Modelling and Optimisation of Integrated Urban Energy Systems for both Heating and Power

Si Chen

Submitted in fulfilment of the requirements for the
Degree of Doctor of Philosophy

James Watt School of Engineering
College of Science and Engineering
University of Glasgow

October 2021

Abstract

Taking into account the rapid increase of renewable energy power generation in the UK, the electrified heating represents an attractive solution for decarbonisation of heat in the long term. However, this will significantly increase the peak power demand in winter and bring further challenges to the grid. Therefore, this work aims to model and optimise a district-level multi-vector integrated energy system for both heating and power through technical and market analysis of using a variety of local renewable energy resources for electricity and heat.

In this thesis, the integrated urban energy system is modelled and optimised in multi processes. As a target system, the heating and electricity demand of the University of Glasgow is used as a case study. In order to model the heating and electricity demand under different weather profiles, the heat demand of the buildings is modelled in an engineering model and a statistical model respectively to predict the hourly heat demand according to weather conditions; while the electricity demand is modelled considering both the building baseload and occupancy rate.

In heat demand modelling, in order to distinguish the heat demand of each building from the data of whole campus provided by the Energy Center when the detailed building parameters are unknown, this work uses a bottom-up building energy model, which uses physical process of heat transfer to simulate the space heating of buildings, and proposes a Bayesian-based calibration method to calibrate the building parameters in the model. The results show that the Bayesian approach-based emulator performs better with fewer calibration times to find the optimal point, which is relable and efficient to calibrate the thermal parameters in building energy models.

Due to the complexity of building a bottom-up building energy model, it is not easy to expand the model to larger areas or add more building samples in the model. Therefore, this work also builds a more general statistical model that can predict the heat demand of different types of buildings simply by giving weather conditions and building characteristics. This work uses artificial neural networks (ANN) technology to simulate the nonlinear relationship between weather conditions, building characteristics and heat demand. In order to improve the training efficiency of ANN, a new sensitivity analysis method is proposed to analyse the correlation between input variables and detect and remove the variables with low importance and the variables that have high importance but contain duplicated features. The result shows the proposed method can re-

duce training time by around 20% while achieving the same training and prediction performance compared with the traditional sensitivity analysis method.

In the electricity demand modelling, the impact of randomness of occupants' activity on power demand forecasting for buildings has been a difficult problem. In order to solve this problem, this work proposes two approaches for fitting and predicting the electricity demand of office buildings by splitting the time horizon for different occupancy rates. The first proposed approach splits the electricity demand data into fixed time periods and using linear regression approach to fit the building baseload and occupancy rate. The second proposed approach uses the ANN and fuzzy logic techniques to fit the building baseload, peak load, and occupancy rate with multi-variables of weather variables. The result shows that the proposed methods reduce the prediction error of electricity demand by 30% and 55% compared with the conventional ANN approach.

To study the impact of electrified heating on buildings and the grid, an Integrated Energy Network (IEN) is established that includes the heat and electric demands of buildings, as well as the generation of local renewable resources and energy storage techniques. In order to rationally plan this new type of IEN based on electric heat pump (HP), this work studies and develops a particle swarm optimisation (PSO) algorithm-based optimisation size method to maximize the decarbonisation on building heating under limited investment cost. According to different source of electric driven, the IEN can be designed as a grid powered HP based heating system and a grid independent renewable heating system (RHS). For the grid powered IEN, this work formulates an operating scheme based on different electricity tariffs to reduce the operational cost of grid power. For the grid independent RHS, this work uses the PSO algorithm to optimise the size of local renewable resources, heat pumps and storage equipment based on the annual investment cost to minimise the total CO_2 emission and reduce the operational cost of natural gas. This work provides a feasible solution for how to invest in RHS to replace the existing gas boiler/CHP based heating system.

In summary, the significance of this study is to use of local renewable energy sources in electric heating taking into account the local weather conditions and the demand of heat and electricity to reduce carbon emissions in heating and electricity supply.

Abbreviation List

AI	Artificial Intelligent
AIC	Annual Investment Cost
ANN	Artificial Neural Network
ASHP	Air Source Heat Pump
BHM	Bayesian History Matching
BMS	Building Management System
BP	Back-propagation
BSS	Battery Storage System
CCS	Carbon Capture and Storage
CHP	Combined Heat and Power
CHP	Combined Heat and Power
CO2	Carbon dioxide
COP	Coefficient Of Performance
DG	Distributed Generation
DHN	District Heating Network
E7	Economy tariff with 7 hours off-peak duration
E10	Economy tariff with 10 hours off-peak duration
EOC	Electric Operational Cost
ESS	Energy Storage System
EU	European Union
GA	Genetic Algorithm
GHG	Greenhouse Gas
GHP	Geothermal Heat Pump
GP	Gaussian Process
HAWT	Horizontal Axis Wind Turbine
HP	Heat Pump
HVAC	Heating, Ventilation, and Air-Conditioning
IEN	Integrated Energy Network
IES-VE	Integrated Environmental Solutions Virtual Environment
LIB	Lithium-Ion Batteries

MAE	Mean Absolute Error
MCMC	Markov Chain Monte Carlo
MSE	Mean Square Error
PDF	Probability Density Function
PSO	Particle Swarm Optimisation
PV	Photovoltaic
QLSR	Quadratic Least Square Regression
RES	Renewable Energy Sources
RHS	Renewable Heating System
RMS	Root Mean Square
SA	Sensitivity Analysis
SARF	Sensitivity Analysis with Reduced Features
SOC	State of Charge
STD	Standard Deviation
SVM	Support Vector Machines
TES	Thermal Energy Storage
UN	United Nations
VAWT	Vertical Axis Wind Turbine
VRB	Vanadium Redox-flow Batteries
WHO	World Health Organization
WMO	World Meteorological Organization

Contents

List of Tables

List of Figures

Acknowledgements

I would like to give my heartfelt thanks to my supervisor, Prof. Zhibin Yu, whose encouragement, guidance and support enabled me to develop a deep understanding of my work. Without his consistent and illuminating instruction, my research work could not proceed to this stage. The research skill, writing skill and presenting skill he taught me will benefit me throughout my life, as well as his inspiring insight into life philosophy and poetry.

I would like to show my gratitude to my associate supervisor, Dr. Daniel Friedrich, from University of Edinburgh for his very kind guidance with his knowledge of mathematical modelling and thermal energy technology. I would like to thank Gillian Brown for providing valuable information on the campus energy consumption.

I offer my regards and blessings to all of the members of James Watt School of Engineering. My thanks also go to Energy Technology Partnership (ETP) and SP Distribution PLC (Scottish Power) for their joint funding of this project [grant number 146].

Special thanks with love to my parents for their unconditional love and support throughout my PhD study and my whole life financially and emotionally. I am also deeply grateful to my boyfriend, for his love and support during my postgraduate study.

Declaration

The author hereby declares that this thesis is a record of work carried out in the James Watt School of Engineering at the University of Glasgow during the period from March 2018 to October 2021. With the exception of the chapters that contain introductory material, all work in this thesis was carried out by the author unless otherwise explicitly stated.

Chapter 1

Introduction

1.1 Background

In current decades, the huge consumption of fossil fuels makes the reserved resource greatly been reduced. Meanwhile, the global climate changes became extreme abnormal. Mitigation of climate change has become an increasingly important topic. According to the United Nations (UN) forecast, the global population will increase to 9.8 billion by 2050 and reach 11.2 billion by 2100 [1]. The increasing population and the continuous use of non-renewable resources will have a serious impact on the climate and the environment [2, 3]. The World Health Organization (WHO) reports that air pollution causes approximately 7 million deaths worldwide every year [4]. In the 2015 Paris Agreement, 197 countries, including the world's three major emitters (China, United States, and European Union), agreed to limit the global average temperature rise to a range below 2°C compared with the pre-industrial level, and worked hard to limit the temperature rise below 1.5°C [5], and climate actions aimed at Carbon Neutral were issued. Carbon Neutral is defined as the reduction of net CO_2 emissions from energy and industrial processes to zero after considering carbon capture and storage (CCS). Net Zero is a similar concept to Carbon Neutral but goes beyound just carbon but all greenhouse gases being emitted into the atmosphere. It indicates that human activities no longer cause global warming. According to forecasts, on a global scale, net zero CO_2 emissions are earlier than net zero greenhouse gas (GHG) emissions, that is, 2°C will be from 2065 to 2080 and 1.5°C will be from 2045 to 2060 [6]. In most cost-optimal scenarios that limit global warming to 2°C compared to pre-industrial levels, there is at least a 66% probability that net greenhouse gas emissions will occur soon after 2085 [6]. The United States and the European Union promise to achieve carbon neutrality by 2050, and China promises to achieve carbon neutrality by 2060. However, the European Climate Foundation stated that although the decarbonisation of European buildings is imminent, the industry is currently not on track to achieve zero greenhouse gas emissions by 2050 and emphasized that current policies are not sufficient to achieve the goal [4].

In 2018, heat demand accounted for half of the world's final energy consumption, 90% of

which came from fossil fuels, which accounted for 40% of global CO_2 emissions [7, 8]. Although industrial heating is the largest consumer, space heating and water heating in buildings account for 46% of the demand [7]. Heating and hot water in buildings account for nearly a quarter of global energy consumption. Heating of residential space is one of the main contributors to GHG emissions, and approximately 90% of the heat comes directly from the combustion of fossil fuels (mainly natural gas). Globally, the type of heating technology installed in the past decade has hardly changed. Fossil fuel equipment and traditional electric resistance heaters accounted for nearly 80% in 2019 [9]. In Europe, 57% of building heating needs are met by natural gas, and another 23% is provided by oil or coal [8]. In many temperate countries, space heating and water heating in buildings account for more than two-thirds of the total heat demand. For example, in the UK, this proportion accounts for 71% [8].

The UK's 2008 Climate Change Act was amended in 2019, stipulating that carbon emissions should be reduced to net zero by 2050 [10]. However, electricity only accounts for 18% of the UK's total energy usage [11]. The space heating and water heating of residential and commercial buildings use 57% more energy than the total electricity generation in the UK each year, and more than 84% of the heat comes from natural gas or oil [8, 12]. In 2017, the UK's annual CO_2 emissions have fallen by 38% since 1990. Emissions from the power sector accounted for 60% of this decline, while the residential sector stagnated. According to the estimates of the Domestic Energy and Carbon Model in the UK, space heating emits about three times as much carbon dioxide as the second source of emissions (lighting and appliances) [13]. Therefore, reducing GHG emissions in the residential sector, especially those related to space heating, is necessary to achieve this net-zero goal.

Since the 1960s, the number of households has almost doubled, but for most of the 1990s and early 2000s, improved heating systems and household thermal efficiency kept annual emissions between 80 and 90 $MtCO_2$/a [14]. In the UK, district heating (DH) accounts for only 0.8% of domestic heating [15]. Nevertheless, between 2013 and 2018, the total number of heat network customers more than doubled, from 211,000 to nearly 500,000 [14]. With the phasing out of inefficient indoor heaters, gas-boiler central heating has become the main method of space heating in the UK. The proportion of residential heating with gas boilers increased from 80% in 1996 to 92% in 2017 [16]. Although the efficiency of gas boilers reduces the use of fossil fuels, they are still the main heating technology in the UK [8]. It is estimated that the combustion of gases in residential and commercial buildings contributes to 18% of the UK's GHG emission [8]. With the high penetration of natural gas in a large number of inefficient buildings, a study showed that without major policy intervention, any move away from natural gas will not happen [17]. Therefore, until the cost of carbon emissions is high enough, natural gas will still be a cost-effective way of heating, and the cost of providing infrastructure for new technologies (such as the large-scale electrification of heat pumps) will be prohibitive. For the UK to achieve net zero emissions, significant changes must be made to the way heat is generated.

At present, natural gas is still the main feature of existing and new housing in the UK, with strong historical, technological, and social barriers to change. With the tightening of system-wide decarbonisation targets in 2050, heating in the residential sector is playing an increasingly important role in reducing emissions. In Scotland, the gas demand is more than three times than that of electricity demand [18], as shown in the yearly pattern of energy consumption in Figure 1.1. Moving away from natural gas has become inevitable, and the long-term development trajectory combines end-use electrification at the household or collective level with supply-side decarbonisation. The UK government issued a heat supply strategy framework in 2012, and extensively proposed long-term electrification [14].
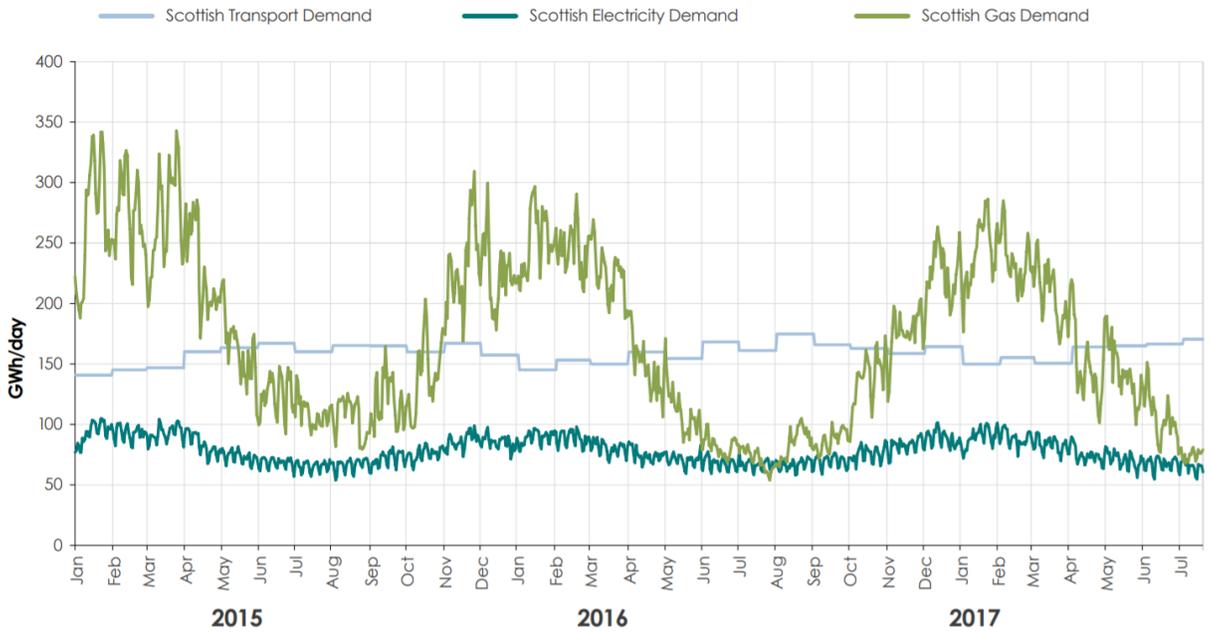
Figure 1.1: Yearly Pattern of Energy Consumption [18].

