Analysis of market incentives on power system planning and operations in liberalised electricity markets

Meng Xu

A Thesis presented for the degree of
Doctor of Philosophy

School of Engineering and Computing Sciences
Durham University
United Kingdom

December 2016

Dedicated to

my parents

Abstract

The design of liberalised electricity markets (e.g., the energy, capacity and ancillary service markets) is a topic of much debate, regarding their ability to trigger adequate investment in generation capacities and to incentivize flexible power system operation.

Long-term generation investment (LTGI) models have been widely used as a decision-support tool for generation investments and design of energy policy. Of particular interest is quantification of uncertainty in model outputs (e.g., generation projections or system reliability) given a particular market design while accounting for uncertainties in input data as well as the discrepancies between the model and the reality. Unfortunately, the standard Monte Carlo based techniques for uncertainty quantification require a very large number of model runs which may be impractical to achieve for a complex LTGI model. In order to enable efficient and fully systematic analysis, it is therefore necessary to create an emulator of the full model, which may be evaluated quickly for any input and which quantifies uncertainty in the output of the full model at inputs where it has not been run. The case study shows results from the Great Britain power system exemplar which is representative of LTGI models used in real policy processes. In particular, it demonstrates the application of Bayesian emulation to a complex LTGI model that requires a formal calibration, uncertainty analysis, and sensitivity analysis.

In power systems with large amounts of variable generation, it is important to provide sufficient incentives for operating reserves as a main source of generation flexibility. In the traditional unit commitment (UC) model, the demand for operating reserves is fixed and inelastic, which does not reflect the marginal value of operating reserves in avoiding the events of load shedding and wind curtailment. Besides, the system-wide reserve constraint assumes that the operating reserve can be delivered to any location freely, which is not true in real-world power system operations. To recognize the value and deliverability of operating reserves, dynamic zonal operating reserve demand curves are introduced to an enhanced deterministic UC model for co-optimizing the day-ahead schedules for energy and operating reserves. In the case study on the RTS-73 test system, comparisons are made be-

tween the choices of reserve policies (e.g., single, seasonal or dynamic zones) and of different reserve zonal partitioning methods. Results suggest that the enhanced deterministic UC model produces on average lower operational cost, higher system reliability and higher energy and reserve revenues than the traditional one.

Finally, we discuss future directions of methodological research arising from current energy system challenges and the computer models developed for better understanding of the impacts of market incentives on power system planning and operations.

Declaration

The work in this thesis is based on research carried out at School of Engineering and Computing Sciences, Durham University, United Kingdom. No part of this thesis has been submitted elsewhere for any other degree or qualification and it is all my own work unless referenced to the contrary in the text.

Acknowledgements

I would like to first thank my PhD supervisors, Dr. Chris Dent, Dr. Amy Wilson and Prof. Janusz Bialek, for giving me the opportunity and guidance in carrying out my research. Dr. Chris Dent has been very supportive throughout my PhD study. He gives me not only tremendous and timely academic support, but also many wonderful opportunities including conferences and workshops. The joy and enthusiasm he has for his research and work is greatly motivational and encouraging for me. Dr. Amy Wilson has provided me with valuable support in statistics, innovative thinking and academic writing. Prof. Janusz Bialek helped me come up with the thesis topic and guided me over almost a year of my research project.

I would not be able to have the opportunity of pursuing a PhD degree without the support and funding provided by the Chinese Scholarship Council and Durham University.

I would like to express particular thanks to Dr. Dan Eager for valuable discussions during the project and for providing the code of his generation investment model. We also acknowledge discussions with Prof. Michael Goldstein and Dr. Alasdair Bruce. I would also like to thank my reviewers, Dr. Hongjian Sun and Dr. Behzad Kazemtabrizi for their time and effort in reviewing my annual reports.

I am very grateful to have great colleagues and friends. Without them, the work place wouldn't be full of joy, interesting discussions and lovely memories. I would also especially thank Dr. Terry Ho, for his tremendous encouragement, comfort, and consideration given to my life.

Lastly, I would like to thank my parents and my family members for all their love and care. No matter how far away I am from them, they have always been doing their best to support me both financially and mentally.

Contents

List of Figures

List of Tables

Nomenclature

Sets and Functions

$\mathcal{T}$	Set of generation investment years of interest, indexed by t.
$\mathcal{P}$	Set of past generation investment years.
$\mathcal{F}$	Set of future generation investment years.
G	Set of generation technologies, indexed by g.
I	Set of input variables of the simulator, indexed by i, j.
J	Subset of input variables of the simulator.
H	Set of hourly time periods in electricity markets, indexed by h.
C	Set of conventional generators, indexed by c, c'.
W	Set of wind generators, indexed by w.
B	Set of buses, indexed by b, b'.
L	Set of transmission lines, indexed by l, l'.
$\mathrm{C_b}, \mathrm{W_b}$	Set of conventional generators and wind generators at bus b, respectively.
Ref	Reference bus.
$\mathrm{Z_h}$	Set of reserve zones at time h, indexed by z_h.
$\mathbb{R}$	Set of real numbers.
$z_h(c)$	Set of conventional generators in reserve zone z_h.
$z_h(w)$	Set of wind generators in reserve zone z_h.
$z_h(b)$	Set of buses in reserve zone z_h.
ρ	Set of operating reserve types $\rho = \{su, sd, nsu, nsd\}$.
K	Set of segments for the curve of expected unserved energy v.s. upward operating reserve, indexed by k.

Q	Set of segments for the curve of expected wind curtailment v.s. downward operating reserve, indexed by q.
$I_{z'_h,z_h}$	Set of interface lines between zone z'_h and zone z_h.
$\mathcal{S}$	Set of post-contingency scenarios, indexed by s.
$f(\cdot), \tilde{f}(\cdot)$	Functions of the simulator and the emulator, respectively.
$h_i(\cdot)$	Functions of the main modules within the LTGI simulator.
$\mathcal{GP}(\cdot,\cdot)$	Gaussian process function.
δ	Model discrepancy function.
$\dot{P()}$	Probability density function of a continuous random variable.
$Pr(A)$	Probability of an event A.
E[·]	Expectation operator.
Cov[·]	Covariance function.
Var[·]	Variance operator.
$\mu(\cdot)$	Mean function.
$c(\cdot,\cdot)$	Correlation function.
$\sigma(\cdot)$	Standard deviation function.
max(·)	Maximum function.
min(·)	Minimum function.
$\phi_{z_h}(\cdot)$	Probability density function of system imbalances at zone z_h at hour h.

Parameters and Variables

N_y	Total number of hours in a year.
x	Vector of input variables.
u, θ, ω	Vector of control inputs, calibration parameters and forcing inputs, respectively.
u_{voll}	Energy price cap.
u_{cone}	Net cost of new entry (Net CONE).
u_{co_2}, ω_{gas}	Multipliers to the central scenario of the long-term projections of carbon and gas prices, respectively.

θ_{markup}	Markup cut-in point where the price markup approaches to zero.
θ_{VaR}	Assumed level of risk aversion of the investor.
$y^B_{g,t}, y^M_{g,t}, y^D_{g,t}$	Investment, mothballing and de-mothballing of generation capacities of type g at year t, respectively.
y^G_t	Installed generation capacity of type g in operation at year t.
y^L_t	Loss-of-load expectation at year t.
β	Hyperparameters (Unknown parameters) in the mean function of the Gaussian Process model.
λ	Scale hyperparameters in the covariance function of the Gaussian Process model.
γ	Correlation hyperparameters in the covariance function of the Gaussian Process model.
y_{obs}	Vector of historical observations of elements $y_{obs,t}$ over $\mathcal{P}$.
y^{LT}	Threshold of loss-of-load expectation.
F_t	Vector of annual fuel (uranium, coal, gas) and carbon prices at year t.
MR_t	Vector of the mean reverting level of annual fuel and carbon prices at year t.
χ_t	Vector of the speed of mean reversion of annual fuel and carbon prices at year t.
q_t	An estimate of annual fuel and carbon prices at year t.
κ	Vector of multipliers that are used to adjust the central DECC estimate.
$P_{markup,t}$	Random hourly price markup payment at year t.
ND_t	Random hourly net demand (demand minus wind generation) at year t.
$AG_{g,t}$	Random hourly available capacity of generation technology g subject to forced outages at year t.
CM_t	Random hourly capacity margin (hourly available thermal capacity minus hourly net demand) at year t.
$P_{e,t}$	Energy prices at year t.

$C_{g,t}$	Generation cost of generation type g at year t.
$RT_{g,t}$	Retirement of existing generators of type g at year t.
V_t	Net present value of an investment at year t.
$V_{VaR,t}$	Value at Risk (VaR) of V_t.
τ_f	The furthest simulation year ahead of the current decision year.
N_I	Total number of input variables.
D	Design points of chosen input variables.
N_b, N_l, N_z	Number of buses, transmission lines, reserve zones, respectively.
$\mathrm{B}_\mathrm{l}^+, \mathrm{B}_\mathrm{l}^-$	The "to" and "from" bus as with transmission line l.
$\mathbf{J}$	Power Transfer Distribution Factors matrix of $N_l \times N_b$.
K_f	Principal component basis vectors of elements $k_1, \cdots, k_{p_f}$.
d_i	The i-th basis function for model discrepancy.
ν_i	The j-th Gaussian Process model of the simulator.
ϑ_j	Weight of the j-th basis function for model discrepancy.
p_δ	Total number of basis functions for model discrepancy.
S_J	Measure of sensitivity to a subset of inputs x_J.
SV_J	Variance of the main effect of a subset of inputs x_J.
$\mathbf{I}$	Square identity matrix.
$\mathbf{K}$	Symmetric similarity matrix of $N_b \times N_b$.
$\mathbf{H}$	Power transfer distribution factors matrix of $N_l \times N_b$.
ω	Vector of line weights, with elements of ω_l.
$p_{c,h}$	Energy cleared for conventional generator c beyond the minimum generation level in day-ahead unit commitment (UC) at hour h.
$r_{c,h}^{\rho}$	Reserve of type ρ from conventional generator c in day-ahead UC at hour h.
$u_{c,h}$	Commitment variable of conventional generator c in day-ahead UC at hour h.
$v_{c,h}$	Start-up variable of conventional generator c in day-ahead UC at hour h.
$y_{c,h}$	Shut-down variable of conventional generator c in day-ahead UC at hour h.

$r^u_{z_h}, r^d_{z_h}$	Upward and downward operating reserve available in zone z_h, respectively.
$r^u_{z_h,z'_h}$	Upward reserve transferred from zone z_h to a different zone z'_h.
$v^{eue}_{z_h}, v^{wc}_{z_h}$	Decision variables representing the zonal values of expected unserved energy and expected curtailed wind corresponding to the level of available upward and downward operating reserves.
$\bar{r}_{z'_h,z_h}$	Sharing bound of upward reserve from zones z'_h to z_h.
$\hat{p}_{c,s}, \hat{p}^{wc}_{w,s}$	Energy cleared for conventional generator c and curtailment for wind generator w in optimal power flow at a contingency s.
$\hat{d}^{ls}_{b,s}, \hat{p}^{f}_{l,s}$	Curtailed load at bus b, and power flow through line l in optimal power flow at a contingency s.
$a_{c,s}$	Availability of conventional generator c in a post-contingency s.
$\hat{\eta}_{l,s}$	Lagrange multiplier of the transmission constraint of line l in the optimal power flow program at a post-contingency scenario s.
ΔM_{z_h}	System imbalances arising from net load (load minus wind generation) forecasting errors and generator outages in zone z_h.
FC_c, VC_c	Fixed and variable cost of generator c, respectively.
SU_c	Start-up cost of capacity provided by generator c.
V^{lol}	Value of loss of load.
V^{wc}	Value of curtailed wind.
$\underline{P}_c, \overline{P}_c$	Minimum and maximum load of generator c.
$\bar{F}_l$	Maximum transmission capacity of line l.
$R^{\tau_\rho}_c$	Ramp-up/-down rate of generator c for response time interval τ_ρ of reserve type ρ.
UT_c, DT_c	Minimum up and down time of conventional generator c.
$P^w_{w,h}, \hat{P}^w_{w,h}$	Day-ahead forecasted power and forecast uncertainty of wind generator w at hour h, respectively.
$LMP_{b,h}$	Locational market price at bus b at hour h.
$RMCP^{\rho}_{b,h}$	Reserve market clearing price of type ρ at bus b at hour h.
$D_{b,h}$	Day-ahead forecasted demand at bus b at hour h.

$R^{\rho,req}, R^{\rho,req}_{z_h}$	Required level of reserve type ρ in the entire system and in zone z_h, respectively.
$O^{\rho}, O^{\rho}_{z_h}$	Ratio of reserve type ρ to the total available upward/downward reserve in the entire system and in zone z_h, respectively.
$A^{eue}_{z_h,k}, B^{eue}_{z_h,k}$	Parameters of the linearized relationship between the value of expected unserved energy and upward reserve in zone z at hour t.
$A^{wc}_{z_h,q}, B^{wc}_{z_h,q}$	Parameters of the linearized relationship between the value of expected curtailed wind and downward reserve in zone z at hour t.
$\hat{D}_{b,s}, \hat{P}^{w}_{w,s}$	Simulated hourly load, wind generation in a post-contingency scenario s.

List of Abbreviations

ADP	Approximate dynamic programming.
ANEM	Australia's National Electricity Market.
BETTA	British Electricity Trading and Transmission Arrangements.
CAISO	California independent system operator.
CCGT	Combined-Cycle Gas Turbines.
CONE	Cost of new entry.
DECC	Department of Energy & Climate Change.
EDUC	Enhanced deterministic unit commitment.
EMR	Electricity market reform.
EU	European Union.
EU ETS	European Emissions Trading Scheme.
EUE	Expected unserved energy.
ERCOT	Electric Reliability Council of Texas.
GB	Great Britain.
GENCOs	Generating companies.
GP	Gaussian Process.
ISO	Independent system operator.
ISO-NE	Independent System Operator-New England.

LMP	Locational marginal price.
LTGI	Long-term generation investment.
LOLE	Loss-of-load expectation.
LOLP	Loss-of-load probability.
MCMC	Markov chain Monte Carlo.
MISO	Midwest independent system operator.
NETA	New Electricity and Trading Arrangements.
NYISO	New York Independent System Operator.
NPV	Net present value.
OCGT	Open-Cycle Gas Turbines.
Ofgem	Office of Gas and Electricity Markets.
ORDC	Operating reserve demand curve.
PJM	Pennsylvania-New Jersery-Maryland Interconnection.
PTDF	Power transfer distribution factors.
RMCP	Reserve market clearing price.
RMSE	Root mean-square error.
RTO	Regional transmission organization.
SA	Sensitivity analysis.
SERC	Southeastern Electric Reliability Corporation.
SRMC	Short-run marginal cost.
TSO	Transmission system operator.
UA	Uncertainty analysis.
UC	Unit commitment.
UK	United Kingdom.
US	United States.
VaR	Value-at-risk.
VG	Variable generation.
VOLL	Value of lost load.
Z-ORDC	Zonal operating reserve demand curve.
Z-VECW	Zonal value of expected curtailed wind.
Z-VEUE	Zonal value of expected unserved energy.

Chapter 1

Introduction

Electricity market reform is being implemented or considered in many power systems in order to meet their own needs. The electricity market (e.g., energy-only market, capacity market and ancillary service market) provides long-term incentives for efficient investment in generation capacity, by means of high energy prices and/or capacity prices and/or capacity payments. There is much debate on the effectiveness of different market designs on incentivizing the adequate amount and the right type of generation capacities. In the short-term, the availability of a flexible resource at times when it is needed will also be affected by the market design; this is an important issue when there is a large amount of variable generation (VG) whose variability and uncertainty require power system flexibility to manage. This chapter will first introduce the research background, and then identify the aims and original contributions of our research. In addition, thesis outline will be presented at the end of this chapter.

1.1 Background

Electricity market liberalisation has been implemented or considered in many countries. Meanwhile, many electricity markets are evolving in response to the needs of power systems [2]. For example, the electricity markets in the United Kingdom (UK) experienced the changes from the Pool (during years 1990 – 2000) to the New Electricity and Trading Arrangements (NETA) (during years 2001 – 2005), and

in year 2005 then the NETA was expanded to the British Electricity Trading and Transmission Arrangements (BETTA) to become the single GB electricity market of England, Wales, and Scotland [3, 4]. The most recent change of of electricity wholesale trading in GB has been the Electricity Market Reform (EMR) which has been running since 2009. The setup of an electricity market aims to facilitate the delivery of a secure, decarbonised and affordable supply of electricity. Delivering a secure electricity supply, which is the primary goal of power systems, not only requires sufficient reliable capacity to minimise the risk of supply shortages, but also calls for diverse energy sources, including renewables, nuclear, Carbon Capture Storage equipped plant, gas and demand side management [5]. Capacity adequacy has become a general concern due to a number of industry trends and regulatory barriers [2, 5–8], including plant closures under Large Combustion Plant Directive in the the European Union (EU), the integration of a large amount of variable generation with low capacity values [9, 10], the establishment of low price caps, and the carbon price floor that drives accelerated plant closures and also introduces regulatory uncertainty.

A reliable and economic power system relies on adequate and efficient planning and operations that happen on different time frames. In planning, it is necessary to account for up to $10 \sim 40$ years into the future, long enough to make and implement investment decisions. Good planning will equip the system with the adequate amount and the right type of resources that supply the energy, and energy prices in the market should usually be relatively stable. Price spikes caused by energy and reserve shortages in the energy market supplemented by additional payments (e.g., capacity payments, uplift payments) if applicable provide signals for investment. By contrast, the operational time frame often refers to day-ahead or real-time and sometimes to longer horizons (e.g., 48-hour). In most power systems, the operational problem is solved by an optimization problem that considers the cost-effective operation of the existing fleet, subject to technical and reliability constraints. In well-designed day-ahead or real-time markets, the pricing signals will provide the incentives for market participants to offer their energy and flexibility for use by the system operator.

Investments in generation are high risk. In the context of liberalised electricity markets as opposed to centralised ones with central planning, it is private generation companies who make investment decisions driven by the maximization of return on investment. The investment decision is often made based on price signal feedbacks and the imperfect foresight of cost estimates, revenue risks, and rate of return [11]. The risks exposed to investors range from policy (e.g., the price cap, CO_2 prices and renewable targets) and market (e.g., fuel cost, demand forecast and electricity price) risks, to technology (e.g., capital cost, construction costs/times and decommissioning costs) and finance (e.g., hurdle rate) risks [12] and they create uncertainty (*i.e.*, imperfect knowledge) in the financial returns of an investment. Various long-term generation investment (LTGI) models have been developed for predicting real-world generation projections and hence guiding investment decisions and the design of energy policy [13–20]. From the perspective of policymakers, who wish to adequately account for uncertainty around future generation projections and system reliability related to the real world, it becomes increasingly important to consider various sources of uncertainty existing in these models.

Market designs, particularly market incentives, play an important role in shaping generation investments and operational decisions under uncertainty [11, 21, 22]. As one of the mainstream electricity market designs, an energy-only market addresses long-term adequacy by rewarding generators for their actual generation at the price of energy. Examples of energy-only markets are Australia's National Electricity Market (ANEM), Alberta, Nordpool and Electric Reliability Council of Texas (ERCOT) [2]. An energy price cap is typically set to protect consumers from extremely high prices due to limited demand side participation. The ideal price cap would reflect the value of lost load (VOLL) which represents the theoretical value attributed to security of electricity supply (or preventing blackouts) by electricity consumers [23]. However, the VOLL are usually very high, which are often not politically attractive, so lower market price caps might be established. If no further intervention is taken, an energy price cap set below the putative scarcity value of energy will distort the perfect price signal for investment to some extent. This may lead to the "missing money" problem. Even without the establishment of an energy

price cap, prices in an energy-only market tend to be very volatile, exposing investors to high risks. Risk-adverse investors may find themselves reluctant to invest based on the possibility of price spikes.

The introduction of an increasing proportion of VG (e.g., wind generators) to future power systems is likely to exacerbate the "missing money" problem. On the one hand, due to the near-zero, zero, or negative bid-based costs of wind producers, electricity markets will experience increased price volatility and declined average price [24]. On the other hand, the limited predictability (i.e., forecasting errors) of VG often leads to the "out-of-merit" dispatch, which means that units are brought online to hold spinning reserve to protect against unexpected outages or rapid ramps in demand and wind generation. The "out-of-merit" dispatch results in not only a substantial reduction in market clearing price that would otherwise increase as demand increases, but also uncertain and reduced capacity factors, increased maintenance costs and shortened expected lifetimes of conventional generators [25, 26]. The profitability of many existing generation units will be under pressure, especially for marginal power plants of which the marginal cost (the additional cost incurred in production of one unit) is the highest among all operational ones, such as gas-fired power plants (Combined-Cycle Gas Turbines/CCGTs, Open-Cycle Gas Turbines/OCGTs), as experienced in some European countries [27].

Maintaining sufficient aggregate revenues in an energy-only market can mainly be achieved by increasing either the price cap or the frequency of scarcity periods [28]. An increase in the frequency of scarcity periods may resort to a change in the shape of the net load profile which is difficult to manage, or to an increase in market concentration which, however, is undesirable because market competition is threatened. In the real-world, the primary measure left to regulators for adjusting investment incentives is increasing the energy price cap. At present, the ANEM operators sets a market price cap of AUS$13,500/MWh, but the case study in [29] finds that the price cap may need to be lifted up to as high as AUS$80,000/MWh to support a 100% renewable market and to maintain the same reliability standard.

A tight energy price cap together with uncertainty around the frequency of scarcity pricing leads to insufficient revenues and increased generation risks, which

makes it unattractive to new entries. It is argued that the wholesale energy-only markets in many countries are sending too volatile or inadequate price signals to existing and new flexible generating recourses including flexible conventional generators and storage devices [17, 30–35]. Some of the main driving factors for concern over energy-only markets, as discussed in [36], include high financial risks in generation investment (referring to Section 2.3.1 for more details), limited demand side participation that creates price spikes [37], market power due to imperfect competition, and distorted energy prices due to current procedures for procurement and use of operating reserves.

Many markets have implemented or considered some hedging mechanisms (e.g., contracts market [28], operating reserve demand curves (ORDCs) [38,39]) to manage price risks and some complementary market designs including capacity payments, capacity markets (installed capacity requirement [40] and capacity demand curve [20]) and capacity subscription [36]. These market incentives are designed to compensate the contribution of capacity and operational flexibility to the system, so as to enhance the robustness of the power system against various sources of uncertainty. In ANEM's energy-only market, market participants may rely on option contracts to hedge against occasional price spikes[1]. Retailers are allowed to participate in a market for call options which provide them with a fixed maximum price for some contracted volume of capacity over a contracted future period in order to cover the majority of their anticipated load demand. On June 1, 2014, ERCOT implemented an ORDC to improve the determination of prices of wholesale energy and operating reserve in scarcity conditions[2]. In the UK, the first capacity market auction was run in 2014 for delivery of capacity from the winter of 2018/19 [23]. In the United States (US), capacity markets are more common and with different design details (cleared auction price, contract length, etc.), such as in Independent System Operator-New England (ISO-NE) and Pennsylvania-New Jersery-Maryland Inter-

[1]http://www.asx.com.au/documents/products/ASX_AustralianElectricityFuturesandOptions_ContractSpecifications_July2015.pdf

[2]https://business.directenergy.com/blog/2014/june/ercot-implements-operating-reserve-demand-curve

connection (PJM), New York Independent System Operator (NYISO) [41]. Ireland, Spain, Ontario are also examples of energy markets with capacity mechanisms. Alternative policies such as the forward and bilateral energy contracts are designed to recover some part of investment cost [3] and financial option contracts can be employed for risk management purposes [33]. For example, in the UK, carbon price support and feed-in tariffs and an emissions performance standard are proposed for supporting low carbon generation [5].

With the growing integration of renewables, the increasing need for flexible capacity requires new market rules and productions. Several options are feasible to improve system flexibility, including investing dispatchable flexible and back-up generation, demand-side participation and storage, interconnections and market tools (e.g. market coupling or capacity remuneration mechanisms) [42]. Generation flexibility, as the most affordable way to provide flexibility [43] is defined by California Independent System Operator (CAISO) as the resource's ramping speed, ability to sustain a ramp, ability to change ramp directions, ability to reduce output and not encounter emission limitations, start time, and ability to cycle on and off frequency [44].

Regarding incentives for generation flexibility in system operations, some mechanisms are in place including efficient centralised scheduling and pricing, 5-min settlements, ancillary service markets, make-whole payments, and day-ahead profit guarantees [2]. A new flexible ramping market which proposes spot market (five minute interval) and forward procurement (integrated day-ahead) products has emerged in California. The idea of incorporating specific requirements for flexible unit operating characteristics in the year-ahead resource adequacy requirements has also been proposed by CAISO, which may eventually develop into five year forward capacity procurement process [45]. More research is required to evaluate the potential value of flexibility to help inform the proper design of markets and the pricing of various flexible products.

1.2 Aims of research

Overall, the research explores the role that a market design (e.g., market incentives or mechanisms) plays in decision-making under uncertainty for both long-term generation investment and short-term (day-ahead and real-time) power system operation.

Regarding the impacts of market incentives on long-term generation investments, it is desirable to provide policymakers with plausible projections of future generation capacity or system reliability under different market designs (e.g., energy-only and capacity market). To understand both uncertainty around these projections and the ways they relate to the real-world, this thesis will carry out calibration (the use of observed data to match a model with the real process), uncertainty analysis (UA) (assessing uncertainty in variables or measurements and exploring the propagation of uncertainty) and sensitivity analysis (SA) (studying how varying values of an independent variable affect a particular dependent variable) of a computationally intensive LTGI model with careful management of various sources of uncertainty.

Regarding the impacts of market incentives on short-term power system operations, this thesis explores how to incorporate the potential value of flexibility in the deterministic unit commitment model to help inform the proper design of operating reserve markets, that is, the pricing of operating reserve products. An alternative is to represent the timing and the locational values of operating reserves by zonal operating reserve demand curves (Z-ORDCs) in order to reward to improve reserve deliverability and adequate incentive for flexible generating resources in a transmission network. It is important to investigate the effectiveness of different reserve policies (e.g., single, seasonal or dynamic reserve zones) that have been proposed to emphasize the locational and timing value of operating reserves.

1.3 Original contributions

The research has a number of original contributions:

1) Use of Bayesian emulation that is based on a Gaussian Process model as an approximation to a LTGI model to systematically manage three major sources

of uncertainty (in Section 2.3). The sources of uncertainty include input uncertainty arising from unknown precise values for model inputs, structural uncertainty due to imperfect science that is used to approximate the underlying true system, and functional uncertainty representing unknown functions when model evaluations take a long time. This is the first time that such emulation techniques have been used to manage these uncertainties associated with generation investments (in Section 4.2.1).

2) Presenting a statistical approach for the calibration of LTGI models, based on a Bayesian update of prior judgments. Given some historical observations of the output of interest, one can infer from observations improved knowledge of uncertain model parameters and the imperfect model structure (in Section 4.2.3);

3) Quantifying a plausible range of model outputs that is consistent with the available knowledge (both historical observations and expert knowledge) based on a Bayesian update of prior judgments, and demonstrating that a failure to account for uncertainties may result in misleading results to investors and policymakers (in Section 4.3);

4) Studying the robustness of electricity market designs (energy-only and capacity markets) against uncertainty through an UA of system reliability metrics (e.g., loss-of-load expectation/LOLE) in a quantitative manner. Also, demonstrating how to determine regions of market design parameter space with a high probability of maintaining the system reliability target (in Section 4.4).

5) Performing an efficient and comprehensive SA of a LTGI model and identifying the most important input or groups of inputs (in Section 4.5).

6) Proposing an enhanced deterministic unit commitment (EDUC) model incorporating dynamic zonal ORDCs for studying the impact of incentives for generation flexibility on system or market performances. (in Chapter 5).

1.4 Thesis outline

The overall structure of the thesis is described below.

Chapter 2 firstly provides a literature review of research results on generation investments in electricity markets, covering the economic, political and technical aspects. Since various sources of uncertainty are involved in the long-term planning period, the key uncertainties that define the quality of generation projections are identified.

Chapter 3 presents a model specification of a LTGI model and then uses it to make projections of operational thermal capacities and LOLEs at each decision year during the planning horizon of interest. LTGI models are generally computationally intensive and so only a limited number of simulations can be carried out. A statistical methodology based on Bayesian emulation for addressing the computational challenge of model evaluations and enabling efficient calibration, UA and SA is presented in this chapter.

Chapter 4 demonstrates a case study relating to the GB power system planning. First, a formal calibration is carried out on the LTGI model that runs through a historical planning horizon. Next, calibration results (the updated information of model parameters and model discrepancy) are applied to future projections of thermal capacities and LOLEs. Then, a robustness index is presented and quantified through an UA of the maximum LOLE over the future planning period. Last, a comprehensive SA is carried out and results of SA imply ways of improving the market robustness most efficiently by focusing on the most important model inputs.

Chapter 5 proposes an enhanced deterministic unit commitment model incorporating hourly updated Z-ORDCs. Reserve zones are defined by the approach of spectral clustering. A case study on system performances using the RTS-73 test system is given. Comparisons are made between the choices of reserve policies (e.g., single, seasonal or dynamic zones) and of reserve zone partitioning methods.

In Chapter 6, conclusions are drawn and suggestions for future work are presented.

Chapter 2

Impact of market incentives on system adequacy

This chapter first gives a review of evolving electricity market designs and system adequacy metrics with the increasing integration of VG. Different market designs will have distinct influences on the bidding behaviour of GENCOs and consumers, on the dynamic behaviour of market prices and investment incentives, and hence on the investment decisions. For guiding the decision-making in energy policies and generation investments or the evaluations of the social effects of some interventions, various types of LTGI models have been developed in academic and industrial areas.

Then, this chapter describes various sources of uncertainty involved in the LTGI models, which poses great challenges in model computations, model applications and the interpretation of model outputs. Assessing uncertainty is necessary to make sense of model results, which requires tasks of calibration, predictions, uncertainty analysis (UA) and sensitivity analysis (SA). Existing work on dealing with these tasks is not comprehensive enough, due to huge burden of model computations and inadequate consideration of uncertainties. In order to fill in the research gap, we will adopt Bayesian emulation in Chapter 3 which can enable these computational tasks by approximating the full model with an emulator while quantifying all major sources of uncertainties.

2.1 Electricity markets in transition

Both research problems – the LTGI problem in the planning framework and the unit commitment (UC) problem in the operational framework, will be studied under the context of liberalised electricity markets. The design of electricity markets evolves with the change in the energy structure (the proportions of different energy suppliers including coal-fired plants, gas-fired plants and renewables) and with the need for new investments. Accordingly, some changes have been seen in the roles that market participants play, in market competitiveness and in the methodology of system adequacy metrics used in practice. Our main focus is to explore the effectiveness and robustness of the mainstream electricity market models in sending the right price signals and investment incentives for maintaining system flexibility and adequacy.

2.1.1 Market participants

The problems of planning and operations are relevant to the interest of private investors, demand customers, policymakers, system operators and other market participants. Different participants play distinct roles in the wholesale electricity market (spot market) where electricity is traded through bids to buy and through offers to sell.

On the supply side, the privatisation and liberalization of electricity markets make the decision of investing in generation capacity transfer from generally state-owned monopolies to competitive and private investors or GENCOs who sell their electricity in the spot markets [46]. In a market-based system, costs cannot be automatically passed on to consumers [46], and investors are exposed to highly volatile market prices.

On the demand side, customers (e.g., retail suppliers) and companies (e.g., large industrial companies) consume electricity they need. In liberalised electricity markets, very large consumers may be exposed to volatile electricity prices and they may wish to participate in the demand management program by modifying their demand pattern in response to the market prices. Long-term price elasticity of demand as a

whole that accounts for the programs of reducing energy consumption over the full period of investment has been considered in some LTGI models, as in [18, 47–49].

Although policymakers cannot make decisions on investing a new plant in liberalized electricity markets, they play a role in encouraging a diverse, low-carbon and flexible generation mix by providing incentives [46]. For shaping the future structure of electricity supply and system reliability, it is important to investigate the anticipated investment behaviour in the electricity market driven by different policy designs.

A system operator takes the main responsibility of managing the electricity network to ensure the security, reliability and efficiency of supply of electricity. The system operator undertakes the task of real-time dispatch of generation, managing security, planning incentives and contracts signed with generation firms for maintaining generation-demand balance and system security during future trading periods. The entity who plays the role of a system operator can be owned by the transmission grid company, or may be fully independent according to the design of electricity markets, and so it may be named differently in different power systems. For example, in Europe, a transmission system operator (TSO) administers the transmission grid on a national or regional basis. In the United States (US), a regional transmission organization (RTO) or an independent system operator (ISO) performs similar functions as the European TSO.

2.1.2 Overview of market designs

Overall, three major electricity market models around the world are discussed here and they are categorized into centralised and liberalised (competitive) markets [2,41]. The first is the centralised vertically integrated structure where most or all assets (i.e., generation, transmission and distribution) within a certain geographical area are owned and operated by a single entity. The second is the competitive energy-only market design where resources are only paid for the energy and ancillary services (optional)[1]. The third is the competitive energy plus capacity market design where an

[1]https://www.hks.harvard.edu/hepg/Papers/MISO_Resource_Adequacy_112305.pdf

additional revenue mechanism is available to reward generators for their installed or available capacity, in addition to revenues from selling energy and ancillary services. The key feature of liberalised electricity markets is that the potentially competitive functions of generation and retail are separated from transmission and distribution; and a wholesale electricity market and a retail electricity market are established.

Centralised vertically integrated structure

During early days of the electric power industry, a vertically integrated structure was favoured by governments. Under a vertically integrated structure, no competition exists between different utilities and ideally the energy price is set by the highest short-run marginal cost/SRMC of running generators (the cost imposed on electricity suppliers by extracting one more unit of supply). The centralised generation capacity expansion takes the objective of minimizing the total social cost (including investment cost, operational cost, reliability cost, environmental cost, etc.) or maximizing the social welfare, as opposed to the process of liberalised generation investments where private GENCOs make decisions with an objective of maximizing the expected profit of a new investment.

The US took the traditional model of the vertically integrated electricity utility for decades. Later, the US adopted the power pools together with interconnections over a larger network (electrical grid) instead of a single utility before moving to liberalisation and ISOs and RTOs in the last decade of the 20th century. The transition from a vertically integrated structure to a deregulated (competitive) structure (e.g, an energy-only or a energy plus capacity design) is primarily driven by the need for a more efficient supply of electricity. A healthy competitive market might also help attract investment in new technologies given right incentives.

In the late 1980s, UK began the preparations for privatisation with the generating companies. Reforms in England and Wales, Scotland, and Northern Ireland were implemented in 1990s and then the liberalisation of different parts of the industry such as British Energy and National Grid happened at different times. The liberalization in the UK was followed by several other Commonwealth countries including Australia and New Zealand, and regional markets such as Alberta. In many rapidly

developing countries, previous vertically integrated utilities (e.g., China's power sector [50] and the state electricity boards in India) are now being restructured.

In the 21st century, market designs are rapidly evolving to address various challenges posed by VG to power system adequacy, operational efficiency and ancillary services. The electricity market reform (EMR) was recently implemented by the UK government for purposes of ensuring security of electricity supply while achieving decarbonisation goals affordably. The GB EMR has introduced support mechanisms including the Feed-in Tariffs with Contract for Differences, Carbon Price Floor and Emissions Performance Standard to support investment in low-carbon generation, as well as the capacity market for supporting security of supply.

Energy-only market

As a central transaction platform in power markets, an energy-only market deals with day-ahead scheduling and real-time dispatch for energy and ancillary services (mainly referring to operating reserve) coordinated by system operators.

In a perfect energy-only market, generators offer their SRMCs as price takers who cannot change market prices. The market-clearing (spot) energy price is determined by the intersection of the generation supply curve and the electricity demand curve (the equilibrium solution). All generators are remunerated at the same hourly market-clearing energy price, set by the SRMC of the marginal plant needed to meet demand, which is the basic principle of liberalized wholesale power markets. The market-clearing energy price reflects the VOLL when there is a shortfall in electricity supply. A perfect market brings sufficient income to generator investors [3] under the assumption that investors are risk neutral, which, however, is not the case in practical world. It is investors' risk aversion that raises concern over simple arguments that energy only markets can deliver security of supply.

A perfect energy-only market does not exist in practical world. On the one hand, the VOLL is very difficult to estimate and highly uncertain due to the inclusion of the consumers' costs and the wider social costs from blackouts [51]. In real-world power systems, regulators usually set the energy price cap lower than the real VOLL in times of scarcity to avoid extremely high market prices [46], which may result

in insufficient revenues for generators, especially the marginal plants who rely on price spikes to recover their costs. Scarcity refers to times when the supply may be insufficient or too costly to deploy to meet the high demand. In order to provide additional incentives for investments in flexible resources, scarcity pricing is often implemented in an energy-only market [2] and it will be discussed in more detail in the following section. On the other hand, it is difficult to administer a perfectly competitive market environment. Market power is exercised when a GENCO is capable of profiting from increasing prices (price markups) above competitive levels (uniform market-clearing prices) for a sustained period of time. The price markup in the market can be monitored as the difference between the spot energy price and the SRMC [52]. Strategic interactions among different GENCOs are also seen both on short-term operational decisions (i.e., strategic bidding in the spot market) and on long-term investment decisions (i.e., strategic investments in an oligopoly situation).

Scarcity pricing

As explained in Section 1.1, increased penetrations of VG reduce the average electricity prices but increase the need for flexible capabilities such as fast ramping, frequent on-off cycling and balancing ORs, and hence increase the operational cost of thermal generators [53, 54]. To ensure capacity adequacy and revenue sufficiency, energy-only market is often combined with scarcity pricing, ancillary service scarcity pricing, or emergency demand-response pricing.

Scarcity pricing refers to arrangements to modify prices in the wholesale electricity market (spot market) when the system operator reduces demand through administrative actions such as emergency load shedding [55]. Scarcity pricing introduces a price floor and a price cap to the spot market, which improves revenue certainty for providers of last resort resources (including generation and demand response).

An alternative to the floor and cap mechanism is ancillary service scarcity pricing. Ancillary service markets include those for provision of spinning contingency reserve, non-spinning contingency reserve and regulating reserve at both short-term (spot)

and longer-term (bilateral) markets. Additional operating reserves and normal balancing reserve such as regulating reserve are needed for balancing the increasing penetration of VG. Hence, the ancillary service clearing prices for those services will be increased, and market participants will face greater uncertainty in ancillary service demands and prices. With effectively and well designed ancillary service markets, the impacts of VG on resource adequacy and revenue sufficiency can be mitigated. A well tested idea is applying ORDCs during scarcity [39] which has been found in operation in some RTOs in the US, such as NYISO, ISONE, MISO and ERCOT. The ORDC is designed to introduce an additional price component to the energy price and to ensure that the resultant energy price reflects the increasing value of electricity as the possibility of rotating outages increases.

Capacity mechanisms

Capacity market and some form of capacity payments are two main capacity mechanisms designed to support the revenue adequacy of electricity suppliers. Capacity market is a bilateral contract and forward market where participants purchase or sell a volume of capacity products which meet reliability requirements[2]. The capacity payments mechanism is a fixed revenue system of payment for participants offering generation capacity. Capacity market is more commonly adopted than the capacity payments mechanism, because it is generally reckoned to be hard to predict the capacity outcome triggered by a given payment.

New designs of forward capacity obligations and associated auction mechanisms, such as the determination of capacity prices and reserve margins, are proposed in [33] to determine capacity prices. The use of a reliability option contract, a new variant on reserve requirements, is gaining support in the EU [56]. In a reliability option contract, a strike price is established via a capacity auction in the day-ahead market. The establishment of the strike price mitigates market power during scarcity because generators must pay the difference between the market-clearing energy price and the strike price and hence do not benefit from price manipulation.

[2]https://www.hks.harvard.edu/hepg/Papers/California.PUC.Capacity.Markets.White.Paper.pdf

2.1.3 A review on the effectiveness study of market designs

Some modelling and simulation studies have been conducted to examine the effectiveness of capacity mechanisms on improving capacity adequacy in both industrial and academic world.

Three market designs, namely, the capacity payment, the competitive capacity market and their joint combination are examined in [18], based upon a long-term system dynamics model in an energy-only market. A variable capacity payment mechanism is modelled and studied in [14], which appears to help stabilize the market price and reserve margin, and may bring extra credits for consumers in the long run. The interactions between neighboring power systems are highlighted in [57], in which the effects of several market designs (e.g, energy-only markets, price-capped market with a forward capacity market, price-capped market with a capacity mechanism) on incentivising capacity adequacy are tested on two simplified interconnected markets. In [48], the investment decisions of private firms are represented by a two-stage game. The model results in the Spanish power system case study show that the two examined regulatory mechanisms, namely, capacity payment and price adder are ineffective and costly in delivering the desired level of capacity in an oligopolistic context.

In the GB EMR process, a suite of models has been developed including a fully integrated power market model, the dynamic dispatch model [58] provided by Department of Energy & Climate Change (DECC), and network modules designed and built by UK National Grid. The dynamic dispatch model covers the GB power market over the medium to long term, which allows the impact of policies on the investment and dispatch decisions to be analyzed, provided with data, assumptions and scenarios. In the dynamic dispatch model, a forward capacity market based on a capacity demand curve (see Section 3.2.1) will be triggered (from 2018/19) if the amount of generation capacity in an energy market where plant is no longer profitable is not enough to meet the security of supply reliability target [58]. To draw some confident conclusions on the effect of the forward capacity market on the GB power system adequacy based on model simulations, it is necessary to calibrate some model parameters that represent our assumptions of the real-world.

The capacity market proposal of PJM is modelled in [20], where the effects of assumptions concerning investor behaviour on the effectiveness of different capacity demand curves (capacity payments as a function of reserve margin) are assessed. The simulation results from the present PJM installed capacity system show that although there is significant uncertainty regarding the future effects of capacity mechanisms on financial and adequacy consequences (generator profits, consumer payments and reserve margins), the use of capacity demand curves reduces risk exposure of generators and consumers, lowers costs to consumers and increases new investments. Again, these results would have been more convincing if model validation or calibration work was done.

In real-world analysis relying on computer models, key issues are how well these models represent the real world and how much uncertainty is involved in these model results. These concerns are often addressed by model calibration and uncertainty quantification. More details on the various sources of uncertainty and the ways of dealing with uncertainty will be presented in Sections 2.3, 2.5.

2.1.4 System adequacy/reliability metrics

Metrics used by system operators to evaluate whether a current or future power system meets adequacy or reliability requirements are evolving with the energy structure of power systems [59]. These metrics include a fixed gross capacity margin or derated capacity margin, probabilistic metrics (e.g., loss-of-load probability - LOLP, loss-of-load expectation - LOLE and expected unserved energy - EUE), or economic standards.

The gross capacity margin is normally expressed as the percentage of the level by which available nameplate or gross generation capacity exceeds the maximum expected level of demand (e.g., the peak demand) to the peak demand. For example, standards of a planning capacity margin of 10% to 20% are used in some regional power systems such as CAISO and SERC/South Carolina Electric & Gas Company (SCE&G) [60].

The derated capacity margin attempts to account in a margin metric for the different availability statistics of different plants in a power system [61]. Derated

conventional capacity is usually defined as expected (in the mathematical sense) available conventional capacity for generation during the period being examined after accounting for unit-specific seasonal availability ratings. For example, the availability probability of a CCGT plant is 85% during the peak time according to the Office of Gas and Electricity Markets/Ofgem's Electricity Capacity Assessment [62]. The derated capacity, also called the capacity value of VG is handled differently by capacity resource planners from conventional generation because of its intermittent nature. Different metrics may be used to quantify the capacity value of an additional generator of VG, such as effective load-carrying capability (the additional load that could be served by the additional generation without increasing the adequacy risk level) or equivalent firm capacity (the completely firm generating capacity that replaces the additional generation and gives the same risk level) [63].

Derated margin is a derived quantity which acts as a proxy for the results of a full risk calculation, such as LOLE. However, the derated capacity margin does not indicate how likely it is to have a shortfall. To assess risk at all, probabilistic metrics - LOLP, LOLE and EUE have been commonly used as the reliability standard in many markets. Ireland, Spain, GB and some regional markets in the US such as MISO and PJM have all opted to conduct their adequacy assessment based on a probabilistic risk metric such as LOLE [60–62, 64]. LOLP is defined as the probability that generation is insufficient to meet demand at a snapshot in time [65]. LOLE is the expected (in the mathematical sense) number of periods (e.g. half hours, hours, daily peaks) of periods of shortfall within the time window under study [66]. EUE is the expected (in the mathematical sense) amount of energy that is unserved by electricity suppliers within the time window under study.

The early history of adequacy assessment using LOLP is reviewed in [67]. Currently in the US, such as in PJM, MISO, ISO-NE markets, a generally accepted generating capacity adequacy criterion is an LOLP of one day in ten years or an LOLE of 0.1 daily peaks per year. The 0.1 LOLE is specified at a time resolution of days (typically the expected number of daily peaks on which there is a shortfall) [60]. In GB, the reliability standard is set by the government at 3 hours LOLE per year at an hourly resolution [68]. Due to the different time resolutions used for

defining the LOLE metric, the one day LOLE in the US is not the same as 24 hours LOLE in GB. It should be noted that the same standard of LOLE does not result in the same level of reliability in different power systems, because the LOLE study varies across a range of different modelling assumptions (e.g., accounting for load forecast uncertainty or not) and reliability modelling tools (e.g., different models of load forecast uncertainty).

Other markets choose both reliability and economic metrics (e.g., a target of minimizing customer costs while maintaining a specified capacity margin in Southeastern Electric Reliability Corporation (SERC)/Southern Company), and a combination of reliability metrics (i.e., the capacity margin and LOLE in the Canadian Maritimes, the LOLP and EUE in Northwest/Bonneville Power Administration) for system adequacy assessment [60].

2.2 Overview of long-term generation investment models

Complex LTGI models have been developed as a decision-support tool of planning energy systems and designing public energy policy. LTGI models differ with respect to their adopted mathematical forms and their underlying economic assumptions, i.e., the mechanism of electricity markets and the behaviour of how each firm makes its investment decisions. Prior to providing a review of basic LTGI model types, the various techniques for developing two critical sub-models, namely the modelling of spot market prices and the economic assessment of an investment project are described.

2.2.1 Modelling of spot energy prices

As a signal for generation dispatch and investment in liberalised market, the spot energy price and its formation deserve a careful study. A well-designed and well-functioning market is supposed to generate prices that do not only compensate for operation costs, but also allow generators to recover the investment costs with low

profitability risks [21]. This long-term consideration should help to discern which of the pricing approaches is more appropriate. However, the full long-run incentive effects of these pricing rules have not been well understood.

Generally, there are two types of short-term spot energy price simulation models. One is completely based on data analysis techniques and historical electricity price data, such as normal distributions of the electricity day-ahead price used in [69–71], time series analysis of historical electricity prices in [72–74], and artificial intelligence techniques which map the relationship between the input parameters (e.g., load levels, time periods) and electricity prices [75,76]. The other one is based on a bottom-up description of the power system, especially the supply-demand balance and the strategic behaviour of GENCOs [77–79]. The bottom-up electricity price model is capable of incorporating the temporal fluctuations and short-term uncertainties in supply and demand, and investigating the effects of varying installed capacity and demand level on future prices. However, computational efficiency becomes one of the major challenges when one considers to calculate the high-resolution (hourly or even half-hourly) spot energy price with the inclusion of generator-level inter-temporal operational constraints during a long-term planning period, which can account for 30 or more years.

A simple bottom-up electricity price model is based on aggregating the demand and generation at all buses in the transmission network. The conventional generators of the same technology can be aggregated and stacked in a merit order. Merit order, as a simplified dispatch engine, provides a ranking of available electrical generation in ascending order of their SRMCs of production. Then, the uniform market-clearing price equals the largest SRMC of all generators online. In GB EMR process, the dynamic dispatch model uses aggregated cost information for each technology, makes dispatch decisions according to the generation merit order and calculates the spot energy price as the SRMC of the marginal plant plus a mark up. In [49, 77, 80], a probabilistic dispatch algorithm based on the Baleriaux-Booth method [81] is used to determine the spot market price while characterizing the short- and long-term uncertainties in demand by discrete probability distributions and modelling forced outages and scheduled maintenance in thermal generators.

In [13], a computationally efficient probabilistic costing method is employed based on a mix of normal distribution approximation to the distribution of net demand as well as a normal distribution approximation to available conventional generation capacity that is subject to forced outages. Since investment cost recovery requires a sufficiently long time horizon beyond the forward market, GENCOs should gain some knowledge of current and future market prices for scheduling and investment purposes. A forward-looking simulation is conducted in [4, 13, 82] to predict future market prices during the dynamic process of investment.

However, the perfect economic merit order is always distorted to some extent due to technical limitations such as minimum stable generation and minimum up/down times. In non-sequential modelling methods, the chronological characteristics of demand level, generation output and wind availability are removed from their probability distributions, and hence the inter-period dynamics and constraints at the unit level cannot be examined. With an increasing wind penetration, the issues of overgeneration, lack of reserve, and high start costs, resulting from system flexibility limitations may become significant in assessing the system reliability and the economic performance of a new investment. Unit construction and commitment is combined with economic dispatch in [83, 84] to produce prices that serve as the control signal for new investments.

The economic dispatch models have well-defined solution for linear or uniform market clearing prices and generators being paid the system marginal price can recover their full operational costs, but the more general UC models may produce no analogous energy prices that fully recover generators' operational costs including start-up/shut-down costs [38]. This is the context where uplift or make-whole payments arise. In many imperfect energy-only markets, uplift payments are added on top of the linear or uniform market clearing prices, which is the so-called non-linear pricing rules, in order to ensure generators are sufficiently compensated for their full operational costs. The analysis in [21] investigates the long-term impact of two major groups (i.e., linear and non-linear) of pricing rules on the optimal energy mix under an energy-only market. It is found that the need for uplift payments is increasingly prominent for a power system with high penetration of renewable energy

sources.

2.2.2 Economic evaluation of a new investment

Investment decisions largely depend on the profitability of a projected investment or project. There exist different aspects of economic assessment, including the NPV analysis, real options theory and utility theory.

