes. To generate results for comparison from this approach the expected result is taken by multiplying the results of each scenario by the probability of that scenario occurring[11].

Stochastic optimization (SO): When running the model as a stochastic optimization problem the model is forced to make a single set of decisions that must be fixed over all potential futures for a selected set of years (by implementing Equation (17)). This reflects real world decision making where investors must decide to make an investment without knowing what important factors like the future levels of demand, fuel prices, or

[11] It is important to note when comparing a treatment (such the introduction of a carbon tax) comparing the probability weighted average results before and after the treatment is the same as calculating the probability weighted average change created by the treatment. That is, the change in the average expansion plan is the same as the average change in expansion plans.

construction costs will be. While one set of initial build decisions are made, results are provided for each of the scenarios, where the level of generation, carbon emissions, and future build decisions will differ for each scenario. Therefore, the model needs to make investments without knowing what the future will be, but the way these investments operate will depend on what future scenario actually occurs and, in this thesis, typically the expected operation output will be examined. As discussed in depth in Section (1.3) and in [10] and [15], under the assumption of perfect competition, the actions of individual investors in a competitive market can oftentimes be modelled as a centralized least-cost generation expansion planning problem, including in the particular case where investors are making decisions on the basis of a stochastic optimization over uncertainty.

It should be noted that a two-stage stochastic optimisation has been implemented, where only the initial set of build decisions must be made without knowledge of the final values for the uncertain parameters. In reality all decisions on the powerplant to build at all points in time should be made in this way (which would be modelled as a multistage stochastic optimisation). However, due to complexity limitations, such a model will be unsolvable with the wide range of sources of uncertainty considered here. Consequently, the results found here can be expected to underestimate the impact of uncertainty.

It is important to be clear about the interpretation of the stochastic optimization in this context. One potential application of the SO GEP for the policy maker could be to find the decarbonization support policy that is optimal given uncertainty (as in [57]), this would be a normative perspective. Instead, the question being investigated is: given that investment decisions are made taking uncertainty into account (without the perfect knowledge of the future), potentially in a decentralised market, what then is the expected effect (in terms of outcomes in the market) of different decarbonization policies?

The stochastic optimisation model is extremely expensive from the computational point of view, and in practice scenario analysis is a more common approach to attempting to capture uncertainty [7], [4]. However, the interpretation of the scenario probability weighted average build decision, or the average effect of a new treatment such as a new energy policy, based upon the modelling of independent scenarios is less straightforward. Given the popularity of GEP modelling for policy decisions, and the popularity of scenario analysis or Monte Carlo simulation, one can imagine this being used as an estimate of the expected build given uncertainty. An example would be a case of introducing a new decarbonisation policy, the effect of this policy is evaluated over a number of scenarios (1000's in the case of Monte-Carlo) and the average effect of the decarbonisation policy reported. Contrasting this to the SO solution, it is important to note that in each of the simulations, the model has perfect knowledge of the future. Figure 2 provides an illustrative example[12] of the difference in how uncertainty is treated in the three modelling approaches and the potential implications for results.

[12] Note that these are fictional illustrative results, however, they are representative of the results found in Chapters 5 and 6. It should also be noted that the descripted labelling of scenarios as "high", "medium", and "low" fuel prices is for simplicity of explanation and is not reflective of the scenarios produced by the scenario generation and reduction process used in this thesis (which may select, for example, a fuel price which is high in the short term but low in the long time in the set of representative scenarios).

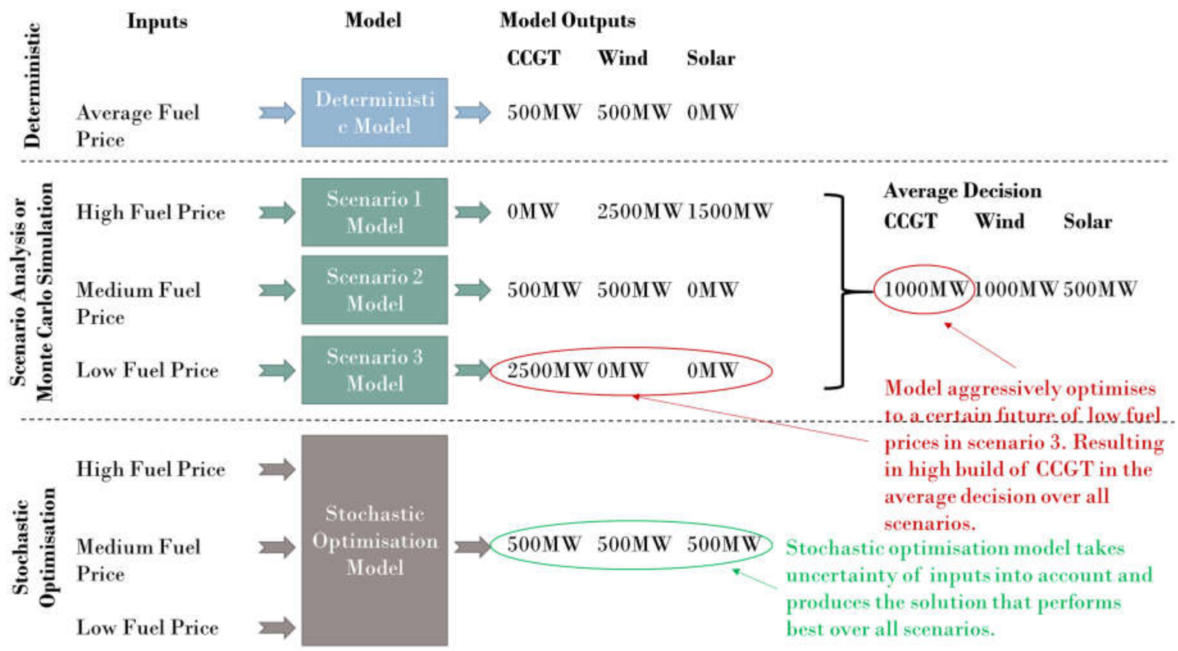


Figure 2 Illustrative example of the different modelling approaches' treatment of uncertainty and the potential implications for results. The model determines the optimal build of three technologies: CCGT, Wind, and Solar, with fuel price as an uncertain input.

In Figure 2 it can be seen that using the expected fuel price value in the deterministic model will result in 500MW of new CCGT and 500MW of new Wind selected as the predicted level of future build. When examining the scenario analysis, however, these results are sensitive to the level of fuel prices. For a lower fuel price that model builds extensive CCGT (2500MW), potentially to replace existing expensive sources of energy such as diesel generators or imported electricity. On the other hand, expensive fossil fuels (in the high price scenario) leads to significant investment in the renewable technologies of Wind and Solar. When the model has perfect knowledge of the future it can aggressively optimise to each situation, which means even on average there is significantly more build (1000MW CCGT, 1000MW wind, 500MW solar). Examining the expected expansion plan over all scenarios the capacity is higher than the deterministic case with 1000MW of CCGT, 1000MW of wind and 500MW of Solar. Finally, the stochastic optimisation model considers all three fuel price scenarios, but selects the expansion plan that would be optimal given that future fuel prices are uncertain. This expansion plan differs to both of the previous results with 500MW of each technology built. This result would indicate that mitigating the costs of the high fuel price scenario justifies the construction of Solar capacity despite this not being justified in the average scenario. However, the high level of build found by averaging the results of individual scenarios is not justified overall. For example, while 2500MW of CCGT may be optimal with low fuel prices the costs of the CCGT are such that if there are high or medium fuel prices these costs appears to outweigh the benefits of this technology, and a more balanced mix of technologies is found to optimise cost given uncertainty in the stochastic optimisation result.

2.3 Generation Expansion Planning Model

In this section, details of the Mixed Integer Linear Programming (MILP) GEP model[13] are provided focusing on the application of the models to the studies included in this thesis. As discussed previously one of the issues with incorporating uncertainty into the model is the increase in problem size and complexity as we increase the number of scenarios included. In order to achieve the goals of this thesis then I focus on model formulations that minimise problem size while retaining representativeness. The approach follows the finding of [10], [58] in determining the trade-offs between computational complexity and the ability of the model to represent a real world power system. The specific insights I take from the literature are:

[13] The model is implemented in C# and utilises the GUROBI mathematical programming solver.

- As with the widely applied LUSYM [59] and MIT GenX [60] GEP models the model here features a clustered unit commitment formulation, which has shown to result in relatively small errors when similar units are clustered together (0.06% in terms of objective function value which is below the typical MIP Gap of 0.1%) in exchange for significant reduction in problem size [61].
- An advantage of the clustered unit commitment formulation shown in [10] and [58] is that the integer commitment decisions can be relaxed with little effect on solution quality (both in terms of objective function value and selected technology mix). In Figure 3 I demonstrate the intuition behind this result. In Figure 3a I demonstrate how 3 CCGT units which would be modelled with binary commitment decisions can be combined into a single unit clustered unit commitment formation with an integer commitment variable. Figure 3b illustrates what happens when the unit commitment variables are relaxed in a situation where only a small amount of capacity is required to be online. In a standard binary model, all three units can violate the binary commitment requirement, however, with the clustered unit commitment formulation at most one unit per technology will be non-integer. This result has implications for start costs, minimum stable levels, reserves provision, and unit ramping. However, both [10] and [58] emphasise that while unit commitment decisions can be linearized, excluding unit commitment constraints altogether does adversely affect results.
- The level of detail included in the unit commitment constraints is also important. In a systematic investigation [10] the author finds a number of constraints found in typical short term unit commitment models add little to the solution quality in long term GEP models which I then exclude here. Specifically, the author finds little benefit from differential power plant efficiency at different load levels, additional ramping costs, minimum up time, minimum down time, minimum start-up time constraints, or start-up or shutdown range constraints. While the

author finds only a small but significant difference for the inclusion of start costs, I continue to include these here, using the same formulation as [10], which does not require an additional start variable but determines starts from the difference in units operating between periods. The author also finds that the type and number of reserves can play an important role in ensuring representative results. Here I only include one form of reserves (nominally spinning reserves), however, the implementation of this reserve matches the more restrictive reserve type used by the author in that it is restricted both by available ramping capability and headroom (not headroom alone). A GEP model based on these simplifications of a detailed unit commitment formulation is concluded by [10] to be "*more than accurate enough*". Removal of these constraints explains much of the difference between the model here and the LUSYM model [59] which was developed by the same author as [10].

- In [58] the authors demonstrate that the incorporation of unit commitment constraints and the contribution ES can make to ramping[14] and reserve challenges into the GEP is important for valuing the contribution of ES. However, the authors also emphasize the role of assumptions as typically having as high or higher an impact on results than the formulation.

[14] As with the application of the models in this thesis battery ES is not considered to be ramping constrained.

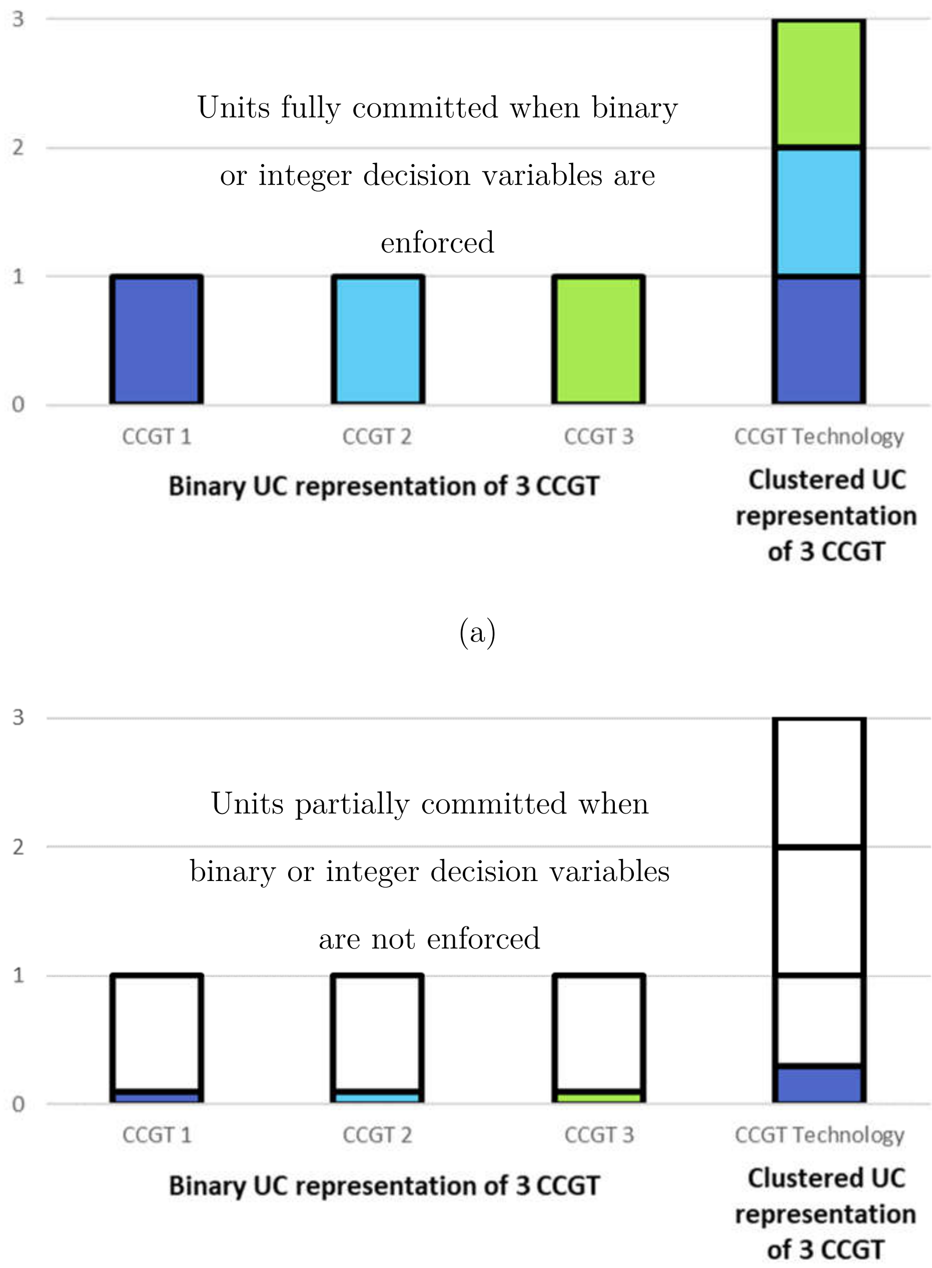


Figure 3 Illustration of clustered unit commitment formulation (a) binary commitment variables for 3 CCGT units vs clustered representation (b) linear unit commitment variables for 3 CCGT units vs clustered representation.

2.3.1 Model objective

Equations (1) - (3) present a simplified view of the objective function of the GEP optimisation under the three means of incorporating uncertainty. Broadly, the model attempts to minimise the expected value of all costs. To do so in the case of the stochastic optimisation the model must optimise over all scenarios together. The term $W_s^{Scenario}$ determines the probability of each scenario (the weight) and is used to trade-off between the costs of the different scenarios (a highly improbable scenario has relatively little weight in determining the outcome). When modelling the deterministic problem, however, only a single scenario is used and W_s^S takes a value of 1 and the model is greatly simplified. Solving the model as independent scenarios the objective function minimises costs for each scenario independently as shown in Equation (2).

Deterministic:

$$min(BC + FC + VC + SC) \tag{1}$$

Scenario Analysis:

$$min\, W_s^S\ (BC_s + FC_s + VC_s + SC_s)\ \ \forall\, s \in S \tag{2}$$

Stochastic Optimisation:

$$min \sum_{s \in S} W_s^S\ (BC_s + FC_s + VC_s + SC_s) \tag{3}$$

The costs considered in the objective include build costs BC_s for new units (Equation (4)), annual fixed costs for all units FC_s (Equation (5)), variable generating costs VC_s (Equation (6)) and finally system costs SC_s including any CO_2 emission price costs and other penalty costs such as unserved energy or reserve violations (Equation (7)).

From Equation (4) the build costs per scenario BC_s is simply the discounted sum of all individual units cost to build at each point in time $\mathrm{c}_{\mathrm{syp}}^{\mathrm{B}}$. The fixed costs are determined by the number of built units $\mathrm{b}_{\mathrm{syp}}$ in each year and the fixed cost for that unit type $\mathrm{C}_{\mathrm{p}}^{\mathrm{F}}$. Variable costs VC_s are defined in (Equation (6)) where $\mathrm{C}_{\mathrm{syp}}^{\mathrm{g}}$ is the cost of generation for each type of generator at given point in time and $\mathrm{g}_{\mathrm{sydtp}}$, $\mathrm{C}_{\mathrm{syp}}^{\mathrm{o}}$the cost for online operating units (including variable costs to be operating at the units minimum stable level) and

o_{sydtp} the number of operating units for each technology, and $C_p^{rc} \cdot rc_{sydtp}$ are the cost to curtail renewable energy and the quantity curtailed respectively, finally the cost of emissions C_y^e and the amount of emission e_{sydt}^{CO2}. Again variable costs must be weighted both by the number of years being represented by each modelled year DF_y^S and the number of days being represented by each modelled day W_d^D. Equation (7) defines system costs as comprising the cost of unserved energy C^{UE} and the quantity of unserved energy v_{sydt}^{UE}, the cost of violating the spinning reserve up constraint C^{rU} and the under provision of this reserve, the cost of violating the spinning reserve up constraint C^{rD} and the underprovision of this reserve v_{sydt}^{rD}.

$$BC_s = \sum_{y \in Y} \sum_{p \in P} DF_y \cdot c_{syp}^B \tag{4}$$

$$FC_s = \sum_{y \in Y} \sum_{p \in P} DF_y^S \cdot C_p^F \cdot b_{syp} \tag{5}$$

$$VC_s = \sum_{y \in Y} DF_y^S \cdot \sum_{d \in D} W_d^D \cdot \left(\sum_{t \in T} \sum_{p \in P} C_{syp}^g \cdot g_{sydtp} + C_{syp}^o \cdot o_{sydtp} + c_{sydtp}^S + C_y^e \cdot e_{sydtp}^{CO2} + C_p^{rc} \cdot rc_{sydtp} \right) \tag{6}$$

$$SC_s = DF_y^S \cdot \sum_{y \in Y} \sum_{d \in D} W_d^D \cdot \left(\sum_{t \in T} C^{UE} \cdot v_{sydt}^{UE} + C^{rU} \cdot v_{sydt}^{rU} + C^{rD} \cdot v_{sydt}^{rD} \right) \tag{7}$$

In this cost minimisation formulation, the model represents a classic GEP problem for each individual scenario. However, as discussed in the introduction and in [22], a perfectly competitive market will achieve the same outcomes found as the cost minimisation under the common GEP assumption of inelastic demand. That is, the perfectly competitive market can be expected to achieve the welfare maximising outcome found by the GEP and conversely, the GEP provides a simple model of a perfectly competitive electricity market.

When the model is extended to the two-stage stochastic optimisation, where all scenarios are solved together as in Equation (3), with the use of non-anticipativity constraints, we can retain a market interpretation. In [21] the authors demonstrate, again under the condition of inelastic demand, that a perfectly competitive market where participants optimise over the same uncertainty assumptions, can be modelled as a two stage GEP stochastic optimisation problem. In a more recent work [15], the authors extend this result to show that even with heterogenous risk preferences, as long as risk trading is possible, the perfectly competitive market can again be represented by the inclusion of risk constraints to the two stage stochastic optimisation model.

Of note are the weights and discount factors used in this equation. For each scenario $\forall\, s \in S$, a scenario weight W_s^S is applied, representing the probability of that scenario occurring (see section 2.5for details). The model is run with a limited set of representative days $d \in D$, each having a set of chronological time periods $t \in T$, and an associated weight W_d^{Day} (broadly the number of days that each 'representative' day is representing) such that $\sum_{d \in D} W_d^{Day} = 365$. See Chapter 2 for in depth discussion of the methodology used to select representative days.

Representative days with chronological time periods have been shown to be an adequate approach to model high levels of renewable penetration [62] while minimising problem size.

The discount factors DF_y and DF_y^{Step} are used to discount future build and annual costs, depending on the number of years included in each time step[15]. For example, in the model applied in Chapter 5 only every fifth year is included (as a set of representative days) but weighted to reflect that it is representing a 'step' of five operational years with the discount factor DF_y^{Step}. If we assume an annual discount rate of df and a step size of 5,

[15] If each year of the planning horizon is modelled, that is a step size of 1, then DF_y is equal to DF_y^{Step}.

the annual factor DF_y and step discount factor DF_y^{Step} are calculated as follows:

$$DF_y = (1 + dr)^{-(y-1)} \tag{8}$$

$$DF_y^S = \sum_{i=y}^{y+4} DF_i \tag{9}$$

2.3.2 Model constraints

As discussed in the introduction the advantage of the GEP model is its ability to represent the operation of the power system in detail. In this section I present the constraints used to do so.

This formulation allows the adoption of a "clustered unit commitment" approach to implementing different technologies within the model which has been found to be highly efficient for GEP modelling [10].

2.3.2.1 System-wide constraints

$$\sum_{p \in P} g_{sydtp} + \underline{G_p} \cdot o_{sydtp} + v_{sydt}^{UE} = D_{sydt} \quad \forall s \in S, y \in Y, d \in D, t \in T \tag{10}$$

$$\sum_{p \in P} r_{sydt}^{U} + v_{sydt}^{rU} \geq R_{sydt}^{U} \quad \forall s \in S, y \in Y, d \in D, t \in T \tag{11}$$

$$\sum_{p \in P} r_{sydt}^{D} + v_{sydt}^{rD} \geq R_{sydt}^{D} \quad \forall s \in S, y \in Y, d \in D, t \in T \tag{12}$$

Equation (10) is the supply-demand constraint allowing for unserved energy, this equation ensures the demand of electricity D_{sydt} at each point in time is met by generation g_{sydtp} or the baseline level of generation $\underline{G_p}$ provided by each operating unit o_{sydtp}. However, the model allows for unserved energy v_{sydt}^{UE} in extreme cases, this variable is typically highly penalised in Equation (7). Additionally, the dual value of this constraint provides an estimate for the expected electricity price (the marginal cost of electricity). Equation (11) and Equation (12) are the spin up and spin down constraints respectively, ensuring spinning reserve requirements R_{sydt}^{U} and R_{sydt}^{D} are met at each point in time from reserves provided by generating units r_{sydt}^{U} or r_{sydt}^{D}, each allows for violation of the respective requirement, v_{sydt}^{rU} or v_{sydt}^{rD}, at a penalty price included in Equation (7).

2.3.2.1 Build constraints

$$o_{sydtp} \leq \begin{pmatrix} b_{syp} & p \in PB \\ B^{E} & p \notin PB \end{pmatrix} \quad \forall p \in PB, s \in S, y \in Y \tag{13}$$

$$b_{syp}, B^{E} \quad \in \mathbb{Z}_0^{+} \tag{14}$$

$$c_{syp}^{B} \geq \begin{pmatrix} C_{syp}^{\boldsymbol{B}} \cdot \left(b_{syp} - b_{sy-1p}\right) & y < L_p \\ C_{syp}^{B} \cdot \left(b_{syp} - b_{sy-1p} + {c_{sy-Lp}^{B}}/{C_{sy-Lp}^{B}}\right) & y \geq L_p \end{pmatrix} \quad \forall s \in S, y \in Y, p \in PB \tag{15}$$

$$b_{syp} \leq \overline{B}_p \quad \forall s \in S, y \in Y, p \in PB \tag{16}$$

$$b_{syp} = b_{s-1yp} \quad \forall\, s > 0 \in S, y \in Y^{NA}, p \in PB \tag{17}$$

Equation (13) ensures that units must be built to operate by constraint the unit's operating variable $\boldsymbol{o_{sydtp}}$, in the case of existing units the build decision variable $\boldsymbol{b_{syp}}$ is replaced with the number of existing units B^{Exist}, in Equation (14) both $\boldsymbol{b_{syp}}$ and B^{Exist} are restricted to positive integer variables. Equation (15) ensures the build cost $\boldsymbol{c_{syp}^{B}}$ is paid for each new build unit in the model depending of the cost to build that unit $\boldsymbol{C_{syp}^{B}}$. Additionally, this constraint ensures that each new unit is only included in the model for its lifetime (or a build cost can be incurred to build a replacement unit).

Equation (16) limits the number of units built for any given technology by the parameter $\overline{B}_p$. Typically, equation (16) is set at a value such that this constraint does not bind. The exceptions are when there is a limited quantity of a new build capacity that can be added to the system (for example the number of sites for new large scale hydro is limited) or when a supply curve is modelled which increases the cost of capacity as more is added to the system.

The inclusion of the non-anticipativity constraint Equation (17) enforces the inclusion of decision making with uncertainty in the case of stochastic optimisation. The non-anticipativity constraint ensures the model is constrained to make one set of build decisions that apply across all scenarios for the non-anticipativity years. In this way the model must take uncertainty into account (by optimising over all possible futures) without knowing which future scenario would occur. This contrasts to the scenario analysis approach, where several future scenarios are modelled, but in each case the model optimises to the values known for that future scenario.

2.3.2.2 Conventional operation constraints

$$o_{sydtp} \in \mathbb{Z}_0^+ \quad \forall p \in PC, PH, s \in S, y \in Y \tag{18}$$

$$c_{sydtp}^S \geq C_p^S \cdot (o_{sydtp} - o_{sydtp-1}) \quad \forall p \in PC, PH, s \in S, y \in Y, d \in D, t \in T \tag{19}$$

$$0 \leq g_{sydtp} + r_{sydt}^U \leq \left(\overline{G_p} - \underline{G_p}\right) \cdot o_{sydtp} \quad \forall p \in PC, PH, y \in Y, d \in D, t \in T \tag{20}$$

$$0 \leq r_{sydt}^D \leq g_{sydtp} \quad \forall p \in PC, PH, y \in Y, d \in D, t \in T \tag{21}$$

$$g_{sydtp} + r_{sydt}^U \leq RL_p^U \cdot o_{sydtp} \quad \forall p \in PC, PS, PH, s \in S, y \in Y, d \in D, t \in T \tag{22}$$

$$g_{sydtp} - g_{sydtp-1} - r_{sydt}^D \geq RL_p^D \cdot o_{sydtp} \quad \forall p \in PC, PS, PH, s \in S, y \in Y, d \in D, t \in T \tag{23}$$

$$ER_p^{CO2} \cdot \left(g_{sydtp} + \underline{G_p} \cdot o_{sydtp}\right) = e_{sydtp}^{CO2} \quad \forall p \in PC, PH, s \in S, y \in Y, d \in D, t \in T \tag{24}$$

Equation (18) ensures that units must be either committed or not committed in full. This constraint is used to ensure that both start costs are fully accounted for and the generator minimum stable level is respected. However, following the analysis in [10] I relax this assumption in a number of cases as it greatly increases the size of the problem without adding significantly to the representativeness of results in the clustered unit commitment formulation. The relaxed model is referred to as the linear relaxation model. It is important to note that only the integer constraint on commitment decisions is relaxed and the build of new units remains an integer decision.

Equation (19) ensures a cost c_{sydtp}^S must be paid when a unit is started, this cost depends on the cost to start a unit of the given type C_p^S. Equation (20) limits the combined provision of generation and ramp up services to the maximum capacity $\overline{G_p}$ less the operation at the minimum stable operating level of the unit $\underline{G_p}$, while ensuring that generation and reserves are provided only when units are committed. The maximum generation level per unit $\overline{G_p}$ is not a nameplate capacity but instead a de-rated capacity level representing average availability (as in the capacity is de-rated for the expected proportion of the time that capacity would be unavailable due to planned or unplanned outages). This provides a stylized method for incorporating the effect of outages as in [10].

Equation (21) ensures spin down reserves are only provided where operational room to the generator minimum stable level is available. Equations (22) and (23) limit the change in generation to that possible given the number of online generator units and the ramp rate in the upwards RL_p^U and downwards RL_p^D directions. Provision of reserves must also be within the ramping limit given the change in generation. It should be noted, following [10] I model a linear relaxation of operation decisions and do not enforce that the variable o_{sydtp} is an integer variable. Equation (24) determines the CO_2 emissions e_{sydtp}^{CO2} from each unit based on their emission rate ER_p^{CO2} and generation.

2.3.2.3 Energy storage constraints

$$0 \leq g_{sydtp} + r_{sydt}^U \leq \overline{G_p} \cdot o_{sydtp} \quad \forall p \in PS, s \in S, y \in Y, d \in D, t \in T \tag{25}$$

$$0 \leq c_{sydtp} \leq R^{cg} \cdot \overline{G_p} \cdot \left(b_{syp} - o_{sydtp}\right) \quad \forall p \in PS, s \in S, y \in Y, d \in D, t \in T \tag{26}$$

$$0 \leq es_{sydtp} \leq \overline{ES} \cdot b_{syp} \quad \forall p \in PS, s \in S, y \in Y, d \in D, t \in T \tag{27}$$

$$EF_p^c \cdot c_{sydtp} - g_{sydtp} = \begin{pmatrix} \Delta es_{sydtp} & t > 0 \\ es_{d,t=0,p} - es_{d,t=T,p} & t = 0 \end{pmatrix} \tag{28}$$

$$0 \leq g_{sydtp} + r_{sydt}^U \leq es_{sydtp} \quad \forall p \in PC, d \in D, t \in T \tag{29}$$

Equation (25) limits the combined provision of generation and ramp up services provided by the storage unit. Equation (26) limits the charge c_{sydtp} of the ES technology based on the maximum charging rate (defined relative generation capacity with the ration R^{cg}) and the number of units built but not currently generating. Equation (27) then limits the stored energy es_{sydtp} given the number of installed units and the energy storage of each unit $\overline{\mathrm{ES}}$. Equation (28) defines the level of energy stored at any given time incorporating the round-trip efficiency of cycling the ES $\mathrm{EF}_{\mathrm{p}}^{\mathrm{c}}$. A daily cycling limit is enforced in Equation (28) by ensuring that the unit starts and finishes that day with the same energy level. This decision reflects the use of representative days in the model, which imply that the possibility of moving energy between days is not considered. Equation (29) ensure sufficient energy for provision of both generation and spin up reserve.

2.3.2.4 Variable renewable energy constraints

$$g_{sydtp} - rc_{sydtp} = CF_{sydtp} \cdot b_{syp} \quad \forall s \in S, y \in Y, d \in D, t \in T, p \in PR \tag{30}$$

$$rc_{sydtp} \geq 0 \quad \forall s \in S, y \in Y, d \in D, t \in T, p \in PR \tag{31}$$

Equation (30) ensures renewable generation is determined by the input capacity factor $\mathrm{CF}_{\mathrm{sydtp}}$, unless a curtailment action is undertaken $\mathrm{rc}_{\mathrm{sydtp}}$ which is penalised in the objective function to reflect subsidies forgone by existing renewable generators. Curtailment is restricted to only reduce renewable generation (Equation (31)).

2.3.2.5 Large-scale hydro constraints

$$\sum_{d \in D} \sum_{t \in T} W_d^D \cdot \left(g_{sydtp} + \underline{G_p} \cdot o_{sydtp} \right) \leq L_p^H \cdot b_{syp} \quad \forall s \in S, y \in Y, p \in PH \tag{32}$$

$$\left(g_{sydtp} + \underline{G_p} \cdot o_{sydtp} \right) \geq \underline{G_p^H} \cdot b_{syp} \quad \forall s \in S, y \in Y, d \in D, t \in T, p \in PH \tag{33}$$

In addition to Eq. (18) to (24) for the hydro units maximum annual production limit $\mathrm{L}_{\mathrm{p}}^{\mathrm{H}}$ is enforced based upon historically observed values in Eq. (32). In the Ghanaian system modelled all existing hydro units produce at a minimum level, to maintain river waterflow, and I capture this behaviour by applying a minimum production level at each point in time $\underline{\mathrm{G}_{\mathrm{p}}^{\mathrm{H}}}$ (Eq. (33)).

2.3.3 Modification to allow for natural gas constraints and generator fuel switching

$$\sum_{d\in D}\sum_{t\in T}\sum_{p\in P} W_d^D \cdot \frac{g_{sydtp}^f + \underline{G_p} \cdot o_{sydtp}^f}{EF_p^F} \leq Limit_y^f \quad \forall y \in Y, f \in F \tag{34}$$

$$\sum_{f\in F_p} o_{sydtp}^f \leq \begin{pmatrix} b_{yp} & p \in PB \\ B^E & p \notin PB \end{pmatrix} \quad \forall p \in PB, s \in S, y \in Y \tag{13F}$$

$$c_{sydtp}^S \geq C_p^S \cdot (o_{sydtp}^f - o_{sydtp-1}^f) \quad \forall p \in PC, f \in F, y \in Y, d \in D, t \in T \tag{19F}$$

$$0 \leq g_{sydtp}^f + r_{sydt}^U \leq \left(\overline{g_p} - \underline{g_p}\right) \cdot o_{sydtp}^f \quad \forall p \in PC, f \in F_p, y \in Y, d \in D, t \in T \tag{20F}$$

Equation (34) limits the annual availability for fuels across all units (accounting for the fact that the model uses representative days) to impose for example an annual limit on available natural gas as modelled in Chapter 4. Equations (13F), (19F), and (20F) replace Equations (13), (19), and (20) respectively to account for that fact that each multi-fuel generator can have different operating regimes with different fuels. These constraints ensure that a unit can only use a single fuel at a time and that it must first switch off and on again to switch fuels, incurring a start cost. However, as the formulation does not model generator up and down times this can happen between one period and the next.

2.3.4 Inclusion of renewable energy support policy constraints

$$\sum_{d\in D}\sum_{t\in T}\sum_{p\in PC} W_d^D \cdot e_{sydtp}^{CO2} \leq L_y^e \quad \forall y \in Y, s \in S \tag{35}$$

$$\sum_{d\in D}\sum_{t\in T}\sum_{p\in PR} W_d^D \cdot \left(g_{sydtp} + \underline{G_p} \cdot o_{sydtp}\right) \geq T_y^{RGP} \cdot \sum_{d\in D}\sum_{t\in T} W_d^D \cdot D_{sydt} \quad \forall y \in Y, s \in S \tag{36}$$

$$\sum_{p\in PR} \overline{G_p} \cdot b_{syp} \geq T_{sy}^{RC} \quad \forall y \in Y, s \in S \tag{37}$$

Equations (35) to (37) define the constraints used to implement decarbonization and renewable support policies. Equation (35) limits the total annual emissions of CO_2 with the parameter L_y^e, Equation (36) ensures a minimum proportion of renewable generation T_y^{RGP}, and Equation (37) ensures a minimum quantity of renewable capacity T_{sy}^{RC}. It is important to note that the policies are applied to each individual scenario. For example, in the case of an emission limit, the limit must be respected in every possible future scenario, as opposed to a limit that in expectation would be achieved but would not necessarily be achieved in every future scenario.

Additional renewable energy support policies considered simply alter the values of model parameters.

2.4 GEP model decomposition

One approach typical to allowing large optimisation problems to be solved is that of problem decomposition. Two common approaches to problem decomposition were tested for their ability to reduce problem complexity and allow for an increased number of scenarios in a single model: benders decomposition and progressive hedging.

The stochastic optimisation GEP problem with linear unit commitment constraints, is particularly well suited to the L-shaped decomposition, as what occurs in each scenario is effectively independent. This means all constraints relating to the operation of the system can be parallelized into sub-problems and only the build constraints remain in the master problem. The benders decomposition and acceleration techniques specific to the expansion planning problems of [63] were tested on the problems of Chapters 5 and 6. While this approach to model formulation significantly reduced solution time in a slight majority of cases it was found to be unreliable. In a sizeable minority of cases the master problem failed to converge in a reasonable timeframe despite testing a number of additional techniques to reduce the number of constraints included, group scenarios, remove unnecessary constraints, or provide a high quality starting set of constraints. In the situations where the master problem did not converge the quality of the solution (as measured by the objective function value) was typically low especially compared to the non-decomposed version of the problem that would also not achieve the target MIP gap. In addition, even where the benders decomposition completed, the solution often had a higher objective function value than the non-decomposed problem (although within the bounds of the given MIP-gap).

An alternative decomposition method tested was that of progressive hedging [64]. Under the progressive hedging approach to stochastic optimisation each scenario (or groups of scenarios) are solved independently, then where decisions differ across scenarios penalty terms are added to each independent optimisation problem to force them to converge together. These steps are then repeated in an iterative process until the scenario problems converge on variable that must match across scenarios. In the stochastic optimisation GEP problem this is only the first stage build decisions which is only a very limited set of the total number of decision variables. An important consideration in the implementation of this algorithm is the determination of the penalty term to apply to first stage variables that differ across scenarios, and several approaches from [65] were tested as well as different approaches to grouping scenarios into subproblems.

Unfortunately one issue with the progressive hedging algorithm is that in the case of integer variables it is not guaranteed to converge to the optimal solution [64] and this was found to be the case in a small but significant number of test cases here. The solutions found in Chapters 5 demonstrate how the SO optimal solution across all uncertainty is often very different to the solution found from solutions where uncertainty is included on one dimension at a time. This situation presents a particular problem for the progressive hedging algorithm, as this approach is effectively trying to combine different solutions into a single solution. Particularly, progressive hedging approaches which try to speed up the progressive hedging algorithm by "locking" decision variables that remain unchanged over a number of iterations will have problems in this case. An illustrative example could be a case in which given demand uncertainty the model would build CCGT, but given fuel cost uncertainty the model would build biomass capacity. Assume that in this case under combined demand and fuel uncertainty, solar capacity would be preferred. This could present a difficult situation for the progressive hedging algorithm. Basically, with each iteration the algorithm will penalise CCGT in the case of demand uncertainty (and increase the value of biomass) and the reverse in the subproblem with biomass uncertainty. It is possible that these penalisations or benefits will result in solar capacity being selected but it is also possible a solution which is a mix of CCGT, and biomass is found (that is inferior to adding solar capacity but is identical across sub problems). This issue does not occur in the case of linear variables and is sensitive to the approach selected to applying penalty weights to the subproblems or speeding up the progressive hedging algorithm.

2.5 Scenario Generation and Reduction

In this section I present the methods used to generate scenarios used to represent uncertainty in the model. Firstly, a modeller must decide if a given source of uncertainty can be described by a continuous random variable (ie. that it is an aleatoric source of uncertainty). This is often the case, however, optimisation to a continuous variable is typically computationally difficult and the modeller will instead aim to work with a discrete approximation in optimisation models [66]. In the case of SO we are also typically limited in the number of discrete scenarios that can be included in the model and so wish to select a minimal but representative set of discrete scenarios to represent the uncertainty in the continuous random variable. This is the aim of the scenario generation and reduction process.

I use the following high level procedure for scenario generation and reduction (broadly following [67]):

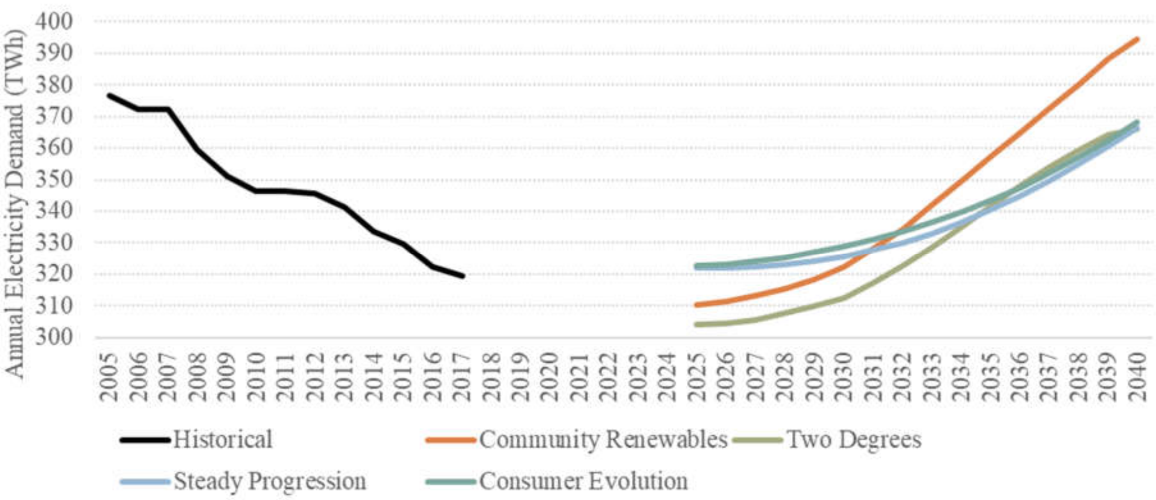
1. Fit a statistical model to the available historical data for a given uncertain input (e.g. energy demand). This fitted model will be used as the data generating process.
2. Generate a range of future scenarios based upon the fitted statistical model. This range of scenarios provides a discrete approximation of the future distribution of the input of interest.
3. Reduce this scenario set to a limited 'representative' set of scenarios based upon a scenario reduction methodology.

However, I modify this standard process to account for the fact that a number of important variables for the GEP problem oftentimes are expected to greatly deviate from their recent historical trends. Two examples, in the form of total electricity demand in the UK and EU Emission Trading Scheme (ETS) CO_2 prices, are presented in Figure 4 [68]. It can be clearly seen that fitting a reasonable statistical model to the historical data would result in a forecast that follows historical trends of decreasing demand and ETS prices. However, more sophisticated forecasts based upon wider domain knowledge (such as plans to increase the electrification of heating and transport and upcoming developments in the ETS scheme) all demonstrate the opposite trend. Ideally, the GEP modeller would like to take advantage of these more sophisticated forecasts that incorporate domain knowledge while being able to derive a given number of future scenarios with a probability weighting for each scenario (as is produced by a typical scenario generation and reduction process). To achieve this goal, I propose the following process:

1. First, a statistical process that contains both a trend and volatility component is fit to the historical data (Geometric Brownian Motion as is common for the GEP problem [69], however a Ornstein Uhlenbeck (OU) process would provide a good stationary alternative [70]).
2. The same process is then fit to the exogenous expert forecast the modeller wishes to base their modelling on.

3. To generate the future scenarios, the parameter value found for historical volatility are used as the volatility parameter and the trend parameter found in fitting the process to the exogenous forecast are used as the trend parameter.
4. As with the typical scenario generation and reduction exercise, a range of future scenarios (at least 1,000 [71], here 10,000) are then generated with the combined statistical model. This range of scenarios represents a discrete approximation of the distribution of future forecast values (based on historical volatility and expected trend).
5. A scenario reduction exercise is then undertaken to select a representative set of scenarios, each with an estimation of the probability of that scenario occurring (the proportion of underlying scenarios that each is representing), see [67] for more detail. Clustering algorithms such as the partitioning around the medoid (PAM) [72] clustering algorithm can be used for scenario reduction, however, simple greedy algorithms such as fast forward selection (FFS) [73] are often applied to power system problems [74].

The above process is repeated for each source of uncertainty. The inputs for which this method is applied (annual demand, natural gas prices, renewable and battery storage construction costs, and ETS CO_2 price) are discussed in section 5.1.



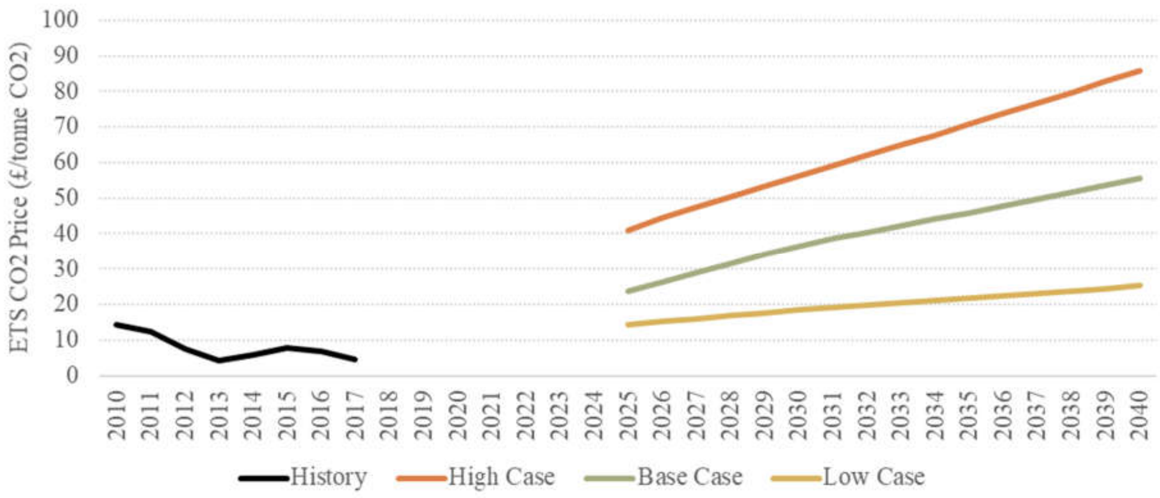


Figure 4 Historical values and forecast scenarios for total electricity demand and ETS CO_2 prices adapted from [68]. Scenarios are named based on their underlying assumptions, see [68] for further details.

It is important to note two assumptions inherent in this methodology. Firstly, the approach assumes that future volatility will be of a similar magnitude to historical volatility. This assumption may seem strong for certain inputs, for example where historical data is relatively sparse and volatility may not be accurately captured. The modeller may wish to test the resilience of conclusions to relaxing this assumption, by for instance increasing the volatility in the scenario generation process where inputs are expected to become more volatile. To test the conclusions in this chapter, I also use an alternative set of scenarios not derived from this methodology but taken directly from the literature.