The full use of electric heat pumps under current conditions in the UK will result in a 25% increase in electricity demand, but a 65% increase in peak demand [19]. One possibility to integrate electricity and heat networks is to use Combined Heat and Power (CHP) units that can produce heat and electricity simultaneously [20]. The CHP units and heat pumps act as the linkage of electricity and heat networks for the district heating systems. Furthermore, heat demand highly depends on seasons and weather condition. The CHP units have strong interdependence between electricity and thermal generation. However, the heat loads are much higher than the electricity demands in winter, especially at off-peak hours. To meet the heating demands, CHP units have to generate large amount of power including both the heat and electricity. The usage of renewable energy is therefore limited as only electricity been generated [21–23]. This occupies the fraction of wind energy generation and causes large curtailment on wind energy [24].

In the past, heating and power sectors were normally tackled separately. The future energy network will include multi-type energy vectors, including renewable energy sources, such as solar, wind, and geothermal, and non-renewable energy sources, such as natural gas. It is much

benefit to consider the multi-vectors in an integrated energy system. The integrated energy systems have been developed rapidly in recent years, and have achieved good performance in energy efficiency improvement, CO_2 emission reduction and renewable energy integration [20, 25–28]. The coupling of heat and electricity represents a typical integrated energy system. The increasing penetration of these kinds of coupling components makes the interaction between the electricity and heat networks more significant. The tight connection of the two networks results in more flexible utilisation of the coupling components (CHP units, electric boilers, heat pumps, circulation pumps, etc.), distributed renewable resources and energy storage systems. Flexible conversion between electrical and thermal energy could reduce the use of traditional heating schemes that burn natural gas for heat, allowing policymakers to see the possibility of carbon-reduced heating. Therefore, a policy called the Renewable Heat Incentive (RHI) has been risen to encourage and promote the technologies of renewable heat sources, such as biomass boilers, solar thermal systems, geothermal systems, etc. [29]. Considering the rapid increase of renewable power generation in the UK, electrifying heating represents an attractive solution for the decarbonising heating sector in the longer term. However, this will bring uncertainties and further challenges to the national grid, in terms of creating extra power generation capacity and balancing the generation and demand. The current energy infrastructure cannot address these challenges.

1.2 Major Contributions

This thesis aims to model and optimise district-level multi-stakeholder, multi-vector integrated energy systems for both heating and power through a technical and market analysis, which utilises multiple sources of local electricity and heat generation from renewables and involves multiple energy users and stakeholders of infrastructure. The first major scientific question that needs to be solved is how to model the heat and electricity demand of buildings, which has nonlinear dependency with weather profile and uncertain occupants' behaviour. The second major scientific question that needs to be solved is how to optimise such integrated urban energy systems and also provide an insight in the opportunity of utilising electricity and heat storage systems to aid the balancing and effectiveness of matching supply and demand with the multiple characteristics on the system.

In order to solve the above two major questions, this work aims to develop modelling approach to estimate the heat and electricity demand of buildings by using the data of University of Glasgow as a case study, and to design a feasible plan to gradually phase out existing gas-fired boilers for building decarbonisation based on the heat demand of the university campus and the distribution of local renewable energy, and to contribute to the realization of the UK's goal of net zero by 2050.

The major objectives of this thesis and main contributions of each chapter has been shown

in Figure 1.2. In order to model the heat and power demands of buildings and find out the most relevant weather variables, the thermal parameter calibration is studied in Chapter 3, the weather dependency is studied in Chapter 4, and the uncertain occupancy rate is studied in Chapter 5. In order to incorporate local renewable resources according to weather conditions in such integrated urban energy systems and optimise the size of each component, the integrated energy system and local renewable resources are modelled in Chapter 6.1 & 6.2, the optimal sizing approach is studied in Chapter 6.3, and the case studies are shown in Chapter 6.4 & 6.5 to consider the reduction of CO_2 emission.
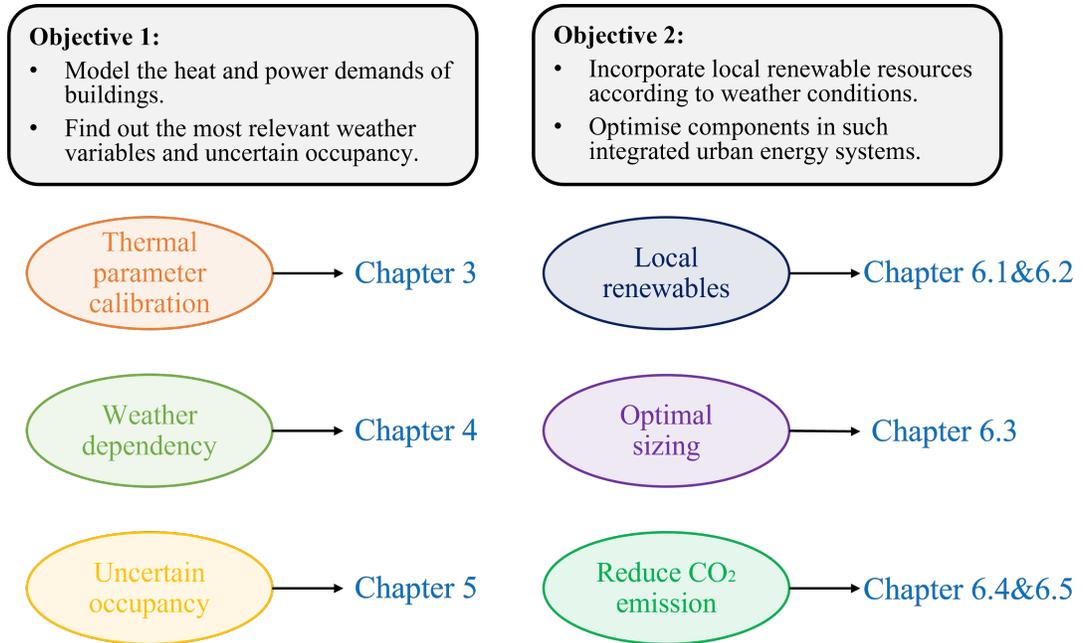
Figure 1.2: Major objectives of each chapter.

The thesis reports the research work undertaken based on modelling and optimisation of integrated urban energy systems for both heating and power. The major contributions of this work can be summarized as follows:

- Detailed thermal parameters are essential for predicting the demand of district heating network using physics-based energy models. But the parameters are not always available or sufficiently accurate. To reduce the simulation time in calibration and the dependency of accurate data of buildings, this work develops a prediction approach using a building energy model whose parameters are calibrated by Bayesian-based calibration method to match the recorded data of monthly heat demand. In the proposed calibration approach, an emulator is established to evaluate the untested values of thermal parameters using Bayesian method, and then use the evaluation results to search for the most suitable parameters value. The designed approach greatly accelerates the calibration speed. The method is used to calibrate a single parameter and multiple parameters of the building

thermal energy models for a district heating network. After it has been verified with mea-
sured data, the developed calibration method is used to calibrate parameters of building
energy models. The output of the calibrated model can predict the hourly building heat
demand in district heating networks.

- Artificial neural network (ANN) has become an important method to model the nonlinear
relationships between weather conditions, building characteristics and its heat demand.
Due to the large amount of training data required for ANN training, data reduction and
feature selection are important to simplify the training. However, in building heat demand
prediction, many weather-related input variables contain duplicated features. This work
develops a sensitivity analysis approach to analyse the correlation between input variables
and to detect the variables that have high importance but contain duplicated features. The
proposed approach is validated in a case study that predicts the heat demand of a dis-
trict heating network containing tens of buildings at a university campus. The proposed
approach can detect and remove several unnecessary input variables and helps the ANN
model to reduce its training time while maintaining the prediction accuracy. It indicates
that the approach can be applied for analysing large number of input variables to help
improving the training efficiency of ANN in district heat demand prediction and other
applications.

- Due to the impact of occupants' activities in buildings, the relationship between elec-
tricity demand and ambient temperature will show different trends in the long-term and
short-term, which show seasonal variation and hourly variation, respectively. This makes
it difficult for conventional data fitting methods to accurately predict the long-term and
short-term power demand of buildings at the same time. In order to solve this problem,
this work proposes two approaches for fitting and predicting the electricity demand of of-
fice buildings. The first proposed approach splits the electricity demand data into fixed
time periods, containing working hours and non-working hours, to reduce the impact of
occupants' activities. After finding the most sensitive weather variable to non-working
hour electricity demand, the building baseload and occupant activities can be predicted
separately. The second proposed approach uses the ANN and fuzzy logic techniques to
fit the building baseload, peak load, and occupancy rate with multi-variables of weather
variables. In this approach, the power demand data is split into a narrower time range
as no-occupancy hours, full-occupancy hours, and fuzzy hours between them, in which
the occupancy rate is varying depending on the time and weather variables. The pro-
posed approaches are verified by the real data from the University of Glasgow as a case
study. Compared with the traditional ANN method, both proposed approaches have less
root-mean-square-error (RMSE) in predicting electricity demand. In addition, the sec-
ond proposed approach can provide more information for building energy management,

including the predicted baseload, peak load, and occupancy rate, without requiring additional building parameters.

- Due to the distributed and seasonal heating demand, the decarbonisation of heating is more challenging, especially for countries that are cold in winters. Electrically powered heat pumps are considered as an attractive solution for decarbonising heating sector. Since grid-powered heat pumps may significantly increase the power demand of the grid, this work considers using local renewable energy to provide power for heat pumps, which is known as the grid independent renewable heating system including photovoltaic, wind turbine, battery storage system and thermal energy storage. This work investigates a complete renewable heating system (RHS) framework and sizing the components to decarbonise building heating. The relationship between the reduction of gas consumption and the requirement of battery storage system under the corresponding installation capacity of renewable components is analysed with their technical requirements. Then, according to different investment plans, this work uses the particle swarm optimisation algorithm for optimal sizing of each component in the RHS to find a solution to minimise CO_2 emissions. The RHS with optimal sizing can minimise CO_2 emissions and reduce the operational cost of natural gas. This work provides a feasible solution of how to invest the RHS to replace the existing heating system based on gas boilers and CHPs.

1.3 Thesis Outline

The thesis is organised as follows.

Chapter 2: Literature Review

This chapter introduce the commonly used approaches in engineering modelling of building heat demand, statistic modelling of heat demand, statistic modelling of electricity demand and integrated heat and power network in literatures. After the review of commonly used approaches, the problems of current approaches and motivation of this research work have been discussed.

Chapter 3: Model Calibration of Building Space Heating

This chapter describes the process of modelling the heating network using a bottom-up building energy model and develops a parameter calibration approach by Bayesian-based calibration method. This aims to solve the problem that the detailed thermal parameters are essential but not always available to be obtained or sufficiently accurate. Section 3.1 describes the modelling of building heat demand in a district considering the physical process of heat transfer. Section 3.2 presents the development process of Bayesian-based model calibration approach. Section 3.3 demonstrates the simulation results of thermal parameter calibration and heat demand prediction of University of Glasgow as a case study. And the chapter is concluded in Section 3.4.

Chapter 4: Statistical Modelling of Heat Demand

This chapter investigates the artificial neural network (ANN) based statistical method to model the nonlinear relationships between weather conditions, building characteristics and its heat demand, and then develops a sensitivity analysis approach to analyse the correlation between input variables and to detect the variables that have high importance but contain duplicated features. Section 4.1 describes the ANN model for building heat demand prediction and data collection for training. Section 4.2 presents the development of feature selection based sensitivity analysis approach. Section 4.3 demonstrates the simulation results of the proposed sensitivity analysis approach and training performance compared with using the traditional approach. And finally the chapter is concluded in Section 4.4.

Chapter 5: Statistical Modelling of Electricity Demand

This chapter investigates the modelling of electricity demand and proposes two approaches for fitting and predicting the electricity demand of office buildings by splitting the time horizon for different occupancy rates. The first proposed approach splits the electricity demand data into fixed time periods and using linear regression approach to fit the building baseload and occupancy rate. The second proposed approach uses the ANN and fuzzy logic techniques to fit the building baseload, peak load, and occupancy rate with multi-variables of weather variables. Section 5.1 describes the data processing and modelling process for electricity demand prediction. Section 5.2 presents the development of two electricity demand prediction approaches. Section 5.3 demonstrates the simulation results of electricity demand prediction using the two proposed approach comparing with the traditional ANN based approach. And finally the chapter is concluded in Section 5.4.

Chapter 6: Integrated Heat and Power Network

This chapter investigates the integration of heat and power network and a heat pump based renewable heating system (RHS), and then develops a PSO algorithm based optimal sizing approach for grid independent RHS for building decarbonisation. This work provides a feasible solution of how to invest the RHS to replace the existing heating system based on gas boilers and CHPs. Section 6.1 describes the traditional gas boiler/CHP based integrated heat and power network in University of Glasgow. Section 6.2 describes the modelling of each components in the renewable heating system. Section 6.3 presents the PSO algorithm based optimal sizing approach for the RHS to reduce the CO_2 emission. Section 6.4 demonstrates the simulation results of operational cost reduction in grid powered heat pump based integrated energy network under different electricity tariffs. Section 6.5 demonstrates the simulation results of minimisation of CO_2 emission and reduction of operational cost of nature gas in the grid independent RHS. And finally the chapter is concluded in Section 6.6.

Chapter 7: Conclusion and Future Works

The thesis has concluded with a summary of the results and several suggestions for future work. The suggestions for future work will highlight the unsolved problems that remained.