The net present value (NPV) analysis/the discounted cash flow method is often used in LTGI models, as in [4,47,57,85]. The NPV provides an indicator of how much value an investment adds to a GENCO, by accounting for the difference between the present (discounted) value of revenues (cash inflows) and the present value of costs (cash outflows). The discount rate used for discounting future cash flows to the present value is a key parameter of this analysis. Some common rates are a firm's weighted average cost of capital, the hurdle rate (the minimum acceptable rate of return), and variable discount rates (varying discount rates with the riskiness of investments). Generally, if the NPV is positive the project is worthwhile to be invested in, while the project should be turned down if the NPV is negative. If the NPV is zero the investor is paying exactly what the project is worth.

In a risky investment environment, there is uncertainty about future cash flows and the expected (in a mathematical sense) value is often estimated. To reflect the investors' risk aversion, the NPV method can be used together with a risk-adjusted discount rate as in [49] or the risk functions of plants' profit, such as value-at-risk (VaR) [86] and conditional value-at-risk [87]. Often, the VaR calculation relies on a Monte-Carlo (MC) simulation-based approach for a complete economic evaluation of an investment under a set of scenarios, as conducted in [13,46,88]. Increasing the number of scenarios generally reduces the sampling errors and hence improves the robustness of the estimation of the true risk function, while the computational time will be unavoidably increased.

Real options theory adopted in [77, 78, 89, 90] is an alternative method of assessing investments under uncertainty. Real option theory considers subsequent decisions that can modify the project once it is undertaken, while the NPV approach does not allow for this flexibility of the project and consequently underval-

ues its benefits. Regarding the LTGI problem, the NPV approach only looks at one possibility (i.e., investing in a new generation capacity or not) in isolation. By contrast, real options theory assesses a full range of possibilities, allowing a flexible and dynamic representation of investment timing and different options (e.g. delaying investments, abandoning the project, or changes in the cost effectiveness of the investment project) [91].

Utility theory models risk preferences in profitability assessment by quantifying the utility function of a decision marker in the objective function. As used in [92,93], risk-averse firms maximize the concave utility function of profit with an exponential form, while a risk-neutral firm maximizes a linear utility function of profit.

2.2.3 Basic model types

Regardless of the market price models and the economic evaluation methods, the modelling techniques vary among different LTGI models. There is always a trade-off between detail (e.g., short-term spot energy price models) and scope (e.g., research aims, assumptions).

Linear programming

Linear programming is one of the simplest modelling approaches applied to the LTGI problem. In [94], a multi-period and multi-criteria model based on deterministic linear programming and the analytical hierarchy process for the generation expansion problem is proposed. Linear portfolio optimisation models are developed for investigating the impact of VG in generation resource planning in [95, 96]. In a linear programming formulation, simplifications must be made on the commitment configurations as well as the ramping and reserve constraints, which could result in overly restricted cycling (i.e., start-up and shut-down) of conventional generators.

Mixed Integer programming

The LTGI problem includes discrete decisions (e.g., the number of new power plants of each technology) and hence is often formulated as a mixed integer programming

problem. Growing attention has been paid to incorporating UC constraints into a such a mixed integer programming program.

A detailed integer block UC formulation within a capacity planing model is proposed in [97]. The idea is to group the generation with similar technical and cost characteristics and transforming the number of plants on-line to the commitment state, so the dimensionality of the UC program can be largely reduced. Ref. [98] further compares the reduction of computation time and the loss of accuracy under different levels of aggregation. Another simplification technique of UC employed in [83] is to reduce the size of the search space by adding physical constraints and additional heuristic constraints – priority ordering among small flexible units with similar technical and cost characteristics.

However, even with simplified UC programs, the computation time can be drastically increased within a multistage planning horizon. One group of approaches to speeding up a multistage program is adopting a rolling-horizon setting. Essentially, the multistage model is solved by solving successively several submodels with shorter and shorter horizons, e.g., [99]. A second group of approaches to accelerating the computational speed is to carefully select a limited number of representative days or weeks to represent possible system states rather than to simulate many years of hourly operations. A vast majority of existing approaches to selecting a representative relies on either simple heuristics or clustering algorithms. In [83], simple heuristics are used to select four representative weeks characterizing the average of the net load profiles of four seasons and one or more weeks representing extreme conditions. A collection of independent days are randomly simulated from the historic load and renewables data using binning strategy in [100], where the sampling approach allows for parallelization to improve computational efficiency. The heuristic approaches lack a consistent criterion for assessing the validity of the approximation. In comparison, clustering algorithms are more advanced by dividing similar load, wind speed and/or solar patterns into clusters with regard to some distance functions and picking one representative period from each of resulting clusters [101]. Scenario reduction techniques, such as the fast-backward method, are also used to select representative periods, e.g., [102]. Another advanced approach is to optimize

the selection of representative periods using a user-defined criterion. Ref. [103] introduces a week selection method of minimising the least square error between the approximate and the actual load or net load (demand minus wind generation) duration curve. All these selection approaches introduce structural uncertainty to be explained in Section 2.3, because there is always a mismatch between the selected days and the real-world operational conditions.

Both a mixed integer program and a linear program are deterministic and they omit a key feature of the real investment problem that involves various sources of uncertainty (see Section 2.3), and these deterministic programs are unlikely to produce results relevant to the real world.

Stochastic programming

Stochastic programming provides a powerful tool for solving the problem in terms of making optimal and sequential decisions on dispatch and investment in time along with a precise description of uncertainty and the folding of uncertainty over time [104]. In the formulation of stochastic programming, uncertainties are very often represented as scenario trees that may be constructed from some expert knowledge or historical observations.

In [105], a two-stage stochastic programming model for investments in thermal generation capacity is developed for studying the impact of demand response at high wind penetration levels. In the objective function of cost minimization, operational costs on an annual basis are approximated as those calculated from several weeks representing high, medium and low load seasons, respectively. In [106], a multistage stochastic mean-variance optimisation model is developed for portfolio optimisation in deregulated electricity markets. A multistage framework enables the modelling of investment decisions at multiple future time points, using the information up to that time point. To gain computational tractability, the mutistage stochastic program is simplified through the stage-aggregation (aggregating the decision stages) and linear decision rules.

The quality of investment decisions largely depend on the selected scenarios that cannot fully represent the uncertain operational conditions. Besides, the solution

time of the optimisation model scale with the size of the scenario tree that expands exponentially with the number of uncertain parameters and decision stages [106].

Dynamic programming

Dynamic programming has the advantage of handling stochastic and multistage planning or investment models while linking different time periods together so that time-linking constraints such as those related to hydro and storage may be considered [107]. In the dynamic programming formulation, the number of generating units of each type of unit is considered as the state variable and the number of new generating units constructed in a year is represented as the control variable.

A dynamic programming based generation planning framework in a competitive, energy-only electricity market is presented in [108], in which price dynamics are simulated by a price-demand model with the consideration of short-term price uncertainties, and then the annual profit is stochastically simulated using Monte Carlo (MC) approach. In [49, 77, 89], stochastic dynamic programming has been used for solving the optimal sequential investment problems where uncertainty in demand and fuel prices is represented as a discrete Markov chain. In these papers, only one investor within a market is studied. The investor is assumed to be a new entrant to the market, and its effect on the profitability of the existing fleet of generation capacity is ignored. A framework combining dynamic programming and game theory is proposed in [109], in which a forward dynamic programming has been used for solving the long-term investment problem faced by each GENCO and the Cournot game is applied to model the strategic interactions among market players.

The greatest concern towards the application of dynamic programming is the curse of dimensionality. In this case, the dimension of the state space grows rapidly with the increasing number of possible capacity mixes at each year and number of planning years, resulting in a substantial increase in computing requirements. In the presence of uncertainty capturing different sources of uncertainty such as energy from wind, demands, prices, and rainfall.

One solution is limiting the number of decision alternatives to be evaluated at a time. In [49], only one new plant of each technology can be constructed within the

planning horizon of 10 years in the case study. A second approach is conducting a year-by-year optimisation with static or dynamic look-ahead algorithms to avoid the problem of short-sighted decision-making [110]. The computational efficiency for the multiple executions of dynamic programming can be significantly improved compared with the global dynamic programming optimisation. A third approach is approximate dynamic programming (ADP) [111] that has recently been proposed to address complex deterministic and stochastic sequential decision-making problems with a large number of decision variables and uncertainties. ADP has been successfully applied to modelling long-term energy policy and investment decisions in [107] with the capability of capturing short-term variations in intermittent energy and demand, as well as storage. The key underlying idea is to replace the true value function with some form of (parametric or non-parametric) statistical approximation. ADP is not used in our current work but it should be explored as one of our directions for future methodological research.

System Dynamics

System dynamics is an approach to understand the nonlinear and dynamic behavior of complex systems using flows, internal feedback loops, and time delays [112].

In LTGI models based on system dynamics, the feedback loops denote the response of new investments to price signals (e.g., energy and capacity prices), and the time delays refer to the lead times of capacity construction. Continuous or discrete differential equations are used to represent the dynamics (change) of installed capacity over a planning horizon [113]. In comparison to some traditional analytical or optimisation methods, system dynamics models are easier to implement and they have more flexibility in modelling the interactions involved in the complex long-term investment decision-making process. Optimization-based models are restricted somehow (e.g., the problem size and complexity) in order to guarantee that there is an optimal solution. By contrast, System dynamics models often choose a simplified response function (e.g., profit versus action/decision) instead of an objective (minimizing cost or maximizing profit) when dealing with decision-making problems. Besides, system dynamics models use dynamic simulation to study the behaviour

of systems and the impact of alternative policies. The simulation technique can efficiently capture linear or non-linear interactions between inputs and outputs, and easily implement delays involved in the investment implementation, bounded rationality of making decisions (e.g., imperfect spot energy price forecasts used for making investment decisions), and imperfect foresight toward future developments (e.g., uncertain load growth and fuel prices).

System dynamics has been applied to addressing issues of market incentives and regulatory effects [14, 17, 18, 114], competitive strategies and market power [4], and most importantly the dynamics of investment decisions and investment cycles [4, 14–16]. A system dynamics model is developed and calibrated (to variables like system capacity and the wholesale spot energy price in both PJM and ISO-NE) in [115], where the simulation results show that liberalisation is a driver of boom-and-bust cycles on generation capacity growth. Simulations in [14, 17, 18] explore the effects of different market mechanisms on the amplitude of boom-and-bust cycles.

System dynamics programs and investment optimisation programs can complement each other in performing a LTGI study [110]. In [70], a mixed integer programming model is formulated for each GENCO to maximize its total expected profit over the whole planning horizon, and system dynamics obtains and updates the information that is needed for individual GENCOs such as the long term electricity demand and electricity price development.

Hierarchical models

Generation investment decision-making requires the coordination between long-term or medium-term generation planning and short-term operation. Mathematical programming with a single objective function might be inadequate for determining both optimal bidding strategies and investment decisions in liberalised markets. For example, one may wish to study the impacts of design of policies determined by policymakers on investment decisions made by GENCOs. To account for multiple objectives of different market participants, the multilevel or hierarchical modelling technique has been employed in recent studies.

Ref. [116] develops a bi-level market equilibrium model where the upper-level

decisions are design of policies made by the leader (policymakers) and the lower-level decisions are investment decisions made by followers (GENCOs). Ref. [117] employs a two-stage stochastic program to explore the effects of different market designs and demand-side bidding (i.e., a price responsive demand) on capacity adequacy, while accounting for uncertainty in demand, fuel costs and transmission capacities. In [118–120], a bi-level optimisation model is formulated with the investment game in the upper level and the market clearing game in the lower level, in order to analyse the competitive behaviour among individual GENCOs considering non-cooperative investment game or incomplete information of rival producers in the energy market.

Hierarchical models can not be directly handled by traditional equilibrium problems due to the mixing of primal and dual variables. Complementarity models are introduced in [116, 121–125] to cast equilibrium problems using the mixed complementary problem format. The implementation of a mixed complementary problem is to firstly formulate an optimisation model for each player, then obtain the first-order optimality conditions (Karush-Kuhn-Tucker/KKT conditions) for each mathematical program, and finally combine all those conditions together with the market clearing conditions, resulting in a (linear or nonlinear) complementarity problem. In an uncertain context, stochastic equilibrium models can be reformulated from a standard two-stage stochastic optimization capacity expansion model, as described in [85].

The scalability of a hierarchical model is a bottleneck for its applications to long-term/multi-period generation investments in large-scale electricity networks. Moreover, it is difficult to deal with a mathematical program with equilibrium constraints because the feasible region of the program is not necessarily convex or or even connected. The lower-level problems in a hierarchical problem are often required to be free from integer variables and to be as simple as possible so that the mixed complementary problem/mathematical program with equilibrium constraints can be solved within reasonable time.

2.2.4 Challenges in the applications of LTGI models

The various types of LTGI models as reviewed in Section 2.2.3 have in common a high dimensional input space, with considerable uncertainty over both the appropriate values of inputs and over the relationship between the model structure and the real world. In real-world applications, it is challenging but valuable to carry out a comprehensive uncertainty analysis that considers all sources of uncertainty.

There exist studies on modelling some stochastic quantities in LTGI models, such as short-term and long-term demand uncertainty described by normal distribution function [17, 18] or a discrete binary Markov chain in [19]. However, uncertainty in some internal model parameters and uncertainty in the model structure (see Section 2.3 for more explanations) are usually not considered in the literature but they will be emphasized in our work.

Another challenge is that the LTGI models are often computationally intensive. A single model run may take many hours [4, 13] or even many days or weeks with more detailed modelling of short-term power system operations [21, 97]. Consequently, the number of possible runs of a complex LTGI model is limited, which makes it difficult to perform a comprehensive UA and SA.

2.3 Sources of uncertainty

The role of risk and uncertainty is of great importance in the LTGI problem. Both risk and uncertainty are based on lack of certainty in a potential fact, an event, an outcome or a scenario, etc. Here, a risk refers to a factor that may result in uncertain financial returns of investment, such as demand, fuel prices and spot prices. Uncertainty refers to a situation which involves imperfect and/or unknown information due to limited observability and/or stochastic behaviour, as well as lack of knowledge.

Inherent uncertainty is involved in projecting revenues (and costs) over the long lifetimes (20 to 50 years) of a new generation capacity, and thus it influences the decision-making on investments and energy policies. Therefore, it is necessary to adequately account for uncertainty around the outcomes (e.g., projections of generation

capacity and system reliability) of a LTGI model related to the real world [126,127].

Three major sources of uncertainty, namely, input uncertainty, structural uncertainty and functional uncertainty, will be taken into account in the LTGI problem. Other sources of uncertainty, such as observation error are assumed not to contribute much extra uncertainty.

2.3.1 Input uncertainty

Input uncertainty comes from the model parameters that are inputs to the model but whose appropriate values are unknown to us. Among these model inputs, there are some parameters whose values cannot be inferred by statistical methods, such as future design of energy policy, as this is totally controlled by policymakers. By contrast, uncertainty in some internal model parameters that represent particular modelling assumptions may be reduced by learning from historical observations, in order to move the model closer to the real process or system [128].

Predictions of the future profit of an investment rely on the outcome of random events (regarding their predictability such as wind speed or generator outages). These random events are often represented by statistical models, such as modelling the wind speed at a certain time snapshot as a normally distributed random variable which is characterized by the parameters of mean and variance. It is commonly assumed in the literature that all of uncertainty in the wind speed has been encompassed by a specified probability model. However, there is still uncertainty in the input parameters and the structure of the probability model regarding its accuracy in representing the real-world wind speed (as will be further discussed in Sections 2.3.2).

Examples of input uncertainty in the LTGI model mainly include risk factors ranging from policy risks, technology risks, finance risks to market risks.

Policy risks

Design of future energy policies is unknown to investors. Some administratively determined values, such as the energy price cap in an energy market have a significant impact on investment decisions. Uncertainty in carbon prices is largely dominated

by carbon policies supporting investment in low-carbon generation including the Carbon Price Support mechanism [129] and the European Emissions Trading Scheme (EU ETS). The Carbon Price Support is proposed by the GB EMR consultation with a carbon price floor made up of the price of CO_2 from the EU ETS and the Carbon Price Support rate per ton CO_2.

Technology risks

Technology risks, such as uncertainty in future capital costs, construction costs and decommissioning costs, affect the type of generation capacity to be invested. In comparison, nuclear and coal power plants are less attractive than CCGT and OCGT power plants because they have higher capital costs and longer lead time.

Finance risks

Private investors make their investment decisions based on their forecasted profitability of a new investment over a long time horizon, during which there is uncertainty of a return and a potential for financial loss. Rate of return, expressed as a proportion of the annual income to the original investment, is one of the finance risks. Generally, investors would require a higher rate of return for choosing a generation technology with a higher cost of capital.

An investor's risk attitude toward finance risks is unknown to other investors and policymakers. There are three alternative risk preferences – risk aversion (the preference for a certain outcome over a risky outcome), risk neutrality (indifference between a certain outcome and a risky outcome), and risk loving (the preference for a risky outcome over a certain outcome).

Market risks

In an electricity market, investors and policymakers are exposed to electricity price risks arising from the unknown bidding behaviour of GENCOs, as well as from future projections on fuel prices and electricity demand that are difficult to predict [13]. Uncertainty in fuel prices is passed on to uncertainty in electricity prices since the fuel cost is included in the energy offer of GENCOs to the spot market. Due to

the recent improvement of energy efficiency, economic downturn and the potential increasing level of demand side response, the demand growth rate has considerable uncertainty, which contributes to long-term electricity price risks.

2.3.2 Structural uncertainty

Structural uncertainty arises from the fact that the model and the reality are almost always not the same. More specifically, there may be unknown functional relationships or imperfect science in the model equations or errors in the structure of the model due to the lack of knowledge of the underlying true system or process. Structural uncertainty is referred to as model inadequacy or model discrepancy between the model and the reality; this is expected even with the best values of model parameters. Structural uncertainty often exists in complex LTGI models and needs to be calibrated and quantified against the real-world observations.

2.3.3 Functional uncertainty

As mentioned in Section 2.2.4, LTGI models with a high-dimensional input space are often computationally intensive. A particular issue in high dimensional models is that while broad knowledge of the model input-output relationship across a range of credible inputs is required when it comes to carry out computational tasks of calibration, UA or SA, it is usually only possible to cover sparsely this input space with model runs. At any untried point (where the full LTGI model has not been run), the function representing the relationship between the model inputs and outputs are unknown to us, which forms functional uncertainty. In order to enable a quick evaluation of the function at any input, a meta-model is often used in replace with the full model. The meta-model gives an evaluation of a model output at any untried input point as an approximation to the output evaluated by the full model. Uncertainty in this approximation (between the meta-model and the full model), in principle, arises from functional uncertainty.

2.4 An illustrative example of uncertainty

An example presented here is used to illustrate the three major sources of uncertainty – input uncertainty, structural uncertainty and functional uncertainty.

Example 1. (*The concentration of a chemical*[3]). Suppose the concentration of a chemical that evolves in time is modeled as $e(r,t)$ satisfying the differential equation,

$$\frac{de(r,t)}{dt} = re(r,t), \tag{2.1}$$

where r is a rate parameter and t is time.

The equation (2.1) above can be written as,

$$e(r,t) = e_0 \exp(rt). \tag{2.2}$$

We wish to learn about the best value of the model parameter r for describing the real system (i.e., the concentration of a chemical) when some observations are available. Assuming the initial conditions are $e_0 = e(r, t = 0) = 1$ and a measurement at $t = 3.5$s, we run the model from $t = 0$s to $t = 5$s at five selected values of the model parameter r. Fig. 2.1 shows the five simulation runs indicated by the colored lines and the measurement at $t = 3.5$s is shown with a black dot, with measurement error given by the error bars.

If the analytic solution to the function $e(r,t)$ as expressed in (2.2) is unknown, which is common for complex functions, the solution would be generated by numerically solving the differential equation. To answer which values of r produce the output $e(r,t)$ that is consistent with the observations, we now focus on the function $e(r, t = 3.5)$ as a function of r only.

Fig. 2.2 presents the five simulation runs (in colored points) and the observation (in solid black horizontal line) at $t = 3.5s$. If the analytical model $e(r, t = 3.5) = \exp(3.5r)$ is known, a single value of r would be easily identified if the measurement error is ignored. However, uncertainty in the observation of $e(r, t = 3.5)$, indicated by the two dotted black horizontal lines, leads to uncertainty in the inferred values of

[3]This toy example is adapted from one of Michael Goldstein's presentations titled "Bayesian uncertainty analysis for complex systems modelled by computer simulators".

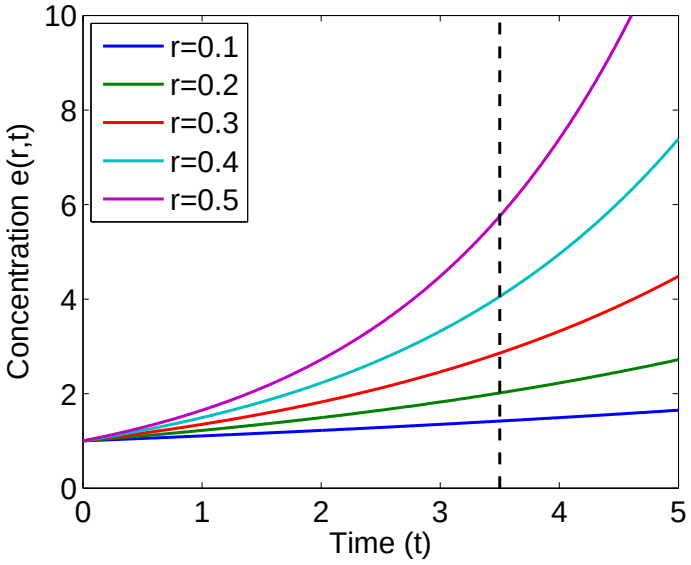

Figure 2.1: Concentration of a chemical that evolves in time at different rates.

r, that is, a range (in green) of possible values of r consistent with the measurements, with all the implausible values of r in red/yellow in Fig. 2.2.

Apart from uncertainty in the model parameter r, structural uncertainty (or model discrepancy) around the model output $e(r)$ itself is indicated by the red dashed lines in Fig. 2.2. The consideration of model discrepancy results in more uncertainty in the estimated value of r, and hence a larger range (in green/yellow) of r values, with all the implausible values of r in red.

Functional uncertainty at untried input data is indicated by the solid purple lines, as shown in Fig. 2.3. Without the consideration of model discrepancy, functional uncertainty and measurement uncertainty result in the plausible values of r in green in Fig. 2.3. Supposing that functional uncertainty is independent from structural uncertainty, the additive effect of these two sources of uncertainty leads to the uncertainty range quantified by the blue solid lines. Hence, the range (in green/yellow) of possible values of r is enlarged, in comparison with that shown in Fig. 2.2.

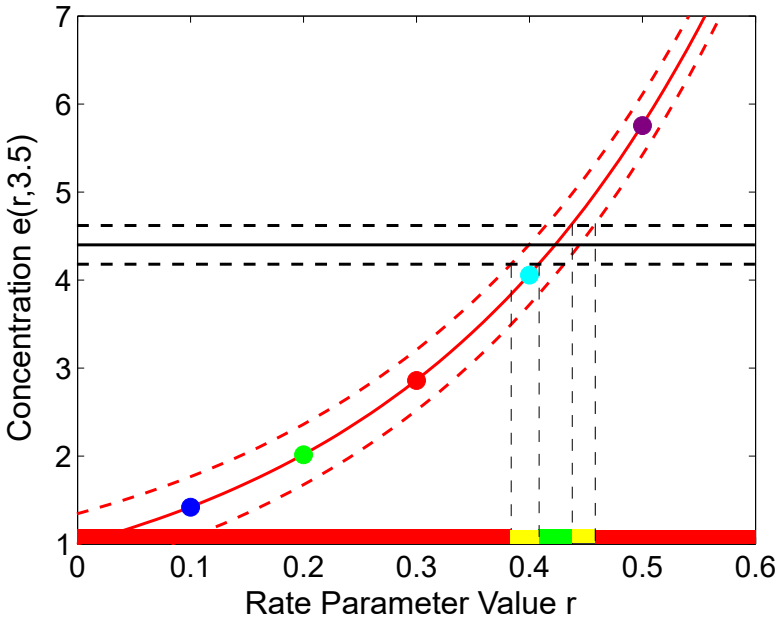

Figure 2.2: Concentration of a chemical as a function of the rate parameter at time $3.5s$*, accounting for structural uncertainty and observation errors.*

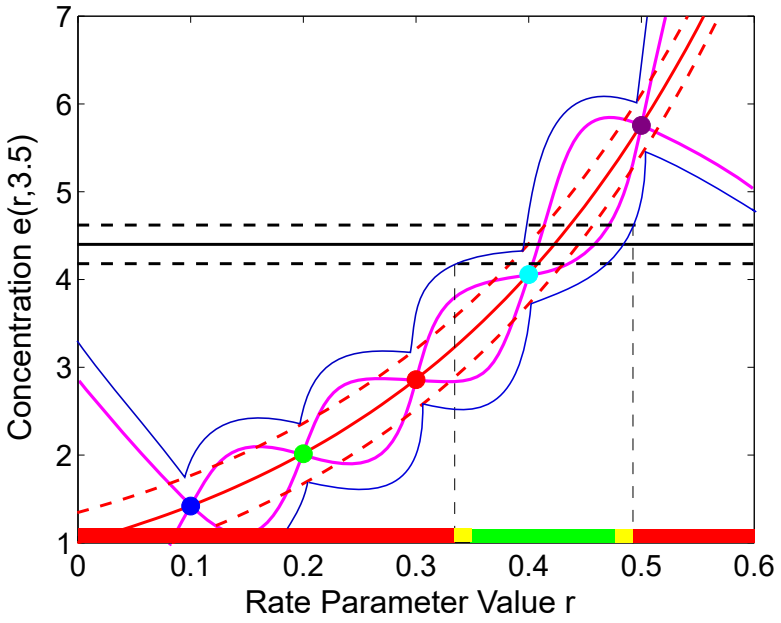

Figure 2.3: Concentration of a chemical as a function of the rate parameter at time $3.5s$*, accounting for structural uncertainty, observation errors, and functional uncertainty.*

2.5 Dealing with uncertainty

For real-world applications of LTGI models such as design of energy policy or determination of investment decisions, it may be necessary to quantify, analyse and reduce uncertainty in model outputs. General tools for dealing with uncertainty are calibration, UA and SA. Different assumptions on the characteristics of uncertainty may provide different implications for policymakers. Therefore, it is important to describe and analyze uncertainty in a systematical way that reflects the use of the model for policy purposes. Meanwhile, it is necessary for policymakers to design a robust energy policy against different sources and assumptions of uncertainty.

2.5.1 Calibration

Calibration or *history matching* refers to the process of learning from observations and tuning a model to the real process to best approximate and predict reality. Calibration is a valuable tool for validating a LTGI model and ensuring the quality of investment decisions to be delivered by the design of energy policy.

Some calibration work is found on the statistical modelling of fuel prices or electricity prices based on historical data [73, 74]. However, there is limited work on calibrating a complex LTGI model. Dan Eager et. al. developed a LTGI model described in [82] and found a set of model parameters for projecting future generation capacities. The model is validated by showing graphical comparisons between the simulated and observed market dynamics (i.e., capacity margin, total installed capacity, installed capacity for each generator, and average monthly spot energy prices) since the introduction of the NETA to England and Wales in GB wholesale electricity market. However, there is no calibration work done regarding tuning the model parameters to better match with the real world. The modellers in [115] manually varied some model parameters that potentially have substantial effect on the price behaviour, gradually achieving a reasonable match between the simulated results and the historical data of peak spot prices for PJM and ISO-NE. The quality of calibration is examined through visual graphs and two statistical measures - the mean absolute error over the mean and the Theil statistics [130]. However,

the calibration method based on trial-and-error used in [115] is very inefficient for resource-demanding models. In addition, it is incapable of capturing and quantifying the structural uncertainty that applies to future projections of model outcomes.

A formal calibration typically involves calibration of a subset of uncertain model parameters against historical observations of the model output whilst modelling the discrepancy between the model and the real system. Uncertainty in calibration parameters whose values are unknown to modellers may be reduced via inference using statistical methods. In Bayesian approaches, uncertainty in model parameters is specified ex ante as a probability distribution based on the prior beliefs of the model user or other experts, and these prior beliefs are updated according to Bayes' rule (*i.e.*, by identifying values of calibration parameters that are plausible with respect to historical observations of the model output). A more detailed description will be given in Section 3.3.4.

In the example presented in Section 2.4, the calibration process involves the inference of the rate parameter and of the model discrepancy based on the available observations. To the best of our knowledge, no such formal calibration of LTGI models has previously been done. If a calibration against historical observations is not performed, this severely limits the conclusions which can be drawn from assumptions based models regarding investment decisions, the capacity requirements and policy design in the real system.

2.5.2 Prediction

Prediction is estimation of model outputs at input configurations that have not been tested. There is uncertainty about how close the true real-world outcomes will be to the prediction that is the outputs simulated by the model. As model users, investors or policymakers should wish to be provided with a plausible range of the possible model outcomes rather than a single value or scenario. It is also of great value to quantify uncertainty around a prediction using a probability interval.

Suppose investors or policymakers are determining the combined values of all decision/control variables that trigger the adequate amount of generation capacity at low costs. Predictions will be made at the selected settings of control inputs. In

this situation, uncertainty in the predicted total installed capacity arises from many sources, particularly from input uncertainty and structural uncertainty.

2.5.3 Uncertainty analysis

Uncertainty analysis focuses on uncertainty quantification and propagation of uncertainty from model inputs and structure to model outputs. Uncertainty quantification is a process of assessing the uncertainty of variables (e.g., model inputs, model outputs and observations) that are used in decision-making problems or the uncertainty of a measurement in experiments. Propagation of uncertainty explores the whole picture of model predictions while accounting for various sources of uncertainty in model inputs and structure. Depending on nature of the variable and the technique we adopt, uncertainty in a variable may be characterized by a (joint) continuous/discrete probability distribution or a set of quantitative/qualitative scenarios.

UA is most relevant when those outputs provide guidance in real-world decision-making problems, such as using a LTGI model for investigating the uncertainty of future LOLEs, and further for suggesting the settings of the energy price cap that has an effect on the range of LOLE. Since there exist various kinds of uncertainties in the lengthy planning process, the lack of UA will lead to the underestimation the risks exposing to the investors.

A number of techniques such as MC simulations, scenario planning and Bayesian statistics can be used for carrying out an UA. In [18], MC simulations are employed to determine the average value and the highest and lowest limits of confidence intervals for variation of reserve margin and generation total price caused by uncertainty in annual demand growth rate. However, MC-based UA is particularly challenging for a high-dimensional energy planning model that is expensive, in time and computational resources to evaluate at any point of the input space. Scenario planning, as the simplest and aggregate form of UA, characterizes uncertainty using a small number (typically three to five) of plausible scenarios (i.e., possible alternative futures), and more often expresses the resultant model outputs in qualitative terms such as "high" or "low" than in quantified terms such as probabilities [131]. Compared with

numerical simulation based UA, scenario planning is relatively easy to implement and to explain but it is not analytically rigorous as the qualitative terms such as "plausible" it relies on do not convey clear and precise meaning. Besides, scenario planning is not comprehensive due to the limited number of scenarios selected that may not represent the full range of potential futures.

A formal UA based on probability distributions provides the analyst with a comprehensive picture, that is, the full range of outcomes giving all extremes, and with an understanding of the outcomes given a specific setting (a real-life scenario) of input variables. In practice, we can collect a small number of quantitative judgements from a group of experts, and then fit a proper distribution to those judgements. Useful individual judgements on an uncertain quantity/event include plausible limits, mean, quantiles (e.g., median and tertiles), and likelihoods using Roulette method (using the bins in the Roulette grid) [132]. However, a probability distribution is more difficult for stakeholders in government to understand than scenarios. They may also be skeptical on the credibility of a probability distribution that is formed based on many assumptions.

2.5.4 Sensitivity analysis

Sensitivity analysis studies the relationship between a model's inputs and outputs. Quantifying the output uncertainty is the role of UA, but there is often interest in identifying which inputs have the strongest or negligible influence on outputs; this can be addressed by ranking the sensitivity of model outputs to variation or uncertainty in model inputs [133]. In addition, one may wish to know how much of the overall uncertainty in the output of a model or system can be attributed to uncertainty in particular inputs or groups of inputs, in which case variance-based methods are widely used for measuring the sensitivity of the output to an input variable by the amount of variance in the output caused by that input.

In the LTGI problem, there can be many parameters of interest, and investors or policymakers might not be fully aware of critical parameters that affect uncertainty in their objectives (e.g., investors' future profits or power system reliability). Therefore, sensitivity information is useful in giving decision-makers an idea of uncertainty

involved with new investments and policy design, and hence implying effective ways to reduce decision uncertainty by focusing on the most important model inputs.

One measurement of sensitivity is partial derivatives, as expressed as $\partial f(x)/\partial x_i$, which measures the change in output $f(x)$ when a particular input x_i was perturbed slightly from the nominal (base-line) value. This measurement is referred to as local SA because it is limited to exploring the impact of an infinitesimal change in an input. When the nominal input values are perturbed slightly, the SA results have limited value in understanding the consequences of input uncertainty. Besides, local SA is unable to account for any nonlinearity in the response to x_i. Alternatively, global SA methods are more commonly used.

One simple global SA method, the one-way or one-at-a-time method has been employed in [13,16,18,20,115,134,135], where each input is perturbed only one at a time from its nominal value while all others are held constant at their nominal values for each run of the model, and then the resulting change in output is regarded as a measure of sensitivity to the varied input. However, the one-way method is incapable of taking into account interactions among different inputs.

Multi-way SA can identify the combined effects of two or more inputs, through varying the inputs together rather than individually using a large and highly structured set of simulator runs [136]. In the one-way and multi-way methods where the inputs are substantially perturbed, there is a question of determining the perturbation or uncertainty range. For example, variations in the inputs are chosen randomly within the upper and lower range of an input, without accounting for the relative credibility of these values.

To avoid the arbitrary definition of variations, probabilistic SA is an alternative approach to multi-way SA that can address interactions and nonlinearities [137,138]. Probabilistic SA emphasizes on the careful and explicit specification of a probability distribution of model inputs. Therefore, it has the unique ability to describe the relationship between input uncertainty and output uncertainty.

2.5.5 Methods of carrying out calibration, uncertainty analysis and sensitivity analysis

A conventional way to conduct a formal calibration or a UA/SA is the MC method of drawing random configurations of inputs from their uncertainty distributions, running the model for each input configuration to obtain the set of outputs, and constructing the output distribution (which can in principle be evaluated to any desired accuracy).

Computationally intensive models associated with large studies tend to have high-dimensional inputs. The MC-based method may require thousands of (if not more) individual evaluations in order to avoid sparse coverage of the model input space. It may be practically impossible for complex models to achieve very dense coverage of input space even if very large computer resource is available [131]. Even where a very large number of runs may be possible by acquiring additional computing resource, the approach adopted in our work allows results to be obtained in a systematic way with a smaller computing resource.

The Bayesian approach is applied in our work to inferring the unknown values of some parameters and quantifying the model discrepancy of the LTGI model. Directly using such a complex computer model for statistical inference produces intractable likelihood functions. There are methods of indirect inference [139], approximate Bayesian computation [140], and likelihood-free Markov chain Monte Carlo (MCMC) [141] proposed to overcome the computational challenge of intractable likelihoods in complex computer models. However, for carrying out a comprehensive UA and SA, there still need a large number of computationally expensive model runs. The underlying idea of resolving the intractable model runs is to use a simpler and more efficient model as an approximation to the original model. Bayesian emulation is such an approach that introduces a statistical emulator to approximate the full computer model and to quantify the functional uncertainty in the output of the model at all points in the input space where it has not been evaluated. In this way, Bayesian emulation can systematically deal with various sources of uncertainty (e.g., input uncertainty, structural uncertainty, functional uncertainty) and enable

very efficient calibration, UA and SA of a complex computer model.

Bayesian emulation has been successfully applied to complex galaxy formation models [142], climate models [143, 144], infectious disease models [145] and DNA population dynamics [146], etc. The Bayesian approach adopted in our work is based on a Gaussian process (GP) emulator, which characterizes uncertainty regarding the simulator as a GP and uncertainty regarding the inputs and outputs by complete probability distributions. More theory and technical details of Bayesian emulation will be presented in Chapter 3.

2.6 Chapter summary

This chapter first gives an overview of electricity market designs in transition for facilitating the integration of VG, including vertically integrated structure, energy-only market, scarcity pricing, and capacity mechanisms. Next, the roles that market participants play in the LTGI problem are explained, followed by the evolution of system adequacy metrics. Then, the modelling studies on the LTGI problem in the literature are reviewed.

The computer models and the real world are not the same, due among other issues to uncertainties in input parameters and discrepancies between the model structure and the real world. An example was presented to illustrate three major sources of uncertainty, namely, input uncertainty, structural uncertainty and functional uncertainty. In order to use LTGI models robustly for decision support, it is necessary to quantify uncertainty in the relationship between model outputs and the real-world equivalents. In order to systematically and efficiently manage uncertainty in a complex LTGI model, it is necessary to create an emulator of the full model, which may be evaluated quickly for any input and which quantifies uncertainty in the output of the full model at inputs where it has not been run. The statistical methodology, Bayesian emulation, will be demonstrated in the next chapter.

Chapter 3

Application of Bayesian calibration to long-term generation investment models

This chapter presents methodology for carrying out formal calibration, uncertainty analysis (UA) and sensitivity analysis (SA) of a computationally intensive long-term generation investment (LTGI) model (*i.e.*, *the simulator*) with careful management of various sources of uncertainty. Calibration links the model with the real world by learning from historical observations the updated information of some uncertain model parameters and model discrepancy. UA and SA are powerful tools for making sense of uncertainty around model inputs and outputs when a LTGI model is used for real-world analysis.

LTGI models, such as the one proposed in [4, 13] and used as an exemplar in our work, are often computationally expensive to run. The traditional MC-based techniques require a very large number of simulator runs that may be impractical to achieve. In comparison, Bayesian emulation enables very efficient calibration, UA and SA. A Bayesian emulator approximates the simulator as a Gaussian process model and quantifies uncertainty in the approximation. A simple example is presented to illustrate the performance of an emulator. Finally, descriptions are given on how to carry out validation, calibration, and SA within a Bayesian framework.

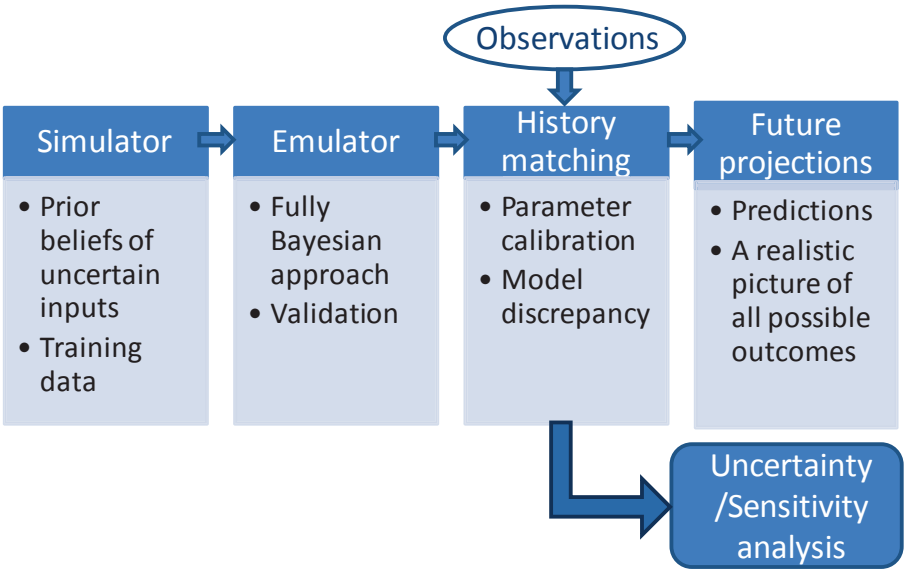

Figure 3.1: Diagram of the Bayesian framework.

3.1 Diagram of the methodology

A highly efficient Bayesian approach described in [133,137,144,147] will be employed to tackle the computational challenge found when performing calibration, prediction, UA and SA on complex computer models. Fig. 3.1 shows a diagram of the proposed Bayesian framework, which is based on a Gaussian process (GP) model (*i.e.*, *the emulator*) that is built as an approximation of the simulator using a limited number of simulation runs (*i.e.*, *training data*). The traditional MC-based method directly uses the simulator for model evaluations at untried model inputs. By contrast, the Bayesian approach uses the emulator for evaluations and only a small number of simulator evaluations are required for the development of the emulator.

In Section 3.2, a brief description of the simulator under study is provided, emphasizing uncertain model inputs and outputs that are of interest. The theoretical foundations of the Bayesian framework are illustrated in Section 3.3.

3.2 The simulator

A brief description of the simulator under study is provided in this section, with emphasis on the uncertain model inputs and the output of interest. The LTGI model developed by Eager, Bialek and Hobbs [4, 13] will be used as an exemplar

to demonstrate the application of the Bayesian approach to a decision-support tool. The modeling technique of this simulator is based on system dynamics which is most commonly used in industrial practice particularly in modeling electricity markets under uncertainty [113]. Whilst this is a specific application, the Bayesian framework as presented here can be generally applied to models in which uncertainty plays a key role and where the link between the model and reality is of great importance and interest to model users. For more discussions, please refer to Section 4.6.

3.2.1 High-level formulation of the simulator

The simulator developed in [4,13] provides projections of installed thermal capacity and LOLE given a scenario of on-shore and off-shore wind capacities over a long planning horizon $\mathcal{T}$. The simulator is system dynamics based and it obtains an investment response as a function of a performance index (e.g, profitability and its likelihood) instead of using optimality or equilibrium criterion as in optimization based models. The simulator is stochastic because it considers annual load distribution, generator availability, annual load growth rate, and stochastic fuel prices. Forward-looking MC simulations are conducted at each decision year to obtain the performance index of an investment. A full description of the simulator may be found in [4, 13].

Fig. 3.2 shows the structure and the main inputs and outputs of the simulator. A high level description of the five main modules within the simulator is provided in eqs. (3.1a)–(3.1e).

$$F_t = h_1\big(\kappa, q_t\big) \tag{3.1a}$$

$$P_{markup,t} = h_2\big(ND_t, \{AG_{g,t}\}, \theta_{markup}, u_{voll}\big), \quad \forall g \in \mathrm{G} \tag{3.1b}$$

$$\big[P_{e,t}, \{C_{g,t}, y_t^L\}\big] = h_3\big(CM_t, F_t, P_{markup,t}\big), \quad \forall g \in \mathrm{G} \tag{3.1c}$$

$$\big[y_{g,t}^B, y_{g,t}^M, y_{g,t}^D\big] = h_4\big(\theta_{VaR}, \{P_{e,\tau}\}, \{C_{g,\tau}\}, u_{cone}\big), \quad \forall \tau \in \{t, \ldots, t+\tau_f - 1\} \tag{3.1d}$$

$$y_{g,t} = h_5\big(c_{g,t-1}, y_{g,t}^B, y_{g,t}^M, y_{g,t}^D, RT_{g,t}\big). \tag{3.1e}$$

The five modules are described in the following sections.

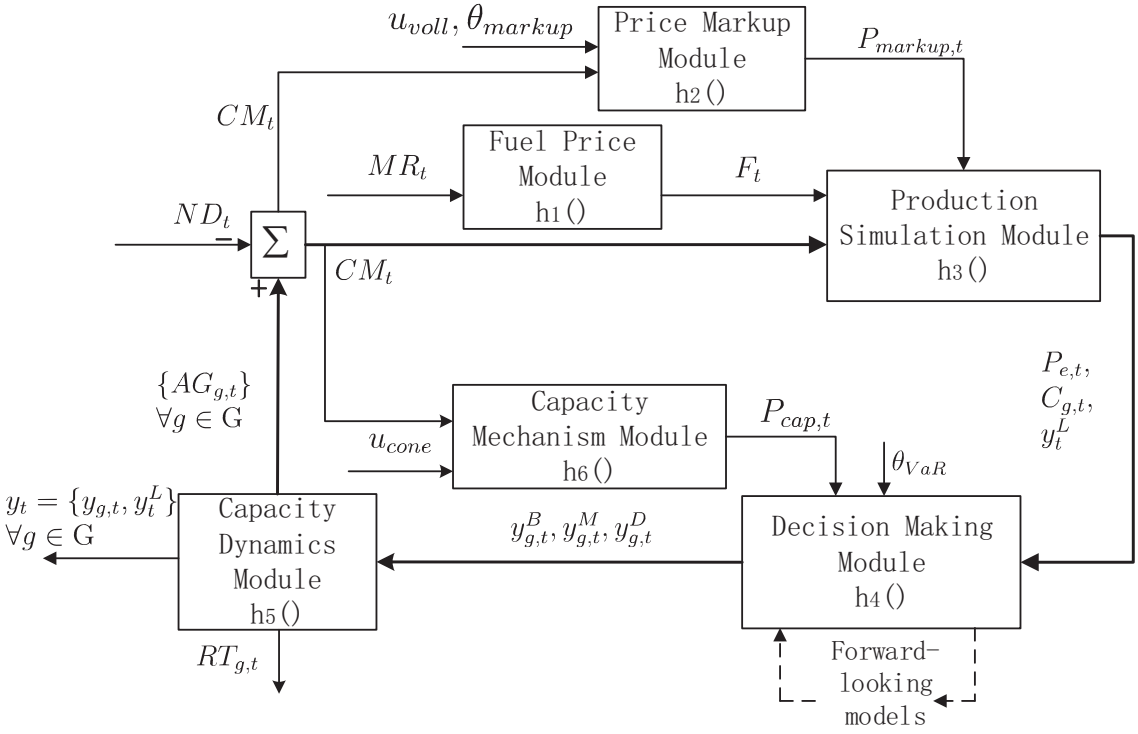

Figure 3.2: The model structure of the simulator.

Fuel price module

In (3.1a), $h_1(\cdot)$ simulates annual fuel and carbon prices (both in the past and future) using mean-reverting stochastic processes [13, 148], as expressed as,

$$dF_t = \chi(MR_t - F_t)d_t + vol(F_t)dW_t, \tag{3.2}$$

$$MR_t = \frac{\kappa}{\chi}\frac{dq_t}{dt} + \chi q_t, \tag{3.3}$$

where F_t is the fuel price at year t, MR_t the reference time dependent mean reverting level, $vol(\cdot)$ characterizes the expected change and the local uncertainty (the year-to-year volatility) of each fuel type, W_t a standard one-dimentional Wiener process, q_t the DECC central estimate, κ a multiplier, and χ is the speed of mean reversion.

Projected fuel prices depend on the reference long-run fuel price projections, referring to the DECC projections here. A multiplier κ is used to adjust the central estimate q_t upwards or downwards, which reflects the long-run or global uncertainty of fuel and carbon prices arising from market changes or political interventions. Applying a multiplier to the trend of future carbon prices, as shown in Fig. 3.3, results in different trend levels of carbon prices (in solid and dashed lines) associated

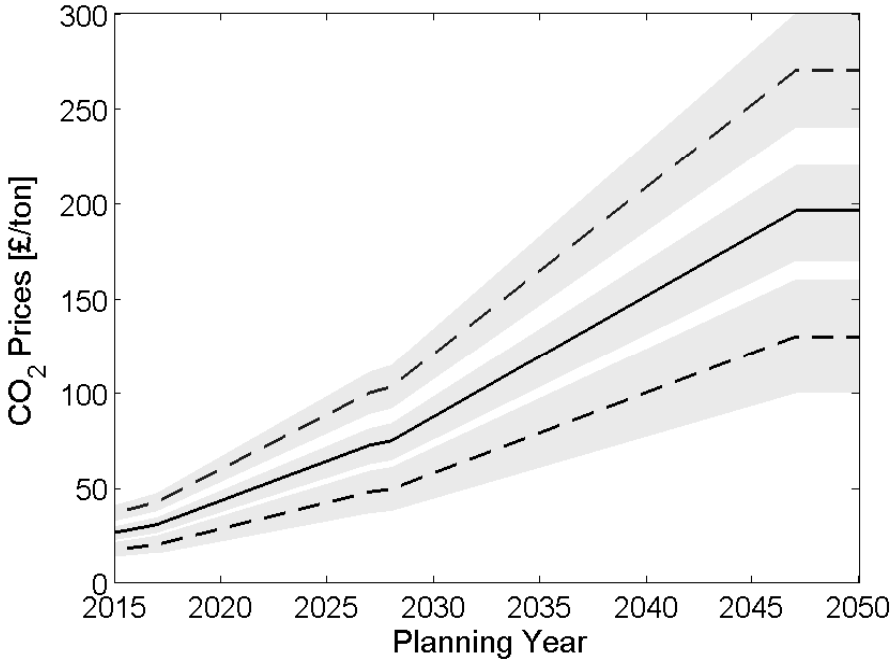

Figure 3.3: Different trend levels of carbon prices (in solid and dashed black lines) using the central estimates of carbon prices published by DECC [1] as a reference (in solid black line); The expanding shaded area (in light grey) reflecting the local uncertainty along the trend level.

with local uncertainty (in shaded areas).