Secondly, this methodology assumes that each source of uncertainty is independent. For correlated sources of uncertainty, the scenario generation and reduction process should be altered to generate correlated samples and maintain covariance in the reduction process, for example as in [67].

Chapter 3

Clustering representative days for power systems generation expansion planning: capturing the future system challenges

This chapter extends previous work on the selection of representative days for power system expansion models.

3.1 Introduction

Power systems are changing at an accelerated pace with the introduction of ever greater proportions of renewable energy, smart grid technologies, electric vehicles, and the wider availability of affordable energy storage (ES). These changes, particularly to the types of technologies used in the production of power, are forecast to introduce a number of new challenges particularly around the area of flexibility [9]. Decision makers, such as energy companies, power system operators, and government policy makers, are then faced with an important problem: how to make decisions on the direction the power system should evolve to meet these challenges without having perfect knowledge of the scope of the challenge or the costs and benefits of potential future technologies.

To make these decisions we turn to the use of models, simplified representations of the system, to forecast and assess different future investment options. However, the future challenges are predicted to be different and more complicated than those seen historically [8]. For example, the problem of integrating variable renewable energy, and the demand for flexibility it places on the system, increases with the proportion of renewable energy in the system and will therefore be significantly higher in the future [8,9]. This flexibility requirement has important implications for the optimal build and operation of all units on the system [75,76]. The study of generation expansion with high levels of renewable energy is a particularly active area of study (see [8] [35] for reviews), and of increasing interest is the role that storage can play [4]. To arrive at robust results in these types of studies, we need to ensure that models used for decision making contain the detail required to represent these future challenges.

When applying a model, we are limited in our ability to add complexity, such as future flexibility requirements, before the model becomes unwieldy. An important modelling task is then to select which areas of the real system we can simplify or abstract out. As the future system challenges require further detail in some areas of the model, reducing the scope in other areas may be required. A common area to reduce detail is the representation of time (the level of temporal granularity) included in the model to allow more complex incorporation of other factors such as operational details.

A trade-off then exists between model detail or model complexity and the level of temporal granularity included and this trade-off is an active area of research. In [77] the authors explored the trade-off between spatial and temporal granularity in a planning model for electricity expansion in California. The authors found the largest change in cost was associated with changes in spatial detail, which suggested a preference for spatial detail over temporal detail in model formulation. On the other hand [62] examines the trade-off between the level of temporal detail and the level of operational detail in energy system planning models and finds that temporal detail, and the method of selecting the temporal resolution, to be important for results in the high renewable penetration systems of the future. Further, an increasingly important trend is the use of stochastic optimisation to attempt to deal with uncertainty in planning model inputs. To facilitate this, [78] explores the trade-off between the level of stochastic and temporal variability and concludes that temporal detail should be sacrificed to improve the level of stochastic representation.

The amount of temporal detail we can include in any model is then severely limited by our need to include detail in other areas. The challenge is to ensure that sufficient temporal detail is included so that the model is still representative. In a GEP model this is complicated by the fact that many of the important challenges of the future, such as flexibility provisions [75] [79], demand shifting and storage operations [77] [79] [80] are dependent on the representation of time. For example a detailed representation of flexibility in the model operational detail (such as ramping and start-up constraints) cannot be expected to produce accurate results without the inclusion of the temporal detail that generates the flexibility challenges as demonstrated in [12].

The task of selecting the level of temporal detail is an important one and is increasingly attracting academic attention. A common initial approach had been to use a number of representative individual periods ([79] [81] [82] for example), often calculated from a load duration curve, which minimised the computational effort [82]. This approach is particularly common in energy system models [83] (although work has been done to extend this approach [80] in this context). The problem with this approach is that the importance of the transition between time periods is lost; on the short-term ramping scale, the daily storage operation scale, and on the longer scales of weekdays vs weekend or summer vs winter.

To preserve these relationships, a number of studies then instead selected larger blocks of representative time periods; for example either representative days [75] [80] [84] [85], or representative weeks [86].

An increasingly popular means to select which periods should be included as representative is that of clustering, where observed historical data is grouped and a representative period (hour, day, or week) is selected to represent each group. A number of different algorithms are used including k-means, k-medoids, principle component analysis, and hierarchal clustering (see [87] for overview and [88] [89] for comparison in the application to this task).

This chapter extends previous work on the selection of representative days for power system expansion models. The work seeks to ensure that the future challenges of providing the flexibility requirement for the incorporation of variable renewable energy sources and the role of ES is accurately represented for the purpose of long-term planning. Towards this end, the main contributions of this work are that we:

- Develop a new methodology for adjusting the weights calculated by clustering algorithms for representative days before using them in the GEP model. The new approach has the advantage of ensuring that both the forecast annual energy requirements and relationship to peak value are represented in GEP and that ramping implications of increased renewable penetration are also accurately captured.

- Explore different approaches to representing historically observed data (i.e. separate time series for load and renewables or a combined time series for net load, the means for determining net demand, and the use of duration curves or ordered series) in the clustering algorithm and find large differences in the quality of GEP modelling results based on this choice.
- Test the inclusion of a number of potential 'extreme points' or 'days of interest' to improve the representation of future flexibility requirements and determine the benefits for accurately modelling the role of ES.
- Compare and assess the technology mix in generation expansion plan results based on the emphasis of different features of the data to explore the implications for future technology selection based on the accurate inclusion of future power system requirements (and the potential biases where these are ignored).

In the remainder of this chapter I outline the clustering methodology used for representative day selection (section 3.2), the different means of including input data and extreme points (3.3), the proposed weighting methodology (section 3.4), and the test GEP model and dataset (section 3.5). The chapter then follows with results (section 3.8) and draws conclusions on the best approach to selecting representative days for GEP (section 3.9).

3.2 Clustering Methodology

As is increasingly common for the task of selecting representative input data for GEP the methodology here makes use of clustering techniques. A number of possible clustering algorithm exist that could be applied to the data as overviewed in [87], with perhaps 'k means' the most common. An important consideration in the selection of algorithm is which output will be used in the GEP model. One possibility would be to use an average of all data in the cluster (a k-means approach is based on minimising the distance to this average). However, an average day load or wind profile will necessarily reduce the hour-to-hour volatility by smoothing [90]. Instead, it is preferable to use an observed day in the GEP model and so use a k-medoids approach where the clustering selects an actual observed day that minimises the distance to the other members of its cluster, the specific algorithm is based upon Partitioning Around Medoids [72]. This approach is validated by the finding that k-medoids clustering can reduce error in long-term planning studies [88] [91].

Of more importance than the clustering algorithm itself, typically, is the treatment and representation of the data to be clustered [89]. I test a number of different approaches to represent the data in the clustering; more specifically, I broadly test across two dimensions:

- **Representation of data** - The data series chosen to represent a day in the clustering algorithm (for example net demand vs separate wind, solar and demand series).
- **Extreme points -** The potential inclusion in the clustering output of extreme points or 'days of interest' assumed to be important to the accuracy of the generation expansion planning model.

Allowing for the manipulation of these dimensions has implications on the implementation of the clustering algorithm as described below.

3.2.1 Algorithm

The clustering algorithm developed for this analysis is based on the Partitioning Around Medoids (PAM) algorithm [72]. For k representative clusters with the inclusion of p preselected medoids, the algorithm can be written in seven steps, as follows:

1. **Normalise data** - The data is a collection of days each with one or more time series of observations (eg. 48 half hourly net demand observations). Each observation in each series is normalised with the series mean and standard deviation.
2. **Force medoids -** Determine p preselected days and set these as medoids for the first p clusters.
3. **Initialise remaining medoids -** Randomly select the remaining $k - p$ medoids from the dataset as initial medoids for the remaining clusters.
4. **Determine clusters -** For each day determine the distance (using the below described distance metric, see 3.2.1.1) to each of the k medoids and find the closest medoid. Groups of days with the same closest medoid are considered 'clusters'.
5. **Update medoids -** For each of the k - p clusters without a pre-selected medoid calculate the distance between each day and all other days (see 3.2.1.1) – select the day with the minimum total distance to all other days in the same cluster as the new medoid for that cluster.
6. **Converge -** Repeat steps 4 – 5 until the there is no change in the closest medoids for each day (or a maximum number of convergence iterations reached).
7. **Report -** For each cluster of days report the selected medoid and the cluster size (to be used as a weight in GEP) as well as the total distance of all days in the cluster to their cluster medoid.

The quality of the solution found may depend on the random selection of initial medoids. The algorithm is typically repeated for several different randomly selected starts and the result with the lowest total distance metric is kept. In this case, the algorithm is repeated 10000 times.

3.2.1.1 Distance metric

The unit of selection here is a day which is then tested with different levels of data, for example either a net demand series with 48 observations or two separate wind and demand series with 48 observations each. The distance metric used is an average squared Euclidean distance which helps ensure consistency across different numbers of series and different series sizes (for example two data series per day vs one):

$$Distance(d, d') = \frac{\sum_{s=1}^{n_s} \frac{\sum_{t=1}^{n_t} \left(d_{s,t} - d'_{s,t}\right)^2}{n_t}}{n_s} \tag{38}$$

where *d* and *d'* are two selected days, each day is represented by n_s data series. With each data series containing n_t observations (here 48 half hour periods). $d_{s,t}$ is then the values of series *s* at time *t* on day *d*.

3.3 Representation of input data

3.3.1 Selection of series for clustering algorithm

Starting with historically observed data for a power system, there are a number of different ways observed information could be included in the clustering algorithm.

Individual series: historical demand, wind capacity factor, and solar capacity factor- The data that determines net demand for each day can be represented as individual series (one series each for load, wind, and solar) as in [79]. Here demand is in MW per period, wind and solar in average fleet capacity factor for the period.

Net demand series – the data for wind, solar, and demand can be combined to determine net load series as is in [86]. The derivation of the net load curve would be different for different levels of future wind and solar capacity (see 3.3.2 for discussion). The advantage of this approach is that the relative importance of the underlying series (demand, wind, solar etc) is considered.

Net demand duration curve - An alternative common representation for a net demand series is that of a duration curve. Here the net demand is ordered from highest to lowest value for each day. The advantage of this approach for clustering is the highest demand period (and second highest, and third etc) for each day is compared to the highest of every other day, which would not happen with the chronological series if they happen to fall in different periods. The disadvantage of the duration curve is that it does not represent the chronological nature of the time series. That is, it loses all information about the size of the transition between periods which can drive flexibility requirements.

Ramp duration curve –[89] introduces the concept of a ramp duration curve. First the ramp rates are calculated (by differencing the net demand series) and then ordered to form a duration curve. This allows an additional series for the algorithm to cluster on that represents the ramping requirements of each day.

This chapter tests these representations in different combinations to attempt to find a representation of the input series that allows the algorithm to cluster against the dynamics that are important for the GEP model.

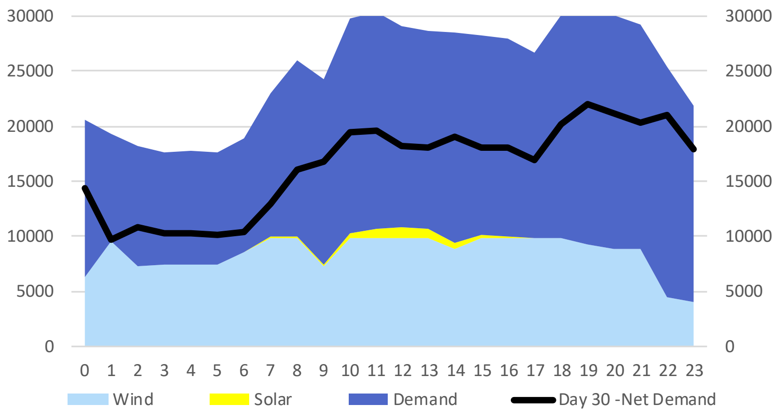
3.3.2 Historical and future net demand

Clustering is often performed on historical observed net demand. The issue with this is demonstrated in Figure 5 where two days are compared: day 30 in winter (a) and day 139 in spring (b). In chart (c) with historical levels of renewable capacity the two net demand levels are reasonably close (solid lines). However, at future levels of renewable uptake the net demand of the two days is drastically different (dotted lines).

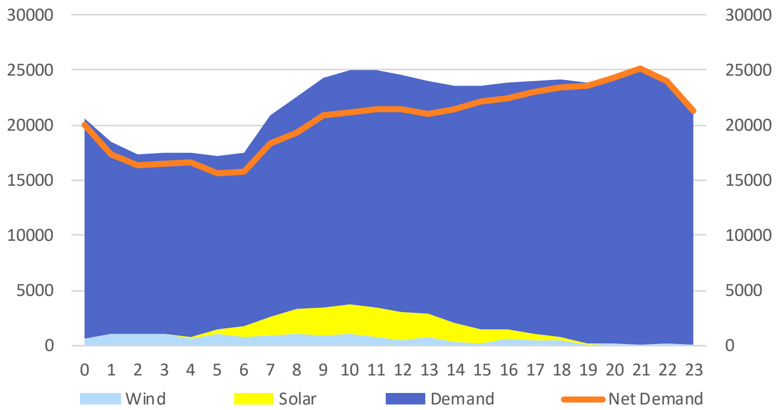
This presents a potential problem when clustering days based on historical net demand as the clustering algorithm will see little difference between the two days (and cluster them together) when for the purpose of the GEP model they are very different.

As a result, it could be expected that clustering on a series based on historical renewable deployment may not produce a net demand series which is representative for the whole period of the expansion planning problem. I therefore compare two representations of net demand:

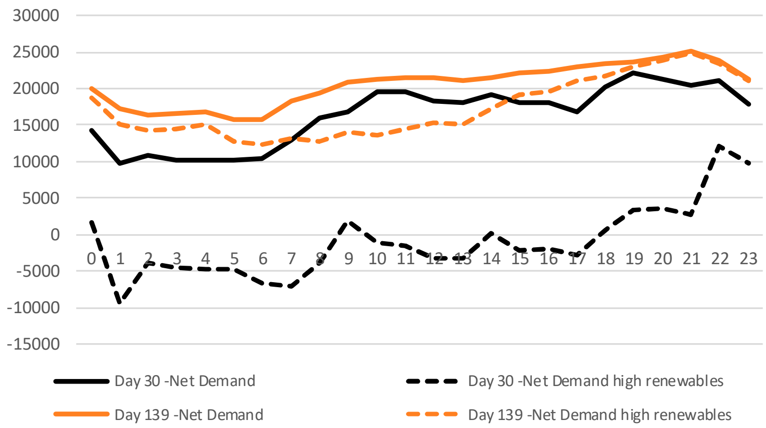
- One where each day has a historical net demand series (**ND**)
- One where each day has two net demand series; one based upon historical renewable deployment and one based upon the future deployment of the modelled year (**NDHF**).



(a)



(b)



(c)

Figure 5 (a) determination of historical net demand day 1 (b) determination of historical net demand day 2 (c) comparison of net demand with different levels of future renewable uptake.

3.3.3 Inclusion of ‘days of interest’ (forced inclusion of selected medoids)

Table 1 Description of days of interest

Days of interest (day with:)	**Description**
Maximum value (peak period) – Always included	Day with the series highest recorded value.
Maximum min-max difference	Day with the greatest difference between the highest and lowest value.
Maximum min-max difference over 6 hours	Day with the greatest difference between the 6 hours with the highest values and the 6 hours with the lowest values.
Maximum consecutive positive change up	Day with the highest positive increase over consecutive periods (largest consecutive ramp up)
Maximum consecutive positive change down	Day with the highest negative change over consecutive periods (largest consecutive ramp down)
Maximum single period change positive (peak period ramp up)	Day with the highest positive difference between two periods
Maximum single period change negative (peak period ramp down)	Day with the highest negative difference between two periods
Minimum value	Day with the series with the lowest recorded value.

For the purpose of modelling power systems clustering can underrepresent extreme points which are important for driving GEP results [81] [92]. Forcing the inclusion of specific observations based on the modellers prior knowledge of the problem improves the quality of results for generation expansion planning. An obvious example is the inclusion of the ‘Peak’ period of the year as this will be a considerable driver of the expansion of capacity and is unlikely to be central to a cluster (as it is an extreme point) and therefor unlikely to be otherwise selected. The peak day is therefore often “manually” added to the selection as additional or set as a cluster centroid [92] [90]. Without this forced inclusion the clustering will take a central point of the cluster that the day with the peak period is within to represent that period and the generation expansion planning model will not contain the exact peak demand required to be met in the future (based on historical peak demand).

As clustering tends to underrepresent extreme points by selecting central points or averages [81] [88] [93], and as these can be important in driving expansion results [81] [92], I examine the inclusion of a number of ‘days of interest’ which contain specific observations that may help to ensure the selected representative days capture the future power system challenges relating to flexibility and storage operation. The extreme points tested are summarised in Table 1.

The additional medoids are chosen based on similar reasoning to that of the peak day, that results can be driven by extreme values. I test two representations of future ramping requirements to ensure that the system builds the flexibility to deal with the greatest flexibility challenge. I test both the greatest ramp (up and down) between two individual periods and the greatest consecutive ramp (up and down).

To test properties that may be important to the deployment of storage devices that operate on a daily cycle I test two ‘days of interest’. The extreme min-max differences is tested to attempt to capture the greatest value a storage can achieve in a single hour and the greatest min max distance over six hours as an estimate of the value that can be achieved by using the full storage capacity.

Finally, minimum values are included as they are likely to be important for representing future renewable curtailment challenges (either through curtailing renewables or forcing conventional generation off and incurring future start-up costs).

Where the input data includes multiple series (e.g. Net demand data for current and future renewable penetration) the selected observation (e.g. Day with peak period) for each of the series is included.

3.4 Weighting of representative days for expansion planning model

The output of the clustering algorithm consists of several selected typical days (cluster medoids) and their corresponding cluster sizes (that quantify the number of actual days that are being represented by each selected medoid day). However, the summed product of these is not guaranteed to equal the total energy (or renewable production) of the observed data which is an important factor in the GEP model.

Careful treatment of this problem is particularly important when comparing different representative day selections. For example, if two representative day selections had different levels of net demand to meet, this would influence the expansion results.

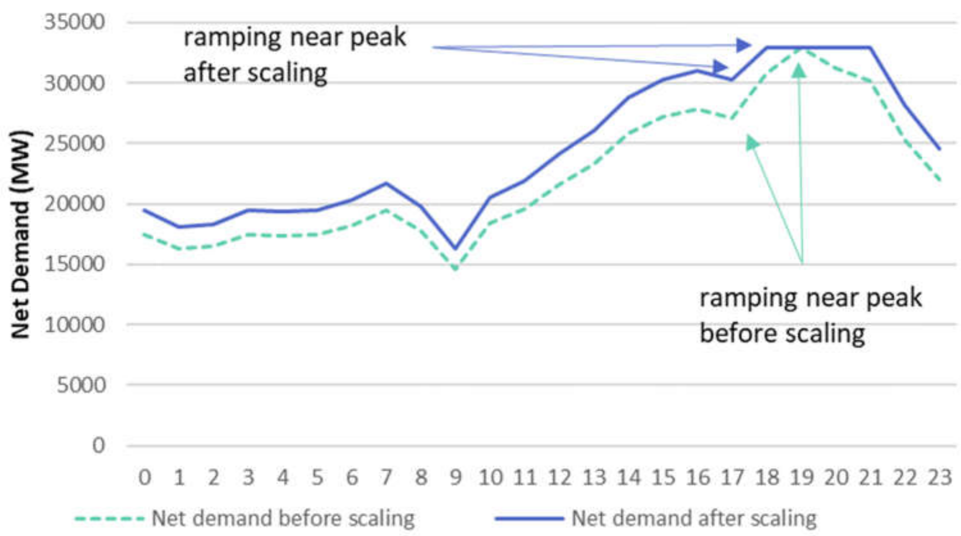


Figure 6 Illustration of Scaling of net demand series, maintain peak value, to increase average net demand 10%, as in [83]

In some approaches the underlying data is scaled[16] so that this error is removed [83] [90] [94]. In practice the scaling to observed values is often done in a way so as to preserve the value of the peak period [83] meaning that some periods are scaled more than others (the peak period is not scaled, and no other period is scaled to be above the peak period).

[16] For example, if we had selected one representative day, we would scale the net demand values for each hour so the total net demand for the day multiplied by the number of days in the year would equal the full dataset annual total.

The scaling is then different for different periods and this will alter the change between periods (the required ramp rate), a key potential metric for driving system costs and ensuring future flexibility challenges are modelled. Ramp rate requirements around the peak period, a potentially extremely challenging period for the model as most capacity is already committed at high load, will potentially be altered the most. An illustrative example where a single representative day is selected, but net demand for that day is 10% lower than the average over the full dataset is included in Figure 6.

Here I propose an alternative methodology that allows us to preserve underlying data series and instead adjusts the weights applied to each day in the GEP. The methodology is as follows:

- Select representative days and determine cluster weightings for these days (as outlined in section 3.2).
- For each year to be modelled determine the net demand associated with each day. The net demand will depend on future demand, renewable modelled capacity[17], expected change in renewable annual capacity factor (if different to historical), and the day's renewable capacity factor.
- Find a set of weights to be used in the GEP (a weight for each cluster medoid, for each year) as close as possible to the cluster weights that ensures total net demand in the GEP.

To determine the new weights a simple quadratic optimisation problem is solved that

[17] Here we focus on the case where renewable capacities are known in advance (potentially as uptake is drive by policy targets as opposed to market dynamics).

calculates a scalar[18] for each cluster size that minimises the scalar squared distance from 1 (Equation (39)) to achieve a final weight as close to the original cluster size as possible. The optimisation is constrained to ensure the total net demand is achieved[19] (Equation (40)) and that the sum of the weights remains unchanged (Equation (41)). Additionally, the formulation requires that no weight can be scaled below 1 or above the total number of days (Equation (42)).

The optimisation problem needs to be solved for each year in the GEP model (unless no change in relative shares of wind, demand, or solar are expected as the level of Net Demand associated with a representative day will change with these shares). The mathematical formulation of the problem is given below.

$$\min_{s} \sum_{d \in D} (1 - s_d)^2 \tag{39}$$

s.t.

$$\sum_{d \in D} W_d^C \cdot ND_d \cdot s_d = ND^A \tag{40}$$

$$\sum_{d \in D} s_d \cdot W_d^C = |D| \tag{41}$$

$$1 \leq W_d^C \cdot s_d \leq |D| \quad \forall\, d \in D \tag{42}$$

[18] A formulation based on adding or subtracting to the weights was also considered. However, this approach tends to adjust the days with smaller weights to a relatively higher degree, which also tend to be the more extreme days, and results in more extreme final weighted representative days.

where s_d is the scalar to be applied to the weight W_d^C (the direct output of the clustering) of each representative day $d \in D$ when used in the GEP model. ND_d is the total net demand of day d and ND^A is the total net demand of the underlying dataset over all $|D|$ days.

While this formulation is not guaranteed to be feasible[20] in general, it was found to be so for all cases analysed here.

Here I focus on the case where net load is known in advance as renewable build is determined by factors exogenous to the model (government policy etc). It is possible to formulate this problem so that not just net demand is achieved but capacity factors for wind and solar are scaled to match historically observed values. However, this is a significantly more restricted problem and results in unreasonably large scalars being calculated (where the problem is feasible) in this case study with the average absolute change to weights greater than 50%.

This formulation, additionally, allows us to either treat renewable production as exogenous and simply use net demand in the GEP model, or alternatively, have endogenous representations of renewable capacities and use the capacity factors associated with the representative days in the model.

[20] For example, if all selected days had higher/lower average net demand than the average over the year.

3.5 Generation expansion planning model

The model outlined in Chapter 2 is applied with the following modifications:

- The model is deterministic (with objective function Equation (1)) and only a single future year is optimised at a time, as such the indices s and y are not required on any of the variables.
- Each existing power station is modelled individually, therefore, the clustered unit commitment formulation is not used.
- Integers are enforced on commitment decisions, therefore the linear relaxation formulation is not used.
- The UK does not feature significant large-scale hydro, as such existing run of river hydro is subtracted from demand and the large-scale hydro constraints are not required.

3.6 Evaluation model

The aim of the representative day selection task is to determine the approach best suited to selecting inputs to a GEP model. To compare the different time period selections I first simulate the selected represented days and weights in a generation expansion planning model to determine the set of expansion decisions and the estimated system costs.

To determine how well the selected days have represented the full data set the expansion plans are then tested in a daily unit commitment (UC) model. The daily UC model allows us to calculate how the expansion plan performs over the entire dataset, i.e. the full time period of the original data, which cannot be solved practically as a single GEP optimisation problem.

3.6.1 Unit commitment (UC) model

To assess the build decisions on the full dataset I restrict the model to a set of sequential single day problems that focus on the unit commitment decisions. To do so I fix the build decisions to the optimal values from the GEP model and remove the investment and fixed costs (so they are not incurred every day).

The model is run sequentially for the full dataset (a full year) optimising one full 48 period day at a time with a 48 period lookahead as in [95]. That is, the model optimises decisions over a full 96 periods representing two days (to ensure naïve start or shutdown decisions are not made at the end of the first day as the model must ensure units are online for a second day or incur start costs) but only the results of the first day are recorded. The model then simulates the following day with the initial commitment decisions and initial generation levels set from the end of the previous day (period 47).

3.6.2 Simulation system

All simulations are performed on a laptop with an Intel Core i7-6500U CPU @ 2.50GHz and 8GB of RAM. Models are implemented in C# and optimisations problems solved with the Gourbi solver. Optimality gap target of 0.1% reached with an average time of 2013 seconds.

3.7 Case study

The United Kingdom (UK) has committed to decarbonisation and the electricity system is expected to feature large quantities of renewable energy in the near future. I make use of this system as a case study and examine two cases with differing levels of renewable penetration:

Medium term system (2020) - medium renewable penetration – the UK market as it currently exists with expected conventional retirements and expected increases in renewable capacities [68].

Future system (2025) – high renewable penetration – In this future scenario there is a significant growth in wind and solar generation capacity and a decrease in conventional dispatchable capacity. From the capacity mix forecast for 2025 in [68] I add 2% growth to demand and allow the model to select the capacity to meet this gap. See Figure 7 for capacity mixes.

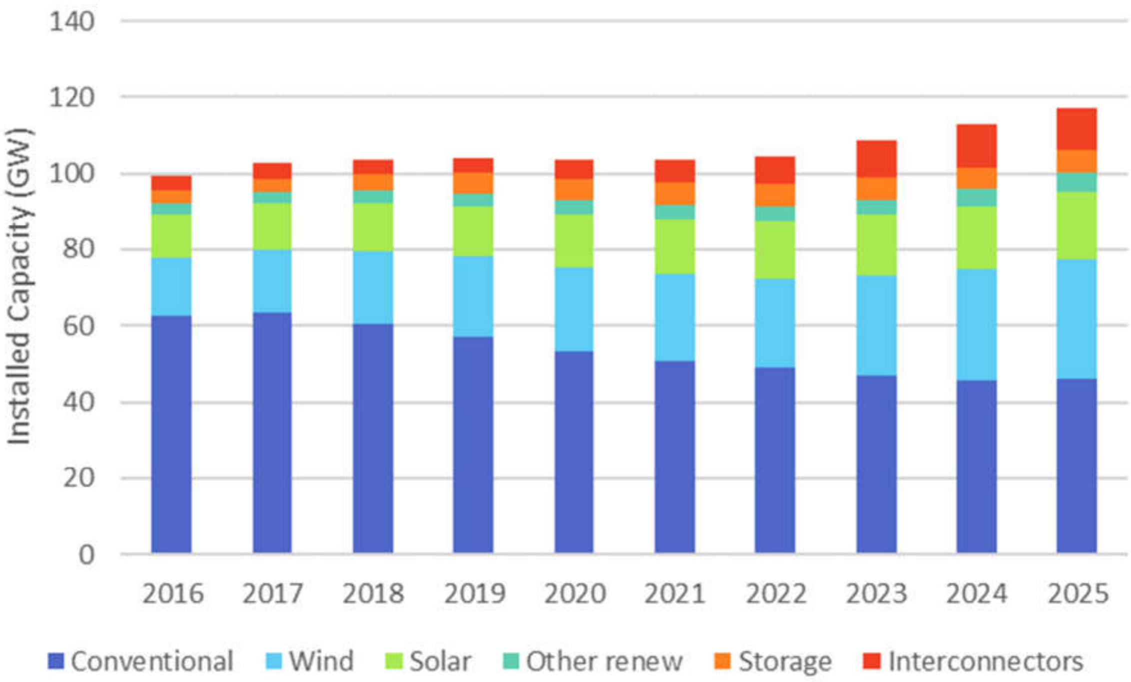


Figure 7 Installed generation capacity adapted from slow progression scenario [68]

For robustness I perform the clustering exercise and all analysis on two separate historical datasets (one for 2015 and one for 2016) based on half-hourly demand, solar, and wind profiles[21].

[21] While we would not recommend selecting on a single year of data to find a representative set of profiles for future expansion planning this approach is sufficient for

3.7.1 Generation expansion planning

Here I examine the ability of a number of technologies to meet future system challenges. In particular I compare two different theoretical battery storage technologies, one with 2 hours of storage capacity and one with 5 hours. While these technologies may not be representative of the current economical storage sizing, they allow us to distinguish the reason storage is being built. If the model is using storage to shift large quantities of load (overnight wind to peak or solar to mid evening) the larger storage capacity would be required. However, if storage is being used to meet ramping challenges it may not be the case that such a large storage is required, and a smaller cheaper storage may be selected by the model.

I compare these two storage technologies to two conventional technologies CCGT and a Peaker (OCGT).

I include a supply curve for each technology by increasing the capital costs of each technology as more of that technology is built[22]. For my purposes this assumption improves the robustness of the comparison as a small change cannot result in the optimal technology mix switching completely from one technology to another and we can be more confident of large technology changes when we observe them.

For more information on unit characteristics see Appendix B.

---

testing the ability of a clustering algorithm to represent and underlying dataset [89].

[22] This assumption can be thought to reflect either a preference for diversification or that as the optimal siting for new technologies is exploited less optimal siting must be used and installation costs increase or renewable production decreases.

3.8 Results

In this section the effect of changing the representation of the input series to the clustering (sections 3.8.1 and3.8.2) is examined, before the inclusion of extreme points (section3.8.3). For all results I examine the ability of the clustering to represent underlying demand, and additionally the ability to perform well in a GEP model. For more details on the assumptions underlying this analysis please see Appendix B.

3.8.1 Representation of data input series

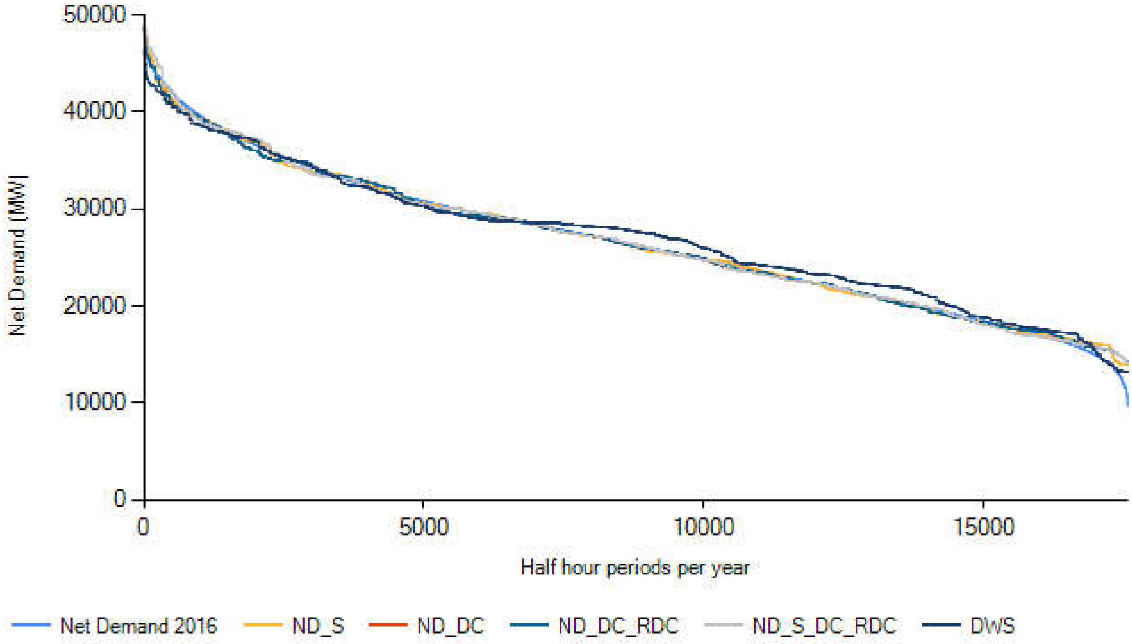
Table 2 Key for case abbreviations

Abbreviation	**Case description**
ND_S	Historical net demand as original chronological series used in clustering
ND_DC	Historical net demand as duration curve used in clustering
ND_DC_RDC	Historical net demand as duration curve and historical net demand ramp duration curve used in clustering
ND_S_DC_RDC	Historical net demand as original series, as duration curve, and as ramp duration curve all included in clustering
DWS	Individual historical demand, wind capacity factor, and solar capacity factor included in clustering.

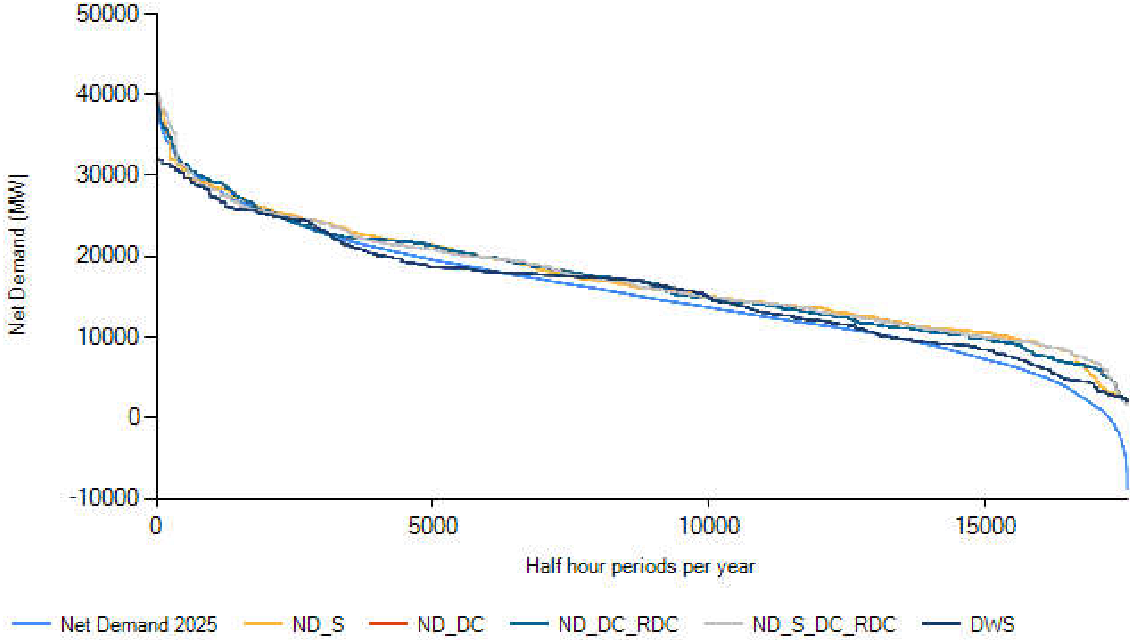
All case report results for 9 selected days with the day including the peak value for each series forced as a centroid for inclusion.

3.8.1.1 Clustering and scaling results

Figure 8 displays the load duration curves constructed for the different representative day selections of 9 days based upon historical data in comparison to the full data set (for the 2016 input data case). The peak value is the same across all selections based upon net demand, as the day including this value is enforced as the maximum of that series. For the selection based upon individual demand, wind, and solar inputs (DWS) this is no longer the case as while the peak demand day is captured the peak net demand is not (as the peak net demand day would have lower renewable production). Further as the weighting placed on this day are similar across all the selections, and no other days with high net demand values are selected the high load portions of the load duration curves are similar across all selections. The ability of the selections to capture the low load portion of the net demand series differs more significantly. This effect becomes more pronounced as I look to model the net demand for the 2025 year with higher levels of renewable penetration as illustrated in Figure 8 (b).



(a)



(b)

Figure 8 Load duration curves for full year data and different selections (a) 2016 Net Demand (b) 2025 Net Demand

Table 3 reports the clustering algorithm average squared Euclidean distance and a number of metrics relating to the fit of the selected data to the full data series as commonly used to assess sections [84] [88]. To assess the difference between the full dataset duration curve and the selected day duration curve I calculate the normalised root mean square deviation (NRMSD) [84] for each point i on representative day duration curve $RDDC$ and the full dataset duration curve DC.

$$NRMSD = \frac{\sqrt{\frac{1}{8760}\sum_{i=1}^{8760}(RDDC_i - DC_i)^2}}{\frac{1}{8760}\sum_{i=1}^{8760} DC_i} \tag{43}$$

The squared Euclidean distance the clustering process attempts to minimise demonstrates that the ordered duration curve data based approaches are significantly easier to cluster than the unordered net demand or individual series data. However, when I look at how well the selections fit the load duration curve as a whole (as measured by NRMSD) I see again the selections ability to represent the data decreasing the further into the future I look (as wind and solar capacity increase in the underlying net demand). The individual series clustering (DWS) is the least susceptible to this issue as it inherently places a high value on capturing wind and solar dynamics. This works well for this case study where wind and solar are important drivers of net demand, however, this may not be the case for all applications. The addition of the ramp duration curve to the net load duration curve does improve its ability to fit the ramp duration curve both in the historical and future data as measured by the NRMSD.

Table 4 demonstrates the error in total net demand that would result from using the representative days and weights from the clustering algorithm directly. As with the other metrics the error increases into the future and reaches significant levels. The average absolute scalar calculated by the quadratic optimisation is also reported here and increases with the size of the error, reaching high values for the furthest into the future year of 2025.

Table 3 Clustering metrics - Euclidean distance, normalised root mean square deviation (NRMSD)

	Average squared Euclidean Distance		**Normalized root mean square deviation (NRMSD) of duration curves – average of 2015 and 2016**								
	2015	2016	Historical Net Demand	2020 Net Demand	2025 Net Demand	Demand	Wind	Solar	Historical Ramp	2020 Ramp	2025 Ramp
ND_S	29.78	26.28	0.05	0.12	0.28	0.18	0.70	0.17	0.08	0.10	0.15
ND_DC	15.30	15.31	0.05	0.11	0.24	0.16	0.46	0.16	0.07	0.09	0.14
ND_DC_RDC	17.87	18.25	0.06	0.11	0.21	0.14	0.53	0.26	0.06	0.07	0.13
ND_S_DC_RDC	23.10	21.95	0.06	0.12	0.26	0.18	0.65	0.14	0.07	0.08	0.13
DWS	72.69	77.67	0.13	0.14	0.15	0.12	0.14	0.11	0.11	0.12	0.16

Table 4 Error in net demand of sum product of representative day net demand and cluster size and average absolute change in weights from calculated scalars

	Error in net demand without scalars						**Average of absolute scalar adjustments to clustering weights calculated to remove error in total net demand**			
	2015 Load shapes			**2016 Load shapes**			**2015 Load shapes**		**2016 Load shapes**	
	2015	**2020**	**2025**	**2016**	**2020**	**2025**	**Scaled to ND 2020**	**Scaled to ND 2025**	**Scaled to ND 2020**	**Scaled to ND 2025**
ND_S	-0.1%	3.3%	13.4%	0.0%	2.9%	4.0%	16.2%	37.6%	11.5%	8.7%
ND_DC	0.0%	1.7%	6.8%	-0.1%	2.4%	4.0%	8.8%	19.6%	10.4%	8.7%
ND_DC_RDC	-0.1%	2.9%	11.5%	-0.1%	1.1%	1.8%	11.7%	27.9%	4.2%	4.3%
ND_S_DC_RDC	0.0%	2.6%	10.6%	-0.3%	3.4%	1.1%	13.1%	29.6%	13.6%	2.6%
DWS	1.8%	2.3%	4.0%	0.5%	1.8%	2.2%	9.7%	8.4%	6.3%	4.5%

3.8.1.2 Expansion planning model results

In this section I compare how the capacity expansion planning model performs using the different sets of representative days. Figure 9 displays the error in the estimate of total cost (as compared to a unit commitment model simulating the full data set) for each of the expansion plans selected by the GEP model. For the most part when using the net demand series to select representative days the GEP model can be shown to overestimate the costs, with the exception of the 2025 model year with the 2016 load shapes. The individual series selection (DWS) greatly underestimates costs across all scenarios, potentially reflecting its lower coverage of the higher portion of the net demand load duration curve.

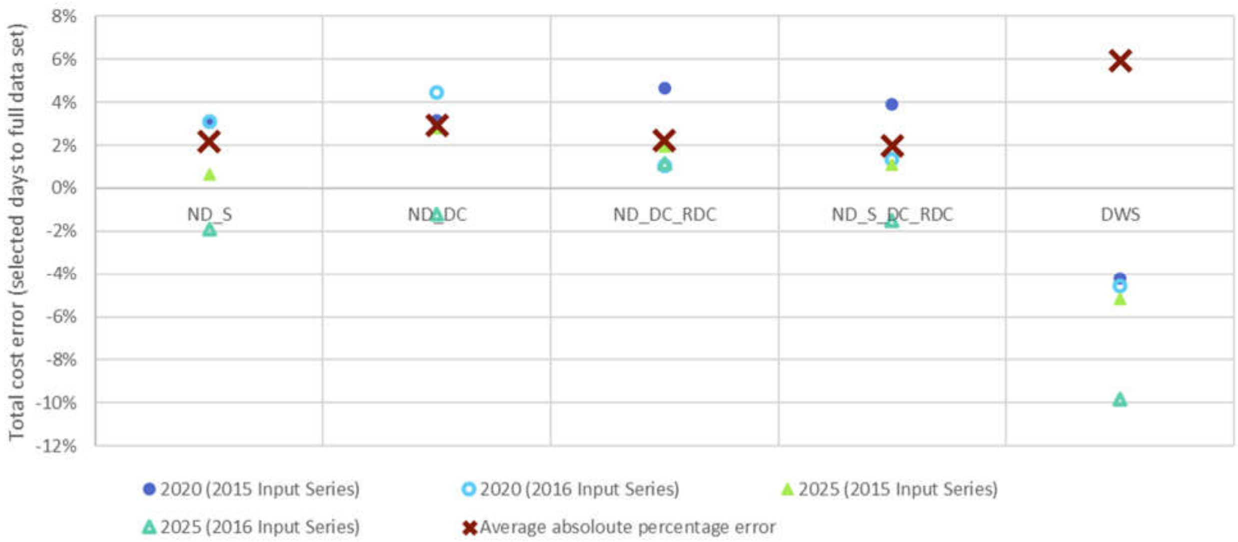


Figure 9 Percentage error in representative day cost estimation between GEP model with representative days and unit commitment results over full dataset by capacity expansion model year (2020 or 2025) and input data series year (2015 or 2016)

In addition to the errors in cost estimation we are interested in how close the solution found is to the optimal one for the entire represented year. While it is not possible to solve the expansion problem for the full year I can compare solutions found to the best solution found as assessed by the unit commitment model for that case (as an approximation of the optimal solution). Figure 10 presents the costs above the best solution found (across all combinations tested here) for both clustering data years and future model years for each representative day selection method.

Figure 10 shows that all representative days perform relatively poorly in the 2020 model year when selecting from 2015 representative days (despite different errors in cost estimation for this case the final expansion plan for the model is very similar across all cases). The two approaches that include the original net demand series (ND_S and ND_S_DC_RDC) perform the best in total with the duration curve approaches (ND_DC and ND_DC_RDC) performing badly based upon the 2015 data set. While the addition of the ramp duration curve improves results over the duration curve alone it is the combination of these with the chronological series that performs the best. The importance of the unordered original series (ND_S) may reflect a number of factors: firstly, the even relative weighting between the duration curve and ramp duration curve may not reflect the relative importance of these features to the GEP model. Alternatively, the time of day that ramping occurs (and its correlation with net demand) may be important and oversimplified by the ramp duration curve. For example, a high ramp up may be easily met during an off-peak period but challenging close to the peak. However, without the unordered series this difference is not captured in the clustering. Additionally, the number of consecutive ramping periods may cause different challenges for the GEP model and result in more starts than alternating ramp up and ramp down periods despite these being similar when converted to a duration curve.