1.4   List of Publication

The publications produced from this research work are listed in this section as follows:

1. Si Chen, Daniel Friedrich, Zhibin Yu, and James Yu. District heating network demand prediction using a physics-based energy model with a Bayesian approach for parameter calibration. *Energies*, 12(18):3408, 2019.

2. Si Chen, Yaxing Ren, Daniel Friedrich, Zhibin Yu, and James Yu. Sensitivity analysis to reduce duplicated features in ANN training for district heat demand prediction. *Energy and AI*, 2:100028, 2020.

3. Si Chen, Yaxing Ren, Daniel Friedrich, Zhibin Yu, and James Yu. Prediction of office building electricity demand using artificial neural network by splitting the time horizon for different occupancy rates. *Energy and AI*, page 100093, 2021.

4. Si Chen, Daniel Friedrich, and Zhibin Yu. Optimal sizing of a grid independent renewable heating system for building decarbonisation. *Frontiers in Energy Research*, page 455, 2021.

5. Si Chen, Yaxing Ren, Daniel Friedrich, Zhibin Yu, and James Yu. District heat demand prediction using artificial neural network with data of sample building. In *International Conference on Applied Energy (ICAE 2019). August 12-15, Västerås, Sweden*, 2019.

Chapter 2

Literature Review

2.1 Engineering Modelling of Building Heat Demand

2.1.1 Building energy simulation approaches

The increasing of energy consumption of using fossil fuel has been causing global warming problems. In recent decades, a lot of efforts have been made to reduce greenhouse gas emissions and the consumption of fossil fuel to mitigate the global environmental degradation and warming [30]. Many countries have deployment targets for low carbon technologies such as photovoltaic panels, electric vehicles and heat pumps with the aim to cut CO_2 emissions [31]. In terms of global energy consumption, buildings are becoming an important sector in current and future energy landscapes [32–34]. In cold regions, the demand of heat is much higher than that of electricity [35, 36]. In the building sector, heat demand is much higher than in other sectors and constitutes 40% of the household energy consumption [37, 38]. Accurate prediction of the heat demand of buildings is needed for the optimisation of heating networks, thermal comfort maintenance, and smart control of intelligent HVAC system [39]. The energy performance of a building depends on a great amount of information regarding weather condition, building materials, number of occupants, usage of equipment inside the building, human social activity influences, etc. In order to reduce energy consumption and GHG emissions in every sector of the economy, building energy consumption estimation has become a key approach to achieve the goals. Thermal performance of buildings is related to three main categories: micro-climate of the environment of the building, building physics, and required thermal comfort inside the building [40]. Energy performance of building is complicated since it depends on multiple variables associated to the building characteristics, equipment and systems, weather, occupants, and sociological influences [41]. Since energy consumption is a function of a great amount of information regarding (a) building characteristics, (b) energy systems characteristics, control and maintenance, (c) weather parameters, and (d) occupants' behaviour, among other sociological parameters, forecasting buildings energy consumption is not an easy task [41].

The building energy models used to predict the heat dynamics of buildings can be categorised into engineering approaches, data-driven approaches, and their hybrids [42, 43], as shown in Figure 2.1 [41, 44].

```
┌─────────────┐        ┌──────────┐         ┌──────────────┐
│  Statistical│        │  Hybrid  │         │  Engineering │
└─────────────┘        └──────────┘         └──────────────┘
      │                                            │
 ┌────┴────┐                              ┌─────────┴─────────┐
┌──────────┐  ┌───────────┐         ┌──────────┐    ┌────────────┐
│Regression│  │Intelligent│         │ Forward  │    │ Calibrated │
└──────────┘  └───────────┘         └──────────┘    └────────────┘
                   │
     ┌─────────────┼─────────────────┐
┌──────────────┐ ┌───────────────┐ ┌────────────────────────┐
│Genetic Algorithm│ Neural Networks│ │Support Vectors Machine │
└──────────────┘ └───────────────┘ └────────────────────────┘
```
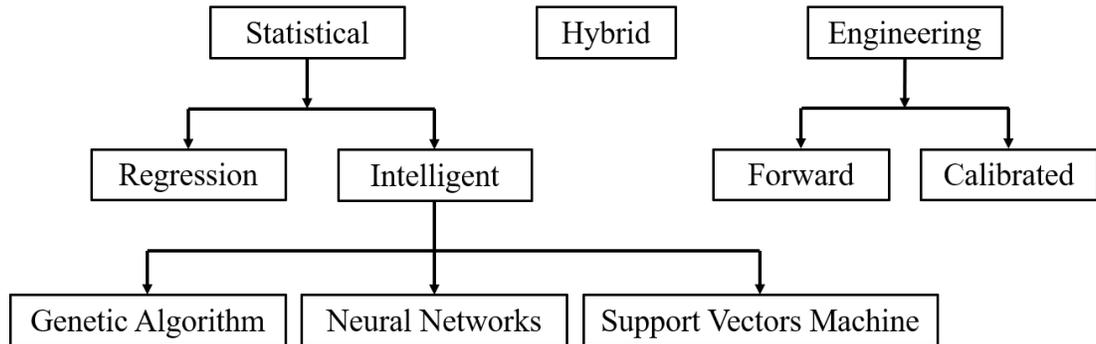
Figure 2.1: Classification of energy modelling approaches [41, 44].

- Statistical methods: These methods aim to use statistic approaches to analyse the historical data of energy consumption by users. Both regression methods and intelligent methods can be used to analyse the historical data. The regression methods can be either linear or multivariate regression to generate the internal relationship from a large amount of measured data. The intelligent methods, which include neural networks, machine learning, and other artificial intelligence approaches, to solve the nonlinear internal relationship even with small quantities of data for training.

- Engineering methods: These methods are developed with building energy simulators, which use bottom-up models to simulate the energy behaviour of a building. Based on the physical principles of heat transfer and engineering expertise, the bottom-up models can simulate the real thermal behaviour of a building according to the given ambient weather conditions, user behaviour and related building parameters including the heating area, thermal coefficients of walls, roof and windows, and solar heat gain coefficient. The thermal performance parameters depend on the thermal properties of various building material and the configurations of the building envelope [45]. Therefore, the building energy model is reliable to predict the heat demand only when the building thermal parameters of the model is well calibrated [46].

- Hybrid models or "grey box": These methods are the combination of the statistical and engineering models to overcome the limitation of physical approaches that need to know detailed building and weather information, and the limitation of statistical approaches that need to be trained using a large amount of measurement data with uncertainties. The hybrid models are developed for the cases of partially known building information and lack of recorded historical data for statistical analysis and training [41].

2.1.2   Model calibration methodologies

Many calibration methodologies have been proposed depending on different analytical methods and strategies [47–50]. Analytical calibration methods are widely used to adjust input parameters until simulation outputs match actual measurements [51]. In these studies, most calibration techniques are manual while only a few are automated. When the complexity of the building model increases, the reliability of the calibration method decreases [52]. In addition, these analytical calibration methods also require significant amounts of data from sensors, such as internal temperature or consumed energy in the target building [53]. On the other side, the input parameters have complicated internal interaction between each other since there are a large number of input parameters [52]. As a result, the analytical calibration method is unable to get available evidence to determine each parameter. Due to this issue, some sensitivity analysis approaches have been introduced to reduce the number of parameters to be adjusted in analytical calibration methods [54, 55]. Alternatively, some studies have investigated simplified models with lumped parameters to reduce computational cost [56, 57]. In many areas, such as in large parts of Scotland, a large number of historical buildings have been maintained and modified several times since they were built. It is a big challenge to measure the exact parameters of these historical buildings or to install sensors to obtain detailed data [58, 59]. Besides, district heating networks (DHNs) usually contain tens to hundreds of different buildings, which increases the difficulty of collecting detailed data of each building for calibration.

2.1.3   Scientific questions

If the measured data is not accurate or the data is affected by uncertain factors, the traditional method could cause overfitting issue to the recorded data. The uncertainty analysis attracts particular research interests as the building energy models represent a simplification of reality. Mustafaraj et al. proposed a calibration method to ensure accuracy and to reduce the likelihood of errors for building energy efficiency simulation [60]. Heo et al. proposed a general theoretical approach to the uncertainty analysis in calibrating building energy simulation models [30]. Yang et al. devised a mesh generation strategy to calibrate the model parameters under high uncertainty by comparing the simulated indoor air temperatures to the experimental measurements [61]. Another possible approach is Bayesian calibration, which performs well in finding the posterior distribution of uncertain building parameters with noisy sensor data sets [43]. Yuan et al. used a Bayesian calibration approach to quantify the uncertainties associated with model parameter for representation improvement [62], and they also proposed a Gaussian Process (GP) based Bayesian method to simultaneously calibrate and rank input parameters of building energy simulation models to enable the efficient use of limited data sources [63]. However, the analysis of uncertainty aims to find the total distribution of the uncertainty statistically. The uncertainties still affect recorded data in time-series. Also, the traditional calibration method runs

the simulation model many times until it matches the accurately recorded data of heat consumption. However, in calibrating the model of a DHN, the simulation time is extremely long as the network includes dozens of building models.

2.2 Statistic Modelling of Heat Demand

2.2.1 Building heat demand

In Europe, the energy consumption of buildings has steadily increased and accounts for nearly 40% of total energy usage [64]. The data shows that heating, ventilation and air-conditioning (HVAC) systems consume the most electrical energy to provide comfort to the working spaces of the building [65]. In most cold areas, such as Scotland, heating accounts for the highest proportion of total building energy consumption and is much higher than other sectors [37] and nature gas is another significant source of heating. Some local communities, such as universities, have their own combined heat and power (CHP) plants and boilers in the district heat supply for dozens of buildings. The building heat demand can be supplied by either the district heat supply, nature gas, or electricity. Load forecasting is essential for energy consumption management and anomaly detection of energy usage in buildings as well as for the integration of variable renewable sources [66]. But at present, forecasting the building electricity power consumption is difficult due to the complexity of the system inside the building and varieties of loads, especially the electric appliances [65]. Therefore, the accurate prediction of building's heat demand, regardless of the supplies of heat demand, becomes the target [67].

The modelling of building heat demand can be classified into three main categories: engineering modelling methods, data-driven methods, and their hybrids [41, 68]. Engineering methods develop bottom-up building models and simulate the heat transfer process based on physical principles [45]. Most building energy simulators use bottom-up building models and heat transfer principles to simulate the energy consumption, e.g. TRNSYS, EnergyPlus and Integrated Environmental Solutions Virtual Environment (IES-VE) [69, 70]. However, building energy simulators require a detailed description of the building to account for the end-use heat demand and have long simulation times, especially for large energy networks. The required number and accuracy of parameters and computational cost are the main drawbacks of these methods. For these reasons, data-driven methods have become popular as an alternative modelling approach to predict the heat demand of buildings.

The data-driven models are developed using statistical methods to fit the input parameters to outputs without any knowledge of their physical relationship, a so called "black-box" model. The input parameters for building energy consumption include both the environmental parameters, such as temperature, solar radiation, humidity, and atmospheric pressure, as well as building design parameters, such as percentage area of windows, thermal properties of walls, and

building orientation and also the occupant behaviour [71]. Apart from these, recent research indicates that the time series data, such as time of day and days of the week, is also important as input variable of data-driven models [33]. That is because the heating time is normally correlated to both the indoor temperature and heating mode, which can be known as a kind of occupant behaviour. Comparing with the conventional regression modelling technique, artificial intelligent (AI) techniques are known to perform more reliable and efficient in many modelling tasks [72–74].

2.2.2   Artificial intelligent methodologies and sensitivity analysis

The commonly used AI techniques include genetic algorithms, support vector machine and artificial neural networks (ANN) [75, 76]. ANN has become one of the most important methods in empirical nonlinear modelling and is widely used to model complex functional relationships between weather conditions and building characteristics as inputs and its heat demand as outputs [71, 77]. The commonly used measurement methods for validating the result of ANN models include mean absolute error (MAE), mean square error (MSE), correlation (R), and coefficient of determination (R^2) [78]. The major advantages of ANN are its very low model construction cost and ability of flexible input-output mapping for complex systems [79]. Whatever complexity of the target system, the ANN is able to use the simplest construction to model its behaviour. This feature of ANN make it widely used to forecast irregular variables, which includes geography information such as wind speed [80] and global solar radiation [81], random building energy usage such as electricity usage [66], cooling load [82] and energy consumption [72, 77], and power generation systems such as photovoltaic [83, 84] and PV/thermal system [78].

Although the construction of ANN has these advantages, its shortcomings are the large amounts of data required for the training, long training time, high risk of overfitting and difficulty of interpreting the knowledge gained by "black-box" models [85]. A common way to reduce these shortcomings is to delete unimportant data components in the training sets to obtain smaller networks, reduced-size data vectors and minimised redundancy in the training data [86, 87]. This can be achieved by analysing the total disturbance of network outputs due to perturbed inputs [87]. Reducing the number of inputs to an ANN model and to select key variables is known as feature selection, which aims at identifying the most relevant input features within a dataset [88–90]. Researchers have analysed different methods of feature selection for ranking and identifying important inputs, such as sensitivity analysis, fuzzy curves, and change of MSE [86].

Sensitivity analysis identifies which input parameters are important for the prediction of the output variable and also quantifies how the changes in the values of the input variables alter the value of the output variable [85, 91]. Several methods have been proposed to explain the contribution of variables in ANN models, including the adaptation of their connection weights [92], a fictitious input matrix considering a successive variation of one input variable while the others

are kept constant [93], the connection weights selected by a randomization approach [94], a perturbation of the input variables [95], the partial derivatives of the output according to the input variables using the connection weights of the ANNs [96, 97]. Typical model sensitivity analyses are "one-at-a-time" simulations that evaluate the impact of each input in turn and ignores the interactions with other input variables [97, 98].

2.2.3   Scientific questions

In predicting the district heat demand with given weather information, it is found that most input variables contain duplicated feature even if they show a high importance in the sensitivity analysis indicating that the training data can be further simplified to reduce the training cost and the risk of overfitting. Therefore, the correlations among inputs also need to be analysed to simplify the training data and remove the duplicated features.