Price markup module

In (3.1b), $h_2(\cdot)$ calculates the energy price markup, as shown in the upper graph in Fig. 3.4. For simplicity, all generators are stacked in merit order and dispatched whenever they are available. The wholesale energy price model is

$$P_{e,t} = SRMC_g + P_{markup,t}, \tag{3.4}$$

where

$$P_{markup,t} := u_{voll} e^{b \cdot CM_t}. \tag{3.5}$$

The price markup $P_{markup,t}$ is represented as an exponential function of the capacity margin. The markup reaches the energy price cap at a capacity margin of zero. The parameter b is calibrated so that the markup approaches zero at a cut-in point of capacity margin, denoted as θ_{markup}. The price markup given as (3.5) becomes prominent when peaking plants $N_g - 1, N_g$ are on the margin. The price

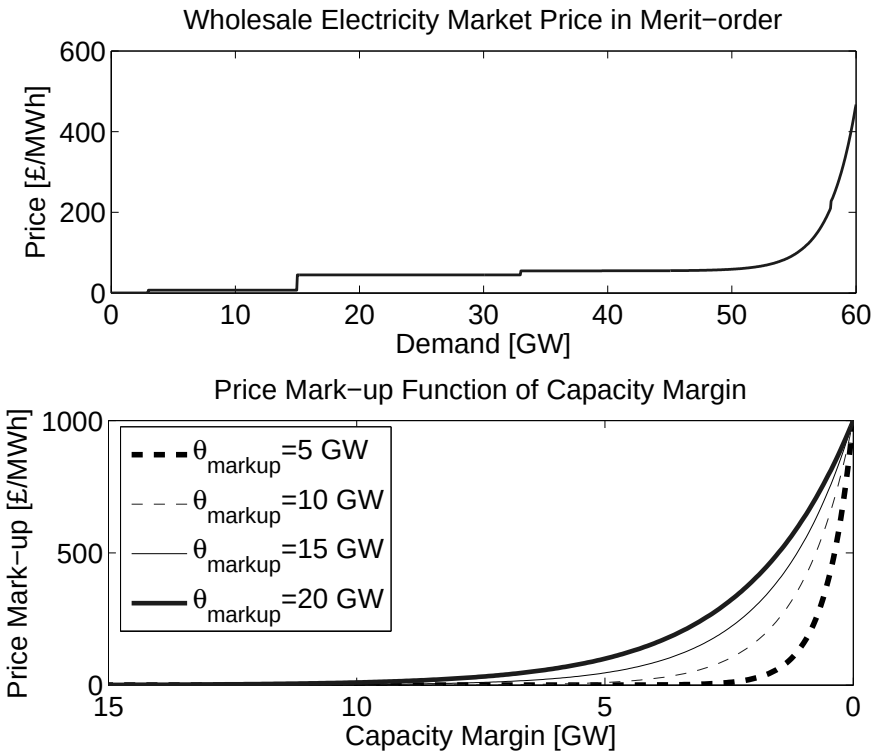

Figure 3.4: Upper: Energy prices cleared in merit-order with price markup; Lower: Price markup functions with different values of θ_{markup}.

markup can be derived from the joint probability distribution of the capacity margin CM_{N_g-1} and CM_{N_g}. Different assumptions on the value of θ_{markup}, that is, on the competitiveness level of the power system, lead to different levels of price markup, as shown in the lower graph in Fig. 3.4.

Production simulation module

In (3.1c), $h_3(\cdot)$ performs the probabilistic production costing method in a nonequilibrium market settlement. The method performs a convolution of generator outages with the annual net load (demand minus wind generation) curve ND_t and calculates the energy prices, costs, revenues and LOLE. Since an existing published model is adopted, discussion about more details on the form of this model is beyond the scope of this thesis.

Decision making module

In (3.1d), $h_4(\cdot)$ is the module of decision-making under uncertainty. At each decision year, the investor assesses the distribution of the value of a project V_t based on the NPV of the first τ_f years of forecasted profitability of new and existing generation capacity. To consider the investor's risk preference, the VaR of the distribution of V_t is selected according to $P_r(V_t \leq V_{VaR,t}) = \theta_{VaR}$, and then $V_{VaR,t}$ is used for making investment decisions. The smaller the value of θ_{VaR} is, the more risk averse the investor is assumed, and the lower level of investment would be.

Given the cost estimations, the energy prices and the capacity price if a capacity market is available, $h_4(\cdot)$ calculates the gross margin of per MW of a plant, as defined as the overall revenue received from the energy and the capacity market minus its variable costs. We refer to Section III in [13] and Section 8.2.4.1 in [4] for more details. Then, the module outputs are annual decisions on the investment in new thermal generators and the mothballing/de-mothballing of existing thermal generators.

Capacity dynamics module

The last function $h_5(\cdot)$ in (3.1e) simulates capacity dynamics that update the generation capacity portfolio at each decision year by adding the newly installed and/or demothballed capacity and removing the mothballed and/or retired capacity. The module inputs include exogenous variables such as plant life time and thermal plant retirements, and endogenous variables such as new builds and new mothballed/demothballed capacity. The module outputs are the installed thermal capacity of each technology and system reliability index (e.g., LOLE) at the decision year.

Capacity mechanisms module

Capacity mechanisms $h_6(\cdot)$ may refer to a forward capacity market, capacity payment, or a reliability option contract. A capacity demand curve derived by DECC and announced in advance of each auction is considered here. Fig. 3.5 shows a representation of supply and demand in the UK capacity market where the intersection of the capacity demand curve and the supply curve sets capacity prices [35]. A ca-

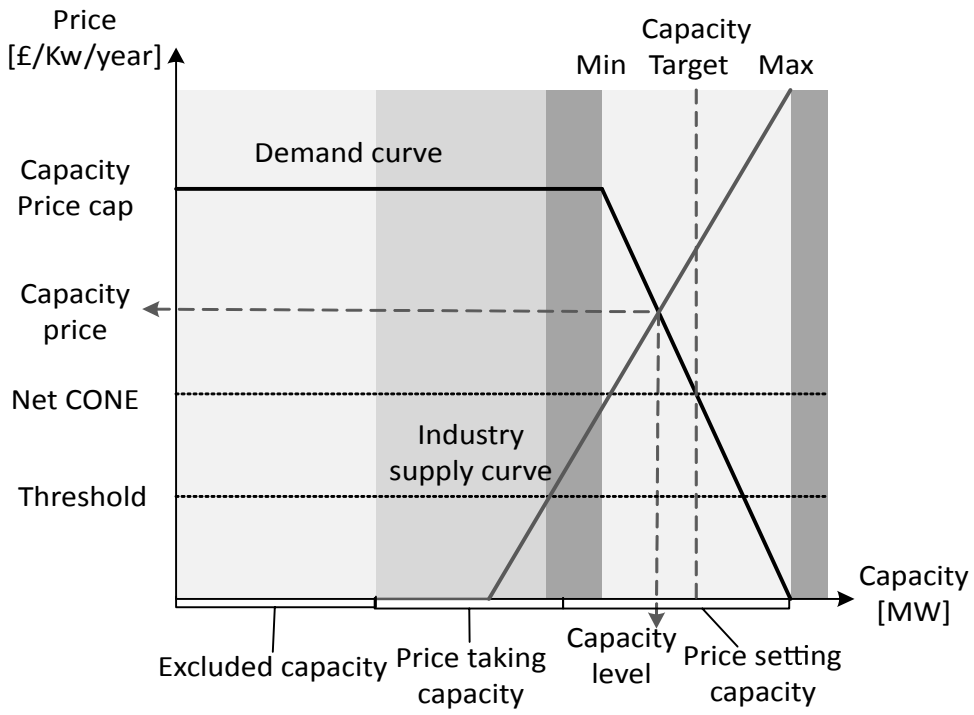

Figure 3.5: Capacity market supply and demand curve.

pacity balance target represents the estimated optimal level of installed capacity in the system. The capacity price at capacity balance target is determined by the net CONE that represents the cheapest cost of a new entrant peaking plant minus its expected annual energy market revenue [23]. In this model of a capacity market, the net CONE, the slope of the demand curve and the capacity price cap are the main design parameters [20], among which the key uncertainty and complexity driving capacity pricing comes from the net CONE. The net CONE is administratively determined in the capacity market (like VOLL). However, the true value of net CONE highly uncertain due to the uncertain cost structure bid by the market participants and estimates of revenues from energy market [23].

3.2.2 Model inputs and outputs

Model inputs

In the decision-making module (3.1d) as described in Section 3.2.1, two common electricity market designs, namely the energy-only market design and the forward capacity market design will be studied. There are too many model parameters to

explore all of them in a reasonable execution time. The inputs of interest here are the source of uncertainty that can result in a substantially different trend in the long-run investment decisions. SA helps identify the influential parameters, and hence can be used for screening out unimportant variables in high-dimensional problems before a full sensitivity analysis is performed. The selection of model inputs in our work is based on the one-way SA results obtained from [4] as well as our research aims of exploring the impacts of market designs.

Among the modules described in Section 3.2.1, six model inputs are identified and categorized into three types from policymakers' perspective: control variables $u := \{u_{cone}, u_{voll}, u_{co_2}\}$ that are determined by the capacity market design, energy price cap policy and the carbon policy, respectively; calibration parameters $\theta := \{\theta_{markup}, \theta_{VaR}\}$ of which unknown values are expected to be learned using historical observations of the model output [133], as further explained in Section 3.3.4; and a forcing input $\omega := \omega_{gas}$ reflecting the global uncertainty (uncertainty in the long-term trend) of future gas prices. Two inputs (calibration parameters) are used for model calibration, five for forecasting under an energy-only market design and six for forecasting under a capacity market design.

We cannot calibrate the control inputs $\{u_{cone}, u_{voll}, u_{co_2}\}$, since they will be administratively determined in the future, nor can we calibrate the forcing input ω_{gas} as the future long-term trend of gas prices might not be independent from the historical trend due to market changes. Calibration parameters are those representing model assumptions such as the market competitiveness and the investors' risk preference, and it should be reasonable to assume that these parameters do not change over time. So the calibration parameters are the only things that we can vary to improve the accuracy of the model, i.e., to match it more closely to the real-world.

The above input categorization is derived from a policymaker's perspective; an investor might have a different categorization of inputs. For example, θ_{VaR} will be a control input since the investor has the freedom of choosing a risk preference, and $\{u_{cone}, u_{voll}, u_{co_2}\}$ will be forcing inputs as these policy-related parameters are uncertain to the investor. Apart from the six model inputs, other model parameters that make the simulator stochastic such as fuel price volatility $vol(\cdot)$ and the variance

of the normally distributed demand growth rate are not assumed to have a long-run effect on investment decisions. The selected model inputs are assumed to be constant over the planning horizon. This approach allows direct exploration of the relationship between one parameter (used to represent a model input) and model outputs. Similar practices can be found in [16,18,20,135]. The dimensionality of the input space can be extended if independent values are needed for an input variable at each year or at each stage (e.g., every five years). Also more model parameters (e.g., the mean or the variability of demand growth rates) can be included in the input space if they are deemed to contribute substantial additional uncertainty.

Model outputs

One of the outputs of interest is the time series of annually installed thermal generation capacity $y^G = \{y_t^G\}, \forall t \in 1 \dots \mathcal{T}$, where $y_t^G = \sum_{g \in \mathrm{G}} y_{g,t}$. The historical observations of operational thermal capacities are available, which allows for calibration in the history matching procedure. The planning time horizon of interest $\mathcal{T}$ is either the past ($t \in \mathcal{P}$), for which observations exist, or the future ($t \in \mathcal{F}$), for which a projection is made.

A second output of interest is the time series of annual LOLE - loss-of-load expectation over the future planning horizon $y^L := \{y_t^L\}, \forall t \in 1 \dots \mathcal{F}$. A third output of interest is the maximum annual LOLE, as expressed as $y := \max(y_t^L), \forall t \in \mathcal{F}$, because we are interested in determining energy policy scenarios within which the maximum value in annual LOLEs over the future planning horizon does not exceed, at any time, a given threshold (e.g. 3 hours per year). After the investment/mothballing/demothballing decisions are made, the LOLE is computed as the number of hours in a year N_t multiplying by the snapshot LOLP, that is, the probability that net load (load minus wind generation) is not met by supply from conventional generators at a randomly chosen point in any hour within the year, as expressed as,

$$y_t^L = N_t \times Pr(ND_t > G_t), \tag{3.6}$$

where $Pr(\cdot)$ is the estimate of the LOLP, estimated using convolution, as described in [4, 13]. Given a specific MC sample of annual demand growth rate, there is a net

load distribution ND_t and hence a solution to an estimate of LOLE at each decision year t according to (3.6). With 100 MC samples of annual demand growth rate at year t, there will be 100 values of y_t^L. The expected (in the mathematical sense) value of y_t^L is chosen as the LOLE at year t, denoted as y_t^L.

The simulator is stochastic because the quantities including future fuel prices and annual demand growth rate are modelled as random variables. However, the expected value of MC simulated annual LOLEs, y_t^L, is a deterministic function of the six selected model inputs if the number of MC simulations is sufficiently large. The simulator returns the same outputs y_t^G, y_t^L if repeatedly executed on the same set of inputs. The statistical metric - LOLE can not be directly observed in the history. The loss-of-load events are rare and the observed loss-of-load time durations in many historical years were zero in the GB power system. In order to properly conduct calibration over $\mathcal{P}$, the time series of historical thermal capacity $y^G := \{y_t^G\}, \forall t \in \mathcal{P}$ are chosen to calibrate against as historical observations are readily available. Given a set of model inputs $x := \{u, \theta, \omega\}$ of N_I elements, the simulator output is $\{y^G, y^L\}$. In this way, the simulator can be described as a deterministic function, $f(\cdot)$,

$$[y^G, y^L] = f(x). \tag{3.7}$$

3.3 Bayesian Approach

In this section, calibration, UA and SA within a Bayesian framework are introduced as ways of dealing with uncertainty in a LTGI simulator. The Bayesian approach is based on a very efficient emulator that is built on top of the simulator as an approximation and estimates the uncertainty in this approximation.

3.3.1 Introduction to the Bayesian approach

As presented in Section 3.2, the real-world problem of generation investments has been modelled as a simulator that is sophisticated and computationally expensive, particularly when there are a large number of stochastic variables taken into account [4, 82, 97] (see Section 4.1.3 for details). Variation or uncertainty in the input values of x propagates into output uncertainty, resulting in a range of projections of

generation capacities and LOLEs. It is important to manage uncertainty in order to provide model users (e.g., policymakers and investors) with the whole picture of model outputs. A sufficient management of uncertainty ensures the proper applications of LTGI models, such as probabilistic predictions of generation projections and robust design of energy policy. The computational complexity of the simulator makes it difficult to perform enough runs to get a dense coverage of the input space and therefore it is impractical to carry out a comprehensive UA and SA.

An alternative is to build a simplified model (a metamodel) as an approximation to the simulator behaviour in order to enable efficient prediction, UA and SA. One example in the area of energy modeling is the use of a non-Bayesian statistical method – multivariate adaptive regression splines, in combination with the LP-based bounding method in representing the input-output relationship of an integrated planning model [149]. Like Bayesian emulation, multivariate adaptive regression splines provides policy analysts a direct view of the multidimensional surface (response surface) of the model as a function of selected inputs. Uncertainty in the response arising from input assumptions is characterized in the form of bounds by the bounding method using multivariate adaptive regression splines, as opposed to by the probability distribution using Bayesian emulation. Bayesian emulation is more general than the metamodelling approach adopted in [149]. Firstly, the approach of multivariate adaptive regression splines is mainly applicable to linear or nonlinear convex optimization models, because the algorithm is heuristic so it does not guarantee optimal parameter estimates when solving a nonconvex optimization problem that may have many local, non-optimal minima. The bounding method proposed in [149] uses mathematical properties of linear programs, and so it cannot be applied to some types of power system planning models, such as those based on system dynamics, mixed integer programming and stochastic programming as reviewed in Section 2.2.3. In addition, Bayesian emulation allows for an explicit incorporation of prior information and consideration of structural and functional uncertainty that are failed to consider in [149].

For the purpose of exploring uncertainty regarding the model output propagated from all major sources of uncertainty, namely, input uncertainty, structural

uncertainty and numerical uncertainty as explained in Section 2.3, Bayesian emulation is employed to enable very efficient calibration, predictions [126], UA [127] and probabilistic SA [137]. Bayesian emulation is based on a statistical meta-model (the emulator) that is built to approximate the LTGI model and to quantify the uncertainty in the approximation. Within the Bayesian framework, all sources of uncertainty are quantified through probabilities [133].

Bayesian statistics takes a much broader definition of probability than frequentists. The frequentist statistics defines the probability as the limit of the frequency of the trials when an event is repeated for a large number of times [133]. The uncertainty in the repeatable event, called aleatory uncertainty, arises from its intrinsic randomness and unpredictability [133]. Examples of aleatory uncertainty include the outcomes of tossing dice and getting a full house in poker [150]. However, most of the uncertain quantities of our interest, such as the population of the city of London in year 1900, are not repeatable. This kind of non-repeatable uncertainty, named as epistemic uncertainty [133], is due to our lack of knowledge or data and in principle it might be reduced by gathering more information (e.g., referring to a reference book or historical data). To encompass epistemic uncertainty, in Bayesian statistics, probability is interpreted as the degree of belief, which is sometimes referred to as personal probability or subjective probability. Hence, uncertainty in some internal model parameters which is categorized into epistemic uncertainty can be managed by Bayesian approaches.

A fully Bayesian approach based on Gaussian Processes (GPs) is employed in our work, in the sense that the (*prior* and *posterior*) beliefs about model inputs, the model structure, and the model outputs of interest are all treated as uncertain and described through probability distributions. The prior distribution formulates our prior beliefs about the value of some parameters that are uncertain before we observe the data. The prior distribution can often be suggested by experts. The posterior distribution specifies our updated beliefs about the parameters after we observe the data. The posterior is achieved according to the Bayes' rule, which combines the prior with the observed data. A Bayesian emulator based on a GP gives a parametric representation between the (joint) probability distribution of model inputs and that

of model outputs. That is, a complete probability distribution instead of moments (e.g., mean, variance, skewness) or scenarios are used to characterize model outputs.

In the existing literature, a small number of scenarios are usually used for modelling input uncertainty in LTGI models. In scenario-based approaches, the results from predictions, UA and SA largely depend on the selected scenarios. Therefore, the choice of scenarios needs to be representative and qualitative. A simple technique is to choose the high, medium and low levels of a relevant quantity based on experience or modelling analysis. Some advanced scenario sampling techniques, such as importance sampling and scenario reduction [151, 152] can be used to limit the number of selected scenarios while preserving as much statistical information as possible. However, these scenarios levels will neither be exhaustive nor cover credible worst cases extensively, as opposed to a probability distribution. In addition, even without explicit probability judgments, choices of scenarios include implicit probability judgments in terms of what is credible. Comparing a probability distribution with scenarios, model results may well be less sensitive to precise choice of probability judgments than to precise choice of a small number of scenarios.

The aim of Bayesian emulation is to evaluate the function (simulator) f given in (3.7) at a small number of carefully configured input points, and to approximate this function as accurately as possible with a statistical representation (emulator) $\tilde{f}$. The emulator is computationally less demanding to evaluate than the simulator. The emulator encodes our uncertainty in the value of the function $f(x)$ where it has not been evaluated. That is, the emulator forms a probability distribution over the simulator and for the simulator outputs at new test points. This uncertainty information is useful in making more robust predictions on new test points.

The Bayesian approach is a two-stage approach, involving emulation of the simulator's response in the first stage and calibration and UA/SA using the emulator in the second stage. The application of Bayesian emulation to the LTGI problem seeks to estimate the relative credibility of different future system outcomes of interest (e.g., investment projections or system adequacy) in a quantitative manner.

3.3.2 Emulation using a Gaussian process

A deterministic simulator is represented by the function $f(\cdot)$ in (3.7), with the simulator inputs comprising the function's argument and the simulator output(s) comprising the function value. $f(\cdot)$ is treated as an *uncertain function*, as the value of $f(x)$ for any value of x is unknown until the simulator is run at x. In the Bayesian approach, our prior uncertainty on the function $f(\cdot)$ is modelled with a GP $\tilde{f}(\cdot) = \mathcal{GP}(\cdot, \cdot)$, with mean function $\mathrm{E}[f(x)|\beta] = m(x)$ and covariance function $\mathrm{Cov}[f(x), f(x')|\lambda, \gamma] = \lambda^{-1}c(x, x')$ [133], so that,

$$f(x)|\beta, \lambda, \gamma \sim \tilde{f}(x) = \mathcal{GP}\Big(m(x); \beta, \lambda^{-1}c(x, x'); \gamma\Big), \quad x, x' \in \mathcal{X}, \tag{3.8}$$

where x, x' are any two points over the standardised input space $\mathcal{X}$; $m(x); \beta$ denotes that $m(x)$ has parameters of a vector of regression coefficients β if a regression function is assumed; λ is a vector of unknown scale hyperparameters; $c(x, x')$ is the correlation function of parameters γ; γ is a diagonal matrix of N_I positive correlation hyperparameters $\gamma := \{\gamma_1, \ldots, \gamma_{N_I}\}$ for each input dimension; and $c(x, x')$ provides spatial correlation across $\mathcal{X}$ using a positive-definite function such that $c(x, x) = 1, \forall x$.

Definition: $\tilde{f}(x)$ is a GP if for any finite subset of input data $x^{(1)}, \ldots, x^{(m)} \subset \mathcal{X}$, the marginal distribution over that finite subset of random variables $\tilde{f}(x^{(1)}), \ldots, \tilde{f}(x^{(m)})$ has a multivariate Gaussian distribution,

$$\begin{bmatrix} \tilde{f}(x^{(1)}) \\ \vdots \\ \tilde{f}(x^{(m)}) \end{bmatrix} = \mathcal{N}\left(\begin{bmatrix} m(x^{(1)}) \\ \vdots \\ m(x^{(m)}) \end{bmatrix}, \begin{bmatrix} \lambda^{-1}c(x^{(1)}, x^{(1)}) & \cdots & \lambda^{-1}c(x^{(1)}, x^{(m)}) \\ \vdots & \ddots & \vdots \\ \lambda^{-1}c(x^{(m)}, x^{(1)}) & \cdots & \lambda^{-1}c(x^{(m)}, x^{(m)}) \end{bmatrix}\right).$$

According to the definition, a GP is a statistical distribution over functions with a continuous domain. This means that at any input point where the simulator is not evaluated, uncertainty around the output evaluated by the emulator is modelled as a one-dimensional Gaussian distribution (which is the marginal distribution at that point) conditional on the emulator parameters.

As discussed in 3.2.2, both the input space $\mathcal{X}$ and the output space $\mathcal{T}$ are a space of vectors in the simulator under study. To deal with a vector output, one simplified approach is to consider the time index as a new input to the model and

use a univariate emulator with a stationary, separable covariance structure. The covariance structure is separable in input x and time t, so that

$$\mathrm{Cov}[f(x,t), f(x',t')] = \mathrm{Cov}[f(x), f(x')]\,\mathrm{Cov}[f(t), f(t')], \tag{3.9}$$

where the covariance function $\mathrm{Cov}[f(t), f(t')]$ depends on planning years t, t' and some additional hyperparameters.

By assuming a stationary and separable covariance structure, the computational demand is alleviated as the dimension of the covariance matrices to be inverted is reduced [153]. However, a separable covariance structure is lack of flexibility in accounting for interactions between different types of correlations and it implies the conditional independence of outputs [147,153]. A multivariate GP with nonseparable covariance functions is generally the best option in general multi-output problems, yet this option may make it infeasible to solve large-scale realistic problems. In our work, a very general dimension-reducing technique, principal component analysis as described in [147], is adopted for defining a new, orthogonal basis for a set of multivariate data. In the new representation, the outputs are transformed and treated as independent, using many univariate (single-output) emulators. In this way, the approach of emulating an univariate output described in [133] can be generalized for emulating multivariate outputs.

The above GP in (3.8) specifies our prior beliefs about the properties of the unknown function $f(\cdot)$ we are modelling. The forms of the prior mean $m(x)$ and covariance function $\lambda^{-1}c(x,x')$ in a univariate GP used here are given by,

$$\begin{aligned} m(x) &= 0, \\ \lambda^{-1}c(x,x') &= \lambda^{-1}\prod_{i}^{N_I}\gamma_i^{4(x_i-x_i')^2} \\ &= \lambda^{-1}\exp(-\sum_{i=1}^{I}(x_i-x_i')^2\rho_i) \end{aligned} \tag{3.10}$$

where the index i denotes the i-th element of the input vector x, and $\rho_i := -4\ln\gamma_i$ is the spatial correlation parameter between the simulation output and the input parameter i.

The prior mean of the GP model is zero because the output is standardised and represented via principal components (see Section 3.3.2), so that the transformed

output in a new representation can be modelled with a mean of 0 and a marginal variance close to 1. Gamma prior distribution $\Gamma(5, 5)$ is specified for the marginal precision λ, and independent Beta prior distributions $\text{Be}(1, 0.1)$ are assigned to the spatial dependence parameters γ_i. The Gamma prior for λ is chosen so that the marginal variance for each GP process is close to 1 with standardised simulator outputs. The Beta prior for γ_i gives substantial prior mass near 1, which makes the correlation parameter ρ_i close to 0; this reflects our prior assumption that the output has a low dependence on each input parameter. The justification for choosing these priors can also be found in [147].

Given the prior GP model (3.10), the following three steps are taken to develop the emulator:

Step 1: Defining the standardised input space of interest through prior knowledge of these parameters; and selecting a small set of well designed input configurations, known as design points, $D := [x^{(1)}, x^{(2)}, \ldots, x^{(d)}]$ of d elements.

Step 2: Running the simulator $f(\cdot)$ at each of these design points, and obtaining the simulator output, $f(D) := (f(x^{(1)}), f(x^{(2)}), \ldots, f(x^{(d)}))^T$ of d elements.

Step 3: Fitting an emulator by combining the training data $(D, f(D))$ with the prior GP model given in (3.8).

In step 3, before fitting an emulator, the input data are normalised to the range [0,1] by subtracting the mean from the data and dividing by the range of the data; and the output data are standardised by subtracting out the mean and dividing by the standard deviation of the data. The prior GP model is updated in the presence of training data by means of a likelihood function, that relates our prior assumptions of hyperparameters to training data based on Bayesian inference, as explained later in Section 3.3.2. This leads to an updated distribution (i.e., a posterior GP model) that can be used, for example, for predicting new test points. In particular, for any point x', the posterior distribution of the vector output $[f(x')|D, f(D), \lambda, \gamma]$ is a multivariate Gaussian distribution according to the definition of GP (for more details on the estimation of $\{\lambda, \gamma\}$ via an MCMC scheme, we refer to [154]).

Validation is required to ensure that an emulator is sufficiently accurate. We expect a good enough estimate given the training data, and that the uncertainty

assessment associated with predictions of function values are reasonable. By replacing the simulator with the emulator, we are able to proceed with Bayesian inference to calibrate the emulator and to carry out MC simulations for UA and SA. This means that computational tasks such as calibration, UA and SA can be carried out efficiently based on the emulator.

Design

Observing runs of the simulator (*i.e.*, the training data) is the first step to build an emulator, as they are used for learning and estimating hyperparameters of the emulator. The set of points in the space of simulator inputs at which the simulator is run is called the design points. Good input configurations to run the simulator at are important for training the emulator in order to get better results.

As suggested by experts, good design points must be spread over the plausible ranges. Given a 6-dimensional input space, for instance, every possible combination of an extreme low (i.e., 0.005-quantile of its prior distribution), medium (i.e., 0.5-quantile of its prior distribution) and an extreme high (i.e., 0.995-quantile of its prior distribution) value for each input might be tried, resulting in $3^6 = 729$ runs of the simulator which can go beyond our computational ability when each model run takes several hours or days. A major concern with this design is that even the 3^6 combinations of the 6 inputs would give a poor coverage across the whole prior parameter space because the design gives much more extreme points for inputs at which we are less likely to predict. However, a uniform coverage of the input space is needed by the emulator in order to predict well for a realistic range of inputs.

To reduce the number of input configurations while ensuring reasonable space-filling properties, a maximin Latin hypercube design is employed here. In a Latin hypercube, there are as many equally distributed levels of each factor (i.e., each input variable) as the number of needed or desired runs in the design, which is also one benefit of this design [155]. Alternatives to Latin hypercubes in experimental designs for computer experiments include orthogonal arrays, Hammersley designs and a combination of different designs [156, 157]. Latin hypercube is chosen because it is simple to implement and is flexible enough to provide data capable of cover small

and large design spaces reasonably well compared to other designs [157]. To get improved space-filling properties, the maximin Latin hypercube design maximises the minimum distance between the design points, subject to that the levels of all factors are evenly spaced from the lower bound to the upper bound [158]. This maximin design can be done by generating a random Latin hypercube and then permuting the entries in each column such that the maximin criteria is satisfied. However, in practice, there is no guarantee of obtaining a globally optimal design when the exhaustive search is not feasible.

To proceed with the Latin hypercube design, one must decide the number of sample points to use. There is no general rule of thumb regarding the relationship between the sample size and the number of inputs because it really depends on the problem. Validation is necessary to demonstrate whether or not the emulator is a good approximation to the simulator (see Section 3.3.3 for more detail).

Dimension Reduction

As described in Section 3.2.2, the simulator output of interest $y := f(x)$ is a vector of $\mathcal{T}$ elements. To cope with the high-dimensional model output space, principal components analysis (PCA), described in [159] is used to project the high dimensional output data into a new lower dimensional representation of the data that contains most of the variance in the data with minimal loss of information. The principal component basis vectors $K_f = [k_1, \ldots, k_{p_f}]$ are are an uncorrelated orthogonal basis set. They can be obtained via singular value decomposition of the standardised simulation output matrix $f(D)$. Each basis vector is scaled so that the output of $\nu_p(x)$ has a mean of 0 and a marginal variance close to 1.

Based on PCA, the $\mathcal{T}$-dimensional simulator output $f(x)$ is modelled using a p_f-dimensional basis representation [147]:

$$f(x) \sim \sum_{p=1}^{p_f} k_p \nu_p(x), p = 1, \ldots, p_f, \tag{3.11}$$

where $\nu_p(x)$ are mean 0 independent GP models as formulated in (3.8) with their priors given in (3.10).

The basis vector maps the data vector from an original space of $\mathcal{T}$ variables

which are possibly correlated to a new space of p_f variables which are uncorrelated over the dataset. Therefore, the problem of building an emulator that maps $[0,1]^{N_I}$ to $\mathbb{R}^{\mathcal{T}}$ is reduced to building p_f independent, univariate GP models for each $\nu_p(x)$.

Bayesian inference

Bayesian inference is the basic technique for developing and utilizing an emulator. Bayes' theorem is used in Bayesian inference in order to provide the posterior probability distribution of the model parameters rather than so-called point estimates. Bayes' theorem is written as,

$$P(\theta|D) = \frac{P(\theta)P(D|\theta)}{P(D)}, \tag{3.12}$$

where $P(\theta|D)$ is the posterior probability density of θ conditional on available data, $P(\theta)$ is the prior probability density of parameter θ, $P(D|\theta)$ describes the likelihood - the probability density of observing the data D given the parameter value θ, and the denominator $P(D)$ is the evidence or marginal likelihood - the total probability of the observed data. The probability $P(D)$ can be simply viewed as a normalizing constant, and hence Bayes' theorem can be expressed as,

$$P(\theta|D) \propto P(\theta)P(D|\theta), \tag{3.13}$$

that is, the posterior probability is proportional to the prior probability times the likelihood.

For the sake of illustrating Bayesian inference, an example of inference on a binomial parameter is presented here; this example is adapted from the example presented in [160].

Example 2. (*Inference on a binomial parameter*). Suppose we have been given data M consisting of a series of m coin flips which contain r positive trials. The data were assumed to be generated by a sequence of independent draws from a Bernoulli distribution with parameter θ_b, which is the probability model of flipping Heads. Let $M_i = 1$ if flip i was Heads, and $M_i = 0$ otherwise. Let $m_H = \sum_{i=1}^{m} M_i$ be the number of heads in m tosses. Then, the likelihood model is $P(M|\theta_b) = \theta_b^{m_H}(1-\theta_b)^{m-m_H}$. Suppose that prior knowledge about θ is described by a Beta

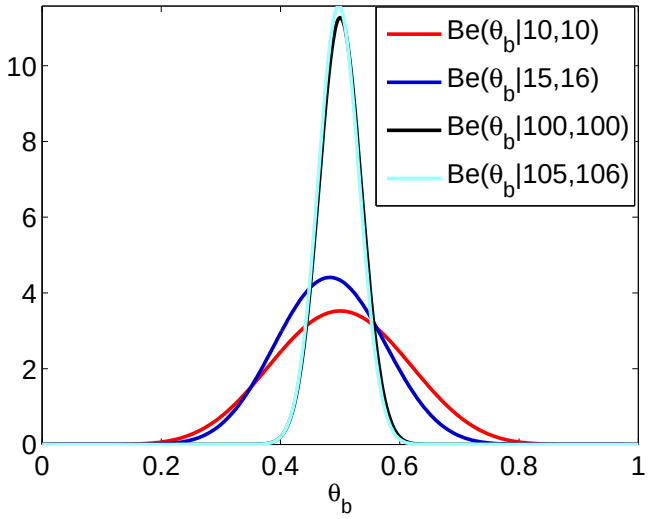

Figure 3.6: Prior and posterior density functions of the parameter θ_b in the probability model of flipping Heads.

distribution $\text{Be}(\theta_b|\alpha_b, \beta_b)$, so that $P(\theta_b|\alpha_b, \beta_b) \propto \theta_b^{\alpha_b - 1}(1-\theta_b)^{\beta_b - 1}$. By varying α_b and β_b, a wide range of possible prior beliefs of θ_b can be obtained.

Applying Bayes' theorem 3.18 gives the posterior density of θ_b as the Beta distribution $\text{Be}(\theta_b|m_H + \alpha_b, m - m_H + \beta_b)$,

$$\begin{aligned} P(\theta_b|m_H, m, \alpha_b, \beta_b) &\propto \theta_b^{m_H}(1-\theta_b)^{m-m_H}\theta_b^{\alpha_b-1}(1-\theta_b)^{\beta_b-1}) \\ &\propto \theta_b^{m_H+\alpha_b-1}(1-\theta_b)^{m-m_H+\beta_b-1} \end{aligned} \tag{3.14}$$

Suppose available information on the parameter θ_b in the probability model of flipping Heads is described by a Beta distribution $\text{Be}(\theta_b|10, 10)$, so that it is judged to be equally likely that the coin flip would be Heads in 20 coin tosses. A random trial of size 11 is conducted, where only 5 flips were Heads. Using the results above, the corresponding posterior distribution is then $\text{Be}(\theta_b|15, 16)$. Fig. 3.6 plots the prior and posterior densities of θ_b. It can be seen from Fig. 3.6 that the initial uncertainty in the value of θ represented by the prior density plot (in red) is significantly reduced by the data, resulting in the posterior density plot (in blue).

Assume a more informative prior, such as a $\text{Be}(\theta_b|100, 100)$ distribution of θ_b. Given the same 11 coin tosses, the corresponding posterior distribution, $\text{Be}(\theta_b|105, 106)$ (in green line in Fig. 3.6, is very different from $\text{Be}(\theta_b|15, 16)$. With the same data set available, the difference in the resultant posterior distribution in the two case

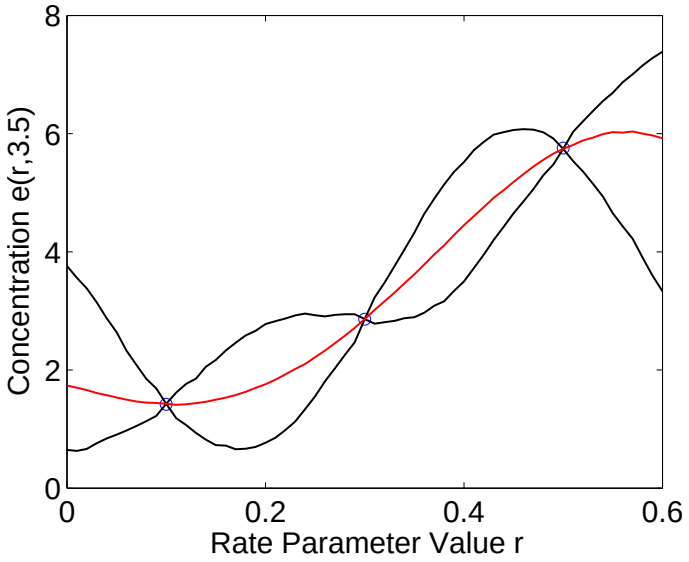

Figure 3.7: The emulator of the concentration function of a chemical $e(r, t = 3.5)$ developed by using 3 training data.

lies in that in the second case the prior is stronger so the data has less effect on it. When there is less data, the posterior becomes more influenced by the prior.

An illustrative example of the emulator

Once an emulator is built, it can be used to make predictions at new input points. Below is an example to illustrate that the emulator quantifies the uncertainty in the output where the simulator is not evaluated, and the quality of the emulator is affected by the number of training data.

Example 1. (*The concentration of a chemical, continued*). The toy example given in Section 2.4 Chapter 2 is used here to explain how the emulator works as an approximation of the simulator.

Three emulators are respectively developed for approximating the f unction $e(r) = \exp(3.5r)$ over $r \in [0.0.6]$ using the three, four and five function evaluations (e.g., training data) given by the blue circles, as shown in figs. 3.7–3.9. In each of the three figures, the purple line gives the true function $e(r)$, the red line gives the posterior mean GP estimate, and the black lines quantify uncertainty bounds of the GP emulator, namely, 5th and 95th percentiles of the emulator output, given the available training data.

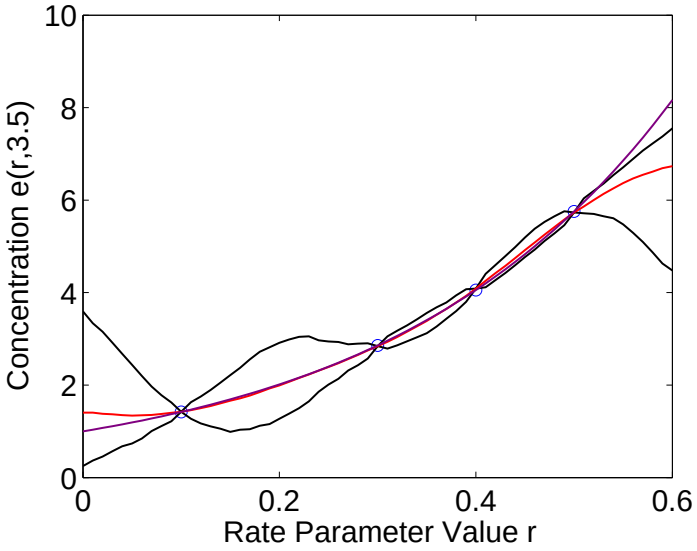

Figure 3.8: The emulator of the concentration function of a chemical $e(r, t = 3.5)$ developed by using 4 training data.

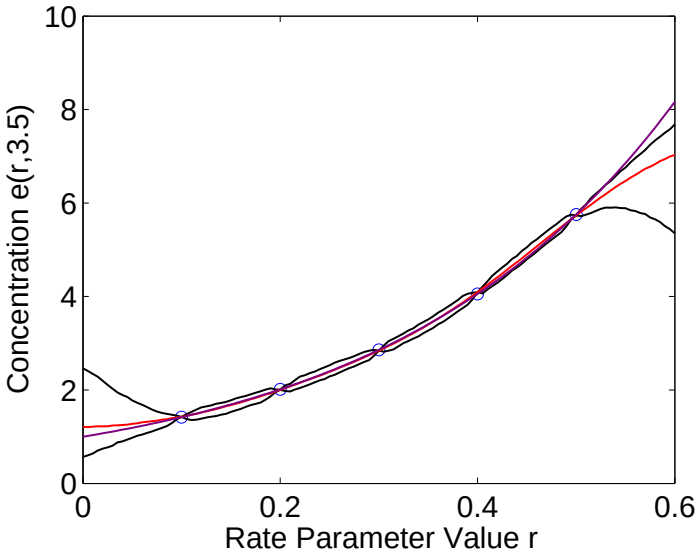

Figure 3.9: The emulator of the concentration function of a chemical $e(r, t = 3.5)$ developed by using 5 training data.

It is clearly seen that the emulator reproduces the training data at the design points with no approximation error. The black lines are formed by draws from the emulator (the posterior GP) at a large number of selected values of the input parameter r, integrating (by sampling from the posterior distributions of) the hyperparameters of the emulator. A comparison made among the three figures reveals that the credible interval narrows, that is, the accuracy of the emulator can be improved with an increasing number of training data. In Fig. 3.8, the purple line plots the analytical expression of the model $e(r, t = 3.5) = \exp(3.5r)$.

Note that in Fig. 3.8 at the rate parameter region of $0.53 - 0.6$, the simulator (the purple line) rises beyond the credible interval quantified by the emulator. This is because that parameter region is beyond the range of training data. An emulator, like any statistical model, is not guaranteed to fit anywhere in the input space. In the region of the input space where there are no training data, the fit of the emulator will depend heavily on the priors and the model assumptions chosen, and these may not be reflective of reality. Hence, it is important to carefully select the design points at which the model is evaluated and to do validation checks on an emulator.

3.3.3 Validation of an emulator

The emulator exactly reproduces the training data because it has been fitted using that data. Hence, the training data cannot be used for validation. The aim of validation is to test the ability of the emulator to predict at untried input points.

Our validation diagnostics are based on comparisons between emulator outputs and simulator runs for a new data, called as test points. The procedure of validating a GP emulator is described as follows.

Step 1: Select a small set of well designed input configurations from the input space, known as validation sample designs, $V := [x^{*(1)}, x^{*(2)}, \ldots, x^{*(N_v)}]$ of N_v elements using an optimised Latin hypercube design procedure. When the simulator is slow to run, it is necessary to keep the size of the validation sample as small as possible, although it might not be a good thing because ideally there would be a large validation sample.

Step 2: Run the simulator at each of the test points to produce the output vector

$f(V) := (f(x^{*(1)}), f(x^{*(2)}), \ldots, f(x^{*(N_v)}))^T$.

Step 3: Evaluate the emulator's predictions for $\tilde{f}(V)$ at the test points and obtaining the means, variances, covariances and 95% credibility intervals for the purposes of our diagnostics.

Step 4: Apply validation diagnostics as an indicator of a validation failure.

A number of validation diagnostics are provided in [161], including the Mahalanobis distance, the pivoted Cholesky decomposition, eigen decomposition, and graphical methods such as credible interval diagnostic (plots of credible intervals of the emulator predictions against the simulator predictions), plots of the individual errors against the emulator predictions, plots of the errors against the index, and plots of error against inputs.

The main diagnostic used in our work is credible interval diagnostic, which plots the predicted values by the emulator with 90% or 99.8% credibility intervals, against the simulator outputs at each validation point. The 90% credible interval is quantified by 5th and 95th percentiles of the emulator output, and the 99.8% credible interval is quantified by 0.01th and 99.9th percentiles of the emulator output. It is expected to see approximately 90% of simulator outputs lie within the 90% credible interval; this is an approximate proportion due to the randomness of the emulator output Having some simulator outputs falling outside of the 99.9% credibility intervals of of the emulator outputs is a clear evidence of failure. If there are no indications of conflict across such diagnostics within the test set that has been used, we then have confidence to say that the emulator represents the simulator accurately. A second diagnostic is to plot the standardised errors (ratio of residual to mean of simulation output) against the index (the planning year), as will be shown in Section 3.3.3. A third diagnostic is to calculate the root mean squared errors of the predictions, which is a widely used indicator of accuracy in computer experiments [156].

3.3.4 Bayesian calibration

The goal of model calibration here is to identify plausible values of calibration parameters whilst simultaneously inferring the model discrepancy using physical

observations of the output over the time period $\mathcal{P}$. Among the inputs $x := \{u, \theta, \omega\}$ described in Section 3.2.2, $\{u, \omega\}$ are already known historically. The calibration parameters θ are assumed to have unknown best values due to our incomplete knowledge of the real-world. If the simulator were run with these best values, it would reproduce the observations plus a model discrepancy term δ plus observation errors if any. With the emulator $\tilde{f}(\cdot)$ as an approximation of the simulator, the relationship between the observations, the model discrepancy and the emulator at the best value of θ can be written as [127],

$$y_{obs} = \tilde{f}(u, \theta, \omega) + \delta + \epsilon, \tag{3.15}$$

where $y_{obs} := \{y_{obs,1}, \ldots, y_{obs,N_P}\}$ is the single time series of historical thermal capacity over the past planning horizon $\mathcal{P}$.

The model discrepancy δ quantifies the mismatch between the model and the observations at the best setting for the calibration parameter θ. The mismatch may arise from inadequacies in the simulator, such as in the model equations, model structure or logic [127]. δ is modelled by a linear combination of basis functions [162]:

$$\delta = \sum_{j=1}^{p_\delta} d_j \vartheta_j, \tag{3.16}$$

where d_j's are basis functions; and the weights ϑ_j's are modelled as independent GP priors over x.

Here the basis functions are independent Normal kernels that are separated along the t direction; this modelling technique is driven by the expectation that the discrepancies, if they exist, should have a strong time persistence. The number of basis functions and the kernel width (or the standard deviation of the kernel) are often specified by the model users; they are chosen so that the kernels of the discrepancy are well separated along the time axis. Therefore, their specification depends on the application (see Section 4.2).

The weights of the discrepancy kernels are modelled as independent mean 0 GP priors,

$$\vartheta_j | \gamma_{\vartheta j}, \lambda_{\vartheta j} \sim \mathcal{GP}(0, \lambda_{\vartheta j}^{-1} c(x, x'); \gamma_{\vartheta j}), j = 1, \ldots, p_\delta \tag{3.17}$$

s where the correlation function $c(x, x')$ is a function of parameters $\gamma_\vartheta := \{\gamma_{\vartheta j}\}, j =$

$1, \ldots, p_\delta$, and independent beta priors $\text{Be}(1, 0.1)$ are assumed for γ_ϑ and gamma priors $\Gamma(1, 0.0001)$ are for $\lambda_\vartheta := \{\lambda_{\vartheta j}\}$. The beta priors for the correlation parameters γ_ϑ assume low dependence between the discrepancy and each input parameter. The gamma priors for the precision parameters λ_ϑ are rather uninformative, which leads to a very small model discrepancy when the historical observations are uninformative, as explained in [147].

To find plausible values for the calibration parameters, θ, alongside inferring the model discrepancy, first suppose that the prior knowledge of the calibration parameters, the discrepancy parameters $\lambda_\vartheta, \gamma_\vartheta$ and the emulator parameters λ_p and γ_p is described by the joint prior distribution $P(\theta, \lambda, \gamma)$, where $\lambda := \{\lambda_p, \lambda_\vartheta\}, \gamma := \{\gamma_p, \gamma_\vartheta\}$. This prior distribution is updated using the observations and the set of training runs obtained from the simulator. This updating is the so-called Bayesian inference (described in Section 3.3.2), implemented according to Bayes' theorem [127]:

$$P(\theta, \lambda, \gamma | y_{obs}, f(D)) \propto P(y_{obs}, f(D) | \theta, \lambda, \gamma) P(\theta, \lambda, \gamma), \tag{3.18}$$

where the left hand term of (3.18) is the posterior distribution of the calibration parameters of interest and $P(y_{obs}, f(D)|\theta, \lambda, \gamma)$ is the joint distribution of the observed data and the training runs $f(D)$, conditional on these calibration parameters. The marginal posterior distribution for each of θ, λ and γ can be obtained through integration over the other parameters.

For more discussions and technical details regarding this Bayesian calibration approach see [127, 162].

3.3.5 Variance-based sensitivity analysis

Calibration helps reduce input uncertainty as described in the last subsection, while the output uncertainty can be reduced most efficiently if those inputs that influence the output most strongly are focused on. SA provides such a tool of exploring how much of the total output uncertainty is attributed to uncertainty in a particular input or a group of inputs, and hence identifying which input parameters are the most influential to the output variations.

SA takes various forms of studying the relationship between a simulator's in-

puts and outputs, such as variance decompositions, partial derivatives in local SA (see Section 3.3.5 for more details) and variance decompositions in variance-based SA. The variance-based SA will be used in our sensitivity analysis for two main reasons. Firstly, the variance-based SA, as a class of probabilistic SA, quantifies the input and output uncertainties as probability distributions, which is consistent with the uncertainty modelling in the GP-based Bayesian approach. Secondly, the variance-based SA is an attractive global SA because it explores the full range of the input space and accounts for nonlinear responses and interactions. To densely cover the whole input space, it may involve many thousands of simulator runs. The Bayesian approach enables the use of a very efficient emulator in replace of an simulator, which substantially reduces computational expense. Moreover, the Bayesian approach quantifies uncertainty where the simulator has not been evaluated, thus allowing understanding of consequences of using limited number of runs.

Variance-based measures of sensitivity prioritizes uncertainties by decomposing the output variance into fractions attributable to inputs and sets of inputs and measuring the sensitivity of the output to an input variable by the amount of variance in the output contributed by that input. One variance-based measurement is known as "main-effect index" that quantifies the contribution to the output variance of the main effect of a subset of inputs x_J, averaged over the joint distribution of all the other input variables x_{-J} [136], as expressed as,

$$S_{\mathrm{M_J}} = \frac{SV_{\mathrm{J}}}{\mathrm{Var}(y)}, \tag{3.19}$$

where

$$\begin{aligned} SV_J &= \mathrm{Var}_{x_J}\left(E_{x_{-J}}(f(x)|x_J)\right) \\ &= \mathrm{Var}_{x_J}\left(\int f(x)P_{-J|J}(x_{-J}|x_J)dx_{-J}\right), \end{aligned} \tag{3.20}$$

$$\mathrm{Var}(y) = \sum_{i=1}^{N_I} SV_i + \sum_{i<j}^{N_I} SV_{ij} + \cdots + SV_{12\ldots N_I}. \tag{3.21}$$

As shown in (3.21), the mean-effect measurement provides a decomposition of the output variation into terms relating to the main effects and various interactions between the input variables. In (3.20) $P_{-J|J}(x_{-J}|x_J)$ denotes the probability density

function of x_{-J} conditional on the value of x_J and $E(f(x)|x_J)$ denotes the expected value of the simulator output at x, conditional on a subset of input variables x_J, averaged over the joint distribution of all the other input variables x_{-J}. The effects of single input parameter and two-input interactions are shown in the following

$$SV_i = \mathrm{Var}(E(f(x)|x_i)) = \mathrm{Var}(\int f(x)p_{-i|i}(x_{-i}|x_i)dx_{-i}), \tag{3.22}$$

$$SV_{ij} = \mathrm{Var}(E(f(x)|x_{ij}))) = \mathrm{Var}(\int f(x)p_{-ij|ij}(x_{-ij}|x_{ij})dx_{-ij}). \tag{3.23}$$

The other variance-based sensitivity measurement is "Total-effect index" or "total-order index". It is often used when the number of input variables is large. "Total-effect index" measures the contribution to the output variance of x_i, including all variance caused by its interactions, of any order, with any other input variables x_{-i}. It is expressed as,

$$\begin{aligned} S_{T_i} &= \frac{E_{x_{-i}}\big(\mathrm{Var}_{x_i}(y|x_{-i})\big)}{\mathrm{Var}(y)} \\ &= 1 - \frac{\mathrm{Var}_{x_{-i}}\big(E_{x_i}(y|x_{-i})\big)}{\mathrm{Var}(y)}. \end{aligned} \tag{3.24}$$

The following steps are used to implement variance-based SA based on an emulator within the Bayesian framework.