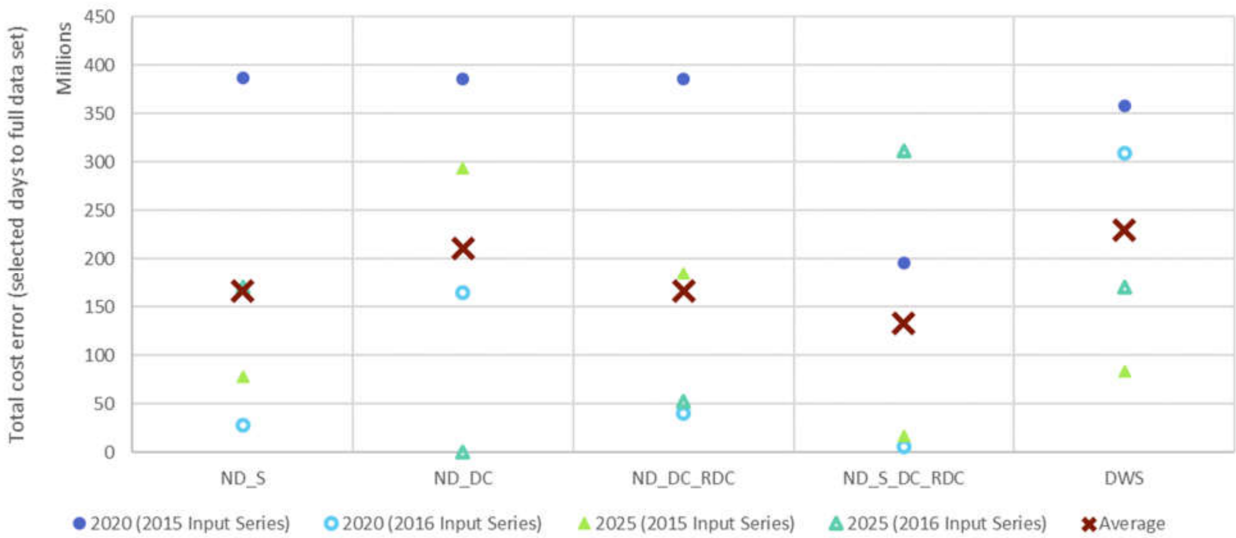


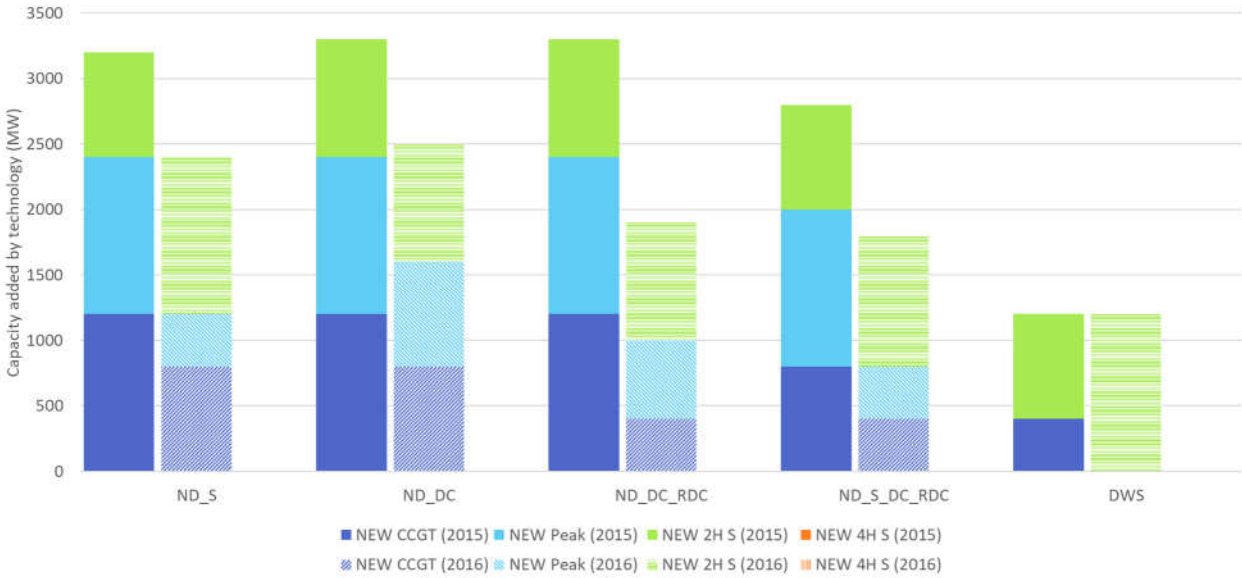
Figure 10 Additional cost of expansion plan evaluated over full dataset over best solution found, all input series year and GEP model year combinations

The differences between performance for the same model year with different datasets

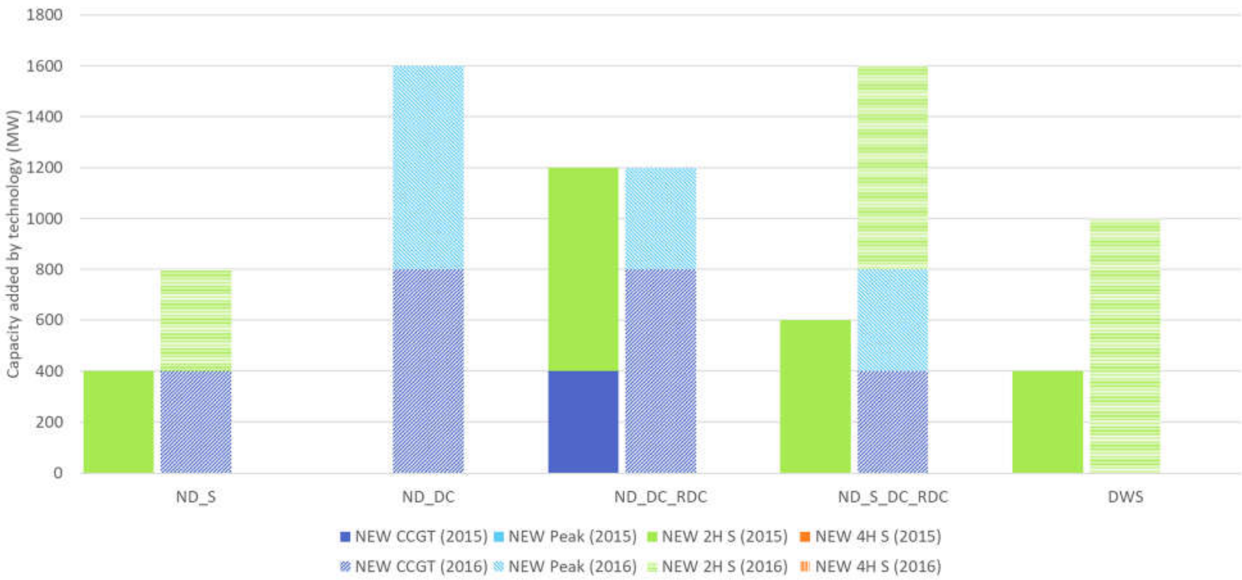
(2015 or 2016 data) demonstrates the importance of testing the selection against different inputs.

Figure 11 compares the capacity expansion decisions for each technology depending on the representative day selection approach. The build plans for the 2020 model year are reasonably similar across net demand clustering approaches but with large differences depending on the input data year. The most significant difference between these approaches is when selecting with the 2016 input data and the approaches that incorporate the original series (ND_S) more storage and less conventional capacity is built (which results in a closer to best solution). The DWS approach leads to significantly lower build of capacity, in particular conventional capacity, and as seen this results in a less optimal solution.

In the 2025 model year there is a more significant difference between the approaches. Based upon the 2015 input data the clustering based on duration curve approaches decides not to build additional capacity, where the other approaches allow a small amount of storage which results in a significantly lower cost solution. When selecting based upon 2016 input data the duration curve approaches build significantly more capacity which results in a lower cost compared to the lower capacity solutions.



(a)



(b)

Figure 11 Capacity expansion plan by representative day selection approach and input dataset (2015 or 2016) (a) 2020 expansion model (b) 2025 expansion model

3.8.2 Future Net demand

Table 5 Key for case abbreviations

Abbreviation	Case description
NDHF_S	Historical and future year net demand as original chronological series used in clustering (two data series total per day)
NDHF_DC	Historical and future year net demand as duration curve used in clustering (two data series total per day)
NDHF_DC_RDC	Historical and future year net demand as duration curve and historical net demand ramp duration curve used in clustering (four data series total per day)
NDHF_S_DC_RDC	Historical and future year net demand as original series, as duration curve, and as ramp duration curve all included in clustering (six data series total per day)

In the previous section I clustered on historical net demand. In this section I report the results of expanding the net demand clustering approaches to include the future net demand (based on known renewable expansion) as well as the historical net demand for each day. This approach allows us to select one set of representative days, based on the realised net demand of each day at different levels of renewable penetration, that could be applied across the multiple years usually included in a GEP model.

3.8.2.1 Clustering

Figure 12 demonstrates how the inclusion of future net demand in the clustering improves the 2025 net demand curve representation at the top and bottom of the load duration curve (highest 10% and lowest 10% of hours) compared to Figure 8, however, the very lowest net demand hours where net demand is negative are still not well represented.

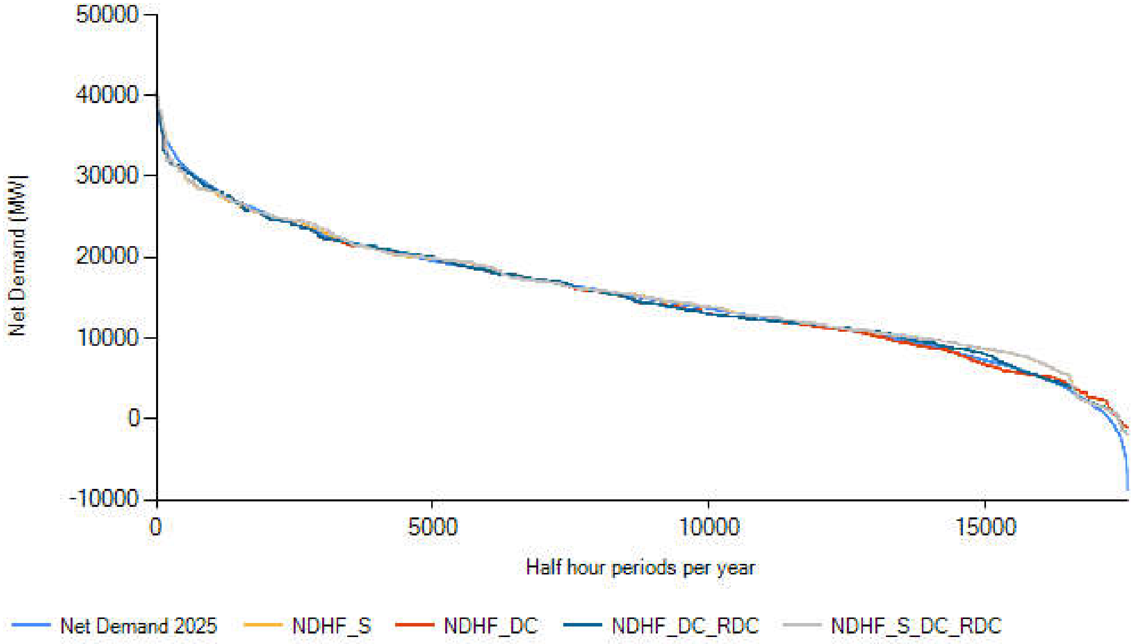


Figure 12 Load duration curves for full year data and different selections 2025 Net Demand

Table 6 reports the changes to the clustering metrics for the new approach. As is to be expected the Euclidean distance the clustering algorithm is attempting to minimise increases significantly as each day contains more data to compare between.

Table 6 Change in clustering metrics resulting from addition of future series to clustering (%)

	Change in Euclidean Distance		Change in Normalized root mean square deviation (NRMSD) of duration curves – average of 2015 and 2016								
	2015	2016	Historical Net Demand	2020 Net Demand	2025 Net Demand	Demand	Wind	Solar	Historical Ramp	2020 Ramp	2025 Ramp
NDHF_S	57%	50%	32%	-25%	-63%	-44%	-50%	-5%	14%	19%	-7%
NDHF_DC	64%	54%	10%	-37%	-70%	-29%	-30%	59%	-6%	-10%	-9%
NDHF_DC_RDC	49%	34%	1%	-28%	-58%	-37%	-53%	-34%	11%	-7%	-31%
NDHF_S_DC_RDC	55%	42%	15%	-29%	-65%	-59%	-52%	11%	-6%	-15%	-25%

The change resulting from the inclusion of future net demand can be seen in the load duration curve distance metrics where the ability to represent the historical load duration curve decreases, however, the representation of the future duration curve improves significantly. The better fit of future demand curves results in lower errors in total demand produced by the clustering algorithm selection. This results in a significant reduction in the scalars required to achieve the correct total demand as shown in Table 7.

Table 7 Error in net demand of sum product of representative day net demand and cluster size and average absolute change in weights from calculated scalars

	Error in net demand without scalars						Average of absolute scalar adjustments to clustering weights calculated to remove error in total net demand			
	2015 Load shapes			2016 Load shapes			2015 Load shapes		2016 Load shapes	
	2015	2020	2025	2016	2020	2025	Scaled to ND 2020	Scaled to ND 2025	Scaled to ND 2020	Scaled to ND 2025
NDHF_S	0.1%	1.1%	0.2%	-0.3%	0.6%	1.8%	5.1%	0.4%	2.2%	4.3%
NDHF_DC	-0.2%	0.5%	-1.2%	0.4%	0.1%	-0.7%	2.6%	3.0%	0.5%	1.7%
NDHF_DC_RDC	-0.3%	-0.3%	-1.1%	0.0%	-0.2%	-0.4%	1.5%	2.7%	0.6%	0.9%
NDHF_S_DC_RDC	-0.1%	0.3%	0.8%	-0.1%	-0.1%	1.4%	1.4%	2.1%	0.3%	3.4%

3.8.2.2 Expansion planning model results

Figure 13 looks at how the average absolute error in cost estimation, and the quality of the build plan as measured by the average cost above the best solution, change as I add the future net demand to the clustering. For both the original series and the duration curve, improvements are seen across both these metrics indicating a robust improvement in the quality of representative days selected for GEP modelling.

The improvements to the duration curve selection previously seen by adding the ramp duration curve (or ramp duration curve and original series) no longer hold, potentially indicating that these were acting as a proxy to select days that were more important for future years which are now captured.

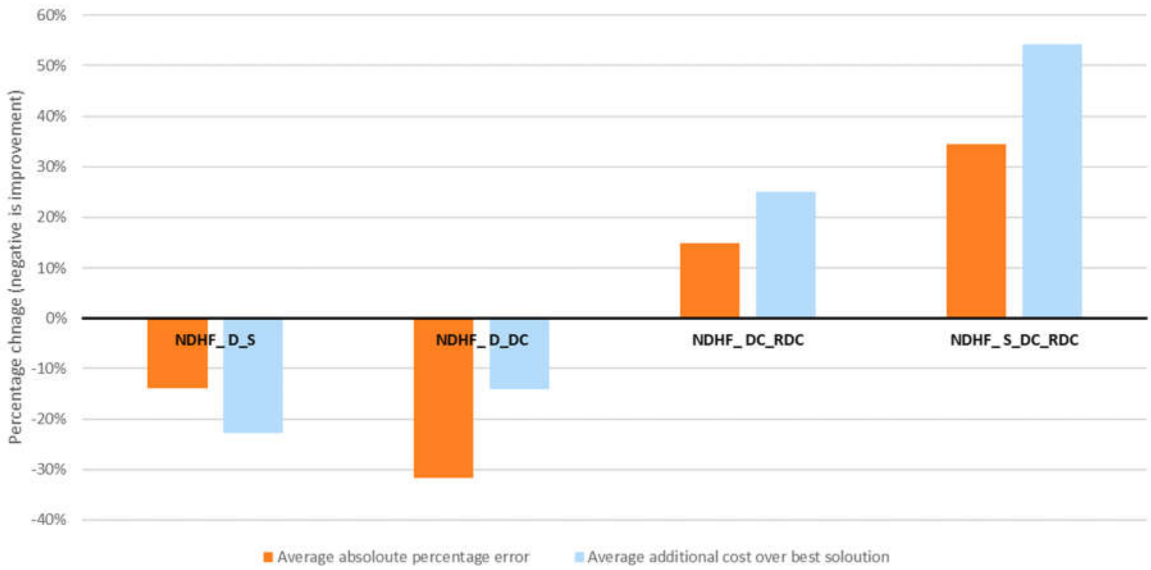
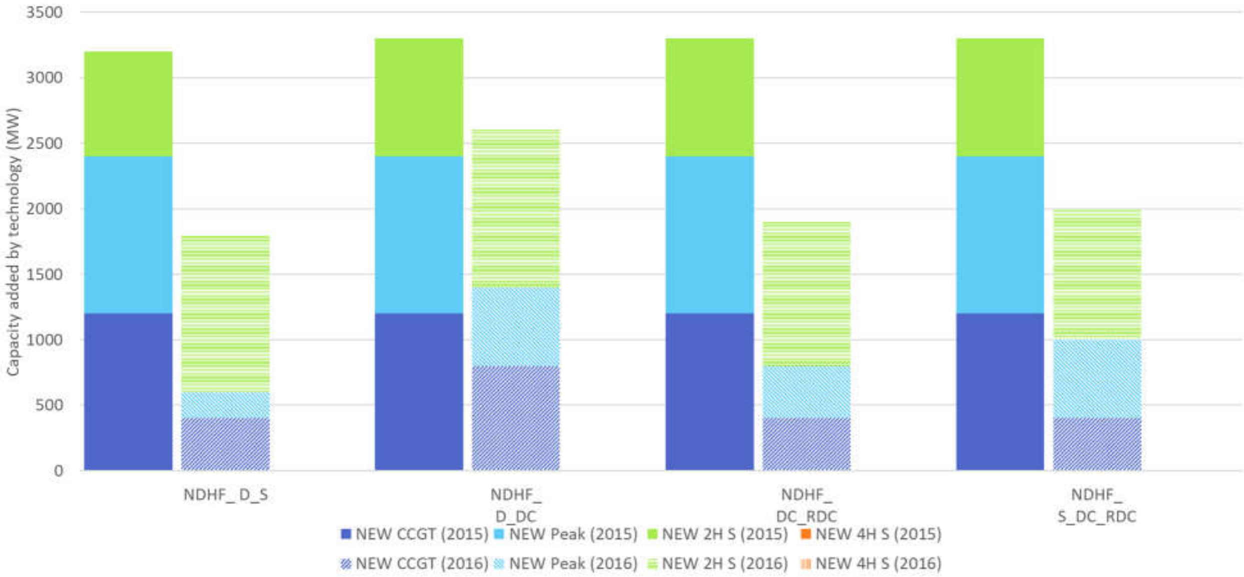
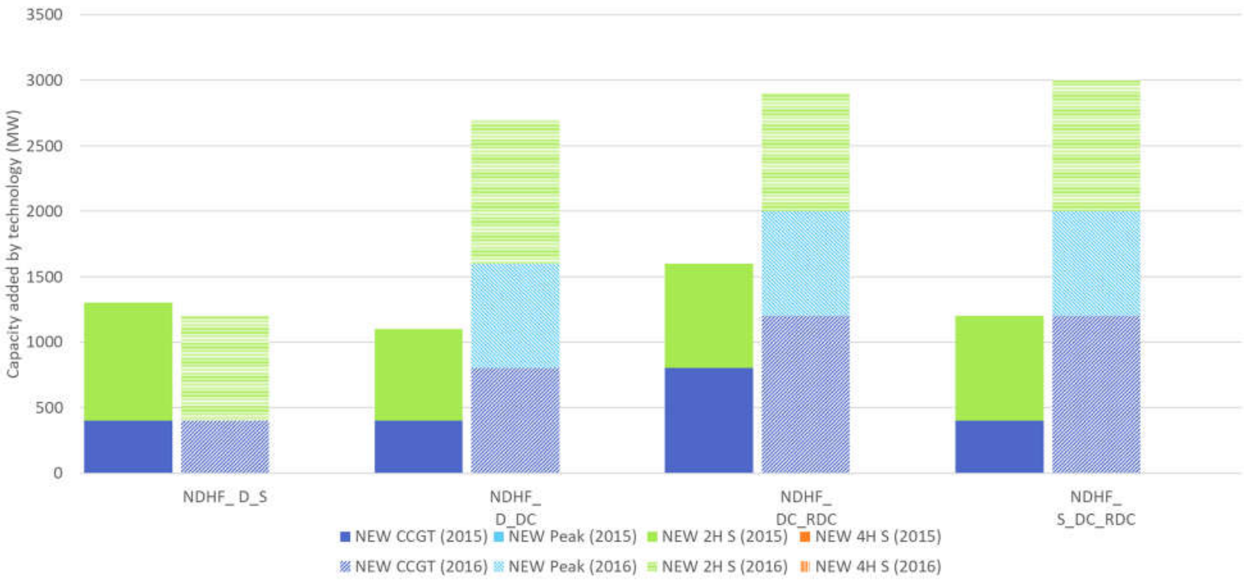


Figure 13 Change in average absolute percentage error of cost estimation, and average additional cost over best solution from the addition of future net demand series to clustering input

The largest difference in expansion plan shown in Figure 14 in comparison to Figure 11 is in the 2025 model year. In all cases significantly more storage capacity is selected in the expansion plan and in several cases additional conventional capacity as well. Where only storage capacity is increased, solution costs are improved, and the average additional solution cost is improved. Where the approaches that include ramp duration curve add additional CCGT capacity the solution quality falls.



(a)



(b)

Figure 14 Capacity expansion plan by representative day selection approach (a) 2020 expansion model (b) 2025 expansion model

3.8.3 Extreme point inclusion

Table 8 Key for case abbreviations

Abbreviation	**Case description**	**Abbreviation**	**Case description**
NDHF_S MRU	Historical and future year net demand as original chronological series with day featuring max ramp up included as medoid	NDHF_DC MRU	Historical and future year net demand duration curve with day featuring max ramp up included as medoid
NDHF_S GMMD	Historical and future year net demand as original chronological series with day with max greatest min max distance included as medoid	NDHF_DC GMMD	Historical and future year net demand duration curve as original chronological series with day with max greatest min max distance included as medoid
NDHF_S GMMD6H	Historical and future year net demand as original chronological series with day with max greatest min max distance over six hours included as medoid	NDHF_DC GMMD6H	Historical and future year net demand duration curve with day with max greatest min max distance over six hours included as medoid
NDHF_S MRD	Historical and future year net demand as original chronological series with day with max ramp down included as medoid	NDHF_DC MRD	Historical and future year net demand duration curve with day with max ramp down included as medoid
NDHF_S Fut MCRU	Historical and future year net demand as original chronological series with day with max consecutive ramp up included as medoid	NDHF_DC Fut MCRU	Historical and future year net demand duration curve with day with max consecutive ramp up included as medoid
NDHF_S Fut MCRD	Historical and future year net demand as original chronological series with day with max consecutive ramp down included as medoid	NDHF_DC Fut MCRD	Historical and future year net demand duration curve with day with max consecutive ramp down included as medoid
NDHF_S MinV	Historical and future year net demand as original chronological series with day with min value included as medoid	NDHF_DC MinV	Historical and future year net demand duration curve with day with min value included as medoid

In this section I report results for the inclusion of additional ‘days of interest’ to attempt to capture the representation of future system challenges. I focus on the case where the clustering is based upon future and historical net demand for the benefits discussed in the previous section.

3.8.3.1 Clustering results

The preselection of a cluster medoid to represent a cluster reduces the ability of the clustering algorithm to minimise the Euclidean distance (as it is now more constrained) although not greatly in most cases. For brevity these results, the NRMSD, average net demand errors, and average absolute scalars are included in 0.

3.8.3.2 Expansion planning model results

Net Demand w/ future shape + extreme points

Figure 15 compares the changes in expansion plan costs and errors in cost estimation between the different GEP model solutions with different ‘days of interest’ added to the clustering as medoids. In four cases both an improvement to cost estimation and to quality of the GEP solution over the full dataset can be seen. Three of these cases relate to ramping requirements (“max ramp down”, “max consecutive ramp down”, “max consecutive ramp up”). While these cases do not necessarily improve the representation of the ramp duration curve as a whole (as measured by the NRMSD) they all result in inclusion of values near or at the low end of the ramp duration curve which is not well represented without these medoids. The other medoid that improves the solution is the min-max value distance which can be seen in Figure 17 to increase slightly the storage built in two of the four cases (and unchanged in the other two). This demonstrates the importance of extreme points, as while including them generally performs poorly on metrics relating to how close is the representative data to the underlying series (the distance metric and RMSD) they can drive more important dynamics in the GEP model. Effectively in this case, the inclusion of these dynamics is significantly more important for the GEP decisions than the clustering would consider.

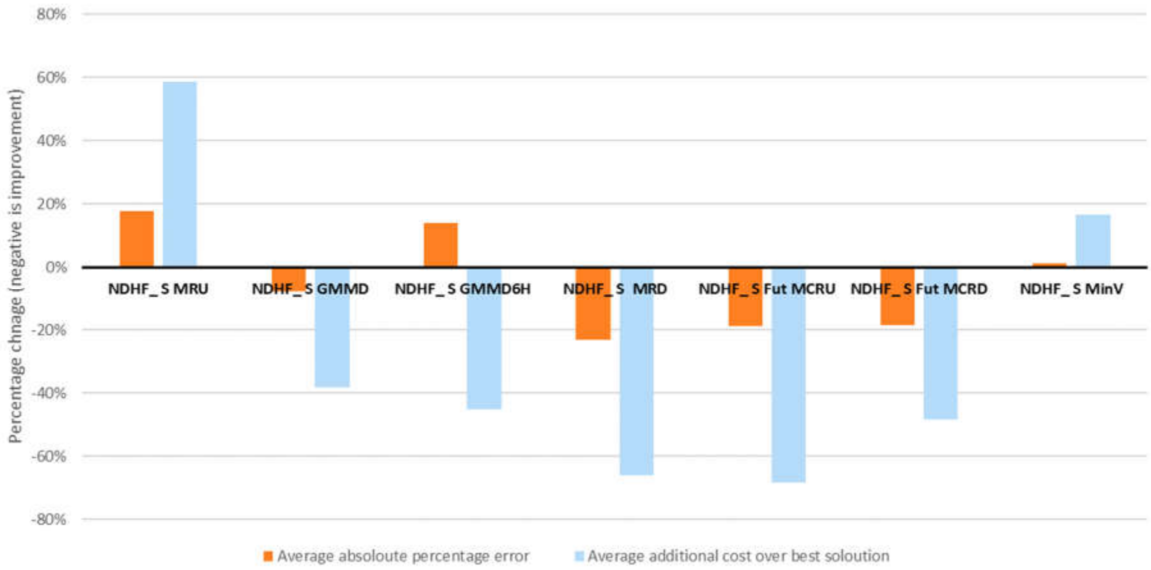


Figure 15 Additional cost over best solution found, all cluster year and model year combinations, with sum of squared cost estimation

Demand duration curve w/ future + extreme points

Turning to the case where I use duration curves to represent the data we can see only one case that improves results across both metrics, again the inclusion of the 'max ramp down', although several cases improve the optimality of the expansion plan. The improvement in solution optimality is seen in the cases using the 2015 load shapes where again this medoid results in the inclusion of the lower values of the ramp duration curve otherwise missed (and despite the lowest values of the net demand duration curve not being included, the addition of which can be seen here to only slightly improve solution optimality at the cost of cost estimation accuracy NDHF_DC MinV).

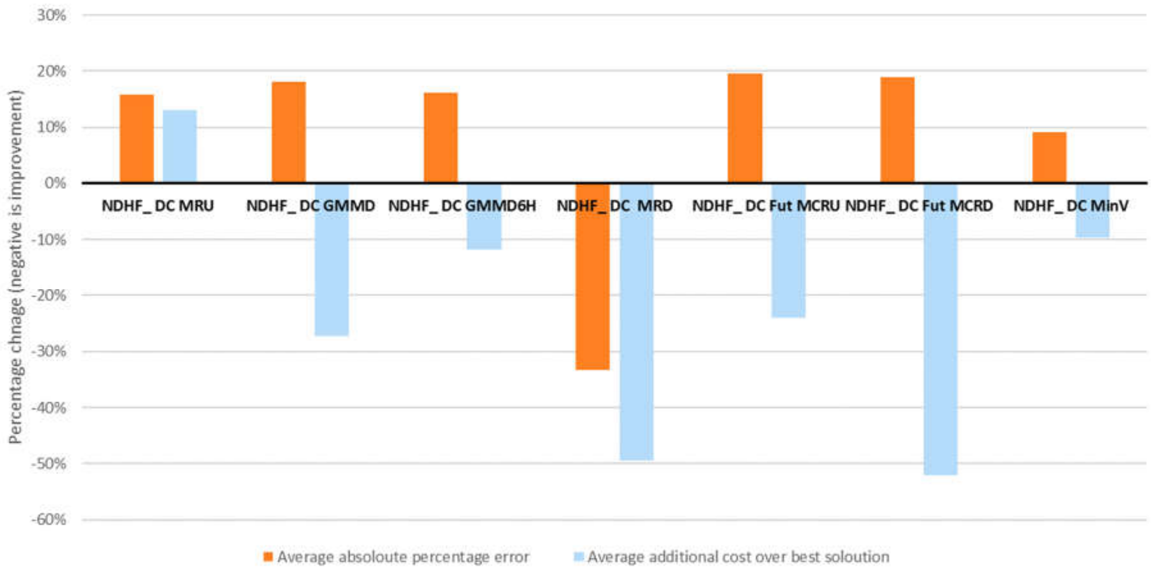
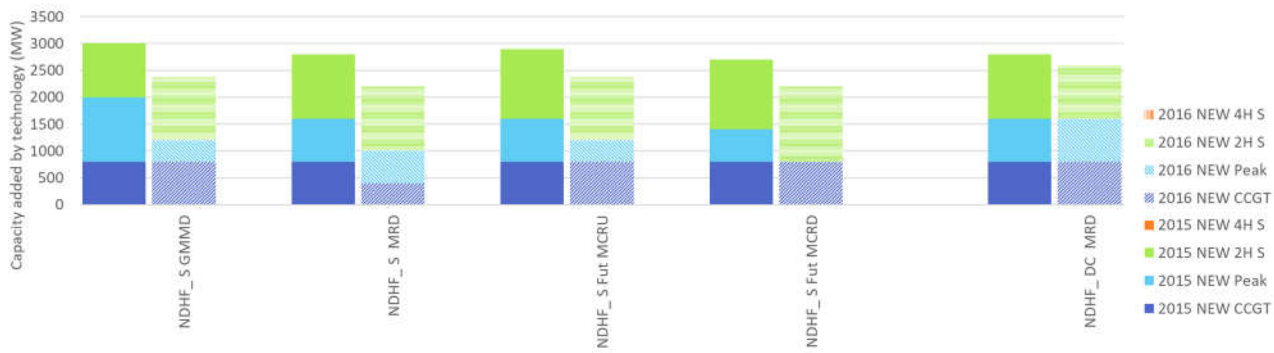


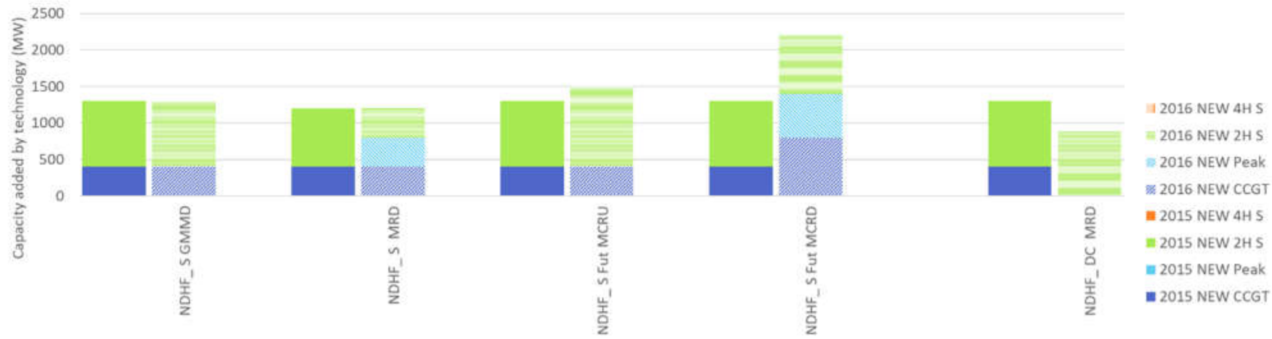
Figure 16 Additional cost over best solution found, all cluster year and model year combinations, with sum of squared cost estimation

3.8.3.3 Expansion planning results – change in build for extreme day inclusion

Figure 17 reports the build plans for the inclusion of extreme days where the solution is improved. For the case where the unordered net demand is used the greatest improvement comes from the 2020 model year based with the 2015 dataset. In all the improved build plans 400MW of CCGT capacity was removed and substituted with an additional 200-500MW of storage capacity depending on the case (with additional reductions in peaking capacity of 0-600MW) with the largest switch to storage associated with the greatest savings. In the duration curve case, the improvements in solution quality are found mostly with the 2015 dataset and are again associated with an addition of or substitution towards storage.



(a)



(b)

Figure 17 Capacity expansion plan by representative day selection approach (a) 2020 expansion model (b) 2025 expansion model

3.9 Discussion and Conclusions

In this chapter I have demonstrated the importance of carefully performing the clustering of representative days for generation expansion planning. While approaches tested based on net demand had the same peak load and total energy to be fulfilled in the GEP model, the generation expansion plans differed greatly among representative day selection methods in terms of both the total capacity to be added and the technology selected. The differences between the best and worst expansion plans range in cost between £340m and £650m (large in comparison to the modest investment in new units of £400m to £2bn).

I first found that clustering based upon net demand performed better than clustering upon the individual demand and renewable capacity factor series in this case study. Additionally, while I found that the addition of a ramp duration curve to the clustering did improve results in comparison to just a duration curve based on historical data, this affect may simply be forcing the selection of days that have high renewable penetration (which becomes more important as renewable penetration increases into the future) and is not the best means to capture ramping dynamics in the clustering.

Instead I found the largest improvement to the clustering was gained by including multiple series in the data used for clustering, specifically both the future (high renewable penetration) and historical (low renewable penetration) net demand associated with each day. When clustering based on historical net demand, the algorithm is prone to group together days that are high load - high renewable production with medium load - low renewable production days. While these are similar in the historical data set as renewable capacity increases into the future the days produce very different net load profiles. In particular, I found the clustering did not select days to represent the extreme low values of the net load duration curve without this approach. The inclusion of this future data improved both the ability of the represented day expansion plan to estimate the costs of the full year simulation and the optimality of the GEP model selected expansion plans.

Some additional gains were also found by ensuring that days containing specific ramping features were included as cluster medoids. I found it important to include the largest down ramp in both the duration curve and non-duration curve cases. I also found, however, no consistent benefit to including a number of other potential extreme points such as the minimum value, which I had expected to be important for driving curtailment costs.

Overall, I found the best results were when clustering was based on the chronological net demand series, as opposed to a duration curve based approach (or any combination of the two or ramp duration curves). This clustering should include both historical and future net demand series for each day and the medoids for peak net demand and peak single period ramp down should be included.

I found important technology mix implications for the expansion plans selected by the GEP model based upon the different approaches to clustering input data. In the majority of cases, improvements to solution optimality are gained where conventional capacity is substituted by storage capacity, indicating that an oversimplification of future system challenges biases the model against ES. Based on the specific ES technology selected it seems probable that this storage is being found beneficial for improving ramping costs as opposed to shifting large quantities of additional renewables between on- and off-peak periods. This finding is reinforced by the fact that the inclusion of medoids relating to ramping (particularly downwards ramping) improved the solution quality and led to a greater preference towards storage over conventional technologies. This result reinforces the importance of ensuring that the modeller accurately capture the challenges of the future system to guarantee that the right technologies are selected.

Finally, I presented a method for ensuring that the results from the clustering algorithm achieved the targeted level of net demand in the system without altering the underlying net demand shapes that are observed or predicted from historical data. This method performed well (achieving accurate net demand, peak net demand, and maintaining underlying net demand shapes) without requiring significant alterations to the output clustering weights in most of the cases (the approach performed less well where errors resulting from the clustering were large). The ability to ensure the underlying net demand shapes and ramp rates are accurately reflected in the GEP model was demonstrated by the importance of ramping related medoids to the results. Future work aims to improve the underlying clustering process, so this weighting approach can be extended to ensure that renewable capacity factors are accurately modelled for endogenous renewable build modelling.

Chapter 4

Assessing the costs of contributing to climate change targets in sub-Saharan Africa: The case of the Ghanaian electricity system

In this chapter, I generate a range of emission reduction costs for Ghana that provide an important benchmark for the relatively under-studied sub-Saharan region and identify drivers of these costs specific to developing countries.

4.1 Introduction

The provision of reliable, secure, sustainable and affordable modern energy for all citizens is key to poverty liberation and economic growth [96,97]. In the sub-Saharan African region alone, which hosts a population of over 950 million, more than 590 million are without access to electricity, making it the most electricity-poor region in the world [96]. It is estimated that most of the countries within the region have an electricity access rate of about 43%, while the average annual residential electricity use hovers around 488 kWh/capita, which represents about 5% of the electricity use in the US [98]. This recognition led to the adoption in 2015 of the new United Nations Sustainable Development Goals (SDGs), signed by 193 developed and developing countries with the objective of ensuring access to affordable and sustainable modern energy for all by 2030 [96,99]. Electricity demand is currently constrained by supply and with a low level of energy security, which is attributed to various factors including poor system reliability, limited infrastructure, fuel import dependence, and heavy reliance on fossil fuels, hydropower, and traditional biomass resources [100]. Indeed, the Paris Agreement also acknowledges this need and further promotes access to sustainable energy especially in Africa.

Ghana is one of the few countries within the sub-region which has been successful in reducing energy poverty, having shown long and strong political commitment since the launch of its National Electrification Scheme (NES) in 1989 [101]. The objective of the NES was to attain 100% universal electricity access by 2020. By 2017, the country had recorded an electrification rate of about 84% [102], which is one of the highest in the sub-region after South Africa. Electricity demand has experienced a steady rise with a projected growth rate of about 6% [103]. In 2017 alone, the country recorded a coincident system peak demand of 2,192 MW and total energy consumption of 14,177 GWh including losses [103]. This trend has been linked to factors such as increasing economic growth, population, electricity access, urbanization and commercial activities [104,105] making Ghana representative of the future of the sub-Saharan region. [23]As of 2017, the population of Ghana stood at 28.83 million, GDP was USD59 billion (current) while the annual GDP growth rate was about 8.14%. On the other hand, the installed capacity grew at a rate of 5.7% between 2000 and 2015 [106]. In 2017, the total installed capacity available for grid power supply stood at 4,310 MW [107] and the total energy supplied including imports was 14,178 GWh [103].

Though the current supply capacity seems enough to meet the system load peak, the country has historically experienced recurrence of power supply shortages as is typical of the region and has had to rely on expensive barge mounted power stations. The country also suffers from limited availability of natural gas and is often required to switch the operation of its gas fleet to alternative expensive and polluting fuels. Energy supply may pose a substantial challenge in the future especially in light of the growing demand, energy insecurity and climate change vulnerabilities. Investments in a few traditional fuel sources or technologies might be unsustainable in the long run looking at the global energy market volatilities. Consequently, fuel diversification is key for future electricity supply planning. Power system planning, which ensures energy policy and investment decision recognizes supply and demand-side management options, is critical to power generation improvement [108].

As is an increasing trend in the energy sector the system planner, grid operator, and electric utilities are faced with significant uncertainty in the decision making. The uncertainties in energy resources, demand dynamics, climate change, generation and

[23] https://data.worldbank.org/country/ghana

emission costs require robust models which support investment and environmental decisions in planning and projecting supply capacity expansion. The face of the traditional electric power system is changing due to the global consensus on cutting down greenhouse gas emissions to reduce the impact of climate change and sustain strategic primary energy security [109]. In this context, the development of indigenous renewable and clean energy sources is key to ensuring energy independence and self-reliance. Renewable energy sources (RES) such as solar and wind are being widely deployed internationally and there exists potential to expand their role in Ghana. RES are characterized by a high level of unpredictability and its integration as part of the electricity system is critical for future supply capacity planning. Ghana is endowed with substantial exploitable RES such as solar, wind, biomass and hydro [110]. The parliament of Ghana promulgated the Renewable Energy Act in 2011 to enhance the development, utilization and adequate renewable energy (RE) supply to meet 10% of electricity demand by 2020 [111,112]. Howbeit, the pace of development of non-conventional RE has been slow. As at the end of 2017, the total on-grid installed capacity stood at 22.5 MW constituting 0.5% of the domestic generation capacity [107].

Ghana is then faced with the challenge of expanding its supply of electricity to the population, providing for the fast growing energy demand of a developing country, addressing existing and future energy security threats, all with a relatively low budget and a high cost of financing (both public and private). To examine the future trajectory of the system we can turn to the use of modelling.

4.1.1 Generation expansion planning in Ghana

In Ghana, studies on supply models are reasonably scarce. Awopone et al. [113] used the Long-range Alternatives Planning (LEAP) tool to compare the existing supply capacity expansion options to a proposed scenario of higher penetration of renewable energy (RE) for a period between 2010-2040. In all, four scenarios were developed namely: Base Case scenario, Coal scenario, Modest renewable scenario and High renewable scenario. The base case scenario was based on Ghana's expansion plan which included characteristics that do not set targets for emissions and hence no restriction on the use of fossil-based fuel. Additionally, LEAP is limited in its ability to measure short term VRES dynamics and the build of VRES is essentially an input selected by the modeller. The results indicated that though the high penetration of RE requires higher capital investment, the fuel cost savings over the span of the study period lead to significant economic benefits. The study also indicated that to meet future energy needs, the existing generation system ought to be expanded more than 5 times from 2014-2040 indicating the scale of the challenge facing Ghana.

Afful-Dadzie et al. [114] developed a multi-period stochastic mixed-integer linear programming (MILP) optimization model for studying long-term generation capacity planning under demand uncertainty in the face of a budget constraint. The results indicated that financing new energy generation capacity with a dedicated periodic budget allocation of 0.75% of GDP equivalent, 95% of annual electricity demand could be met adequately for the period 2016-2035. Overall, it was established that as financial constraints get tighter, planning decisions on new generation capacities become costlier in comparison to when adequate funds are available. Nevertheless, the uncertainties in fuel prices, hydro inflows, as well as a proper representation of load curve to allow for base and peak plants, VRES dynamics, and ramping challenges were all not considered.

Diawuo and Kaminski [115] designed a short-run cost minimization model of Ghana's energy system developed in the General Algebraic Modelling System (GAMS) as a linear programming problem. The model sought to estimate the generation cost and emissions generated when different primary energy resources are utilized based on different created scenarios anchored on RE penetration, demand dynamics, emission cost and fuel prices. A comparative analysis of the scenarios indicated that the introduction of more RE had the highest average unit cost of generation of close to 0.14 USD/kWh, representing a marginal increase of 115% with a CO_2 reduction of 99.89% relative to the reference scenario.

Relative to the time frame and resolution of the developed models studied in Ghana, the following two situations prevail:

1) some of the models focused on a multi-year analysis of the energy systems without considering the short-term dynamics of the model resources thereby having low temporal resolution;
2) other models also focused on only hourly temporal resolution for a period of one year and capture resource dynamics providing high temporal resolution, but were unable to make long-term decisions;
3) the models do not directly address the cost of decarbonisation (although they test decarbonisation costs of different assumptions) or deal with the large range of uncertainties present in the energy sector.

In this study, the GEP problem is integrated with an hourly UC problem with the aim of integrating RE and selecting future generation plants which have the least discounted cost and delivers low emissions. An investment decision on supply capacity expansion on a long-term is thus optimized while taking into consideration the dynamics of hourly demand and granularities in electricity supply resources. The model is additionally amended to incorporate the practice of fuel switching, common historically in Ghana due to the limitation in the supply of natural gas and neglected in the above studies. Usually, long-term models which ignore short-term variability can result in an underestimated RE integration cost, which in this case could have critical financial ramifications in planning purposes, especially for developing countries.

4.1.2 Contribution of the study

Long-term energy planning models have been carried out for different countries around the globe, exploring different objectives but with the goal of enhancing effective policy decision making (see [11] for a review). As I have shown specific scientific research dedicated to Ghana in relation to a future supply capacity model is limited, sporadic, or lacking in detail, particularly with regards to the introduction of RES and quantification of the cost for reducing carbon dioxide (CO_2) emissions from electricity generation. Consequently, this study aims to assess how the electricity system must evolve to meet the dynamic characteristics of future electricity demand and determine the role of RES in the energy system.

Additionally, the situation seen in Ghana and many developing countries represents a challenge for the aim of reducing worldwide greenhouse gas emissions. The challenge of meeting a fast-growing demand for energy with challenging financial conditions may not be conducive to this goal. Therefore, the second aim of this study is to quantify the different costs of fulfilling specific CO_2 emission targets to provide insight into the cost of decarbonisation in fast growing developing countries. This analysis allows insight into the unique opportunities and challenges for reducing CO_2 emissions that may not be represented in studies of developed countries.

Finally, to provide a broader understanding of these questions the study aims to examine the level of uncertainty in these estimates by examining different scenarios. This analysis will then provide an understanding of the broad drivers of both the transition of the energy system and the costs of decarbonisation, where differences from typical results for developed countries may be found[24].

---

[24] Exploring a range of scenarios also helps for comparability to other studies in general as it becomes clear how differences in assumptions between this study and the literature influence the results of this work.

4.2 Generation expansion planning model

For the purpose of this study, the mixed-integer linear programming (MILP) generation expansion planning model (GEP) described in Chapter 2 was adapted and validated for the Ghanaian system, in order to test six scenarios between 2020-2040 under different CO_2 emission targets to assess their impact on future supply capacity planning.