2.3   Statistic Modelling of Electricity Demand

2.3.1   Building electricity demand

The energy consumption of a building is affected by many factors, including weather variables, especially the dry-bulb temperature, the structure of the building and the thermal properties of physical materials used, occupancy and human behaviour, and secondary components, such as lighting system [64]. In the literature, many studies have used engineering methods [32, 99] or statistical methods [33, 100] to model and predict the weather-related thermal energy requirements of buildings. However, due to the influence of occupants' activities, predicting the electricity demand is more complicated than predicting the heat demand. In the literature, different prediction methods are used for the thermal load and electrical load of various buildings. The heat load model is usually based on regression analysis, and the power load model is usually based on the probability distribution of hourly daytime analysis [101]. Statistical models can be used to study the effects of temperature, and time series models can be used to predict daily power demand [102].

The data of building energy consumption shows that the basic electricity consumption of buildings includes emergency lighting, service and safety electricity, which are basically constant throughout the day [103]; the variable electricity consumption of buildings includes heating/cooling loads, household air-conditioning equipment, hot water, and other power consumption [64], which are obviously affected by weather variables [102]. In most buildings, the HVAC (heating, ventilation, and air conditioning) systems consume the most electric energy, which can provide a sense of comfort for the working space of the building [65]. However, the simple understanding of occupancy in current research has led to a huge performance discrepancy between estimated energy consumption and measured energy consumption [67, 104]. Due to the overall

increase in per capita building area, power consumption indicators based on building area are no longer suitable for predicting the energy demand of buildings [103]. In [105], Newsham and Birt especially emphasized the influence of occupancy rate, which obviously can improve the accuracy of the model. In the modelling of occupancy rate, it is difficult to collect information about equipment occupancy and operation and, thus, time indicators are usually selected as input related to the timetable to indicate occupancy and equipment usage [106]. In order to solve the influence of seasonal changes and human activities on the power demand forecast of buildings, some studies have established separate models for different seasons or months to predict the power demand including human activities [107, 108].

2.3.2 Methods in occupancy rate prediction

As building energy behaviour is affected by many factors, this complex situation makes it difficult to accurately predict building energy consumption [64]. In the literature, researchers use data mining techniques to discover and summarise electricity consumption patterns hidden in data [109]. Reference [75] reviews the research on the development of data-driven building energy consumption prediction models, and outlines the data-driven building energy prediction research using machine learning algorithms, including support vector machines (SVM), ANN, decision trees and other statistical algorithms have identified existing research gaps and highlighted future research directions in the field of data-driven building energy consumption prediction [75]. Literature [110] demonstrates a comparative study of four data-driven methods used in online building energy consumption forecasting. These methods involve large-scale data extracted from several types of buildings, compared with other methods with less calculations. ANN has higher accuracy. Literature [109] proposed an effective classification model using SVM as the basic optimisation framework, and verified its feasibility. In the literature [111], the researcher proposed an SVM model that only uses the temperature and occupancy rate of the building as attributes to predict the power load in non-residential buildings, which can achieve accuracy and calculation cost in various situations The best balance [111]. Nizami and Al-Garni [112] tried a simple feedforward neural network that correlated electrical energy consumption with the number of residents and weather data. Literature [113] proposed a framework based on deep learning to predict electricity demand by taking care of long-term historical dependence. Literature [114] proposed and evaluated a novel RNN technology that can predict the energy use of non-residential large buildings consisting of 562 rooms. Literature [106] proposed a deep feedforward neural network architecture determined by genetic algorithm, which is used for the daily hourly and daily power consumption of campus buildings in the real world in the UK. Literature [115] proposed a recurrent neural network model, which can predict the electricity consumption of commercial and residential buildings in the medium and long term (more than one year) at a resolution of one hour [115]. Literature [67] proposed an occupancy estimation method based on blind system identification, and estimated the number of occupants based

on artificial neural network and using BSI, and developed and reported a power consumption forecast model for air conditioning systems [67].

Currently, a shallow understanding of occupancy rates has led to a huge performance gap between estimated and measured energy consumption [67]. Due to the overall increase in per capita building area, electricity consumption indicators based on building area are no longer suitable for predicting energy demand occupied by buildings [103]. In [105], Newsham and Birt particularly emphasized the influence of occupancy rate, which obviously can improve the accuracy of the model. Residential activities will greatly affect building performance, so dynamic modelling should be used to faithfully represent reality [104]. There is no existing method that can simulate the time dependence of random activities based on physical quantities [104]. The most important factors are occupied time and free time, which will greatly affect the use of energy, because the occupants will use heating, electric plugs, office equipment and air-conditioning equipment [110]. As it is difficult to collect information on the occupancy and operation of equipment, time indicators (such as hour of the day, day of the week, month of the year) are usually selected as timetable-related inputs to represent the situation of occupancy and operation of equipment [106]. Time dependence is crucial for the use of dynamic building simulation, because the probability of performing an activity changes significantly with time [104]. Statistical models can be used to study the effects of temperature and time series models can be used to predict the daily load of electricity demand [102]. According to the time-related probability of starting an activity and its corresponding duration distribution, a modelling method based on residential activities is proposed [104]. To cope with this challenge, some studies have divided building electricity consumption into two forms, 'basic' and 'variable', and established a building electricity consumption prediction model based on human behavior, which is divided into two parts [103]. Basic electricity consumption is related to building area, while variable electricity consumption is related to building occupancy rate [103]. Research in [116] aims to study a short-term, real-time energy demand forecasting method to cope with changing loads to effectively operate and manage buildings. In some buildings where complex functions are applied, such as hotels and shopping centers, the randomness of personnel activities is higher, which has a greater impact on the reliability of data and the prediction of building energy use, so it will greatly reduce the prediction model [117]. In order to solve the influence of seasonal changes and man-made activities on the power demand forecast of buildings, some studies have established separate models for different seasons or months to predict the power demand including man-made activities [107–109]. However, using seasonal models to build models separately will increase the complexity of the model and reduce the continuity of the model. Therefore, using seasonal models to predict electricity demand is not the best solution.

2.3.3   Scientific questions

These methods are suitable for long-term forecasting of average daily electricity demand to reduce the impact of personnel activities on forecast accuracy, or for short-term forecasting of electricity demand to reduce seasonal effects. However, these methods cannot help to solve some typical cases in specific area, such as office/education buildings heated by electricity in cold areas. In this type of buildings, the heating load is much higher than the cooling load, and the long-term and short-term correlation between electricity demand and temperature are in opposite direction. That is, in the trend of seasonal power consumption, temperature and power demand are in negative correlation, while in the trend of hourly power consumption, temperature and power demand are in positive correlation. This causes difficulties in data fitting between power demand and weather variables. At present, no method can solve this problem effectively. Therefore, this work aims to seek a simple and reliable method to accurately predict the electricity demand of buildings in seasonal and hourly simultaneously by considering the impact of occupancy rates.

2.4   Integrated Heat and Power Network

2.4.1   Integrated energy network

The rapid consumption of fossil fuels has greatly reduced global resource reserves and has led to global climate abnormalities [118–120]. In order to reduce the risk of climate change caused by global warming, it is very important to reduce energy consumption and greenhouse gas (GHG) emissions [121]. As part of the European Green Agreement, the European Union (EU) has set a goal to reduce its GHG emissions by 55% by 2030 and become the world's first fully climate-neutral continent by 2050 [122]. Therefore, the EU has supported many energy projects in the past few years to enhance the role of renewable energy in the European energy plan [123].To achieve these challenging goals, the energy structure of most countries needs to undergo major changes to reduce dependence on fossil fuels and the associated GHG emissions [124].

In recent years, the integrated energy network (IEN) has developed rapidly in terms of energy efficiency improvement, carbon dioxide emission reduction, and renewable energy integration. Considering the natural gas and electricity power system together, the power flow-based analysis of integrated gas-electricity networks has been developed in [125–129]. An extended energy hub approach has been presented as a general and flexible framework of highly integrated electricity and heating networks [130]. The steady state flow and a case study have been discussed in [131]. The decomposed and integrated analysis methods have been developed in [132] to investigate the performance of combined electricity and heat networks. In addition, to combine the electricity, heat and gas systems together in an integrated distribution network, a simulation model is presented in [133] to perform an integrated analysis of electricity, heat, and gas distri-

bution networks. With the growth of renewable energy sources (RES) and CHP units integrated into the power system, the interaction between the electricity networks and other energy networks, such as gas and heat, become tighter [36, 132]. On the other hand, the usage of RES is significantly curtailed in some urban region that has high heat load and low power load. The flexibility of RES integration can be increased with providing electrical boilers and heat storage tanks in a centralised dispatch model [134, 135] and using pipeline energy storage in a district heating network with CHP dispatch [136, 137].

2.4.2 Electric heating with renewable energies

In countries with cold winters, heating accounts for a large part of energy consumption and is usually heavily dependent on the burning of fossil fuels such as natural gas and coal [138]. The current energy demand in the building sector, such as space heating and domestic hot water, accounts for 40% of the total energy demand in the EU [139, 140]. Nearly half of the total energy consumption in the United Kingdom is used for heating, and this proportion is even higher in Scotland [141]. While it is crucial to decarbonise heating, it is challenging due to the distributed demand and large seasonal variations. In 2017, the GHG emission of UK equals to 460 million tons of CO_2, of which nearly 40% came from natural gas used for heating [142]. In order to reduce GHG to net zero by 2050 as part of the government's carbon plan, the UK has pledged to establish more district heating networks and develop large-scale electrifying heating [143]. Nearly 90% of the overall heat demand in the UK is provided by natural gas boilers and combined heat and power (CHP) [143]. Since renewable energy can only generate electricity while CHP units have a strong interdependence between power generation and thermal power generation, CHPs are widely used to provide domestic hot water and space heating in district heating systems [140]. However, CHP technology is entirely focused on efficiency improvement and cost optimisation and rarely considers carbon minimisation. In the past, the natural gas CHP was a good alternative to coal as it could significantly reduce carbon emissions. However, with the electricity grid carbon intensity dropping below 300 kg CO_2/MWh, natural gas CHP can only reduce carbon emissions in a few edge cases.

Increased use of renewable energy sources will make important contributions, such as wind energy, solar energy, hydropower, tides, waves, geothermal energy, environmental thermal energy, biofuels and municipal waste [144]. Due to the non-dispatchable nature of many renewables, energy storage is also required. In the UK, as the installation of renewable energy increases, the feasibility of fossil fuel cogeneration has gradually declined. It is estimated that by 2035, the carbon intensity of electricity will be lower than that of natural gas [143]. This will encourage thermoelectricity, shifting from natural gas boilers to direct electric heating and electric heat pumps.

Heat pumps (HPs) are considered to be the key technology for decarbonisation in the heating industry [145, 146]. The HP uses a small amount of work energy to convert a low-grade heat

source into a higher-grade heat source. The technical details of HP will not be discussed in this article since a lot of literature has previously discussed it [147–153]. If the heat source is outside air, the system is called an air source heat pump. If the heat source is underground soil or groundwater, it is called a geothermal heat pump. Geothermal heat pump is one of the fastest growing renewable energy sources in the world [154]. In the long-term, electrifying heating based on HPs is an attractive solution for the decarbonised heating sector. However, this will significantly increase the peak power demand in winter and will bring further challenges to the national grid in terms of creating additional power generation capacity and balancing power generation and demand. In addition, if electric power is generated from fossil fuels, the decarbonisation effect will not be fully realised. Therefore, this work considers using local renewable energy, which is grid-independent, to power the heat pump for building heating and reduce the usage of traditional gas boiler for decarbonisation.

The development of IEN is to integrate electricity, heat and gas systems into an integrated power distribution network. However, the further integration of heating and power networks has also expanded the opportunities for demand-side management to integrate more variable renewable energy generation into the energy system. Due to the rapid growth, the renewable power generation has been extensively studied in the past decades [155, 156]. In particular, solar energy and wind energy are used as sources to supply power to the grid [157–159]. However, the main disadvantage of most renewable energy sources is their intermittent nature, as they fluctuate on a daily, weekly, and seasonal basis. In order to filter these changes, battery energy storage systems have been widely accepted as one of the potential solutions to shift the electrical load from peak hours to off-peak hours [160, 161]. The idea of combining power generation cycle and HP cycle has been extensively studied, such as solar collector - HP [148], PV - HP [162], HP - thermal energy storage (TES) [141], organic Rankine cycle (ORC) - HP water heaters [138].In the research on the optimal sizing of renewable energy systems, many papers have considered WT, PV, ESS and fuel cells [163, 164].

2.4.3 Optimisation of multi-sector energy network

After the heat demand can be predicted accurately in advance, the building energy consumption can be optimised for a higher energy efficiency to achieve better energy performance. The operational planning and optimisation of the integrated electricity and gas networks have been presented in [129, 165, 166]. To investigate the integrated electricity and heat networks, the optimal power flow model of heat and electricity networks has been introduced in [167–169, 169]. Based on an integrated dispatch model of electricity and heat networks, a generalised approach has been presented for network capacity and sitting of CHP-based Distributed generations (DGs) in [170], and an optimal allocation has been analysed in [171]. To optimise benefits for the daily operation of electricity and gas networks considering the impact of RES, a coordinated scheduling strategy has been presented in [172] to extend the synergies of the two networks, and a

two-stage optimisation algorithm has been proposed in [173] to solve the corresponding mixed integer nonlinear programming problem. In addition, due to the response time constant of the heating system is normally much longer than that of the electricity system, the thermal inertia can be used as a buffer to optimise the operation of heat and electricity dispatch. The thermal inertia of a building has been considered in a new district heating system model for integrated heat and electricity dispatch [174, 175] and the integrated model with embedded thermal inertia has been used to maximise the usage of RES [23]. A network flow optimisation model is developed for addressing the economic and physical integrity of the national electric power system focusing on transmission level [166].

From the literature in the past years, four main optimisation approaches have been considered including the direct search, calculus-based optimisation, genetic algorithm (GA), and particle swarm optimisation (PSO).

- Direct Search Direct search is a straightforward optimisation method that does not consider time delay and derivatives. Thus, the direct search method can be applied in optimising many nonlinear functions that have less dependence on its derivatives [176]. Because of which, the direct search method can be used in minimising the thermal discomfort as well as heating loads of buildings.