- Quantify the uncertainty in each input using probability distributions that identify in detail how the inputs might be varied.
- Sample from the distributions of model inputs as design of experiments.
- Run the emulator plus the discrepancy term as an approximation to the simulator a number of times at sampled designs.
- Using the resulting model outputs obtained in the last step to calculate the sensitivity measurements of interest.

3.4 Chapter summary

This chapter has briefly presented the mathematical model of a computationally expensive LTGI model as the simulator. The relationship between the inputs and

outputs of interest represented by the simulator (i.e., the LTGI model) is approximated by a Bayesian emulator while adequately accounting for uncertainty in real-world applications. The emulator is capable of managing uncertainties arising from different sources, including its stochastic inputs, imperfect science and the limited number of evaluations of the LTGI model.

At any input point where the simulator has not been run, the prediction produced by the emulator is an approximation to the simulator evaluation. The emulator's evaluation is not a point estimate or a single scenario, but a probability distribution with uncertainty information that covers a range of plausible values or scenarios. Besides, this Bayesian method allows for calibrating uncertain model parameters and quantifying the model discrepancy, so that the probabilistic predictions made by the emulator are consistent with historical observations of the model output. Moreover, Bayesian emulation enables efficient UA for uncertainty quantification and variance-based SA for prioritizing uncertainty in model inputs according to their contributions to variations in model outputs.

Chapter 4

Case study on GB power system

In this chapter, a case study on the GB power system is presented, where Bayesian emulation is employed to enable calibration, prediction, uncertainty analysis (UA) and sensitivity analysis (SA) regarding a complex LTGI model.

Bayesian calibration reduces uncertainty in calibration parameters and quantifies the model discrepancy using available historical observations (e.g., historically installed thermal capacity in operation). Then, the model discrepancy is applied to making future projections on thermal capacities. Using the Bayesian method, policymakers can be provided with a realistic picture of the possible model outcomes with probabilities attached. The robustness of different electricity market designs (i.e., energy-only and capacity markets) against uncertainties (e.g., input uncertainty, structural uncertainty and functional uncertainty) is studied in terms of the probability of meeting a LOLE threshold given a scenario of market design parameters; this is achieved by performing an UA of the maximum LOLE over a future planning horizon. Finally, a probabilistic SA is conducted to identify the most influential inputs (a particular input or a group of inputs) to the variations of the model output of interest.

4.1 Data, assumptions and computational time

4.1.1 Data

The data used here are consistent with those provided in [4], including the initial capacity mix, wind and demand data, as well as financial and technical assumptions for generators. Our assumed prior distributions of the six model inputs are listed in Table 4.1.

Table 4.1: Summary of prior distributions of model inputs

Inputs	Prior	Unit	Description
u_{cone}	U(31.8, 66.3)	£/kW-year	Net cost of new entry
u_{voll}	U(1000, 30000)	£/MWh	Energy price cap
u_{co_2}	U(0.8, 1.2)	N/A	Multiplier to the central projection of carbon prices
ω_{gas}	N(1, 0.06^2)	N/A	Multiplier to the central projection of gas prices
θ_{VaR}	U(0.005, 0.55)	N/A	Investor's value-at-risk
θ_{markup}	U(0, 25)	GW	Energy price markup parameter

Expert knowledge can be incorporated in the model by assigning prior distributions to model inputs. When there is little knowledge about an uncertain parameter in the model, a uniform distribution is commonly used as a prior carrying little information, but with credible upper and lower bounds. The bounds of control and forcing inputs can be reasonably wide as weak priors over which we may wish to achieve the desired understanding through UA and SA.

The net CONE only takes effect under the capacity market design, and its prior range is consistent with the low, central and high estimates provided in [62]. The prior range of the energy price cap, u_{voll}, is referred to that of its benchmark - VOLL. According to the study in [51], $(1000 - 30000)$ is a reasonable range for VOLL that is estimated for domestic, industrial and commercial electricity consumers in GB. Hence, the energy price cap u_{voll} is assumed to sample from U(1000, 30000)£/MWh; a narrower prior range with a lower upper limit of the energy price cap (e.g., 20000£/MWh) may be chosen when the energy price cap is administratively set below the VOLL or when there is a lower estimate of VOLL. The risk attitude θ_{VaR}

is sampled from the uniform distribution U(0.5%, 55%) reflecting a range of investment assumptions from extremely risk averse to risk neutral. In the GB power system with 60 GW peak demand, θ_{markup} is sampled from U(0, 25) [GW]. The chosen range of θ_{markup} implies that the energy price uplift function is used under system conditions with a fairly tight capacity margin (range $0 - 41\%$). A prior uniform distribution U(0.80, 1.20) is assigned to u_{co_2}. In the case that $u_{co_2} = 1.05$, the reference trend of carbon prices, which takes DECC's central forecast [163], would be shifted upwards by 5%. The uncertainty range of future carbon prices resulting from the chosen range of u_{co_2} is broadly in line with the range of DECC's carbon projections in [163]. The prior belief for the forcing input ω_{gas} is a normal distribution $N(1, 0.06^2)$, which indicates bias over the reference gas price level, and results in a range consistent with that estimated by DECC [1].

4.1.2 Assumptions

In the application of Bayesian emulation to the LTGI problem, our main assumptions are as follows:

(1) Assume that known sources of uncertainty (e.g., observation errors) not included in the LTGI emulator are not likely to contribute much extra uncertainty, in comparison with the contribution of the three major sources of uncertainty (e.g., input uncertainty, structural uncertainty and functional uncertainty) that are accounted for;

(2) Assume that structural uncertainty obtained from model calibration is a good proxy for structural error to be applied in future projections;

(3) Assume that the model outputs depend upon a set of selected model inputs that are explained in Section 3.2.2 and these inputs are judged to be important determinants of the output of model simulations;

(4) Assume that the risk preference and the investment logic of a representative investor do not change over time in the long-term decision-making process.

In the GB case study, the robustness of two mainstream market designs, namely, energy-only and capacity markets will be compared in Section 4.4. Regarding the capacity demand curve adopted in the UK capacity market, as shown in Fig. 3.5 in Section 3.2.1, there are some main features and modelling assumptions according to some UK capacity market proposals [23, 164], which are as follows:

(1) A 4-year ahead auction for capacity will be held every year since year 2014 (e.g. an auction in 2020 would be for delivery in 2024/2025).

(2) If successful at auction, an existing plant will be awarded a one-year contract at the clearing price and new entrants will have access to fifteen-year contract;

(3) The capacity price cap is set as 1.5 times of the net CONE;

(4) The choice of the net CONE is based on the assumptions on the projected level of revenue from a capacity market agreement that is required to permit investment in new generation capacity.

(5) A 1.5 GW range is set above and below the capacity target level [23], reflecting an increasing appetite for capacity at lower prices. This capacity range together with the capacity price cap determines the slope of the capacity demand curve.

4.1.3 Computational time

The simulator was run in the Matlab/Simulink R2012a environment using an Intel(R) Core(TM) $i5-3470$ 3.20GHz processor with 8.00GB RAM. The run time for a single simulation of the 30-year generation planning varied between 140 and 600 minutes, with 100 MC simulations of the forward-looking years of operation for each investment decision. The Gassian Process Model/Sensitivity Analysis (GPM/SA) code package that was developed by Los Alamos National Laboratory[1] has been

[1]The Matlab code for implementing Bayesian emulation is available online from https://github.com/libqueso/gpmsa-matlab/tree/master/Examples

adapted in our work for carrying out emulation, UA and SA (as provided in Appendix A.1). In comparison with simulator runs, one emulator evaluation in the same environment took approximately 10^{-4} seconds, a speed ratio in the order of $10^7 \sim 10^9$. Note that the simulator output is a single scenario or a certain value while the emulator output is a probability distribution. The time needed for developing and validating an emulator includes that for obtaining training data and validation data by running the simulator, and that for fitting the emulator to the training data which takes about 6 minutes.

Traditional MC-based probabilistic SA that is directly applied to the simulator would take several months or even over one year for a thousand simulations, which would be required to give a sense coverage of the input space. However, with the developed emulator, sensitivity analysis can be achieved within several seconds. The advantage of the emulator-based approach in saving computational time is clearly seen. It is noted that the simulated simulator outputs are deterministic as explained in Section 3.2.2, while the emulator outputs are probabilistic with uncertainty bounds because for any input, the emulator outputs is modelled by a multivariate normal distribution conditional on the emulator parameters.

4.2 Emulation, validation and calibration

It will be shown here how the simulator can be emulated and validated, and then calibrated against historical observations of the model output. In practice a good simulator is needed in that it is able to reasonably well reproduce the dynamics of observations (at least for some parameter values).

4.2.1 Emulating the long-term generation investment model

Prior to carrying out model calibration against the observations of installed thermal capacity, an emulator is needed for approximating the relationship between the calibration parameters and the installed thermal capacity. Since the introduction of NETA in 2001, an energy-only market has been implemented in the GB power system. Before 2001, the Pool market in England and Wales included capacity

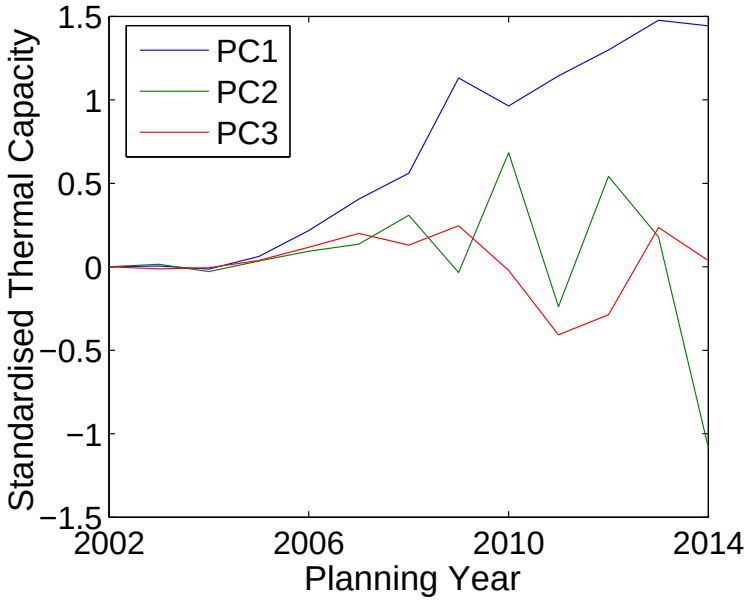

Figure 4.1: Three principal components of the standardised simulation data (historical thermal capacities in operation).

payments. To ensure that the LTGI model keeps the same energy-only structure during the simulation process, the model is run from year 2000 to year 2014 for calibration purposes. During the past time period, the energy price cap was set as 10000£/MWh. The long-term gas price level ω_{gas} was set as 1, implying that historical gas prices have been chosen during the decision-making process.

The emulator is built using 12 training data that are composed of 12 design points over the two-dimensional input space $(\theta_{VaR}, \theta_{markup})$ and the corresponding 12 scenarios of annual outcome of the total thermal capacity over the planning horizon $\mathcal{P}$. The observation data y^G_{obs}, against which the simulator is calibrated, consists of a single time series of total thermal capacity from year 2003 to year 2014. Note that the observations at years $2001 - 2002$ are omitted from calibration because the investment decisions do not take effect until 2003 due to construction delays associated with thermal power plants. Fig. 4.1 shows 3 principal components used in capturing 98% variability of the time series of standardised thermal capacity. Note that the first two principal component (in blue and green lines) has much bigger vertical scale than the third one (in red line); this confirms that most of the variation in the simulation data is captured by the first two principal components.

Note that the only model output for calibration is the time series of total thermal capacity in operation. For comparison purposes, Fig. 4.2 gives a breakdown of simulated (in lines) and observed (in circles) installed thermal generation capacities at all design points, and also historically observed on-shore and off-shore wind capacities. Only 0.5 GW new nuclear plant was built in year 2014 at one of the 12 simulated scenarios, and no new coal plant was chosen to be invested across all scenarios. The top two graphs show that the simulated nuclear and coal capacities are very close to the observations. The most attractive generation type is CCGT, which is revealed by both simulation results and observations. In 2015, there is a significant decline in the amount of OCGT across all simulated scenarios; this is due to plant retirements. The input data of onshore and offshore wind capacities are obtained from rounding the actual historical data. About 9 GW solar capacities had been installed by 2015 but they are not considered in our simulation. These approximations do not affect the simulation results much, because wind and solar have small capacity values [9, 10, 165] and they have insignificant contribution to power system reliability when the total capacity is low.

4.2.2 Validation results

Apart from the training data, 6 additional model runs on a maximin Latin hypercube design are used for validation. Fig. 4.3 shows the 12 design points (in red circles) and the 6 test points (in blue stars) that are sampled from the prior ranges of two calibration parameters, respectively. Each axis is divided into the same number of intervals as the desired sample size, and the points (between which the minimum distance is maximised) have been sampled within those intervals.

The validation procedure described in Section 3.3.3 is carried out.

Each graph in Fig. 4.4 presents the reference plausible range (probabilistic prediction) of the emulator's prediction compared with the evaluation of the simulator at a validation point. A common indicator of a validation failure is to have significantly more than about 10% of simulator's evaluations lying outside of the 90% credible interval (the 5th and 95th percentiles) of the reference distribution, or to have significantly more than 0.2% simulator's evaluations falling outside of the 99.8%

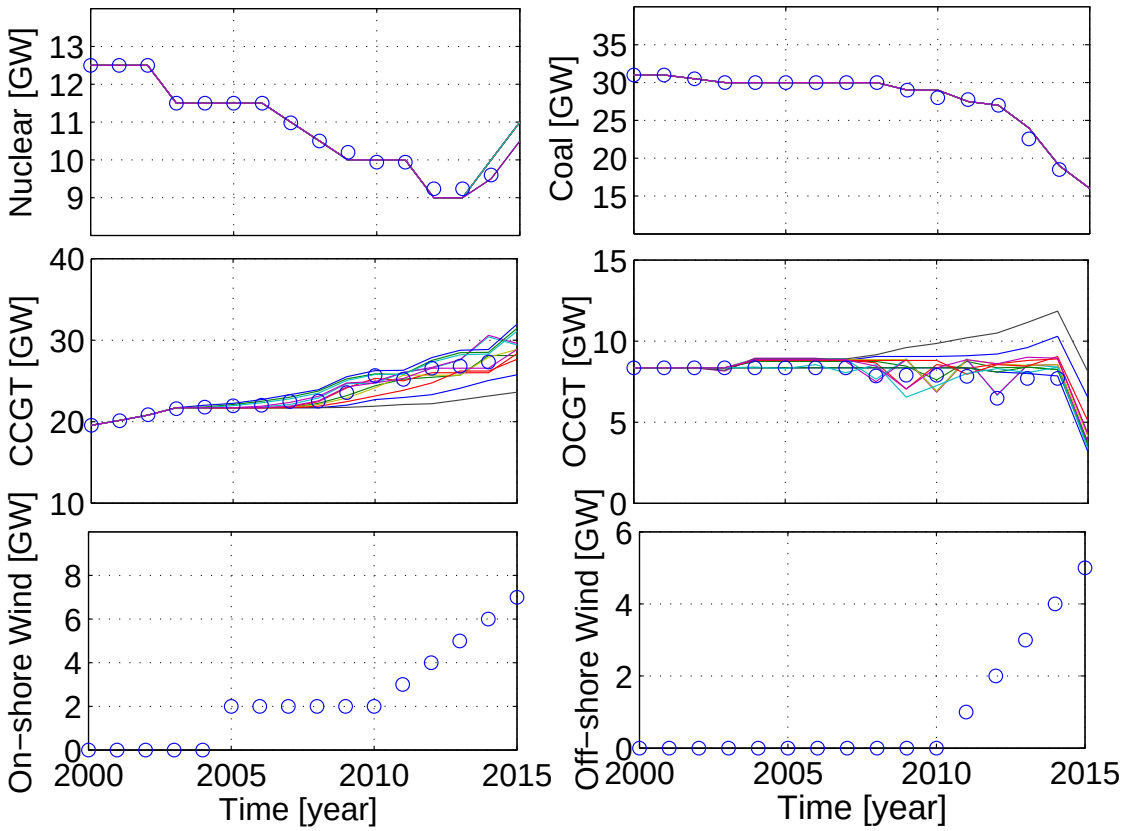

Figure 4.2: Simulated (in colored lines) and observed (in circles) capacities of each generation technology in operation over the past planning horizon.

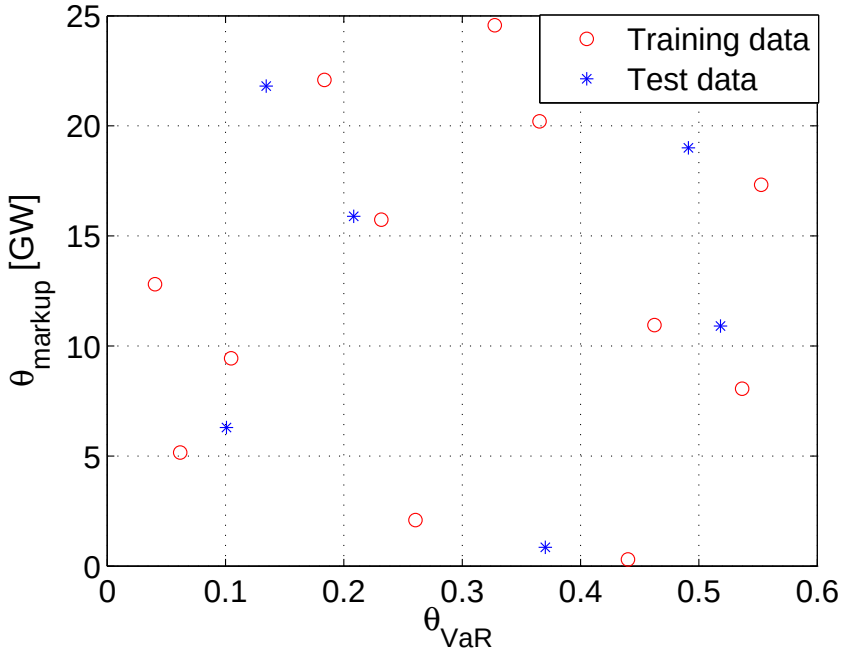

Figure 4.3: Distribution of training data and test data in a 2-dimensional Latin hypercube design.

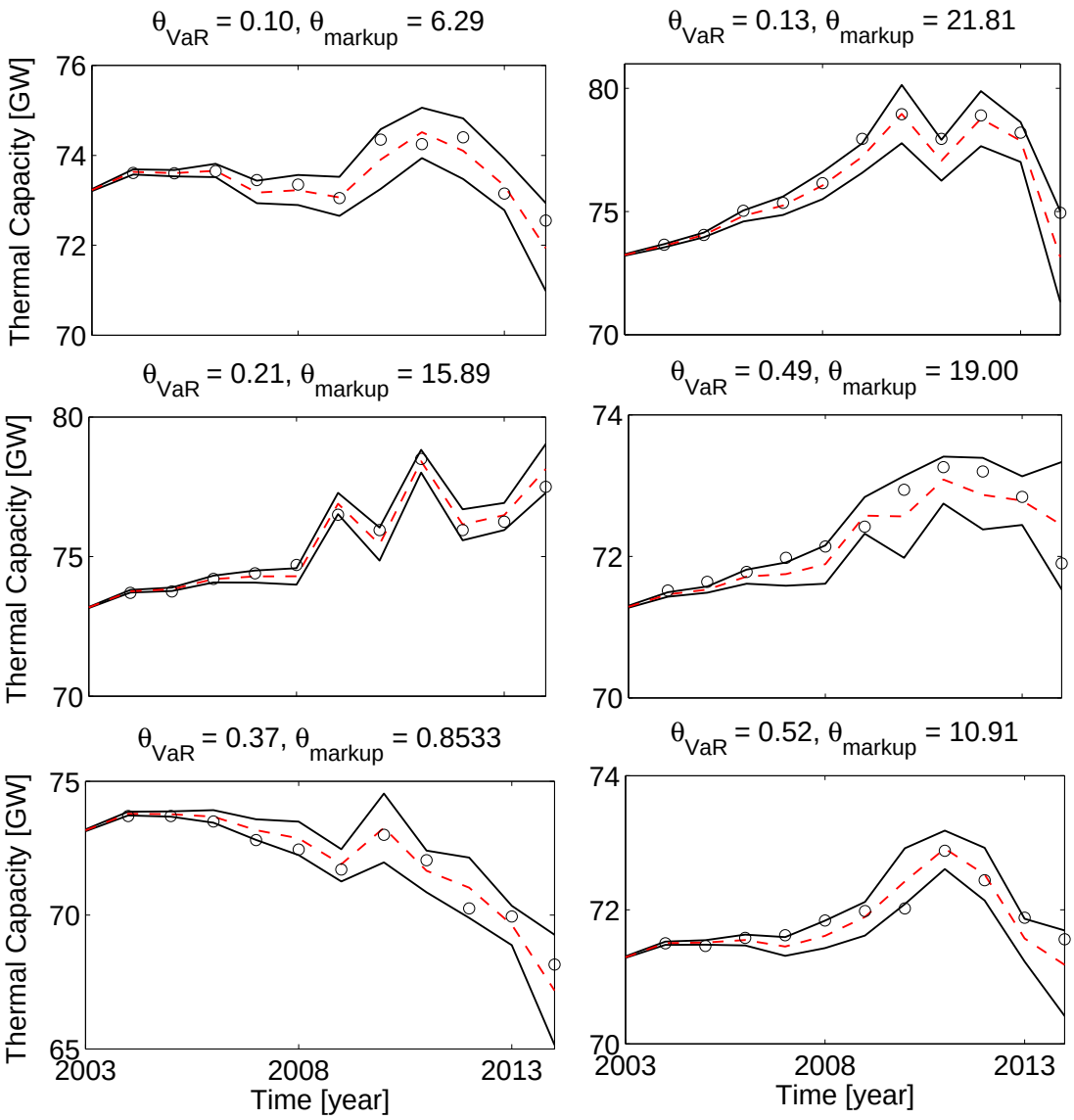

Figure 4.4: Predictions of historical thermal capacities at test points for validation; Circles show the output (i.e., installed thermal capacity) simulated by the simulator; Dashed red and solid black lines indicate the mean, and the 5th and 95th percentiles of the output predicted by the emulator.

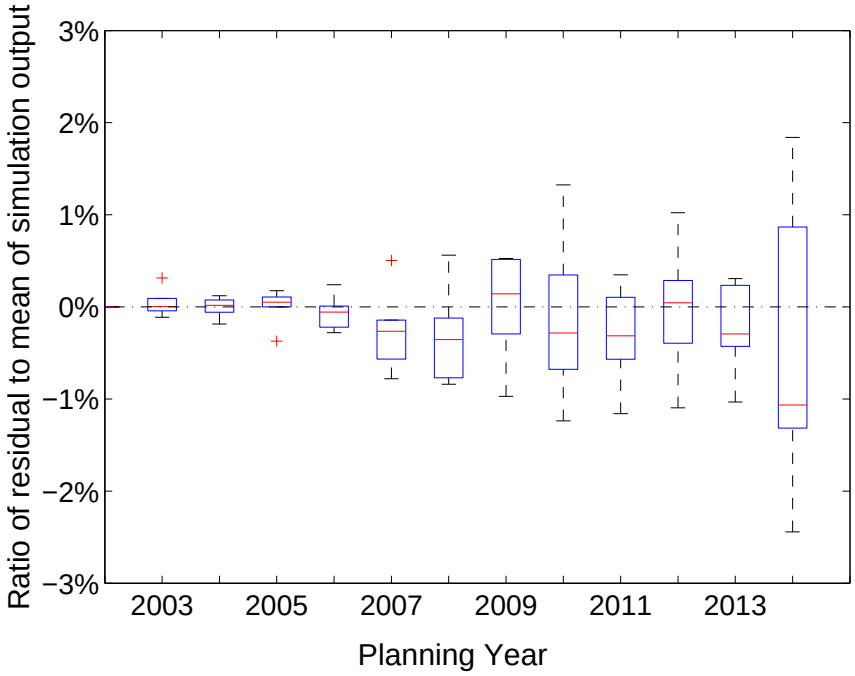

Figure 4.5: Boxplots of residuals at each planning year.

credible interval as a clear evidence of failure. Fig. 4.4 shows that almost all the results produced by the simulator are located within the 90% credibility intervals (in solid black lines) of model output predicted by the emulator, suggesting that the emulator performs well. At a test point (e.g., $\theta_{VaR} = 0.21, \theta_{markup} = 15.89$) which is closer to a design point (see Fig. 4.3), there is less uncertainty in the emulator output.

The analysis of prediction errors is also used as a diagnostic. Fig. 4.5 shows a boxplot of residuals (*i.e.*, differences between predicted mean values from the emulator and observed values from the simulator) at each planning year. The root mean-square error (RMSE) between the emulator's mean of evaluation and the simulator's evaluation is 46.4 MW. The ratio of RMSE to the mean value of the simulated output is 0.62%; this number is small enough to indicate a good mean prediction by the emulator. If the results from diagnostics are not satisfactory, more training data may be chosen for developing the emulator.

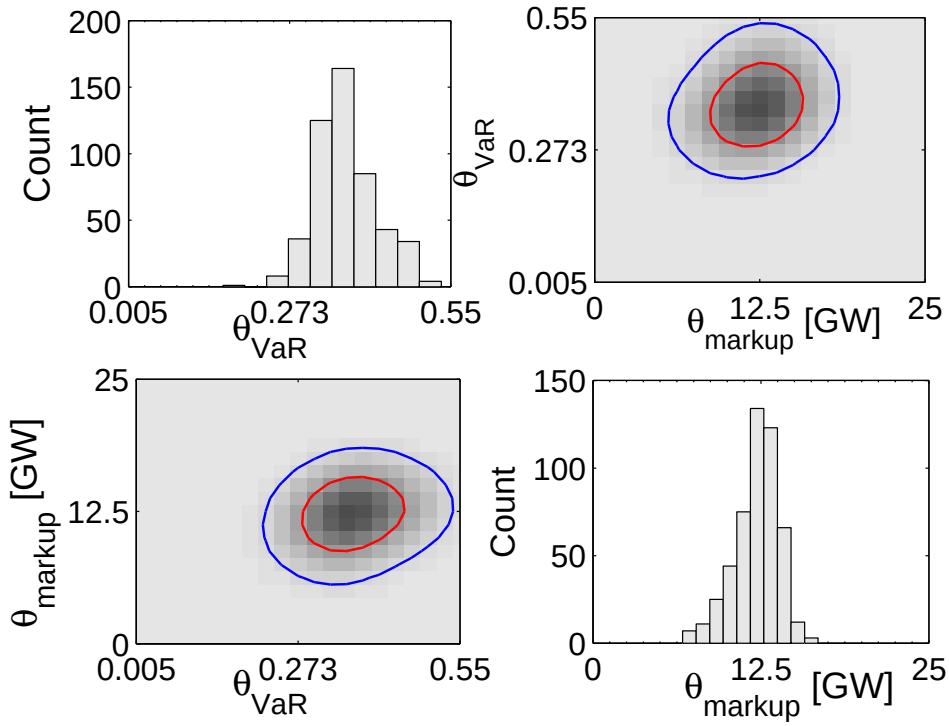

Figure 4.6: Marginal and bivariate posterior distributions of θ on the original scale; Red lines: the 50th percentile; Black lines: the 90th percentile.

4.2.3 Calibration results

Once the emulator is validated, we can then proceed to use it to calibrate the simulator. Calibration obtains the posterior distributions of the calibration parameters by combining the priors with observed data, and meanwhile infers the model discrepancy, so that these may be applied to future projections. The formal calibration approach makes significant improvement on the validation work presented in [82]. In [82], no formal calibration work has been done on uncertain model parameters, and the simulation results at the assumed 'good' values of model parameters are graphically compared against observations without giving any quantitative information, and issues such as structural uncertainty are not accounted for.

Parameter Calibration

Fig. 4.6 shows probability density functions for the marginal (diagonal bar plots) and bivariate (off-diagonal contour plots) posterior distributions of the two calibration parameters on the original scale. The red and blue contour lines represent the

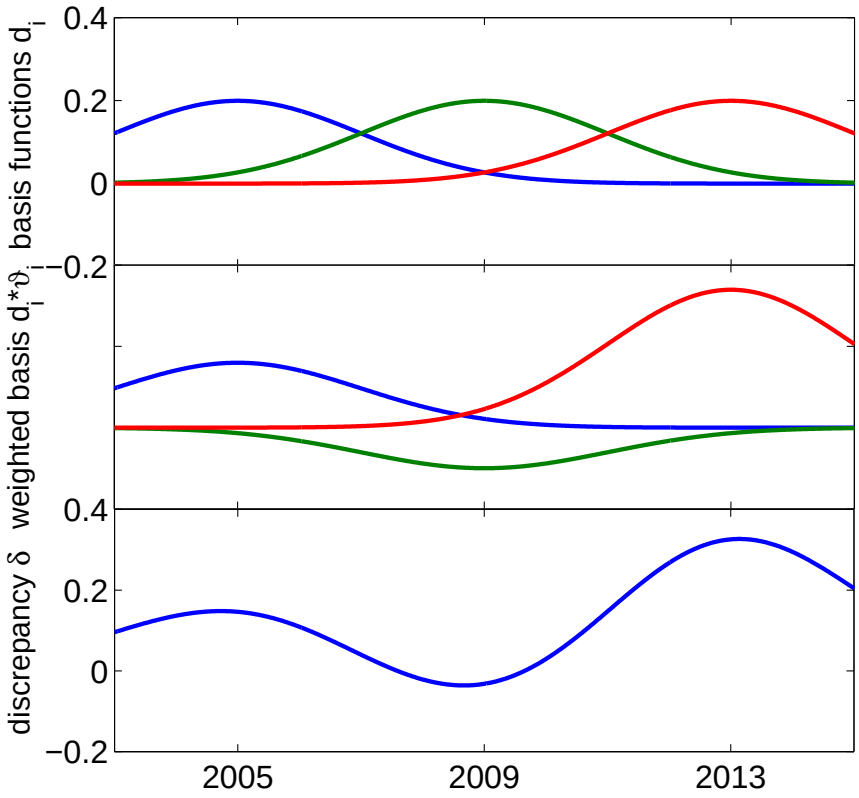

Figure 4.7: An illustrative example of the model discrepancy consisting of three normal kernels.

50th and 90th percentiles of the bivariate posterior distributions, respectively. As compared with their prior uniform distributions specified in Section 4.1, the posterior distributions are constrained in these two dimensions by removing input values that result in implausible outputs.

The posterior distribution of θ_{VaR} is constrained within the range of $0.25 - 0.5$, showing that the investor tends to be risk-averse when they are faced with uncertainties in the future such as fuel prices and demand growth rates. The posterior distribution of θ_{markup} indicates that it is plausible for generators to receive an uplift payment when the capacity margin falls into the range of $6 - 17$ GW (on the original scale).

Calibrated and Discrepancy-adjusted Simulator

Assumptions have to be made about the model discrepancy that adjusts the simulated thermal capacity over the time axis. It is expected that the number of normal kernels should be small. A large number of normal kernels would result in too many

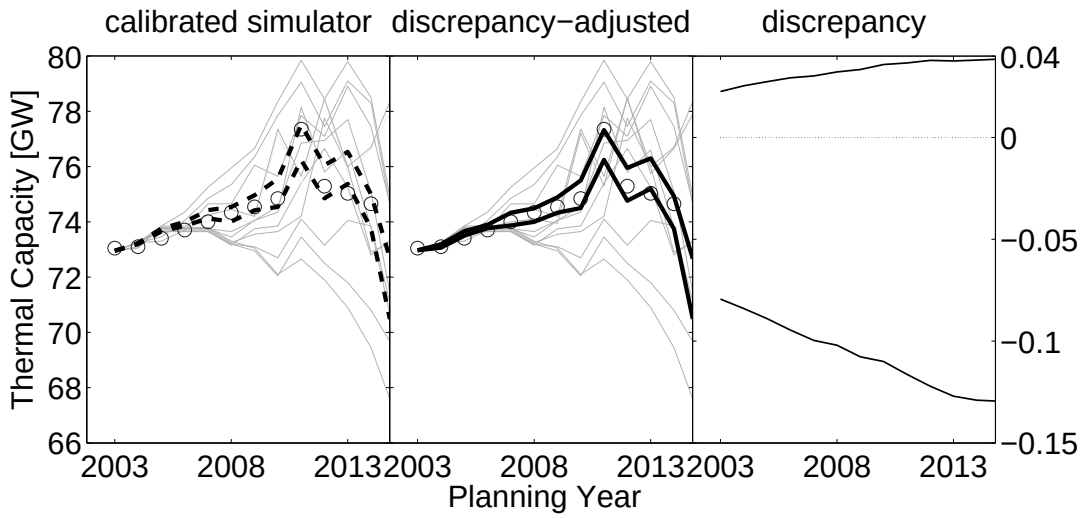

Figure 4.8: The effect of calibration on model accuracy. Circles: the historical data of operational thermal capacities; Left: Calibrated simulations; Center: Discrepancy-adjusted calibrated simulations after adding the discrepancy term to the calibrated simulations; Right: The discrepancy term between emulated and observed values.

uncertain parameters $\sigma_\vartheta^2, \lambda_\vartheta$ that need calibrating and there will be insufficient data (i.e., the limited number of observations available) to reduce uncertainty in these parameters. It is also expected that the discrepancy between simulated and actual thermal capacity has a strong time persistence over the planning horizon. However, only one or two normal kernels may not be flexible enough to model the discrepancy over a long planning horizon. Therefore, it is justified to use a reasonably small number of wide normal kernels to model δ, as long as the kernels fully cover the time horizon. It is simpler to model the discrepancy using a straight line than normal kernels, but a straight line has less flexibility as it cannot capture the potentially non-linear pattern of the discrepancy on the time axis. As explained in Section 3.3.4, each kernel weight is a GP which is a smooth function of x with both positive and negative values, because the model discrepancy is expected to change smoothly with the input condition x.

Here, the model discrepancy is represented by a linear combination of 3 weighted normal kernels centered on years $(2005, 2009, 2013)$, each with a standard deviation of 2. The three kernels are evenly distributed over time and they cover the whole historical planning horizon. Fig. 4.7 shows an illustrative example of the discrepancy

model. The top graph shows the three normal kernels, the middle graph shows the weighted normal kernels (by assigning a real number to each weight) and the bottom graph gives the resultant model discrepancy. The resultant discrepancy term in our case study, shown as the right graph in Fig. 4.8, is very small compared with the total thermal capacity. Since there is only one observation available which is not very informative, the posterior of the discrepancy largely depends on its prior which have very small values, as explained in Section 3.3.4.

The uncertainty in the calibration parameters, represented by the prior or posterior distributions, propagates into the output uncertainty, resulting in a probability distribution over the outputs. Fig. 4.8 shows how the calibration and the discrepancy term reduce the plausible output space when observations are available. A credible interval of $[5\%, 95\%]$ is taken as the plausible range in our case study. Without calibration, a wide range of simulator outputs is simulated by the simulator, as shown by the light grey lines in both the left and the centered graphs. With calibration, a much narrower plausible range of simulator outputs is evaluated by the emulator by the dashed black lines in the left column. After adding the discrepancy term to the calibrated evaluations, the predicted range of simulator outputs, as quantified by the black lines in the centered graph, more closely matches the observation data compared with that predicted by the calibrated simulator. The plots in the right column quantify the plausible range of the model discrepancy, which has a much lower order than the model output and grows slightly over time.

4.3 Predictions using the discrepancy-adjusted and calibrated emulator

This section uses the results of calibration (i.e., the posterior distributions of calibration parameters and model discrepancy) for obtaining plausible future projections of thermal capacity and LOLEs under the energy-only market design. Apart from uncertainty in calibration parameters and model discrepancy, additional uncertainty on the inputs (u, ω) takes effect in making future projections. Following a Bayesian approach allows for the combination of the simulator runs, the posterior distribu-

tions of the calibration parameters θ, and the model discrepancy.

4.3.1 Future projections on operational thermal capacities in an energy-only market

As for calibration, the first stage for obtaining plausible future projections is building an emulator; this is trained using a group of 25 design points over the 5-dimensional input space $x = \{u_{voll}, u_{co_2}, \theta_{VaR}, \theta_{markup}, \omega_{gas}\}$ and the corresponding 25 scenarios of annual installed thermal capacity over period of years $2013 - 2040$. The same validation approach described in Section 3.3.3 was employed by using 6 additional model runs and similar results were obtained suggesting a good fit. All these design and test points are sampled using a Latin hypercube design over their prior ranges given in Table 4.1. A relatively narrow range $U(1000, 20000)$£/MWh is chosen for u_{voll} in this study. Alternatively, during the sampling process, the priors of the two calibration parameters may be replaced with the posteriors obtained during the calibration stage; this can potentially improve the quality of emulator built upon the limited number of training data. In order to compare the future projections made by the pre-calibrated simulator with the calibrated one, the prior distributions of calibration parameters are used in our study.

Fig. 4.9 gives a scatterplot matrix of the 25 normalised design points (in red circles) and 6 normalised test points (in blue points) in the five-dimensional input space. Some parts of the input space have not been covered by the 25 points, such as the top-right corner and the bottom-left corner in the two-dimensional scatterplot of $(u_{voll}, \theta_{markup})$. More design points in a Latin hypercube would give a better exploration of the input space at the expense of longer computational time of the simulator. Three principal components are chosen to capture 95% of the standardised simulation output (future operational thermal capacities) under the energy-only market design, as shown in Fig. 4.10. The first two principal components (in blue and green lines) capture most of the variation in the simulation data than the third one (in red line). Given 5 model inputs and 3 principal components, there are in total 18 hyperparameters in the emulator fitted to 25 training data. Then, there are 7 degrees of freedom left in the estimation. It would be desirable to have more

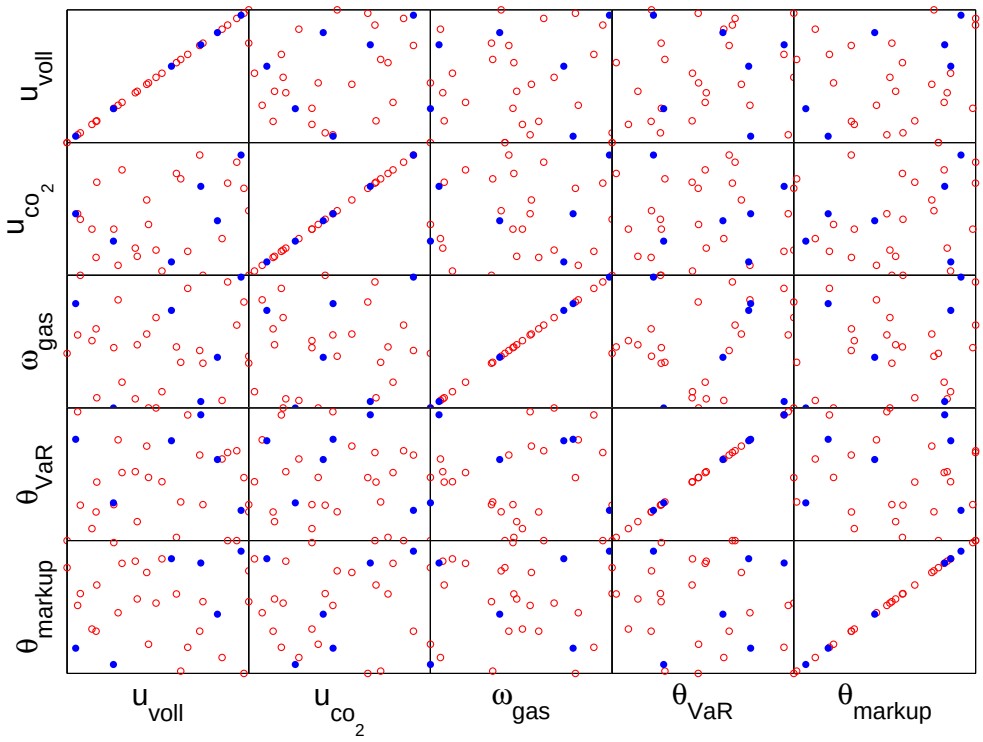

Figure 4.9: A scatterplot matrix of the design points and test points on the [0 1] scale in the five-dimensional input space; Red circles represent design points, and blue points show test points.

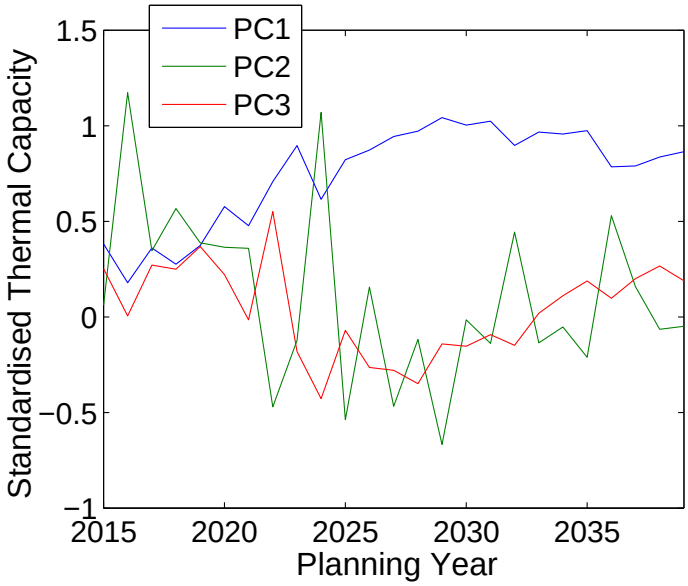

Figure 4.10: Three principal components of the standardised simulation data (future thermal capacities in operation).

training data so that the fitted emulator resulting from more degrees of freedom carries more merits.

Fig. 4.11 shows a breakdown of simulated scenarios of operational thermal capacities given a projection of future on-shore and off-shore wind capacity. On average, the projected total thermal capacity into the future is falling for three main reasons. First, GB system has a fast pace of power plant retirements. Almost all coal-fired power plants will be closed by 2025, combined with the retirement of the majority of the UK's ageing nuclear fleet and over 5 GW of OCGT fleet within a decade from 2015. The electricity supply gap is too big to close by only building enough new CCGT plants. Second, some thermal capacities will be replaced by wind capacities subsidized by the government. The projected wind capacity will reach the target of 45 GW by 2030. Meanwhile, the uncertain demand growth rate used in the simulator is assumed to have a mean of 0 [4] and so the residual demand to be supplied by thermal plants is declining. Third, there may be insufficient incentives for private GENCOs to invest in any type of thermal capacity, particulary when there is a low energy price cap under the energy-only market design.

Fig. 4.12 presents validation results of the emulator that is used for predicting future thermal capacities under an energy-only market design. The red, black and blue solid lines in each graph quantify the 50%, 90% and 99.8% credibility intervals of the emulator's prediction respectively, compared against the simulator's evaluations (in circles) at test points. It can be seen from Fig. 4.12 that most of the simulation output lie within the 90% credibility intervals and all lie within the 99.8% credibility intervals of the emulator output, suggesting that the emulator performs well. Fig. 4.13 shows a boxplot of the ratio of residuals to the mean of simulated thermal capacity at each planning year. The residual in year 2024 is larger than others because the surge in the thermal capacity (lying within 95th and 99.9th percentiles) shown in the middle-right graph in Fig. 4.12 is not fully captured by the principal components. Such an isolated outlier might be ignored because there are no large residuals systematically observed at that particular test point. The ratio of RMSE to the mean value of the simulated output is 1.5%, implying a good mean prediction by the emulator. It is acknowledged that the validation set is not large enough to

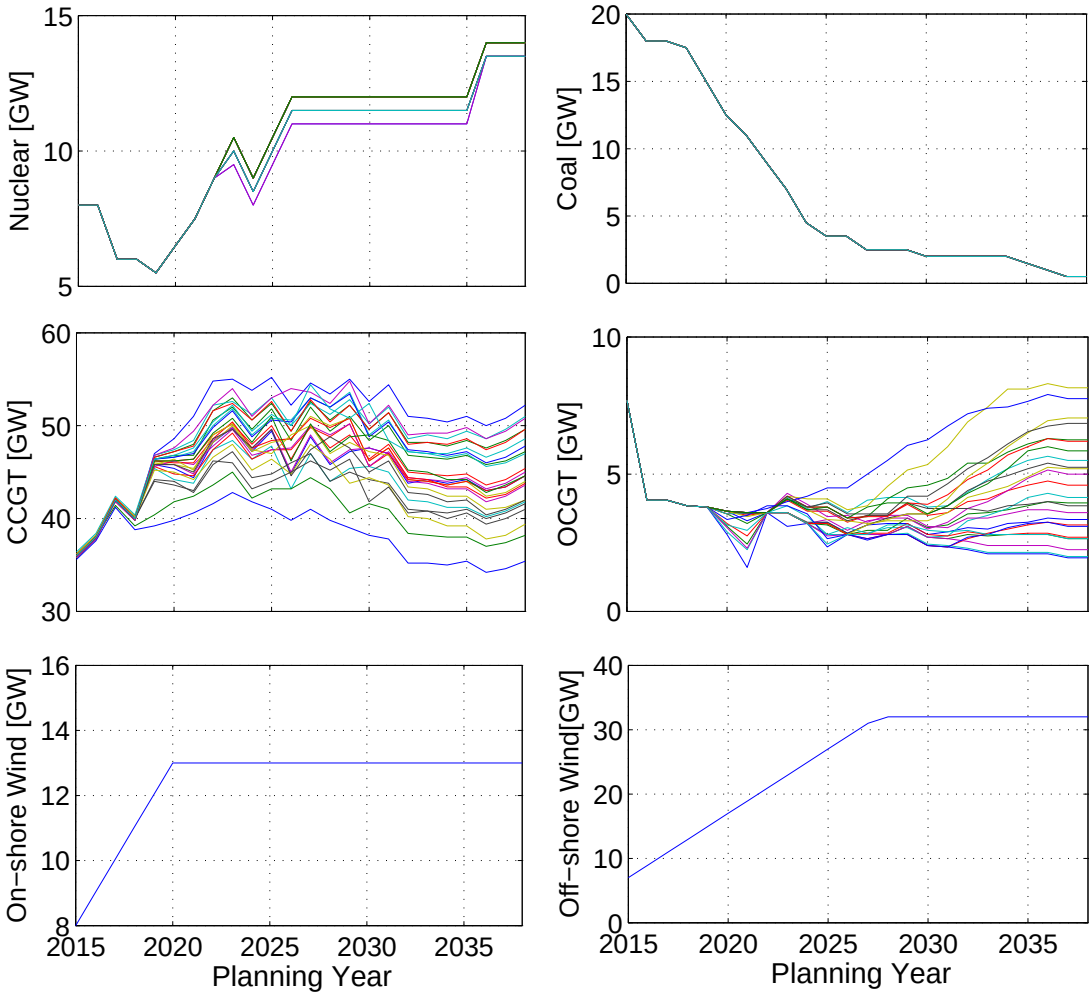

Figure 4.11: A breakdown of simulated generation projections across all design points.

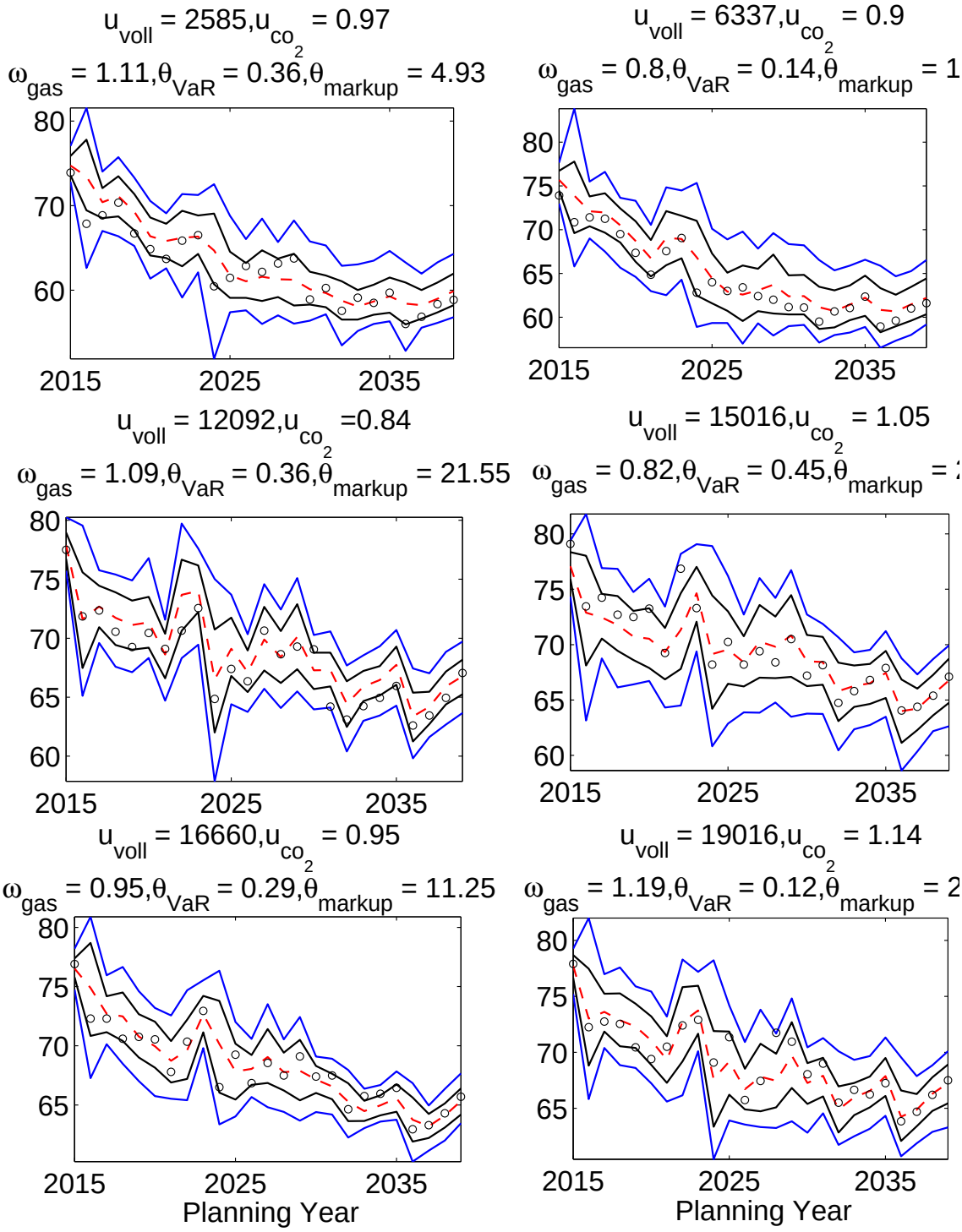

Figure 4.12: Predictions of future thermal capacities at test points for validation; Circles show the thermal capacities simulated by the simulator; Dashed red line, solid black lines and solid blue lines indicate the mean, the 5th and 95th percentiles, and the 0.1th and 99.9th percentiles of the LOLE profile predicted by the emulator.