The model outlined in Chapter 2 is applied with the following modifications:

- The model is deterministic (with objective function Equation (1)), as such the indices *s* is not required on any of the variables. Multiple scenarios are simulated independently in this way, however, as I do not assess the probability of these scenarios it is not possible to combine them into a probability weighted expected result.
- Each existing power station is modelled individually, therefore, the clustered unit commitment formulation is not used for these units which are heterogenous in their use of fuel, efficiency, and in the case of hydro annual production.
- Integers variables are enforced on build and commitment decisions, therefore the linear relaxation formulation is not used.
- The extension of the model to allow for fuel switching power plant and the constraint on available fuel discussed in section 2.3.3 is included.
- The emission constraint from section 2.3.4 is also applied to test the costs of achieving different emission reduction targets.

4.3 Scenario modelling and assumptions

Scenarios are created to project an alternative image of a possible future outlook. The scenarios developed in this study are intended to explore the likely developments of market, financial, and environmental forces that are important for decision making. Six scenarios are developed but classified under four broad scenario frameworks, namely (1) Base case, (2) cost of RE, (3) gas price and availability, and (4) discount rates. Three emission target pathways are developed, namely no emission ceiling (no limit) and restricted emission targets (150g/kWh and 100g/kWh), emission restrictions are enforced starting in 2030. In addition, two areas are tested for sensitivity; the level of energy demand, and the availability of additional non-indigenous technologies. The different possible scenario combinations are shown in Figure 18 and the broad scenario scheme is discussed next.

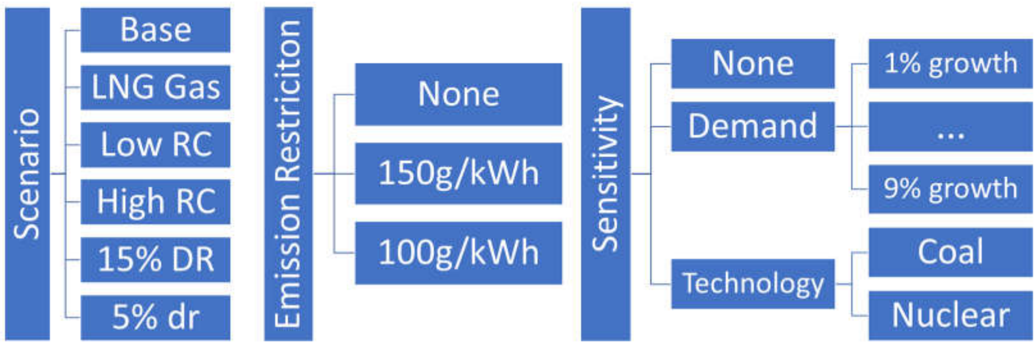


Figure 18. Scenario and sensitivity combinations

4.3.1 Base case

The existing generation technologies summaries are outlined in detail in Appendix B, and are included in the model with fuel efficiencies and availabilities estimated from historical fuel consumption and generation output [103]. Other operating characteristics are based on [116], and [117]. An important driver of generation share in Figure 19 is the current constraint on the availability of natural gas. However, over the period of the study this constraint will be eased significantly by projects under construction to come online in the 2022-2023. Table 9 shows the technologies considered for expansion planning.

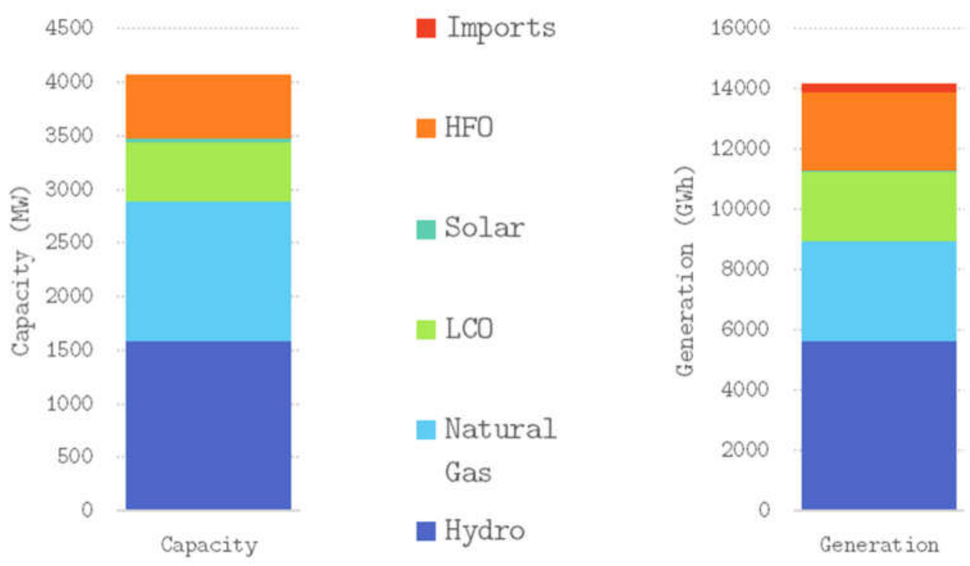
Table 9 Expansion planning technologies

	Unit Size (MW)	Efficiency	Ramp rate (MW/min)	Start Cost ($/MW)	Variable O&M Cost ($/MWh)	Fixed Cost ($/kW-yr.)	Annual availability	Fuel
CCGT	400	59%	4.7	56.7	2.8	10.6	82%	Natural Gas, LCO
OCGT	200	34%	7	56.7	7.1	12.2	92%	Natural Gas
Peaker	210	39%	7.3	36	7.1	12.2	90%	HFO
Li-ion battery storage	100 (90 while charging)	90% (round trip)	N/A	0	0	0	100%	4-hour energy storage
Solar	100	N/A	N/A	0	0	19.9	N/A	N/A
Onshore wind	100	N/A	N/A	0	0	43.6	N/A	N/A
Biomass**	100	25%	1.5	78.5	5.6	111.7	84%	Biomass
Hydro**	100	N/A	N/A	N/A	0	37.1	22%*	N/A
Coal***	500	34%	3	78.5	5.0	33.0	85%	Coal
Nuclear***	1000	100%	N/A	78.5	2.3	100.9	90%	Uranium

* Annual production limit based upon historical existing Bui hydro power station generation.

** A maximum of 600MW of new biomass and 900MW of new Hydro is allowed based upon estimates of potential capacity from USAID [118].

*** The inclusion of coal or nuclear as a technology option is tested as a sensitivity and not included in the base case.



(a) (b)

Figure 19 Ghana existing (2017) Generation capacity (a) and generation (b) by fuel.

4.3.2 Renewable energy costs scenarios

With the historical trend in the cost of RE, a higher or lower presence of indigenous RE resources ought to be investigated as part of the overall solution for energy affordability and security. I explore high and low renewable and storage construction costs paths based upon the analysis and costs provided by [117]. Cost assumptions used for the scenarios developed under renewable energy cost are included in Appendix B. The model increases the construction costs slightly for renewable technologies as uptake increases, creating a supply curve, reflecting the fact that as the most beneficial sites are developed either development costs will increase, capacity factors fall, or transmission connection costs increase (otherwise not included in this analysis). This approach is preferred to that of enforcing hard limits on available renewable wind and solar capacity at around 1GW of [118].

4.3.3 Gas price and availability scenario

Next, I explore a scenario where Ghanaian gas becomes tied to world markets through liquified natural gas (LNG) imports and exports. This removes the limit on gas supply but increases the price to match that of a forecast world price of LNG. The gas prices assumed are shown in Appendix B, the data sources used are [103,118,119].

4.3.4 Discount rate scenarios

The discount rate to be applied in generation expansion planning for developing countries is an important consideration. For a base case, I use 10% for all generation types in line with the discount rate used in the Integrated Power System Master Plan for Ghana document [118] and current government bond rates in Ghana, I additionally test an increase to 15% and decrease to 5%.

Poncelet [10] discusses the role of discount rate in GEP models. When attempting to use the model for descriptive or predictive analysis the discount rate faced by investors (either centralised utilities or market participants) should be used so that results reflect market outcomes. For normative analysis a social discount rate, which is typically lower, would instead be used.

4.3.5 Demand projection

The demand projections are based on the demand forecast by the utility company for the period 2020-2028 as reported in the 2019 Ghana supply plan [103][25]. For the period 2028-2040, the average forecast growth rate of 5.6% [118] is used. The demand projection

[25] The utility company uses a log-log linear regression model based on historical annual GDP as the only independent variable and historical annual electricity consumption in GWh as the dependent variable [118]. The projected demand is comparable and within the ranges of related studies carried out by Energy Commission & USAID [118] and Awopone et al. [113].

is compared to alternatives found in the literature in Appendix B.

Historically, in Ghana, the measured electricity demand is not reflective of the full grid connected load primarily due to power theft, metre malfunction, non-metering, self-generation, power outages due to generation, transmission or distribution disruptions, and electricity price variations which affect electricity consumption over time, etc. Although, the projected demand might not fully reflect the entire demand in the system, it captures to some extent the supressed demand over time. This important assumption is tested for sensitivity, both to quantify the contribution that energy efficiency can make and to determine the potential effect of unsatisfied demand increasing the energy requirement over time.

4.4 Results and discussion

In this section, the model is first validated to ascertain the accuracy in representing current system behaviour before being applied to the problem of system expansion between for the period 2020-2040.

4.4.1 Model validation and accuracy

The model is validated for its accuracy in determining the evolution of electricity generation for 2017 and 2018[26]. To assess the accuracy, the output of the model for the year 2017 and 2018 on electricity generation were compared to: the actual data of 2017; and a generation plan by the utility/grid operator, GRIDCo, in 2018; as presented in Figure 20. The model reproduces the generation by fuel share within acceptable accuracy,

[26] Actual annual demand and annual hydro availability are used in the back-cast model.

after the imposition of the fuel constraint. The model particularly well reflects the results of the existing planning model, meaning the results should be directly comparable to existing forecasts provided by the grid operator. However, the model's use of average availability to derate capacity factors is unable to reflect situations where multiple power plants may be on an outage at similar times, which in the real world leads to slightly higher generation from the LCO units or even expensive diesel units.

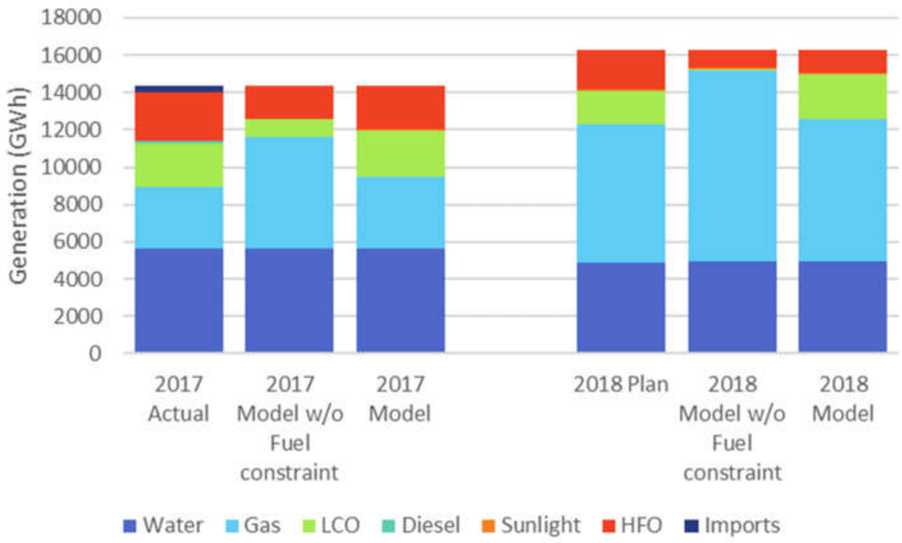


Figure 20 Comparison between modelled and actual electricity generation

4.4.2 Evolution of power generation capacity under different CO_2 emission targets

The additional capacity added by the model is presented in Figure 21 under the assumptions of the base case. This capacity is in addition to the existing dependable generation capacity which as of 2017 stood at 3868 MW [103]. To meet the rapidly growing demand under CO_2 restriction and no restriction (Base_no limit, Base_150g/kWh, Base_100g/kWh), the model estimates additional capacity required to be 4000 MW, 6300 MW and 9500 MW by 2030, and then to 9850 MW, 14600 MW and 21800 MW by 2040, for the base scenarios respectively.

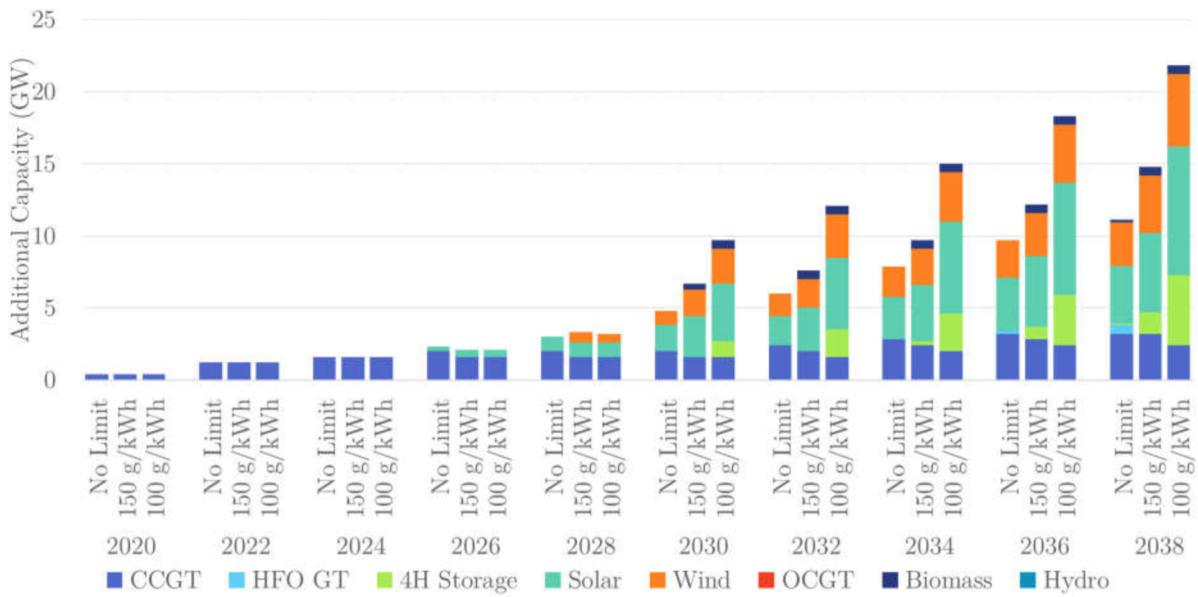


Figure 21 Additional generation capacity in the base case scenario under different emission restrictions

Natural gas was previously wholly imported but currently available domestically and has typically been the fuel of choice for thermal generation. While there remains available gas sources, the trend continues in the most part. However, new generation capacities are diversified with a combination of renewable energy technologies and storage systems in the longer term. In all scenarios, the share of CCGT decreases over time, though with a gradual increase in its capacity. In this base case the model waits until 2030 to add a modest amount of new VRES, which is wholly a mix of solar and wind, and never building more than can be handled by the flexibility of the current hydro stations and so no place is found for energy storage. In the final years of the study, without any emission restriction, a small amount of additional peaker HFO capacity is added, indicating the limit on natural gas supply is reached despite the additional natural gas supply becoming available during the study.

Examining the imposition of a maximum emission intensity in the year 2030 it can be seen that this has a significant effect on results. Firstly, it brings forward slightly the investment in renewable technology to the year 2028, even before the constraint, which displaces investment in an additional CCGT. As would be expected to meet emission targets the investment in RES is greatly increased with biomass and ES added to the technology mix.

Figure 22 to 24 provide the build plan for the period to 2040 under the six different scenarios under the assumptions of no emission restrictions, a 150g/kWh restriction starting 2030, and a 100g/kWh restriction starting 2030 respectively.

Figure 22 provides a comparison of the expected build under the current policy where there are no specific carbon reduction targets. In this case the total capacity and proportion and type of RES depends highly on the assumptions. Perhaps unsurprisingly the cost of RES is the greatest driver of these results with the total quantity of RES built by 2040 varying by a factor of 5 (between 2000MW in High RC and 10300MW in Low RC). The discount rate is also an important driver of results. The higher the discount rate, the lower the economic viability of all types of additional capacity but especially high capital cost technologies such as RES. Comparing the 5% and 15% discount rate scenario, we can see in practice that for developing countries, high discount rates in the short run will mean they invest less in (or delay the construction of) conventional capacity and in the long run less renewable capacity. This should be noted as while high discount rates certainly prevent the adoption of renewable technologies, attempting to remedy this issue for emission reduction reasons can have unintended consequences. As seen here, targeting reduction in borrowing rates (or increasing access to capital markets) can potentially “lock in” the build of additional conventional capacity by developing countries in the short term; this result remains even when a moderate emission target is imposed as in Figure 23 and was known by investors since the start of the study. This result is achieved in a model that has perfect foresight from the year 2020 that a constraint will occur in 2030 and may be even higher in a scenario where a policy to restrict emissions is implemented with less notice. The introduction of additional gas availability, even at a relatively high price, increases the quantity of CCGT capacity built, particularly in the short run.

Figure 23 shows the changes that result from the introduction of a modest emission target starting in 2030. Again, the results do vary depending on the assumptions but that as expected the quantity of RES developed increases significantly across all scenarios starting in 2030. In two scenarios there are changes in capacity up to four years earlier (5% DR and LNG Gas) than the introduction of the target, with both increases in RES and decreases in conventional capacity. This provides evidence of the value of announcing a credible emission target as early as possible as the effect over a lifetime of a generation unit can change investment decisions, and increase the proportion of RES, and reduce carbon emissions in anticipation of the target.

Figure 24 indicates the new capacity requirement in the six scenarios under the most restrictive emission target of 100g/kWh. As is expected the strict target results in the highest build of new RES, by 2030 the Low RC produces the highest RE capacity estimated at 7300 MW (52% is solar, 17% wind and 6% biomass). Interestingly despite the fact that RES require a high capital cost investment the proportion of RES is relatively high in the high discount rate scenario in 2030, as a result of lower investment in CCGT capacity through the pre-2030 period. By 2040 the build plan required to reach the emission target is relatively similar across all scenarios, with slight differences based on the cost of renewables which either lead to a slightly higher renewable share (low costs) or a slightly lower total capacity and the development of additional hydro power (high costs). This result indicates that emission abatement costs will be driven more by the counterfactual and may drive interesting differences compared to developed countries. Broadly, the optimal system for a restrictive emission target may be similar across countries, but the cost to the country to achieve it will depend on what they would have done in the absence of an emission obligation.

Interestingly, despite some available capacity, the only scenario that suggested the construction of new hydro plant was the High RC scenarios where wind and solar costs are relatively high. However, all scenarios with emission constraints suggest levels of RES development where the economic benefits of backup/battery storage capacity become important. This result indicates that simple cost comparisons (such as comparing levelized cost of energy) are not appropriate between VRES and conventional technology as the cost of storage capacity must be included with higher levels of VRES integration. The Low RC_100g/kWh scenario required the largest capacity of storage of 1200 MW and 5200 MW for 2030 and until 2040 respectively. The magnitude and time for battery discharge vary across different technology mixes but the period for charging the battery storage ranges from 9 am to 3 pm especially when there is the availability of solar and wind resources (see Appendix A).

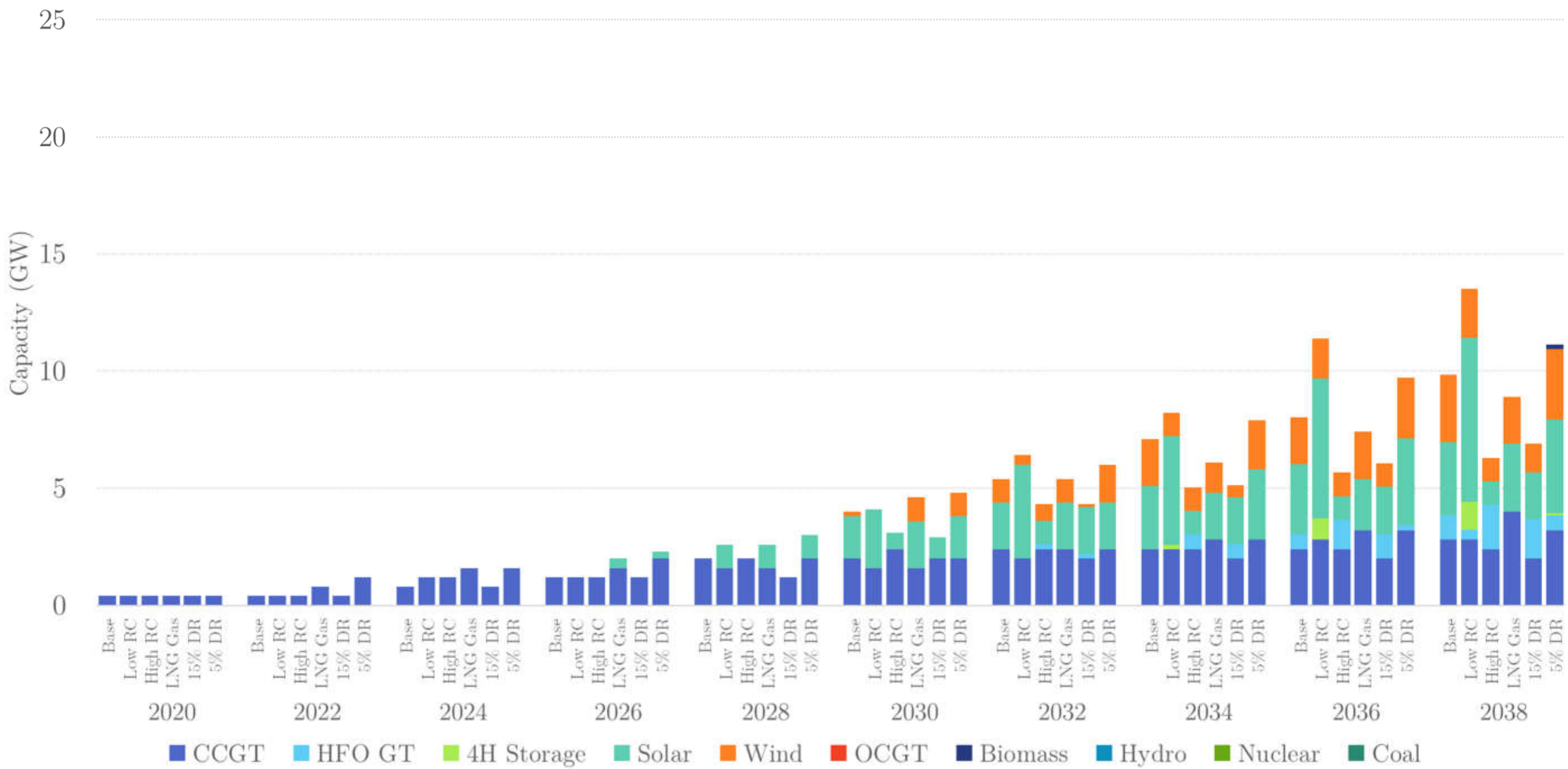


Figure 22 Additional generation capacity in the all scenarios without emission restrictions

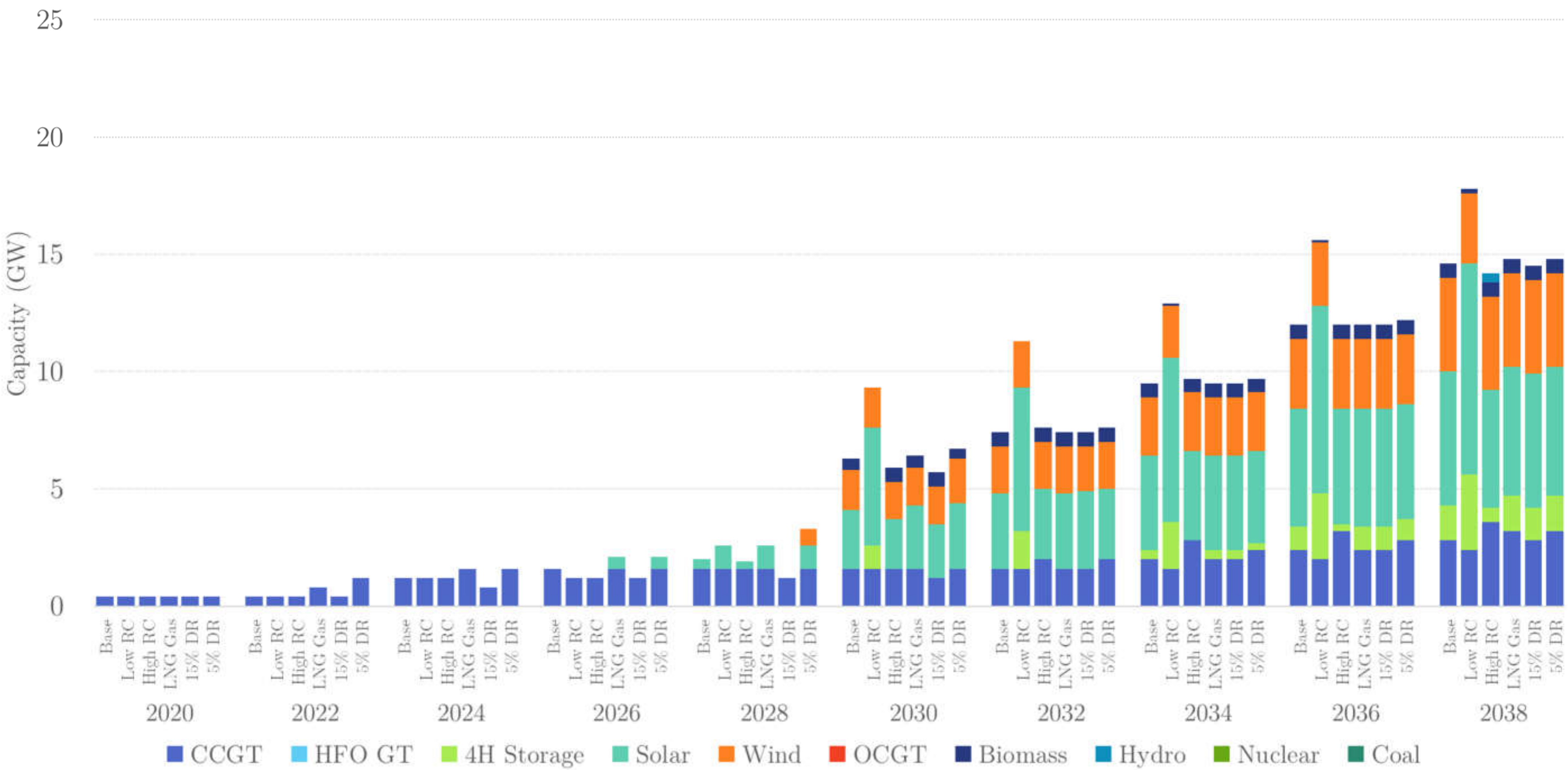


Figure 23 Additional generation capacity in the all scenarios with 150g/kWh emission restriction starting 2030

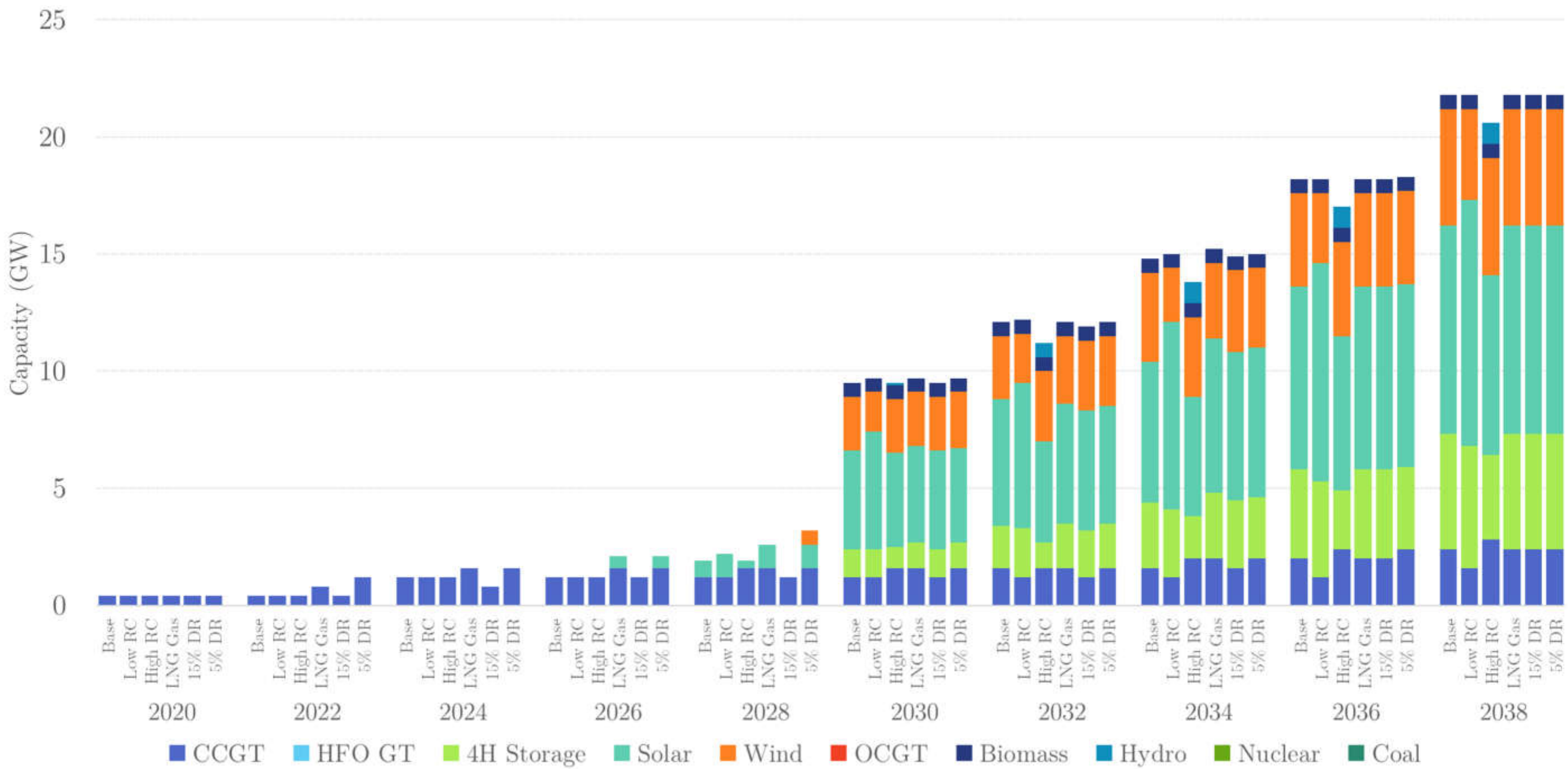


Figure 24 Additional generation capacity in the all scenarios with 100g/kWh emission restriction starting 2030

4.4.3 Generation mix

The operationalization of generation stations depends on the variabilities of fuel availability, generation cost, type of needed load (peak and off-peak), and starting costs or ramping restriction, and thus provide a varying quantity of energy to match demand[27]. Figure 25 shows the electricity generation mix for each scenario in the years 2030 and 2040.

In the base scenarios, by 2030, the share of gas and electricity production in Base_no limit scenario increases slightly over current values and is high at 66.3% while with the introduction of emission targets this value falls to 40.8% and 25.7% in the Base_150g/kWh and Base_100g/kWh scenarios respectively. The share of new RES sources (RES excluding existing hydro) is increased varying from 10.9 to 47.8% under the base case. The government's 10% RE integration target policy is achieved then during this period with or without emission targets, however this result is undermined by either high renewable costs or a high discount rate.

Across all scenarios the imposition of emission targets leads to a similar level of conventional generation (a maximum possible level), however, the mix of RES that make up the difference does vary across conditions. The exception to this result is in the case of the import of LNG gas without any emission restriction, which results in a lower share of natural gas generation in the year 2030, due to higher prices, and a higher share in 2040 as a result of increased availability. In a number of scenarios without emission

[27] Demand here includes additional demand required to meet losses induced by charging and discharging energy storage. Storage generation is included here (not net storage generation which would be negative taking charging into account).

restrictions and without the availability of imported gas, so much additional CCGT capacity is built by 2040 that fuel switching is again required and a small portion of generation is met by generation from LCO fuel, which has significantly higher CO_2 emissions.

Figure 25 Generation mix by year (2030, 2040) and emission target

4.4.4 Discounted total cost for all scenarios

The discounted cost, which is the objective function of the model formulation is the sum of the investment, fixed, and variable costs related to power supply discounted to the base year 2020, and measures the cost-effectiveness and economic performance of the scenarios over the entire model time span. The lower the cost, the more cost-effective the scenario is[28].

Figure 26 presents the total discounted cost and shows that when using the base discount rate in both unrestricted and restricted situations, the Low RC scenario had the least cost followed by the base case, High RC, and LNG gas scenarios. Although with the strict emission target the cost of the High RC scenario approaches that of the LNG gas scenario. It also can be seen that the additional costs of meeting emission targets are significant but low compared to the total cost of operating the energy system over the 20-year study period. It is important to remember, however, that part of this result is due to discounting, with the majority of the additional costs starting around the year 2030, the contribution to present value will be significantly increased where discounted rates are lower.

The costs of the scenarios where discount rate changes cannot be directly compared to the remaining scenarios here, as the discount rate used to calculate present value changes. For example, the 15% DR scenario has the highest undiscounted costs, as the system invests less in new technologies that would be economic at lower discount rates, however the value shown here is the lowest. I control for this difference in the calculation of abatement costs in the next section. However, it can be seen that the proportion of costs associated with investment depends highly on the discount rate as would be expected. With a low discount rate, and a high level of emission restriction, investment costs approach nearly 50% of total system as costs as more RES capacity with no or low variable cost is introduced.

[28] Comparison between scenarios with different discount rates cannot be made easily here as the scenario discount rate is used to

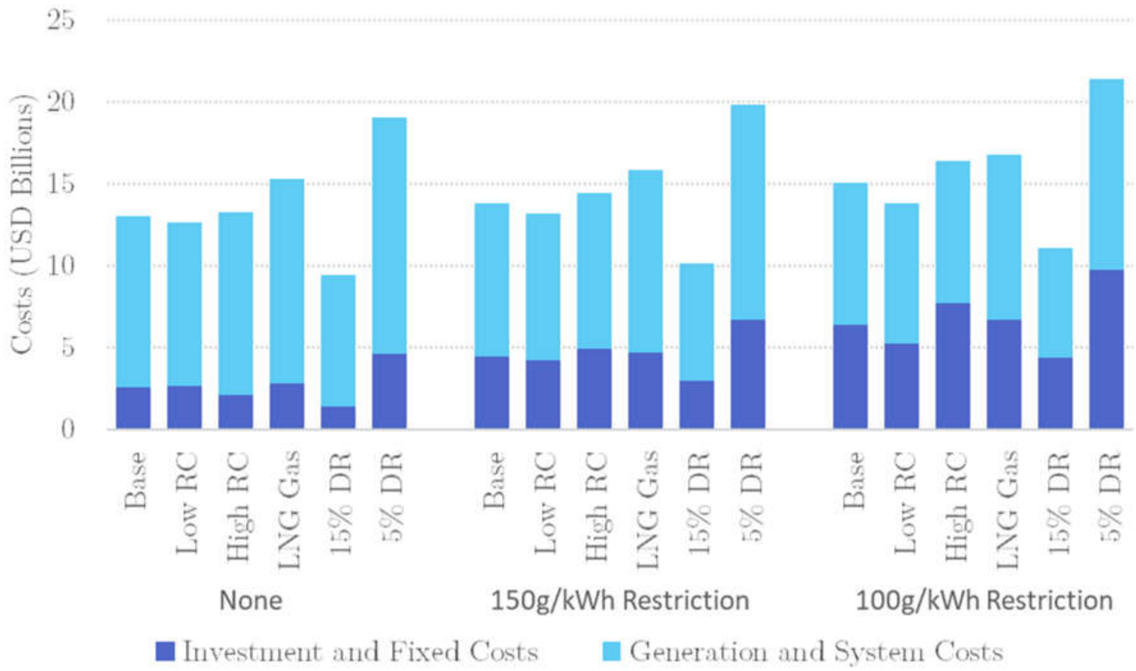


Figure 26. Discounted cost

4.4.5 CO_2 Emissions and Abatement Costs

Ghana, and Africa in general, contribute least to greenhouse gas emissions but they are most vulnerable to climate change impacts like droughts [98]. Ghana's emission is less than 0.1%[29] of the global CO_2 emissions. However, the growing demand for energy will likely lead to increases in Ghana's contribution to global CO_2. Yet this growing demand also presents an opportunity, the incremental cost of investing in RES as opposed to conventional capacity is lower than full cost of RES. So, countries that are growing can reduce emissions at a lower marginal cost than countries where demand is flat or falling.

Table 10 shows the total CO_2 emissions from the generation of electricity over the study period. In the Base scenarios (Base_no limit, Base_150g/kWh, Base_100g/kWh), the

[29] https://en.wikipedia.org/wiki/List_of_countries_by_greenhouse_gas_emissions

cumulative CO_2 emission until 2040 varied between 103 and 164 mt. Compared to the other scenarios, Euro gas scenarios ranged from 100-157 mt, Low RC scenarios, 103-149 mt, High RC scenarios, 103-189 mt, 5% DR scenarios, 99-148 mt and 105-188 mt for 15% DR scenarios. For the most part the total emissions when constrained after 2030 is similar across all scenarios, as there is no scenario where it is either cost effective to reduce emissions below the target or greatly reduce emission in the period before the target is imposed.

The main difference between scenarios exists then in the case of no emission restrictions which then importantly drives much of the result for the abatement costs also found in Table 10. Across all scenarios the 5% DR produces the highest cost to achieve the initial modest emission target. This result initially seems counter intuitive, that low financing costs which may be thought of as crucial to investing in RES actually leads to high abatement costs. When looking at the levels of greenhouse gas emissions it is easy to see why this is the case, the low DR leads to higher investment in RES without any restriction and then a lower level of emissions. So, while the cost of meeting the 150g/kWh target is similar the relative emission reduction is lower, and the marginal abatement cost is higher. A similar result is found when comparing the 15% DR to the base case or 5% DR case. This has an important implication for reducing worldwide carbon emissions. Because developing countries face high financing rates the level of RES, they can be expected to achieve is low and a small incentive may be able to increase this deployment of RES. This may be more cost effective than the same incentive in a developed country which already has a high base level of RES installation.

A similar effect is seen in the case of high RC where the cost to meet the initial target appears low, simply due to a high level of conventional capacity and emissions in the counterfactual case. In this scenario the incremental cost to increase to the higher target is extremely high, reflecting the cost of renewable energy.

Across all scenarios the cost to meet the move from the 150g/kWh to the 100g/kWh target is significantly higher than the imposition of the initial target. This result is driven by a number of effects, firstly moving to the initial target some conventional capacity that would be built is displaced and replaced by RES. The additional cost of the substitution of technology to be built is cheaper than the straight additional cost of RES making meeting this initial target more economic. In addition, some RES technologies (such as biomass here or hydro) are limited in the additional capacity that they can provide which is developed entirely in meeting the initial target. Finally, the integration costs of VRES increase with the total quantity of VRES on the system. For example, here I find that beyond a certain level of VRES more ES needs to be added to the system when adding VRES which results in a significant cost. All of these results mean that reducing emissions is incrementally more costly in a given system, but particularly, a growing system as found in Ghana. On a worldwide scale this means that it is likely more cost effective that all systems seek to reduce CO_2 emission, as opposed to this effort being concentrated in developing countries as is currently the case. Again, this result means that relatively small incentives could achieve much greater emission reduction in developing countries, which have less ability to invest in even small amounts of RES, than in developed countries.

One additional interesting result is that exposure to the international gas market in the long term may reduce the relative costs for the developing country to decarbonise as it becomes more costly to either import gas or use indigenous gas for national consumption.

Table 10 Cost of reduction for scenarios

Scenario	Total Emissions (million tonne) None	150g/kWh	100g/kWh	Discounted cost of reduction (USD/tonne)* None to 150g/kWh	None to 100g/kWh	150g/KWh to 100g/kWh
Base	164	122	103	19	33	64
Low RC	149	122	103	18	25	35
High RC	189	123	103	18	36	99
LNG Gas	157	120	100	15	27	50
5% DR	148	118	99	25	39	62
15% DR	188	124	105	11	24	64

* Costs for this calculation are all discounted at the same 10% rate to improve comparability between scenarios where discount rates are adjusted.

4.5 Sensitivity analysis

The sensitivity analysis is conducted for demand as a model parameter and exogeneous based power plant technologies focusing on nuclear and coal. These sensitivities are intended to test the potential uncertainties and its economic and emission impact on power sector planning.

4.5.1 Demand sensitivities

The demand sensitivities test the effect on decarbonisation results of high and low energy intensities. Here, the high energy intensity (6%, 7%, 8%) is assumed to be characterised by a relatively rapid surge or trend in population, economic growth, electrification rate, urbanisation rate, appliance ownership rate, latent demand, new energy uses such as the use of electric vehicles, etc. and the impact of "rebound" or "backfire" effect of energy efficiency implementation strategies.

On the other hand, the sensitivity analysis based on the demand growth rate below the baseline (1%, 2%, 3%, 4% and 5%) representing low energy intensity is used to assess the long-term impact of demand-side management (DSM) measures on power system capacity and generation. Ghana has a history of promoting DSM and energy efficiency and it's one of the two countries within the sub-Saharan region which has successfully established and implemented the minimum energy performance standards and labelling for household electric products such as lighting (CFL), air conditioners, refrigerators and freezers. The introduction of DSM strategies to reduce electricity demand in Ghana can have a critical impact on the investment strategy in building new generation capacity, fuel mix diversification and greenhouse gas policy considerations. With DSM some of the decision on building additional generation capacity can be delayed.

This suggests that if more investments were made *in situ*, some of the generation plants would have to be shut down, thereby reducing their capacity factors and hence making them less cost-effective and ultimately becoming economically not viable. Table 11 shows that with more aggressiveness in the implementation of DSM, the capacity needed to be installed is reduced significantly, reducing the discounted cost by 10-45% while emission is reduced by 9-44% in comparison to the base scenario.

The overall effect on abatement costs is interesting. At high levels of growth more conventional capacity is required in the base case which can be substituted for renewable technology and decreases overall abatement costs. At slightly lower levels of growth (5%, 4%) this substitution effect is lower and so abatement costs are higher, however, at very low levels of growth it becomes relatively low cost to achieve targets (given Ghana features a high level of existing hydro generation) and abatement costs fall. These results demonstrate that energy efficiency and DSR can play an important role in reducing both cost and emissions in growing developing countries.

Table 11 Comparison of Base scenario vs. demand growth sensitivity by 2040 (% comparative difference)

	Unrestricted Total Discount Cost (% change)	Total Unrestricted Total Discount Cost (% change)	Discounted cost of reduction (% change)		
			None to 150g/kWh	None to 100g/kWh	150g/KWh to 100g/kWh
8% Growth	38%	53%	-11%	-5%	27%
7% Growth	19%	29%	-20%	-7%	20%
6% Growth	3%	5%	-7%	0%	7%
5% Growth	-10%	-9%	12%	-1%	-15%
4% Growth	-21%	-19%	7%	-14%	-31%
3% Growth	-30%	-25%	-24%	-37%	-42%
2% Growth	-38%	-36%	-42%	-45%	-48%
1% Growth	-45%	-44%	-65%	-54%	-51%

4.5.2 Technology sensitivities

One way in which the results are dependent on assumptions is a focus on the development of indigenous supplies of energy as is currently the trend in Ghana. However, an interesting question is how this assumption affects results and what role the introduction of new non-indigenous energy sources can play in terms of decarbonisation costs. Table 12 examines the effect of the addition of nuclear plant as an option to the main scenarios. I believe that there a number of barriers to the introduction of nuclear technology to Ghana that are not included here, however, these results can shed some light on the benefits of introducing this technology and whether these are sufficient to justify the costs of overcoming any barriers.

In the unrestricted scenarios the introduction of a new option should only decrease costs, and discounted costs are slightly lower as presented in Table 12[30]. In three of the scenarios

[30] The slight increase in the case of the 5% DR scenario reflects the adjustment made to

(Base, High RC, 5% DR) nuclear technology is built, however, the change in total cost is almost insignificant indicating the value nuclear can play in reducing costs is marginal at best. It is worth noting in the base case the build of nuclear is not displacing the build of CCGT units while natural gas is available, only in the last year of the study is nuclear build replacing the expensive HFO units that are required when natural gas supply reaches its limit. However, the availability of nuclear technology in these scenarios does have the side effect of reducing the total quantity of CO_2 emissions.

Table 12 Sensitivity of Nuclear technology addition to reference scenarios by 2040 *(% comparative difference)*

	Unrestricted Total Discount Cost (% change)*	**Total Unrestricted Total Discount Cost (% change)**	**Discounted cost of reduction (% change)***		
			None to 150g/kWh	None to 100g/kWh	150g/KWh to 100g/kWh
Base	0%	-5%	-17%	-38%	-54%
Low RC	0%	0%	0%	-20%	-35%
High RC	0%	-17%	-3%	-48%	-78%
LNG Gas	0%	0%	-52%	-66%	-73%
5% DR	1%	-12%	5%	-34%	-57%
15% DR	0%	0%	-38%	-49%	-54%

* Costs for this calculation are all discounted at the same 10% rate to improve comparability between scenarios where discount rates are adjusted.