- Calculus-Based Optimisation The most commonly used calculus-based optimisation method is to set the gradient of the objective function to zero [40]. Another calculus-based optimisation method is Newton's method, which is similar to the Steepest Descent Method that uses an iterative process from an initial guessed starting point to finally converge to the optimum point.

- Genetic Algorithm (GA) The genetic algorithm is a population-based algorithm to search for the global-optimum solution. The iterative process of GA will converge to better solutions based on the breeding of the parents with higher performance [40, 177]. However, the GA has shortages that it requires a large amount of non-optimal data as the requirement of global search [40].

- Particle Swarm Optimisation (PSO) PSO method is a widely used optimisation method developed from social life behaviours, such as bird flocking and fish schooling [40]. The PSO method uses a large number of the swarm to search the optimum point globally and share the information among all swarm for the next search step.

2.4.4 Scientific questions

Due to the large differences in local weather conditions, occupancy, energy prices, government subsidies and building types, this optimisation must be done for each component. When different

energy vectors are intertwined in future smart energy systems, optimal sizing of each component is especially important. The previous published papers only considered the perspective of electrical power use, and do not consider both heat demand and electricity generation together as well as the impact of heating demand that increases the electrical energy use under the trend of electrified heating in the future. In addition, most of the literature considers sizing from the aspect of technical requirements and does not incorporate future technology and price trends. Therefore, within the author's knowledge, the past research has not done feasibility analysis and optimisation for different power generation, thermal systems and the investment allocation of each system.

Chapter 3

Model Calibration of Building Space Heating

3.1 Building Heat Demand Model

3.1.1 Building heat loss

Heat transfer in a building is the transport of thermal energy, due to the spatial temperature difference. Building envelope exposed to the external environmental conditions such as temperature (due to solar radiation and air temperature), air velocity (due to convection) and moisture content in air (humidity). Heat and mass transfer would occur through the envelope, generally from the outside environment to the inside environment [178]. Materials porosity plays a significant role in heat transfer [179]. Actual heat loss depends on countless variables including wall surface area, aging (normally oxidation or weathering), composition, condition, and construction technology [46]. Heat loss or gain in a building is down to conduction, convection, and radiation, in building conduction occurs through envelope assemblies, radiant heat transfer is primarily from the sun and convection is usually the result of wind or pressure-driven air movement. There is a widespread perception of the reduced thermal performance of historic structures compared to more recent buildings although researchers have shown this to be untrue [58, 180].

Buildings external elements (walls and roofs) are continuously exposed to the atmosphere. Therefore, the radiative heat flux (q_{rad}) is significant and it can be described by the equation [180]:

$$q_{rad} = A \varepsilon \sigma \left(T_s^4 - T_{air}^4 \right) \tag{3.1}$$

where A is the area of the external element, ε is the surface emissivity, σ is the Stefan-Boltzmann constant, T_s is the wall surface temperature and T_{sky} is the fictitious sky temperature. T_{sky} represents the temperature of an equivalent atmosphere, considering that the atmosphere is not characterised by uniform temperature and it emits only at particular wavelengths. Here, the

Dreyfus assumption is provided $T_{sky} = T_{air}$ [180]. Rewrite the radiative heat flux as

$$q_{rad} = Ah_{rad}(T_s - T_{air}) \tag{3.2}$$

where h_{rad} is the radiative heat transfer coefficient, which can be obtained from (3.1) and (3.2) as [180]

$$h_{rad} = \varepsilon\sigma(T_s + T_{air})(T_s^2 + T_{air}^2) \tag{3.3}$$

Heat flow can be measured as a U-value and is defined as the heat flow through one square meter of a structure when the temperature on either side of the structure differs by one degree Celsius.

Heat loss from a unit surface area of external wall or ceiling given by the following equation:

$$q = U\Delta T \tag{3.4}$$

where U is the total heat transfer coefficient (or the wall conductance).

The following equation gives the wall conductance U for a typical wall:

$$U = \frac{1}{R_i + R_w + R_o} \tag{3.5}$$

where R_i and R_o are the inside and outside air film thermal resistance; R_w is thermal resistance of the wall. R_w takes values according to the structural components of the wall [37].

The higher the U-value, the worse the thermal performance of a building. Besides, the actual U-value of a wall is dependent on several parameters that are difficult to accurately measure including moisture content, physical properties of the brick such as density and composition, mortar proportion, thickness and presence of air cavities [59]. Heat loss occurs through the building fabric envelope in different quantities; generally, the roof is 25 percent, windows 10 percent, 15 percent ground, 15 percent draughts, and 35 percent walls.

3.1.2 Modelling of district heat demand

The district heating network (DHN) at the main campus of the University of Glasgow is used as a case study, which has thirty-four buildings. It is modelled in a building energy simulator to predict its heat demand over the year. For this purpose, the campus map is used to determine the floor area and shape of each building. The buildings whose heat are supplied by the DHN are simulated in the model.

There are several commonly used simulators for this purpose. This chapter chose Integrated Environmental Solutions Virtual Environment (IES-VE) to develop bottom-up models of buildings for thermal design and analysis as well as heating and cooling load simulation. IES-VE is an integrated system to simulate the building thermal behaviour using a proprietary engine called

Apache. Compared with other building energy simulators, the strength of IES-VE is to correctly express the relationship between thermal zones and spaces defined in the building information bottom-up models [181].

The first step is to draw the building envelope either in the inbuilt toolbox or in an external building bottom-up modelling software. We established a building model of all buildings in SketchUp based on the campus map [182], as shown in Figure 3.1, and then imported the bottom-up models into IES-VE for simulation of heating energy consumption. Apart from the floor area information given from the 2D campus map, the 3D information, such as the height of each building, shape of the roof, windows area, are measured from the 3D satellite image of Google global. The comparison between the satellite image and the completed building model of 34 buildings is shown in Figure 3.2(a) and 3.2(b). Then the completed model of buildings is imported into IES-VE, which can automatically detect the external walls, roofs, internal walls, and windows of each building from the bottom-up model.

Figure 3.1: Campus map of University of Glasgow [182].

There are many free parameters to be set in the IESVE simulation software. The most effective parameters are the building material and property parameters, which include the heat transfer coefficient (also called U-value) of external wall, windows, ground floor, roof, internal ceiling/carpeted floor and internal partitions. The building envelope of the bottom-up model separates the environment inside the building from the outside of the building and provides heat

(a)

(b)

Figure 3.2: (a) The 3D satellite image obtained from Google global; (b) bottom-up model of buildings in a heating network built in SketchUp.

transfer resistance. The U-value is calculated in the software based on the basic building parameters, such as the thickness and materials. The external walls and roofs are the main surfaces for heat transfer between inside and outside environment, and there are many different types of materials in the structure [178, 183]. The most common applied materials in construction are brick and blockwork, concretes, insulating materials, and plasters, as shown in Figure 3.3. In the exampled construction of the external wall, the better thermal resistance of the material causes faster temperature decreasing across layers.

In contrast to the walls that have a low heat transfer coefficient, the windows are transparent and have a higher heat transfer coefficient. Therefore, the size of windows is a significant parameter affecting the average U-values of the building envelope as well as the solar radiative coefficient in building heat demand estimation [180]. The window area can be easily measured

Condensation Prediction for External Wall
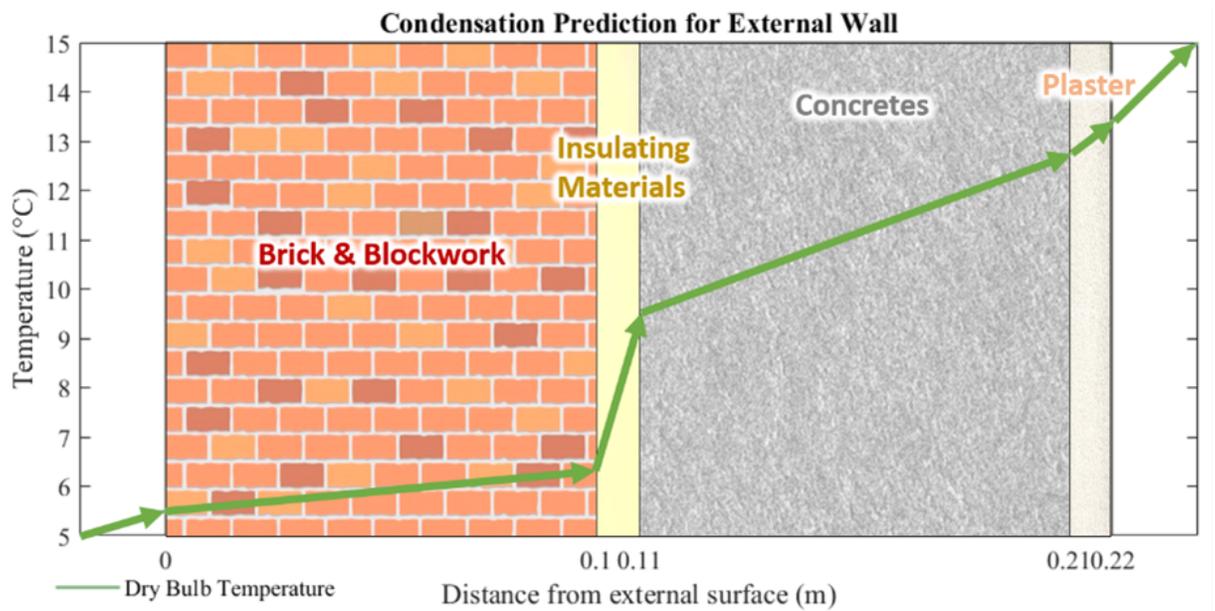
Figure 3.3: Temperature and condensation estimation for external building wall in IES.

for each building and is normally not a parameter that will be changed over time. For example, after hundreds of years, the thermal resistances of historical buildings could have significantly changed due to retrofitting, but their basic building information, such as floor area, building height and windows area, normally remains unchanged. The window area is a significant parameter affecting the building's thermal behaviour, but it does not need to be calibrated like other parameters, such as the U-value of the ground floor, roof, internal ceiling/carpeted floor and internal partitions.

In the University of Glasgow, some buildings have more than hundreds of years of age. It is difficult to find out the accurate measurement of its material and thickness at each part of its external wall. Thus, the initial parameter setting is based on the approximate material and thickness of the external wall of the main buildings. It must be pointed out that each building in the university campus has different building material and using age. However, it is not available to record the heat demand of each building. Thus, to predict the district level building energy model of the whole campus, the lumped parameters considering all the buildings are used in the model.

In addition to the building envelope parameters, the thermal gain also affects the building heat demand, including the thermal template of the HVAC system, the internal gain based on the number of people per unit area, the domestic hot water usage and the heating equipment operating hours. However, they are normally unpredictable and are also challenging to be accurately measured, especially in old buildings. In the IES-VE simulations, the parameters of these uncertain thermal gains were set as a constant ratio of the total heat demand of each building. The building management department provides the setting value from the documentation of plan. There are extra thermal gains inside the building as heat generated from people, computers, and

other heating equipment. In addition, the building operating time also affects the total heat demand. For example, libraries are heated 24 h every day while office buildings usually are heated from 8 am to 6 pm. These parameters are variable depending on human activities and thus are difficult to predict accurately. The values set in IES-VE are chosen from the average usage of building template.

After defining the parameters, IES-VE will calculate the heat transfer coefficient of each part of the buildings from the above information and calculate the heat demand of every building with the imported weather profile and operating mode. The outdoor environment, especially the temperature, is the main factor that affects the building energy demand, as shown in Figure 3.4. In IES-VE model, the impact of weather is calculated together with the building heat transfer coefficient to get the heat demand as an output. The weather profile is obtained from the recorded database from the nearest weather station in Glasgow International Airport, including solar position, hourly temperature, humidity, and wind speed and direction. Since the campus is just around 7 miles away from the airport, the difference between the real weather condition around the heating network and the recorded weather profile from the airport is negligible. But in practice, the differences for wind speed and direction between airport and campus can be significant due to the altitude, topography, occlusion by surrounding buildings and trees, etc. The wind speed can affect the ventilation of target buildings and thus the final estimated heat demand. This is a limitation of using this method and is difficult to address due to the lack of more detailed weather data. Thus, the difference in weather profiles between the airport and the campus are neglacted in this thesis.
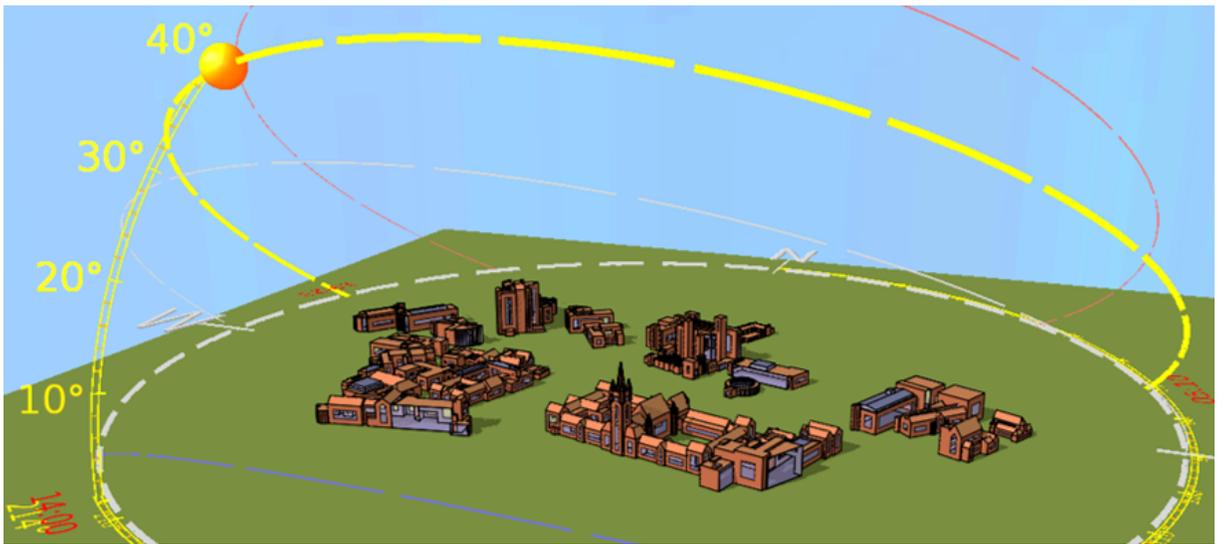
Figure 3.4: Import weather profile and suncast in the building model.