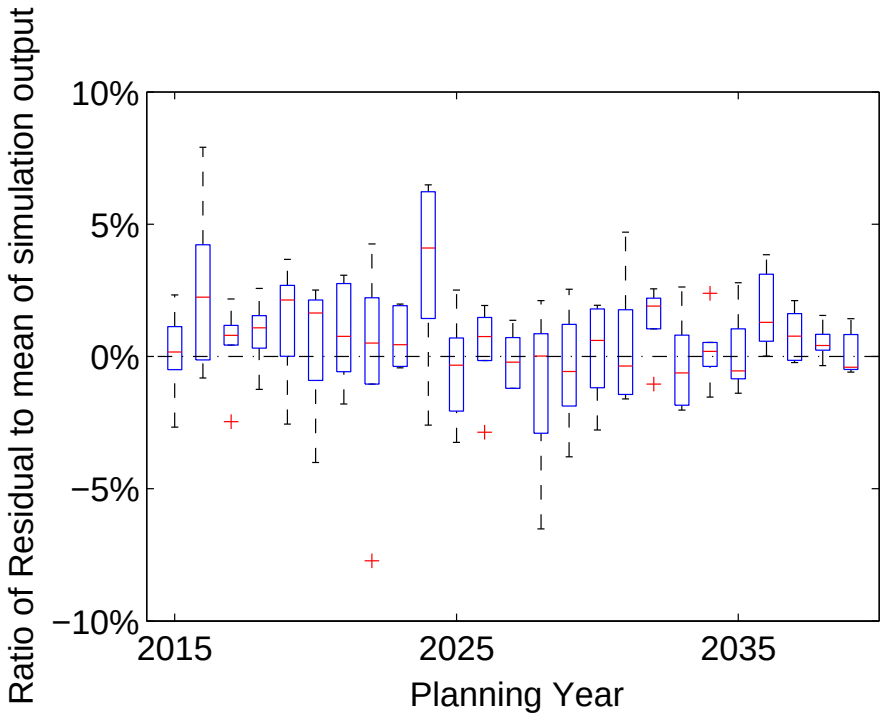

Figure 4.13: Boxplots of residuals at each future planning year.

have great certainty over the accuracy of the emulator that has 5 input variables. More validation data are generally desired if one wants to gain greater certainty.

The validated emulator is then used for making predictions in the second stage, where the control variables are fixed but the calibration and forcing parameters remain uncertain. In this case, the output uncertainty results from uncertainty in forcing and calibration parameters, functional uncertainty, and structural uncertainty. As a consequence, a probabilistic prediction (with uncertain bounds) of generation projections at an input point is estimated by the emulator here as shown in Fig. 4.16, whereas in [82] a deterministic scenario (path) is produced by the simulator which uses a specific set of plausible values of calibration parameters. Without taking into account uncertainty in calibration parameters and structural uncertainty, the simulator's estimation may be inconsistent with historical observations.

Fig. 4.14 shows a probabilistic prediction of future thermal capacities at fixed values of $u_{voll} := 10000$ and $u_{co_2} := 1$. The grey lines in the left and center column show all the simulations obtained from the simulator. A comparison is made between the plausible range of thermal capacities predicted by the simulator before

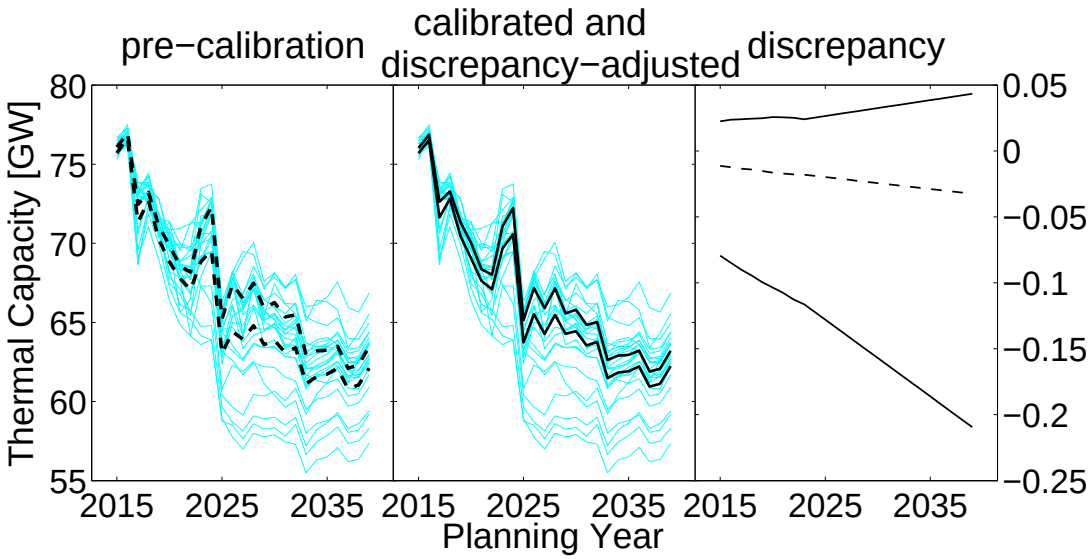

Figure 4.14: Uncertainty quantification in future generation projections; Left: Simulator before calibration; Center: Discrepancy-adjusted calibrated simulations; Right: Discrepancy term applied to future projections.

calibration (in dashed lines in the left column), and the calibrated and discrepancy-adjusted simulator (in black lines in the center column). As expected, a much wider plausible range of future thermal capacities is predicted by the emulator before calibration; in this case, the prior distributions of calibration parameters are used instead of their posterior distributions and no model discrepancy is accounted for. The right column in Fig. 4.14 shows the model discrepancy applied to future projections. The discrepancy is consistent with that inferred from history matching (in Fig. 4.8) for the first 12 simulation years of interest (2015 - 2026). In order to reflect the increasing uncertainty far into the future, the model discrepancy is assumed to increase by 5% per year from 2026 onwards. The discrepancy term used here is simply an illustration of mitigating the risk of making overconfident projections. In reality, a much larger model discrepancy is expected than what is used here, when projecting thermal capacities over 25 years into the future. To obtain a large model discrepancy, the prior of the discrepancy model can be modified (by assigning an informative prior of the marginal precision model in (3.17)). Alternatively, a larger growth rate of the model discrepancy projected into the future can be assumed.

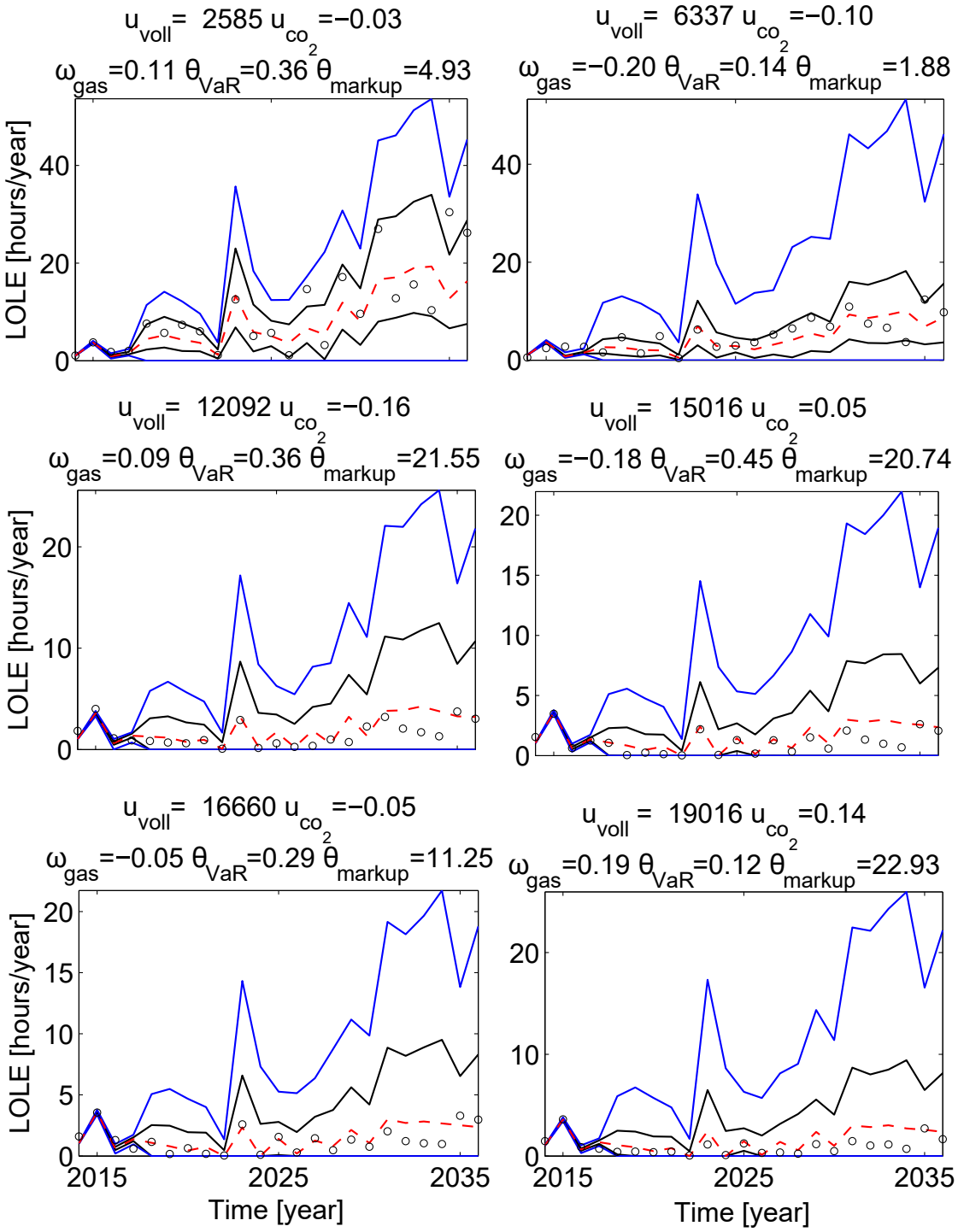

Figure 4.15: Predictions of LOLEs at data points for validation; Circles show the annual LOLE simulated by the simulator; Dashed red line, solid black lines and solid blue lines indicate the mean, the 5th and 95th percentiles, and the 0.1th and 99.9th percentiles of the LOLE profile predicted by the emulator.

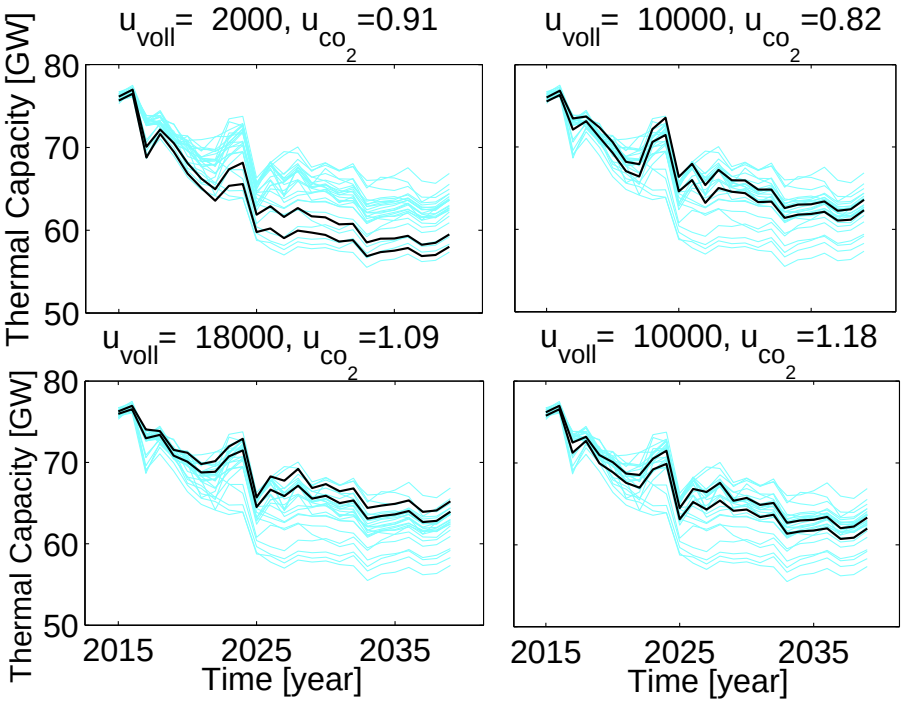

Figure 4.16: Predictions of thermal capacities at different values of (u_{voll}, u_{co_2}). Crane lines are simulation output and black lines indicate the plausible range of projected thermal capacities.

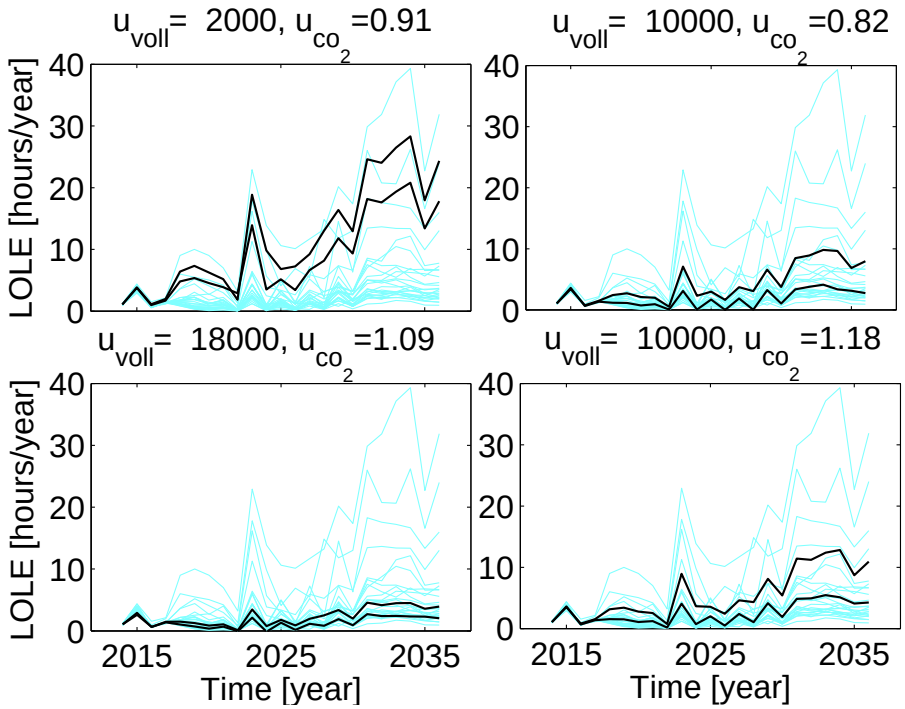

Figure 4.17: Predictions of LOLE profiles at different values of (u_{voll}, u_{co_2}). Crane lines are simulation output and black lines indicate the plausible range of the projected annual LOLEs.

4.3.2 Future projections on annual LOLEs

The time series of LOLE are calculated in (3.1c) as a metric for system reliability using the same set of design points as used for emulating thermal capacities in Section 4.3.1. The LOLE profiles under the energy-only market design are emulated and validated using the Bayesian approach as described in Section 3.3.2.

The validation results presented in Fig. 4.15 indicate that the mean of the emulator's evaluation reasonably matches the simulator's output of the LOLE profile and almost all the test data lie in the 90% credible intervals predicted by the emulator. The original credible intervals of LOLE predicted by the emulator contain negative values that have no physical meaning. This is because the emulator output is a continuous response of the input variables. When the uncertainty range is large and the LOLE in a year is close to zero at some input settings, the emulator will produce negative evaluations. All the negative values that are drawn from the emulator are replaced with zero before calculating the credibility intervals. Although this approach introduces discontinuity at zero, the results are still usefulness because the larger values of LOLE is of more importance to power system reliability and hence of more interest to policymakers.

For exploring the combined effects of (u_{voll}, u_{co_2}), the probabilistic predictions of thermal capacities and LOLE are presented in Fig. 4.16 and Fig. 4.17 respectively, at selected combinations of "high", "middle" and "low" values of both variables. In comparison with the GB standard of 3 hours per year LOLE [61], the risk of security of supply from year 2023 onwards for some choices of input settings can be very high, as shown by the grey lines in Fig. 4.17. The right-hand graphs in Fig. 4.16 and Fig. 4.17 show that thermal capacities decline and the LOLE increases as the value of u_{co_2} increases (*i.e.*, the trend level of carbon prices increases) but the effect is small. One of the advantages of the probabilistic predictions is that it is natural and computationally efficient to determine the combination of (u_{voll}, u_{co_2}) with a high probability of keeping the $LOLE$ at each planning year below some set threshold.

4.4 A study on the robustness of market designs using the calibrated emulator

This subsection presents a study on the robustness of two market designs, namely energy-only market and capacity market in the context of uncertainty. The robustness of a specific setting of a market design is indicated by the probability of not exceeding a certain threshold of LOLE, denoted as y^{LT} (e.g., 3 hours per year), given as,

$$R(u) = Pr(y^L < y^{LT}|u). \tag{4.1}$$

Obtaining future projections of the maximum LOLE is necessary for exploring scenarios of energy policies that can plausibly be applied to the GB power system without increasing LOLE by more than a certain threshold. The calibration procedure as shown in Section 4.2 produces the posterior distribution of calibration parameters θ that will be applied to obtain plausible predictions of LOLE over the future planning years. Following the Bayesian emulation technique allows for the combination of the simulator runs and the posterior distributions of calibration parameters.

4.4.1 Emulating the maximum LOLE under two market designs

In the study on the robustness of market designs, a wider prior range $\mathrm{U}(1000, 30000)$ has been chosen for the energy price cap u_{voll} under both market designs. The prior distributions of all other model inputs are the same as those provided in 4.1. The simulator is run forward to year 2040 to produce the output of interest - the maximum LOLE, and then an emulator is built for each of the two market designs. An emulator is trained in an energy-only market design using a group of 35 design points over the 5-D input space and the corresponding 35 scenarios of the maximum LOLE over period $\mathcal{F}$. The other emulator is trained in a capacity market design using a new group of 40 design points over the 6-D input space because the capacity market design needs an extra control parameter u_{cone}.

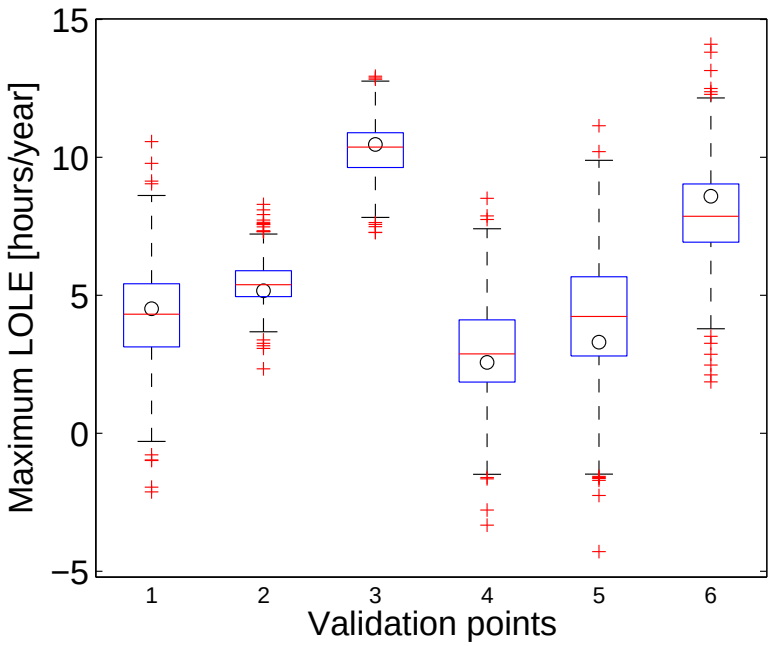

Figure 4.18: Predictions of the maximum LOLE under an energy-only market design at test points; Circles show the simulator's evaluations; Horizontal red line, blue lines indicate the mean, the 5th and 95th percentiles of the maximum LOLE predicted by the emulator, respectively.

Apart from training data, a handful of extra design points are used for validating the emulators for both market designs. The boxplots in Fig. 4.18 and Fig. 4.19 show that all test points (in circles) lie in the 5th and 95th percentiles (in horizontal blue lines) of the distribution of the maximum LOLE. The validation results indicate that the emulators are accurate enough for carrying out the robustness study.

4.4.2 Uncertainty analysis of the maximum LOLE

In the LTGI model described in Section 3.2.1, there is uncertainty in the input values of x, which propagates into output uncertainty, resulting in a range of the maximum LOLE. An UA will derive the uncertainty range of the maximum LOLE under energy-only or capacity market design, given that the model inputs including the parameters of market designs have specified probability distributions (see Table 4.1). A full UA is achieved here through exploring the posterior distributions of all model inputs together with the emulator via Markov chain Monte Carlo [154, 166].

Fig. 4.20 shows the uncertainty range of the maximum LOLE under an energy-

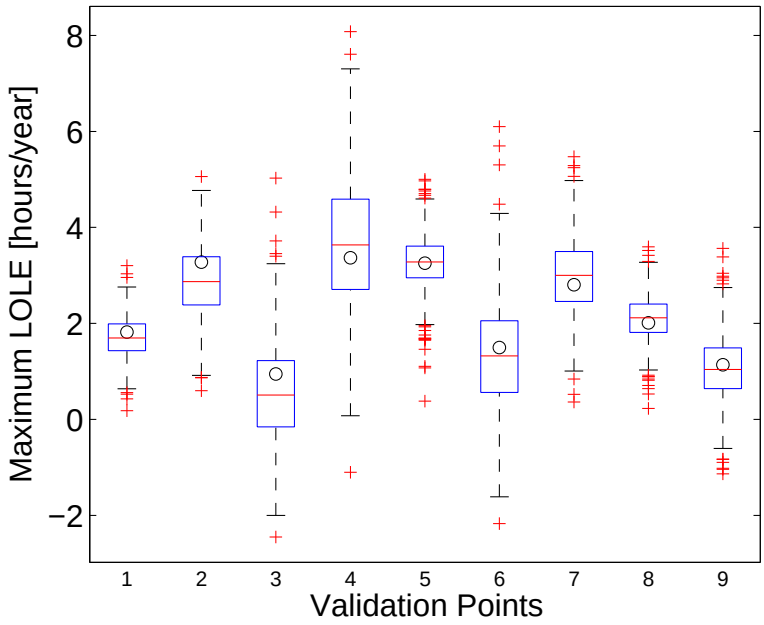

Figure 4.19: Predictions of the maximum LOLE at test points under a capacity market design; Circles show the simulator's evaluations; Horizontal red line and blue lines indicate the mean, 5th and 95th percentiles of the maximum LOLE predicted by the emulator, respectively.

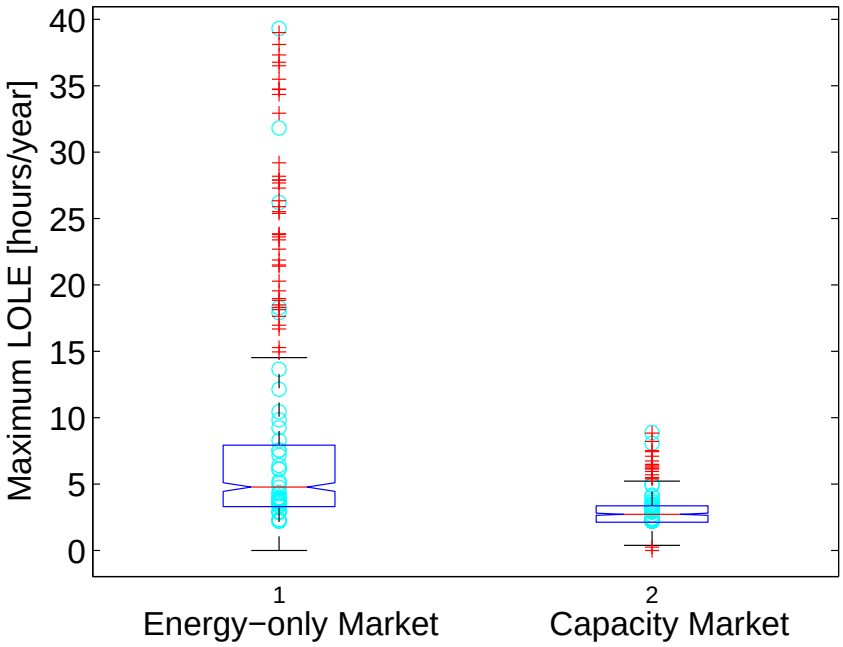

Figure 4.20: Uncertainty range of the maximum LOLE given all exogenous uncertain inputs of interest under two market designs. Crane circles show the simulated data of the maximum LOLE at design points. Red crosses '+' display data beyond the whiskers.

only market design (on the left column) and a capacity market design (on the right column). The box shows the interquartile range, which accounts for 25th – 75th percentile. The whiskers add 1.5 times the interquartile range to the 75th percentile and subtract 1.5 times the interquartile range from the 25th percentile. The central line shows the median (the 50th percentile) of the data. The center of the box gives the mean of the data. As can be seen in Fig. 4.20, the reliability of the system under an energy-only market design appears much riskier than under a capacity market design, implying that capacity adequacy will be a great concern at lower settings of LOLE and the system is less robust to market risks and investors' risk attitude in the context of energy-only markets.

4.4.3 Algorithm and results of the robustness study

The procedure of quantifying $R(u)$ at a given set of control variables is described as follows:

Step 1: Choose a set of values from control variables $u := \{u_{cone}, u_{voll}, u_{co_2}\}$, obtained from their prior distributions;

Step 2: Take a sample from the posterior distribution of calibration parameters $\theta := \{\theta_{VaR}, \theta_{markup}\}$, obtained after calibration;

Step 3: Take a sample from the forcing parameter $\omega := \{\omega_{gas}\}$, obtained from its prior distribution;

Step 4: Take a sample from the posterior distribution of GP model parameters, denoted as λ, γ, obtained after calibration;

Step 5: For each combination $(\theta, \omega, \lambda, \gamma)$, use the emulator to predict the probability density function $P(y^L|u, \theta, \omega, \lambda, \gamma)$, and compute the probability $Pr(y^L < y^{LT}|u, \theta, \omega, \lambda, \gamma)$;

Step 6: Average out over $(\theta, \omega, \lambda, \gamma)$ to get the probability $Pr(y^L < y^{LT}|u)$.

Applying the above procedure to the calibrated emulators developed in 4.3, Table 4.2 shows the probabilities of the maximum LOLE exceeding 3-hour per year under different settings of design parameters in energy-only and capacity markets. Note that the variable u_{cone} only applies in capacity markets. Comparing the results between the second and the third columns, at the same setting of u_{voll}, u_{co_2}

Table 4.2: A comparison of the probabilities of exceeding the threshold (3-hour per year) of LOLE between energy-only market and capacity market

Control Variables u			Energy-only Market			Capacity Market		
u_{cone}	u_{voll}	u_{co_2}	$R(u)$	$E(y^L)$	$\text{Var}(y^L)$	$R(u)$	$E(y^L)$	$\text{Var}(y^L)$
£/kW/yr	£/MWh	N/A	N/A	hrs/yr	$(\text{hrs/yr})^2$	N/A	hrs/yr	$(\text{hrs/yr})^2$
31.89	1000	1	1	44.16	2.83	0.97	4.17	0.69
47.18	5000	1.2	1	11.61	1.46	0.71	3.37	0.75
47.18	5000	1	1	11.45	1.43	0.53	3.02	0.34
47.18	5000	0.8	1	11.33	1.40	0.20	2.50	0.67
66.21	2000	1	1	29.92	2.36	0.75	3.54	0.89
31.89	10000	0.8	0.98	6.87	1.61	0.768	3.50	1.14
47.18	15000	1	0.83	3.90	1.22	0.09	2.20	0.62
66.21	19500	1	0.57	3.21	1.68	0.04	1.51	0.45

in both markets, the expected value of the maximum LOLE, $E(y^L)$, informing the expected level of the security of supply, considerably reduces with the introduction of capacity payment. Although some combinations of the control variables in the two market designs may result in similar expectations of the maximum LOLE, for example, setting $u_{voll} = 25000$£/MWh in an energy-only market, and setting $u_{cone} = 47.18$£/kW/yr, $u_{voll} = 5000$£/MWh in a capacity market, the variances of the maximum LOLE, $\text{Var}(y^L)$, however, are generally larger in energy-only markets. A comparison between $Pr(y^L < 3|u)$ in both markets shows that a capacity market design is generally more robust to uncertainties than energy-only market. The results from the 2-nd and 5-th rows reveal that the increase in carbon prices may discourage thermal investments and hence reduce the level of security of supply in both markets. However, the effect of the control parameter u_{co_2} is insignificant.

A clearer way to show the impact of the capacity and energy price caps on the long-term system reliability is making a large number of predictions on the maximum LOLE at possible scenarios of two control parameters u_{cone}, u_{voll}. Fig. 4.21 displays the probability map of the maximum LOLE not exceeding 3 hours a year given 100 combinations of u_{cone} and u_{voll} sampled from their prior distributions while averaging over the joint distribution of all other input variables. This quantitative information

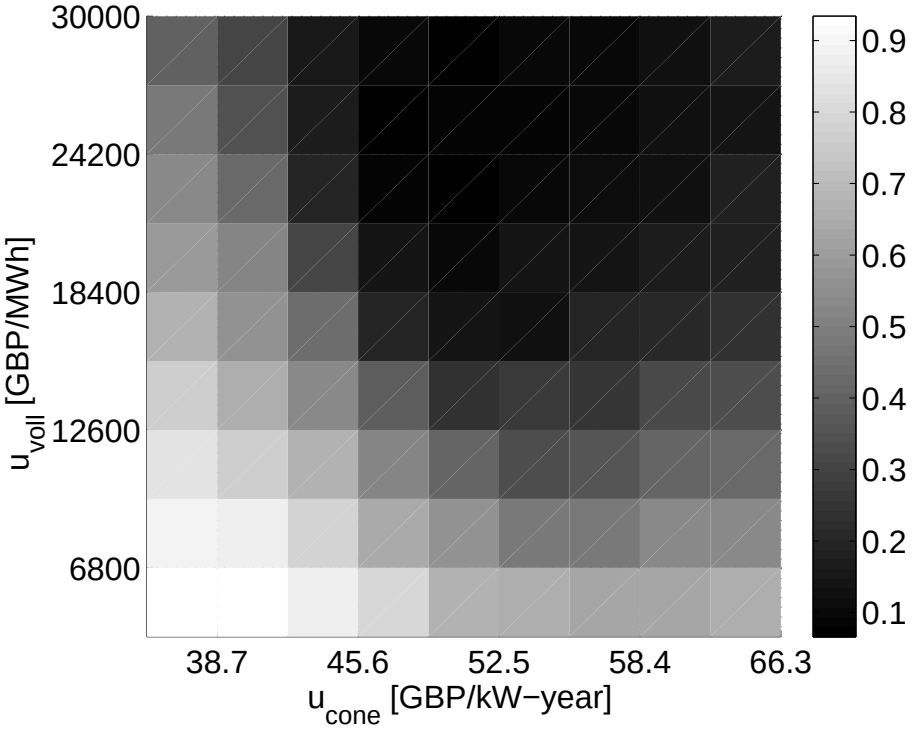

Figure 4.21: Probability map of the maximum LOLE not exceeding 3 hours a year given different combinations of u_{cone} and u_{voll} in the capacity market.

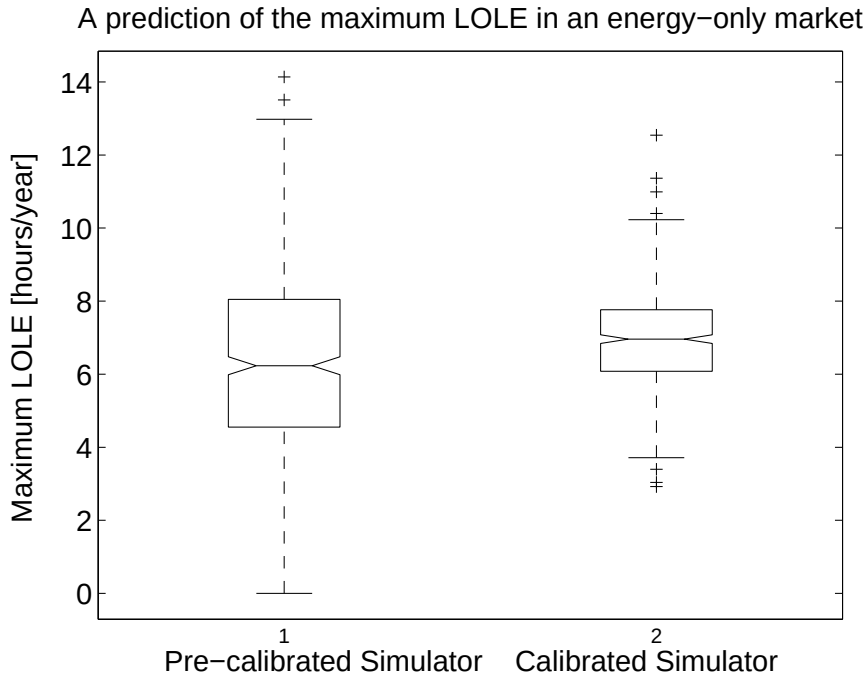

Figure 4.22: Uncertainty range of the maximum LOLE in an energy-only market, conditional on values of $u_{voll} = 10000£/MWh$ and $u_{co_2} = 1$ and averaging on all other input variables.

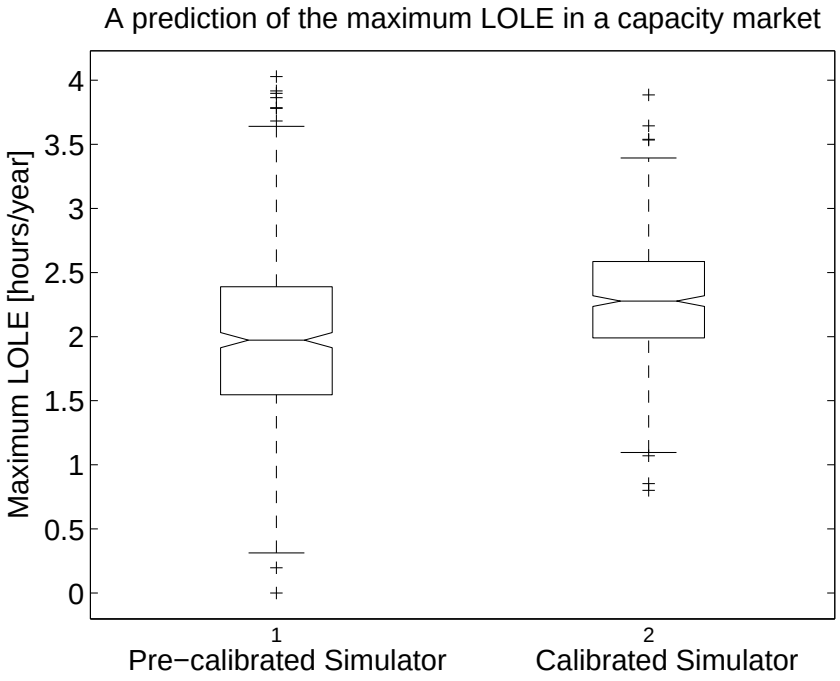

Figure 4.23: Uncertainty range of the maximum LOLE in a capacity market, conditional on values of $u_{cone} = 47.18£/kW/yr$, $u_{voll} = 10000£/MWh$ and $u_{co_2} = 1$ and averaging on all other input variables.

would be very helpful for choosing good values of u_{cone}, u_{voll}.

To help policymakers understand the use of Bayesian emulation for design of energy policy, the plausible ranges (*i.e.*, the probabilistic prediction) of the maximum LOLE at a specific setting of control variables under an energy-only market and a capacity market are graphically shown in Fig. 4.22 and Fig. 4.23, respectively. In each figure, a comparison is made between the output predicted by the simulator before calibration, shown as the left-hand box plot, and by the calibrated simulator, shown as the right-hand box plot. In this case, uncertainty in the maximum LOLE propagates from uncertainty in forcing and calibration parameters as well as the approximation error between the emulator and the simulator. Since uncertainty in calibration parameters has been reduced through the calibration procedure (as shown in Fig. 4.6), the plausible range of maximum LOLE produced by the calibrated simulator is much narrower than that produced by the pre-calibrated simulator as shown by the two box plots in both figures.

The probabilistic prediction made by the calibrated emulator has a lower risk

of overfitting than the deterministic prediction made by the simulator which uses a single set of values (point estimates) of model inputs as in [82]. Using a posterior distribution covering a range of possible values for calibration parameters leads to a range of output evaluations that are consistent with historical observations. Whereas, using point estimates of pre-calibrated parameters produces a point estimate of LOLE lying somewhere in between the left-hand box plot in Fig. 4.23; this is a conservative estimation in an uncertain environment.

4.5 Sensitivity analysis results

A comprehensive probabilistic SA can be efficiently conducted based upon the emulators validated and calibrated in Sections 4.2–4.4. The implementation requires a distribution for the uncertain inputs given in Table 4.1. The SA approach based on an emulator has benefits of allowing full exploration of the input space, estimating variance-based measures at a low computational cost, and accounting for interactions and nonlinear responses. By contrast, directly using a limited number of training data in SA, is far less informative, because it does not fully explore the input space. It is acknowledged that although an emulator allows SA to explore the whole input space, it does so by extrapolation from the training data.

4.5.1 Sensitivity to generation projections

A probabilistic SA is firstly conducted on the emulator developed in 4.2 to study the individual and combined effects of calibration parameters on the historical thermal capacities.

Firstly, the following mean-effect functions for each input parameter x_i are explored,

$$M(x_i) = E_{x_{-i}}(y^G|x_i) = \int yP_{-i|i}(x_{-i}|x_i)dx_{-i}, \tag{4.2}$$

where $E_{x_{-i}}(y^G|x_i)$ quantifies the expected value of the simulator output y^G to varying an individual input x_i, averaged over the probability distribution of all the other input variables x_{-i} and conditional on the value of x_i.

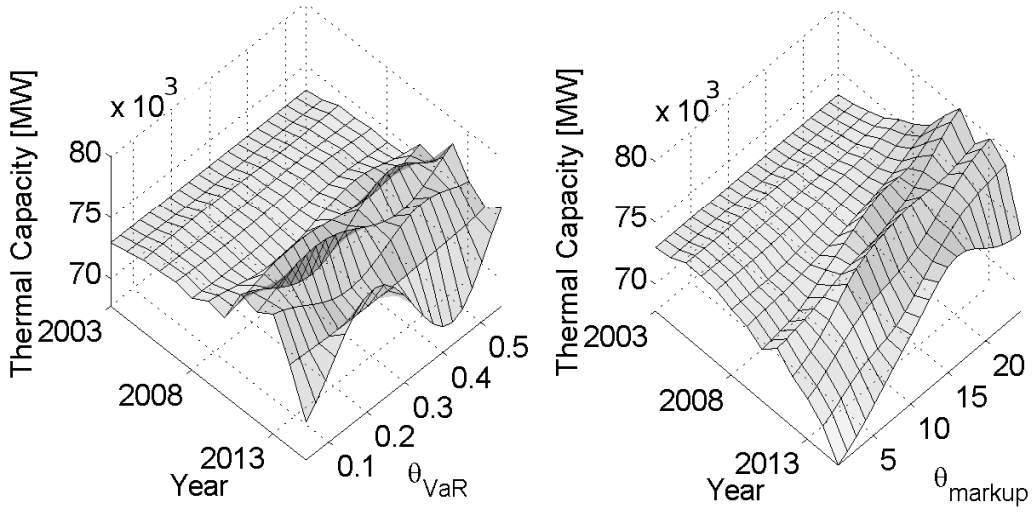

Figure 4.24: Sensitivity analysis of $\theta = \{\theta_{VaR}, \theta_{markup}\}$ in history matching

Fig. 4.24 plots the average response surface of the time series of historical generation capacities to each of the two calibration parameters. The two graphs in Fig. 4.24 indicate that the two parameters have time-delay effects, that is, it takes some time (a few years) to make changes to the installed thermal capacity due to construction delays of new investments. In addition, the time series of operational thermal capacities is not changing smoothly with the values of calibration parameters, because long-term generation investments naturally display boom-and-bust cycles in a liberalised electricity market [14, 17, 18, 115]. This is also one of the reasons of using principal components (an orthogonal basis set) in representing the time series of thermal capacities, so that the weight of a basis are relatively smooth functions of the model inputs. The clearly visible sinusoidal pattern across 2014 suggest that there is a nonlinear relationship between the calibration parameters and the annual thermal capacities.

Table 4.3 provides the variance contributions of individual and joint calibration parameters to the total variance of the model output in terms of the main effect index defined in (3.19). The quantitative information delivers a clear message that the price markup model parameter θ_{markup} alone is dominating variations in installed thermal capacity.

Table 4.3: Main-effect sensitivity measures of single and joint calibration parameter effects on generation projections (% total variation) in history matching

Input variables	θ_{VaR}	θ_{markup}	$(\theta_{VaR}, \theta_{markup})$
% of total variation	6	85	9

Table 4.4: Main-effect sensitivity measures of single and joint parameter effects on generation projections (% total variation) in the energy-only market

Inputs	u_{voll}	u_{co_2}	ω_{gas}	θ_{VaR}	θ_{markup}
u_{voll}	57.6	0.46	0.13	0.01	0.53
u_{co_2}	0.46	2.3	0.12	0.02	0.24
ω_{gas}	0.13	0.12	0.8	0.02	0.33
θ_{VaR}	0.01	0.02	0.02	0.6	0.13
θ_{markup}	1.84	0.53	0.33	0.13	34.9

A probabilistic SA is then carried out on the emulator developed in Section 4.3 for analyzing the effects of the five model inputs $\{u_{voll}, u_{co_2}, \omega_{gas}, \theta_{VaR}, \theta_{markup}\}$ on future generation projections in energy-only markets. Table 4.4 provide the variance contributions of individual model inputs and those of two-input interactions to the overall variance associated with the projections of thermal capacity. It is clearly seen that the most important input parameters are u_{voll} and θ_{markup}, as varying these contributes the most to variations in projected future thermal capacities. An index of 57.6% indicates that uncertainty about u_{voll} accounts for over half of the overall uncertainty in the output. Uncertainty in θ_{markup} accounts for 34.9% of the overall uncertainty in the output. Both parameters take effect on scarcity pricing, that is, modifying the uniform market-clearing energy prices in scarcity situations. The results from SA highlight the significant long-term effect of scarcity pricing on incentivizing the investment of conventional generation capacity in future electricity markets with a high penetration of wind power.

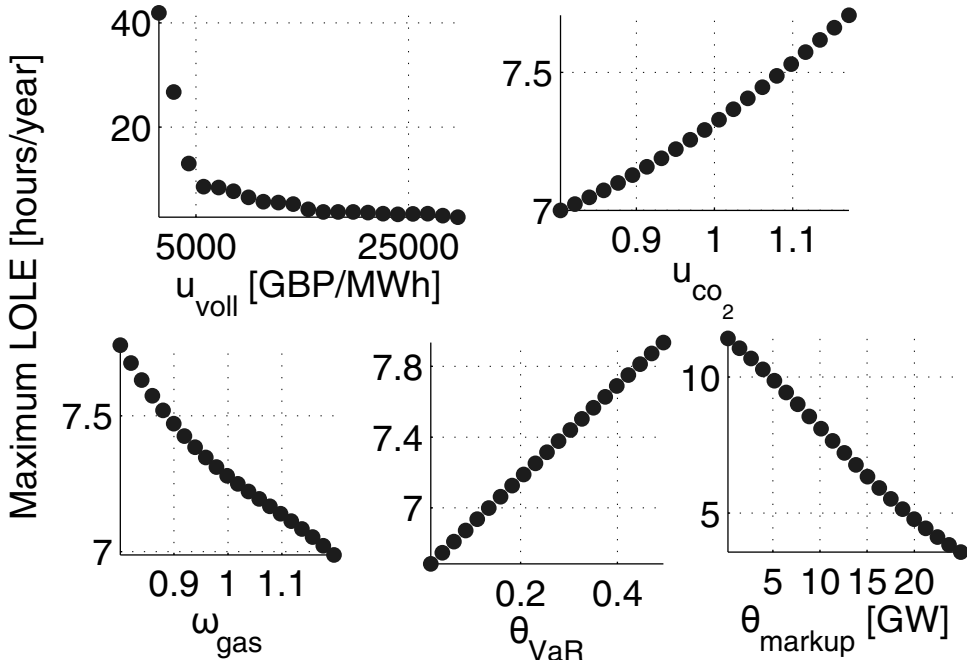

Figure 4.25: Main-effect sensitivity plots for each parameter individually varied whilst averaging over others in the energy-only market.

4.5.2 Sensitivity to system reliability

In this section, a probabilistic SA is conducted on the emulators developed for predicting the maximum LOLE under an energy-only and a capacity market design in Section 4.3. Three methods are considered for assessing the influence of the input parameters on system reliability.

Fig 4.25 and Fig 4.26 plot the mean-effect functions of each input x_i, representing the expected value of the maximum LOLE against the value of each input in the energy-only and capacity markets. These plots show the input that has the greatest effect on the maximum LOLE is the energy price cap u_{voll} in both market designs. Besides, an increase in θ_{markup}, that is, a decrease in the capacity margin where the energy price markup applies, considerably reduces the maximum LOLE, justifying scarcity rents as investment incentives. On the one hand, it is noted that u_{voll} is assumed to have a rather wide range $(1000, 30000)$ over which it is varied. Similar to all other SA approaches discussed in Section 2.5.4, the probabilistic SA approach has the same issue of often observing larger impact on the output in response to larger variations in an input. Changing the range of variation will change the SA measures and may alter the ranking of inputs resulting from their contribution to variation

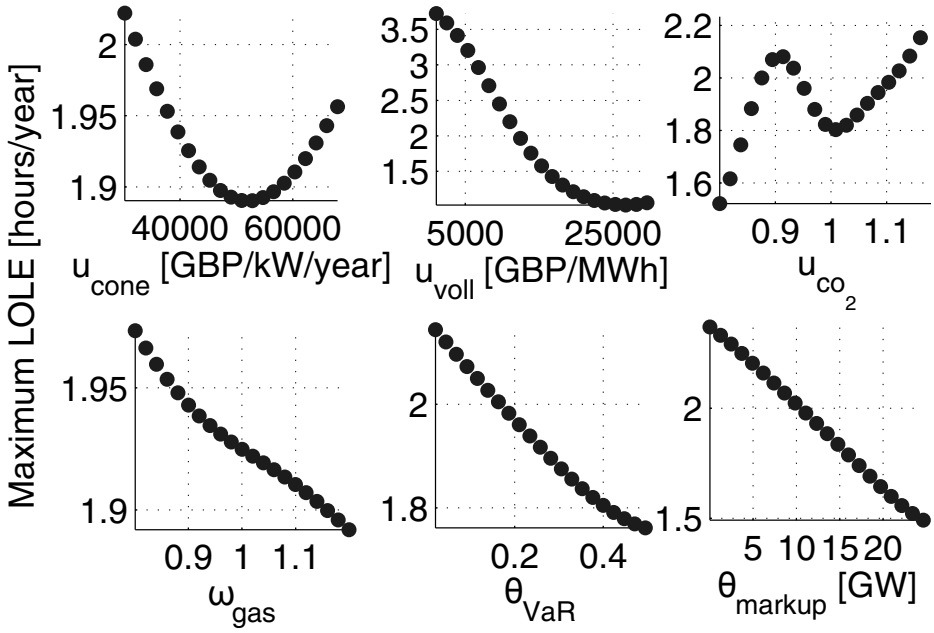

Figure 4.26: Main-effect sensitivity plots for each parameter individually varied whilst averaging over others in the capacity market.

in the output. On the other hand, different sets of training data were gathered for testing the consistency of these mean-effect plots. The mean-effect results for the gas price parameter ω_{gas} in Fig. 4.26 are associated with a large degree of uncertainty due to its insignificant effect. The plots of all other five model inputs do not vary much among different sets of training data.

Table 4.5: Main-effect sensitivity measures of single and joint parameter effects on the maximum LOLE (% total variation) in the energy-only market

Inputs	u_{voll}	u_{co_2}	ω_{gas}	θ_{VaR}	θ_{markup}
u_{voll}	57.52	0.64	0.07	3.00	7.86
u_{co_2}	0.64	0.62	$2.24e-3$	0.11	0.24
ω_{gas}	0.07	$2.24e-3$	0.09	0.04	0.04
θ_{VaR}	3.00	0.11	0.04	1.35	5.7
θ_{markup}	7.86	0.24	0.04	5.7	12.47

Secondly, based on the eqs. (3.19)–(3.21), Table 4.5 and Table 4.6 show the main effect of particular inputs and two-input interactions on the output of interest in energy-only and capacity markets, respectively. In each table, the diagonal data

Table 4.6: Main-effect sensitivity measures of single and joint parameter effects on the maximum LOLE (% total variation) in the capacity market

Inputs	u_{cone}	u_{voll}	u_{co_2}	ω_{gas}	θ_{VaR}	θ_{markup}
u_{cone}	0.71	1.32	1.89	0.10	0.40	0.69
u_{voll}	1.32	53.75	7.54	0.35	0.65	1.73
u_{co_2}	1.89	7.54	2.96	1.25	2.13	11.06
ω_{gas}	0.10	0.35	1.25	0.43	0.08	0.24
θ_{VaR}	0.40	0.65	2.13	0.08	1.47	1.04
θ_{markup}	0.69	1.73	11.06	0.24	1.04	5.41

Table 4.7: Total effect sensitivity measures (% of total variation) in the energy-only and capacity markets

Market	u_{cone}	u_{voll}	u_{co_2}	ω_{gas}	θ_{VaR}	θ_{markup}
Energy-only	N/A	79.2	0.22	0.36	2.02	36.38
Capacity	7.18	74.52	29.96	4.2	8.87	22.35

show the percentage of variance that is due to the corresponding input alone, and the remaining data are the percentage of variance that is due to the second-order interactions. The total two-input interactions have less influence on the output variations than the main effect of the single parameter u_{voll} in both markets. For the capacity market, Table 4.6 shows that future carbon price level u_{co_2} has limited effects on its own and most of its contribution to the output variance comes through interactions with the other parameters.