The introduction of nuclear does cut the cost of meeting emission goals, in some cases

---

ensure all scenarios are compared on the basis of the same discount rate. In the 5% DR scenario investment in nuclear energy was made on the basis of a 5% discount rate but costs on this table are discounted at a 10% discount rate under which the nuclear added would not have been economic.

significantly. The only exception is the case of meeting a moderate emission target with low renewable costs where nuclear is still not economic, or in the case of the 5% DR where emissions are reduced significantly where there is not restriction, increasing the average cost of meeting the moderate target.

Table 13 Sensitivity of Coal technology addition to reference scenarios by 2040 *(% comparative difference)*

	Unrestricted Total Discount Cost (% change)	Total Unrestricted Total Discount Cost (% change)	Discounted cost of reduction (% change)		
			None to 150g/kWh	None to 100g/kWh	150g/KWh to 100g/kWh
Base	-0.4%	16%	-34%	-28%	0%
Low RC	-0.1%	7%	-25%	-17%	0%
High RC	-1.1%	18%	-26%	-26%	0%
LNG Gas	-0.1%	18%	-42%	-32%	0%
5% DR	-0.6%	21%	-46%	-36%	0%
15% DR	0.0%	0%	0%	0%	0%

* Costs for this calculation are all discounted at the same 10% rate to improve comparability between scenarios where discount rates are adjusted.

Looking at the coal sensitivity, the results for the discounted costs for no emission restriction scenarios are slightly lower across all scenarios except the 15% DR scenario where the high capital costs of coal units make them not economic. While total cost savings are minimal, the effect on emissions is large. This results in significant reductions in the cost of carbon abatement, and we can conclude that any country that is considering the build of new coal units offers an extremely cost-effective opportunity for reducing global CO_2 emissions. Given the cost effectiveness, I believe this option (new coal fire power plant) may be more tempting for developing countries which may not face as significant political pressure against this option. The one situation where this may not be the case is where developing countries face high costs of borrowing, while these costs also act as a barrier to investment in RES, here the results show that improving access to financial markets for a country with a high cost of borrowing may significantly increase its carbon emissions through investment in coal fired power stations.

4.6 Conclusions

Sustainable electricity supply is key in supporting the socio-economic development of a country, however, reliable access to energy comes with a cost. This study evaluates how electricity supply will evolve in Ghana in the future to meet the future demand under and the cost of contributing to world-wide decarbonisation efforts for the period 2020-2040. Developing countries face both unique challenges in decarbonising their energy systems but at the same time present important opportunities that have been quantified here. In addition, the inherent uncertainty in the energy system was explored across five key dimensions including renewable energy costs, financing costs, natural gas costs and availability, energy demand, and access to non-indigenous resources. The model developed provides insight over current studies by capturing the dynamic short-term operational challenges of variable renewable electricity supply and demand in combination with long-term investment decisions and the unique fuel limits and fuel switching behaviour common in Ghana. The model was validated by comparing its output to the actual annual generation data to assess its accuracy. Summary results indicate that the total discounted cost in expanding generation to meet the demand for all scenarios range USD 9-21 billion while the expected emissions range 99-189 $mtCO_2$.

Over time the penetration and share of renewable energy become more pronounced in all the scenarios, suggesting the cost-effectiveness of RE in the future, as construction costs fall or fuel cost rise. This reinforces the suggestion that RE deployment can be economical relative to the conventional fossil-based technologies in the long-term. However, I still find significant opportunities for additional decarbonisation in the model results and barriers to this occurring.

An important difference between developed and developing countries can be the rate at which either the private or public sector can raise finance. I find here that the higher discount rate associated with developing countries leads to significantly less investment in renewable energy absent some incentive. The high discount rate faced by a developing country has a mixed effect on build decisions. Firstly, it increases the benefit of delaying build decisions meaning that decisions are made later when renewable costs have decreased relative to conventional technologies. In this way, it can avoid “locking in” the build of conventional generation, however, the additional reliance on existing (inefficient and high emission intensity) units mean that CO_2 emissions can be expected to be very high where discount rates are high. In addition, these high rates also reduce the ability to build renewable technologies where a greater portion of the costs is incurred upfront. However, this means the relative cost of reducing carbon emissions in this case is lower in developing countries with a decarbonisation cost less than half that of a country with a low discount rate (11 USD/tonne vs 25 USD/tonne). This seemingly unintuitive result, that higher financing rates, that make it more costly to invest in renewable energy, actually reduce the cost of decarbonisation implies that developing countries provide a unique opportunity to reduce emissions at a lower cost. While the ability to finance high capital cost RES is often cited as a challenge faced by developing countries, I find that addressing this issue directly can have negative consequences. I find reducing financing rates, by for instance improving developing countries access to financial markets, can lead to additional investment in conventional technology in the short run (particularly where the country is considering coal as an option) and as such reduce the benefit in terms of decarbonisation. Interventions that would either subsidise renewable energy directly for developing countries or targeted access to finance for renewable energy projects only would in this case be a preferred policy.

Another key difference between the developing world and developed countries is the pace of growth in energy demand. I find this plays an important role in decarbonisation costs as the cost of displacing the build of conventional capacity is significantly less the cost of adding renewable capacity to a system that wouldn't otherwise need more capacity. In this way again the developing countries offer a unique opportunity to reduce carbon emissions at a lower cost than would be expected to be found in a developed country, although, again this opportunity exists as developing countries cannot be expected to bear the costs of CO_2 reduction on their own.

Finally, I find in general that a small reduction in emissions is more cost effective relatively than a large cut in emissions in a given system. I find a number of drivers for this result including that the larger the share of variable renewable sources of energy the greater the required level of energy storage. On a worldwide scale this means that it is likely more cost effective that all systems seek to reduce CO_2 emissions as opposed to this effort being concentrated in developing countries as is currently the case. Again, this result means that relatively small incentives could achieve much greater emission reduction in developing countries, which have less ability to invest in even small amounts of RES, than in developed countries.

Finally, I find that the technology mix considered also matters. I find that a country considering the addition of a new coal fired power plant (as is more likely for developing countries without political opposition) offers the lowest cost opportunity to reduce emissions over all scenarios with costs up to 66% lower. I additionally find that nuclear power can play a role in reducing the costs of reducing emissions by up to 46%.

This study is subject to some limitations mainly because of data difficulties. In situations where data was unavailable reasonable data proxies and assumptions were made. The assumptions and data used in relation to the input parameter have a significant impact on the quantitative results. As has been articulated previously, the pattern of the daily load curve till 2040 is based on the 2013 actual load curve data which essentially have one peak demand in the evening from 6-10 pm. Howbeit, as the economy progresses and behavioural patterns change over time, the shape of the load curve is expected to change accordingly to probably reflect what pertains in the developed world and in advanced countries, where the peak is found in the morning and/or afternoon and the evening. Unfortunately, such data is currently unavailable to effectively incorporate such variabilities in the model. Further studies could be done to assess the impact of such load curve variation. Furthermore, adequate studies specifically on the impact of DSM schemes on future generation and grid network expansion will be interesting. On the whole, the findings of this study provide insights which can inform policy decisions in electricity planning in Ghana and again provides a useful learning curve for other developing countries within the sub-Saharan region who are pursuing renewable energy integration.

Chapter 5

Long-term uncertainties in generation expansion planning: Implications for electricity market modelling and policy

In this chapter, I demonstrate the importance of the inclusion and careful treatment of uncertainty across all relevant inputs.

Investment decisions in the energy sector are capital intensive and require returns over long periods of time. Investor decision making is a complex process due to the large amount of inherent uncertainty across a range of important economic and physical factors over this extended planning horizon.

In the previous chapter I found that results from the GEP model for quantification of decarbonization costs are highly sensitive to uncertain input parameters. This is no surprise given the range and size of uncertainty found in the energy market as discussed in the introduction. However, as discussed in section 2.2 the scenario-based approach to GEP modelling of uncertainty, as used in this previous chapter, does not capture the effect that uncertainty has on the decisions made by market participants and so does not capture all the effects of uncertainty.

When modelling a competitive electricity market or the actions of a central planner there is an important distinction between the different approaches to incorporating uncertainty into a model. A stochastic optimisation model can be used to represent the fact that investors do not know the future of key inputs to their decision making, and this uncertainty will affect their decisions [21]. An approach such as Monte-Carlo simulation (or scenario analysis as seen in the previous chapter) instead represents uncertainty as a number of different scenarios, in each scenario investors in effect have perfect information of the future, and so uncertainty does not affect their decision making. However, in this case a range of different potential futures are assessed and from this range an estimate of the final distribution of outcomes is produced.

In this chapter I investigate the importance of incorporating into the GEP model the reality that decision makers must consider a wide range of economic and physical sources of uncertainty in their investment decision making. In the context of a national generation expansion planning problem we:

- Provide a simple method for generating uncertainty scenarios for GEP modelling when the modeller has only a single central forecast (as is typically provided by government or forecasting agencies). This method allows the modeller to take advantage of sophisticated expert forecasts for key variables, while also being able to generate any number of desired scenarios and determine a probability value for these scenarios as is required for stochastic optimisation.

- Demonstrate and quantify the importance of including a wide range of sources of uncertainty in combination. In particular, I find that the often-ignored sources of economic uncertainty can be more important than the traditional uncertainty in the physical quantity of energy required in future years.
- Find that it is important to include multiple sources of uncertainty in combination. I find that assessing each source individually can both underestimate the importance of uncertainty in that source, and can lead to investment decisions that actually increase costs when compared to simply ignoring uncertainty with a deterministic solution.
- Demonstrate the benefits of stochastic optimisation for dealing with uncertainty in either normative or descriptive modelling. I find that the stochastic optimal solution differs significantly from the average of scenarios modelled individually. I find that in a number of cases even a simple deterministic model provides a solution closer to the stochastic optimal than the average of individual scenarios.

5.1 Generation expansion planning model

The model outlined in Chapter 2 is applied with the following modifications:

- The model is simulated with the deterministic, individual scenario, and stochastic optimisation objective function. The results of these approaches are compared, to demonstrate differences in the approach to incorporating uncertainty into the model.
- The clustered unit commitment formulation is used for existing and new build

technologies and integers are relaxed on unit commitment decisions, meaning the linear relaxation formulation is used.

- The UK does not feature significant large-scale hydro, as such existing run of river hydro is subtracted from demand and the large-scale hydro constraints are not required.
- For a number of inputs, the forecast future trend differs significantly from that seen in the historical data (as discussed in 2.4). As such, the scenario generation and reduction approach described in 2.4 is applied to generate scenarios and weighting around central forecasts. Exogenously sourced forecasts are also used to test the sensitivity of results to this approach.
- In this case I use the partitioning around the medoid (PAM) [72] clustering algorithm as the scenario reduction method.

5.2 Case study

5.2.1 Current power system and expansion technologies

I take a stylized representation of the United Kingdom (UK) power system as a case study. The UK has been pro-active in the decarbonisation of the system to date, which means it already features a substantial level of renewables in the system. Additionally, it features both an ageing and retiring fleet of existing units along with a forecast for demand growth over the medium term, meaning we expect that new capacity will need to be developed, which we can model with the GEP model presented above. For simplicity and generality, I do not include a number of policies specific to the UK: the UK capacity mechanism, the ban on new onshore wind, the UK carbon price floor, or future new build of renewables based on current government policies, instead allowing the model to select the optimal combination of new build technologies.

The technologies considered as expansion candidates are included in Appendix B.

Table 14. For brevity, additional assumptions relating to model parameters and the case study are provided in Appendix B.

Table 14 Key characteristics of generation technology candidates for expansion

	Size (MW)	**Efficiency (%)**	**Start cost ($/MW)**	**Maximum annual capacity factor**	**VOM ($/MWh)**	**Build Costs ($/kW)**			**Fixed costs ($/kW-yr.)**		
						2025	2030	2035	2025	2030	2035
CCGT	500	53%	44	85%	2.72	1048			10.4		
CCGT CCS**	500	45%	56	85%		2079	1991	1918	33.1		
OCGT	500	34%	61	93%	7.03	896			12.0		
Li-Ion Storage (8 hours storage)	500	90% round trip	0	100%	0	2564*	2291*	2048*	6.1*	5.5*	4.9*
Solar PV	500		0	13%***	0	911*	858*	811*	7.2*	6.8*	6.4*
Onshore Wind	500		0	36%***	0	1599*	1602*	1601*	50.6*	50.7*	50.7*

* Value dependent on scenario. Scenario weighted average of generated values reported here.

** Carbon Capture and Storage (CCS) reduces CO_2 emissions by 90%.

*** Annual production reported here, hourly production based upon historical values.

5.2.2 Sources of uncertainty

I include uncertainty across four key inputs that cover economic and technical sources of long-term uncertainty:

- Total annual energy demand – Forecast (2020-2040) and historical values (2005-2017) for the total UK energy demand are taken from [68].
- Natural gas prices – Forecast (2020-2040) values for the total UK natural gas prices are taken from [68] with historical values from 1993 to 2018 sourced from [120] .
- ETS CO_2 Price – Forecast (2020-2040) values for ETS CO_2 prices are taken from [68], due to limited quantity of available historical data for ETS I use natural gas price volatility as a proxy for ETS CO_2 Price volatility.

- Technology costs – construction and fixed costs for new onshore wind, solar PV, and Li-ion batteries. To reduce the number of scenarios required[31] it is assumed that drivers of technology costs are perfectly correlated (i.e. High solar construction costs occur with high onshore wind and Li-ion storage construction costs). This would likely be the case where cost reductions are driven by research and development funding for renewable technologies or demand for renewable technologies. Forecast values for all technologies are taken from [121]. Historical price volatility based on solar PV construction cost price volatility 2001-2017 are sourced from [122].

For the majority on the analysis I include 3 scenarios for each source of uncertainty for a total of 81 scenarios[32]. I test the effect of increasing or decreasing this number in section 5.3.5.

5.3 Results

5.3.1 Scenario generation and reduction

Here I present the results of the scenario generation and reduction process for the four sources of uncertainty: energy demand; ETS Price; Natural Gas Price; and Constructions

[31] Modelling technology costs as independent would require 729 scenarios as opposed to 81 in the case of dependent scenarios.

[32] The value of 3 scenarios is selected for reasons of practical limitations, moving to 4 scenarios per source and 256 in total proved too large of a stochastic optimisation problem on available hardware without decomposition.

costs for Solar PV, Onshore Wind, Li-ion storage[33]. I compare these scenarios with the exogenous central forecast scenario and additional high and low forecast scenarios produced by the transmission system operator and commonly used in government policy analysis in [68], and NREL [121] for construction costs. In Table 15 I report the fitted GBM parameters.

Table 15 Estimated GBM Parameters for historical data and exogenous forecasts, annual time step.

	Historical		Forecast
	μ	σ	μ
Energy Demand	-1.4%	1.1%	0.8%
Natural Gas Price	4.0%	15.2%	1.7%
ETS CO_2 Price		*	7.0%
Onshore wind		**	-0.2%
Solar PV	-7.5%	7.0%	-1.4%
Li-Ion Storage		**	-2.5%

* Historical Natural gas price volatility used as proxy for ETS volatility due to short historical time series.
** Solar PV historical construction cost volatility

Examining the generated and exogenous scenarios presented in Figure 27 to Figure 30, I see that the scenarios generated cover reasonably well the range of exogenous scenarios, despite the completely different generation techniques. It should be noted that only the central scenario is an input to the scenario generation process and the range of generated scenarios is based entirely upon historical volatility, where the exogenous scenarios are of the form high, low and central. It is unclear how the range for these scenarios is generated.

The major exception is that given the historical volatility in energy demand, combined

[33] Here for brevity we report only solar constructions costs, however, Li-ion and onshore wind show a similar trend as the same value for price volatility is used.

with the forecast growth, the generated forecasts for demand growth are significantly higher in the short term and more uncertain than those provided by National Grid in [68]. With the ETS price and natural gas prices, the scenario reduction process provides one relatively high scenario, however, this is assigned a small probability. The importance of extreme points in the clustering of inputs has been demonstrated for the GEP model in other inputs [81] and this result is likely an advantage over relying on scenarios built from bottom up assumptions as in [68]. However, the appropriateness of these scenarios relies on the assumption of GBM as the underlying statistical process for the commodities. Finally, the construction cost scenarios show a similar fit to the scenario range developed by [121] with a slight bias towards higher costs. To ensure the GEP model results are not specific to the scenario generation process, I additionally report results using the exogenous scenarios with an equal probability weight assigned to each.

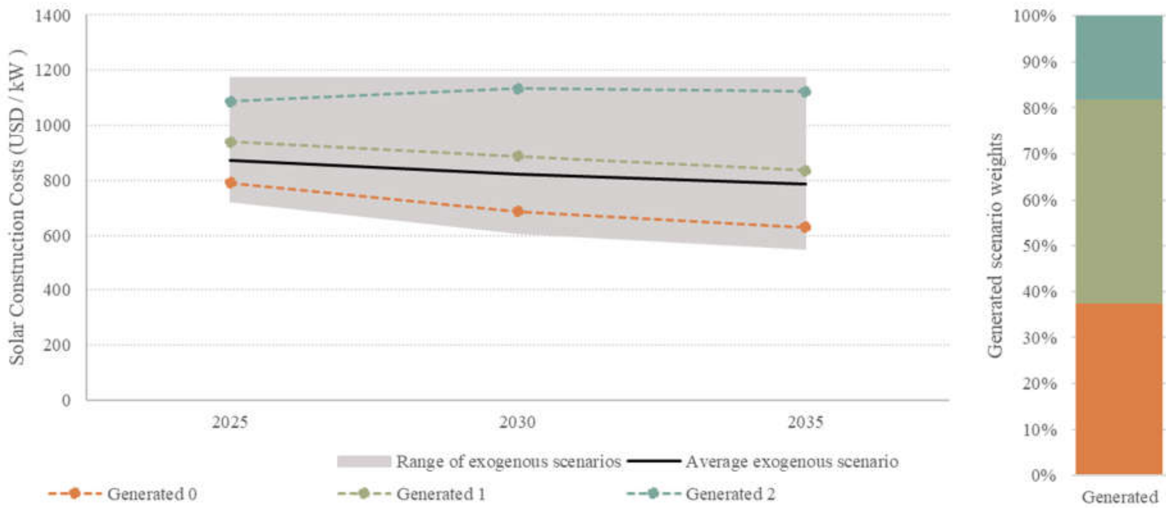


Figure 27 Generated scenarios and scenario weights compared to exogenous scenario range for solar PV construction costs.

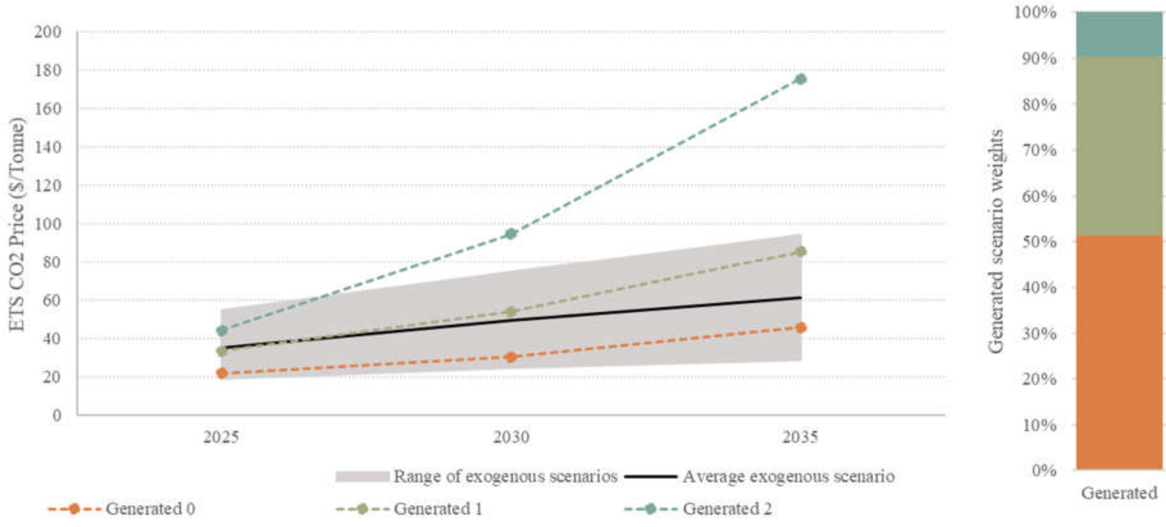


Figure 28 Generated scenarios and scenario weights compared to exogenous scenario range for ETS CO_2 price.

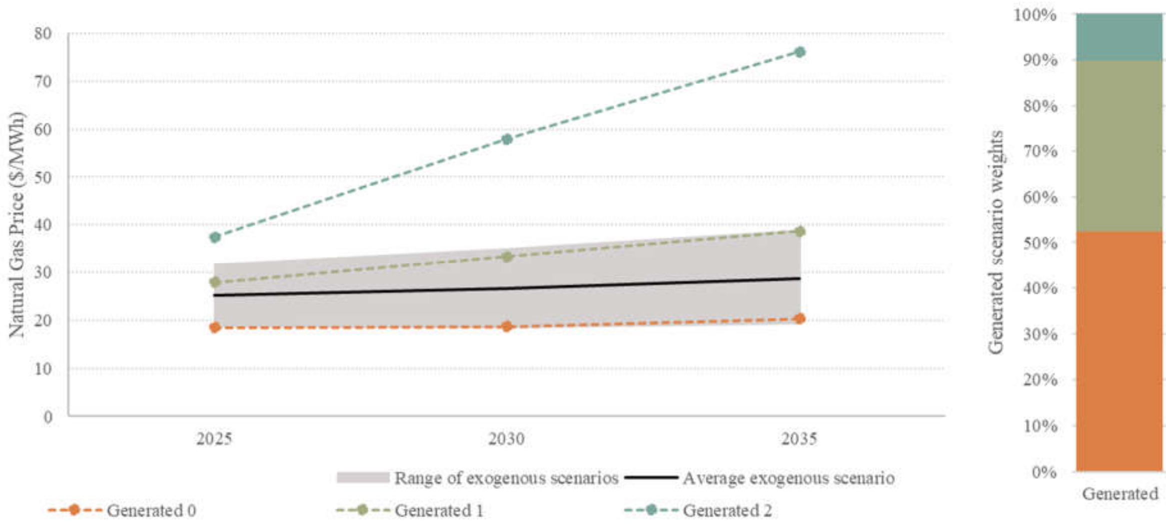


Figure 29 Generated scenarios and scenario weights compared to exogenous scenario range for natural gas price.

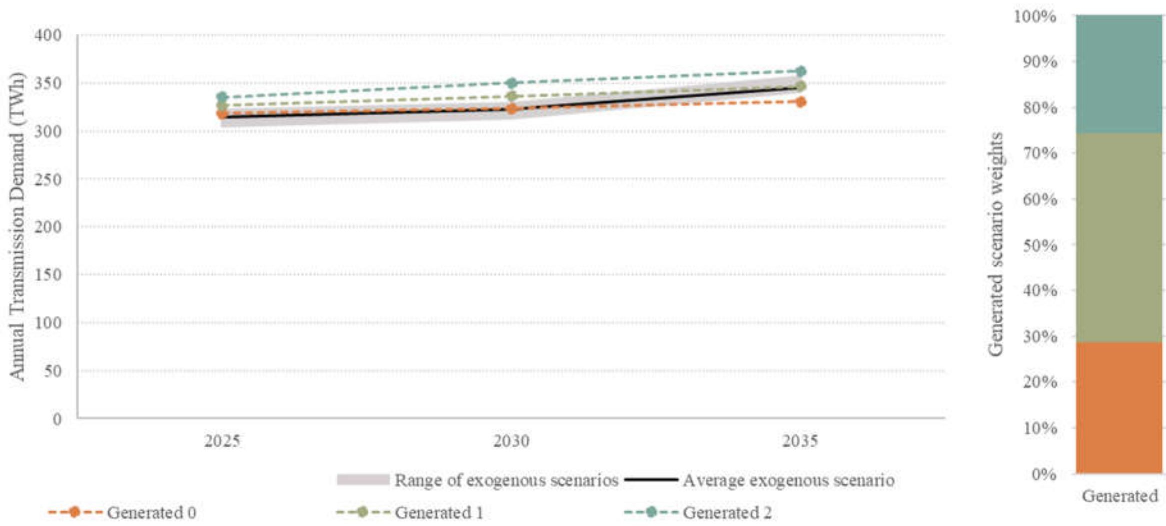


Figure 30 Generated scenarios and scenario weights compared to exogenous scenario range for annual UK energy demand.

5.3.2 GEP modelling results under individual and combined uncertainty

In this section I present the results for the GEP model under different representations of uncertainty, namely including uncertainty individually on each input, or all sources together. Firstly, in Figure 31 I present the build decisions made by the deterministic model which forms a benchmark to which I will compare other results. In this model all uncertain variables take the probability weighted average value from the forecasts derived. In this way the uncertain inputs have the same expected value across all approaches to incorporating uncertainty.

The deterministic model finds that the optimal mix of power plants include most technologies. The exception is CCGT CCS, as the model does not find the additional cost of carbon sequestration to be justified[34]. The model also finds significant build is

[34] The economic build is higher than that found in [68] reflecting the higher average energy demand produced by the scenario generation and reduction process.

required in the 2020-2025 period to replace retiring units.

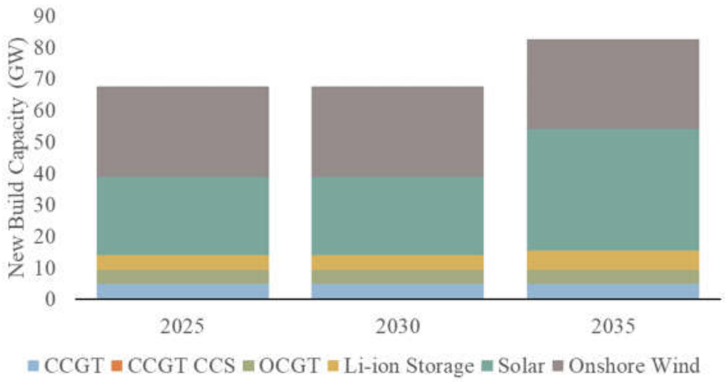


Figure 31 GEP model build decisions for deterministic model by technology

To evaluate the effect of adding uncertainty, I make use of the common metrics: Value of Stochastic Solution (VSS) [123], Expected Value of Perfect Information (EVPI) [123], and introduce the metric Value of Scenario Result Average Solution (VSRAS).

$$VSS = SO - EV \tag{44}$$

$$VSRAS = SRA - EV \tag{45}$$

$$EVPI = \sum_{s \in S} (w_s \cdot SS_s) - SO \tag{46}$$

The VSS is defined as the difference between the objective function value of the stochastic optimisation problem with the first stage variables (2025 build decisions) fixed at the values found in the deterministic model (called the expected value (*EV)* solution) and the objective function value of the stochastic optimisation model SO[35]. One way to interpret the VSS is the expected cost savings a central planner would gain by following

[35] We have defined the VSS such that a positive value represents an improvement in the solution value (reduction in the value found in the cost minimisation).

the build decisions found in a stochastic optimisation model as opposed to the simple deterministic model. In this way it offers a measure of the importance of treating uncertainty correctly in the model.

Here I introduce the VSRAS, where I compare the objective value of the scenario result average (SRA) to the EV. The SRA is calculated by optimising each $s \in S$ scenarios individually, which produces individual results for the first stage build decisions. I then take the scenario weighted average build decision for each technology giving us one single build decision per technology. While the VSS metric has the property VSS > 0 [123] this is not necessarily the case for the VSRAS, where the value of the average solution SRA performs worse than the deterministic solution, I find VSRAS < 0.

As with the calculation of VSS, I evaluate the VSRAS in the full SO model, in this case by using the decision found by averaging individual scenario solutions. This average build decision is used to fix the first stage build decisions and the resulting optimisation model is solved with these constraints to provide an objective function value SRA. This metric is intended to provide a rough estimate of the value of modelling uncertainty outside the model (as with scenario analysis), where the modeller generates a range of solutions and must decide how to combine these to make a decision (in this case I use the scenario weighted average). Comparing the VSS and VSRAS will allow us to provide an estimate of how much of the benefit of looking at uncertainty comes simply from the greater range of values from having multiple scenarios (VSRAS) and how much of the value comes from optimising against this set of scenarios (VSS). If I find the VSRAS to be close in value to the VSS, I can conclude that making decisions based on individual scenarios (or the average of a Monte-Carlo simulation) is a sufficient means of dealing with uncertainty in the modelling, without requiring the significant increase in model complexity created by solving all scenarios in a single optimisation problem, as with the stochastic optimisation approach[36].

Finally, I present the EVPI, which compares the probability weighted sum of the objective function value of each scenario SS_s with the stochastic optimisation result SO. This provides an alternative estimate of the uncertainty in the model, in this case: how much does the fact that the central planner (or market) has to make a single set of first stage decisions, without knowledge of the future cost. Or, alternatively, how much lower would costs be on average if the decision makers had perfect foresight of which scenario was going to occur.

Table 16 reports the values of the metrics discussed above in the GEP for the UK with the generated scenarios. Firstly, when evaluated against a single source of uncertainty, the VSS is significantly higher than the VSRAS across all sources of uncertainty and the VSRAS is often negative. This indicates that simply taking the average of a number of independent scenarios performs poorly for making investment decisions (and likely poorly reflects market decisions). Examining Figure 32, we can see the differences in capacity build between the SO and SRA solutions. Firstly, looking at the case of demand uncertainty, we can see the average build of the scenarios is close to the deterministic case, with little change. However, the stochastic optimisation increases slightly the build of wind and solar technologies, while the average result of the individual scenarios is to both decrease solar and CCGT capacity. These results provide evidence of the sub-optimal decision making when each scenario is treated individually in the capacity expansion problem. In fact, the build plan found by the SRA model performs worse over uncertainty, than the deterministic model that completely ignores uncertainty in four of

[36] On a 24 Thread 4Ghz CPU the 81 individual scenarios take a total time of 35 minutes to solve, the stochastic optimisation model takes over 8 hours.

the five situations, as indicated by a negative VSRAS in Table 16. For a central planner this would mean implementing the average build decisions from the SRA approach, would result in a more costly system than if they were to ignore uncertainty in their modelling when deciding on which plant to build. For a policy maker looking to model a competitive market, this result implies that the average result of the individual scenarios does not reflect what build we would expect in the actual market when uncertainty is taken into account by participants.

Table 16 Values of stochastic solution (VSS), value of scenario results average solution (VSRAS), and expected value of perfect information (EVPI) with generated scenarios

USD Millions		Inputs for which uncertainty is applied				
		Ren. Build Cost	Demand	ETS CO_2 Price	Natural Gas Price	All combined
Assessed against individual source of uncertainty	VSS	23	15	36	58	
	VSRAS	-82	3	-641	-286	
Assessed against all sources of uncertainty	VSS	109	-93	-145	-8	275
	VSRAS	67	-27	-619	-216	-289
EVPI		215	466	503	327	1380

When evaluated against a single source of uncertainty, the VSS is always greater than zero [123]. However, when evaluating the build decisions selected by adding uncertainty to a single input in a model that contains all sources of uncertainty, there are a number of large negative VSS values up to -\$145 million in the case of ETS CO_2 Price. Figure 32 shows the cause of this result: when considering only ETS CO_2 price uncertainty, the SO model selects to build less CCGT and more renewable, OCGT and storage technologies, in total increasing the build of capacity. In contrast, in the case of the SO model considering all sources of uncertainty combined it selects build plans with significantly less renewable technologies and less capacity overall. Hence, the SO model finds that these are not economically justified when all sources of uncertainty are considered in the optimization model. In other words, considering ETS CO_2 price uncertainty alone reduces the quality of the solution both in terms of system cost for the planner, and the ability to forecast which technology the market would be expected to produce.

Finally, from Table 16, it can be seen that the VSS for all sources of uncertainty combined is twice as high as the sum of the VSS of each individual source of uncertainty. Moreover, the EVPI of the combined sources of uncertainty is 75% higher than the sum of the individual calculated EVPI for each source. These values indicate that there is an interaction between the sources of uncertainty that cannot be captured without combining them in the same model. Without combining uncertainty, the modeller is likely to underestimate the importance of addressing this crucial characteristic. It is also worth noting that the results show that the economic sources of uncertainty, which are less featured in the literature, have a greater influence on both the VSS and the expected technology mix.

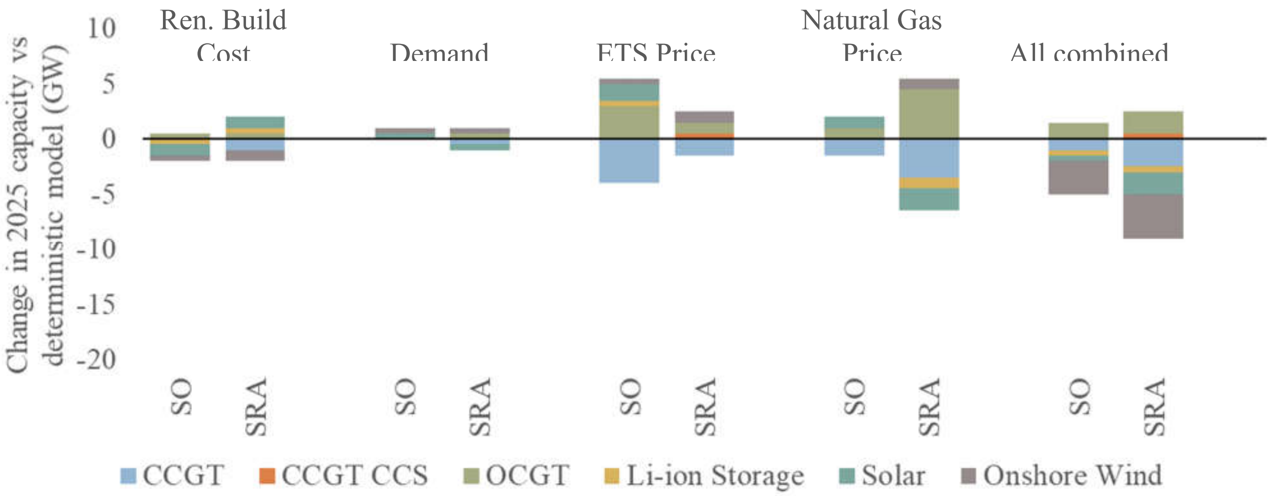


Figure 32 Change in 2025 new build capacity for Stochastic Optimisation (SO) and average of individual scenario results (SRA)

It is important to consider several other results from Figure 32. Firstly, the average of the build decisions of individual scenarios (SRA) differs significantly from the stochastic optimisation solution in all cases. Additionally, the set of build decisions found to be optimal when combining all sources of uncertainty is very different to what appears optimal for any individual source of uncertainty. For example, the SO result when introducing uncertainty on any given input increases the amount of renewable capacity built in three of the four cases, however, when combining all sources of uncertainty in the model the SO finds a significant reduction in renewable capacities optimal. A modeller that evaluated sources of uncertainty individually would likely conclude that building more renewable technology is optimal with regards to uncertainty, however, the opposite is the case. This result again reinforces the conclusion that evaluating sources of uncertainty in isolation, or in simplified treatments such as scenario analysis, leads to significantly sub-optimal results and is inappropriate for market descriptive analysis.

It is worth emphasising the implications for a policy maker due to the popularity of GEP for evaluating policy. If we consider the SO result of the combined uncertainty to be representative of the actual market outcome as previously discussed, a policy maker looking at results from either a deterministic model, or a model containing uncertainty on any one dimension, would significantly overestimate the amount of renewable capacity the market will provide. A policy maker focused on decarbonisation may conclude that renewable support policies are not required due to the results of their modelling. Where in fact a model combining all these sources of uncertainty, and thereby better representing the full set of uncertainties faced by participants in the market, favours reducing the overall capacity committed to build, especially of expensive renewable technologies and instead favours the cheapest (although environmentally most damaging) OCGT technology. In such a case, policy makers could be surprised by the result of the market selecting environmentally damaging technologies against the range of modelling results used to inform their decision.

Policy makers are also typically interested in additional results from their modelling including electricity prices and carbon emission. I present these results, considering all sources of uncertainty, in Figure 33. Firstly, it can be seen that the expected annual 2025 CO_2 emissions using stochastic optimisation are higher than the deterministic case and lower than the average over individual scenarios for the year 2025. Of more interest is the spread of individual scenario outcomes as policy makers are often interested in worst (or best-case scenarios). Modelling scenarios individually greatly overestimates this spread, for example in a scenario with high demand, low gas prices, low ETS prices, and high renewable construction costs, the model considering only this scenario, build no new renewable capacity and instead adds conventional capacity. However, in the real market, decision makers must make a set of build decisions under uncertainty (without knowing that this exact scenario will occur) and build a more balanced mix of technology, as captured by the stochastic optimisation model. In this case only the outcome level of demand affects the final CO_2 emissions. The distribution of electricity prices demonstrates the opposite trend, as the individual scenarios optimises to each set of circumstances, each scenario is modelled with the cost optimal build plan, and there is a lower spread of possible outcomes. However, given build decisions must be made in advance they will not be cost optimal against the majority of scenarios individually (just optimal on average) and more extreme pricing outcomes can result.

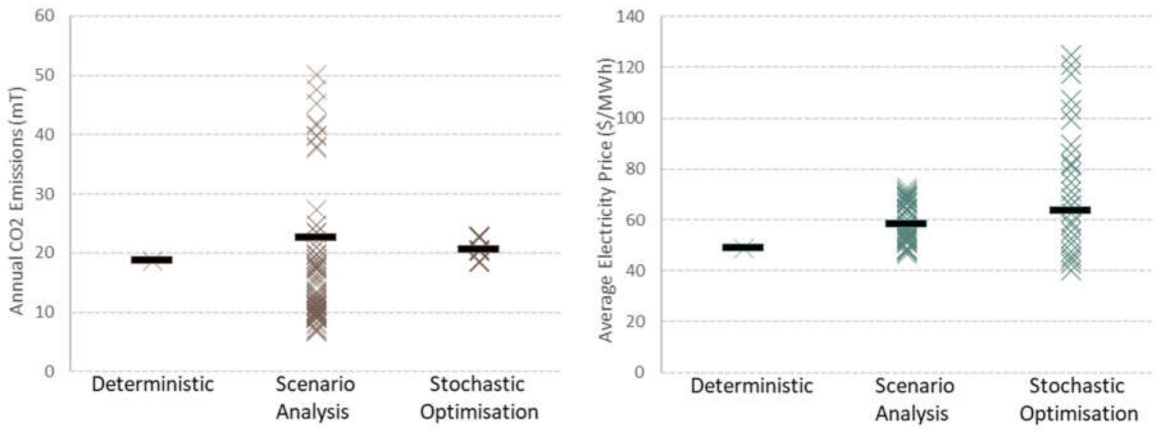


Figure 33 Scenario and expected (black bar) 2025 annual CO_2 emissions [left]. Scenario and expected (black bar) 2025 average electricity prices [right].

5.3.3 GEP modelling results under alternative scenarios

To test that the results found are not specific to the scenario generation and reduction process used in the analysis so far, I test an alternative quantification of uncertainty. Each of the data sources for uncertain inputs (see section 5.2.2) provides a range of forecast scenarios (typically high, medium, low) and here I input these scenarios explicitly into the model and provide each an equal weight[37].

In Table 17 and Figure 34, a similar pattern of results to those presented in section 3.2 can be seen. Here the results show that the VSS for all sources combined to be 67% higher than the sum of the individually assessed sources, again indicating interaction effects. In this case, the renewable build cost is the most important source of individual uncertainty. The expansion plan derived from this uncertainty is also relatively close to the optimal build decisions given all four sources of uncertainty. This is an important result if we consider that this representation of uncertainty is realistic, as renewable build costs are not a commonly considered a source of uncertainty.

[37] Scenarios found in government forecasts or the literature are often not provided with specific weights, making it difficult to use them for stochastic optimisation or market modelling. The equal weight assumption here would reflect a market where all investors use these forecasts for decision making and consider each to have an equal probability.

Table 17 Values of stochastic solution (VOSS), value of scenario results average solution (VSRAS), and value of perfect information (VOPI) with generated scenarios

USD Millions		Inputs for which uncertainty is applied				
		Ren. Build Cost	Demand	ETS CO_2 Price	Natural Gas Price	All combined
Assessed against individual source of uncertainty	VSS	671	46	29	46	
	VSRAS	223	-101	-40	-121	
Assessed against all sources of uncertainty	VSS	1027	40	-184	382	1320
	VSRAS	310	404	-37	61	641
EVPI		215	3090	170	142	568

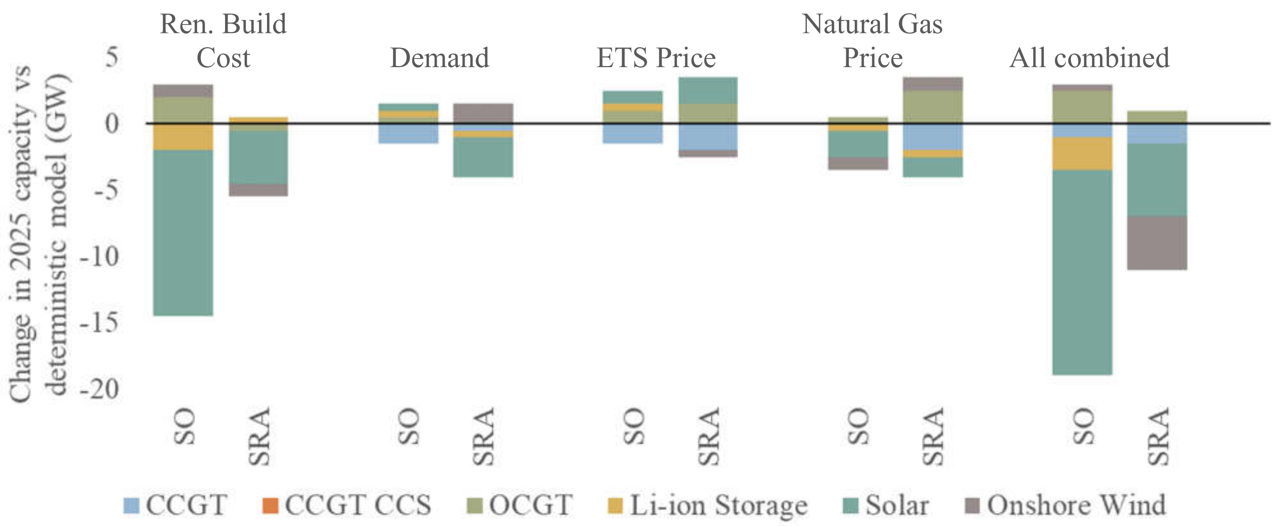


Figure 34 Change in 2025 new build capacity for Stochastic Optimisation (SO) and average of individual scenario results (SRA) with exogenous sources of uncertainty

5.3.4 Exploring uncertainty interactions in GEP Results

In this section, I run the model with different combinations of the generated sources of uncertainty (as in section 5.3.2). I initially found that the VSS for all sources combined was 100% higher than the sum of VSS for the individual sources. When different combinations of uncertainty are considered, the difference between the combined and the sum of individual VSS, ranges from -70% (in the case where the sum of the individual VSS of demand and ETS is higher than the combined value of these two sources) to 125% in the case of natural gas price, renewable build cost and ETS CO_2 price. Interestingly, the combination of natural gas price with renewable build cost performs relatively well, as do all cases where 3 sources of uncertainty are included as long as natural gas price is included.

Table 18 Value of stochastic solution and expected value of perfect information for different combinations of uncertainty

Uncertainty Sources			**VSS Evaluated against included uncertainty**	**VSS Evaluated against all uncertainty**	**EVPI**
Ren. Build Cost	Demand		43	144	925
Ren. Build Cost	ETS CO_2 Price		68	-21	1033
Demand	ETS CO_2 Price		15	21	809
Natural Gas Price	Ren. Build Cost		126	211	2910
Natural Gas Price	Demand		41	90	1864
Natural Gas Price	ETS CO_2 Price		117	102	2567
Ren. Build Cost	Demand	ETS CO_2 Price	47	162	990
Natural Gas Price	Ren. Build Cost	Demand	154	254	1979
Natural Gas Price	Ren. Build Cost	ETS CO_2 Price	264	257	3213
Natural Gas Price	Demand	ETS CO_2 Price	101	237	2741
All combined				275	3657

5.3.5 Comparing wider sources of uncertainty to higher detailed uncertainty

In the previous sections the results have shown that combining uncertainty across different sources is important in the GEP. However, a trade-off exists, and we must decide if we should put scenarios in all sources of uncertainty, or a greater number of scenarios for any given single source. Here I test this trade-off, by generating four scenarios for each source of uncertainty for a combined total of 256 scenarios. This stochastic optimistion problem over all 256 scenarios cannot be solved[38] to find the optimal decision, but I can assess how a set of build decisions performs over this larger uncertainty set. Here I compare build decisions made incorporating all the scenario detail on a single variable (four scenarios), against including combined uncertainty across all inputs with either two or three scenarios on each variable. In addition, I perform an out of sample assesment where I assess the build decisions against 1000 scenarios generated in the scenario generation procedure (before the scenario reduction procedure).

Table 19 demonstrates that even just allowing two scenarios on all sources of uncertainty in combination outperforms capturing a single source of uncertainty in more detail (in this case all four scenarios represented in the model were used for assesment). The values for looking at each individual source of uncertainty broadly follow those seen previously in Table 16, with the exception that detailed uncertainty in natural gas prices performs significantly better.

Table 19 Value of stochastic solution evaluated over 256 potential scenarios (all combinations of 4 representative scenarios across 4 sources of uncertainty) and 1000 randomly selected out of sample scenarios.