In the IES-VE simulation, the variation of the building envelope parameters, such as wall and glass thickness as well as their material and structure, will significantly affect the thermal behaviour of the building model and subsequently the heat demand. Therefore, the heat transfers related building parameters need to be carefully calibrated to force the model to match the

thermal behaviour of the real buildings.

The ideal approach is to calibrate the thermal parameters of each building according to recorded data. However, this will require a large number of sensors to be installed in each building. It is not efficient for calibrating the models of multi-buildings due to extremely high costs. An alternative method is to use lumped parameters for all buildings, and only one sensor is required to record the heat demand of the whole heating network. While the usage of lumped parameters could reduce the accuracy of the heat demand of a single building, in a DHN with tens to hundreds of buildings, the accuracy of a single building will not have a significant impact on the overall heat demand of the network. What is more, the complexity and computational load in model calibration will be significantly reduced in calibrating the lumped parameters for all the buildings. Also, the average relative heat loss of district heating network is variable within the range of 9% to 30% [184–186] depending on its insulation, pipe dimension, supply and return temperature, and geographical distribution [187]. Comparing with the heat demand of the whole network, heat loss on pipes is relatively low. Therefore, the heat loss in pipes is regarded as an uncertain disturbance and not considered in detail in the thesis.

3.2 Development of Model Calibration Approach

3.2.1 Bayesian methodology

The model calibration aims to find reliable values of a set of model parameters from historical data to adjust the model to match the real data. Consider the case that the computer model is fast to run with a known function y_m as model output. The modelling assumptions are formalised with the statement of '*reality = model + bias*' [188]:

$$y_r(x) = y_m(x, u) + b(x) \tag{3.6}$$

where x is the input; u is a tuning parameter; $y_r(x)$ is the value of the real process at input; $b(x)$ is the unknown bias function. Then the field data $y_f(x)$ is represented as:

$$y_f(x) = y_r(x) + \varepsilon_f = y_m(x, u) + b(x) + \varepsilon_f \tag{3.7}$$

where ε_f is independent normal random error. In the Bayesian approach, all unknown variables are presented by random values. The independent error can be presented as a normal distribution with mean zero and variance σ_f^2 as:

$$\varepsilon_f \sim \mathcal{N}\left(0, \sigma_f^2\right) \tag{3.8}$$

A multivariate normal density has been produced for collecting all field data, y_f. The field data can be denoted by $f(y_f | u, \sigma_f, b)$, while the elements u, σ_f, b are unknown from the model,

and their prior distribution will be presented by $p(u, \sigma_f, b)$. The Bayesian theorem yields the posterior density $p(u, \sigma_f, b|y_f)$ of the unknowns with the given field data as shown below [188]:

$$p\left(u, \sigma_f, b \mid y_f\right) \propto f\left(y_f \mid u, \sigma_f, b\right) p\left(u, \sigma_f, b\right) \tag{3.9}$$

To compute their posterior density, the normalizing constant needs to be determined by integrating the right-hand side in the previous equation. However, in practice, it is challenging to calculate all the parameters to get the posterior distribution [189]. The posterior density is typically worked out by Markov chain Monte Carlo (MCMC) analysis [190]. Thus, the result will be N draws from the posterior distribution to describe u, σ_f, and b, the unknowns. These unknowns can then be sampled using u_i, σ_{fi}, and b_i ($i = 1, \ldots, n$), then the posterior distribution can be estimated from these samples.

3.2.2 Emulator design in general Gaussian process model

An emulator represents the belief about the behaviour of an unknown function [186]. The emulator is designed using Bayesian approach, which is used to evaluate the untested points for providing evidence in searching for the most probable point. In the calibration process of building heat demand prediction model, the emulator is designed to represent the belief of estimation error with plenty of parameters to determine the characteristics of its expected distribution. The main function of the emulator is to calculate the conditional mean and conditional variance of simulators output at an untested input point [191]. The emulator could also be called surrogate model or metamodel in literature.

Assuming $g(x)$ is a general Gaussian process (GP) written as:

$$g(x) \sim \mathcal{N}\left(\beta, \sigma^2 c\left(x, x'\right)\right) \tag{3.10}$$

where β is the parameter for moving the mean; σ^2 is the parameter for the scaling the variance; x and x' are the dimensional vectors of different inputs written as $[x_1, x_2, \ldots, x_p]$, where p indicates the number of dimension of the inputs; $c(x, x')$ is the correlation function of x and x' to describe how they are related, which plays a fundamental role in GPs [192]. It determines the relativity of $g(x)$ and $g(x')$ depending on the distance between x and x'. The common correlation functions are known as Linear, Exponential, Gaussian, and Matern functions. The Gaussian correlation function can be represented as:

$$c(x, x') = \exp\left(-W^2\right) \tag{3.11}$$

where W denotes the weighted distance between x and x' and defined as:

$$W = \text{sqrt}\left[\sum_{i=1}^{p}\left(\frac{x_i - x_i'}{\delta_i}\right)^2\right] \tag{3.12}$$

where δ_i are the correlation parameters to weigh the distance between x and x' in different dimensions. Assuming $g(D)$ is the observation of $n (n > 2)$ tested points to train the emulator, its likelihood of a Gaussian distribution [191]:

$$
\begin{aligned}
p\left(g(D)\,|\beta,\sigma^2,\delta\right) &= \mathcal{N}\left(H\beta,\sigma^2 A\right) \\
&= \frac{|A|^{-1/2}}{(2\pi\sigma^2)^{n/2}}\exp\left[-\frac{1}{2\sigma^2}\left(g(D)-H\beta\right)^T A^{-1}\left(g(D)-H\beta\right)\right]
\end{aligned}
\tag{3.13}
$$

where H is a $n \times 1$ column vector with n unit values; A is the covariance matrix of $g(D)$ defined as:

$$
A = \begin{bmatrix}
1 & c(x_1,x_2) & \cdots & c(x_1,x_n) \\
c(x_2,x_1) & 1 & \cdots & c(x_2,x_n) \\
\vdots & \vdots & \ddots & \vdots \\
c(x_n,x_1) & c(x_n,x_2) & \cdots & 1
\end{bmatrix}
\tag{3.14}
$$

Assuming a prior distribution $p\left(\beta,\sigma^2\right) \propto \sigma^2$ to marginalise of σ^2 and β, and then (3.13) can be approximated to:

$$
p(g(D)|\delta) \propto \frac{|A|^{-1/2}\left|H^T A^{-1}H\right|^{-1/2}}{(\sigma^2)^{n/2}}
\tag{3.15}
$$

The correlation parameters can be determined by maximising the equation of $p(g(D)|\delta)$ as:

$$
\hat{\delta} = \arg\max_{\delta}\left[p(g(D)|\delta)\right]
\tag{3.16}
$$

Finally, the posterior distribution for the emulator can be obtained as:

$$
p\left(g(x)\,\middle|\,g(D),\hat{\delta}\right) = \mathcal{N}\left(\mathrm{E}^*\left[g(x)\right],\mathrm{Var}^*\left[g(x)\right]\right)
\tag{3.17}
$$

with its mean and variance as:

$$
\mathrm{E}^*\left[g(x)\right] = \hat{\beta} - c^T(x)A^{-1}\left(g(D)-H\hat{\beta}\right)
\tag{3.18}
$$

$$
\mathrm{Var}^*\left[g(x)\right] = \frac{\hat{\sigma}^2}{n-2}c_1\left(x,x'\right)
\tag{3.19}
$$

where

$$
\hat{\beta} = \left(H^T A^{-1}H\right)^{-1}H^T A^{-1}g(D)
\tag{3.20}
$$

$$
\hat{\sigma}^2 = g(D)^T\left[A^{-1}-A^{-1}H\left(H^T A^{-1}H\right)^{-1}H^T A^{-1}\right]g(D)
\tag{3.21}
$$

$$
c(x) = \left[c(x,x_1),c(x,x_2),\ldots,c(x,x_n)\right]^T
\tag{3.22}
$$

$$
c_1(x,x') = 1 - c^T(x)A^{-1}c(x) + \left(1-c^T(x)A^{-1}H\right)\left(H^T A^{-1}H\right)^{-1}\left(1-c^T(x)A^{-1}H\right)^T
\tag{3.23}
$$

The function of the emulator is to estimate the expected result at untested points without simulating them, which can rapidly provide evidence via searching for the most expected value of parameters. The building simulator is used to calculate the energy consumption of buildings with given ambient information depending on the physical behaviour of heat transfer. The emulator aims to use the data at tested points to generate a data map and then estimate the expected value at untested points. So, the real energy consumption is normally different from the theoretical result. This also leads to difficulties in a building model and parameter calibration.

3.2.3 Model calibration process

The external wall thickness is one of the most significant factors of building envelops to resist heat transfer. The structure and material of the external building wall are chosen from the building information of the University of Glasgow. With different wall thickness, the IES simulated results of the heat demand of the whole year according to the ambient temperature in Figure 3.5(a). In the calibration approach, the initial tests are chosen with different wall thickness at 20 mm, 720 mm, and 1160 mm, respectively. It shows that different wall thickness and glass thickness cause the variation in heat demand. A wall thickness of 20mm is usually not possible in practice, but for the initial reference points of the surrogate model, we aim to provide as wide an initial range as possible to avoid it missing any possible parameter values. Figure 3.5(b) shows that different wall thickness causes the variation in heat demand. Comparing with the recorded heat consumption, the calibrated building wall and glass thickness, which finally affects the building thermal insulation, can be obtained to match the historical data. The result of 20 mm is obviously higher than that of the recorded data while the result of 1160 mm is lower than the recorded data. Then its best suitable parameter value can be detected from the calibration approach. Therefore, the objective of this work is to calibrate the building envelope parameters to detect the most suitable parameters to get the minimum estimation error from the IES simulation comparing with the recorded heat demand.

Window area will affect the overall heat transfer coefficient of the building and gain in solar heat transfer. It is easy to measure accurate window area values from each building. Therefore, the window area is an important parameter but does not require calibration. In addition to wall thickness, four other key parameters of the building envelope also have a significant impact on building heating demand, including glass thickness, roof thickness, floor thickness and internal partition thickness. In addition to building envelope parameters related to thermal performance, human activities also have a huge impact on building thermal demand, such as building managers turning on/off heating systems, people opening/closing doors and windows, and domestic hot water usage. But these are extremely random, making it difficult to use a model to track accurately.

The building model in the IES-VE simulator contains detailed information about each building, such as floor area, building height, the material of wall and roof, wall thickness, windows
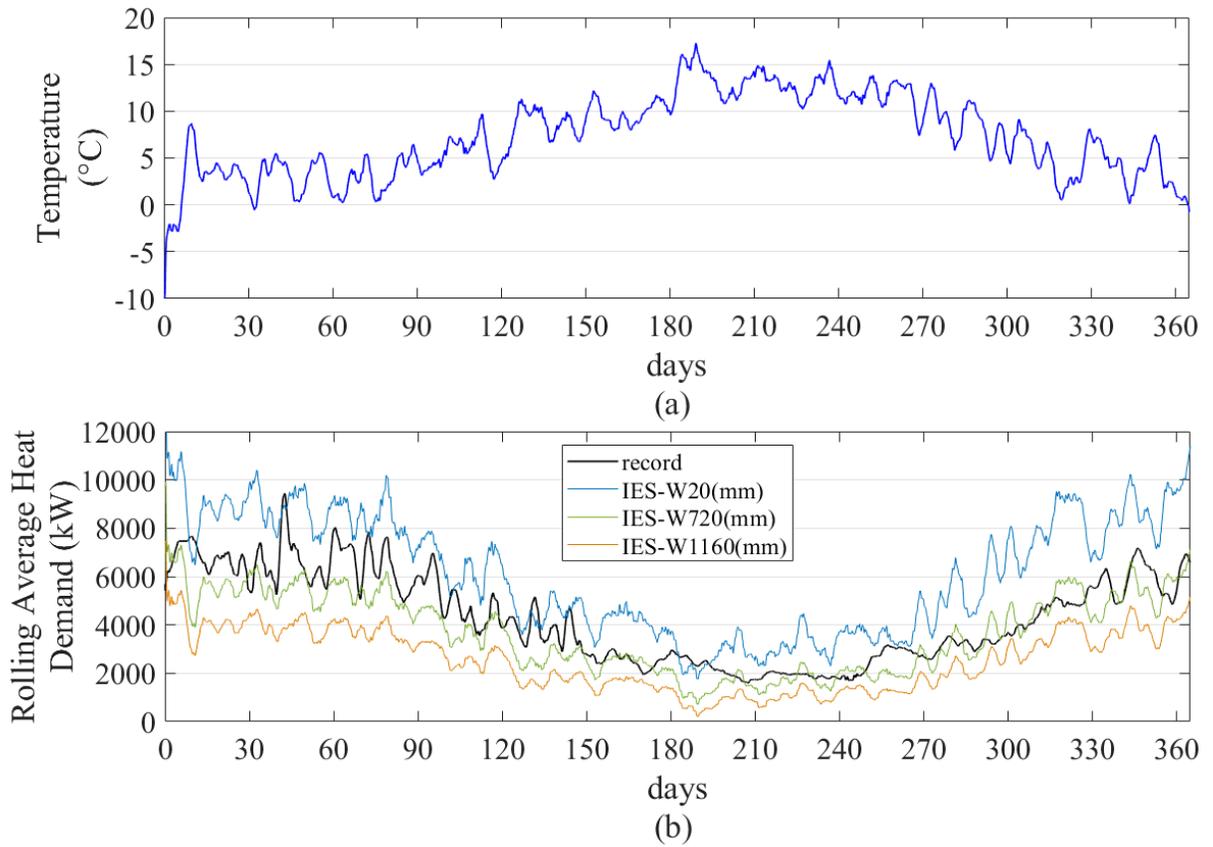
Figure 3.5: Temperature and estimated building heat demand of 24-hour moving average.

area, as well as solar angle and period. Besides, the weather condition for building energy simulation includes the temperature of dry and wet air, relative humidity, wind speed and direction, solar radiation, cloud cover, and atmospheric pressure. Thus, to simulate the building energy consumption with one set of parameters will take at least several minutes. The traditional calibration methods that require thousands of iterations will take too much computational time. The model calibration iteration process of the developed Bayesian-based calibration approach is shown in Figure 3.6.