Thirdly, Table 4.7 gives the total-effect sensitivity indices of each input parameter under the two market designs, calculated from (3.24). In both markets, variations in the maximum LOLE mostly arise from change in the energy price cap u_{voll} and its interactions with other input parameters. The second most important input in terms of its total-effect on variation in the maximum LOLE is the energy price markup parameter θ_{markup} in an energy-only market, while in a capacity market, it is the long-run level of carbon prices u_{co_2}.

Overall, both the main-effect and the total-effect sensitivity indices show that in terms of impact on the maximum LOLE, the administratively determined parameter

u_{voll} is the most important model input. The setting of VOLL can affect whether a capacity market has a net cost or benefit. The choice of u_{cone} also has an effect on the economic impact of a capacity market. This is relevant to the question of how to achieve the target system reliability at the lowest cost. Estimates of LOLE are very dependent on estimates assigned to the valuations of VOLL, net CONE and carbon prices, indicating the role that energy policy plays in shaping investment decisions. From the perspective of a private investor, to reduce uncertainty in system reliability, it seems that reducing future policy risks faced by private companies in power generation would be a good strategy. Moreover, uncertainty in the price markup θ_{markup}, reflecting the competitiveness of electricity market or the level of scarcity rent at peak demand hours, deserves a careful study.

4.6 Generality of the Bayesian framework

The GB case study has provided a full exemplar of how the Bayesian framework can be applied in a large, complex problem, showing all important tasks of calibration, validation, uncertainty analysis and sensitivity analysis. As mentioned in Section 3.3.1, the Bayesian framework is very useful for those who rely on models to understand complex processes involving uncertainty, and those who wish to know how much they can trust the model outputs. The unique characteristics found in different models suggest that the emulator must be tuned for each particular model, e.g. the dimensionality of the input and the output space, and prior distributions for inputs, the emulator and the discrepancy. Besides, different parametric forms can be chosen for the emulator and the model discrepancy to reflect the characteristics of a particular model.

Examples of power system models to which Bayesian emulation could be applied include: models discussed in [58, 167], which are used in GB Electricity Market Reform; a model presented in [20] to study the capacity market proposal of PJM; and models described in [14, 16–19] which are used for academic and industrial studies of electricity markets. The particular uncertainty modelling methods used in these studies are different to those used in the LTGI model described in our work. For ex-

ample, methods for the modelling of uncertainties within the model (e.g., load, wind generation, fuel prices, policy options), energy dispatch (e.g., Dynamic Dispatch Model (DDM) [58], PLEXOS [167]) and investment logic may differ. A particular challenge for fitting an emulator to the DDM in [58] is the high-dimensional input space of this model, meaning that a careful selection of the inputs to be included in any emulator must be carried out. A key similarity between the DDM and the LTGI exemplar is that outputs for both models are in the form of a time-series, so the dimension reduction techniques and the form used for the model discrepancy described here could be applied directly to the DDM. The PJM capacity market model in [20] can be enhanced by employing the Bayesian emulation method to provide a more realistic and confident assessment of capacity market designs. Some model parameters (e.g., parameters in the utility function) that were originally chosen based on behavioral assumptions, can be more rigorously calibrated against historical observations of model outputs of interest by fitting an emulator to the model retaining the same main features but with an energy-only market design where historical data were observed.

4.7 Chapter summary

In our GB case study, for model inputs, uncertainty in regulatory decisions (e.g., energy price cap, net CONE and carbon prices), future gas prices, and calibration parameters (e.g., investor's risk preference and energy price markup) has been accounted for and specified with prior distributions. For model outputs, both the real-world observations and the simulated outputs of operational thermal capacities as well as simulated LOLE profiles have been considered. Uncertainty in model outputs explored in our work arises from not only input uncertainty but also uncertainty associated with the model (i.e., structural uncertainty and functional uncertainty) that are often ignored in the existing literature.

While considering these three major sources of uncertainty, we apply a Bayesian framework to the GB case study. Calibration has been carried out first to reduce uncertainty in some of the inputs whist quantifying the structural uncertainty by

learning from historical observations. Then the UA explores the output distribution of interest with reduced risks. Future projections of the long-term generation projections and LOLEs, as well as the maximum LOLE are obtained using a combination of the simulator, calibrated model parameters and model discrepancy. The calibration results and the predictions in the case study have been compared the validation performed by Eager *et al.* in [4, 13, 82], as discussed in Section 4.2.2. Moreover, we proposed a robustness index of market designs in the uncertain context and then conducted a quantitative robustness study by exploring the distribution of the maximum LOLE. SA results quantitatively show that the investment decisions and the maximum LOLE are most sensitive to the two factors affecting scarcity pricing – the energy price cap and the price markup, which imply by the model that the scarcity pricing policy plays an important role in shaping generation investment decisions and hence the long-term system reliability.

Chapter 5

Implications of operating reserve market designs

In this chapter, we discuss the market implications of the short-term operating reserve market design. The timing and locational values of operating reserves need to be rewarded to improve the deliverability and the price signal of operating reserves that are mostly provided by flexible generating resources in a transmission network. An enhanced deterministic unit commitment (EDUC) model incorporating hourly updated zonal operating reserve demand curves (Z-ORDCs) is proposed in this chapter. Reserve zones are hourly determined by the approach of spectral clustering. A case study on the implications of operating reserve market designs using the RTS-73 test system is given. Comparisons are made between the choices of reserve policies (e.g., single, seasonal or dynamic zones) and of reserve zone partitioning methods.

5.1 Introduction

To encourage a low-carbon and secure supply of electricity, it is argued that market incentives are needed for rewarding the ancillary services provided by relatively clean and flexible sources. Operating reserve is among the larger class of ancillary services. In terms of its value, operating reserve is a crucial service for ensuring that the day-ahead operational planning of generators' schedule can accommodate the uncertainty arising from unexpected variations in the load profile or faults in

generators, transformers and transmission links. In terms of its cost, operating reserve keeps spinning generators partially loaded (running at non-optimal operating points) in order to respond to system contingencies and variations or unpredictability in demand and variable generation (VG), such as wind and solar. A large portion of the operating reserve requirement is provided by thermal and hydropower plants in many countries. Alternative sources are demand response and energy storage, but currently they only contribute to a small fraction of the requirement. The additional variability and uncertainty arising from VG units may increase the total amount of and more frequent procurement of operating reserve. Therefore, a proper pricing mechanism for operating reserves is required when it is economic to do so. To achieve this, the potential value of these operating reserve services under uncertain system conditions must first be recognized.

5.1.1 Operating reserve

Different power systems use different terms and definitions for operating reserve, but the functionality is very similar. The term operating reserve is defined in our work as the capability above the firm demand needed by the system operator for regulation, balancing forecasting errors of load and VG (such as wind and solar), and dealing with faults of equipment (e.g., generators and transmission lines). Our definition of operating reserve is similar to that given by NERC[1].

According to the required response time (ramp rate and start time), operating reserves is classified as spinning reserves (10 minutes) and non-spinning reserves ($10 \sim 30$ minutes) with both upward and downward directions respectively [168]. The spinning reserve is the on-line reserve capacity that is provided by synchronized (committed) generators. The non-spinning reserve is the off-line generation capacity that can be synchronized to the grid within 10 minutes of an order by the system operator and that is able to maintain that output for at least two hours. Upward reserves from generators that are available to increase output and load that is allowed to curtail have historically been more often required for reliability purposes, due

[1]The North American Electric Reliability Corporation

to the sudden failure of large generators and transmission lines. As opposed to upward reserves, downward reserves have historically been less needed for reliability since it is far less likely to see that large loads suddenly disconnect. However, in power systems with a high penetration of variable generation/VG that increases and decreases output stochastically, both upward and downward reserves are valuable in assisting in generation and load balance. The value of operating reserve is illustrated at two dimensions - time and space, in Sections 5.1.2, 5.1.3.

5.1.2 Timing value of operating reserve

The timing value (the value depending on time) of operating reserves changes with the time-varying conditions of power system operation. Traditionally, in power systems with a low penetrations of VG, spinning and non-spinning reserves are designed with deterministic rules such as a percentage of the aggregate load and/plus the largest unit in the system [169]. However, new operating reserve methodologies may be required for coping with the additional variability and uncertainty of VG.

Probabilistic methods have been proposed in [170–172] for determining operating reserve requirements in the presence of wind power. The common approach is to explore the variability and the forecasting error distributions of the net load (load net of wind) in setting the operating reserve requirements. In the wind integration study carried out in Minnesota [171], more operating reserve margin was allocated for hourly forecasting errors in the net load, in order to account for the high variability of wind when wind generation is in the middle range of the rated capacity. In the "All Island Grid Study" in Ireland [173], the additional contribution for wind generation in addition to the size of the largest on-line unit is included in the spinning reserve requirement. The 3σ rule - 3 times the standard deviation of 10-minute changes (variability) in the net load has been suggested to cover the majority of expected variability [174]. The 3σ rule has a limitation of accounting for the tail of the net load distribution. The study in [174] explored the 3+5 rule, which requires the system to carry hourly spinning reserve no less than 3% of hourly forecasted load plus 5% of hourly forecasted wind generation, and the results show that the $3 + 5$ rule usually works well and the quantity of reserves held was always no less than

that by the 3σ rule. These deterministic operating reserve requirements based on some simple probabilistic analysis can easily been incorporated into a deterministic UC model. However, it is acknowledged here that these simple rules lack a rigorous scientific basis, and in particular that the standard deviation will not characterise the tail unless the net load distribution can reasonably be assumed normal.

An alternative to planning operating reserves is using stochastic programming. In stochastic programming, reserve requirements and generation schedules are co-optimized for achieving an economical objective that includes the cost/benefit in day-ahead and real-time markets. A two-stage scenario-based stochastic programming model, such as the one developed in [175], is a stylized stochastic UC model. In the first stage, the day-ahead operational costs and constraints are considered. In the second stage, some scenarios of real-time system operational conditions (e.g., realized load, wind generation and generator availability) are considered, so that the scenario-based operational cost included in the objective function acts as an approximation to the real-time operational and reliability costs. The case study in [175] shows that the results outperform those from peak-load-based reserve requirements and the 3+5 rule. However, stochastic programming-based UC models are often computationally expensive to solve. In addition, it is challenging to select and properly weight qualitative scenarios serving as inputs to a stochastic programming formulation.

In order to achieve a balance between considering the real-time system conditions and ensuring the computational efficiency in the UC model, a multi-segment operating reserve demand curve (ORDC), has been tested and found in operations in some US RTOs (e.g., MISO, ERCOT, PJM) [39]. The ORDC is used to represent the value of operating reserves by avoiding load shedding events [39, 176]. In ERCOT, the price (marginal value) of available operating reserves is calculated as the product of value of lost load (VOLL) and loss-of-load probability (LOLP). A dynamic (hourly or daily) system-wide ORDC has been designed for addressing the time-varying operating conditions in [177], where it has been found that in comparison with the traditional deterministic UC program, the one incorporating the ORDC produces less volatile energy and operating reserve prices and on aver-

age higher prices for the days without price spikes, which indicates that generators are exposed to less price risks. As opposed to the scenario-based stochastic programming, the ORDC approximates the change in real-time system reliability cost based on some probability analysis methods instead of scenarios. The reliability cost makes a significant proportion to the total system cost in real-time markets due to the very high VOLL. Therefore, an UC model incorporating an ORDC produces closer results to the stochastic program solution than a traditional UC model.

5.1.3 Locational value of operating reserve

The locational value of operating reserves emphasizes the deliverability of available operating reserves across a transmission network. Traditionally, operating reserve requirements incorporated in deterministic unit commitment models, such as the largest unit and the $3 + 5$ rule aiming at different types of operating reserve, are simply imposed to the entire transmission network, rather than a regional or a bus-level. However, it may happen that the procured reserve is not deliverable to a desired location within the network. To address this issue, the concept of reserve zones has been proposed and implemented in many day-ahead electricity markets (e.g., Midwest ISO (MISO), Electric Reliability Council of Texas (ERCOT), California ISO (CAISO) [178]), aiming for infrequent intrazonal congestion [179] by enforcing zonal reserve requirements rather than system reserve requirements [180].

The industrial practice of zonal configuration conducted by SOs varies due to lack of consensus on the methodology of reserve zone partitioning. Two popular partitioning methods are based on identifying zones with similar nodal prices (i.e., LMPs) and linear sensitivity factors (e.g., Power Transfer Distribution Factors (PTDFs)) [181].

In the LMP-based method, buses are separated into higher and lower LMP areas if congestion occurs, and then the buses with similar LMPs are clustered to define a zone [182, 183]. The LMP represents the cost of supplying extra 1 MW of energy to a particular bus in the transmission system. Without any transmission congestion, the LMPs across the transmission network will be equal (uniform) because the cost of producing the energy from a marginal power plant will be the same at a particular

time (at the same system-level of demand). When there is transmission congestion on the way to deliver energy, there is a cost component (congestion cost) added to the formation of LMP (as will be explained in Section 5.2.3), which makes the LMP at the bus where energy is delivered to different from any other LMP. The difference among LMPs arising from congestion cost makes LMPs a good basis for determining reserve zones.

In the PTDF-based method, the common procedure is divided into two steps. First, critical lines that are likely to be congested are identified based on expert knowledge [184], power flow information [185,186] or customized off-line studies [187]. Changes of injection in different buses (e.g., generation level, demand level) will influence the flows through critical lines to various levels. Second, buses with similar effects on the flows through critical lines are clustered into one zone. Since the PTDF represents the sensitivity relation between changes in injections or loads at a bus and changes in flows through a line, buses characterized by similar PTDFs with respect to critical lines desirably define a region of similar sensitivity to line congestion under various system conditions [188]. The sign of the PTDF provides a clear distinction whether a power injection in a particular bus increases or decreases the flow through a line. In the case that there is only one congested line that is identified to be critical, the two buses at the two ends of a congested line have a negative and a positive PTDF with respect to this line, respectively. The congested line becomes the zonal interface (the line connecting two zones) and the two buses will be separated into two zones (zone 1,2). If there is a sudden increase in the injection of a load bus in zone 1, extra energy needs to be transferred from generators that are located in zone 1, in order to prevent the congestion situation from worsening. The intra-zonal deliverability can be governed by intra-zonal congestion management. A desired zonal partition is supposed to lead to infrequent intra-zonal congestion.

Different variants of the PTDF-based method of zonal partition have been proposed in existing literature. In [189], the zones have been formed based on the combined effect of real and reactive line power flow sensitivity indexes. A procedure of sequential network partition and congestion contribution identification is

presented in [187], where connectivity of the partitioned zones in complex networks is guaranteed. However, the process of zonal partition in [187] involves solving a complex mixed integer non-linear programming model, which takes a large amount of computational time. In [185,186], the reserve zone partitioning method employed the k-means algorithm to group buses of similar values of weighted PTDFs. The weight is used for emphasizing the importance of a line according to its absolute level of power flow, with larger values assigned to lines that are more likely to be congested. The k-means clustering chooses a random rather than the optimal initialization. To address this limitation, the authors in [181] designed an algorithm that performs the zonal partition starting from a pre-partition with a reasonably large number of zones and then continuing to merge two of the closest zones until any desired number of zones is produced. A second limitation of k-means clustering is that the choice of k, denoting the number of clusters, is an input parameter. An inappropriate choice of k may yield poor results. A third limitation of k-means clustering is that it makes strong assumptions that the resulting clusters form disjoint convex sets of the underlying space (or, to be precise, the clustered sets are compact and have convex boundaries). In the situation where nested reserve zones are preferred as studied in [190], k-means clustering is incapable of delivering such form of the clusters. Therefore, a more general clustering method is needed for performing zone partition.

5.1.4 Contributions of this chapter

We propose an enhanced deterministic unit commitment (EDUC) model to incentivise both the timing and the locational value of operating reserves. This is achieved by introducing hourly zonal ORDCs (Z-ORDCs) into a deterministic UC model. The Z-ORDCs truthfully represent the marginal benefit of zonal operating reserves by considering the VOLL and the expected unserved energy or the expected curtailed wind while tackling the complexity of inter-zonal reserve sharing. The EDUC model produces co-optimized schedules for energy, intrazonal and inter-zonal reserves, for purposes of reducing out-of-market corrections and stabilizing market revenues for flexible generators as opposed to the traditional single-zone deterministic UC model.

The contributions of this chapter are three-fold.

- An EDUC model incorporating dynamic Z-ORDCs is developed in Section 5.2. This is an extension to the single-zone deterministic UC model incorporating a system-wide ORDC proposed in [177].
- A spectral clustering method is applied to reserve zone partitioning, as in Section 5.4.1. Spectral clustering is more general than k-means clustering used in [185] in the sense that spectral clustering does not make strong assumptions on the form of the clusters and hence can solve very general problems like nested reserve zones. Moreover, there is a general procedure for choosing the number of clusters in the spectral clustering method.
- The formulation of Z-ORDCs in Section 5.4.2 depends on decision variables of intrazonal reserve and inter-zonal reserve sharing as well as uncertain parameters of load, wind generation and conventional generator availability. By introducing the variable of inter-zonal reserve sharing to the deterministic UC model, a tradeoff is made between the value and the cost for each zone with the optimal amount of reserve sharing.

The remainder of this chapter is organized as follows. Section 5.2 presents the model formulation of the EDUC and the pricing of energy and operating reserves. Section 5.3 describes the uncertainty modelling of load, wind, and generator availability. Section 5.4 discusses the reserve zone partitioning and the development of Z-ORDCs. Section 5.5 shows a case study on the RTS-73 test system. Section 5.6 concludes this chapter.

5.2 Model formulation and pricing

5.2.1 Formulation of the enhanced deterministic UC model

The model formulation of the EDUC is expressed below in (5.1), with an objective of minimizing the market-wide social cost as in (5.1a) subject to generator-level, zonal-level, and system-level constraints as in (5.1b)–(5.1y).

$$\min \sum_{h\in \mathrm{H}} \sum_{c\in \mathrm{C}} \left[VC_c(u_{c,h}\underline{P}_c + p_{c,h}) + FC_c u_{c,h} + SU_c v_{c,h} + v^{wc}_{z_h} + v^{eue}_{z_h} \right], \tag{5.1a}$$

$$\text{s.t.} \quad (\lambda_{b,h} \geq 0) : \sum_{l\in \mathrm{L}|\mathrm{B}^+_l = b} f_{l,h} - \sum_{l'\in \mathrm{L}|\mathrm{B}^-_{l'} = b} f_{l',h} + \sum_{c\in \mathrm{C_b}} (\underline{P}_c u_{c,h}$$

$$+ p_{c,h}) + \sum_{w\in \mathrm{W_b}} P^w_{w,h} = D_{b,h}, \forall h \in \mathrm{H}, b \in \mathrm{B}, \tag{5.1b}$$

$$f_{l,h} = B_l(\theta_{b,h} - \theta_{b',h}), \forall h \in \mathrm{H}, l = (b, b') \in \mathrm{L}, \tag{5.1c}$$

$$(\eta_{l,h} \leq 0) : -\bar{F}_l \leq f_{l,h} \leq \bar{F}_l, \forall l \in \mathrm{L}, \tag{5.1d}$$

$$\theta^{min}_b \leq \theta_{b,h} \leq \theta^{max}_b, \forall h \in \mathrm{H}, b \in \mathrm{B}, \tag{5.1e}$$

$$\theta_{\mathrm{Ref},h} = 0, \forall h \in \mathrm{H}, \tag{5.1f}$$

$$0 \leq p_{c,h} \leq u_{c,h}\bar{P}_c - \underline{P}_c, \forall h \in \mathrm{H}, c \in \mathrm{C}, \tag{5.1g}$$

$$v_{c,h} - y_{c,h} = u_{c,h} - u_{c,t-1}, \forall h \in \mathrm{H}, c \in \mathrm{C}, \tag{5.1h}$$

$$v_{c,h} + y_{c,h} <= 1, \forall h \in \mathrm{H}, c \in \mathrm{C}, \tag{5.1i}$$

$$\sum_{k=t-UT_c+1}^{t} v_{c,k} \leq u_{c,h}, \forall t \in (UT_c, \ldots, \mathrm{T}), c \in \mathrm{C}, \tag{5.1j}$$

$$\sum_{k=t-DT_c+1}^{t} y_{c,k} \leq 1 - u_{c,h}, \forall t \in (DT_c, \ldots, \mathrm{T}), c \in \mathrm{C} \tag{5.1k}$$

$$-60R^{60}_c \leq p_{c,t-1} - p_{c,h} \leq 60R^{60}_c, \forall h \in \mathrm{H}, c \in \mathrm{C}, \tag{5.1l}$$

$$r^{\rho}_{c,h} \leq \tau_\rho R^{\tau_\rho}_c, \forall h \in \mathrm{H}, c \in \mathrm{C}, \rho \in \{su, sd, nsu, nsd\}, \tag{5.1m}$$

$$p_{c,h} + r^{su}_{c,h} + r^{nsu}_{c,h} \leq (\overline{P}_c - \underline{P}_c)(u_{c,h} - y_{c,t+1}), \forall h \in \mathrm{H}, c \in \mathrm{C}, \tag{5.1n}$$

$$p_{c,h} - r^{sd}_{c,h} - r^{nsd}_{c,h} \geq 0, \forall h \in \mathrm{H}, c \in \mathrm{C}, \tag{5.1o}$$

$$(\alpha^{\rho,zq}_{z_h} \geq 0) : r^{\rho}_{z_h} \geq R^{\rho,req}_{z_h}, \forall h \in \mathrm{H}, \rho \in \{su, sd, nsu, nsd\}, z_h \in \mathrm{Z_h}, \tag{5.1p}$$

$$(\alpha^{u,sq}_{c,h} \geq 0) : \sum_{z_h\in \mathrm{Z_h}} r^u_{z_h} \geq \underline{P}_c u_{c,h} + p_{c,h} + r^{su}_{c,h} + r^{nsu}_{c,h}, \forall h \in \mathrm{H}, c \in \mathrm{C}, \tag{5.1q}$$

$$(\alpha^{\rho,ro}_{z_h} \leq 0) : O^\rho r^u_{z_h} <= \sum_{c\in z_h(c)} r^{\rho}_{c,h}, \forall h \in \mathrm{H}, z_h \in \mathrm{Z_h}, \rho \in \{su, nsu\}, \tag{5.1r}$$

$$(\alpha^{eue}_{z_h,k} \geq 0) : v^{eue}_{z_h} \geq A^{eue}_{z_h,k} r^u_{z_h} + B^{eue}_{z_h,k}, \forall h \in \mathrm{H}, k \in \mathrm{K}, z_h \in \mathrm{Z_h}, \tag{5.1s}$$

$$(\alpha^{wc}_{z_h,q} \geq 0) : v^{wc}_{z_h} \geq A^{wc}_{z_h,q} r^d_{z_h} + B^{wc}_{z_h,q}, \forall h \in \mathrm{H}, q \in \mathrm{Q}, z_h \in \mathrm{Z_h}, \tag{5.1t}$$

$$r^u_{z'_h,z_h} \leq \bar{r}_{z'_h,z_h}, \forall h \in \mathrm{H}, z_h, z'_h \in \mathrm{Z_h}, z'_h < z_h, \tag{5.1u}$$

$$r^u_{z'_h,z_h} \geq -\bar{r}_{z_h,z'_h}, \forall h \in \mathrm{H}, z_h, z'_h \in \mathrm{Z_h}, z'_h < z_h, \tag{5.1v}$$

$$(\alpha^u_{z_h} \leq 0) : r^u_{z_h} = \sum_{c \in z_h(c)} (r^{su}_{c,h} + r^{nsu}_{c,h}) + \sum_{z'_h \in \mathrm{Z_h} | z'_h < z_h} r^u_{z'_h,z_h} - \sum_{z'_h \in \mathrm{Z_h} | z'_h > z_h} r^u_{z_h,z'_h}, \forall z_h \in \mathrm{Z_h}, \tag{5.1w}$$

$$(\alpha^d_{z_h} \leq 0) : r^d_{z_h} = \sum_{c \in z_h(c)} (r^{sd}_{c,h} + r^{nsd}_{c,h}), \forall z_h \in \mathrm{Z_h}, \tag{5.1x}$$

$$(\alpha^{zz}_{z'_h,z_h} \leq 0) : \bar{r}_{z'_h,z_h} = \sum_{l \in \mathrm{I}_{z'_h,z_h}} \left(\bar{F}_l - \delta_l f_{l,h} + (1-\delta_l) f_{l,h}\right), \forall z_h, z'_h \in \mathrm{Z_h}, \tag{5.1y}$$

where,

$$A^{eue}_{z_h,k} = \frac{v^{eue}_{h,k} - v^{eue}_{h,k+1}}{r^u_{z_h,k} - r^u_{z_h,k+1}}, B^{eue}_{z_h,k} = -r^u_{z_h,k} A^{eue}_{z_h,k} + v^{eue}_{h,k}, \tag{5.2a}$$

$$A^{wc}_{z_h,q} = \frac{v^{wc}_{h,q} - v^{wc}_{h,q+1}}{r^d_{z_h,q} - r^d_{z_h,q+1}}, B^{wc}_{z_h,q} = -r^d_{z_h,q} A^{wc}_{z_h,q} + v^{wc}_{h,q}. \tag{5.2b}$$

$$p_{c,h}, r^{\rho}_{c,h}, r^u_{z_h}, r^d_{z_h} \geq 0, r^u_{z'_h,z_h} \in \mathbb{R}, u_{c,h}, v_{c,h}, y_{c,h} \in \{0,1\}, \tag{5.2c}$$

The social cost in a traditional deterministic UC has a major component of the operational cost that includes variable, no-load, and start-up costs of conventional generators as formulated as the first three terms in (5.1a). By contrast, two new objective terms $v^{eue}_{z_h}, v^{wc}_{z_h}$ representing the zonal values of upward and downward operating reserves has been added to the EDUC model. The Z-ORDCs are represented by the piecewise linear opportunity cost functions of zonal reserves; these functions are further replaced by decision variables $v^{eue}_{z_h}, v^{wc}_{z_h}$ and a set of linear constraints (5.1s) (5.1t) that form a convex basin using the constrained cost variable technique described in [191]. The parameters in the linearised constraints are expressed in (5.2a) (5.2b), and further explanations are given in Section 5.4.2. Similar practice in formulating the relationship between system-wide reserves and risk costs can be found in [192].

Apart from the new constraints of formulating piecewise linear Z-ORDCs, constraint (5.1p) places a lower limit (e.g., 6% of forecasted load plus 10% wind generation) on the upward operating reserve (both spinning and non-spinning reserve) at each zone in order to protect the local area from contingencies [193]. Constraint (5.1q) is the system-wide N-1 reserve requirement enforcing that the total available upward reserve is larger than the output plus the upward reserve provided

by any generating unit. Constraint (5.1r) ensures that both the spinning and non-spinning reserve supplied by generators accounts for a certain ratio (e.g., 50%) of the total available upward operating reserve at each zone. Equations (5.1w) define that the total available upward operating reserves in a zone equal those provided by conventional generators located in the same zone plus the imported reserve and minus the exported reserve to any other zone. Equations (5.1x) define that the downward operating reserves in a zone all come from conventional generators located in the same zone. Constraints (5.1u) (5.1v) set limits on the inter-zonal upward reserve sharing $r^u_{z'_h,z_h}$ between z'_h and z_h with a positive or negative direction. We define reserve delivered out of zone z'_h into zone z_h as positive, and reserve delivered into zone z'_h out of zone z_h as negative. The limit of reserve sharing from zone z'_h to z_h, denoted as $\bar{r}_{z'_h,z_h}$, can rely on off-line analysis or expert suggestions [186, 194]. Here, $\bar{r}_{z'_h,z_h}$ is treated as a variable dependent on the line flows given in (5.1y), where δ_l equals 1 if $B_l^+ \in z_h(b)$, otherwise δ_l equals 0 if $B_l^- \in z_h(b)$.

The remaining constraints have been commonly included in the traditional single-zone deterministic UC formulation. Constraint (5.1b) is enforced to restore the nodal power balance at any time. Equation (5.1c) are linearised power flow equations. Transmission constraints (5.1d) prevent over-loaded lines. Constraints (5.1e) set the limits of the voltage angles, and constraint (5.1f) sets the voltage angle at the reference bus as zero.

Constraints (5.1g)–(5.1l) include generation limits, on/off switch, on/off state, minimum-up time, minimum-down time, and the hourly ramping constraints, as in [195]. Constraints (5.1m) specify the limit of each research type ρ at the generator level, denoted as $R_c^{\tau_\rho}$. For example, the maximum amount of the upward and downward spinning reserve is set as one sixth of its hourly ramping limit if the unit is committed. The unit-level limit of the upward and downward non-spinning reserve is set as half of its hourly ramping limit, for a fast generator at all times and for a slow generator only when it is committed. Constraints (5.1n) and (5.1o) ensure that the total amount of power output and upward/downward operating reserves meets the minimum and maximum generation limitations as in [195, 196].

In the formulation (5.1), η, α are the dual variables for constraints (5.1b) and (5.1p)–

(5.1y), respectively. The decision variables are reserve sharing $r^u_{z_h,z'_h}$, as well as the generation output, the reserve provision, the commitment, the start-up and the shut-down variables of conventional generators, denoted as $p_{c,h}, r^\rho_{c,h}, u_{c,h}, v_{c,h}, y_{c,h}$, respectively.

5.2.2 Traditional deterministic UC model

The model formulation of a traditional single-zone deterministic UC model is expressed below in (5.3), with an objective of minimizing the market-wide operational cost, subject to generator-level and system-level constraints.

$$\min \sum_{h\in \mathrm{H}} \sum_{c\in \mathrm{C}} \left[VC_c(u_{c,h}\underline{P}_c + p_{c,h}) + FC_c u_{c,h} + SU_c v_{c,h}\right], \tag{5.3a}$$

$$\text{s.t. } (5.1\text{b}) - -(5.1\text{o}) \tag{5.3b}$$

$$(\alpha^{\rho,s} \geq 0) : r^{\rho,s} \geq R^{\rho,req}, \forall h \in \mathrm{H}, \rho \in \{su, sd, nsu, nsd\} \tag{5.3c}$$

$$(\alpha^s_{c,h} \geq 0) : \sum_{c'\in \mathrm{C}} r^{su}_{c',h} + r^{nsu}_{c',h} \geq \underline{P}_c u_{c,h} + p_{c,h} + r^{su}_{c,h} + r^{nsu}_{c,h}, \forall h \in \mathrm{H}, c \in \mathrm{C} \tag{5.3d}$$

$$(\alpha^{\rho,so} \leq 0) : O^\rho \sum_{c\in \mathrm{C}} (r^{su}_{c,h} + r^{nsu}_{c,h}) <= \sum_{c\in \mathrm{C}} r^\rho_{c,h}, \forall h \in \mathrm{H}, \rho \in \{su, nsu\} \tag{5.3e}$$

5.2.3 Energy and operating reserve pricing

Both the EDUC model formulated as (5.1) and the traditional deterministic UC model as in (5.3) are mixed integer programs. After solving the mixed integer program, integer variables are fixed as constants and so the mixed integer program is replaced with a linear program, from which the energy and operating reserve prices are derived. In the day-ahead energy and operating reserve market, the locational market prices (LMP) at each bus b and the reserve market clearing price (RMCP) of each type ρ provided by resource c are cleared together and calculated on an hourly basis.

The LMP is defined as the marginal cost (i.e., the change in the objective function) of supplying an extra unit (e.g., MWh) of electric energy at a specific bus b while considering the marginal cost of electric suppliers and the physical transmission network [197]. Similarly, the RMCP is the marginal cost of providing an extra unit (e.g., MWh) of a specific reserve product ρ from a resource (here referring to a

conventional generator) c while considering the reserve requirements. Based on the marginal pricing concept [198], the LMPs or the RMCPs are derived from the dual variables (i.e., the Lagrange multipliers) of energy or reserve related constraints in the linear program in replace of the UC program, respectively. The dual variable of a constraint gives the change in the optimal value of the objective function (profit or cost) due to the relaxation of a given constraint (i.e., through a small change in the constraint parameter).

Regarding the EDUC model, the LMPs and RMCPs are calculated using,

$$LMP_{b,h} = [\lambda_{b,h}] + \left[\sum_{l} (-\eta_{l,h} \frac{\partial f_{l,h}}{\partial D_{b,h}}) + \sum_{z_h, z'_h \in \mathrm{Z}} \left(\alpha^{zz}_{z'_h, z_h} \sum_{l' \in \mathrm{I}_{z'_h, z_h}} \frac{\partial f_{l',h}}{\partial D_{b,h}} \right) \right], \tag{5.4}$$

$$RMCP^{\rho}_{c,h} = \begin{cases} \sum\limits_{z_h | c \in z_h(c)} \left[\sum\limits_{c' \in \mathrm{C}} \alpha^{u,sq}_{c',h} - O^{\rho} \alpha^{\rho,ro}_{z_h} + \alpha^{\rho,zq}_{z_h} - \alpha^{u}_{z_h} + \sum_k (\alpha^{eue}_{z_h,k} A^{eue}_{z_h,k}) \right], \\ \text{if } \rho = \{su, nsu\}, \\ \sum\limits_{z_h | c \in z_h(c)} \left[\alpha^{\rho,zq}_{z_h} - \alpha^{d}_{z_h} + \sum_q (\alpha^{wc}_{z_h,q} A^{wc}_{z_h,q}) \right], \text{if } \rho = \{sd, nsd\}. \end{cases} \tag{5.5}$$

The LMP shown in (5.4) is defined for each bus location b. Ignoring marginal losses, the LMP includes two cost components of energy and congestion [190,199] as shown in the two pairs of square brackets. The dual variable of the power balance equation (5.1b), $\lambda_{b,h}$, represents the additional energy cost of supplying one extra unit of load demand at bus b. The derivative $\frac{\partial f_{l,h}}{\partial D_{b,h}} = J_{l,b}$ represents the PTDF between bus b and line l, that is, sensitivity of the power flow through the lth transmission line to the load level at bus b. The dual variable of (5.1d), $\eta_{l,h}$, denotes sensitivity of the congestion cost to the transmission limit of line l. The dual variable of (5.1y), $\alpha^{zz}_{z'_h, z_h}$, indicates sensitivity of the congestion cost to the transmission limit of the interface between two reserve zones. When one or more of power flows are constrained, LMPs will vary among buses. Otherwise, LMPs are the same at all buses.

The RMCP of each generator defined in (5.5) is determined by the dual variables of constrains (5.1p)–(5.1x) relating to the decision variables $r^{u}_{z_h}, r^{d}_{z_h}$ as well as the coefficients of $r^{u}_{z_h}, r^{d}_{z_h}$ in these constraints. It is shown that $RMCP^{\rho}_{c,h}$ are identical when operating reserve resources c are located in the same reserve zone. Even when

there are sufficient supply of reserves (no reserve shortage or energy re-dispatch), reserve prices might not be zero due to the cost of risks reflected by the Z-ORDCs, as expressed as $\sum_k(\alpha^{eue}_{z_h,k}A^{eue}_{z_h,k})$ in (5.5).

Regarding the traditional single-zone deterministic UC model, the LMPs and RMCPs are derived from the following equations,

$$LMP^s_{b,h} = [\lambda_{b,h}] + \left[\sum_l(-\eta_{l,h}\frac{\partial f_{l,h}}{\partial D_{b,h}})\right], \tag{5.6}$$

$$RMCP^{\rho,s}_{c,h} = \begin{cases} \left[\sum\limits_{c'\in \mathrm{C}} \alpha^s_{c',h}\right] + \alpha^{\rho,s} - O^\rho\alpha^{\rho,so}, \text{if } \rho = \{su, nsu\} \\ \alpha^{\rho,s}, \text{if } \rho = \{sd, nsd\} \end{cases} \tag{5.7}$$

The LMP shown in (5.6) includes the energy cost and the congestion cost that are the same as the first two components in (5.4). The price of upward spinning/nonspinning reserve shown in (5.7) is determined by the dual variables of the system-wide requirement (5.3c), the N-1 constraint (5.3d) and the ratio of upward reserve (5.3e). The price of downward spinning/nonspinning reserve shown in (5.7) is determined by the dual variables of the system-wide requirement (5.3c).

5.3 Uncertainty modelling

Modelling uncertainty in load, wind generation and generator availability is the first stage of this work. The uncertainty models will further be used in the zonal partition and the development of Z-ORDCs.

Load uncertainty mainly arises from the day-head forecasting errors. The normalized day-ahead forecasting error at each bus usually has been modelled as independent, identically and normally distributed random variables with mean zero and a given standard deviation, such as in [177, 200]. This uncertainty model of load forecasting errors is sufficient for the purpose of our work, because our intention is to present the methodology of developing zonal, dynamic and elastic reserve demand curves that is general to all types of forecasting models of load and wind generation.

Uncertainty in the availability of a conventional generator is mainly due to forced outages. The cumulative distribution function of the total available conventional generation is discrete and represented by a capacity outage probability table

(COPT). The COPT is used for the probability density distribution of the random change in the power system while developing the Z-ORDCs (see Section 5.4.2). The COPT does not capture the chronological characteristics of generator outages, while a chronological model such as a two-state Markov process model of generator outages produces more realistic results in short-term power system operational planning. To account for the mismatch between the real-world and the forecasted generator availability, a two-state Markov process model is chosen for generating post-contingency scenarios of generator outages in the real-time market, as described in Section 5.4.1. For more details on the Markov model of generator outages, we refer to [201].

The day-ahead forecast uncertainty of wind power has been characterized as a normal distribution in [202, 203]. In comparison with the normal distribution, a beta distribution is more appropriate to model the normalized wind generation $\hat{p}^w_{w,h}$ that lies within the interval $[0, 1]$ [204], because the beta distribution accounts for its bounded and fat-tailed nature [204–206]. Here, the mean is assumed to be the forecasted wind power generation and the standard deviation depends on the forecasted level of wind power with respect to the rated wind power, as in [204],

$$\begin{aligned} \mu(\hat{p}^w_{w,h}) &= \hat{p}^w_{w,h} = P^w_{w,h}/C^w, \\ \sigma(\hat{p}^w_{w,h}) &= 0.03 + 0.15(P^w_{w,h}/C^w). \end{aligned} \tag{5.8}$$

Correlation, including that between load/wind uncertainty at different locations and that between load and wind uncertainty is not considered in our current work, but it will be explored in terms of improving the method of reserve zone partitioning by accounting for the correlation of post-contingency power flows (as described in Section 5.4.1) in order to match closer to the real-world.

5.4 Development of ZORDCs

5.4.1 Determination of reserve zones

The matrix $\mathbf{K}$ formed by PTDFs can be used by SOs for defining reserve zones, each comprising the buses of similar PTDFs and hence similar sensitivity to line flows. A finer metric "Weighted PTDF" was proposed in [185] to emphasize critical lines

(i.e., transmission bottlenecks) and diminish the least important lines, especially in the context of stochastic operating conditions due to large amounts of wind. A new partitioning method of reserve zones is presented here.

Spectral clustering approach

Spectral clustering, as one of the graph theory-based methods [207] is employed to separate points that is represented in form of the similarity graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with vertices $\mathcal{V}$ and edges $\mathcal{E}$. In the reserve zone partitioning, there are a set of buses $b_1, \ldots, b_{N_b}$ representing vertices and some notion of similarity (weights of the edge of two vertices are connected) $K_{bb'} := K(b, b') \geq 0$ between all pairs of buses b_i and b_j. The goal of clustering is to divide the buses into several groups such that buses in the same group are similar and buses in different groups are dissimilar to each other. Spectral clustering uses the eigenvectors of a graph Laplacian matrix derived from the data, which changes the representation of a large amount of data from the high-dimensional original space to a low-dimensional space so that clusters can be easily detected in the new representation. Whereas, traditional k-means clustering deals with the data points directly in the original data space. By using a low-dimensional representation of the data space, spectral clustering often performs more efficient than traditional k-means clustering, even for large data sets [208].

The difficulty in implementing the spectral clustering lies in choosing a sparse similarity graph. Once the similarity graph is in hand, the remaining work of spectral clustering is to solve a simple linear problem, as opposed to k-means, where there are problems of getting stuck in local minima or finding the optimal initializations by restarting the algorithm for several times. Spectral clustering can be used in very general problems with non-convex data that k-means is incapable of dealing with; this is particularly useful in delivering nested reserve zones.

Spectral clustering uses eigenvectors of matrices derived from the similarity matrix of points, thereby projecting the original high-dimensional data into a low-dimensional space that can be easily clustered. For details on the theory of spectral partitioning, see [208]. The similarity matrix $\mathbf{K} := (K_{bb'})_{b,b' \in \mathrm{B}}$ used in spectral

clustering is symmetric with its values defined as:

$$K_{bb'} = \begin{cases} \exp\big(-\omega\|\mathbf{K}_b - \mathbf{K}_{b'}\|^2/2\big) & \text{if } b \neq b' \\ 0 & \text{if } b = b', \end{cases}$$

where,

$$\omega\|\mathbf{K}_b - \mathbf{K}_{b'}\|^2 = \sum_{l\in L} \omega_l (K_{bl} - K_{b'l})^2, \tag{5.9}$$

where $\|\mathbf{K}_b - \mathbf{K}_{b'}\|$ is the Euclidean distance between any two buses b, b'; and vectors $\mathbf{K}_b$ and $\mathbf{K}_{b'}$ denote the corresponding columns of $\mathbf{K}$ in N_l-dimensional space. Therefore, similarity of buses will be examined by comparing their distances using the metric of Euclidean distance. The line vector $\omega := \{\omega_l\}$ assigns larger weights to critical lines. Small values of $K_{bb'}$ will be given to two nodes whose injections have similar influences on those identified critical lines. The above definition of $\mathbf{K}$ follows Gaussian kernel similarity function [208].

The objectives of clustering are to find a partition of the similarity graph such that 1) the edges between different groups have a very low weight (which means that the PTDFs of buses in different zones are dissimilar to each other); and that 2) the edges within a group have high weight (the PTDFs of buses in the same cluster are similar to each other). To satisfy both objectives, a spectral clustering algorithm based on the normalized graph Laplacian described in [208] is employed (see Appendix A.2 for more details). The normalized graph Laplacian matrix is defined as

$$\mathbf{L_{sym}} := \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{K}\mathbf{D}^{-1/2}, \tag{5.10}$$

where $\mathbf{D}$ is the degree matrix defined as the diagonal matrix with the degrees $d_1, \ldots, d_{N_b}$ with $d_i = \sum_{j=1} N_b K_{ij}$, and $\mathbf{L} := \mathbf{D} - \mathbf{K}$ is the unnormalized graph Laplacian matrix.

Given the input of the similarity matrix $\mathbf{K} \in R^{N_b \times N_b}$, the outputs are bus clusters $B_1, \ldots, B_{N_z}$ with $B_z = \{b | y_b \in C_z\}$. To enforce clustering stability, the number of clusters N_z is chosen such that all the first N_z eigenvalues $\lambda_1, \cdots, \lambda_{N_z}$ of the normalized Laplacian matrix are very small but the λ_{N_z+1} is relatively large. That is, find N_z that maximizes the eigengap $|\lambda_{N_z+1} - \lambda_{N_z}|$; this is justified by perturbation theory or spectral graph theory [208]. Basically, the first N_z eigenvectors of the

Laplacian matrix carry all the information about the N_z connected components in the graph that we wish to cluster, exactly one eigenvector possessed by each connected component. Hence, the first N_k eigenvalues represent the features of the N_z connected components. According to the perturbation theory, spectral clustering work better with larger eigengap, because the eigenvectors of the idea case (where the graph has N_z connected components) and the perturbed case (where the N_z connected components are not completely disconnected) will be closer. The eigengap heuristic works well when the data contains very well pronounced clusters. In ambiguous cases where there is no well-defined gap (e.g., there exist approximately the same eigengap between all eigenvalues), the eigengap heuristic returns ambiguous results as other clustering methods do.

Determination of line weights

Prior to implementing spectral clustering, the line weight vector ω in (5.4.1) and (5.9) is determined. The procedure of calculating line weights starts from generating a considerable number of post-contingency scenarios of hourly load and wind generation from their day-ahead forecast distributions as well as generator outages from two-state Markov process models using sequential Monte Carlo simulations, then performs the optimal power flow that determines the least-cost conventional generation scheme for each post-contingency scenario $s \in \mathrm{S}$, and finally derives the post-contingency power flow $\hat{f}_{l,s}$ and the lagrangian multiplier $\hat{\eta}_{l,s}$ assigned to each line constraint.

The optimal power flow procedure used here is a linear program, as follows,

$$\min \quad \sum_{c\in\mathrm{C}} VC_c\hat{p}_{c,s} + V^{wc}\sum_{w\in\mathrm{W}}\hat{p}^{wc}_{w,s} - V^{lol}\sum_{b\in\mathrm{B}}\hat{p}^{ls}_{b,s}, \tag{5.11a}$$

$$\text{s.t.} \quad \sum_{l\in\mathrm{L}|\mathrm{B}_\mathrm{l}^+=b}\hat{f}_{l,s} - \sum_{l\in\mathrm{L}|\mathrm{B}_\mathrm{l}^-=b}\hat{f}_{l',s} + \sum_{c\in\mathrm{C_b}}\hat{p}_{c,s} - \sum_{w\in\mathrm{W_b}}(P^w_w - \hat{p}^{wc}_{w,s}) = D_b - \hat{d}^{ls}_{b,s}, \forall b \in \mathrm{B} \tag{5.11b}$$

$$0 \le \hat{p}^{wc}_{w,s} \le P^w_w, \forall h \in \mathrm{H}, w \in \mathrm{W} \tag{5.11c}$$

$$\hat{f}_{l,s} = B_l(\theta_b - \theta_{b'}), \forall h \in \mathrm{H}, l = (b, b') \in \mathrm{L} \tag{5.11d}$$

$$-\bar{F}_l \le \hat{f}_{l,s} \le \bar{F}_l, \forall l \in \mathrm{L} \tag{5.11e}$$

$$\bar{P}_c \leq \hat{p}_{c,s} \leq \underline{P}_c, \tag{5.11f}$$

where the objective of optimal power flow is to minimize the generation costs and the penalty costs of involuntary load shedding and wind curtailment. Note that only committed generators in the scheduling stage are accounted for in the optimal power flow procedure.

Instead of defining a utilization metric (e.g., $\omega_l := \mu(s, \frac{|\hat{f}_{l,s}|}{\bar{F}_l}) + \sigma(s, \frac{|\hat{f}_{l,s}|}{\bar{F}_l})$) as line weights as in [185, 186], both line utilization and congestion cost are considered in the definition of line weights in (5.13). The use of line weights identifies not only the often-congested lines but also the critical lines whose congestion is relatively expensive to manage. The first term in (5.13) defines the line utilization as the ratio of post-contingency power flow to the transmission limit. The second term in (5.13) uses the dual variable $\hat{\eta}_{l,s}$ to describe the total added cost of shifting generation necessary to alleviate the congestion on line l. Like the line utilization, the congestion cost term is also transformed to the range $(0, 1)$ dividing by the maximum congestion cost on all lines at any hour. To reflect the relative importance of line utilization and congestion cost in the definition of line weights, a reasonable (or preferred) proportion $\varepsilon \in (0, 1)$ is assumed for line utilization and the remaining proportion $(1 - \varepsilon)$ for congestion cost.

$$\omega_l = \mu(s, \hat{\omega}_{l,s}) + \sigma(s, \hat{\omega}_{l,s}), \tag{5.12}$$

where

$$\hat{\omega}_{l,s} = \varepsilon \frac{|\hat{f}_{l,s}|}{\bar{F}_l} + (1 - \varepsilon) \frac{|\hat{\eta}_{l,s}|}{max(l \in \mathrm{L}, |\hat{\eta}_{l,s}|)}, \forall l \in \mathrm{L}, s \in \mathrm{S}. \tag{5.13}$$

5.4.2 Development of zonal operating reserve demand curves

The Z-ORDCs are developed here to represent the incremental values of upward and downward operating reserves in terms of their contributions to the reduction in the zonal value of expected unserved energy (Z-VEUE) and the zonal value of expected curtailed wind (Z-VECW).

The Z-VEUE is the product of VOLL and the zonal expected unserved energy

(ZEUE), expressed as a function of the total upward available reserve,

$$\begin{aligned} \text{Z-VEUE}_{z_h}(r^u_{z_h}) &= V^{lol} E\left[max(0, \Delta M_{z_h} - r^u_{z_h})\right] \\ &= V^{lol} \int_0^{r^u_{z_h}} (x - r^u_{z_h}) \phi_{z_h}(x) dx, \end{aligned} \tag{5.14}$$

where $\phi_{z_h}(\cdot)$ denotes the probability mass function of the system imbalances ΔM_{z_h} at zone z_h, that is, the forecasting error of net load minus generator outages, and $E\left[max(0, \Delta M_{z_h} - r^u_{z_h})\right]$ accounts for the excess of positive system imbalances over the amount of upward operating reserves provided by conventional generators.

Similarly, the Z-VECW denotes the product of the value of curtailed wind and the expected curtailed wind energy and is expressed as a function of the total downward available reserve as follows,

$$\begin{aligned} \text{Z-VECW}_{z_h}(r^d_{z_h}) &= V^{wc} E\left[max(0, -\Delta M_{z_h} - r^d_{z_h})\right] \\ &= V^{wc} \int_{-\infty}^{-r^d_{z_h}} (-x - r^d_{z_h}) \phi_{z_h}(x) dx, \end{aligned} \tag{5.15}$$

where $E\left[max(0, -\Delta M_{z_h} - r^d_{z_h})\right]$ denotes the expected curtailment of wind generation, that is, the excess of negative system imbalances M_{z_h} over the amount of downward operating reserves $r^d_{z_h}$ provided by conventional generators.