	4 Scenarios				2 Scenarios	3 Scenarios
	Ren. Build Cost	**Demand**	**ETS CO2 Price**	**Natural Gas Price**	**All combined**	**All combined**
VSS evaluated across 256 representative scenarios	10	-30	-210	205	331	315
VSS evaluated across 1000 out of sample scenarios	-25	-124	-394	446	569	693

[38] On available or standard hardware in a reasonable timeframe.

5.4 Conclusion

Uncertainty is an increasingly important aspect of decision-making relating to the electricity systems of the future. Over the long-term time horizons required for investment decisions and government policy, history indicates that forecasts tend to be varied and inaccurate. However, I do not find it typically to be the case that all sources of uncertainty are included in the modelling used for decision making when reviewing the literature.

In this work I demonstrate the importance to the GEP model of including uncertainty across a wide range of key inputs. Further, I find the value of this uncertainty and the effect on model outputs can only accurately be assessed when the sources of uncertainty are considered together. In particular, I find an important role for economic sources of uncertainty such as natural gas prices, ETS CO_2 prices, and technology costs to be considered in addition to physical uncertainties such as demand.

Finally, I find that simplistic treatments of uncertainty that represent uncertainty outside of the optimization model, such as scenario analysis or Monte-Carlo simulation, are unlikely to perform well for descriptive modelling analysis of competitive electricity markets, as for policy analysis. I find that stochastic optimisation, which combines all uncertainty into a single model and represents the fact that investors must make decisions without knowledge of the future, generates results significantly different to modelling approaches that incorporate uncertainty as individual scenarios. This has important implications for policy makers as the GEP model is a common approach to modelling competitive electricity markets for policy analysis.

I demonstrated in a UK case study the importance of combining uncertainty across different inputs, finding that the value of the stochastic solution increases non-linearly when uncertainty inputs are combined, and is 67% to 109% higher than when uncertainty across individual inputs are superimposed. However, I also stress the importance of these results for policy makers. Building on the insight that the stochastic optimisation result reflects the expected outcome in a perfectly competitive market, I demonstrate that the policy maker would see a large mismatch between the technology mix that emerges in a competitive market and the results of a deterministic planning model or the average of individual scenarios. Of particular importance is the fact that in the UK case study the quantity of renewable technologies is significantly reduced when uncertainty is taken into account, and the cheap but environmentally most damaging OCGT technology is preferred. These results indicate that policy makers may underestimate the support required to incentivise a given level of renewable investment if they are not considering uncertainty in the model using stochastic optimisation. In addition, I find that individual scenarios overestimate the range of levels of CO_2 emissions, but underestimate the range of potential electricity price outcomes. In the following chapter I build on these results to compare the performance of different decarbonisation policies under different modelling treatments of long-term uncertainty.

Chapter 6

Renewable energy support policy evaluation: the role of long-term uncertainty in market modelling

In this chapter, I investigate the importance of the approach to representing long-term uncertainty in the modelling used to evaluate different renewable support policies.

6.1 Introduction

In this chapter, I focus on the problem faced by policy makers who are tasked with encouraging the decarbonization of the energy sector to mitigate the effect of CO_2 emissions on the environment. For the practical purpose of evaluating long-term policy in the electricity sector, Generation Expansion Planning (GEP) models are the approach favoured by policy makers [11]. However, I demonstrated in the previous chapter that both the sources of uncertainty, and approach to modelling uncertainty, have important implication for the GEP model.

As it turns out, uncertainty has played a key role in the debate about policy instruments for incentivising emissions abatement. In a seminal work [46], Weitzman showed that the two approaches of carbon tax and cap-and-trade are equivalent in terms of social welfare under a set of simple assumptions, however, they differ in the case where carbon abatement costs are unknown. In such a situation the preferred policy depends on the relative slope of the marginal benefit and cost of abatement curves at the equilibrium.

However, this result relies on a high-level theoretical analysis, for example, a generic choice is available between supplying to the market with carbon abatement or supplying to the market without carbon abatement. This approach also relies critically on the modelled marginal abatement curves [44]. However, investment decisions in the electricity market are significantly more complex due to the complexity of transmission constraints, the sizing constraints of new entrants, the difficulty of storing electricity inter-temporarily, the requirement to provide ancillary services, the variability and uncertainty of variable renewable generation sources, among other concerns [25].

For this reason GEP models are then the approach favoured by policy makers [11]. In fact a report by the European Commission finds that the majority of the pressing electricity market design questions for competitive markets can be represented in optimisation models such as the GEP [23] and has selected this approach for the model it has commissioned to perform policy analysis [48]. In fact, 90% of energy system models available commercially or academically aim to address the questions asked by policy makers [49].

Unfortunately, as discussed in Chapter 1 the representation of uncertainty in GEP models is often simplified, likely due to the curse of dimensionality. However, in the previous chapter I demonstrated the importance of the incorporation of this uncertainty for outcomes in a GEP model. The main contribution of this chapter is then that I demonstrate the importance of the inclusion and careful treatment of uncertainty across all relevant dimensions when assessing renewable energy policies. More specifically, in the context of a real-world generation expansion planning problem we:

- Compare the effectiveness of six different policy options; a cap on CO_2 emissions (as with a cap and trade scheme), a CO_2 price, a renewable capacity target, a green certificate scheme, a renewable generation subsidy, and a renewable capital grant under different modelling treatments of uncertainty.
- Demonstrate the importance of the inclusion of uncertainty into the GEP model for policy assessment, with many policies designed in a deterministic model not achieving target levels of carbon reduction or significantly underestimating costs.
- Demonstrate that common modelling approaches that attempt to capture uncertainty as multiple different individual scenarios (such as scenario analysis or Monte-Carlo simulation), perform poorly at representing the reaction of a competitive electricity market. These approaches can underestimate the expected carbon abatement response to decarbonisation policies and significantly underestimate costs.
- Demonstrate the importance of incorporating uncertainty for determining the technology expected to be built by a competitive market. With a scenario analysis approach both underestimating the quantity of conventional capacity that the market

would construct without renewable support policies and underestimating the benefits of solar capacity over wind capacity in an uncertain market environment.

6.2 Generation expansion planning model

The model outlined in Chapter 2 is applied with the following modifications:

- The model is simulated with the deterministic, individual scenario, and stochastic optimisation objective function. The results of these approaches are compared to demonstrate differences in the approach to incorporating uncertainty into the model.
- The clustered unit commitment formulation is used for existing and new build technologies and integers are relaxed on unit commitment decisions, meaning the linear relaxation formulation is used.
- The UK does not feature significant large-scale hydro, as such existing run of river hydro is subtracted from demand and the large-scale hydro constraints are not required.
- For a number of inputs the forecast future trend differs significantly from that seen in the historical data (as discussed in section 2.4). As such, the scenario generation and reduction approach described in 2.4 is applied to generate scenarios and weighting around central forecasts. Exogenously sourced forecasts are also used to test the sensitivity of results to this approach.
- I use the fast forward selection (FFS) [73] technique with the Euclidean norm set as the distance metric for the scenario reduction method, as this approach is often applied to power system problems [74].

6.3 Case study

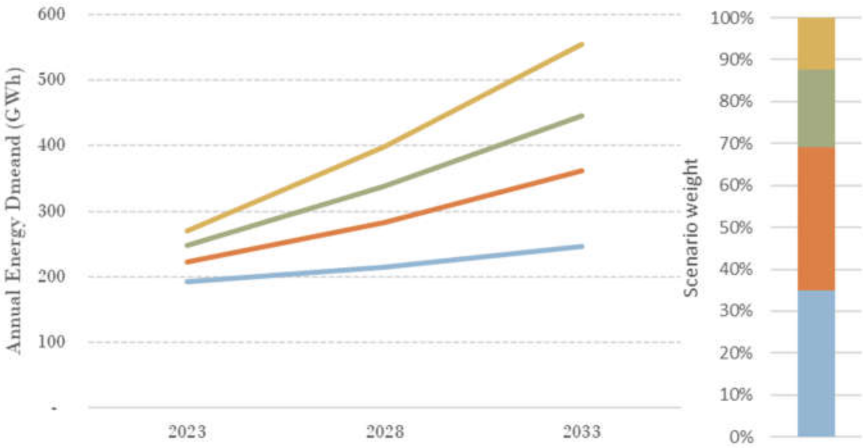
As previously discussed, including such a large number of scenarios in a single optimization creates major challenges. I therefore select a case study where I can minimize the problem size while still drawing representative results. For this exercise I model the island of Terceira in Portugal which has several characteristics that are useful here. Firstly, it is a small isolated system (effectively a microgrid), however, energy demand is growing and future investment in generation capacity is required. The island also features a relatively high level of variable renewable penetration which currently results in occasional curtailment actions making it representative of the high renewable generation systems of the future. The model selects between four representative technologies in 5MW increments: small scale diesel engines, Li-ion storage, PV solar, and small-scale onshore wind.

Details on the case study parameters are included in 0.

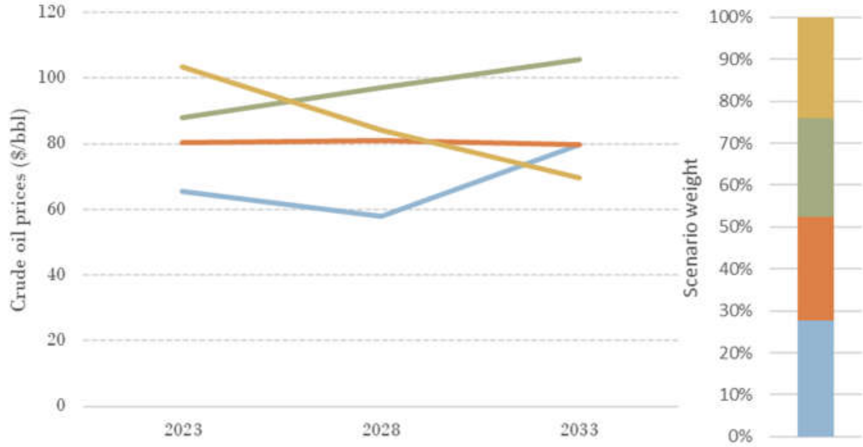
6.3.1 Sources of uncertainty

I apply uncertainty to three key model inputs: fuel prices, energy demand, and renewable technology capital and operating costs, as presented in Figure 35. For fuel prices and demand it is reasonably common to consider these sources of uncertainty aleatoric and consider that a stochastic process can be considered a reasonable representation (as characterized in [5]). For these sources of uncertainty I follow a typical scenario generation and reduction approach, as discussed in [40] and detailed in section 1.2.1.

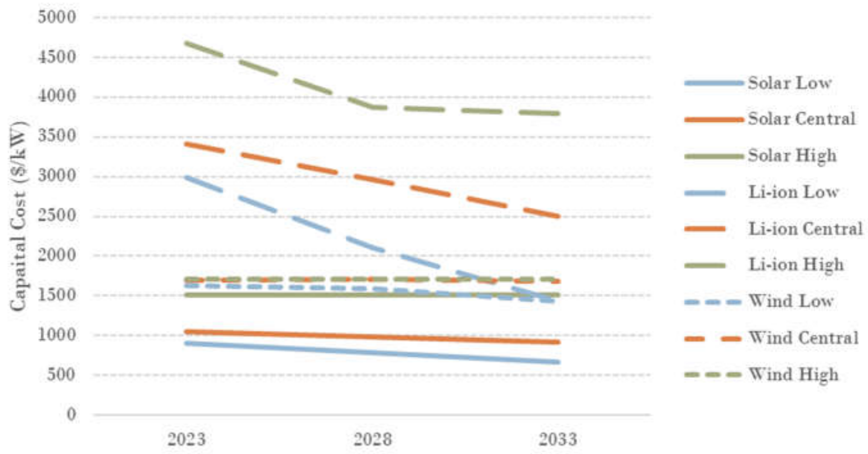
For technology costs I take a different approach and apply equal weights (equal probabilities) to the high, medium, and low forecasts developed by the National Renewable Energy Laboratory (NREL) [124]. In this way the evolution of costs is uncertain but exogenous to the model (as opposed to endogenously determined, as in [125]). Technology costs can be considered to entail both epistemic and aleatoric uncertainty [5] and the application of expert judgement to define and assign a weighting to such scenarios for modelling is one potential approach, as in [126]. An alternative proposed approach would be to fit a triangle distribution over the range of scenarios as in [127], however, this assumes the extreme scenarios represent the most



(a)



(b)



(c)

Figure 35 Generated representative scenarios and weight-ings for annual energy demand (a) and crude oil prices (b). Exogenous technology construction cost scenarios (low,central, high) each with equal weight (c).

extreme possible values the series could take, which I do not believe to be the case. In this chapter, I use the model to represent the decisions of actors in the market, and the treatment of technology costs should reflect how they incorporate this information into their decisions. While the application of equal weights is an assumption, I believe this is a reasonable treatment for the question of estimating investors' long-term views.

6.3.2 Decarbonization / Renewable Energy Support Policy Modelling

Here I include six different common decarbonization or renewable support policies [128] for assessment under different assumptions about uncertainty, as detailed in Table 20. The policies are chosen to be representative of the measure currently in use or under consideration for electricity markets and can broadly be classified under two different criteria: does the policy target prices or quantities? Does the policy target CO_2 emissions or renewable generation?

The difference between price-based and quantity-based policies has been discussed in the introduction, here I also look to compare direct emission targeting policies against renewable technology targeting policies with both price- and quantity-based measures. The final goal of all the policies is the reduction of CO_2 emission, so in this sense the renewable targeting policies are an indirect approach. Where these renewable targeting policies lead to more renewable generation, they will displace conventional generation and the associated CO_2 emissions. However, they do not incentivise other ways in which decisions can reduce CO_2 emissions, for example, building higher efficiency conventional units that produce less CO_2 compared to existing units [128].

The policies are also chosen so as to ensure that the solution to the centralised GEP model remains representative of a competitive electricity market. In the case for the cost-based measures, these policies mean adjusting model parameters and the stochastic optimisation model remains reflective of a competitive market. This result has also been shown to be true in the case of capacity targets [21] and generation and CO_2 constraints

[10], [23]. As discussed in the introduction the majority of energy policy questions can be modelled in the cost optimisation framework [23], one notable exception would be a proportional subsidy on energy prices (e.g. a 10% additional payment to renewable suppliers) which would require the formulation of a mixed complimentary problem (MCP) [10]. However, such policies are relatively uncommon [129], potentially due to the uncertainty in subsidy cost and uncertainty in revenue to generators.

The values for the carbon policies are set in such a way as to achieve the same target carbon emissions in the deterministic case for all policies. This provides an equal basis on which to compare policies under the inclusion of uncertainty.

Quantity based policies:

CO2 Cap - An annual limit on the carbon emission in each year. The policy is implemented as a constraint on annual total emissions (Equation 11). This implementation assumes that no individual generator has a cap on emissions, but the system as a whole does face a restriction. The allocation of CO_2 emissions between units will then be determined by what is cost optimal (welfare maximising), as would be seen with a cap and trade policy.

Renewable capacity target - In this policy a minimum quantity of renewable capacity is targeted by the policy maker. One way in which this is achieved in practice is through a tender or auction process, where a government sets a required level of renewable capacity and different projects compete to be selected to provide this capacity.

Green certificates (or renewable portfolio standard) - In this policy a minimum proportion of generation relative to demand must be provided from renewable energy sources. In order to achieve this, typically retail suppliers must purchase green certificates at a set proportion to the quantity of energy they provide to end users. Renewable generators are awarded these certificates in proportion to their generation and are able to sell these to the retail suppliers. By setting the required proportion of green certificates the policy maker can then achieve their renewable generation target and renewable investors are encouraged to invest as a result of the additional revenue created from the sale of green certificates. For the purpose of modelling it is effectively assumed that the price of these certificates can rise to whatever level is sufficient to achieve renewable target goals and no additional rents above perfectly competitive levels are achieved by suppliers. This mechanism is also similar to the renewable portfolio standard (RPS).

Price based policies:

CO2 price - A price for the emission of CO_2 produced in the supply of electricity. The price on carbon emissions is included in the objective function (Equation 1) as a parameter similar to other variable costs. As with other variable costs this means this cost is taken into account in both generation and build decisions.

Generation subsidy - A subsidy for renewable energy generation provided to the system. This subsidy should be interpreted as additional to market revenue in a market environment (as opposed to a replacement payment for market revenue such as the current UK feed in tariff). The subsidy is implemented as a negative variable cost for new built renewable technologies.

Capital grant – A grant towards the capital cost of new renewable technologies (alternatively the grant can be considered a tax incentive). This policy is implemented as a reduction in capital costs for renewable technologies.

Table 20 Renewable support policy values, calibrated to achieve 350g/kwh CO_2 intensity in deterministic model

Renewable support policy	2023	2028	2033
CO_2 limit	6.2 ktonne/y	7.8 ktonne/y	9.9 ktonne/y
CO_2 price	85 $/tonne	40 $/tonne	50 $/tonne
Renewable Capacity Target	50 MW	65 MW	80 MW
Green Certificates	53 %	54%	53%
Renewable subsidy		25 $/MWh	
Renewable grant		650 $/kW	

6.4 Results

In Figure 36 the expected capacity expansion plan without any decarbonization or renewable support policy is presented under the three treatments of uncertainty (D, SA, and SO).

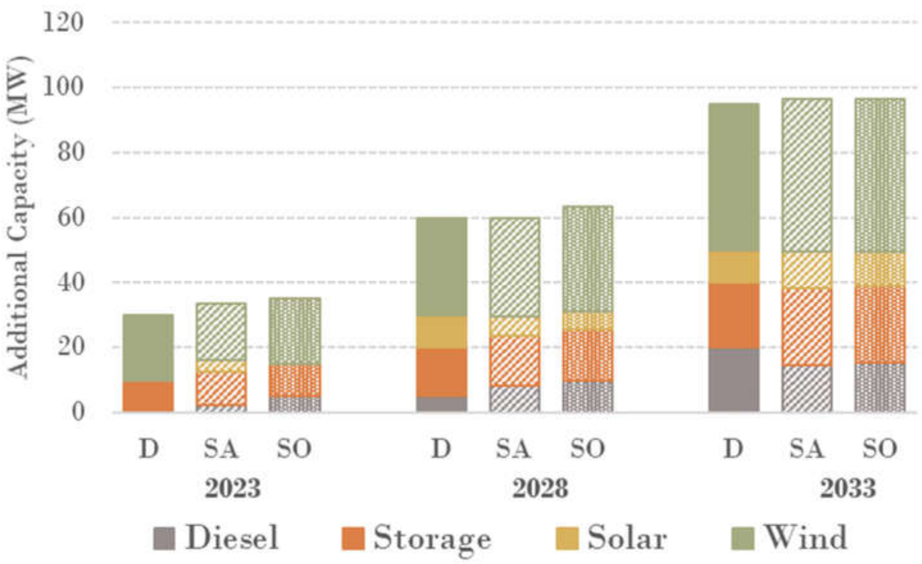


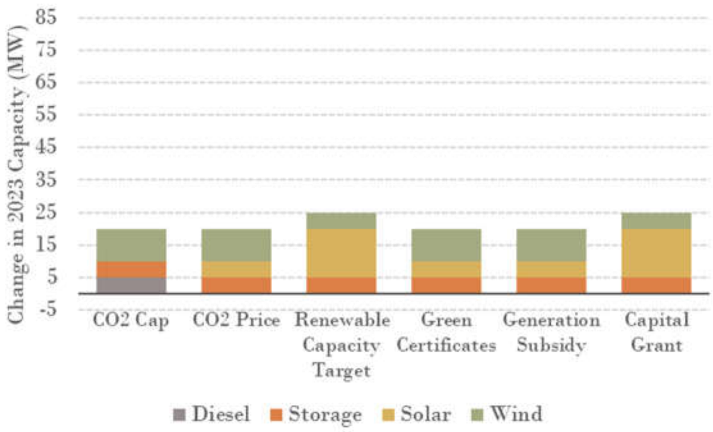
Figure 36 Base case new build results of GEP model under the three approaches to modelling uncertainty: deterministic(D), scenario average (SA), and stochasic optimisation (SO) without the application of any renewable support policy.

Firstly, it is interesting to note that under all three treatments the expansion plans are relatively similar. The first year is of the most interest as this is the year where the stochastic optimization selects decisions under uncertainty (the non-anticipativity constrained year). In this year the deterministic model selects additional onshore wind capacity and a number of new Li-ion storage units. When comparing the deterministic and scenario average results for this year, we see that diesel and solar power are selected in only a subset of scenarios, so that for the scenario weighed average build, the quantity is less than one full 5MW unit of either (if I were to round to the nearest unit, the scenario average result and the deterministic result would be the same here).

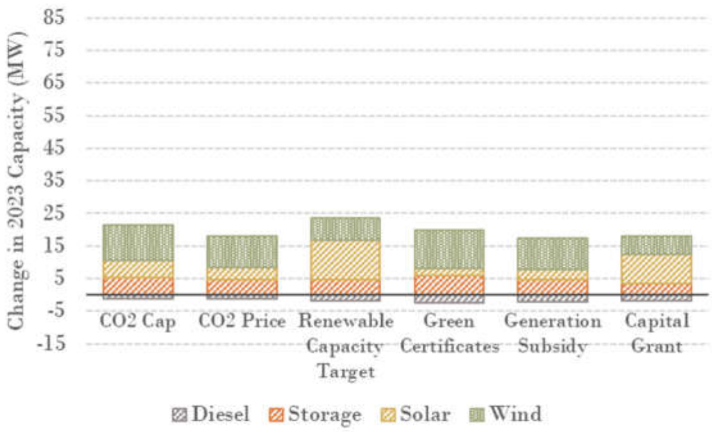
Comparing to the stochastic optimization solution, it can be seen that even though the diesel generator is only needed in a select few scenarios, we would expect that a perfectly competitive market (or a central planner) would build this generator if fully considering uncertainty in a stochastic optimisation. In practice, this means that the upside of building a diesel unit in the relatively few scenarios where it is required outweighs the downside in the majority of others.

These results demonstrate important differences between the modelling results under different treatments of uncertainty. In this situation, were policy makers to rely on a deterministic or scenario based model they may conclude that it is unlikely that any new conventional generation will be built, and perhaps delay decarbonization or renewable policy believing incorrectly that the market without any policy action would add additional conventional capacity in the short term. This result provides the first evidence that the individual scenario analysis approach is insufficient for capturing the effect of uncertainty on market participant decisions for market modelling.

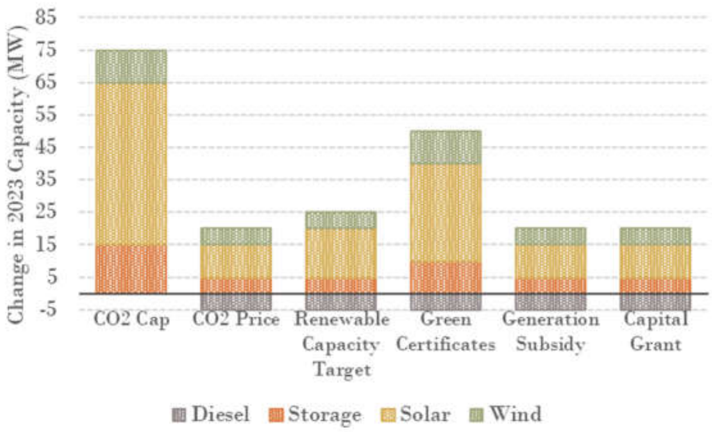
Figure 37 demonstrates the change in expansion plan in the first model year (2023) for each decarbonization support policy, as assessed by the model under the three different treatments of uncertainty. Examining first the deterministic model it is clear that the effect of the policies is reasonably similar. The notable differences are that the CO_2 Cap policy actually incentivizes the build of a new diesel unit (which would otherwise have been built later in 2028) due to the benefit of improved efficiency on CO_2 emissions. This effect is important to capture if the policy maker is concerned with the effect of "locking in" additional conventional capacity in the short run, which can be an important factor for driving system emissions over the long term [130]. Note that introducing uncertainty as individual scenarios removes this effect from the results. However, a similar effect exists with the stochastic optimisation, where the CO_2 Cap policy does not reduce Diesel build, unlike the other policies. These results illustrate an example where the inclusion of uncertainty as individual scenarios performs worse for policy comparison than the deterministic model.



(a)



(b)



(c)

Figure 37 Change in the build of new units in 2023 model year after the implementation of renewable support policies under different treatments of uncertainty.

In comparing the deterministic results one other difference is clear, the two policies that target renewable capacity (Renewable Capacity Target and Capital Grant) result in additional solar capacity being built and less onshore wind compared to the other scenarios. This difference results from the fact that the renewable generation technologies differ in the amount of their capital cost intensity. With lower capital costs, solar benefits more from this form of policy.

The results for the average of the individual scenarios remain relatively close to that of the deterministic model, indicating that the incorporation of uncertainty into the modelling with this approach would have little effect on conclusions the policy maker would draw. The main difference is a slightly higher preference for wind capacity over solar capacity, indicating that while solar capacity performs well in the average scenario (the deterministic case), wind capacity is more economic to build in a greater number of cases such that on average it is preferred to solar. The scenario average approach does find some small reduction in the build of diesel units under each of the policies, although only in limited scenarios so that on average the effect is small.

Finally, the previous results can be compared to the output of the stochastic optimisation model which I use to represent how a competitive electricity market would be expected to respond to the different policy options, given participants are taking uncertainty into account in their decisions. The most significant difference is in the CO_2 Cap policy which sets a hard limit on the quantity of CO_2 emissions the system can produce, regardless of the scenario. In the higher demand scenarios this requires the construction of significant quantities of renewable capacity, however, as decision makers must make one set of decisions given the uncertain demand, they are forced to build this high quantity of renewable energy. This contrasts to how the SA model represents this uncertainty: while in a small number of scenarios significant renewable capacity is built, on average the quantity is close to that of the deterministic model in the SA case. This difference in modelling results is at the heart of the issue with attempting to model how an electricity market will react to uncertainty as the average of a range of individual known scenarios, where all inputs for each scenario are in fact certain.

The same effect is seen with the Green Certificates policy although to a lesser extent. This policy is naturally hedged against uncertainty in demand by setting the emission target as a proportion of demand. However, again decision makers must make one set of build decisions that are capable of meeting the more extreme cases. In both of these cases a relatively high proportion of Li-ion storage capacity is also found to be economic. One advantage of this storage capacity is to help manage the renewable generation in the case that demand is low. Again, this effect is not picked up by treating uncertainty as individual scenarios, as in each of these scenarios the optimal quantity of renewable technology is constructed. The other capacity based mechanism Renewable Capacity Target does not display this effect, instead as can be seen in Figure 38 the capacity target set on the basis of a deterministic model fails to achieve the required reduction in CO_2 emissions in roughly half of all scenarios.

Overall, the price-based mechanisms are relatively similar across modelling approaches in terms of the total quantity of capacity induced by implementation of the policy. However, the mix between new solar and new wind changes with more solar expected to be built by the SO model. The SA approach finds that if, in each individual scenario, you know the future costs and demand requirement, new onshore wind is typically the most cost-efficient technology to respond to decarbonization policies, and so this technology is built often on average. However, decision makers in the market do not know the future and must decide what to build given this uncertainty. One way in which solar generation is preferable, is that in the case that demand is lower than expected on average, solar generation is better correlated with the peak of the day and less curtailment is required.

The changes in model build decisions discussed so far are the mechanism through which different policies change the CO_2 intensity of the system. Although, in practice it is typically the actual change in carbon intensity and the cost of achieving this change that the policy maker is interested in; which I present in Figure 38 and Figure 39 respectively.

Figure 38 Carbon intensity levels in the 2023 model year for all scenarios under different carbon support policies and treatments of uncertainty (D deterministic, SA scenario average, SO stochastic optimisation).

In Figure 38 I present the effectiveness of the policies at reducing carbon emissions to reach the target carbon intensity level. All policies were calibrated to achieve the target emission intensity level in the deterministic case of 250 g / kWh, however, when uncertainty is incorporated into the model this is not necessarily the case. In the case of the SA model a different set of build decisions is optimised for each individual scenario resulting in different emission intensities for each individual scenario. Both the maximum emission intensity and the expected emission intensity are higher than both the D and

SO models. This is an important result for a policy maker looking to use modelling to calibrate the levels for the different policies (eg. set the value of the capital grant). The expected value of the intensity level is up to 55% higher than the SO model (under the CO_2 Cap policy) and the maximum up to 29% higher than the SO model (under the Capital Grant policy).

This result means that a policy maker using the average of individual scenarios (as with a Monte-Carlo based model) to make decisions would set a policy 55% more restrictive than required to meet its target. In the case the policy maker is conservative and is interested in making sure the policy achieves the target in the worst-case scenario they could set the level 29% too high incurring unnecessary cost. This result in the worst-case scenario under the Capital Grant policy identifies a dynamic not yet discussed so far. In the scenario where renewable generation is the least cost effective (low fuel price, high renewable construction costs) the model only builds 5MW of additional onshore wind capacity as the subsidy is not sufficient to make additional capacity economic. In this scenario the emission intensity level is high. However, as reflected in the SO model a decision maker does not know this worst case scenario will occur, and given the uncertainty over all costs it is still economic (in expectation) to build significant renewable capacity given the proposed level of capital grant.

In Figure 39, I compare the expected costs of the policies under different model treatments of uncertainties. For the purpose of these comparisons, subsidy costs / carbon revenues are accounted for to ensure an applicable comparison[39]. When looking at the

[39] 6That is, a subsidy will reduce the costs found by the model so the cost of providing that subsidy must be added back to the modelled cost to find a value comparable with other policies. In the case of the CO_2 price revenue raised must be subtracted from the

deterministic results all policies have very similar costs, with the CO_2 targeting policies having slightly lower costs than the indirect renewable subsidy policies. The CO_2 cap policy actually has the lowest cost to achieve the target level of carbon intensity, however, the CO_2 price policy actually achieves additional reductions in CO_2 emission for a relatively small additional price, making it the cheapest on an abatement costs basis.

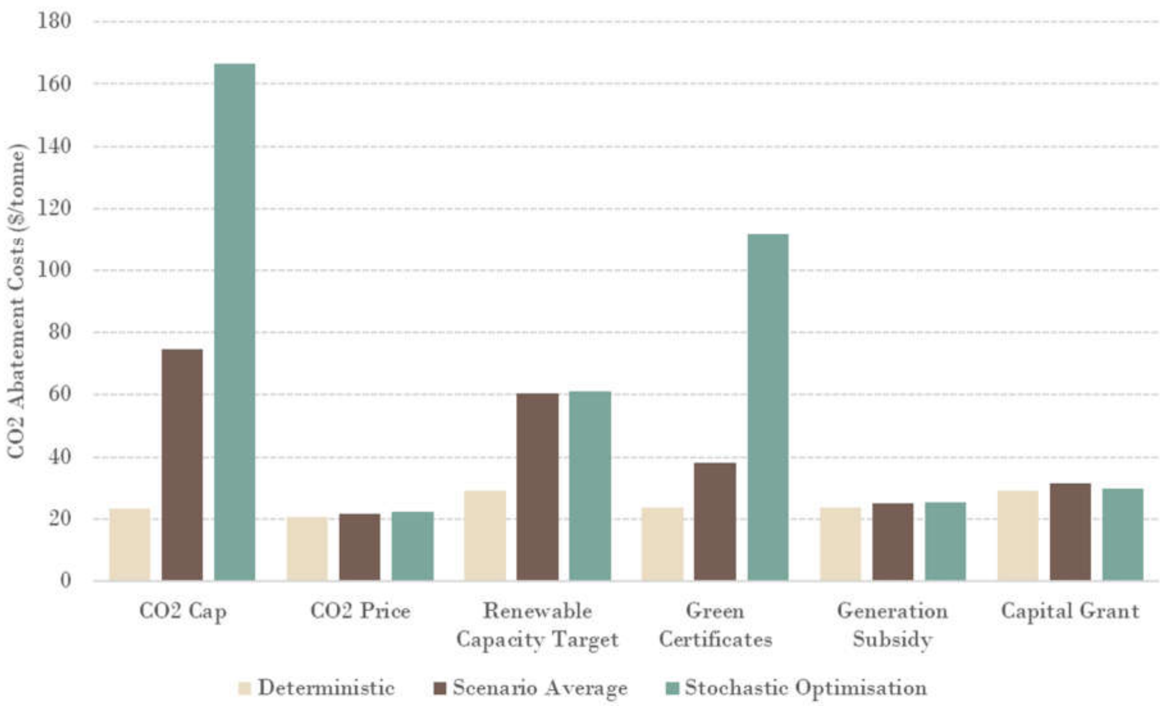


Figure 39 Expected carbon abatement costs (change in discounted total system costs resulting from renewable support policy / reduction in emissions) for all renewable support policies and treatments of uncertainty.

However, when I incorporate uncertainty this result changes. In particular, the quantity targeting policies are significantly more expensive, with costs up to seven times higher,

---

modelled cost.

as these policies enforce sufficient changes to meet the target over all scenarios, where the other price-based approaches are relatively cheaper. However, it should be noted that these quantity-based policies are the only ones that achieve the emission intensity target (in expectation). These results reinforce the fact that a carbon limit provides greater certainty around emissions, where a carbon price produces greater certainty around costs.

While it is clear that the deterministic model underestimates costs as it underestimates the amount of carbon reduction required in higher demand scenarios, I also find again that using individual scenarios to represent uncertainty generates results unrepresentative of a competitive market. In particular, the SA approach compared to the SO approach underestimates the cost of quantity targeting policies by up to 66% (Green Certificates) and overestimates cost based policies by up to 4%. Again, this results from the fact that modelling scenarios individually never captures situations where sub-optimal outcomes occur because decision makers could not see the future. For example, as discussed previously under the Capital Grant policy there are a number of situations where it is uneconomic to build additional renewable capacity (the grant is insufficient). However, given that building renewable capacity is expected to be economic considering the uncertainty in the market, building of this capacity should be expected to happen. This means that the case could occur that renewable capacity is built even though doing so is not economic and the possibility of this happening increases expected costs, however, this is effect is not captured by the SA model. In the case of overestimating costs the opposite is happening in price based policies, the SA model builds significant renewable build in cases where it is very economic (low construction costs, high fuel costs) resulting in more abatement than necessary to reach the target and in this way incurring additional costs.

6.5 Conclusions

Investment decisions in the energy sector are complex and require a return over long periods of time. Moreover, these decisions are complicated by the large amount of uncertainty inherent in the forecasts of inputs that are important for making these decisions. In this work I investigated the importance of representing uncertainty in the modelling used to evaluate different decarbonization and renewable support policies intended to influence investor decisions.

I compared six different policies options; a cap on CO_2 emissions (as with a cap and trade scheme), a CO_2 price, a renewable capacity target, a green certificates scheme, a renewable generation subsidy, and a renewable capital grant in a small power system case study. These policies reflect both direct and indirect mechanisms for reducing CO_2 intensity as well as price and quantity-based approaches.

All the policies are roughly equivalent when evaluated with a deterministic model but the technology mix, CO_2 emissions, and cost, all differ significantly when uncertainty is included inside the model. The abatement cost of the CO_2 cap policy, the most economical when assessed without uncertainty, is seven times higher in a model that takes the market's reaction to uncertainty into account.

I find that using common modelling approaches that attempt to capture uncertainty as multiple different individual scenarios (such as scenario analysis or Monte-Carlo simulation), which each individually has certain knowledge of the future, performs poorly at representing the reaction of a competitive electricity market as measured by a stochastic optimisation model. I find that incorporating uncertainty in this way can underestimate the expected carbon abatement response to policy. A policy maker using the average of individual scenarios (as with a Monte-Carlo based model) to make decisions could set a policy 55% more restrictive than required to meet its target. In the case the policy maker is conservative and is interested in making sure the policy achieves the target in the worst scenario they could set the level 29% too high incurring additional unnecessary costs.

Additionally, I find that a deterministic model that ignores uncertainty can underestimate costs by up to 85%. Incorporating uncertainty as individual scenarios only slightly improves this result. This approach to uncertainty underestimates the cost of quantity targeting policies by up to 66% (Green Certificates) and overestimates cost based policies by up to 4%.

Finally, I find important technology mix differences between the uncertainty modelling approaches. When not modelling uncertainty as a stochastic optimisation, which reflects a perfectly competitive market, the policy maker can underestimate the level of conventional capacity built without any de-carbonisation policy. In addition, I show that incorporating uncertainty as individual scenarios finds that in each individual scenario, where investors know the future costs and demand requirement, new onshore wind is typically the most cost-efficient technology to respond to decarbonization policies. However, decision makers in the market do not know the future and must decide what to build given this uncertainty.

Overall, I find a strong case for including long-term uncertainty into electricity system models in the form of a stochastic optimisation model. I find all outcomes of interest differ significantly when uncertainty is excluded, or modelled as individual scenarios, biasing the choice between different renewable energy policies. However, these results are based on a relatively small system as a result of the computational burden of stochastic optimisation modelling. In future work I extend these results to a full-size power system and explore the effect of adding market risk preferences to the modelling. This work additionally relies heavily on the assumption of perfect competition and the extension of the modelling to include market power remains an important area for additional study.

Chapter 7

Conclusions

This chapter finalises the work, summarising conclusions and considers aspects to be developed in future work.

7.1 Summary

Uncertainty is an increasingly important aspect of decision-making relating to the electricity systems of the future. Over the long-term time horizons required for investment decisions and government policy making, history indicates that forecasts tend to be varied and uncertain. Hence, long-term forecast uncertainty should be an important aspect of any GEP modelling exercise. Surprisingly, I found this to typically not be the case in the literature, so in the thesis I started by exploring the role of long-term uncertainty in decision making based upon GEP modelling in a number of conditions.

The first challenge associated with incorporating uncertainty into the GEP model is to deal with the curse of dimensionality. This problem emerges from the complexity of the model which increase exponentially with the number of random variable sources to be considered. To incorporate the wide range of random sources I intended to explore and model their uncertainties, I presented in Chapter 2 a GEP model formulation that draws on conclusions from the literature to minimise problem size while maintaining representativeness. This model was further amended to incorporate the fuel switching behaviour required to model the Ghanaian system in Chapter 4, and the decarbonisation policies examined in Chapter 6.

While this formulation allowed us to complete the analysis undertaken in this thesis, the model still remained limited to a two-stage stochastic optimisation formulation (as opposed to multistage) and to three to four scenarios per source of uncertainty modelled. One usual approach to allowing large optimisation problems to be solved efficiently is that of problem decomposition. While in this thesis I implemented and tested two decomposition approaches they were not found to be consistent in improving solution time or reducing the value of the objective function.

An additional challenge for a modeller seeking to represent uncertainty in GEP is how to obtain the scenarios to be included in the model. While it is often the case that a modeller can find forecasts for uncertain inputs such as demand or fuel prices, it is often difficult to obtain the intrinsic uncertainty of such forecasts, even when multiple scenarios are given, as usually the probability of each scenario occurring is not provided. In Chapter 2, a practical methodology for a modeller to apply historical volatility to an existing forecast of a future variable to derive a number of scenarios, along with the probability of each occurring, is provided. In Chapter 5, I find that this method produces results similar to a range of exogenous forecasts with the added benefit of providing the probability of these weightings.

In the quest to reduce model size I also investigated one of the more promising possibilities to reducing complexity within each individual scenario, the model representation of time, in Chapter 3. The challenge in attempting to reduce the level of temporal detail in the model is to ensure that sufficient detail is included so that the model is still representative. For example, a detailed representation of flexibility in the model operational detail (such as ramping and start-up constraints) cannot be expected to produce accurate results without the inclusion of the temporal detail that generates the flexibility challenges. In this chapter I examined the popular clustering approach to selecting a limited set of representative days for the model.

In Chapter 3, I initially investigated different approaches to representing the data to be clustered. The results demonstrated that the approach to representing the net demand of the power system in the clustering algorithm was important for deriving representative results. The results did not support the hypothesis that adding a representation of the ramping characteristics, in the form of a ramp duration curve found in the literature as a benchmark of ramping profile, to the clustering improved results as had originally been expected.

Instead, I found that the most significant improvement to the clustering was gained by including multiple series in the data used for clustering, specifically both the future (high renewable penetration) and historical (low renewable penetration) net demand associated with each day. When clustering was based on historical net demand, the algorithm was prone to group together days that were high load - high renewable production with medium load - low renewable production days. While these are similar in the historical data set, as renewable capacity increases into the future, the days produce very different net load profiles. In particular, I found that the clustering did not select days to represent the extreme low values of the full series net load duration curve without this approach. Both of these results demonstrated how clustering, which seeks to maximise the representativeness based on the underlying data, could provide results that do not maximise the representativeness of GEP model results.

Some additional gains were also found by ensuring that days containing specific ramping features were included as cluster medoids. I found it important to include the largest down ramp in both the duration curve and non-duration curve cases. This result was consistent across a number of clustering approaches and the four cases used in the results. I found, however, no consistent benefit to including a number of other potential extreme points such as the minimum value, which I had expected to be important for driving curtailment costs. The fact that ramping extreme points can be important to driving results can differ from some of the literature on GEP modelling [62].

In fact, the results are largely in line with this finding and I believe where I differ it reflects the size of the ramping challenge found in this case study. It is worth noticing that the UK system in the 2025 model year has high levels of renewable penetration, reasonably inflexible nuclear power, medium flexibility large biomass units, and relatively lower level of interconnection than comparable European countries. Ramping challenges, and single period ramps, clearly increase significantly with these factors [131]. The UK market also works on a 30-minute timeframe, where, a number of studies looking at ramping dynamics in GEP model use an hourly timeframe over which ramping constraints are unlikely to bind and cycling is likely to be the main issue associated with variability. For example, [95] finds in an Irish system with approximately 50% flexible capacity (vs 28% in the 2025 UK case study here) that moving from hourly to half hourly timeframes increase system costs 0.13% (or more accurately the hourly time step underestimates costs to this degree), due to ramping and additional variability effects. While small, a 0.13% increase system costs could potentially justify about £65m in investment to relieve these extra otherwise missed challenges. To ensure the robustness of this result, I also explored a number of potential causes to ensure that it was being driven by ramping challenges and not another artefact of the data on the selected 'extreme day´. I found this result to be robust and an example case where additional ES capacity was beneficial to the system in a way that can be missed by standard clustering techniques.

Chapter 3 also demonstrated that there exists important technology mix implications for the expansion plans selected by the GEP model, based upon the different approaches to clustering input data. In the majority of cases, improvements to solution optimality were gained where conventional capacity was replaced by storage capacity indicating that an oversimplification of future system challenges biased the model against ES. Based on the specific ES technology selected it seems probable that this storage is being found beneficial for improving ramping costs as opposed to shifting large quantities of additional renewables between peak and off-peak periods. This finding is reinforced by the fact that the inclusion of medoids relating to ramping (particularly downwards ramping) improved the solution quality and led to a greater preference towards storage over conventional technologies. This result reinforces the importance of ensuring that the GEP modeller accurately capture the challenges of the future system to guarantee that the right technologies are selected.

Finally, I presented a method for ensuring that the results from the clustering algorithm achieve the targeted level of net demand in the system without altering the underlying net demand shapes that are observed or predicted from historical data. This method performed well without requiring significant alterations to the output clustering weights in most of the cases (the approach performed less well where errors resulting from the clustering were large). The method allowed achieving accurate net demand and peak net demand while maintaining the underlying net demand shapes. However, the errors in total (or average) demand produced by the clustering in this thesis were small compared to many seen in the literature. I believe this simple additional step to processing clustering results could provide significant benefits in cases where the clustering produces larger errors (such as where fewer representative days are selected, or representative weeks are used).

Having investigated the approach to selecting the representation of time and short-term dynamics in an investment model, I then looked to apply this model. In the fourth chapter, the GEP model is applied to the important issue of quantifying decarbonization costs and, in doing so, explores the role of uncertainty. Sustainable electricity supply is key in supporting the socio-economic development of a country. However, the expansion of the electricity system to provide reliable access to energy comes with both a financial and environmental cost. Research into the introduction of VRES and the quantification of the cost for reducing carbon dioxide emissions from electricity generation is typically based on case studies from the developed world, whereas developing countries such as those found in the West African region are particularly under served. This creates an important gap in the literature, as opportunities and challenges of contributing to efforts to reduce carbon emissions may differ in these less developed areas.

To investigate this question, I applied the GEP model to the West African country of Ghana for the purpose of quantifying the costs of contributing to worldwide decarbonization efforts. This represented an extension to the current literature on expansion planning in the region due to the detailed incorporation of unit commitment and operational constraints. The formulation also needed to be extended to account for the unique challenges of fuel switching, a common occurrence in Ghana, where natural gas supply can be limited. In addition, the inherent uncertainty in the energy system was explored across five key dimensions, including renewable energy costs, financing costs, natural gas costs and availability, energy demand, and access to non-indigenous resources.