In the model calibration process, the first step is to choose the initial values of parameters from their likely range and test them in IES-VE for the detailed building energy simulation. Then compare the simulated monthly heat demand with the recorded data to get the accumulated difference for a whole year with the tested parameters. Based on the result of difference using tested parameters, the conditional mean and variance of the difference at untested parameters are calculated from the designed emulator. After that, the emulated result of untested parameters can be analysed to calculate the reduced likely range of parameters. Then choose the parameter with the maximum likelihood from the reduced range as the suggested parameters to be evaluated in IES-VE simulation model in the next iteration. Repeat the model calibration process until the stopping criteria are met.

In the developed approach, the emulator is used to analyse the current simulation results to
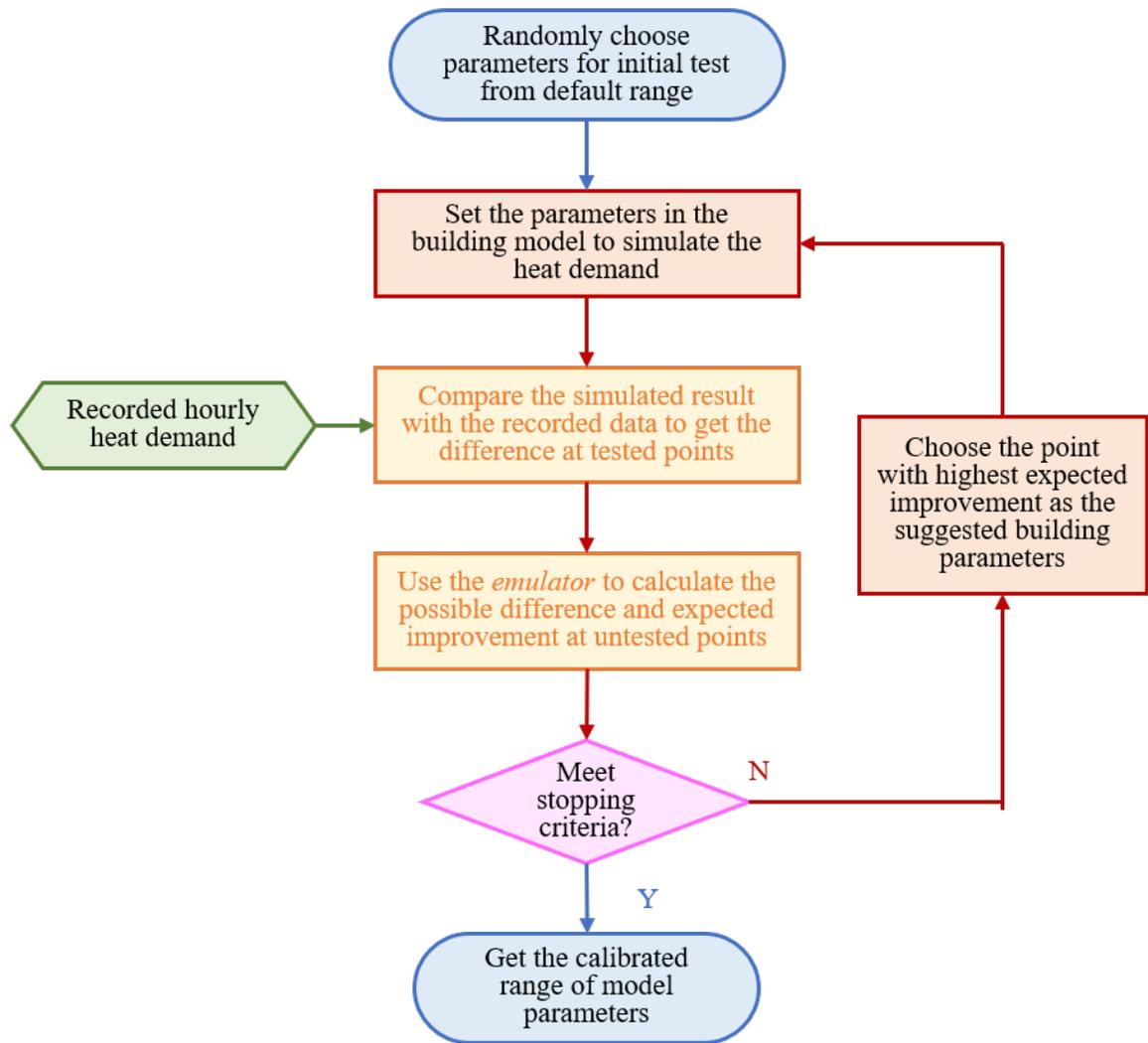
Figure 3.6: Flow chart of history matching calibration approach.

guess the parameters with the maximum likelihood as a suggested output. When the new simulation result is received, the emulator calculates the new conditional expectation and variance, and updates the suggested point of parameters. Therefore, even with fewer simulation results, the developed calibration approach can also provide a set of suggested parameters depending on current results.

3.3   Simulation Result and Discussion

In this section, a case study has been conducted to validate the proposed method in calibrating the thermal parameters of the building envelope. To verify the performance of the developed calibration approach, the case study is separated into three cases. The first case is to calibrate the single parameter, in which the wall U-value is chosen to be calibrated. The detailed explanation and result of each iteration are provided. Then the case is followed by two other cases to verify the approach in calibrating the two-parameters and multi-parameters simultaneously.

In the test, the calibration aims to find the most suitable parameter values of building model from historical data, including heat energy consumption and weather profile. However, if we use the model to predict the future heat demand of buildings, the prediction of long-term weather profile from historical data is difficult due to the uncertainty of unpredictable weather condition. In addition, human activities are also randomly happened events that cause the uncertainties in recorded real-time heat demand. Thus, the monthly heat demand contains the lead contribution of heat demand, and the impact of uncertainties can be significantly reduced in the monthly data. The other parameter settings in the IES-VE model for simulation is shown in Table 3.1. The values of the parameters are either from the building management office or estimated from the average daily usage.

Table 3.1: Settings of other parameters in IES-VE model.

Parameters	Value & Unit
Room heating set-point	21.8 °C
Occupant density of people	2 m^2/people
Internal gains of people	90 W/people
Lighting and equipment power density	45 W/m^2

3.3.1 Wall heat resistance coefficient calibration

The thermal behaviour of external wall is one of the most significant factors of building envelope, which affects the heat demand for a given input weather profile. Thus, in the first case, the thermal transmittance coefficient, or U-value, of the external wall is chosen as the only thermal parameter to be calibrated. The error criteria for calibration are the relative RMS error of monthly heat demand, based on hourly estimation results and the recorded heat demand data (derived from the gas consumption). However, the hourly recorded heat consumption includes the uncertainty of human activities, which is known as disturbance to the heat demand affected by weather conditions. In order to reduce their impact, the calibration uses the data of monthly heat demand to calculate the estimation error. Then their relative RMS error can be obtained as ration of the RMS error (i.e., the difference between the simulated results with the recorded data) and the recorded monthly heat demand. The estimation error is finally known as the percentage error of predicted monthly heat demand to the real value.

In the building consumed energy calibration, the conditional mean and variance are calculated from the emulator based on the training from the tested points, as shown in Figure 3.7(a). The black circles indicate the difference of heat demand between recorded data and simulated result at the tested U-value of the external wall in IES-VE simulation. Based on the initial tests, the point that causes the minimum heat demand difference needs to be obtained.

In the Gaussian distribution, its probability value at different variance is given as an explanation of the emulator's process. In Figure 3.7(b), the σ indicates the standard deviation, which is
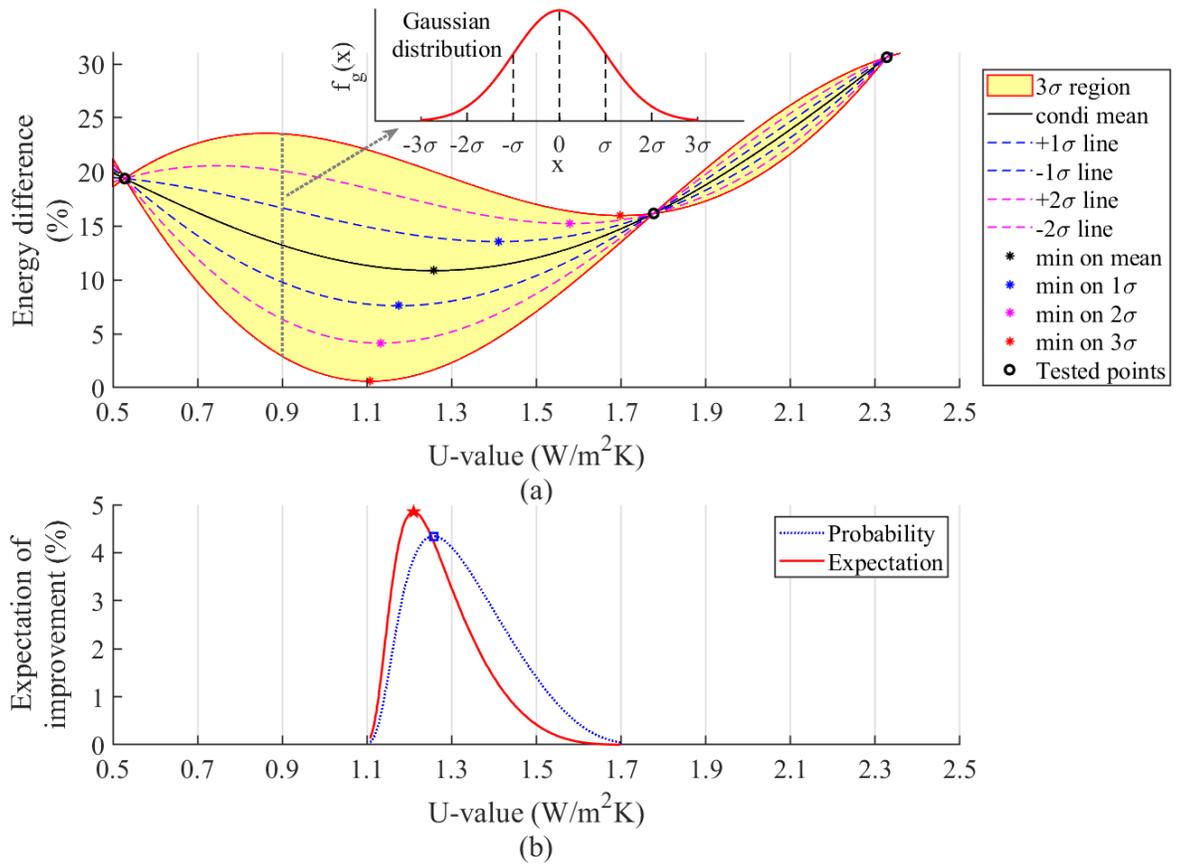
Figure 3.7: The process of emulator for model parameter calibration. (a) Emulator's process with Gaussian distribution; (b) the probability density function and expectation of improvement.

the square root of the variance. As the 'three-sigma rule' described, the points beyond the $\pm 3\sigma$ limit are assumed as a small probability event and can be neglected, as seen in the Gaussian distribution. The yellow region is the 3σ region at the untested points to cover all the probability of the event. As the case is to find the point with the lowest energy difference, only the bottom differences are considered in the calculation. The black line shows the curve of conditional mean at untested points and indicates the expectation with the highest probability. The dashed blue line shows the lower probable curve of $\pm 1\sigma$ away from the conditional mean line, while the dashed red line shows that of $\pm 2\sigma$. The star points show the min value of each line at $\pm 1\sigma$, $\pm 2\sigma$, and $\pm 3\sigma$. These can be used to emulate the expected global minimum point.

The probability density function (PDF) of the event that the minimum points match the real history data is shown as the blue dash line in Figure 3.7(b). The wall U-value below the lowest value and above the highest value of the curve is deemed to the implausible points, whose expectation improvement is negative, or its probability of available improvement is too small. If only the highest probability is considered, it is too conservative in choosing the suggested point for the next iteration, and that will cause many more iteration steps. Therefore, in consideration of the benefit of energy difference reduction. The probability and expectation curves are shown in the red line in Figure 3.7(b). The probability curve is calculated from its variance in Gaussian