With sufficient regularity assumptions, the Z-VEUE as a function of zonal upward operating reserve is linearised at some input points $r^u_{z_h,k}$ where the corresponding output points are $v^{eue}_{z_h,k}$. The same practice is applied to the Z-VECW. These piecewise linear functions are expressed as,

$$\begin{aligned} v^{eue}_{z_h,k} &= A^{eue}_{z_h,k} r^u_{z_h,k} + B^{eue}_{z_h,k}, \quad \text{if} r^d_{z_h,k} \le r^d_{z_h} \le r^d_{z_h,k+1}, \forall k \in K \\ v^{wc}_{z_h,q} &= A^{wc}_{z_h,q} r^d_{z_h,q} + B^{wc}_{z_h,q}, \quad \text{if} r^d_{z_h,q} \le r^d_{z_h} \le r^d_{z_h,q+1}, \forall q \in Q. \end{aligned} \tag{5.16}$$

from where the coefficient and constant parameters are derived as shown in (5.2a) (5.2b).

The linearised constraints (5.1s) (5.1t) are derived by employing the constrained cost variable technique as described in [191, 192]. By incorporating the Z-ORDCs into the EDUC model, a tradeoff is made between normal energy dispatch and reservation of generator capacities as well as interface capacities.

5.5 Case study

This section evaluates system performances given different reserve policies (single, seasonal and dynamic zones). The proposed EDUC model is tested on a modified IEEE RTS-73 bus 3-area test system used in [186,194]. The system has 73 nodes, 96 conventional units, 9 wind farms, 117 lines, and 51 loads. Wind farms with 3900 MW of capacity are placed at nodes $2, 14, 16, 20, 23, 26, 37, 44, 47$. Modifications to each of the three identical areas of the test system follow [194,209]: line $(11-13)$ is removed; 480 MW of load is shifted from nodes $14, 15, 19$ and 20 to node 13; and the capacity of line $(14-16)$ is decreased to 350 MW. A study on the system performances is conducted at each day of the peak load week (week 51). A parameter ε of 0.5 is chosen in the determination of line weights. The hourly day-ahead load forecasting error at each bus is standardised as the proportion of the central forecast, and the standardised load forecasting error is modelled as $N(0, 0.05^2)$. Given the probability distribution functions of uncertainty in load and wind generation (as in (5.8)) at each bus as well as the probability mass function of available conventional generating capacity (described by the COPT), the probability mass function of the aggregated system imbalances at each zone, $\phi_{z_h}(\cdot)$, can be calculated using convolution under the independence assumption.

5.5.1 Market setup

In the day-ahead market of energy and operating reserves, reserve zones are firstly determined using the spectral clustering approach as described in Section 5.4.1, the EDUC is solved to produce the schedules, the cost information as well as the energy and reserve prices (i.e., LMPs and RMCPs) calculated on an hourly basis using (5.4) and (5.5). Afterwards, the operations are simplified to a 24-hour security-constrained unit commitment (by removing all the relevant objective terms and constraints to operating reserves in (5.1)) that produces the dispatch of generators and the additional commitments of fast generators in the real-time market. A total number of 500 post-contingency scenarios (i.e., hourly load, wind generation, and generator outages) are simulated on each day, representing the stochastic op-

erational conditions in the real-time market. A pricing model is then run by fixing the commitment variables of fast generators to derive the real-time LMPs.

In order to investigate the impact of some reserve policies on system performances, comparisons are made between the case of traditional single-zone deterministic UC without incorporating ORDCs (C-T), of EDUC with seasonal zones (C-S), of EDUC with dynamic (hourly) zones determined by the spectral clustering approach (C-DS), and of EDUC with dynamic (hourly) zones determined by the k-means clustering (C-DK) approach that has been used in [185, 186]. In case C-S, seasonal zones are determined using the k-means clustering approach applied to weights that are based on average line utilizations over a season, as in [210].

5.5.2 Comparison of reserve zone partition between spectral clustering and k-means clustering

Different reserve zone partitioning methods may result in different clustering results. First, seasonal zones are determined by k-means clustering, in which the weighted PTDF is used as the centrality measurement and the weights are based on average line utilizations over a season as defined in [185]. Second, the k-means clustering based method proposed in [185] is used to determine hourly zones, where the number of clusters is chosen as the number of congested zonal links. Third, the spectral clustering based method described in Section 5.4.1 is carried out and is compared with k-means clustering.

Fig. 5.1 shows the three seasonal reserve zones. Figures 5.2–5.3 show different reserve zone partitioning results by spectral clustering and k-means clustering at selected two hours in peak load week. At hour 7 on day 351, the wind penetration is as high as 56.1%, the 3 reserve zones identified by spectral clustering and k-means clustering are slightly different, as shown by Figs. 5.2, 5.3.

The advantage of our proposed reserve zone partitioning method is not clearly seen from these clustering results, but will take effect in system performances produced by the EDUC model, as will be discussed in the following subsection.

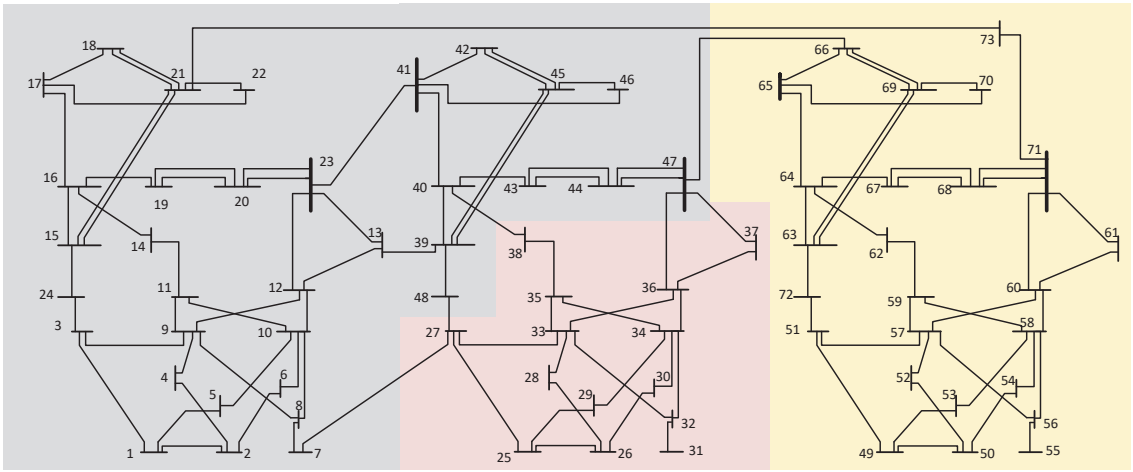

Figure 5.1: Three seasonal reserve zones (in colored shades).

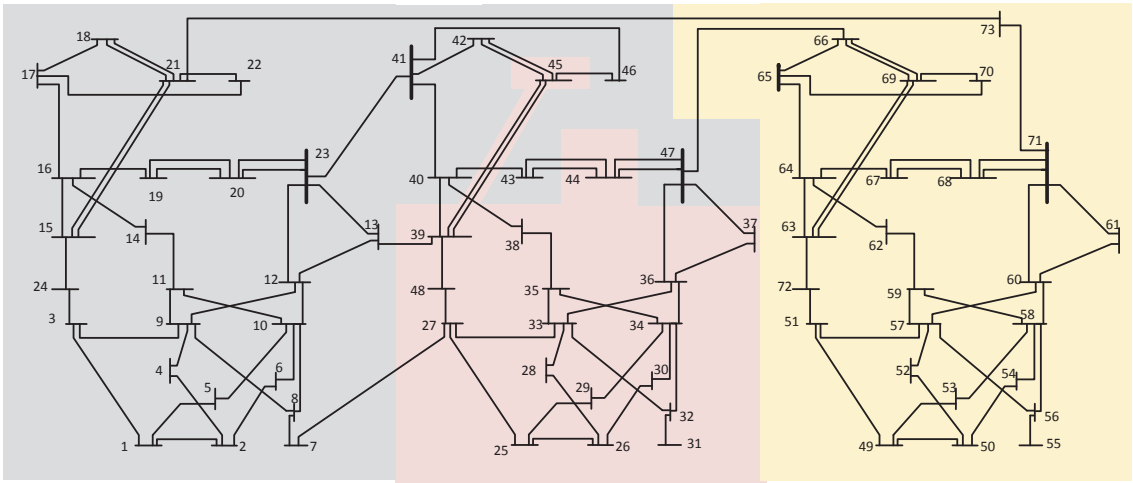

Figure 5.2: Three reserve zones (in colored shades) identified by spectral clustering at hour 7 *on day* 351.

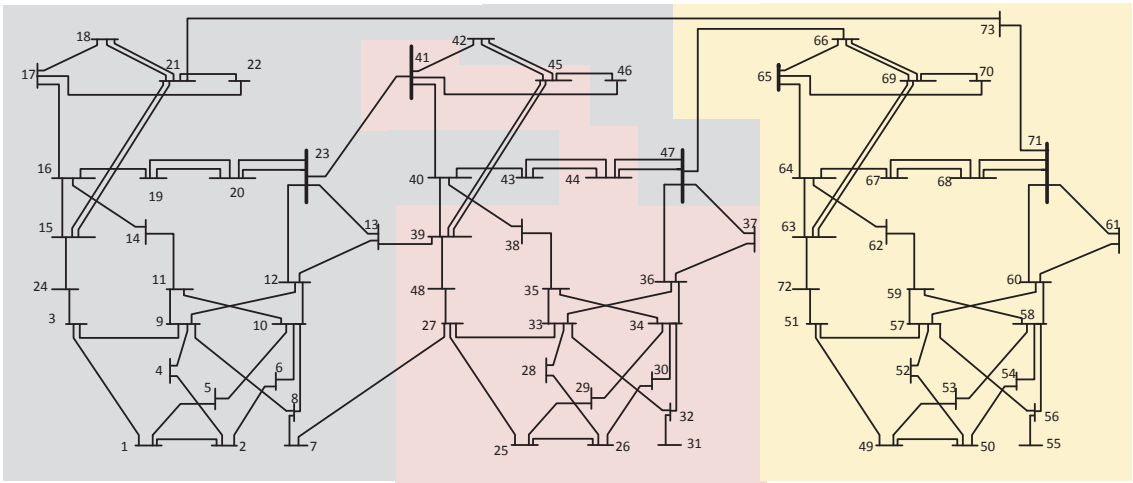

Figure 5.3: Three reserve zones (in colored shades) identified by k-means clustering at hour 7 *on day* 351.

5.5.3 Impact of reserve policies on system cost and reliability

Simulations are run on each day of the peak load week in order to study the impact of reserve policies (e.g., single, seasonal or dynamic reserve zones) on the operational cost of conventional generators, and on system reliability indicated by the amount of load shedding.

Table 5.1 presents averaged daily results of operational costs and load shedding over 500 scenarios obtained in the four cases (C-T, C-S, C-DS and C-DK). The operational cost of conventional generators indicates system economics while the information of load shedding is used as an indicator of system reliability.

In the peak load week, the operational costs of the traditional UC and the model proposed in our work (cases C-T and C-DS) are the lowest among the four cases, followed by the operational cost achieved in case C-DK which is on average lower than that in case C-S. The operational cost here accounts for the start-up costs of conventional generators committed in the day-ahead market and those of fast generators committed in the real-time market, as well as the variable and fixed operational costs of conventional generators in the real-time markets. In terms of the daily operational cost, the proposed reserve zone partitioning method based on spectral clustering is more favorable than the seasonal zones and the dynamic zones based on k-means clustering as used in [185, 186].

In the column of load shedding in each case, the first number gives the expected amount of load shedding, the second number in the bracket the largest amount of load shedding, and the index number shows the number of contingencies where the event of load shedding happens. Due to the introduction of reserve zones and the incorporation of Z-ORDCs, both the expected and the largest amount of load shedding have been significantly reduced in cases using the EDUC model (i.e., C-S, C-DS, C-DK), compared with case C-T using the traditional UC model. In the traditional UC, the low operational cost is achieved on the sacrifice of system reliability. On the three cases C-S, C-DS, C-DK, there seems no direct answer to the question of which reserve policy adopted is the best in terms of the resultant

load shedding, because the reliability indicator depends on the day and the wind penetration. However, on average, the results show that in cases C-DS, C-DK where a dynamic reserve policy is used, the expected and the largest amount of load shedding are among the lowest.

A comparison of operational cost and load shedding between the four cases C-T, C-S, C-DS, and C-DK shows that case C-DS achieves an overall better operational performance in terms of its on average lower operational cost as well as higher system reliability.

Table 5.1: Averaged daily results of operational cost and load shedding over 500 *scenarios*

Day	WP [a]	Operational Cost				Load Shedding			
		C-T	C-S	C-DS	C-DK	C-T	C-S	C-DS	C-DK
	%	M$	M$	M$	M$	MWh	MWh	MWh	MWh
351	42.0	1.520	1.521	1.514	1.529	$4.3(476)^{8}$	$2.1(296)^{12}$	$0.9(173)^{10}$	$1.1(129)^{9}$
352	18.5	2.779	2.789	2.775	2.790	$2.1(233)^{30}$	$0.5(44)^{19}$	$0.4(43)^{16}$	$0.4(43)^{26}$
353	11.1	3.058	3.055	3.051	3.056	$0.8(394)^{2}$	$0.6(292)^{2}$	$0.2(84)^{2}$	$0.3(86)^{3}$
354	27.2	2.335	2.448	2.333	2.354	$0.5(229)^{3}$	$0.1(50)^{1}$	$0.0(0.4)^{1}$	$0.1(49)^{1}$
355	22.1	2.402	2.401	2.401	2.402	$1.1(211)^{9}$	$0.1(19)^{4}$	$0.0(0.7)^{2}$	$0.0(0.7)^{2}$
356	7.7	2.222	2.234	2.217	2.242	$0.0(0)^{0}$	$0.0(0)^{0}$	$0.0(0)^{0}$	$0.0(0)^{0}$
357	8.7	2.109	2.111	2.096	2.114	$0.0(0)^{0}$	$0.0(0)^{0}$	$0.0(0)^{0}$	$0.0(0)^{0}$
Ave	20.9	2.346	2.366	2.341	2.355	1.3(220)	0.5(100)	0.2(43)	0.3(44)

[a] WP is the abbreviation of wind penetration that is defined as the ratio of forecasted daily wind energy to forecasted daily load.

5.5.4 Market implications of reserve policies

After running the UC models and the pricing models in the day-ahead market and the real-time market, day-ahead LMPs and RMCPs as well as contingency-specific real-time LMPs are determined. A two-settlement policy as described in [190, 193, 211] is used to calculate energy and reserve revenues, load payments and uplift payments. After the generators' offers are cleared in the day-ahead market, the extra amount of energy cleared in the real-time market is charged from load consumers and paid to suppliers at the real-time LMP. Vice versa, the reduced amount of

energy cleared in the real-time market is charged from suppliers and paid back to load consumers at the real-time LMP. The uplift component is paid if at the end of a day, a generator's start-up and no-load costs are not covered by any energy revenue in addition with reserve revenue it receives. The operating reserve providers are paid at the RMCP if cleared in the day-ahead market, but may not be dispatched in the real-time market. Contingency-specific disqualification can be made to operating reserves that are not deliverable in the real-time market as described in [186].

Table 5.2: Averaged daily results of energy and reserve revenues over 500 *scenarios (millions $)*

Day	WP	Energy Revenue				Reserve Revenue			
		C-T	C-S	C-DS	C-DK	C-T	C-S	C-DS	C-DK
351	42.0	2.252	2.287	2.396	2.322	0.081	0.199	0.235	0.221
352	18.5	3.816	4.427	4.244	4.415	0.119	0.234	0.217	0.223
353	11.1	5.133	6.022	5.661	6.156	0.094	0.361	0.261	0.376
354	27.2	3.154	3.369	3.224	3.286	0.090	0.143	0.128	0.142
355	22.1	3.385	4.067	3.606	3.886	0.103	0.275	0.236	0.247
356	7.7	3.246	3.330	3.252	3.336	0.056	0.261	0.234	0.258
357	8.7	2.774	3.219	3.084	3.124	0.095	0.254	0.171	0.217
Ave	20.9	3.394	3.817	3.638	3.789	0.091	0.247	0.212	0.241

Table 5.2 compares the impacts of different reserve policies (i.e., single, seasonal and dynamic reserve zones) and UC models on averaged daily results of conventional generators' energy and reserve revenues. The three cases using the EDUC model incorporating Z-ORDCs produce higher energy and reserve revenues than the case using a traditional deterministic UC model. However, there is no general conclusion on whether a dynamic reserve policy or a certain reserve partitioning method increases or decreases wholesale prices. The relative outcomes of energy and reserve revenues among the three cases where an EDUC is adopted depend on the day and the level of wind penetration. On the peak load week, the case adopting our proposed reserve partitioning method (i.e., C-DS) produces on average lower energy and reserve reserves than the cases adopting seasonal zones or dynamic zones partitioned by k-means clustering.

In the EDUC model, the demand curve of operating reserves is represented by piecewise linear cost functions that approximate the value of expected unserved energy or the value of expected curtailed wind corresponding to the level of operating reserves. These cost functions create a day-ahead price adder to wholesale prices (including energy and reserve prices) in the day-ahead energy market, aiming to improve the scarcity price signal. More specifically, wholesale prices will increase automatically as available operating reserves decrease, which happens when electricity demand becomes extremely high or for other reasons (e.g., generator outages), there is not enough generation to maintain needed operating reserves. A distinct feature of ORDCs is that even when the reserve requirements are not violated, there may be a price paid to available operating reserves. By contrast, in the traditional deterministic UC model, the reserve demand curve is a step function and the reserve price of a certain type equals to the dual variables of the corresponding system-wide reserve requirement; this price does not reflect the marginal value of operating reserves in avoiding events of load shedding and wind curtailment. Energy prices are coupled with reserve prices in UC models with a co-optimization of energy and operating reserve. In general, the inclusion of the Z-ORDCs increased the LMPs at nearly all buses for the peak load week considered.

Table 5.3: Averaged daily results of load and uplift payments over 500 *scenarios (millions $)*

Day	WP	Load Payment				Uplift Payment			
		C-T	C-S	C-DS	C-DK	C-T	C-S	C-DS	C-DK
351	42.0	3.628	3.716	3.965	3.758	0.056	0.015	0.009	0.016
352	18.5	4.624	5.405	5.255	5.375	0.278	0.176	0.087	0.110
353	11.1	4.632	5.687	5.548	5.813	0.144	0.017	0.012	0.023
354	27.2	4.308	4.573	4.365	4.464	0.093	0.088	0.044	0.065
355	22.1	4.254	4.817	4.264	4.641	0.091	0.020	0.018	0.018
356	7.7	3.533	3.601	3.515	3.605	0.279	0.002	0.001	0.008
357	8.7	3.223	3.503	3.516	3.508	0.319	0.018	0.015	0.018
Ave	20.9	4.029	4.459	4.347	4.452	0.165	0.048	0.027	0.037

Table 5.3 compares the load payment and uplift payment in the four cases. Generally, the load payment is increased and the uplift payment for conventional generators is reduced when Z-ORDCs are implemented in system operations, because of the increased average energy and reserve prices. Among the three cases where the EDUC model is used in the day-ahead market operations, the one with dynamic reserve zones partitioned by spectral clustering as proposed in our work produces on average lower load payment and uplift payment, which is preferable from the perspective of market efficiency.

5.6 Chapter summary

Flexibility of conventional generators provided by standby operating reserves is valuable in balancing random disturbances, particularly in a transmission network integrated with a considerable amount of uncertain wind power. The zonal operating reserve demand curves are incorporated in the EDUC model to reflect the forecasted values in reducing the cost of expected unserved energy and curtailed wind of operating reserve. The new reserve pricing scheme for the EDUC model is proposed; this settlement scheme includes a price component for zonal reserves introduced by operating reserve demand curves. A spectral clustering approach is employed to determine the optimal number of zones and to define reserve zones at each scheduling hour in the day-ahead market.

Simulations on the IEEE RTS-73 test case show introducing dynamic reserve zones improves the deliverability of operating reserves compared with using the single zone and seasonal zones in terms of the improved system reliability. Implementing appropriate Z-ORDCs in market operations incentivises the availability of operating reserves in terms of increased reserve revenues, which contributes to the flexible and reliable system operations, especially in power systems with large amounts of VG. It is also shown that the spectral clustering approach proposed in our study works well in delivering reserve zones which contribute to improving system reliability and operational efficiency.

Chapter 6

Discussion and Conclusions

In this chapter, we briefly summarize main research results regarding the effects of different market designs or incentives on long-term generation investments and short-term power system operations. To achieve these results, our work has been divided into two parts – the application of Bayesian methodology to a LTGI model and the introduction of zonal operating reserve demand curves to a deterministic UC model. We argue that much can be gained by combining the two models and improving the Bayesian methodology for addressing the resultant computational challenge. We will discuss future directions of the methodological research arising from current power system challenges and the computer models developed for better understanding of these issues and guiding of electricity market designs.

6.1 General discussion of results

6.1.1 Bayesian application to the LTGI problem

In the first part of the thesis (Chapter 2-4), we provided a review of evolving electricity markets and LTGI models, and emphasized the key role that different sources of uncertainty play in LTGI models that are used as a decision-support tool for real-world analysis. We then presented a Bayesian framework for evaluating the impacts of two mainstream wholesale market designs (e.g., energy-only market, capacity market) on generation investments and system adequacy based on a LTGI model

developed in [4, 13]. In particular, we showed for the first time how Bayesian emulation enables very efficient model calibration, uncertainty analysis and sensitivity analysis of a complex LTGI model while systematically managing various sources of uncertainty in the model.

LTGI models, incorporating very complex formulations of the electricity price formation and the investment decision-making logic (and even more complex models that include detailed operational modelling) have been widely used for making generation projections and designing energy policy to maintain system reliability. A great concern of model users (e.g., investors and policymakers) is how well the model represents the real-world and how much credibility can be assigned to the model outputs. A second concern is that running such models of decision-making under uncertainty is often very computationally demanding, yet hugely important for energy policy-making. Particularly, a large number of model evaluations are needed for uncertainty analysis and sensitivity analysis in order to give robust answer to questions about the impact of energy policies on future power system investments and reliability. The Bayesian method allows these two concerns to be addressed.

In the Bayesian approach, an emulator (a Gaussian Process model) is fitted to a limited number of model evaluations at a sparse coverage of input space that are possible to obtain. To validate and assess the adequacy of a Gaussian process emulator, some diagnostics are carried out by comparing simulator outputs and emulator outputs at some test points (a new set of design points used for validation purposes). The validated emulator represents uncertainty in the relationship between a set of uncertain inputs and the outputs of interest of the LTGI model. More specifically, the emulator quantifies the uncertainty in the output where the simulator is not evaluated. In principle, the Bayesian method is capable of addressing all sources of uncertainty that propagate into uncertainty in the model output. Three major sources of uncertainty, namely, input uncertainty (unknown precise values for model inputs), structural uncertainty (imperfect science in the model equations, and errors in model structure or logic) and functional uncertainty (unknown model outputs at inputs where the model has not been run) have been accounted for in our work.

All of the following were accomplished in case studies on the representative Great

Britain exemplar:

(1) In model calibration, the posterior distributions of calibration parameters are derived from the product of their prior beliefs and the likelihood of historical observations given the calibration parameters. Meanwhile, the model discrepancy is inferred from observations and is then applied to future projections.

(2) In making predictions, probabilistic predictions of model outputs (e.g., generation projections and LOLEs) are made by the calibrated and discrepancy-adjusted emulator while accounting for input uncertainty, structural uncertainty and functional uncertainty, as opposed to a point estimate (or a deterministic scenario) of model outputs produced by the simulator.

(3) Through uncertainty analysis of the maximum LOLE over a future planning horizon, the robustness of two market designs (i.e., an energy-only market design and a capacity market design) are compared given different scenarios of market design parameters. The robustness is indicated by the probability of the maximum LOLE meeting a reliability standard (e.g., no larger than 3 hours per year). Simulation results show that a capacity market design generally is more robust to uncertainties than an energy-only market design.

(4) Through variance-based sensitivity analysis, under both an energy-only market and a capacity market, the energy price cap is identified to be the most important model input that affects investment decisions and system reliability in terms of how much of the output variance is attributed to each contributory source of uncertainty.

6.1.2 Zonal operating reserve demand curves incorporated in the enhanced deterministic unit commitment model

In the second part, we developed an enhanced deterministic UC model incorporating zonal operating reserve demand curves and compared system performances resulted from different choices of reserve policies (e.g., single, seasonal or dynamic zones) and

of reserve zone partitioning methods. The enhanced deterministic UC model is tailored for a transmission network integrated with a considerable amount of uncertain wind power, and is aiming to produce better signals of energy and operating reserve prices as incentives for flexible resources.

An improved dynamic zone partitioning method based on spectral clustering is employed to determine the optimal number of zones and to define reserve zones at each hour, depending on the post-contingency line flows. The clusters of reserve zones produced by spectral clustering are compared with that derived from k-means clustering in the case study of the IEEE-73 bus system. In general, the proposed determination of reserve zones is flexible in varying the number and the geographic feature of zones at each hour according to the forecasted day-ahead operating conditions. Once the reserve zones are identified, the zonal operating reserve demand curves can be developed as piece-wise linear approximations to the zonal value of expected unserved energy and/or expected curtailed wind that is a function of upward and/or downward operating reserves.

In the IEEE-73 case study, the simulation results on the peak load week are compared between four cases, namely, the case of traditional single-zone deterministic UC without incorporating ORDCs, the case of enhanced deterministic UC with seasonal zones, the case of enhanced deterministic UC with dynamic (hourly) zones determined by the spectral clustering approach proposed in our work, and the case of enhanced deterministic UC with dynamic (hourly) zones determined by the k-means clustering approach proposed in [185].

The indicators for operational efficiency include the amount of load shedding and the operational cost. The case using our proposed enhanced deterministic UC model together with the reserve zone partitioning method has an improved operational efficiency in terms of its on average lower operational cost and higher system reliability than the other three cases. The simulation results on the operational efficiency imply that dynamic and zonal operating reserve demand curves enable more efficient utilization of upward and downward operating reserves than the methods of single zone, seasonal zones and dynamic zones produced by k-means clustering.

In terms of the market implications of reserve polices, the energy revenue, reserve

revenue and uplift payment for conventional generators, as well as load payment are compared among the four cases. Generally, the introduction of zonal operating reserve demand curves into the enhanced deterministic UC model improves scarcity pricing of energy and locational operating reserves since the produced reserve market clearing prices reflect the risk cost of the system. Therefore, there is an increase in the energy and reserve revenue and a decrease in the uplift payment for conventional generators in cases where the enhanced deterministic UC model is employed. The decline in the uplift payment implies less distorted market signals, which is desirable in terms of market efficiency. On average, the enhanced deterministic UC model produces higher load payment but provides improved reliability services for consumers.

6.2 Future directions of the methodological research

6.2.1 Extensions and challenges of Bayesian emulation

Long-term generation investment models of multiple outputs

In the Bayeisan application to the LTGI problem, it is assumed that the simulator produces three different outputs (e.g., time series of installed thermal capacity, time series of LOLEs and the maximum LOLE), or that we are only interested in three outputs. Covariances between these outputs are ignored, and so the outputs are emulated individually using independent or uncorrelated emulators. The independence assumption or transformation is one of the most common simplification techniques to cope with multiple outputs of the simulator. Under the independence assumption, each output will be predicted separately, reducing the multi-output problem to several individual single-output problems.

However, joint uncertainty, represented by covariances as well as variances, in several outputs, sometimes, is of importance for real-world analysis. Improper independence assumption could lead to significant over-estimation of output uncertainty in situations of strong correlation between outputs. For instance, the outputs of in-

terest in the LTGI model are the installed capacity of each generation technology and there are correlations between investments in different technologies. Or, when we are interested in using a combination of the expected unserved energy, the loss-of-load expectation and consumer costs as the reliability metric, it is desirable to develop an emulator of multiple outputs and to account for their joint uncertainty. For questions involving correlated multiple outputs, a multivariate emulator based on a multivariate GP is an alternative [144]; this is one of the interesting research directions.

Heterogeneous behaviour of long-term generation investment models

In our existing work on the long-term generation investment problem, the adopted simulator produces relatively smooth outputs, because there assumes no sudden change of long-term fuel prices or a political intervention that are enforced as a model input. When these sudden changes or interventions are taken into account in the model specification, the simulators may have discontinuous response such as sudden shifts, or have sharp changes of gradient [133].

The Bayesian emulation approach is widely applicable to various types of simulators where the model output is deterministic [133], stochastic [146], dynamic [212] or multi-variate [144]. However, the basic GP emulator as used in our work will very likely smooth out these discontinuities or sharpness, resulting in a locally poor representation of the simulator. This is because the basic GP emulator assumes that the response of simulator outputs to its inputs is smooth. Besides, it is a prior expectation that the simulator does not respond much more significantly to changes in an input over some parts of the input space than over others. In order to adapt to the non-smooth or unstable response, GPs with less smooth covariance structures may be used and they need large numbers of well-designed training data in order to be developed. Alternatively, new kinds of emulator may be explored to allow for heterogeneity including discontinuities and heterogeneous variance.

Dynamic LTGI models

For making long-term decisions on generation investments and energy policy, a common model type is a dynamic LTGI simulator that produces a time series of the output of interest. A practical challenge in dealing with a full dynamic model lies in the high dimensionality of the input space. In the GB case study presented in Chapter 4, we have assumed that the input variables are constants over the long-term planning horizon in order to limit the dimensionality of the input space, or number of uncertainties considered. However, the values of control or forcing inputs at each year or at each stage (e.g., every five years) may vary independently and hence they require independent variables to represent which results in a higher dimensionality of the input space. For instance, the energy price cap might change with the increasing integration of variable generation and the evolving market designs during the future planning period. The long-term trend of forcing gas prices may be vulnerable to a sudden change that may aries from the production of shale gas in the international market.

To deal with the input dimensionality in a dynamic model, we may emulate the *full simulator* based on emulating the *single-step function* by recursively updating a state vector. The full simulator can be represented as $\{y_1, \ldots, y_T\} = f(y_0, u_1, \ldots, u_T, \omega_1, \ldots, \omega_T, \theta)$, where the inputs are initial conditions y_0, a time series of external control and forcing inputs $\{u_1, \ldots, u_T, \omega_1, \ldots, \omega_T\}$ and constant parameters θ; and the outputs are a time series $\{y_1, \ldots, y_T\}$. By contrast, the single-step function inputs is expressed as $y_t = f(y_{t-1}, u_t, \omega_t, \theta)$, where the inputs are the current value of the state variable (i.e., the value of the full simulator output at the previous time point) y_{t-1}, the associated control and forcing input $\{u_t, \omega_t\}$ and constant parameters θ; and the output is the value of the state variable at the next time point y_t.

It can be seen that the dimensionality of the input space of interest is significantly reduced for emulating the single-input function, as opposed to emulating the full function. Moreover, a single-step emulator achieves more accurate numerical solutions than the full one [213] and is able to make predictions over arbitrary time steps [133].

The single-step function approach also offers time savings for dealing with very computationally expensive simulators in situations where emulation through the full emulator is infeasible, such as for long-term generation planning models with detailed UC programs [83, 98]. In Chapter 3, we have adopted a reasonably complex LTGI model to demonstrate the systematic method of calibration, uncertainty analysis and sensitivity analysis. A relatively efficient production costing model was built in the model and no chronological features in operational constraints were captured. Such a model provides limited implication for problems regarding system flexibility and its incentives. The single-step function approach opens an opportunity of utilizing a more comprehensive LTGI model. This research direction will be further discussed in Section 6.2.2.

High-dimensional LTGI models

For practical policy making, there are situations where the single-step function approach is not applicable to a LTGI model or where the model has tens or hundreds of input variables even after using the single-step function approach. With a high dimensional input space, the number of design points will be considerably large in order to cover the whole input space. Then, fitting a Bayesian emulator to a large set of design points will be computationally expensive because there is a large size of covariance matrix to be dealt with. Correspondingly, there will be a large number of hyperparameters in the emulator, and so the quality of the fitted emulator may not be satisfactory.

Efficient algorithms for dimensionality reduction of the input space are of paramount importance to allow effective probabilistic analysis. Screening, which is also called feature selection in the machine learning area, can be applied prior to calibration and uncertainty analysis as part of the emulator construction. Some screening methods [214], such as screening design methods, ranking methods and wrapper methods can be used to identify variables with negligible total effects on the output variables. Screening enables efficient experimental design (fewer design points filling up a lower dimensional input space) and improves interpretability (less quantities represented by the remaining variables to be estimated or measured in the future).

6.2.2 Studying the impacts of spot market designs on investment incentives

Chapter 5 has studied how reserve policies in a spot market (day-ahead or real-time market) affect the price signals of energy and operating reserves, and hence influence system and market efficiency at the operational level. In a planning framework, the long-term effects of reserve policies on investment decisions or system reliability remain to be explored. More detailed operational or pricing modelling (e.g., ramping, commitment, storage, correlations of renewables and load, advanced pricing rules) can be included in the LTGI model, and will be associated with higher dimensional uncertainty (i.e., input variables) and operational constraints. This would permit study of research questions such as the effectiveness of some new market designs (e.g., green electricity certificates, CO_2 certificates [215]), as well as the impacts on system reliability and flexibility of the use of energy storages or the demand side management in the decentralized electricity generation.

An interesting direction of our future work is to employ the single-step function approach of Bayesian emulation in the framework of approximate dynamic programming (ADP) proposed in [111] to address very complex multi-period optimization (identifying the values of one or more inputs that will minimise or maximise the output) or decision analysis (finding an optimal decision according to a formal description of decision criteria). A relevant question is to identify the optimal scenario of energy policy that delivers the target reliability for consumers with a certain credibility level at the lowest level of cost over the long term. Analyses of energy policies in electricity markets as a sequential decision under uncertainty have been severely restricted by dimensionality and computational burdens. Some types of uncertainty are affected by decisions in energy policies or investments. For instance, the investment decision made by one investor today will change the probability distribution of next period's decisions of rival investors if more than one investors are considered. These rival decisions depend on (largely exogenous) demand levels, fuel and carbon prices and the (endogenous) plant mix. Hence, there is a need for capturing endogenous uncertainties in a flexible way rather than using the typical

modelling techniques such as exogenous scenario trees or probability distributions. multistage, multi-dimensional stochastic dynamic programming is an natural choice for modelling the dynamics of power systems as well as endogenous uncertainties, but it suffers from the so-called curse of dimensionality [216]. In this situation, the ADP framework can be adopted as a tractable alternative solution to the large-scale decision-making problems that take place over multiple time periods under uncertainty.

The central challenge of ADP lies in fitting the value function approximation for making decisions. The value function captures the information about the future state of the system while making the decisions for the current time [107]. Lookup tables, parametric representations (e.g., a linear or nonlinear function), and nonparametric representations (e.g., neural networks) can all function as the value function. The emulator, as a parametric statistical model, can also be used to approximate the true value function that covers the entire state space. The emulation approach is appealing because of its high efficiency and flexibility in approximating linear and non-linear functions as well as in dealing with uncertainty, but care has to be put into the assignments of prior beliefs and the quality of the value function approximation.

6.3 Thesis conclusion

This thesis has for the first time demonstrated the application of Bayesian emulation to the calibration and interpretation of power system planning models. In addition, this thesis has developed an enhanced deterministic UC model for investigating the effects of reserve policies on the short-term operational and market efficiency. These adopted techniques include 1) uncertainty modelling in both the LTGI problem and the deterministic UC problem; 2) Bayesian emulation, calibration, uncertainty analysis and probabilistic sensitivity analysis when a LTGI model is employed for real-world applications; and 3) co-optimization of energy and operating reserve products in the day-ahead electricity market considering a multi-zone transmission network.

Our work has introduced methodologies of analyzing the role that market incen-

tives play on power system reliability and flexibility. Regarding the complex real-world power system planning and corresponding mathematical models, this thesis has provided a unified Bayesian framework which accounts for the major sources of uncertainty and quantifying uncertainty in model outputs in order to make it transparent to modellers and policymakers alike. The Bayesian framework has a radical effect on real-world power system modelling and model usage. A prediction from a Bayesian emulator is a plausible range of a model output with probability attached rather than a single deterministic scenario of the model output produced by a deterministic simulator. The emulator provides generally orders of magnitude higher computational efficiency than the simulator. Thereby, emulators enable formal calibration, uncertainty analysis and sensitivity analysis that require a large number of model evaluations at a dense coverage of the input space.

Regarding the day-ahead power system operational planning, this thesis has proposed a novel enhanced deterministic UC model that incorporates zonal operating reserve demand curves and co-optimizes the energy and the zonal operating reserves. The enhanced deterministic UC model is developed for more efficient utilization and remuneration of operating reserves - one of the major sources of system flexibility, which facilitates the integration and the consumption of variable generation in a transmission network. To explore the long-term effects of reserve policies on system investment and adequacy, a more comprehensive LTGI model with operational details and advanced reserve polices may be used in our future research, which poses challenges and hence requires extensions to the Bayesian emulation approach.

Bibliography

[1] "DECC fossil fuel price projections," 2013. [Online]. Available: http://www.gov.uk/government/uploads/system/uploads/attachment_data/file/212521/130718_decc-fossil-fuel-price-projections.pdf

[2] J. Cochran, M. Miller, M. Milligan, E. Ela, D. Arent, and A. Bloom, "Market evolution: Wholesale electricity market design for 21st century power systems," National Renewable Energy Laboratory, Tech. Rep., October 2013. [Online]. Available: http://www.nrel.gov/docs/fy14osti/57477.pdf

[3] D. M. N. Fabien Roques and W. J. Nuttall, "Generation adequacy and investment incentives in britain: from the pool to neta." University Of Cambridge Department of Applied Economics, 2004, cambridge Working Papers in Economics CWPE 0459.

[4] D. Eager, "Dynamic modelling of generation capacity investment in electricity markets with high wind penetration," Ph.D. dissertation, The University of Edinburgh, 2012. [Online]. Available: http://hdl.handle.net/1842/6264

[5] "Electricity market reform: Delivering UK investment," Department of Energy & Climate Change, Tech. Rep., June 2013. [Online]. Available: https://www.gov.uk/government/publications/electricity-market-reform-delivering-uk-investment

[6] L. Hancher, A. de Hauteclocque, and M. Sadowska, *Capacity Mechanisms in the EU Energy Market: Law, Policy, and Economics*. Oxford University Press, 2015.

[7] C. Smith, "Wind and solar energy: variable generation across the land," *IEEE power & energy*, vol. 13, no. 6, pp. 16–20, November/December 2015.

[8] C. Vázquez, M. Rivier, and I. J. Pérez-Arriaga, "A market approach to long-term security of supply," *IEEE Transactions on Power Systems*, vol. 17, no. 2, pp. 349–357, May 2002.

[9] A. Keane, M. Milligan, C. J. Dent, B. Hasche, C. D'Annunzio, K. Dragoon, H. Holttinen, N. Samaan, L. Soder, and M. O'Malley, "Capacity value of wind power," *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 564–572, May 2011.

[10] F. D. Munoz and A. D. Mills, "Endogenous assessment of the capacity value of solar pv in generation investment planning studies," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 4, pp. 1574–1585, Oct 2015.

[11] R. Gross, P. Heptonstall, and W. Blyth, "Investment in electricity generation–the role of costs, incentives, and risks," UK Energy Research Centre, Tech. Rep., 2007. [Online]. Available: http://www.ukerc.ac.uk/publications/investment-in-electricity-generation-the-role-of-costs-incentives-and-risks.html

[12] R. Gross, W. Blyth, and P. Heptonstall, "Risks, revenues and investment in electricity generation: Why policy needs to look beyond costs," *Energy Economics*, vol. 32, no. 4, pp. 796–804, 2010.

[13] D. Eager, B. Hobbs, and J. Bialek, "Dynamic modeling of thermal generation capacity investment: Application to markets with high wind penetration," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2127–2137, Nov 2012.

[14] M. Assili, M. H. D.B., and R. Ghazi, "An improved mechanism for capacity payment based on system dynamics modeling for investment planning in competitive electricity environment," *Energy Policy*, vol. 36, no. 10, pp. 3703–3713, 2008.

[15] S. Arango and E. Larsen, "Cycles in deregulated electricity markets: Empirical evidence from two decades," *Energy Policy*, vol. 39, no. 5, pp. 2457–2466, 2011.

[16] F. Olsina, F. Garcés, and H.-J. Haubrich, "Modeling long-term dynamics of electricity markets," *Energy Policy*, vol. 34, no. 12, pp. 1411–1433, 2006.

[17] L. de Vries and P. Heijnen, "The impact of electricity market design upon investment under uncertainty: The effectiveness of capacity mechanisms," *Utilities Policy*, vol. 16, no. 3, pp. 215–227, 2008.

[18] M. Hasani and S. H. Hosseini, "Dynamic assessment of capacity investment in electricity market considering complementary capacity mechanisms," *Energy*, vol. 36, no. 1, pp. 277–293, 2011.

[19] G. Doorman and A. Botterud, "Analysis of generation investment under different market designs," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 859–867, 2008.

[20] B. Hobbs, M.-C. Hu, J. Inon, S. Stoft, and M. Bhavaraju, "A dynamic analysis of a demand curve-based capacity market proposal: The PJM reliability pricing model," *IEEE Transactions on Power Systems*, vol. 22, no. 1, pp. 3–14, Feb 2007.

[21] I. Herrero, P. Rodilla, and C. Batlle, "Electricity market-clearing prices and investment incentives: The role of pricing rules," *Energy Economics*, vol. 47, pp. 42–51, 2015.

[22] A. Veiga, P. Rodilla, I. Herrero, and C. Batlle, "Intermittent RES-E, cycling and spot prices: The role of pricing rules," *Electric Power Systems Research*, vol. 121, pp. 134–144, 2015.

[23] "Electricity market reform: Capacity market–detailed design proposals," Department of Energy & Climate Change, Tech. Rep., 2013. [Online]. Available: http://www.gov.uk/government/uploads/system/uploads/attachment_data/file/209280/15398_TSO_Cm_8637_DECC_Electricity_Market_Reform_web_optimised.pdf

[24] W. Steggals, R. Gross, and P. Heptonstall, "Winds of change: How high wind penetrations will affect investment incentives in the GB electricity sector," *Energy Policy*, vol. 39, no. 3, pp. 1389–1396, 2011.

[25] L. Xie, P. Carvalho, L. Ferreira, J. Liu, B. Krogh, N. Popli, and M. Ilic, "Wind integration in power systems: Operational challenges and possible solutions," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 214–232, Jan 2011.

[26] S. jie Deng and S. S. Oren, "Incorporating operational characteristics and startup costs in option-based valuation of power generation capacity," in *Probability in the Engineering and Informational Sciences*, 2002, pp. 155–181.

[27] D. Auverlot, E. Beeker, G. Hossie, L. Oriol, and A. Rigard-Cerison, "the crisis of the european electricity system diagnosis and possible ways forward," 2014. [Online]. Available: www.strategie.gouv.fr/sites/strategie.gouv.fr/files/archives/CGSP_Report_European_Electricity_System_030220141.pdf

[28] J. Carstairs and I. Pope, "The case for a new capacity mechanism in the UK electricity market–lessons from Australia and New Zealand," *Energy Policy*, vol. 39, no. 9, pp. 5096–5098, 2011.

[29] J. Riesz, J. Gilmore, and I. MacGill, "Assessing the viability of energy-only markets with 100 renewables: an Australian National Electricity Market case study." *Econ. Energy Environ. Policy*, vol. 5, no. 1, pp. 105–130, 2016.

[30] R. Baldick, U. Helman, B. Hobbs, and R. O'Neill, "Design of efficient generation markets," *Proceedings of the IEEE*, vol. 93, no. 11, pp. 1998–2012, Nov 2005.

[31] K. Sakellaris, "The greek capacity adequacy mechanism: Design, incentives, strategic behavior and regulatory remedies," in *Energy Market, 2009. EEM 2009. 6th International Conference on the European*, 2009, pp. 1–6.

[32] O. Ozdemir, J. de Joode, P. Koutstaal, and M. van Hout, "Financing investment in new electricity generation capacity: The impact of a german capacity

market on northwest europe," in *European Energy Market (EEM), 2013 10th International Conference on the*, May 2013, pp. 1–8.

[33] P. L. Joskow, "Capacity payments in imperfect electricity markets: Need and design," *Utilities Policy*, vol. 16, no. 3, pp. 159–170, 2008.

[34] G. Muratore, "Incentive based energy market design," *European Journal of Operational Research*, vol. 213, no. 2, pp. 422–429, 2011.

[35] P. Cramton, A. Ockenfels, and S. Stoft, "Capacity market fundamentals," *Economics of Energy & Environmental Policy*, vol. 2, no. 2, pp. 27–46, 2013.

[36] A. Botterud and G. Doorman, "Generation investment and capacity adequacy in electricity markets," *International Association for Energy Economics*, 2008, newletter Article. [Online]. Available: http://www.iaee.org/documents/newsletterarticles/208Botterud.pdf

[37] S. S. M. Venkata and N. Hatziargyriou, "Grid resilience: Elasticity is needed when facing catastrophes," *IEEE Power and Energy Magazine*, vol. 13, no. 3, pp. 16–23, May 2015.

[38] W. W. Hogan, "Market-clearing electricity prices and energy uplift," in *5th Conference on The Economics of Energy Markets, Toulouse, France*, 2008, pp. 4–7.

[39] W. Hogan, "Electricity scarcity pricing through operating reserves," *Economics of Energy & Environmental Polic*, vol. 2, no. 2, pp. 65–86, 2013.

[40] B. F. Hobbs, J. I. nón, and S. E. Stoft, "Installed capacity requirements and price caps: Oil on the water, or fuel on the fire?" *The Electricity Journal*, vol. 14, no. 6, pp. 23–34, 2001.

[41] E. Ela, M. Milligan, A. Bloom, A. Botterud, A. Townsend, and T. Levin, "Evolution of wholesale electricity market design with increasing levels of renewable generation," National Renewable Energy Laboratory, Tech. Rep., September 2014. [Online]. Available: www.nrel.gov/docs/fy14osti/61765.pdf

[42] "Flexible generation: Backing up renewables," Electric Renewables Action Plan, Tech. Rep., October 2011. [Online]. Available: Online: www.eurelectric.org/media/61388/flexibility_report_final-2011-102-0003-01-e.pdf

[43] "The power of transformation - wind, sun and the economics of flexible power systems," International Energy Agency, Tech. Rep., 2014. [Online]. Available: https://www.iea.org/publications/freepublications/publication/The_power_of_Transformation.pdf

[44] "Flexible capacity procurement phase 1: Risk of retirement," California ISO, Tech. Rep., September 2012. [Online]. Available: http://www.caiso.com/Documents/SecondRevisedDraftFinalProposal-FlexibleCapacityProcurement.pdf

[45] I. KEMA, "How to manage future grid dynamics: Quantifying smart power generation benefits," Tech. Rep., October 2012. [Online]. Available: http://www.smartpowergeneration.com/content-center/white-papers/how-to-manage-future-grid-dynamics

[46] "Power generation investment in electricity markets," International Energy Agency, Tech. Rep., 2003. [Online]. Available: http://www.hks.harvard.edu/hepg/Papers/Fraser.gen.invest.elec.mkts.1203.pdf

[47] M. Hasani-Marzooni and S. H. Hosseini, "Dynamic model for market-based capacity investment decision considering stochastic characteristic of wind power," *Renewable Energy*, vol. 36, no. 8, pp. 2205–2219, 2011.

[48] F. Castro-Rodriguez, P. L. Marín, and G. Siotis, "Capacity choices in liberalised electricity markets," *Energy Policy*, vol. 37, no. 7, pp. 2574–2581, 2009.

[49] A. Botterud, M. Ilic, and I. Wangensteen, "Optimal investments in power generation under centralized and decentralized decision making," *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 254–263, 2005.

[50] "China's power sector reforms - where to next?" International Energy Agency, Tech. Rep., 2006. [Online]. Available: https://www.iea.org/publications/freepublications/publication/chinapower.pdf

[51] "The value of lost load (VoLL) for electricity in Great Britain," London Economics, Tech. Rep., 2013. [Online]. Available: http://www.gov.uk/government/uploads/system/uploads/attachment_data/file/224028/value_lost_load_electricty_gb.pdf

[52] E. Hausman, R. Fagan, D. White, K. Takahashi, and A. Napoleon, "LMP electricity markets: Market operations, market power and value for consumers," Synapse Energy Economics, Tech. Rep., February 2006. [Online]. Available: https://www.publicpower.org/files/PDFs/SynapseLMPElectricityMarkets013107.pdf

[53] K. Neuhoff and L. D. Vries, "Insufficient incentives for investment in electricity generations," *Utilities Policy*, vol. 12, no. 4, pp. 253–267, 2004, infrastructure Regulation and Investment for the Long-Term.

[54] T. Levin and A. Botterud, "Capacity adequacy and revenue sufficiency in electricity markets with wind power," *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1644–1653, May 2015.

[55] S. Stoft, "The demand for operating reserves: key to price spikes and investment," *IEEE Transactions on Power Systems*, vol. 18, no. 2, pp. 470–477, May 2003.

[56] M. Keay, J. Rhys, and D. Robinson, "Chapter 2 - electricity market reform in britain: Central planning versus free markets?" in *Evolution of Global Electricity Markets*. Boston: Academic Press, 2013, pp. 31–57.

[57] M. Cepeda and D. Finon, "Generation capacity adequacy in interdependent electricity markets," *Energy Policy*, vol. 39, no. 6, pp. 3128–3143, 2011.

[58] "DECC dynamic dispatch model (DDM)," Department of Energy & Climate Change, Tech. Rep., May 2012. [Online]. Available:

http://www.gov.uk/government/uploads/system/uploads/attachment_data/file/65709/5425-decc-dynamic-dispatch-model-ddm.pdf

[59] E. Ibanez and M. Milligan, "Comparing resource adequacy metrics and their influence on capacity value," in *Probabilistic Methods Applied to Power Systems (PMAPS), 2014 International Conference on*, July 2014, pp. 1–6.