I found evidence for a number of important drivers of decarbonisation costs that are likely to differ between developing and developed countries. Firstly, an important difference between developed and developing countries can be the rate at which either the private or public sector can raise finance. I found that a higher discount rate as is associated with developing countries leads to significantly less investment in renewable energy when absent of some incentive. However, this means the relative cost of reducing carbon emissions in this case is lower in developing countries, with the additional cost to reduce carbon emissions in a high discount rate country less than half that of a country with a low discount rate[40] (11 USD/tonne vs 25 USD/tonne). This seemingly unintuitive result that higher financing rates, that make it more costly to invest in renewable energy, actually reduce the cost of decarbonisation, means that developing countries provide a unique opportunity to reduce emissions at a lower cost. While the ability to finance high capital cost RES is often cited as a challenge faced by developing countries, I found that addressing this issue directly can have negative consequences. I found reducing financing rates, by for instance improving developing countries access to financial markets, can lead to additional investment in conventional technology in the short-run (particularly where the country is considering coal as an option) and as such reduce the benefit in terms of decarbonisation. Interventions that would either subsidise renewable energy directly for developing countries or targeted access to finance for renewable energy

[40] All costs here were evaluated with the same discount rate.

projects only would in this case be a preferred policy.

Another key difference between the developing world and developed countries is the pace of growth in energy demand. I found this plays an important role in decarbonisation costs as the cost of displacing the build of conventional capacity is significantly less than the cost of adding renewable capacity to a system that would not otherwise need more capacity. In this way again the developing countries offer a unique opportunity to reduce carbon emissions at a lower cost than would be expected to be found in a developed country, although, again this opportunity exists as developing countries cannot be expected to bear the costs of CO2 reduction on their own.

Additionally, I found that a small reduction in emissions is more cost effective than a large cut in emissions in a given system. I found a number of drivers for this result including that the larger the share of variable renewable sources of energy the greater the required level of energy storage. On a worldwide scale this means that it is likely more cost effective that all systems seek to reduce CO2 emission as opposed to this effort being concentrated in developing countries as is currently the case. Again, this result means that relatively small incentives could achieve much greater emission reduction in developing countries, which have less ability to invest in even small amounts of RES, than in developed countries.

Finally, I found that the technology mix considered also matters. I found that a country considering the addition of new coal fired power plant (as is more likely for developing countries without political opposition) offers the lowest cost opportunity to reduce emissions over all scenarios with decarbonisation costs up to 66% lower. I additionally found that nuclear power could play a role in reducing the costs of decarbonisation by up to 46%, although the introduction of nuclear energy would involve significant challenges not considered here to implement in practice.

These results provide valuable insight into the contribution that developing countries can play in worldwide decarbonisation efforts and the hurdles they face in doing so. One clear result of this work, however, is that a lot of uncertainty exists around the future path of the energy system, which is part driven by uncertain parameters. This study found that the different assumptions around uncertain factors, such as technology costs and load growth, resulted in significantly different power generation systems and CO2 abatement costs. The decarbonisation costs of a modest target varied between 7 and 26 USD/tonne of carbon removed, a variation of more than 300%. However, the treatment of uncertainty while typical for a GEP modelling exercise may in fact underestimate the overall effect it may have on results. I found that in the case of low renewable energy costs, the investment in conventional generation was reduced in the period before any renewable energy was developed. As the model for each scenario was completely deterministic, it effectively assumed that investors could see in advance that RES prices would be low far into the future and would make decisions on this basis. Additionally, as the scenarios were descriptive an interested policy maker or system planner has no reliable way to combine the results or choose between them. The strength of the arguments I present that developing countries present an opportunity for cost effective decarbonisation depends partly on the question of which of the scenarios occurs. As I do not know even the probability that each scenario would occur, I cannot easily combine them to provide a useful metric such as the expected cost of decarbonisation given this uncertainty. Finally, as is typical in the GEP literature, I looked at each area of uncertainty individually and not in combination. These results may then miss important interaction effects, for example, it could be the case that if renewable technology costs are low at the same time that a developing country faces a lower discount rate, there may be very little value in subsiding decarbonisation in developing countries in comparison to developed ones.

These insights on the importance of long-term uncertainty, and the limitations of the way it is typically incorporated in the GEP literature, motivated the study undertaken in Chapter 5. It is clear from the results discussed above and the discussion in Section

2.2 that the scenario approach to GEP modelling did not capture the effect that uncertainty has on the decisions made by market participants. This is particularly important when modelling a competitive electricity market as there is an important distinction between approaches to incorporating uncertainty into a model. A stochastic optimisation model can be used to represent the fact that investors do not know the future of key inputs to their decision making, and this uncertainty will affect their decisions [21]. An approach such as Monte-Carlo simulation (or scenario analysis as seen in the previous chapter) instead represents uncertainty as a number of different scenarios; in each scenario investors in effect have perfect information of the future, and so uncertainty does not affect their decision making. Instead the scenario analysis approach offers a range of different potential futures and provides the range of decisions investors would make with perfect knowledge of what will occur in each case as an approximation of the distribution of actual outcomes.

In Chapter 5, I then investigated the importance of incorporating into the GEP model the reality that decision makers must take into account a wide range of economic and physical sources of uncertainty in their investment decision making. Specifically, I focused on the question of whether the fact that investors must make decisions under uncertainty, that is without knowledge of the future, is important to incorporate to produce representative descriptive modelling results. I additionally investigated if the effect of these uncertainties can be assessed independently of each other, or if there were interaction effects between them so that they must be included in combination. This meant assessing the GEP model with a stochastic optimisation formulation, which included uncertainty on all important dimensions simultaneously.

I found that, in fact, it was the case that uncertainty was important over a number of key inputs, both physical and economic, and further, the value of this uncertainty and the effect on model outputs can only accurately be assessed when the sources of uncertainty are considered together. In particular, I found an important role for the often-ignored economic sources of uncertainty such as natural gas prices, ETS CO2 prices, and technology costs. These should be considered in addition to physical uncertainties such as the level of peak demand or annual energy requirement.

In addition, I found that the simplistic treatments of uncertainty that represent uncertainty outside of the optimization model, such as scenario analysis or Monte-Carlo simulation, are unlikely to perform well for descriptive modelling analysis of competitive electricity markets for either investment or policy analysis. I found that stochastic optimisation, which combines all uncertainty into a single model and represents the fact that investors must make decisions without knowledge of the future, generates results significantly different to modelling approaches that incorporate uncertainty as individual scenarios. As discussed at length in Section 1.3 of the Introduction, it is only the stochastic optimisation that can be considered to reflect a market where investors take uncertainty into account in their decisions. As such, the fact that the results of the other approaches do not match the stochastic optimisation result has important implications for investors and policy makers as the GEP model is a common approach to modelling competitive electricity markets.

Specifically, I found that in a UK case study, the importance of combining uncertainty across different inputs was demonstrated by increases in the measured value of uncertainty (VOSS) between 67% and 109%, when compared to when the uncertainty across individual inputs was superimposed. I also stressed the importance of these results for policy makers. In this case, I demonstrated that a policy maker would see a large mismatch between the technology mix that emerges in a competitive market and the results of a deterministic planning model or the average of individual scenarios. Of importance is the fact that in the UK case study the quantity of renewable technologies is significantly reduced when uncertainty is taken into account as a SO, and the cheap but environmentally most damaging OCGT technology is preferred. These results indicate that policy makers may underestimate the support required to incentivise a given level of renewable investment if they are not using stochastic optimisation. In addition, I found that individual scenarios overestimate the range of levels of CO_2 emissions, but underestimate the range of potential electricity price outcomes.

Overall the results of Chapter 5 provided strong evidence that the simplification of the important dimension of long-term uncertainty in electricity market modelling can lead to bias in the assessment of policy. This result motivated the study in Chapter 6 where I examined this effect on the important policy topic of comparing different decarbonising policies for the electricity system.

Interestingly, as discussed in Section 1.6.1, uncertainty has played a key role in the economics literature on the comparison of different approaches to achieving decarbonisation. Specifically, the two main high-level approaches typically considered, quantity-based or price-based interventions, are equivalent in terms of social welfare under a set of simple assumptions. However, they differ in the case where carbon abatement costs are unknown. The result relies on a high-level theoretical analysis and in this thesis I extended this question into specific decarbonisation measures assessed in the preferred tool of policy makers, the GEP model.

In Chapter 6, I compared six different policies options: a cap on CO_2 emissions (as with a cap and trade scheme), a CO_2 price, a renewable capacity target, a green certificates scheme (or renewable portfolio standard), a renewable generation subsidy, and a renewable capital grant, in a small power system case study. These policies reflect both direct and indirect mechanisms for reducing CO_2 intensity as well as price- and quantity-based approaches.

To assess the value of uncertainty, I examined policies that were all roughly equivalent when evaluated with a deterministic model. I found that the technology mix, level of CO_2 emissions (or carbon abatement produced by the policy), modelled energy prices, and total cost, all differ significantly when uncertainty is included in the model. The abatement cost of the CO_2 cap policy, the most economical when assessed without uncertainty, is seven times higher than in a model that takes the market's reaction to uncertainty into account.

As with Chapter 5, I again found that common modelling approaches that attempt to capture uncertainty as multiple different individual scenarios (such as scenario analysis or Monte-Carlo simulation), where in each scenario individually the model has certain knowledge of the future, performs poorly at representing the reaction of a competitive electricity market, as measured by a stochastic optimisation model. I found that incorporating uncertainty as independent scenarios can significantly underestimate the expected carbon abatement response to policy. A policy maker using the average of individual scenarios (as with a Monte-Carlo based model) to make decisions could set a policy 55% more restrictive than required to meet its target. In the case the policy maker is conservative and is interested in making sure the policy achieves the target in the worst scenario, it could set the level 29% higher than necessary, this way incurring additional unnecessary costs.

Additionally, I found that a deterministic model that ignores uncertainty can underestimate costs by up to 85%. Incorporating uncertainty as individual scenarios only slightly improves this result. This approach to uncertainty underestimates the cost of quantity targeting policies by up to 66% (Green Certificates) and overestimates price-based policies by up to 4%.

Finally, I also found important technology mix differences between the uncertainty modelling approaches. When not modelling uncertainty as a stochastic optimisation, which reflects a perfectly competitive market, the policy maker can underestimate the level of conventional capacity build without any de-carbonisation policy. In addition, I showed that incorporating uncertainty as individual scenarios finds that in each individual scenario, where investors know the future costs and demand requirement, new onshore wind is typically the most cost-efficient technology to respond to decarbonization policies. However, decision makers in the market do not know the future and must decide what to build given this uncertainty.

Overall, I again found a strong case for including long-term uncertainty into electricity system models in the form of a stochastic optimisation model. I found all outcomes of interest differ significantly when uncertainty is excluded, or modelled as individual scenarios, biasing the choice between different renewable energy policies.

Throughout this thesis, I presented a continuous set of results that make the case for long-term uncertainty being a critical modelling consideration for the energy systems' planner, energy policy maker, or market modeller. I additionally made contributions to the difficult problem of how to reasonably represent such uncertainty into a GEP model while maintaining a level of complexity that the model could be reliably solved on standard hardware. I trust this research provides motivation for decision makers to dedicate greater effort into considering this important variable and helps to stimulate continued academic research into this subject.

7.2 Future work

The findings of this thesis are clear; that long-term uncertainty is an important driver of electricity market investment and system planning outcomes and should be incorporated into the GEP models used for decision making. However, these results are only the beginning of the attempt to achieve this goal and give rise to a number of questions and potentially fruitful avenues of future research.

The first questions that arise relate to the practical aspect of solving the large optimisation problems created by incorporating stochastic optimisation into the GEP model. While I did not find a consistent benefit to problem decomposition, it remains possible that either advances in this area, or further tuning could find a decomposition approach that would benefit this problem and allow a greater degree of uncertainty to be incorporated. In particular given the importance of uncertainty found in results it may be worth attempting to model uncertainty across an even wider set of input assumptions and this would require the problem to be decomposed to be solved on a standard desktop or laptop. Alternatively, high-performance computing could be used to at least test the value of investing in developing models that can incorporate greater uncertainty. Although, the problem before decomposition is limited in the degree to which parallelized computing power can contribute to reducing solution time, solving the model on a typical high-performance computing cluster may be possible, but still requires a long solution time.

To help decision makers incorporate uncertainty into their modelling I provided a simple method for converting a single forecasted scenario into a range of scenarios using historical volatility. The results were benchmarked against exogenous scenarios found in the literature. However, as the relative probability of these exogenous scenarios is unknown this comparison does not provide a complete benchmarking of this methodology. In future work I intend to look to benchmark this approach to a more sophisticated scenario generation model as used by the exogenous forecasters where I can generate both a range of scenarios and their relative probability for comparison. In this future research I will also look to explore a number of different underlying processes to find which are well suited to this approach. The intention is to find a simple and reliable means for converting a single exogenous forecast into a range of scenarios without requiring the GEP modeller to repeat the full forecasting exercise required to generate the original forecast.

An important step in the scenario selection process is the scenario reduction, i.e., the selection of representative scenarios from a wide range of generated scenarios. Two popular techniques were used in this thesis: fast forward selection, and clustering. However, to my knowledge no rigorous comparison of scenario reduction approaches exists and this presents an important gap in the literature. It is clear that, in particular, the fast-forward selection method tends to underrepresent the full range of the distribution of future scenarios (due to basically being a greedy algorithm as opposed to an optimisation one) and there is good reason to think clustering will perform superiorly[41]. But even given this insight there is much that can be explored in terms of different clustering techniques to find which is likely to perform best in GEP modelling. As clustering is inherently trying to represent the underlying distribution, it may also under-represent scenarios that would drive results in the GEP model such as the more extreme values. Additionally, there appears to be little research into determining the cost to model representativeness of the limited number of scenarios typically included, and there may be value in quantifying this relationship for different uncertain inputs and trying to derive rules of thumb based upon the underlying volatility of the source of uncertainty and the sensitivity of the GEP model to this input. Inherently, this research draws on the findings of Chapter 3, that extreme points are important in GEP modelling and underrepresented by clustering and applies it to a new domain.

The results of Chapter 3 provided important evidence to the question on how to best select representative days. However, overall, this study also left open a number of important questions. Firstly, the study was limited to the case where the quantity of renewable energy was known in advance and so a net demand series could be derived. While, the clustering based upon individual wind, solar, and demand series performed relatively well it is not necessarily possible to determine the important extreme days (even the peak net demand day) when this is all the information that is available. In

[41] The clustering provides superior performance in terms of representing the underlying set of scenarios based on the distance metric each approach is attempting to minimise. Which approach provides the superior performance when deriving inputs for the GEP model is an open question.

future work I look to extend these results to the case where renewable energy deployment is endogenous to the GEP model and so net demand is not known in advance.

One particularly important insight for this extension is that while the different input series (demand, wind, solar) are all normalised in some way they still have different effects on the clustering algorithm. Basically, the underlying structure in the demand and solar series (higher in the day, lower in the night) in a given day means that these series contribute less to the penalty function being minimised by the clustering. This means then that the clustering is significantly more focused on representing the wind series, as to not do so would add more to the penalty function. However, it may not be the case that maximising the representation of daily wind patterns, at the expense of the representation of demand or solar, is optimal in terms of maximising the representativeness of the GEP model based on a limited set of days. It may be the case the UK system used for this case study is more sensitive to the input wind series (due to very high onshore and offshore wind development) and this effect may be more problematic in other systems. In future work, I look to test a number of approaches to addressing this problem based on either adjusting the contribution of the different series to the penalty function or adding an additional feature to each day in the clustering to provide the clustering information on the effect of the given day in the GEP model. This second approach also provides a promising method to determine the 'extreme' days that the clustering may otherwise overlook.

In the clustering I made a strong argument for the use of medoids (and then the PAM algorithm) over the use of means (and a k-means algorithm) to represent the underlying data in the GEP model. While all these arguments remain valid, later work [132] has shown that the GEP model can be very sensitive to minor differences in the medoid selection process which is not the case with the k means selection process. This problem results from the fact that the solutions space for the medoid clustering algorithm appears to be relatively flat, meaning that locally optimal medoids can be selected, with very small difference to the globally optimal medoids in terms of the clustering metric, but with large differences in relation to the representativeness in a GEP model. The k-means approach does not suffer from this issue as there is significantly less heterogeneity in the near optimal mean points in comparison to the near optimal medoids. Chapter 3 attempts to solve this issue by testing a large number of starting points for the clustering algorithm, and in this way ensuring a consistent selection of medoids. However, this result actually reinforces the importance of using real world data, medoids, instead of means, in the GEP as the objective function is sensitive between these differences (where the clustering is not) and an open question remains in determining the clustering method that can select medoids that maximise representativeness in relation to the domain of the GEP model results.

The results of Chapter 5 provide evidence for the value of improving the treatment of long-term uncertainty in GEP models. However, there remains a number of open questions that can be explored further based on this result. Firstly, this result focuses on long term sources of uncertainty, however, there is also increasing uncertainty in the short-term dynamics of the market in the form of uncertainty in daily load patterns, and the production of VRES. These short-term sources of uncertainty are the focus of considerable research, however, the trade-off of increasing complexity by improving model representation of long- or short-term sources of uncertainty has not been explored. In future work I intend to extend the analysis of Chapter 5 to determine if modellers are best to focus on long term or short term uncertainty in their modelling, and research the situations under which either is the case. In addition, I may find complementarity between these sources of uncertainty in that including either source of uncertainty reduces the optimal quantity of VRES and so focusing on either may underestimate the total reduction in VRES that is optimal.

Another important aspect of uncertainty in the long term ignored here is that of risk. The stochastic optimisation model used in this thesis inherently assumes that either the central planner or investors in a perfectly competitive market are completely risk neutral, which may not be the case in reality. The introduction of risk preferences, either in the objective function or as risk constraints, to the modelling is likely to affect the importance of including different long-term sources of uncertainty. The results of Chapter 5 tend to suggest that the modelling of the combination of different uncertainties is likely to be even more important in this case as it may lead to more extreme costs values both positive and negative, and the more extreme potential losses will influence risk adverse investors. In future work I intend to look to examine this question by introducing risk constraints as in [15] to reflect market participant risk preferences.

A final extension that seems to follow naturally from the work of Chapter 5 is the extension from the two-stage optimisation problem to a multistage problem. Effectively, in the two stage GEP model, I am assuming investors must take uncertainty into account in their immediate investment decisions, but then all future investment decisions are made with perfect foresight of the future (no uncertainty). Given I found that taking uncertainty into account effects immediate investor decisions, it follows it should also affect their future investment decisions. As we know the future investment decisions in reality, or in a multi-stage model, could be different to those found in the two-stage model used here and this could in turn affect the immediate investment decisions. For example, in a two stage model there is often an incentive for the model to decide to build capacity later in the model, as opposed to immediately in the first stage, as there is no uncertainty and the model can perfectly optimise to a known future. When this simplification is removed it may be the case that a different capacity mix should be built immediately given this bias towards waiting for uncertainty to resolve, is reduced. Additionally, just the general result that the future generation mix will change (to a mix optimised against uncertainty) will likely affect the initial build decisions.

Chapter 6 extends the findings of Chapter 5 to the problem of assessing different policy approaches to encouraging power system decarbonisation. In future work I extend these results to a full-size power system with a greater mix of candidate build technologies to select from. This analysis may provide greater insight into the way different policies favour different renewable technologies given uncertainty. When combined with endogenous technology learning, that installation costs fall as a country develops expertise in deploying a given technology, insight into the benefits of targeting novel technologies may be derived. In addition I intend to explore the relationship between the timing of the policies which, in the analysis of Chapter 6, were announced effectively 10 years in advance for a decision making period where uncertainty is not known, that this may underestimate the relationship between the different policies and uncertainty.

Relatedly, this approach allows the modeller to examine uncertain policy proposals. For instance, the modeller could examine what level of credibility (probability of occurring) is required for an announced new energy policy to have an effect on results. The relationship between the importance of credibility in policy announcements, different types of policy, and market risk preferences could be explored. One particularly interesting application of this approach would be to look at uncertain carbon emission prices. Currently, forecasts for the EU ETS are reasonably uncertain with a wide range of values considered reasonable (see Figure 4) and this uncertainty likely causes inefficiencies in market investment decisions. It may, or may not, be the case that a lower, but more certain, EU ETS price would lead to a greater reduction in CO_2 emissions and this question would be an interesting one to investigate on an EU wide basis.

As with the conclusions of Chapter 5 I also intend to explore the effect of adding market risk preferences to the modelling. Different policy approaches to decarbonisation are often compared on the basis of the risks they impose on renewable energy investors, and the inclusion of risk preferences provides an approach to better compare policies on this dimension. The work in this thesis, additionally, relies heavily on the assumption of perfect competition and the extension of the modelling to include market power remains an important area for additional study. While this complicates significantly the modelling, the inclusion of market power may better reflect the reality of current electricity markets.

Overall, this thesis has provided evidence to demonstrate the importance of long-term uncertainty in representing market and system planning outcomes with the GEP model. The fact this area of research appears underserved in the literature means this thesis has also produced plentiful research questions which I believe can also contribute to our understanding. I hope this work does then stimulate interest and research effort so progress can continue in ensuring long-term uncertainty is properly reflected in our decision making in the electricity system

Appendix A

Additional results

Additional results not included in the main thesis.

A.1 Clustering representative days for power systems generation expansion planning: capturing the future system challenges

A.1.1 Extreme point inclusion – historical net demand - Chronological net demand

Table A.1. Key for case abbreviations

Abbreviation	Case description	Abbreviation	Case description
ND_S MRU	Historical net demand as original chronological series with day featuring max ramp up included as medoid	ND_DC MRU	Historical net demand duration curve with day featuring max ramp up included as medoid
ND_S GMMD	Historical net demand as original chronological series with day with max greatest min max distance included as medoid	ND_DC GMMD	Historical net demand duration curve as original chronological series with day with max greatest min max distance included as medoid
ND_S GMMD6H	Historical net demand as original chronological series with day with max greatest min max distance over six hours included as medoid	ND_DC GMMD6H	Historical net demand duration curve with day with max greatest min max distance over six hours included as medoid
ND_S MRD	Historical net demand as original chronological series with day with max ramp down included as medoid	ND_DC MRD	Historical net demand duration curve with day with max ramp down included as medoid
ND_S Fut MCRU	Historical net demand as original chronological series with day with max consecutive ramp up included as medoid	ND_DC Fut MCRU	Historical net demand duration curve with day with max consecutive ramp up included as medoid
ND_S Fut MCRD	Historical net demand as original chronological series with day with max consecutive ramp down included as medoid	ND_DC Fut MCRD	Historical net demand duration curve with day with max consecutive ramp down included as medoid
ND_S MinV	Historical net demand as original chronological series with day with min value included as medoid	ND_DC MinV	Historical net demand duration curve with day with min value included as medoid

Table A.2. Clustering metrics - Euclidean distance, normalised root mean square deviation (NRMSD)

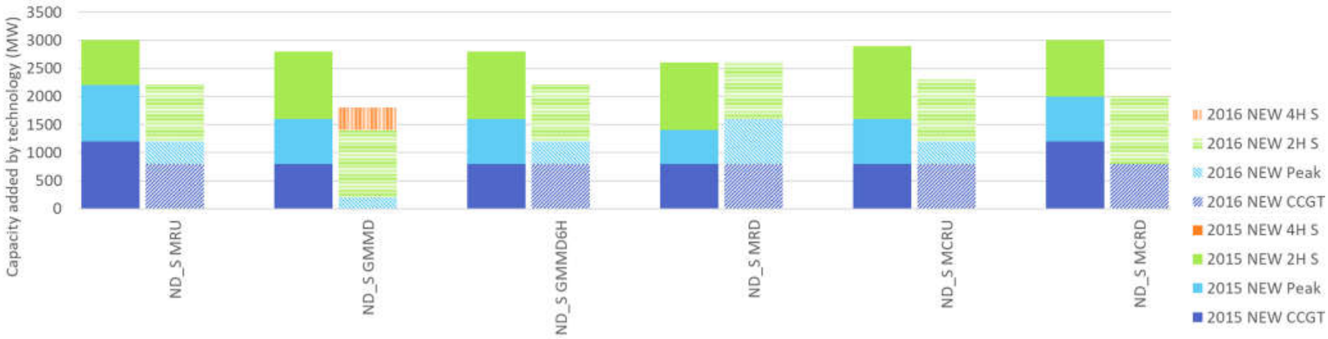
	Average squared Euclidean Distance		Normalised root mean square deviation (NRMSD) of duration curves – average of 2015 and 2016								
	2015	2016	Historical Net Demand	2020 Net Demand	2025 Net Demand	Demand	Wind	Solar	Historical Ramp	2020 Ramp	2025 Ramp

ND_S MRU	30.50	27.78	5.98 %	11.97 %	25.50 %	16.68 %	54.77 %	20.67 %	9.43 %	10.91 %	15.79 %
ND_S GMMD	30.84	27.88	5.80 %	11.20 %	21.77 %	13.63 %	42.46 %	20.47 %	8.30 %	10.20 %	15.53 %
ND_S GMMD6H	30.84	27.78	6.13 %	11.21 %	22.44 %	15.31 %	47.11 %	20.35 %	8.98 %	10.54 %	15.74 %
ND_S MRD	30.27	26.62	5.60 %	12.20 %	27.35 %	17.95 %	63.12 %	19.23 %	9.47 %	10.97 %	15.82 %
ND_S Fut MCRU	30.72	27.37	5.63 %	11.05 %	22.03 %	9.85 %	42.72 %	18.89 %	7.35 %	9.15 %	14.40 %
ND_S Fut MCRD	30.71	27.56	5.91 %	14.54 %	32.78 %	21.86 %	79.82 %	21.99 %	8.86 %	10.53 %	15.50 %
ND_S Min V	31.32	27.83	5.92 %	12.81 %	29.12 %	18.42 %	63.32 %	14.36 %	10.20 %	12.12 %	17.17 %

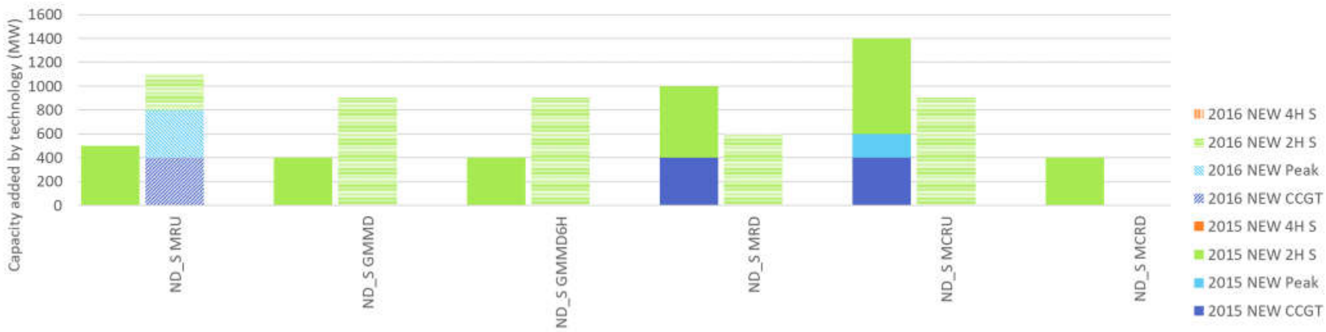
Table A.3. Average absolute change in weights from calculated scalars

	Error in net demand without scalars						Average of absolute scalar adjustments to clustering weights calculated to remove error in total net demand			
	2015 Load shapes			2016 Load shapes			2015 Load shapes		2016 Load shapes	
	2015	2020	2025	2016	2020	2025	Scaled to ND 2020	Scaled to ND 2025	Scaled to ND 2020	Scaled to ND 2025
ND_S MRU	0.2%	2.6 %	9.9%	0.0%	2.7 %	10.7 %	13.6%	29.7%	10.6%	29.3%
ND_S GMMD	- 0.2%	1.0 %	4.6%	0.1%	2.2 %	8.4%	5.1%	14.6%	8.9%	22.9%
ND_S GMMD6H	- 0.2%	1.0 %	4.6%	0.0%	2.7 %	10.7 %	5.1%	14.6%	10.6%	29.3%
ND_S MRD	- 0.1%	2.7 %	11.0 %	0.0%	3.1 %	12.2 %	13.9%	34.4%	13.4%	34.3%
ND_S Fut MCRU	- 0.1%	1.7 %	6.8%	0.4%	1.9 %	6.3%	8.2%	22.7%	7.8%	17.3%

ND_S Fut MCRD	0.2%	4.9 %	19.2 %	0.1%	2.7 %	10.3 %	22.8%	57.8%	10.8%	28.7%
ND_S MinV	0.3%	3.4 %	12.8 %	0.1%	2.9 %	11.2 %	17.3%	40.7%	11.2%	30.0%



(a)



(b)

Figure A.1. Capacity expansion plan by representative day selection approach (a) 2020 expansion model (b) 2025 expansion model

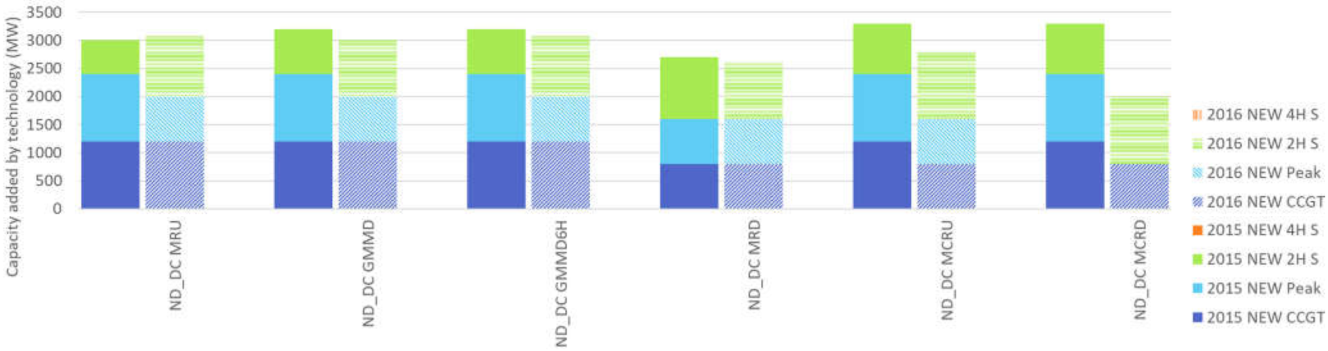
A.1.2 Extreme point inclusion – historical net demand – Net Demand duration curve

Table A.4. Clustering metrics - Euclidean distance, normalised root mean square deviation (NRMSD)

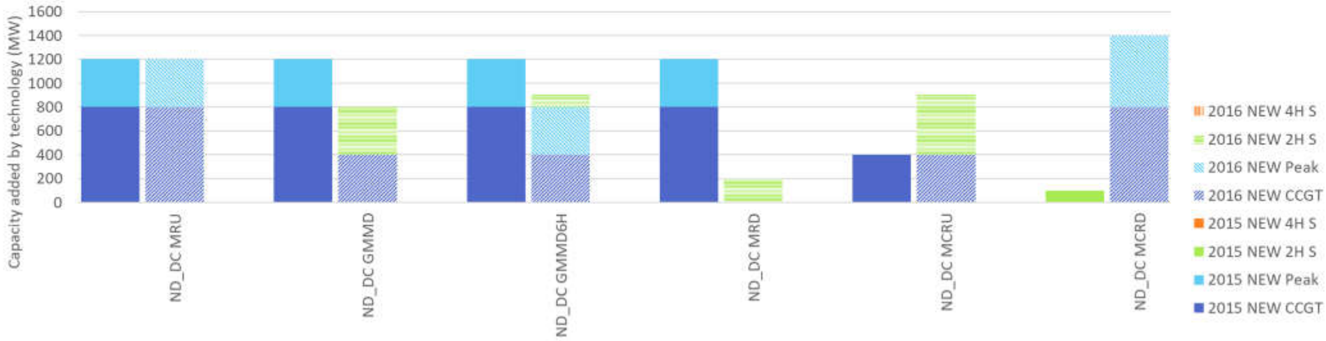
	Average squared Euclidean Distance		Normalized root mean square deviation (NRMSD) of duration curves – average of 2015 and 2016								
	2015	2016	Historical Net Demand	2020 Net Demand	2025 Net Demand	Demand	Wind	Solar	Historical Ramp	2020 Ramp	2025 Ramp
ND_DC MRU	15.89	16.08	5.56%	10.19%	19.57%	11.76%	35.60%	17.16%	7.23%	8.44%	12.60%
ND_DC GMMD	16.10	15.97	5.82%	10.67%	20.96%	14.98%	39.45%	22.75%	7.86%	9.97%	14.89%
ND_DC GMMD6H	16.10	16.08	5.56%	11.14%	21.43%	14.11%	39.46%	20.97%	8.58%	10.86%	15.89%
ND_DC MRD	15.57	15.44	5.46%	10.63%	21.30%	12.46%	34.67%	17.41%	7.31%	8.51%	12.96%
ND_DC Fut MCRU	16.07	15.84	5.54%	11.82%	24.75%	16.58%	47.93%	19.71%	8.17%	10.38%	15.37%
ND_DC Fut MCRD	15.81	15.60	5.13%	10.62%	22.43%	13.23%	36.24%	14.73%	6.26%	7.32%	11.62%
ND_DC MinV	16.26	15.83	4.79%	9.37%	17.40%	12.60%	36.33%	15.75%	7.52%	9.04%	13.73%

Table A.5. Average absolute change in weights from calculated scalars

	Error in net demand without scalars						Average scalar of weights			
	2015 Load shapes			2016 Load shapes			2015 Load shapes		2016 Load shapes	
	2015	2020	2025	2016	2020	2025	Scaled to ND 2020	Scaled to ND 2025	Scaled to ND 2020	Scaled to ND 2025
ND_DC MRU	0.0%	1.2%	4.9%	-0.1%	0.9%	4.1%	6.4%	15.6%	4.1%	11.5%
ND_DC GMMD	0.0%	2.6%	10.4 %	-0.3%	0.7%	3.6%	12.5%	31.9%	3.0%	9.6%
ND_DC GMMD6H	0.0%	2.5%	9.6%	-0.1%	0.9%	4.1%	11.5%	28.9%	4.1%	11.5%
ND_DC MRD	0.0%	1.1%	4.4%	-0.2%	0.8%	3.6%	5.7%	14.1%	3.4%	9.8%
ND_DC Fut MCRU	0.0%	2.8%	11.1 %	0.0%	2.0%	8.0%	13.6%	33.6%	8.4%	21.9%
ND_DC Fut MCRD	0.0%	1.8%	7.3%	-0.1%	0.1%	0.8%	9.1%	20.6%	0.5%	2.3%
ND_DC MinV	0.2%	2.3%	8.8%	-0.2%	0.4%	2.0%	11.4%	24.6%	1.5%	4.7%



(a)



(b)

Figure A.2. Capacity expansion plan by representative day selection approach (a) 2020 expansion model (b) 2025 expansion model

A.1.3 Extreme point inclusion – clustering with combined historical and future net demand

Table A.6. Key for case abbreviations

Abbreviation	Case description	Abbreviation	Case description
NDHF_S MRU	Historical and future year net demand as original chronological series with day featuring max ramp up included as medoid	NDHF_DC MRU	Historical and future year net demand duration curve with day featuring max ramp up included as medoid
NDHF_S GMMD	Historical and future year net demand as original chronological series with day with max greatest min max distance included as medoid	NDHF_DC GMMD	Historical and future year net demand duration curve as original chronological series with day with max greatest min max distance included as medoid
NDHF_S GMMD6H	Historical and future year net demand as original chronological series with day with max greatest min max distance over six hours	NDHF_DC GMMD6H	Historical and future year net demand duration curve with day with max greatest min max distance over six hours included as medoid

	included as medoid		
NDHF_S MRD	Historical and future year net demand as original chronological series with day with max ramp down included as medoid	NDHF_DC MRD	Historical and future year net demand duration curve with day with max ramp down included as medoid
NDHF_S Fut MCRU	Historical and future year net demand as original chronological series with day with max consecutive ramp up included as medoid	NDHF_DC Fut MCRU	Historical and future year net demand duration curve with day with max consecutive ramp up included as medoid
NDHF_S Fut MCRD	Historical and future year net demand as original chronological series with day with max consecutive ramp down included as medoid	NDHF_DC Fut MCRD	Historical and future year net demand duration curve with day with max consecutive ramp down included as medoid
NDHF_S MinV	Historical and future year net demand as original chronological series with day with min value included as medoid	NDHF_DC MinV	Historical and future year net demand duration curve with day with min value included as medoid

Clustering and scaling results - Historical and future year net demand as original chronological series

Table A.7. Clustering metrics - Euclidean distance, normalised root mean square deviation (NRMSD)

	Average squared Euclidean Distance		Normalized root mean square deviation (NRMSD) of duration curves – average of 2015 and 2016								
	2015	2016	Historical Net Demand	2020 Net Demand	2025 Net Demand	Demand	Wind	Solar	Historical Ramp	2020 Ramp	2025 Ramp
NDHF_S MRU	50.38	41.36	7.32%	9.03%	10.27%	9.45%	36.27%	17.45%	8.96%	12.02%	12.35%
NDHF_S GMMD	48.33	41.48	6.62%	8.54%	9.65%	9.84%	35.50%	19.44%	9.74%	12.03%	13.65%
NDHF_S GMMD6H	48.20	44.22	6.81%	8.51%	11.19%	9.11%	37.36%	25.25%	9.00%	11.35%	13.82%
NDHF_S MRD	49.92	41.98	7.48%	8.96%	10.40%	9.33%	32.05%	14.78%	9.08%	13.21%	19.92%
NDHF_S Fut MCRU	49.88	42.82	7.31%	7.30%	10.06%	10.58%	32.48%	17.71%	7.92%	10.71%	13.72%

NDHF_S Fut MCRD	50.59	41.23	7.10%	9.60%	9.53%	11.12%	37.65%	19.85%	8.44%	11.28%	11.29%
NDHF_S MinV	50.32	41.12	6.49%	8.54%	9.20%	10.34%	36.82%	19.46%	9.24%	12.34%	13.57%

Table A.8. Average absolute change in weights from calculated scalars

	Error in net demand without scalars						Average of absolute scalar adjustments to clustering weights calculated to remove error in total net demand			
	2015 Load shapes			2016 Load shapes			2015 Load shapes		2016 Load shapes	
	2015	2020	2025	2016	2020	2025	Scaled to ND 2020	Scaled to ND 2025	Scaled to ND 2020	Scaled to ND 2025
NDHF_S MRU	-0.7%	0.7%	1.6%	-0.4%	0.9%	1.3%	3.1%	3.6%	3.2%	3.1%
NDHF_S GMMD	0.4%	1.3%	0.0%	-0.4%	0.9%	1.1%	5.8%	0.1%	3.2%	2.6%
NDHF_S GMMD6H	0.4%	1.3%	0.7%	-0.3%	1.2%	2.2%	5.8%	1.7%	4.0%	4.5%
NDHF_S MRD	0.2%	0.7%	0.5%	-0.6%	0.7%	1.2%	3.0%	1.4%	3.0%	3.1%
NDHF_S Fut MCRU	-0.2%	0.4%	2.3%	-0.4%	0.9%	1.5%	1.8%	5.0%	3.3%	3.7%
NDHF_S Fut MCRD	-1.0%	1.1%	-0.1%	-0.3%	0.6%	1.3%	5.3%	0.1%	2.4%	3.3%
NDHF_S MinV	-0.6%	1.3%	-0.3%	-0.4%	0.9%	1.3%	5.9%	0.7%	3.1%	3.0%

Clustering and scaling results - Historical and future year net demand duration curve

Table A.9. Clustering metrics - Euclidean distance, normalised root mean square deviation (NRMSD)

	Average squared Euclidean Distance		Normalized root mean square deviation (NRMSD) of duration curves – average of 2015 and 2016								
	2015	2016	Historical Net Demand	2020 Net Demand	2025 Net Demand	Demand	Wind	Solar	Historical Ramp	2020 Ramp	2025 Ramp
NDHF_S MRU	27.23	24.62	5.91%	7.99%	7.96%	9.09%	32.78%	24.47%	7.85%	10.34%	17.49%
NDHF_S GMMD	26.28	24.57	5.52%	6.90%	6.97%	9.24%	24.84%	18.58%	7.93%	9.86%	15.04%

NDHF _S GMMD 6H	26.26	26.33	5.34%	7.11%	7.83%	8.73%	26.63%	16.75%	7.73%	9.75%	13.93%
NDHF _S MRD	26.80	25.28	6.30%	7.66%	8.57%	9.35%	26.44%	17.10%	8.18%	11.73%	22.20%
NDHF _S Fut MCRU	27.25	25.50	5.89%	7.75%	7.33%	10.87%	28.02%	21.74%	8.25%	11.66%	13.79%
NDHF _S Fut MCRD	26.99	24.44	5.62%	7.01%	6.84%	8.94%	29.65%	22.88%	9.28%	9.92%	13.75%
NDHF _S MinV	28.36	24.30	6.37%	7.08%	7.15%	10.82%	31.84%	15.45%	8.24%	9.01%	13.63%

Table A.10. Average absolute change in weights from calculated scalars

	Error in net demand without scalars						Average of absolute scalar adjustments to clustering weights calculated to remove error in total net demand			
	2015 Load shapes			2016 Load shapes			2015 Load shapes		2016 Load shapes	
	2015	2020	2025	2016	2020	2025	Scaled to ND 2020	Scaled to ND 2025	Scaled to ND 2020	Scaled to ND 2025
NDHF_S MRU	-0.6%	-0.2%	-1.3%	0.2%	-0.3%	-0.1%	0.7%	2.9%	1.1%	0.2%
NDHF_S GMMD	-0.3%	-0.1%	-1.2%	0.2%	0.2%	-0.7%	0.5%	2.8%	0.9%	1.5%
NDHF_S GMMD6H	-0.3%	-0.1%	-1.2%	0.1%	-0.1%	0.7%	0.5%	2.8%	0.2%	1.3%
NDHF_S MRD	0.1%	0.0%	-1.8%	-0.1%	0.3%	-0.4%	0.0%	4.3%	1.2%	0.9%
NDHF_S Fut MCRU	-0.6%	0.1%	-0.4%	0.2%	0.4%	0.4%	0.2%	0.9%	1.6%	0.9%
NDHF_S Fut MCRD	-0.3%	0.2%	-1.0%	0.4%	-0.3%	0.1%	0.9%	2.4%	1.1%	0.1%
NDHF_S MinV	-0.5%	0.1%	-1.4%	0.2%	0.3%	0.0%	0.3%	3.2%	1.2%	0.1%

A.2 Assessing the costs of contributing to climate change targets in sub-Saharan Africa: The case of the Ghanaian electricity system

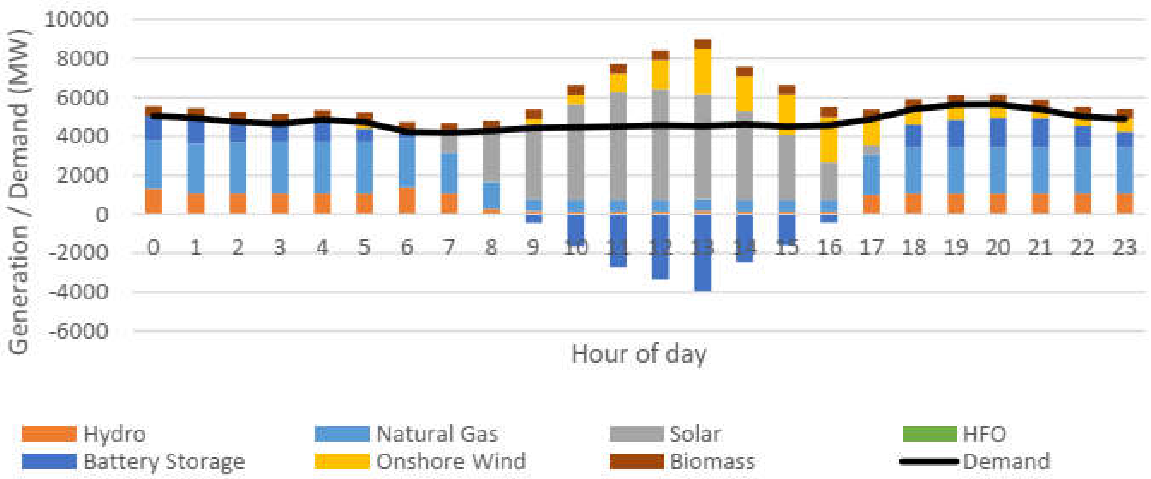


Figure A.3. Typical generation by time of day 2040

Appendix B

Additional case study details

Details of case study parameters and data.

B.1 Clustering representative days for power systems generation expansion planning: capturing the future system challenges

B.1.1 Existing capacity

Existing units capacity and technology sourced from [133] are all implemented individually with binary commitment variables, additional characteristics adapted from [116,134,135].

Table B.1. Existing technology capacities

	2020 Capacity (MW)	2025 Capacity (MW)
Solar	13842	17871
Onshore Wind	12793	16459
Offshore Wind	9428	14873
CCGT	27169	24524
Coal	0	0
OCGT	773	564
Nuclear	8918	4768
Biomass	2706	2706
Storage	3589	5910
STOR units	2500	2500

B.1.2 Candidate expansion technologies

The following technologies are considered for expansion planning with the following associated costs (adapted from [68,116,134,135]) :

Table B.2. New build technology characteristics (a) conventional (b) storage

(a)

	Capacity Unit (MW)	Max Units per station	Maximum stations built	Build Cost (£/kW)	Annual fixed cost (£/kW-yr.)	Minimum Stable Level (%)	Efficiency (%)	Start-up cost (£/MW)
New CCGT	400	1	6	763.46	7.77	25%	50.5%	56.65
New Peaker	210	2	6	632.99	9.32	25%	37.5%	36.05

(b)

	Capacity Unit (MW)	Max units per Station	Maximum stations built	Build Cost 2020 (£/kWh)	Round trip efficiency
Energy storage (2hour and 5-hour storages modelled)	100	4	7	116.5	90%

An annual discount of factor of 7.5% is assumed and all technologies are assumed to have the same economic lifetime of 20 years for simplicity. I implement a supply curve for new technologies by increasing the capital costs of each technology by 20% as for each additional station of a given technology that is built, a station can have multiple units built at the same cost[42]. For my purposes this assumption improves the robustness of the comparison as a small change cannot result in the optimal technology mix switching completely from one technology to another and we can be more confident of large technology changes when we observe them.