[60] J. P. Pfeifenberger, K. Spees, K. Carden, and N. Wintermantel, "Resource adequacy requirements: Reliability and economic implications," The Brattle Group and Astrape Consulting, Tech. Rep., September 2013. [Online]. Available: http://www.ferc.gov/legal/staff-reports/2014/02-07-14-consultant-report.pdf

[61] "GB electricity capacity margin," Royal Academy of Engineering, Tech. Rep., 2013. [Online]. Available: www.raeng.org.uk/publications/reports/gb-electricity-capacity-margin

[62] "Annex C: Reliability standard methodology," Department of Energy & Climate Change, Tech. Rep., 2013. [Online]. Available: www.gov.uk/government/uploads/system/uploads/attachment_data/file/267613/Annex_C_-_reliability_standard_methodology.pdf

[63] C. J. Dent, A. Hernandez-Ortiz, S. R. Blake, D. Miller, and D. Roberts, "Defining and evaluating the capacity value of distributed generation," *IEEE Transactions on Power Systems*, vol. 30, no. 5, pp. 2329–2337, Sept 2015.

[64] "The resilience of the electricity system," House Of Lords, Tech. Rep., 2015. [Online]. Available: www.publications.parliament.uk/pa/ld201415/ldselect/ldsctech/121/121.pdf

[65] C. Dent, K. R. W. Bell, A. W. Richards, S. Zachary, D. Eager, G. Harrison, and J. Bialek, "The role of risk modelling in the Great Britain transmission planning and operational standards," in *Probabilistic Methods Applied to Power Systems (PMAPS), 2010 IEEE 11th International Conference on*, 2010, pp. 325–330.

[66] C. J. Dent, A. Keane, and J. W. Bialek, "Simplified methods for renewable generation capacity credit calculation: A critical review," in *IEEE PES General Meeting*, July 2010, pp. 1–8.

[67] R. Billinton and K. Chu, "Early evolution of LOLP: Evaluating generating capacity requirements [history]," *IEEE Power and Energy Magazine*, vol. 13, no. 4, pp. 88–98, July 2015.

[68] "Electricity security of supply," The Office of Gas and Electricity Markets (Ofgem), Tech. Rep., July 2015. [Online]. Available: https://www.ofgem.gov.uk/sites/default/files/docs/2015/07/electricitysecurityofsupplyreport_final_0.pdf

[69] A. J. Pereira and J. ao Tomé Saraiva, "A decision support system for generation expansion planning in competitive electricity markets," *Electric Power Systems Research*, vol. 80, no. 7, pp. 778–787, 2010.

[70] ——, "Generation expansion planning (GEP)? a long-term approach using system dynamics and genetic algorithms (GAs)," *Energy*, vol. 36, no. 8, pp. 5180–5199, 2011, PRES 2010.

[71] ——, "A long term generation expansion planning model using system dynamics - Case study using data from the Portuguese/Spanish generation system," *Electric Power Systems Research*, vol. 97, pp. 41–50, 2013.

[72] L. Wehinger, G. Hug-Glanzmann, M. Galus, and G. Andersson, "Modeling electricity wholesale markets with model predictive and profit maximizing agents," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 868–876, May 2013.

[73] S. Islyaev and P. Date, "Electricity futures price models: Calibration and forecasting," *European Journal of Operational Research*, vol. 247, no. 1, pp. 144–154, 2015.

[74] O. Féron and E. Daboussi, *Commodities, Energy and Environmental Finance*. New York, NY: Springer New York, 2015, ch. Calibration of Electricity Price Models, pp. 183–210.

[75] Y. Y. Hong and C. Y. Hsiao, "Locational marginal price forecasting in deregulated electricity markets using artificial intelligence," *IEE Proceedings–Generation, Transmission and Distribution*, vol. 149, no. 5, pp. 621–626, Sep 2002.

[76] X. Chen, Z. Y. Dong, K. Meng, Y. Xu, K. P. Wong, and H. W. Ngan, "Electricity price forecasting with extreme learning machine and bootstrapping," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2055–2062, Nov 2012.

[77] A. Botterud and M. Korpas, "Modelling of power generation investment incentives under uncertainty in liberalised electricity markets," in *Proceedings of the Sixth IAEE European Conference 2004*, 2004, pp. 1–3.

[78] C. Barria and H. Rudnick, "Investment under uncertainty in power generation: Integrated electricity prices modeling and real options approach," *Latin America Transactions, IEEE (Revista IEEE America Latina)*, vol. 9, no. 5, pp. 785–792, 2011.

[79] L. Wang, M. Mazumdar, M. D. Bailey, and J. Valenzuela, "Oligopoly models for market price of electricity under demand uncertainty and unit reliability," *European Journal of Operational Research*, vol. 181, no. 3, pp. 1309–1321, 2007.

[80] A. Botterud, M. Mahalik, T. Veselka, H.-S. Ryu, and K.-W. Sohn, "Multi-agent simulation of generation expansion in electricity markets," in *Power Engineering Society General Meeting, 2007. IEEE*, 2007, pp. 1–8.

[81] A. Ford, "Waiting for the boom:: a simulation study of power plant construction in California," *Energy Policy*, vol. 29, no. 11, pp. 847–869, 2001.

[82] D. Eager, J. Bialek, and T. Johnson, "Validation of a dynamic control model to simulate investment cycles in electricity generating capacity," in *Power and Energy Society General Meeting, 2010 IEEE*, 2010, pp. 1–8.

[83] J. Ma, V. Silva, R. Belhomme, D. Kirschen, and L. Ochoa, "Evaluating and planning flexibility in sustainable power systems," *Sustainable Energy, IEEE Transactions on*, vol. 4, no. 1, pp. 200–209, 2013.

[84] C. D. Jonghe, E. Delarue, R. Belmans, and W. D'haeseleer, "Determining optimal electricity technology mix with high level of wind power penetration," *Applied Energy*, vol. 88, no. 6, pp. 2231–2238, 2011.

[85] A. Ehrenmann and Y. Smeers, "Generation capacity expansion in a risky environment: A stochastic equilibrium analysis," *Operations Research*, vol. 59, no. 6, pp. 1332–1346, 2011.

[86] H. Mausser and D. Rosen, "Beyond var: from measuring risk to managing risk," in *Conference on Computational Intelligence for Financial Engineering*, 1999, pp. 163–178.

[87] P. Konstantin and U. Stan, "CVaR norm and applications in optimization," *Optimization Letters*, vol. 8, no. 7, pp. 1999–2020, 2014.

[88] N. Rau and F. Zeng, "Dynamic optimizations of monte carlo simulation to assess locational capacity, transmission, and market parameters," *IEEE Transactions on Power Systems*, vol. 21, no. 1, pp. 34–42, Feb 2006.

[89] H. Khalfallah, "Long-term capacity adequacy in electricity markets: Reliability contracts vs capacity obligations," *Energy Studies Review*, vol. 16, no. 2, 2009.

[90] Y. Wang, F. Wen, C. Chung, X. Luo, and R. Xu, "A real option based approach for generation investment decision-making and generation capacity adequacy analysis," in *Power System Technology, 2006. PowerCon 2006. International Conference on*, Oct 2006, pp. 1–7.

[91] C. Gollier, D. Proult, F. Thais, and G. Walgenwitz, "Choice of nuclear power investments under price uncertainty: Valuing modularity," *Energy Economics*, vol. 27, no. 4, pp. 667–685, 2005.

[92] L. Fan, B. F. Hobbs, and C. S. Norman, "Risk aversion and CO_2 regulatory uncertainty in power generation investment: Policy and modeling implications," *Journal of Environmental Economics and Management*, vol. 60, no. 3, pp. 193–208, 2010.

[93] L. Fan, C. S. Norman, and A. G. Patt, "Electricity capacity investment under risk aversion: A case study of coal, gas, and concentrated solar power," *Energy Economics*, vol. 34, no. 1, pp. 54–61, 2012.

[94] J. Meza, M. Yildirim, and A. Masud, "A model for the multiperiod multiobjective power generation expansion problem," *IEEE Transactions on Power Systems*, vol. 22, no. 2, pp. 871–878, May 2007.

[95] A. Shortt and M. O'Malley, "Impact of variable generation in generation resource planning models," in *Power and Energy Society General Meeting, 2010 IEEE*, July 2010, pp. 1–6.

[96] M. Lynch, A. Shortt, R. Tol, and M. O'Malley, "The effect of operational considerations on the return of electricity generation investment," in *Power and Energy Society General Meeting (PES), 2013 IEEE*, July 2013, pp. 1–5.

[97] B. Palmintier and M. Webster, "Impact of unit commitment constraints on generation expansion planning with renewables," in *Power and Energy Society General Meeting, 2011 IEEE*, 2011, pp. 1–7.

[98] ——, "Heterogeneous unit clustering for efficient operational flexibility modeling," *IEEE Transactions on Power Systems*, vol. 29, no. 3, pp. 1089–1098, May 2014.

[99] V. Guigues and C. Sagastizbal, "The value of rolling-horizon policies for risk-averse hydro-thermal planning," *European Journal of Operational Research*, vol. 217, no. 1, pp. 129–140, 2012.

[100] J. Hargreaves, E. K. Hart, R. Jones, and A. Olson, "REFLEX: An adapted production simulation methodology for flexible capacity planning," *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1306–1315, May 2015.

[101] M. S. ElNozahy, M. M. A. Salama, and R. Seethapathy, "A probabilistic load modelling approach using clustering algorithms," in *2013 IEEE Power Energy Society General Meeting*, July 2013, pp. 1–5.

[102] S. Jin, A. Botterud, and S. Ryan, "Temporal versus stochastic granularity in thermal generation capacity planning with wind power," *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2033–2041, Sept 2014.

[103] M. D. W. Fernando J. de Sisternes, "Optimal selection of sample weeks for approximating the net load in generation planning problems," Tech. Rep., Jan 2013. [Online]. Available: http://web.mit.edu/mort/www/SisternesWebster2012.pdf

[104] P. Kall and S. W. Wallace, *Stochastic Programming*. John Wiley & Sons, Chichester, 1994.

[105] S. Jin, A. Botterud, and S. Ryan, "Impact of demand response on thermal generation investment with high wind penetration," *Smart Grid, IEEE Transactions on*, vol. 4, no. 4, pp. 2374–2383, Dec 2013.

[106] P. Rocha and D. Kuhn, "Multistage stochastic portfolio optimisation in deregulated electricity markets using linear decision rules," *European Journal of Operational Research*, vol. 216, no. 2, pp. 397–408, 2012.

[107] W. B. Powell, A. George, H. Simão, W. Scott, A. Lamont, and J. Stewart, "Smart: A stochastic multiscale model for the analysis of energy resources, technology, and policy," *INFORMS J. on Computing*, vol. 24, no. 4, pp. 665–682, oct 2012.

[108] P. A. J. Fonseka, T. K. Saha, and Z. Y. Dong, "A price-based approach to generation investment planning in electricity markets," *IEEE Transactions on Power Systems*, vol. 23, no. 4, pp. 1859–1870, Nov 2008.

[109] T. Barforoushi, M. Moghaddam, M. Javidi, and M. Sheikh-El-Eslami, "Evaluation of regulatory impacts on dynamic behavior of investments in electricity markets: A new hybrid DP/GAME framework," *IEEE Transactions on Power Systems*, vol. 25, no. 4, pp. 1978–1986, 2010.

[110] "Expansion planning for electrical generating systems: A guidebook," International Atomic Energy Agency, Tech. Rep., 1984. [Online]. Available: http://home.eng.iastate.edu/ jdm/ee552/TRS241_Web.pdf

[111] W. B. Powell, "Merging AI and OR to solve high-dimensional stochastic optimization problems using approximate dynamic programming," *INFORMS Journal on Computing*, vol. 22, no. 1, pp. 2–17.

[112] J. D. Sterman, "System dynamics modeling: Tools for learning in a complex world," *California Management Review*, vol. 43, no. 4, pp. 8–25, 2001.

[113] F. Teufel, M. Miller, M. Genoese, and W. Fichtner, "Review of system dynamics models for electricity market simulations," in *Working paper series in production and energy*, 2013, p. 34 p.

[114] D. W. Bunn and E. R. Larsen, "Assessment of the uncertainty in future UK electricity investment using an industry simulation model," *Utilities Policy*, vol. 4, no. 3, pp. 229–236, 1994.

[115] T. Kadoya, T. Sasaki, S. Ihara, E. Larose, M. Sanford, A. Graham, C. Stephens, and C. Eubanks, "Utilizing system dynamics modeling to examine impact of deregulation on generation capacity growth," *Proceedings of the IEEE*, vol. 93, no. 11, pp. 2060–2069, 2005.

[116] S. Jin and S. M. Ryan, "A tri-level model of centralized transmission and decentralized generation expansion planning for an electricity market–Part I," *IEEE Transactions on Power Systems*, vol. 29, no. 1, pp. 132–141, Jan 2014.

[117] G. Gurkan, O. Ozdemir, and Y. Smeers, "Generation capacity investments in electricity markets: Perfect competition," in *CentER Discussion Paper Series*, no. 2013-2045, May 2013.

[118] V. Nanduri, T. Das, and P. Rocha, "Generation capacity expansion in energy markets using a two-level game-theoretic model," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1165–1172, 2009.

[119] J. Wang, M. Shahidehpour, Z. Li, and A. Botterud, "Strategic generation capacity expansion planning with incomplete information," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 1002–1010, 2009.

[120] S. Kazempour, A. Conejo, and C. Ruiz, "Strategic generation investment using a complementarity approach," *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 940–948, 2011.

[121] F. H. Murphy and Y. Smeers, "Generation capacity expansion in imperfectly competitive restructured electricity markets," *Operations Research*, vol. 53, no. 4, pp. 646–661, 2005.

[122] M. Ventosa, R. Denis, and C. Redondo, "Expansion planning in electricity markets, two different approaches," in *Proc. 14th PSCC Conf., Seville*, July 2002.

[123] S. Kazempour, A. Conejo, and C. Ruiz, "Strategic generation investment using a complementarity approach," *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 940–948, 2011.

[124] S. A. Gabriel, A. J. Conejo, J. D. Fuller, B. F. Hobbs, and C. Ruiz, *Complementarity Modeling in Energy Markets*. Springer Publishing Company, Incorporated, 2012.

[125] A. F. Taha and J. H. Panchal, "Decision-making in energy systems with multiple technologies and uncertain preferences," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 44, no. 7, pp. 894–907, July 2014.

[126] M. Goldstein and J. Rougier, "Bayes linear calibrated prediction for complex systems," *Journal of the American Statistical Association*, vol. 101, no. 475, pp. 1132–1143, 2006.

[127] M. C. Kennedy and A. O'Hagan, "Bayesian calibration of computer models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 63, no. 3, pp. 425–464, 2001.

[128] L. Uusitalo, A. Lehikoinen, I. Helle, and K. Myrberg, "An overview of methods to evaluate uncertainty of deterministic models in decision support," *Environmental Modelling & Software*, vol. 63, pp. 24–31, 2015.

[129] E. Ares, "Carbon price floor," 2014. [Online]. Available: http://www.parliament.uk/briefing-papers/sn05927.pdf

[130] A. Serebrenik and M. van den Brand, "Theil index for aggregation of software metrics values," in *2010 IEEE International Conference on Software Maintenance*, Sept 2010, pp. 1–9.

[131] A. Borison, "Electric power resource planning under uncertainty: Critical review and best practices," 2014. [Online]. Available: www.thinkbrg.com/media/publication514_Borison_ResourcePlanningUncertainty_WP_20140121.pdf

[132] D. E. Morris, J. E. Oakley, and J. A. Crowe, "A web-based tool for eliciting probability distributions from experts," *Environmental Modelling & Software*, vol. 52, pp. 1–4, 2014.

[133] A. O'Hagan, "Bayesian analysis of computer code outputs: A tutorial," *Reliability Engineering & System Safety*, vol. 91, no. 10C11, pp. 1290–1300, 2006.

[134] A. J. Pereira and J. ao Tomé Saraiva, "Generation expansion planning (GEP) – a long-term approach using system dynamics and genetic algorithms (GAs)," *Energy*, vol. 36, no. 8, pp. 5180–5199, 2011.

[135] C.-K. Han, D. Hur, J.-M. Sohn, and J.-K. Park, "Assessing the impacts of capacity mechanisms on generation adequacy with dynamic simulations," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 1788–1797, Nov 2011.

[136] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola, "Variance based sensitivity analysis of model output. design and estimator

for the total sensitivity index," *Computer Physics Communications*, vol. 181, no. 2, pp. 259–270, 2010.

[137] J. E. Oakley and A. O'Hagan, "Probabilistic sensitivity analysis of complex models: A Bayesian approach," *Journal of the Royal Statistical Society, Series B*, vol. 66, pp. 751–769, 2002.

[138] H. C. Boshuizen and P. H. van Baal, "Probabilistic sensitivity analysis: Be a bayesian," *Value in Health*, vol. 12, no. 8, pp. 1210–1214, 2009.

[139] C. Gourieroux, A. Montfort, and E. Renault, "Indirect inference," *Journal of Applied Econometrics*, vol. 8, pp. S85–S118, 1993.

[140] M. A. Beaumont, W. Zhang, and D. J. Balding, "Approximate bayesian computation in population genetics," *Genetics*, vol. 162, pp. 2025–2035, 2002.

[141] P. Marjoram, J. Molitor, V. Plagnol, and S. Tavar, "Markov chain monte carlo without likelihoods," *Proceedings of the National Academy of Sciences*, vol. 100, no. 26, pp. 15 324–15 328, 2003.

[142] I. Vernon, M. Goldstein, and R. G. Bower, "Galaxy formation: A bayesian uncertainty analysis," *Bayesian Analysis*, vol. 05, no. 4, pp. 619–670, 2010.

[143] P. Holden, N. Edwards, K. Oliver, T. Lenton, and R. Wilkinson, "A probabilistic calibration of climate sensitivity and terrestrial carbon change in genie-1," *Climate Dynamics*, vol. 35, no. 5, pp. 785–806, 2010.

[144] S. Conti and A. O'Hagan, "Bayesian emulation of complex multi-output and dynamic computer models," *Journal of Statistical Planning and Inference*, vol. 140, no. 3, pp. 640–651, 2010.

[145] I. Andrianakis, I. R. Vernon, N. McCreesh, T. J. McKinley, J. E. Oakley, R. N. Nsubuga, M. Goldstein, and R. G. White, "Bayesian history matching of complex infectious disease models using emulation: A tutorial and a case study on hiv in uganda," *PLoS Compututational Biology*, vol. 11, no. 1, 2015.

[146] D. A. Henderson, R. J. Boys, K. J. Krishnan, C. Lawless, and D. J. Wilkinson, "Bayesian emulation and calibration of a stochastic computer model of mitochondrial DNA deletions in substantia nigra neurons," *Journal of the American Statistical Association*, vol. 104, no. 485, pp. 76–87, 2009.

[147] D. Higdon, J. Gattiker, B. Williams, and M. Rightley, "Computer model calibration using high-dimensional output," *Journal of the American Statistical Association*, vol. 103, no. 482, pp. 570–583, 2008.

[148] R. S. Pindyck, "The long-run evolution of energy prices," *The Energy Journal*, vol. 20, no. 2, pp. 1–27, 1999.

[149] B. F. H. Jeremy J. Hargreaves, "Metamodeling of inputcoutput relationships for complex power market models," *Energy Systems*, vol. 4, no. 1, pp. 25–45, 2013.

[150] A. O'Hagan, "Dicing with the unknown," *Significance*, vol. 1, pp. 132–133, 2004.

[151] J. Dupačová, N. Gröwe-Kuska, and W. Römisch, "Scenario reduction in stochastic programming," *Mathematical Programming*, vol. 95, no. 3, pp. 493–511, 2003.

[152] J. Morales, S. Pineda, A. Conejo, and M. Carriń, "Scenario reduction for futures market trading in electricity markets," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 878–888, 2009.

[153] B. Zhang, B. A. Konomi, H. Sang, G. Karagiannis, and G. Lin, "Full scale multi-output gaussian process emulator with nonseparable auto-covariance functions," *Journal of Computational Physics*, vol. 300, pp. 623–642, 2015.

[154] D. Gamerman and H. F. Lopes, *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, 2nd ed. CRC Press, 2006, ch. 7.

[155] F. A. C. Viana, "Things you wanted to know about the latin hypercube design and were afraid to ask," in *10th World Congress on Structural and Multidisciplinary Optimization*, May 2013, pp. 1–9.

[156] J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn, "Design and analysis of computer experiments," *Statistical Science*, vol. 4, no. 4, pp. 409–423, 1989.

[157] G. Pistone and G. Vicario, *Design for Computer Experiments: Comparing and Generating Designs in Kriging Models.* Milano: Springer Milan, 2009, pp. 91–102.

[158] M. Morris and T. Mitchell, "Exploratory designs for computer experiments," *Journal of Statistical Planning and Inference*, vol. 43, pp. 381–402, 1995.

[159] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 4, pp. 433–459, 2010.

[160] F. Ronquist, "Bayesian inference," 2005. [Online]. Available: http://people.sc.fsu.edu/ pbeerli/BSC-5936/10-05-05/Lecture_11.pdf

[161] L. S. Bastos and A. O'Hagan, "Diagnostics for gaussian process emulators," *Technometrics*, vol. 51, no. 4, pp. 425–438, 2009.

[162] D. Higdon, *Space and Space-Time Modeling using Process Convolutions.* London: Springer London, 2002, pp. 37–56.

[163] "Carbon values used in DECC's energy modelling," Department of Energy & Climate Change, Tech. Rep., 2011. [Online]. Available: http://www.gov.uk/government/uploads/system/uploads/attachment_data/file/48185/3138-carbon-values-decc-energy-modelling.pdf

[164] A. Barillas, R. Cohen, and N. Sukthaworn, "Lessons from the first capacity market auction in Great Britain," in *European Energy Market (EEM), 2015 12th International Conference on the*, May 2015, pp. 1–6.

[165] B. Hasche, A. Keane, and M. O'Malley, "Capacity value of wind power, calculation, and data requirements: the irish power system case," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 420–430, Feb 2011.

[166] C. P. Robert and G. Casella, *Monte Carlo statistical methods.* Springer-Verlag Inc, 1999.

[167] "Dynamics of GB generation investment," Redpoint Energy, Tech. Rep., 2006. [Online]. Available: http://www.redpointenergy.co.uk/files/final_report_final.pdf

[168] M. R. Hummon, P. Denholm, J. Jorgenson, D. Palchak, B. Kirby, and O. Ma, "Fundamental drivers of the cost and price of operating reserves," National Renewable Energy Laboratory (NREL), Tech. Rep., July 2013. [Online]. Available: http://www.nrel.gov/docs/fy13osti/58491.pdf

[169] O'Neill, R. P., K. W. Hedman, E. A. Krall, A. Papavasiliou, and S. S. Oren, "Economic analysis of the N-1 reliable unit commitment and transmission switching problem using duality concepts," *Energy Systems*, vol. 1, no. 2, pp. 165–195, 2010.

[170] R. Loulou, U. Remne, A. Kanudia, A. Lehtila, and G. Goldstein, "Documentation for the times model part I," Energy Technology Systems Analysis Programme, Tech. Rep., 2005. [Online]. Available: http://www.iea-etsap.org/web/docs/timesdoc-intro.pdf

[171] "2006 Minnesota wind integration study volume I," Enernex Corporation, Tech. Rep., November 2006, prepared for The New York State Energy Research and Development Authority. [Online]. Available: http://www.uwig.org/windrpt_vol201.pdf

[172] M. D. la Torre, "Managing uncertainty at REE (Spanish TSO)," Utility Wind Integration Group (UWIG), Tech. Rep., 2010. [Online]. Available: http://www.uwig.org

[173] E. Ela, B. Kirby, E. Lannoye, M. Milligan, D. Flynn, B. Zavadil, and M. O'Malley, "Evolution of operating reserve determination in wind power integration studies," in *IEEE PES General Meeting*, July 2010, pp. 1–8.

[174] "Western wind and solar integration study," GE Energy, Tech. Rep., May 2010, prepared for the National Renewable Energy Laboratory. [Online]. Available: http://www.nrel.gov/docs/fy10osti/47434.pdf

[175] A. Papavasiliou, S. Oren, and R. O'Neill, "Reserve requirements for wind power integration: A scenario-based stochastic programming framework," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2197–2206, Nov 2011.

[176] E. Ela, B. Kirby, N. Navid, and J. Smith, "Effective ancillary services market designs on high wind power penetration systems," in *Power and Energy Society General Meeting, 2012 IEEE*, July 2012, pp. 1–8.

[177] Z. Zhou and A. Botterud, "Dynamic scheduling of operating reserves in co-optimized electricity markets with wind power," *IEEE Transactions on Power Systems*, vol. 29, no. 1, pp. 160–171, Jan 2014.

[178] B. Hellrich-Dawson, "Price formation in organized wholesale electricity markets," Federal Energy Regulatory Commission, Tech. Rep., October 2014. [Online]. Available: https://www.ferc.gov/legal/staff-reports/2014/AD14-14-pricing-rto-iso-markets.pdf

[179] Z. Alaywan, T. Wu, and A. Papalexopoulos, "Transitioning the california market from a zonal to a nodal framework: an operational perspective," in *Power Systems Conference and Exposition, 2004. IEEE PES*, vol. 2, Oct 2004, pp. 862–867.

[180] A. Negash, T. Haring, and D. Kirschen, "Allocating the cost of demand response compensation in wholesale energy markets," *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1528–1535, May 2015.

[181] M. Klos, K. Wawrzyniak, M. Jakubek, and G. Orynczak, "The scheme of a novel methodology for zonal division based on power transfer distribution factors," in *Industrial Electronics Society, IECON 2014 - 40th Annual Conference of the IEEE*, 2014, pp. 3598–3604.

[182] C. Breuer, N. Seeger, and A. Moser, "Determination of alternative bidding areas based on a full nodal pricing approach," in *Power and Energy Society General Meeting (PES), 2013 IEEE*, July 2013, pp. 1–5.

[183] C. Breuer and A. Moser, "Optimized bidding area delimitations and their impact on electricity markets and congestion management," in *European Energy Market (EEM), 2014 11th International Conference on the*, May 2014, pp. 1–5.

[184] "Bidding zones review process," European Network of Transmission System Operators for Electricity (ENTSO-E), Tech. Rep., January 2014, https://www.entsoe.eu/Documents/MC documents/140123_Technical_Report_-_Bidding_Zones_Review__Process.pdf.

[185] F. Wang and K. Hedman, "Dynamic reserve zones for day-ahead unit commitment with renewable resources," *IEEE Transactions on Power Systems*, vol. 30, no. 2, pp. 612–620, March 2015.

[186] J. Lyon, F. Wang, K. Hedman, and M. Zhang, "Market implications and pricing of dynamic reserve policies for systems with renewables," *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1593–1602, May 2015.

[187] C. Kang, Q. Chen, W. Lin, Y. Hong, Q. Xia, Z. Chen, Y. Wu, and J. Xin, "Zonal marginal pricing approach based on sequential network partition and congestion contribution identification," *International Journal of Electrical Power & Energy Systems*, vol. 51, pp. 321–328, 2013.

[188] M. Klos, K. Wawrzyniak, M. Jakubek, and G. Oryńczak, "The scheme of a novel methodology for zonal division based on power transfer distribution factors," in *IECON 2014 - 40th Annual Conference of the IEEE Industrial Electronics Society*, Oct 2014, pp. 3598–3604.

[189] A. Kumar, S. Srivastava, and S. Singh, "A zonal congestion management approach using real and reactive power rescheduling," *IEEE Transactions on Power Systems*, vol. 19, no. 1, pp. 554–562, Feb 2004.

[190] T. Zheng and E. Litvinov, "Contingency-based zonal reserve modeling and pricing in a co-optimized energy and reserve market," *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 277–286, May 2008.

[191] R. Zimmerman, C. Murillo-Sanchez, and R. Thomas, "Matpower's extensible optimal power flow architecture," in *IEEE Power Energy Society General Meeting (PES)*, July 2009, pp. 1–7.

[192] N. Zhang, C. Kang, Q. Xia, Y. Ding, Y. Huang, R. Sun, J. Huang, and J. Bai, "A convex model of risk-based unit commitment for day-ahead market clearing considering wind power uncertainty," *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1582–1592, May 2015.

[193] "Market optimization details," California ISO, Tech. Rep., 2009. [Online]. Available: www.caiso.com/23cf/23cfe2c91d880.pdf

[194] J. Lyon, M. Zhang, and K. Hedman, "Locational reserve disqualification for distinct scenarios," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 357–364, Jan 2015.

[195] G. Morales-Espana, J. Latorre, and A. Ramos, "Tight and compact milp formulation for the thermal unit commitment problem," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4897–4908, Nov 2013.

[196] ——, "Tight and compact milp formulation of start-up and shut-down ramping in unit commitment," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1288–1296, May 2013.

[197] Y. Fu and Z. Li, "Different models and properties on lmp calculations," in *Power Engineering Society General Meeting, IEEE*, 2006, pp. 1–11.

[198] F. C. Schweppe, M. C. Caramanis, R. D. Tabors, and R. E. Bohn, *Spot Pricing of Electricity*. Springer US, 1988.

[199] F. Li and R. Bo, "Dcopf-based lmp simulation: Algorithm, comparison with acopf, and sensitivity," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 1475–1485, Nov 2007.

[200] B.-M. Hodge, D. Lew, and M. Milligan, "Short-term load forecasting error distributions and implications for renewable integration studies," in *IEEE Green Technologies Conference*, January 2013.

[201] R. Billinton and R. N. Allan, *Reliability evaluation of power systems*, 1st ed. Springer US, 1996.

[202] E. D. Castronuovo and J. A. P. Lopes, "On the optimization of the daily operation of a wind-hydro power plant," *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1599–1606, Aug 2004.

[203] M. Milligan, M. Schwartz, and Y. Wan, "Statistical wind power forecasting models: Results for U.S. wind farms," in *WINDPOWER 2003*. National Renewable Energy Laboratory, May 2003, pp. 18–21.

[204] A. Fabbri, T. G. S. Roman, J. R. Abbad, and V. H. M. Quezada, "Assessment of the cost associated with wind generation prediction errors in a liberalized electricity market," *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1440–1446, Aug 2005.

[205] J. Usaola, "Probabilistic load flow in systems with wind generation," *Generation, Transmission Distribution, IET*, vol. 3, no. 12, pp. 1031–1041, December 2009.

[206] H. Bludszuweit, J. Dominguez-Navarro, and A. Llombart, "Statistical analysis of wind power forecast error," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 983–991, Aug 2008.

[207] E. Cotilla-Sanchez, P. Hines, C. Barrows, S. Blumsack, and M. Patel, "Multi-attribute partitioning of power networks based on electrical distance," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4979–4987, Nov 2013.

[208] U. von Luxburg, "A tutorial on spectral clustering," *Statistics and Computing*, vol. 17, no. 4, pp. 395–416, 2007.

[209] R. Christie, "Power systems test case archive," http://www2.ee.washington.edu/research/pstca/pf118/pg_tca118bus.htm, 1999, University of Washington.

[210] F. Wang and K. Hedman, "Reserve zone determination based on statistical clustering methods," in *North American Power Symposium (NAPS), 2012*, Sept 2012, pp. 1–6.

[211] A. L. Ott, "Experience with PJM market operation, system design, and implementation," *IEEE Transactions on Power Systems*, vol. 18, no. 2, pp. 528–534, May 2003.

[212] S. Bhattacharya, "A simulation approach to Bayesian emulation of complex dynamic computer models," *Bayesian Anal.*, vol. 2, no. 4, pp. 783–815, 12 2007.

[213] S. Conti, J. P. Gosling, J. E. Oakley, and A. O'Hagan, "Gaussian process emulation of dynamic computer codes," *Biometrika*, vol. 96, no. 3, pp. 663–676, 2009.

[214] G. Isabelle. and A. Elisseeff, "An introduction to variable and feature selection," *Journal of Machine Learning Research*, vol. 3, pp. 1157–1182, 2003.

[215] P. Linares, F. Javier Santos, M. Ventosa, and L. Lapiedra, "Incorporating oligopoly, CO_2 emissions trading and green certificates into a power generation expansion model," *Automatica*, vol. 44, no. 6, pp. 1608–1620, Jun. 2008.

[216] M. Webster, N. Santen, P. Parpas, M. Webster, N. S. Panos, M. Webster, N. Santen, and P. Parpas, "An approximate dynamic programming framework for modeling global climate policy under decision-dependent uncertainty," *Comput Manag Sci*, vol. 9, no. 3, pp. 339–362, 2012.

Appendix A

Basic and Auxiliary Results

Section A.1 provides the Matlab source code for three major functions used in the application of Bayesian emulation to a LTGI model. The first function is the main routine consisting of subroutines for implementing calibration, uncertainty analysis, sensitivity analysis based on Bayesian emulation. The second function is a subroutine for processing and packaging main inputs including simulation data, observations and the model discrepancy. The third function is a subroutine for plotting main outputs including the posterior distribution of calibration parameters, model predictions estimated by a calibrated and discrepancy-adjusted emulator, and results from variance-based sensitivity analysis.

Section A.2 provides the Matlab source code for spectral clustering which is used in reserve zone partitioning. Section A.3 gives the code for developing piecewised linearised zonal operating reserve demand curves given the probability mass function of real-time system imbalances. All comments to the code are in *emph* italics. Section A.4 lists the publications over the course of PhD.

A.1 Matlab source code for Bayesian applications

A.1.1 The main function

% read data

dat=fd(1,'pcpct',0.99);

% run setupModel.m, which structures simulation data and observations, and assigns defaults for priors in the emulator, and MCMC parameters.

params=setupModel(dat.obsData,dat.simData);

% run automated step-size selection algorithm; nlev: number of candidate step sizes for each parameter; nburn: number of MCMC iterations at each candidate step size.

nburn=500; nlev=21;

params=stepsize(params,nburn,nlev);

% run MCMC subroutine; nmcmc: the number of draws in one call; pout: containing everything from set-up and MCMC results, including the posterior samples for calibration parameters kept in the pout.pvals.

nmcmc=10000;

pout=gpmmcmc(params,nmcmc,'step',1);

save pout pout;

nmcmc=nmcmc+nburn*nlev;

pvec=floor(linspace(nburn*nlev+1,nmcmc,500));

pout.pvec=pvec; *% pvec: subset of MCMC runs to be used for prediction*

% run main output subroutine for generating plots.

fdPlots(pout,pvec,1:3);

theta = [pout.pvals.theta]';*% theta: posterior samples of calibration parameters*

% variance-based sensitivity analysis; rn: the perturbation ranges of six model inputs.

rn=[0 1;0 1;0 1;0 1;0 1;0 1];

% gSens: computes main and total effect sensitivity indices (two-factor interaction effect indices optional).

sens=gSens(pout,'pvec',pvec,'varlist','all','rg',rn);

pout.sens=sens; *% sens: structure storing all sensitivity output*

save pout pout; *% pout: containing everything from set-up and MCMC, calibration and sensitivity results*

StmPm=[sens.smePm;sens.stePm]'; *% smePm: holds mean of main effect sensitivity indices; stePm: holds mean of total effect sensitivity indices*

SiePm=sens.siePm'; *% siePm: holds mean of two-factor interaction effect sensitivity indices*

A.1.2 Main inputs

% Read data file: fd.m

function params=fd(doPlot,varargin);

pcpct=0.95; *% pcpct: proportion of total variance explained by principal components*

nkern=3; *% number of normal kernals used for modeling the discrepancy*

% read in design

design=textread('design.txt'); m=size(design,1);

% standardize inputs to unit hypercube

xmin=min(design); xrange=max(design)-xmin;

design=(design-repmat(xmin,m,1))./repmat(xrange,m,1);

% read in simulation data

simdata=textread('sim_outputs'); *% simdata: first column is "time/year," remaining columns are model outputs*

ysim=simdata(:,2:end); tsim=simdata(:,1); *% ysim: time series of model runs; tsim: planning year*

% read in observation data

n=1; *% number of observations*

for ii=1:n inf=['obs_outputs' int2str(ii)];

obsdataii=textread(inf); *% obsdata: observation data*

obsdata{ii}=obsdataii(1:end-1,:); ydat{ii}=obsdataii(:,2);

tdat{ii}=obsdataii(:,1)-min(obsdataii(:,1)); Sigy{ii}=1.0^2.*eye(length(ydatii));

end

% summary statistics from simulation data

ysimmean=mean(ysim,2); ysimStd=ysim-repmat(ysimmean,1,m);

% ysimStd: standardized simulation data - subtracting row means ("time") and divided by overall standard deviation

ysimsd=std(ysimStd(:)); ysimStd=ysimStd/ysimsd;

% interpolate to data grid and standardize observation data consistent with the stan-

dardization of simulation data

for ii=1:n

yobs(ii).y=ydatii; yobs(ii).t=tdatii;

yobs(ii).ymean=interp1(tsim,ysimmean,yobs(ii).t,'linear','extrap');

yobs(ii).yStd=(yobs(ii).y-yobs(ii).ymean)/ysimsd;

yobs(ii).Sigy=Sigyii;

end

% compute principal components Ksim on simulations ysimStd using singular value decomposition (SVD);S is an n-by-p rectangular diagonal matrix of non-negative singular values of ysimStd; U is an n-by-n matrix whose columns are the left singular vectors of ysimStd; and W is a p-by-p whose columns are the right singular vectors of ysimStd.

[U,S,V]=svd(ysimStd,0);

lam=diag(S).^2/sum(diag(S).^2); lam=cumsum(lam); *% lam is a vector of singular values that are the square roots of the eigenvalues of the matrix ysimStd^TysimStd*

pu=sum(lam<pcpct)+1; *% pu is the specific rank of the truncated matrix Ksim; only the pu largest singular values are considered in the low-rank matrix approximation*

Ksim=U(:,1:pu)*S(1:pu,1:pu)./sqrt(m); % a truncated n-by-pu score matrix Ksim that is the nearest possible matrix of rank pu to the original matrix

% linear interpolation of eigenvectors constructed from simulator runs to obtain their representation on the data grid

for ii=1:n yobs(ii).Kobs=zeros(length(yobs(ii).yStd),pu);

for jj=1:pu yobs(ii).Kobs(:,jj)=interp1(tsim,Ksim(:,jj),yobs(ii).t,'linear','extrap'); end

end

% construct kernel basis for discrepancy: 3 equally-spaced kernels with bandwidth equal to separation between centers

Dgrid = [3 7 11]; *% user-specified kernel centers*

Dwidth = Dgrid(2)-Dgrid(1); *% user-specified widths of model discrepancy*

pv=length(Dgrid); *% user-specified number of kernels for discre pancy*

% compute the kernel function map, for each kernel

```
Dsim=zeros(size(ysimStd,1),pv);
for jj=1:pv
% first the observations
for ii=1:n
yobs(ii).Dobs(:,jj)=normpdf(yobs(ii).t,Dgrid(jj),Dwidth);
end
% now the simulations
Dsim(:,jj)=normpdf(tsim,Dgrid(jj),Dwidth);
end
% normalize discrepancy kernel basis for numerical stability
Dmax=max(max(Dsim*Dsim'));
Dsim=Dsim/sqrt(Dmax);
for ii=1:n; yobs(ii).Dobs=yobs(ii).Dobs/sqrt(Dmax); end
```

A.1.3 Main outputs

```
%% Plot main results fdPlots.m; plotnum indicates which plots are to be drawn
function fdPlots(pout,pvec,plotnum,varargin)
xlabs={};% control and forcing input labels for plotting
thlabs={'θ_VaR','θ_markup'};% calibration parameter labels for plotting
labs=[xlabs thlabs]; nxlabs=length(xlabs); nlabs=length(labs);
model=pout.model; data=pout.data; pvals=pout.pvals(pvec); nreal=length(pvals);
pu=model.pu; % pu: number of principal components
pv=model.pv; % pv: number of kernel basis functions
p=model.p; % p: number of x parameters (control and forcing)
q=model.q; % q: number of θ parameters (calibration)

doPlot(1:3)=0;
if exist('plotnum'); doPlot(plotnum)=1; end

% process input arguments
ngrid=21; subset=1:nlabs; lsub=length(subset);
parseAssignVarargs('ngrid','subset');
```

```
if nxlabs, thsub=subset(subset>p)-p; else thsub=subset; end

% Set up prediction grid
grid=linspace(0,1,ngrid);

if doPlot(1)
% plot density estimates for univariate and bivariate marginal posterior distributions
of calibration parameters θ
figure(1); clf;
if length(pvec)>1000, pvec2=pvec(floor(linspace(1,length(pvec),1000)));
else pvec2=pvec; end
t=zeros(length(pvec),q);
for i=1:q for j=1:length(pvec) t(j,i)=pout.pvals(pvec2(j)).theta(i); end end
t=t(:,thsub); gPlotMatrix(t,'Pcontours',[0.5 0.9],'ustyle','imcont',...
'lstyle','imcont','ngrid',ngrid,'ksd',0.1,'labels',thlabs(thsub));
figure(1); print -depsc2 fdPost;
end

if doPlot(2)
% plot predictions produced by the calibrated emulator and by the calibrated and
discrepancy-adjusted emulator, as well as the model discrepancy
h=[]; ctr=0;
if ctr,
for ii=1:pout.model.m ysim(:,ii)=pout.simData.orig.y(:,ii)-pout.simData.orig.ymean;
end
for ii=1:pout.model.n yobsii=pout.obsData(ii).orig.y-pout.obsData(ii).orig.ymean;
end
else
ysim=pout.simData.orig.y;
for ii=1:pout.model.n, yobsii=pout.obsData(ii).orig.y; end
end
reps=0;
if reps ind=1; nind=length(ind);
```

else ind=1:pout.model.n; nind=pout.model.n; end

tr=[min(pout.simData.orig.t) max(pout.simData.orig.t)];

rtr=range(tr);

for ii=1:nind jj=ind(ii); figure(2); clf;

outF=strcat('fdPreds',int2str(ii));

% calibrated prediction eta

pred=gPred(pout.obsData(jj).x,pvals,model,data,'uvpred');

% pred.u: array of posterior realizations from simulator model weight processes evaluated at calibrated values of θ

% eta: prediction produced by the emulator (statistical representation of simulator) at each θ*, on original scale*

eta=pout.simData.Ksim*pred.u'.*pout.simData.orig.ysd; save 'etapred' eta '-ascii';

% etabounds: 5*% and* 95*% credible intervals from eta*

etabounds=prctile(eta,[5 95],2);

% meanmat: matrix with columns equal to the mean of the original simulations across the planning horizon if ctr, meanmat=0;

else meanmat=repmat(pout.simData.orig.ymean,[1 2]); end

% delta: discrepancy

% pred.v: array of posterior realizations from discrepancy model weight processes

% deltaR: discrepancy realizations on original scale

deltaR=pout.simData.orig.Dsim*pred.v'.*pout.simData.orig.ysd;

save 'deltapred' deltaR '-ascii';

deltaRbounds=prctile(deltaR,[5 95],2); *% deltaRbounds:* $5-95$*% credible intervals from deltaR*

yhat=deltaR+eta;*% yhat: calibrated simulator (eta) + discrepancy (deltaR)*

yhatbounds=prctile(yhat,[5 95],2); *% yhatbounds:* $5-95$*% credible intervals from yhat*

h(1)=gPackSubplot(1,3,1,1); plot(pout.simData.orig.t,ysim,'y'); hold on;

if ii==nind, ll=pout.model.n; else ll=ind(ii+1)-1; end

for kk=jj:ll, plot(pout.obsData(kk).orig.t,yobskk,'bo'); hold on; end

```
plot(pout.simData.orig.t,etabounds+meanmat,'g','LineWidth',1);
ylabel('Thermal Capacity [GW]','FontSize',15);
title('calibrated simulator','Fontsize',15);

h(2)=gPackSubplot(1,3,1,2); plot(pout.simData.orig.t,ysim,'y'); hold on;
for kk=jj:ll, plot(pout.obsData(kk).orig.t,yobskk,'bo'); hold on; end
plot(pout.simData.orig.t,yhatbounds+meanmat,'k','LineWidth',1);
set(gca,'YtickLabel','');xlabel('Time [year]','FontSize',12);
title('discrepancy-adjusted','Fontsize',12);

h(3)=gPackSubplot(1,3,1,3);plot(pout.simData.orig.t,deltaRbounds,'c','LineWidth',1);
hold on;
line(tr(1):rtr/100:tr(2),0,'LineStyle','--');set(gca,'YaxisLocation','right');
title('discrepancy','Fontsize',12); figure(2); print('-depsc2',outF);
end end

if doPlot(3)
h=[]; ctr=0;
me=pout.sens.tmef.m; npred=size(me,3);
if ctr, meanmat=repmat(pout.simData.orig.ymean,[1 npred]); else meanmat=0; end
figure(3); clf; colormap('copper');
tdat=0:1.0/(npred-1):1.0; AzEl=[45 55]; np=min(lsub,4);
for ii=1:lsub jj=subset(ii);
if nxlabs tt=pout.simData.orig.xrange(jj)*tdat+pout.simData.orig.xmin(jj);
else tt=pout.simData.orig.xrange(p+jj)*tdat+pout.simData.orig.xmin(p+jj); end
r=squeeze(me(jj,:,:))-meanmat;
h(ii)=gPackSubplot(ceil(lsub/np),np,ceil(ii/np),mod(ii-1,np)+1,0.6);
surf(repmat(pout.simData.orig.t,size(tdat)),...
repmat(tt,size(pout.simData.orig.t)),r); colormap hsv; view(AzEl);
xlabel('Time [year]'); ylabel(labs(jj)); zlabel('Thermal Capacity'); alpha(0.25);
axis tight; set(gca,'Xgrid','on','Ygrid','on','Zgrid','on'); end
figure(3); print -depsc2 fdMeSens; end
```

A.2 Matlab source code for spectral clustering

%% Run SpectralCluster.m for hourly zone determination using spectral clustering % PTDF: the PTDF matrix of the transmission network; w: the matrix of line weights.

function [idx k] = SpectralCluster(PTDF,w)

[numl numb] = size(PTDF); *% numl: number of lines; numb: number of buses.*

WPTDFD = zeros(numb,numb);*% WPTDFD: the numb-by-numb weighted Euclidean distance between any two buses.*

SimMat = zero(numb,numb);*% SimMat: the numb-by-numb similarity matrix of two buses.*

D = zeros(numb,numb);*% D: the numb-by-numb diagonal elements of the adjacency matrix.*

for i = 1:numb

for j = 1:numb if i =j WPTDFD(i,j) = w*abs(PTDF(:,i) - PTDF(:,j)).ˆ2; SimMat(i,j) = exp(-WPTDFD(i,j)/2*2ˆ2); end end

D(i,i) = sum(SimMat(i,:));*% D: the diagonal degree matrix.*

end

k = 8;*% k: the initial number of eigenvectors to be taken in spectral clustering;*

W = SimMat;*% W: the weighted adjacency matrix.*

L = D-W;*% L: numb-by-numb unnormalized graph Laplacian matrix.*

Lsym = D$\hat{}$(-0.5)*L*Dˆ(-0.5);*% Lsym: numb-by-numb normalized graph Laplacians.*

Lnorm = (Lsym - min(Lsym(:)))/(max(Lsym(:)) - min(Lsym(:)));*% Lnorm: numb-by-numb normalized graph Laplacians in unit hypercube.*

[V,E] = eigs(Lnorm,k,'sm');*% Compute the first k numb-by-1 eigenvectors V(:,1),...,V(:,k) and eigenvalues E(1,1),...,E(k,k) of Lnorm.*

lambda = diags(E);*% lambda: the first k eigenvalues*

lambdaD = zeros(k-1,1)*;% lambdaD: the eigengap;*

for i=2:k lambdaD(i) = abs(lambda(i)-lambda(i-1)); end

k = find(lambdaD==max(lambdaD))*;% k: the revised number of reserve zone that maximizes the eigengap.*

for i = 1:numb for j = 1:k

T(i,j) = V(i,j)/(sum(V(i,1:k).^2)^(1/2));% *T: the matrix containing the normalized vectors V(:,1),. . . ,V(:,k) as columns.*

end end

opts = statset('Display','final');

% Cluster the points T(i,:)_i = 1,...,numb in R^k into k clusters with the k-means algorithm, returning an numb-by-1 vector idx containing the cluster indices of each point clusters and the k cluster centroid locations in the k-by-k matrix ctrs.

[idx,ctrs] = kmeans(T,k,...'Distance','city',...'Replicates',5,...'Options',opts); end

A.3 Matlab source code for developing Z-ORDCs

% Construct Z-ORDCs using the discrete probability density of discrete levels of system imbalances stored in matrix of [cap,prob] and calculate the parameters [A,b,wA,wb] in the constrained cost variable constraints of the enhance unit commitment model. %Suppose r(i), r(i+1), c(i), c(i+1) as one of the cost segments, then the corresponding constraint on r and Y is,

$$Y \geq c(i) + m * (r - r(i)), \quad m = \frac{c(i-1) - c(i)}{r(i-1) - r(i)},$$

this becomes $m * r - Y <= m * r(i) - c(i)$*; this is the constrained variable technique.*

function[A,b] = CalORDC(cap,prob)

global K *%number of segments of linearized operating reserve demand curve*

rumax = max(-min(cap),0);% *the maximum amount of upward reserve*

A = zeros(K,1); b = zeros(K,1); deltar = rumax/K;

if deltar>0 r = 0:deltar:rumax; id = find(cap+r(1)<=0);

Y(1) = sum((-cap(id)-r(1)*ones(length(id),1)).*prob(id));

if Y(1)>0 for i=1:K ns = i+1; id = find(cap+r(ns)<=0);

if isempty(id) break;

else Y(ns) = sum((-cap(id)-r(ns)*ones(length(id),1)).*prob(id));

m = diff(Y(ns-1:ns))./ deltar;% *slopes*

b = - m * r(ns-1) + Y(ns-1); b(i) = b; A(i) = m;

end end end

A.4 List of publications

1 Meng Xu, Amy Wilson, Chris Dent, Calibration and sensitivity analysis of long-term generation investment models using Bayesian emulation, *Sustainable Energy, Grids and Networks*, vol. 5, pp. 58–69, 2016.

2 Meng Xu, Amy Wilson, Chris Dent, Uncertainty quantification in power system reliability using a Bayesian framework, *Probabilistic Methods Applied to Power Systems (PMAPS), 2016 IEEE 14th International Conference on*, 2016, (accepted).

3 Meng Xu, Amy Wilson, Chris Dent, Zonal operating reserve demand curve applied to day-ahead deterministic unit commitment, *Probabilistic Methods Applied to Power Systems (PMAPS), 2016 IEEE 14th International Conference on*, 2016, (accepted).