[42] This assumption can be thought to reflect either a preference for diversification or that as the optimal siting for new technologies is exploited less optimal siting must be used and installation costs increase.

B.2 Assessing the costs of contributing to climate change targets in sub-Saharan Africa: The case of the Ghanaian electricity system.

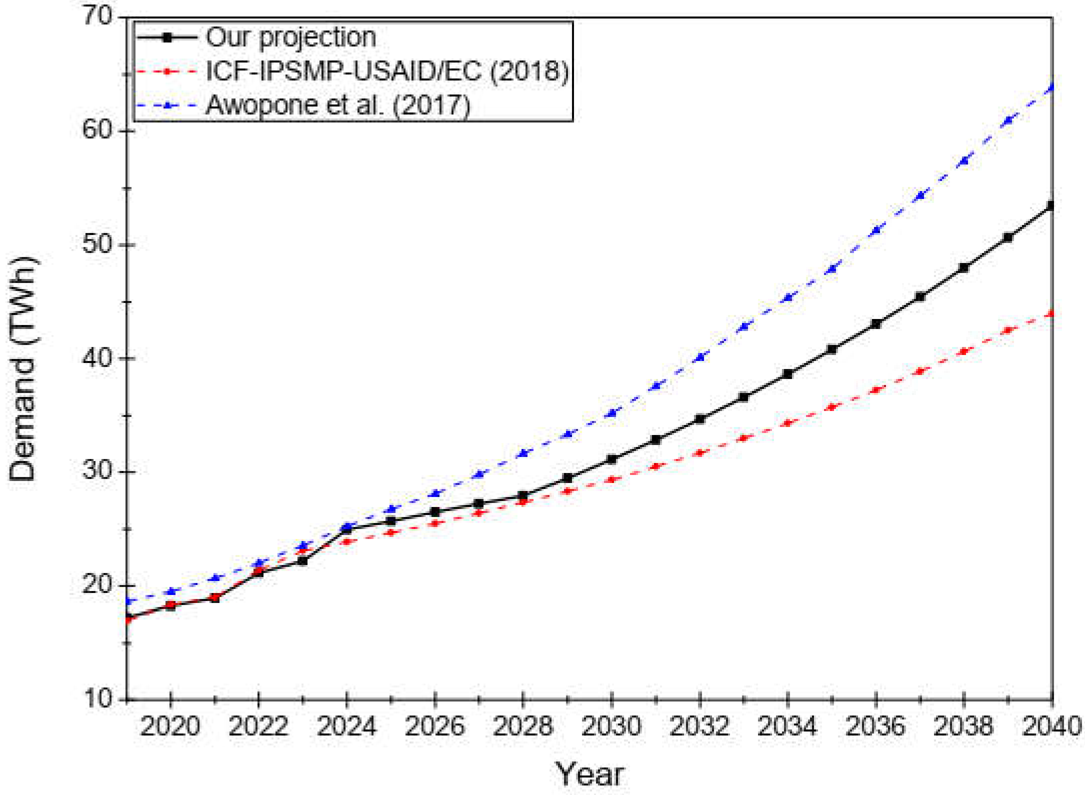


Figure B.1. Demand projection vs alternative found in literature

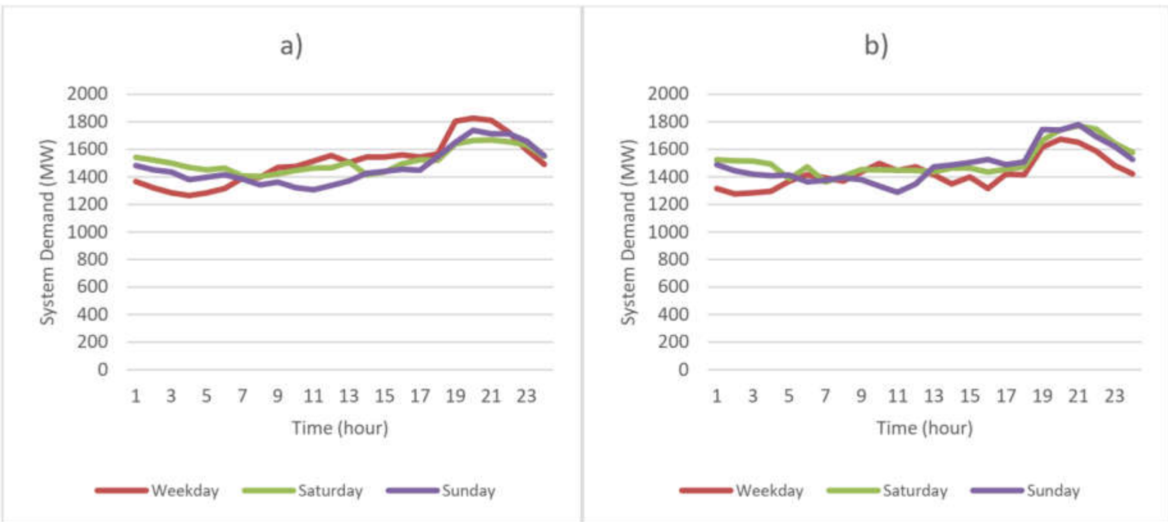


Figure B.2. Typical hourly profile by a) January-driest month and b) June wettest month in the year 2013.

Table B.3. Existing technology capacities

GENERATION PLANT		FUEL TYPE	CAPACITY (MW)				TOTAL GENERATION		
			Installed (name plate)	*% Share*	Average Dependable	Average Available	GWh	*% Share (incl.embedd)*	*% Share (exl.embedd)*
Hydro Power Plants	Akosombo	Hydro	1,020		900	505	4,282	*30.5*	*30.6*
	Bui	Hydro	400		340	205	582	*4.1*	*4.2*
	Kpong	Hydro	160		140	115	752	*5.3*	*5.4*
		Sub-Total	***1,580***	***35.9^{+} 36.7***	***1,380***	***825***	***5,616***	***39.9***	***40.2***
Thermal Power Plants									
Takoradi Power Company (TAPCO)		Oil/NG	330		300	200	686	*4.9*	*4.9*
Takoradi Inter. Company (TICO)		Oil/NG	340		320	260	1,880	*13.4*	*13.4*
Sunon–Asogli Power (SAPP)		NG	560		520	180	1,417	*10.1*	*10.1*
Kpone Thermal Power Plant(KTPP)		Oil/DFO	220		200	20	124	*0.9*	*0.9*
Tema Thermal Plant1 (TT1P)		Oil/NG	110		100	70	365	*2.6*	*2.6*
Tema Thermal Plant2 (TT2P)		Oil/NG	80		70	1	0.5	*0.0*	*0*
CENIT Energy Ltd (CEL)		Oil/NG	110		100	30	59	*0.4*	*0.4*
	AMERI	NG	250		230	200	1,229	*8.7*	*8.8*
	Karpower	HFO	470		450	225	1,814	*12.9*	*13.0*
	AKSA	HFO	260		220	100	799	*5.7*	*5.7*
		Sub – Total	***2,730***	***63.3***	**2,510**	**1,286**	**8,373.5**		
	Trojan*	Diesel/NG	44		40	30	52	*0.4*	-
	Genser*	Coal/LPG	22		18	0	0	*0*	-
Sub-total (including embedded generation)			***2,796***	***63.6***	***2,568***	***1,316***	***8,425.5***	***59.9***	
Renewables*	VRA Solar	Solar	2.5		1.5	1.5	3.0	0.02	
	BXC Solar	Solar	20		16	10	25	0.18	
		Sub – Total	***22.5***	***0.5***	11.5	11.5	***28.0***	***0.2***	
Total *(including embedded generation+ Solar)*			**4,398.5**		**3966**	**2,198**	**14,069**		
Total *(excluding embedded generation and solar)*			**4,310**		**3,890**	**2,156**	**13,989**		

Note: *NG is Natural gas. * Sub-transmission (primary embedded) connection. $^{+}$Including embedded generation and solar.*

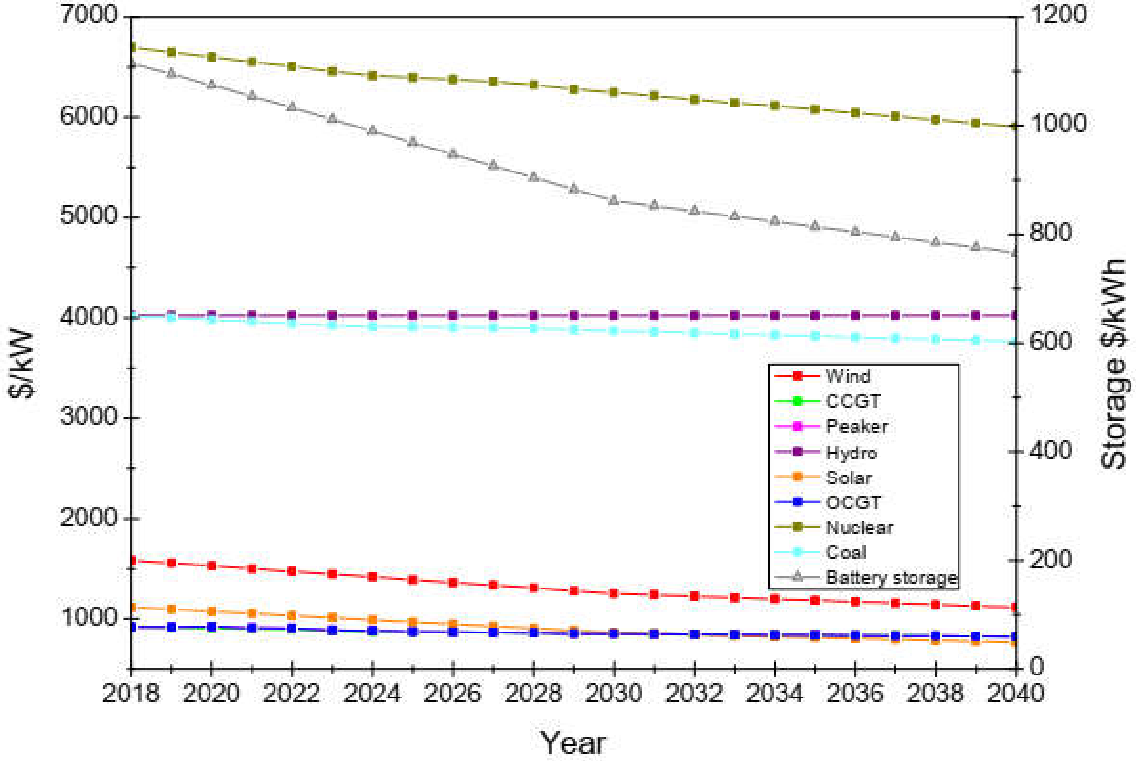


Figure B.3. Construction costs for base case

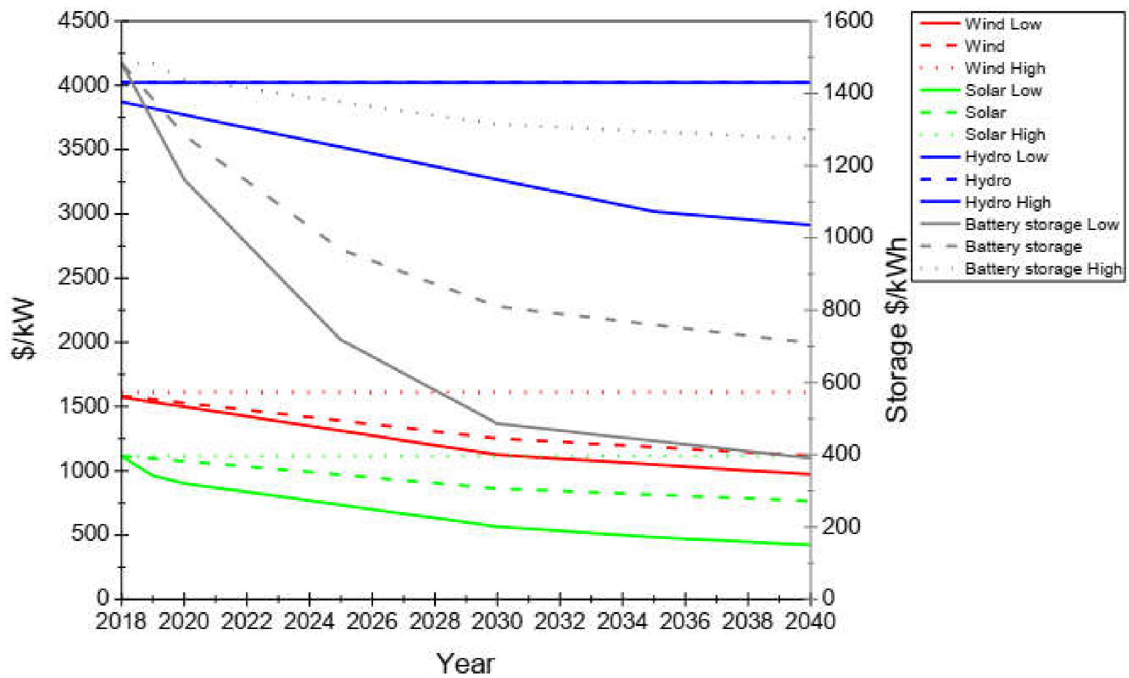


Figure B.4. Renewable Energy Construction costs by Scenario

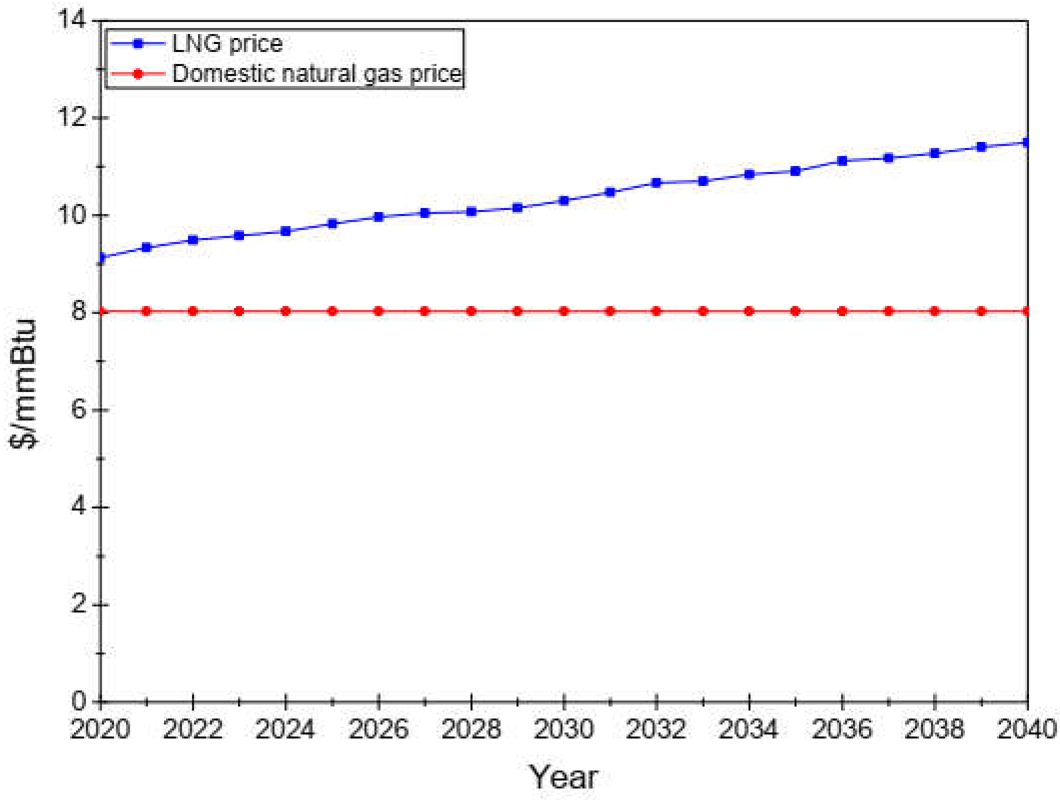


Figure B.5. Gas price by scenario

B.3 Long-term uncertainties in generation expansion planning: Implications for electricity market modelling and policy

Table B.4. Existing generation technologies. Characteristics derived largely from [121]. Capacity (with expected newbuild removed) sourced from [136].

	Solar	Onshore Wind	Offshore Wind	Biomass	OCGT	CCGT	Nuclear	Other Renewables	Other Thermal	Storage
Efficiency (%)	100%	100%	100%	35%	33%	53%	33%	25%	29%	80%*
Unit size (MW)	500	500	500	500	500	500	500	500	500	2940
Total capacity (GW)	12.5	11.5	6.5	5.5	2.5	27.0	5.0	3.0	1.5	2.9
Outage de-rating (%)	0%	0%	0%	78%	30%	85%	81%	64%	30%	100%

Minimum stable level (%)	0%	0%	0%	25%	25%	33%	80%	25%	25%	0%
Variable operating cost ($/MWh)	-50.00	-50.00	-50.00	5.47	7.03	2.72	2.27	5.47	8.43	2.50
Start cost ($/MW)	0.00	0.00	0.00	61.00	28.00	44.00	56.00	61.00	28.00	0.00
Operating cost ($/hr.)	0	0	0	2145	1038	1157	924	1752	1246	0
Fixed operating & maintenance cost ($/kW)	14.03	51.33	137.16	32.53	12.02	10.39	99.20	109.84	14.42	15.00

* Round trip

Nine representative days (representing demand, wind, and solar) with hourly time steps are selected from historical demand and generation data.

B.4 Renewable energy support policy evaluation: the role of long-term uncertainty in market modelling

Table B.5. Existing generation technologies

Name	Fuel	Capacity MW
Wind	None	12.5
Grupo 5-6	Diesel	12.2
Grupo 7-8	Diesel	12.2
Grupo 9-10	Diesel	24.6

References

[1] IPCC. Climate Change 2014: Synthesis Report. vol. 218. 2014. doi:10.1017/CBO9781107415324.

[2] Capros P, Paroussos L, Fragkos P, Tsani S, Boitier B, Wagner F, et al. European decarbonisation pathways under alternative technological and policy choices: A multi-model analysis. Energy Strateg Rev 2014;2:231–45. doi:10.1016/j.esr.2013.12.007.

[3] Craig P, Gadgil A, Koomey JG. What Can History Teach Us? A Retrospective Examination of Long-Term Energy Forecasts for the United States*. Ssrn 2011:83–118. doi:10.1146/annurev.energy.27.122001.083425.

[4] Haas J, Cebulla F, Cao K, Nowak W, Palma-behnke R, Rahmann C, et al. Challenges and trends of energy storage expansion planning for fl exibility provision in low-carbon power systems – a review. Renew Sustain Energy Rev 2017;80:603–19. doi:10.1016/j.rser.2017.05.201.

[5] Moret S, Codina Gironès V, Bierlaire M, Maréchal F. Characterization of input uncertainties in strategic energy planning models. Appl Energy 2017;202:597–617. doi:10.1016/j.apenergy.2017.05.106.

[6] Pfenninger S, Hawkes A, Keirstead J. Energy systems modeling for twenty-first century energy challenges. Renew Sustain Energy Rev 2014;33:74–86. doi:10.1016/j.rser.2014.02.003.

[7] Sadeghi H, Rashidinejad M, Abdollahi A. A comprehensive sequential review study through the generation expansion planning. Renew Sustain Energy Rev 2017;67:1369–94. doi:10.1016/j.rser.2016.09.046.

[8] Oree V, Sayed Hassen SZ, Fleming PJ. Generation expansion planning optimisation with renewable energy integration: A review. Renew Sustain Energy Rev 2017;69:790–803. doi:10.1016/j.rser.2016.11.120.

[9] Frew BA, Becker S, Dvorak MJ, Andresen GB, Jacobson MZ. Flexibility mechanisms and pathways to a highly renewable US electricity future. Energy 2016;101. doi:10.1016/j.energy.2016.01.079.

[10] Poncelet K. Long-term energy-system optimization models Capturing the challenges of integrating intermittent renewable energy sources and assessing the suitability for descriptive scenario analyses. KU Leuven, 2018. doi:10.13140/RG.2.2.21586.45762.

[11] Gacitua L, Gallegos P, Henriquez-Auba R, Lorca, Negrete-Pincetic M, Olivares D, et al. A comprehensive review on expansion planning: Models and tools for energy policy analysis. Renew Sustain Energy Rev 2018;98:346–60. doi:10.1016/j.rser.2018.08.043.

[12] Pereira S, Ferreira P, Vaz AIF. Generation expansion planning with high share of renewables of variable output. Appl Energy 2017;190:1275–88. doi:10.1016/j.apenergy.2017.01.025.

[13] Babatunde OM, Munda JL, Hamam Y. Generation Expansion Planning: A Survey. 2018 IEEE PES/IAS PowerAfrica, PowerAfrica 2018 2018:307–12. doi:10.1109/PowerAfrica.2018.8521173.

[14] Ringkjøb HK, Haugan PM, Solbrekke IM. A review of modelling tools for energy and electricity systems with large shares of variable renewables. Renew Sustain Energy Rev 2018;96:440–59. doi:10.1016/j.rser.2018.08.002.

[15] de Maere d'Aertrycke G, Ehrenmann A, Ralph D, Smeers Y. Risk trading in capacity equilibrium models. Cambridge Work Pap Econ 2018. doi:10.17863/CAM.17552.

[16] Schweppe FC, Caramanis MC, Tabors RD, Bohn RE. Spot Pricing of Electricity. 1988. doi:10.1007/978-1-4613-1683-1.

[17] Feldman AM, Serrano R. Welfare economics and social choice theory, 2nd edition. 2006. doi:10.1007/0-387-29368-X.

[18] Smith A. An inquiry into the wealth of nations. Strahan and Cadell, London 1776.

[19] Hammond PJ. The Efficiency Theorems and Market Failure. Depatment Econ Stanford Univ 1997:1–57.

[20] Boucher J, Smeers Y. Alternative models of restructured electricity systems, Part 1: No market power. Oper Res 2001;49:821–38. doi:10.1287/opre.49.6.821.10017.

[21] Gurkan G, Ozdemir O, Smeers Y. Generation Capacity Investments in Electricity Markets: Perfect Competition. SSRN Electron J 2013. doi:10.2139/ssrn.2314862.

[22] Murphy F, Pierru A, Smeers Y, Murphy F, Pierru A, Smeers Y. A Tutorial on Building Policy Models as Mixed- Complementarity Problems. INFORMS J Appl Anal Publ 2016;46:465–81. doi:https://doi.org/10.1287/inte.2016.0842.

[23] Tennbakk B. Electricity Market Functioning : Current Distortions , and How to Model Their Removal. 2016.

[24] European Union. METIS Technical Note T4: Overview of European Electricity Markets 2016.

[25] Koltsaklis NE, Dagoumas AS. State-of-the-art generation expansion planning: A review. Appl Energy 2018;230:563–89. doi:10.1016/j.apenergy.2018.08.087.

[26] Kannan R. Uncertainties in key low carbon power generation technologies - Implication for UK decarbonisation targets. Appl Energy 2009;86:1873–86. doi:10.1016/j.apenergy.2009.02.014.

[27] Castillo AC. Cost Methodologies Project CCGT Case Study. UKERC Technol Policy Assess 2013.

[28] Morris JF, Srikrishnan V, Webster MD, Reilly JM. Hedging strategies: Electricity investment decisions under policy uncertainty. Energy J 2018;39:101–22. doi:10.5547/01956574.39.1.jmor.

[29] Eryilmaz D, Homans FR. How does uncertainty in renewable energy policy affect decisions to invest in wind energy? Electr J 2016;29:64–71. doi:10.1016/j.tej.2015.12.002.

[30] Boomsma TK, Linnerud K. Market and policy risk under different renewable electricity support schemes. Energy 2015;89:435–48. doi:10.1016/j.energy.2015.05.114.

[31] Stremel JP. Generation system planning under load forecast uncertainty. IEEE Trans Power Appar Syst 1981. doi:10.1109/TPAS.1981.316867.

[32] Cole W, Mai T, Logan J, Steinberg D, Mccall J, Richards J, et al. 2017 Standard Scenarios Report : A U . S . Electricity Sector Outlook 2016 Standard Scenarios Report : A U . S . Electricity Sector Outlook 2017. doi:10.1371/journal.pone.0144005.

[33] Heuberger CF, Staffell I, Shah N, Mac Dowell N. Impact of myopic decision-making and disruptive events in power systems planning. Nat Energy 2018. doi:10.1038/s41560-018-0159-3.

[34] Lumbreras S, Ramos A. The new challenges to transmission expansion planning. Survey of recent practice and literature review. Electr Power Syst Res 2016;134:19–29. doi:10.1016/j.epsr.2015.10.013.

[35] Lund PD, Lindgren J, Mikkola J, Salpakari J. Review of energy system flexibility measures to enable high levels of variable renewable electricity. Renew Sustain Energy Rev 2015;45:785–807. doi:10.1016/j.rser.2015.01.057.

[36] L.M.V.G PMV. P. Application of Decomposition Techniques to the Mid-and Short-Term Scheduling of Hydrothermal Systems. IEEE Power Eng Rev

1983;PER-3:37–8. doi:10.1109/MPER.1983.5520203.

[37] Pereira MVF, Pinto LMVG. Stochastic Optimization of a Multireservoir Hydroelectric System: A Decomposition Approach. Water Resour Res 1985;21:779–92. doi:10.1029/WR021i006p00779.

[38] Pereira MVF, Pinto LMVG. Multi-stage stochastic optimization applied to energy planning. Math Program 1991;52:359–75. doi:10.1007/BF01582895.

[39] Alqurashi A, Etemadi AH, Khodaei A. Treatment of uncertainty for next generation power systems: State-of-the-art in stochastic optimization. Electr Power Syst Res 2016;141:233–45. doi:10.1016/j.epsr.2016.08.009.

[40] Conejo AJ, Carrion M, Morales JM. Decision Making Under Uncertainty in Electricity Markets. 2010. doi:10.1007/978-1-4419-7421-1.

[41] Dai H, Zhang N, Su W. A Literature Review of Stochastic Programming and Unit Commitment. J Power Energy Eng 2015;03:206–14. doi:10.4236/jpee.2015.34029.

[42] Zhan Y. Optimization Approaches for Electricity Generation Expansion Planning Under Uncertainty 2017.

[43] Pigou AC. The Economics Welfare. 1932.

[44] Tang BJ, Wang XY, Wei YM. Quantities versus prices for best social welfare in carbon reduction: A literature review. Appl Energy 2019;233–234:554–64. doi:10.1016/j.apenergy.2018.10.018.

[45] Keohane NO. Cap and trade, rehabilitated: Using tradable permits to control U.S. greenhouse gases. Rev. Environ. Econ. Policy, 2009. doi:10.1093/reep/ren021.

[46] Weitzman ML. Prices vs. Quantities. Rev Econ Stud 1974. doi:10.2307/2296698.

[47] Botterud A, Auer H. Resource Adequacy with Increasing Shares of Wind and Solar Power: A Comparison of European and U.S. Electricity Market Designs. MIT CEEPR Work Pap Ser 2018;9:1–26. doi:10.1093/jnci/djv313.

[48] Sakellaris K, Canton J, Zafeiratou E, Fournié L. METIS – An energy modelling tool to support transparent policy making. Energy Strateg Rev 2018. doi:10.1016/j.esr.2018.08.013.

[49] Müller B, Gardumi F, Hülk L. Comprehensive representation of models for energy system analyses: Insights from the Energy Modelling Platform for Europe (EMP-E) 2017. Energy Strateg Rev 2018;21:82–7. doi:10.1016/j.esr.2018.03.006.

[50] Scott IJ, Carvalho PMS, Botterud A, Silva CA. Clustering representative days for power systems generation expansion planning: Capturing the effects of variable renewables and energy storage. Appl Energy 2019;253:113603. doi:10.1016/j.apenergy.2019.113603.

[51] Diawuo FA, Scott IJ, Baptista PC, Silva CA. Assessing the costs of contributing to climate change targets in sub-Saharan Africa: The case of the Ghanaian electricity system. Energy Sustain Dev 2020. doi:10.1016/j.esd.2020.05.001.

[52] Scott IJ, Carvalho P, Botterud A, Santos Silva CA. Long-term uncertainties in generation expansion planning: Implications for electricity market modelling and policy. Energy 2021:120371. doi:10.1016/j.energy.2021.120371.

[53] Scott IJ, Botterud A, Carvalho PMS, Silva CA. Renewable support policy evaluation : The importance of uncertainty. Appl Energy Symp MIT A+B 2019;1.

[54] Scott IJ, Botterud A, Carvalho PMS, Silva CAS. Renewable energy support policy evaluation: The role of long-term uncertainty in market modelling. Appl Energy 2020. doi:10.1016/j.apenergy.2020.115643.

[55] Neves D, Scott I, Silva CA. Peer-to-peer energy trading potential: An assessment for the residential sector under different technology and tariff availabilities. Energy 2020. doi:10.1016/j.energy.2020.118023.

[56] Jordehi AR. How to deal with uncertainties in electric power systems? A review. Renew Sustain Energy Rev 2018;96:145–55. doi:10.1016/j.rser.2018.07.056.

[57] Zhou Y, Wang L, McCalley JD. Designing effective and efficient incentive policies for renewable energy in generation expansion planning. Appl Energy 2011;88:2201–9. doi:10.1016/j.apenergy.2010.12.022.

[58] Collins S, Deane JP, Poncelet K, Panos E, Pietzcker RC, Delarue E, et al. Integrating short term variations of the power system into integrated energy system models : A methodological review. Renew Sustain Energy Rev 2017;76:839–56. doi:10.1016/j.rser.2017.03.090.

[59] Poncelet K, Van Stiphout A, Meus J, Delarue E, D'haeseleer W. LUSYM Invest : a generation expansion planning model with a high level of temporal and technical detail 2018.

[60] Jenkins J, Sepulveda N. Enhanced Decision Support for a Changing Electricity Landscape: the GenX Configurable Electricity Resource Capacity Expansion Model. MIT Energy Initiat Work Pap 2017:1–40.

[61] Meus J, Poncelet K, Delarue E. Applicability of a Clustered Unit Commitment Model in Power System Modeling. IEEE Trans Power Syst 2018;33:2195–204. doi:10.1109/TPWRS.2017.2736441.

[62] Poncelet K, Delarue E, Six D, Duerinck J, D'haeseleer W. Impact of the level of temporal and operational detail in energy-system planning models. Appl Energy 2016;162. doi:10.1016/j.apenergy.2015.10.100.

[63] Ramos A, Lumbreras S. How to solve the transmission expansion planning problem faster: acceleration techniques applied to Benders' decomposition. IET Gener Transm Distrib 2016;10:2351–9. doi:10.1049/iet-gtd.2015.1075.

[64] Watson JP, Woodruff DL. Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems. Comput Manag Sci 2011. doi:10.1007/s10287-010-0125-4.

[65] Ordoudis C, Pinson P, Zugno M, Morales JM. Stochastic unit commitment via

Progressive Hedging - Extensive analysis of solution methods. 2015 IEEE Eindhoven PowerTech, PowerTech 2015 2015. doi:10.1109/PTC.2015.7232629.

[66] Feng Y, Ryan SM. Scenario construction and reduction applied to stochastic power generation expansion planning. Comput Oper Res 2013;40:9–23. doi:10.1016/j.cor.2012.05.005.

[67] Feng Y. Scenario generation and reduction for long-term and short-term power system generation planning under uncertainties 2014.

[68] National Grid. Future Energy Scenarios. Natl Grid 2016:220. doi:10.1016/B978-0-08-099424-6.00029-6.

[69] Jin S, Ryan SM, Watson J-P, Woodruff DL. Modeling and solving a large-scale generation expansion planning problem under uncertainty. Energy Syst 2011;2:209–42. doi:10.1007/s12667-011-0042-9.

[70] Lundmark R. The Economics of Power Generation Technology Choice and Investment Timing in the Presence of Policy Uncertainty. Low Carbon Econ 2012;3:1–10. doi:10.4236/lce.2012.31001.

[71] Kaut M, Wallace SW. Evaluation of scenario-generation methods for stochastic programming. Pacific J Optim 2007;3:257–71. doi:---.

[72] Kaufman L, Rousseeuw PJ. Partitioning Around Medoids. Find. Groups Data, 1990, p. 68–125. doi:10.1002/9780470316801.ch1.

[73] Heitsch H, Römisch W. Scenario reduction algorithms in stochastic programming. Comput Optim Appl 2003;24:187–206. doi:10.1023/A:1021805924152.

[74] Lin C, Fang C, Chen Y, Liu S, Bie Z. Scenario Generation and Reduction Methods for Power Flow Examination of Transmission Expansion Planning 2017:90–5. doi:10.1109/ICPESYS.2017.8215927.

[75] Koltsaklis NE, Georgiadis MC. A multi-period, multi-regional generation

expansion planning model incorporating unit commitment constraints. Appl Energy 2015. doi:10.1016/j.apenergy.2015.08.054.

[76] Pina A, Silva CA, Ferrão P. High-resolution modeling framework for planning electricity systems with high penetration of renewables. Appl Energy 2013. doi:10.1016/j.apenergy.2013.05.074.

[77] Frew BA, Jacobson MZ. Temporal and spatial tradeoffs in power system modeling with assumptions about storage: An application of the POWER model. Energy 2016;117:198–213. doi:10.1016/j.energy.2016.10.074.

[78] Jin S, Botterud A, Ryan SM. Temporal versus stochastic granularity in thermal generation capacity planning with wind power. IEEE Trans Power Syst 2014;29. doi:10.1109/TPWRS.2014.2299760.

[79] Merrick JH. On representation of temporal variability in electricity capacity planning models. Energy Econ 2016;59. doi:10.1016/j.eneco.2016.08.001.

[80] Pina A, Silva C, Ferrão P. Modeling hourly electricity dynamics for policy making in long-term scenarios. Energy Policy 2011. doi:10.1016/j.enpol.2011.06.062.

[81] Blanford GJ, Merrick JH, Bistline JET, Young DT. Simulating Annual Variation in Load, Wind, and Solar by Representative Hour Selection. Energy J 2018;39:189–212.

[82] Ehrenmann a., Smeers Y. Generation Capacity Expansion in a Risky Environment: A Stochastic Equilibrium Analysis. Oper Res 2011;59:1332–46. doi:10.1287/opre.1110.0992.

[83] Nahmmacher P, Schmid E, Hirth L, Knopf B. Carpe diem: A novel approach to select representative days for long-term power system modeling. Energy 2016;112:430–42. doi:10.1016/j.energy.2016.06.081.

[84] Liu Y, Sioshansi R, Conejo AJ. Hierarchical Clustering to Find Representative

Operating Periods for Capacity-Expansion Modeling. IEEE Trans Power Syst 2017;33:3029–39. doi:10.1109/TPWRS.2017.2746379.

[85] Ludig S, Haller M, Schmid E, Bauer N. Fluctuating renewables in a long-term climate change mitigation strategy. Energy 2011;36:6674–85. doi:10.1016/j.energy.2011.08.021.

[86] Sisternes FJ De, Webster MD. Optimal Selection of Sample Weeks for Approximating the Net Load in Generation Planning Problems. ESD Work Pap Ser 2013.

[87] ElNozahy MS, Salama MMA, Seethapathy R. A probabilistic load modelling approach using clustering algorithms. 2013 IEEE Power Energy Soc. Gen. Meet., 2013, p. 1–5. doi:10.1109/PESMG.2013.6672073.

[88] Härtel P, Kristiansen M, Korpås M. Assessing the impact of sampling and clustering techniques on offshore grid expansion planning. Energy Procedia 2017;137:152–61. doi:10.1016/j.egypro.2017.10.342.

[89] Poncelet K, Hoschle H, Delarue E, Virag A, D'haeseleer W. Selecting representative days for capturing the implications of integrating intermittent renewables in generation expansion planning problems. IEEE Trans Power Syst 2016;PP. doi:10.1109/TPWRS.2016.2596803.

[90] Green R, Staffell I, Vasilakos N. Divide and Conquer? k-means clustering of demand data allows rapid and accurate simulations of the British electricity system. IEEE Trans Eng Manag 2014;61:251–60. doi:10.1109/TEM.2013.2284386.

[91] Kristiansen M, Korp M. Sensitivity analysis of sampling and clustering techniques in expansion planning models 2017;609795. doi:10.1109/EEEIC.2017.7977727.

[92] Kotzur L, Markewitz P, Robinius M, Stolten D. Impact of different time series aggregation methods on optimal energy system design 2017.

[93] Krad, I., Gao DW, Member S. An Assessment of Flexibility Reserves in Stochastic Modeling at Multiple Timescales 2017;3:84–92. doi:10.17775/CSEEJPES.2017.0011.

[94] Pfenninger S. Dealing with multiple decades of hourly wind and PV time series in energy models: A comparison of methods to reduce time resolution and the planning implications of inter-annual variability. Appl Energy 2017;197:1–13. doi:10.1016/j.apenergy.2017.03.051.

[95] Deane JP, Drayton G, Ó Gallachóir BP. The impact of sub-hourly modelling in power systems with significant levels of renewable generation. Appl Energy 2014;113:152–8. doi:10.1016/j.apenergy.2013.07.027.

[96] International Energy Agency. Energy Access Outlook 2017: From poverty to prosperity. vol. 94. 2017. doi:10.1787/9789264285569-en.

[97] Mentis D, Welsch M, Fuso Nerini F, Broad O, Howells M, Bazilian M, et al. A GIS-based approach for electrification planning-A case study on Nigeria. Energy Sustain Dev 2015;29:142–50. doi:10.1016/j.esd.2015.09.007.

[98] Avila N, Carvallo JP, Shaw B, Kammen DM. The energy challenge in sub-Saharan Africa : A guide for advocates and policy makers. vol. Part 1. 2017. doi:10.13140/RG.2.2.25037.44001.

[99] Riva F, Ahlborg H, Hartvigsson E, Pachauri S, Colombo E. Electricity access and rural development: Review of complex socio-economic dynamics and causal diagrams for more appropriate energy modelling. Energy Sustain Dev 2018;43:203–23. doi:10.1016/j.esd.2018.02.003.

[100] ECREEE. ECOWAS Renewable energy and energy efficiency status report. vol. 34. 2014.

[101] International Energy Agency. Africa Energy Outlook-A focus on energy prospects in Sub-Saharan Africa. 2014.

[102] Nunoo AR. Ghana's energy sector policies presentation, 2017.

[103] GRIDCo. Electricity supply plan for Ghana. vol. 53. 2018. doi:10.1017/CBO9781107415324.004.

[104] Diawuo FA, Pina A, Baptista PC, Silva CA. Energy efficiency deployment: A pathway to sustainable electrification in Ghana. J Clean Prod 2018;186:544–57. doi:10.1016/j.jclepro.2018.03.088.

[105] Sakah M, De la Rue du Can S, Diawuo FA, Delight M, Kuhn C. A study of appliance ownership and electricity consumption determinants in urban Ghanaian households. Sustain Cities Soc 2019;44:559–81. doi:10.1016/j.scs.2018.10.019.

[106] Energy Commission. National energy statistics (2000-2013). 2014.

[107] Energy Commission. ENERGY SUPPLY AND DEMAND OUTLOOK FOR GHANA. 2018.

[108] Ouedraogo NS. Modeling sustainable long-term electricity supply-demand in Africa. Appl Energy 2017;190:1047–67. doi:10.1016/j.apenergy.2016.12.162.

[109] Delarue E, Cattrysse D, D'haeseleer W. Enhanced priority list unit commitment method for power systems with a high share of renewables. Electr Power Syst Res 2013;105:115–23. doi:10.1016/j.epsr.2013.07.014.

[110] Gyamfi S, Modjinou M, Djordjevic S. Improving electricity supply security in Ghana - The potential of renewable energy. Renew Sustain Energy Rev 2015;43:1035–45. doi:10.1016/j.rser.2014.11.102.

[111] Sakah M, Diawuo FA, Katzenbach R, Gyamfi S. Towards a sustainable electrification in Ghana: A review of renewable energy deployment policies. Renew Sustain Energy Rev 2017;79:544–57. doi:10.1016/j.rser.2017.05.090.

[112] Mensah LD, Yamoah JO, Adaramola MS. Performance evaluation of a utility-scale grid-tied solar photovoltaic (PV) installation in Ghana. Energy Sustain Dev

2019;48:82–7. doi:10.1016/j.esd.2018.11.003.

[113] Awopone AK, Zobaa AF, Banuenumah W. Techno-economic and environmental analysis of power generation expansion plan of Ghana. Energy Policy 2017;104:13–22. doi:10.1016/j.enpol.2017.01.034.

[114] Afful-Dadzie A, Afful-Dadzie E, Awudu I, Banuro JK. Power generation capacity planning under budget constraint in developing countries. Appl Energy 2017;188:71–82. doi:10.1016/j.apenergy.2016.11.090.

[115] Diawuo FA, Kaminski J. An analysis of the Ghanaian power generation sector using an optimization model. J Power Technol 2017;97:15–27.

[116] Kumar N, Besuner P, Lefton S, Agan D, Hilleman D. Power Plant Cycling Costs 2012. doi:10.2172/1046269.

[117] NREL. Annual Technology Baseline (ATB) Cost and Performance Data for Electricity Generation Technologies. Natl Renew Energy Lab 2019.

[118] Energy Commission, USAID. Integrated Power System Master Plan for Ghana. 2018.

[119] IEA. World Energy Outlook 2017 Excerpt - Outlook for Natural Gas. 2017. doi:https://doi.org/10.1787/20725302.

[120] Prices of fuels purchased by major power producers - GOV.UK n.d. https://www.gov.uk/government/statistical-data-sets/prices-of-fuels-purchased-by-major-power-producers (accessed January 7, 2020).

[121] Vimmerstedt LJ, Augustine CR, Beiter PC, Cole WJ, Feldman DJ, Kurup P, et al. 2018 Annual Technology Baseline (ATB). 2018.

[122] Transparent Cost Database | Transparent Cost Database n.d. https://openei.org/apps/TCDB/ (accessed January 7, 2020).

[123] Escudero LF, Garín A, Merino M, Pérez G. The value of the stochastic solution in

multistage problems. Top 2007;15:48–64. doi:10.1007/s11750-007-0005-4.

[124] NREL. Annual Technology Baseline (ATB) Data. 2017 2017. https://atb.nrel.gov/electricity/2017/index.html?t=in (accessed November 26, 2018).

[125] Heuberger CF, Rubin ES, Staffell I, Shah N, Mac Dowell N. Power capacity expansion planning considering endogenous technology cost learning. Appl Energy 2017;204:831–45. doi:10.1016/j.apenergy.2017.07.075.

[126] He Y. Analysis of investment decision making in power systems under environmental regulations and uncertainties 2013.

[127] Pye S, Sabio N, Strachan N. An integrated systematic analysis of uncertainties in UK energy transition pathways. Energy Policy 2015;87:673–84. doi:10.1016/j.enpol.2014.12.031.

[128] Levin T, Kwon J, Botterud A. The long-term impacts of carbon and variable renewable energy policies on electricity markets. Energy Policy 2019;131:53–71. doi:10.1016/j.enpol.2019.02.070.

[129] Abdmouleh Z, Alammari RAM, Gastli A. Review of policies encouraging renewable energy integration & best practices. Renew Sustain Energy Rev 2015;45:249–62. doi:10.1016/j.rser.2015.01.035.

[130] Haoqi Q, Libo W, Weiqi T. “Lock-in” effect of emission standard and its impact on the choice of market based instruments. Energy Econ 2017;63:41–50. doi:10.1016/j.eneco.2017.01.005.

[131] Huber M, Dimkova D, Hamacher T. Integration of wind and solar power in Europe: Assessment of flexibility requirements. Energy 2014;69:236–46. doi:10.1016/j.energy.2014.02.109.

[132] Teichgraeber H, Brandt AR. Clustering methods to find representative periods for

the optimization of energy systems: An initial framework and comparison. Appl Energy 2019;239:1283–93. doi:10.1016/j.apenergy.2019.02.012.

[133] Department for Business E& IS. Digest of United Kingdom Energy Statistics (DUKES) 2016: main chapters and annexes - GOV.UK 2016. https://www.gov.uk/government/statistics/digest-of-united-kingdom-energy-statistics-dukes-2016-main-chapters-and-annexes (accessed October 25, 2017).

[134] Gonzalez-Salazar MA, Kirsten T, Prchlik L. Review of the operational flexibility and emissions of gas- and coal-fired power plants in a future with growing renewables. Renew Sustain Energy Rev 2018;82:1497–513. doi:10.1016/j.rser.2017.05.278.

[135] Grid N. Electricity scenarios illustrator n.d.

[136] National Grid. Future Energy Scenarios. 2018.

[137] Chao HP, Peck S. A market mechanism for electric power transmission. J Regul Econ 1996. doi:10.1007/BF00133357.

[138] Aien M, Hajebrahimi A, Fotuhi-Firuzabad M. A comprehensive review on uncertainty modeling techniques in power system studies. Renew Sustain Energy Rev 2016;57:1077–89. doi:10.1016/j.rser.2015.12.070.