distribution, while the expectation curve is the product of probability and improvements. Use expectation to represent the next suggested point taking into account not only the likelihood but also the degree of possible improvement. Thus, it provides a better guess to narrow down the parameters range for the next iteration. The suggested point shown as the red pentagram is not the highest probability point but is expected to get the best benefit.
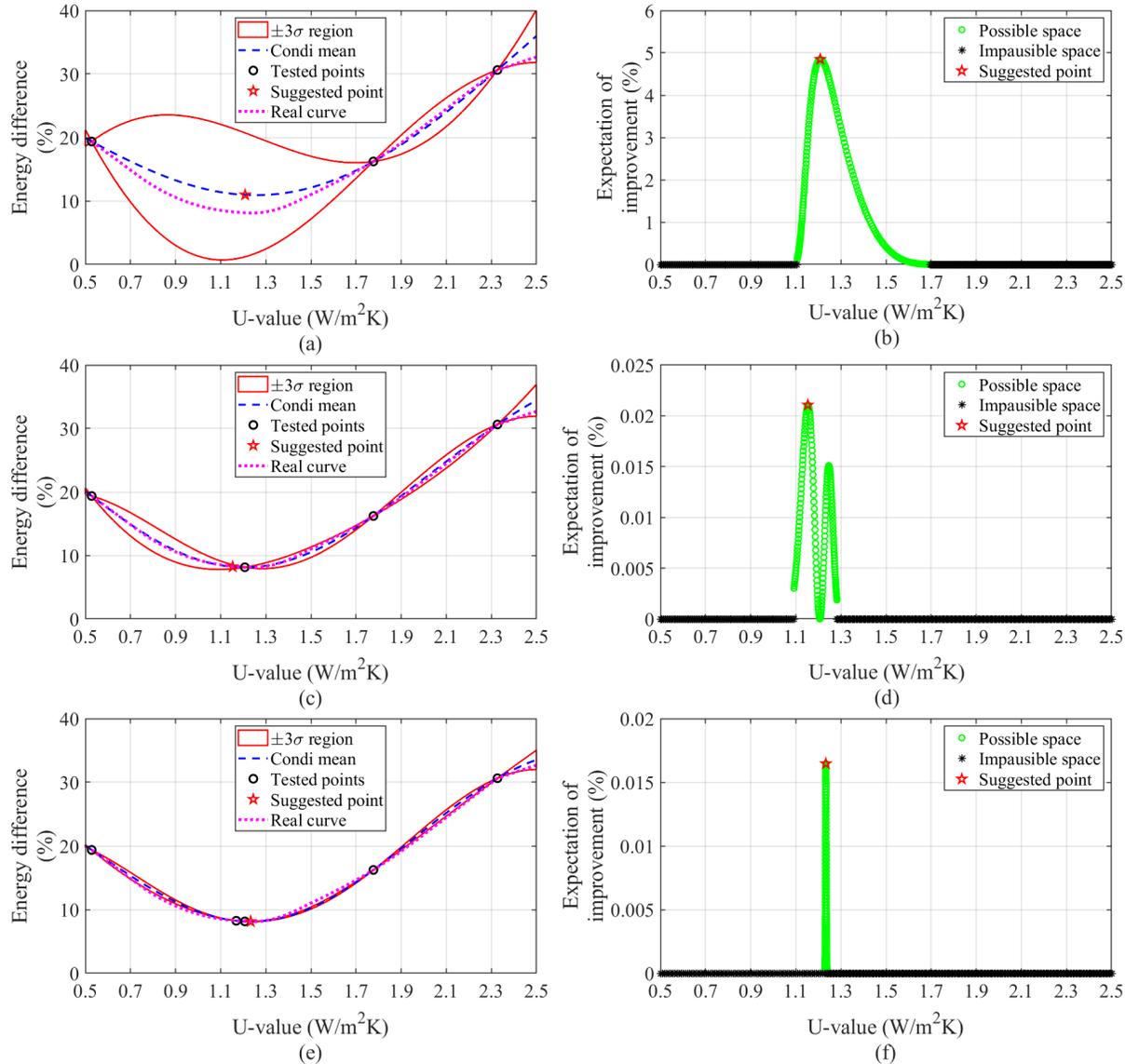
Figure 3.8: Results of one parameter calibration after each iteration. (a) Emulator's result in the 1st iteration; (b) expected input space and suggested result in the 1st iteration; (c) emulator's result in the 2nd iteration; (d) expected input space and suggested result in the 2nd iteration; (e) emulator's result in the 3rd iteration; (f) expected input space and suggested result in the 3rd iteration.

Depending on the output of the emulator, the x value of the suggested point can be used as the next input for simulator testing. Figure 3.8 shows the calibration process in three iterations. The green area at the bottom on the right side indicates the expected input space of the wall U-value. The initial parameter range of wall U-value is chosen from 0.5 W/m^{2}K to 2.5 W/m^{2}K,

which covers most building types. The three initial tested points are randomly chosen from the original parameter ranges and tested in the simulator to get the result of energy difference to recorded data. In the figure, when the U value of external wall is 2.32 W/m^{2}K the simulated heat demand is 32% different from that of the recorded data, which means the wall heat resistance coefficient is too high. In the same way, when the wall U-value is 0.52 and 1.78 W/m^{2}K, the heat demand differences are 20% and 16%, respectively. In the first iteration, the potential input space is within the range of 1.1 W/m^{2}K to 1.7 W/m^{2}K as presented with the green line in Figure 3.8(b), around 30% its initial input space. The first suggested parameter is at the wall U-value of 1.2 W/m^{2}K with about 5% improvement in expectation. Then, IES-VE is used to simulate the building heat demand with the wall U-value at the suggested value.

The result is shown in the second iteration, with the new point of heat demand difference being less than 10% as in Figure 3.8(c). Based on the information of the new tested point, the suggested parameter is updated to the wall U-value of 1.15 W/m^{2}K, which is close to the last suggested point. The figure indicates that the emulator estimates the minimum point has higher probability within the range from 1.1 W/m^{2}K to 1.28 W/m^{2}K, approximate 9% of its initial input space. The suggested parameter at 1.15 W/m^{2}K has the expectation of 20% improvement. The new suggested wall U-value is tested in the simulator for the next iteration. After the simulation test, it shows that the suggested U-value at 1.15 W/m^{2}K was a bad guess with an even higher heat demand difference than the previous point. However, the new result updates the range of valid wall U-value from 1.21 W/m^{2}K to 1.25 W/m^{2}K, approximate 2% of its initial input space. The final suggested value of the wall U-value is 1.24 W/m^{2}K. Thus, when the most expected improvement of the suggested point is less than the threshold, for example, 0.02% in this case, it is assumed that no more iteration is required. After the stopping criteria are met, the calibration gives the final range of the wall U-value to the building model. The calibration results of parameter range and simulation results of the suggested parameter in each iteration have been given in Table 3.2. The parameter range is reduced to 2%. With the calibrated parameter value, the estimation error of hourly heat demand is reduced from 16.1% to 8.11%, and the final suggested value of wall U-value is 1.24 W/m^{2}K. The calibrated wall U-value is higher than in modern domestic buildings. The latest Scottish building regulations standards require a maximum U-value of 0.27 W/m^{2}K for walls of new buildings as discribed in "Building Standards Technical Handbook, 2019" by Scottish goverment. Most buildings in the campus were constructed before the 1950s, before such regulations were introduced. The cavity of building external wall was left un-insulated and historical buildings were constructed with solid walls, which have even worse thermal resistance coefficient.

There are many optimisation methods for searching for the global minimum. However, most of the methods require the information of gradient or even second-order derivatives at tested point, such as the Gradient Descent and Newton's method. In the present case study, the gradient at tested point is not available from the physics-based energy model. Thus, to make

Table 3.2: Calibration results of each iteration in single parameter calibration.

Parameters	Iteration	Calibrated Range (W/m^{2}K)	Range to Its Initial (%)	Suggested Value (W/m^{2}K)	Estimation Error (%)
External wall U-value (W/m^{2}K	0th	0.5-2.5	100	1.75	16.24
	1st	1.1-1.7	28.2	1.2	8.16
	2nd	1.1-1.28	9.1	1.15	8.27
	3rd	1.21-1.25	1.9	1.24	8.11

a meaningful comparison, the proposed calibration method is compared with the golden search algorithm that does not require gradient information. The comparison is shown in Figure 3.9(a), the BHM method only needs four iteration times to reduce the initial parameter ranges into less than 2% while the golden search method needs ten iteration times to reduce the parameter ranges, showing that the proposed method only needs much fewer iteration times to run the simulation in calibrating a model parameter. The comparison of modelling error between golden search and BHM method after each iteration is given in Figure 3.9(b). Although both methods find the point of least modelling error before the final iteration, the wide range of possible parameter space also needs to be reduced to provide enough evidence. The BHM not only obtains a much faster speed in finding the optimised point, but also solves the issue that the Golden search method will reduce the search speed when it gets close to the target point. The improved efficiency of the approach increases the calibration speed and reduces the computational cost to run the large simulation model of a district-level heating network with multiple buildings.
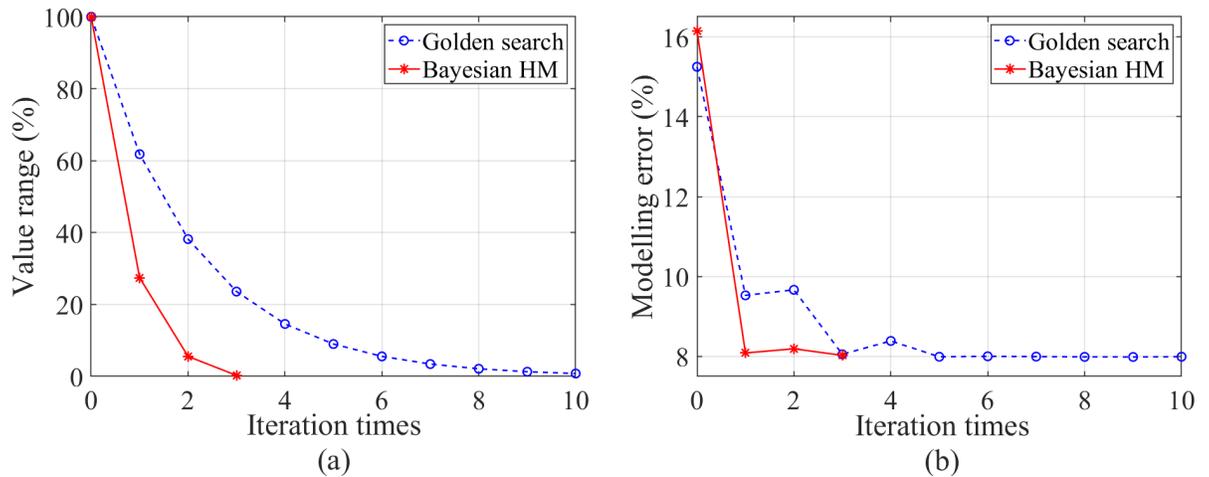
Figure 3.9: Comparison of BHM calibration method with traditional golden search method. (a) The reduction of value range; (b) the reduction of modelling error.

3.3.2 Dual-parameters calibration

Apart from the external wall, the thermal transmittance coefficient of windows is also a significant parameter in the building heat transfer. The windows' U-value depends on its thickness,

cavity and types, such as single or double-glazed window. The parameters of windows' thermal behaviour should also consider the average human activities to windows, such as opening a window that reduces the total thermal resistance of the building envelope. If the U-value of external wall and windows are calibrated individually using the single parameter calibration approach, it could not calculate their internal interaction between the parameters, and the calibration processes have to be operated many times. Therefore, an approach of calibrating dual parameters using BHM approach can be implemented. The results of calibrating both the wall U-value and windows U-value together are shown in Figure 3.10.
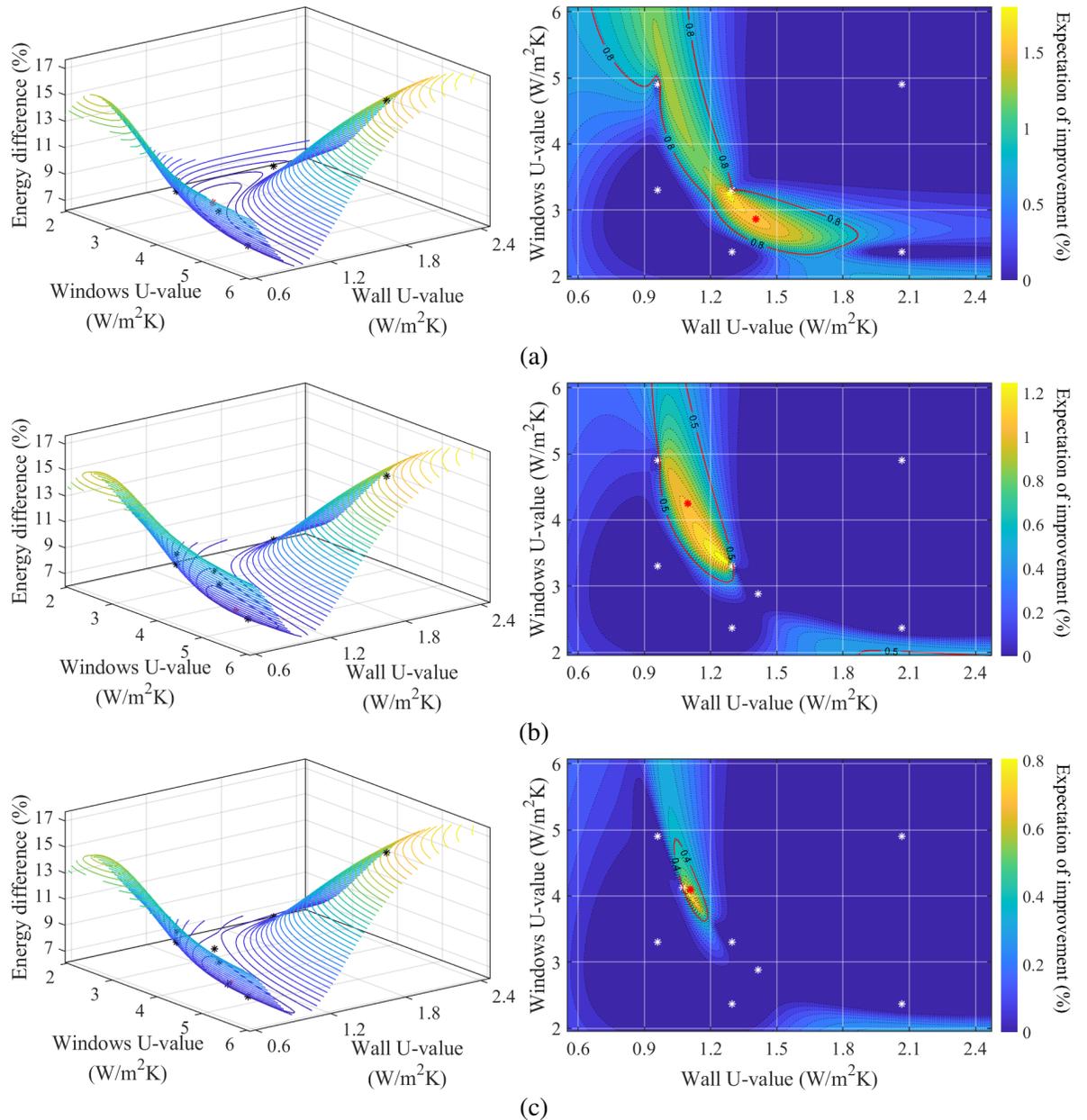
Figure 3.10: Two parameters calibration result after (a) the first iteration, (b) the second iterations and (c) the third iterations.

The left figure in Figure 3.10a shows the surface of the simulated error to the recorded

data regarding the variation of the two calibration parameters. The figures on the right show the likelihood density function that indicates the expectation of improvement at each point. From the colour bar, the lighter area shows higher likelihood while darker area shows the lower probability. The density map indicates the expected parameter input area to be calibrated in the two-parameter map. In the first iteration, the parameter input space has been reduced to 19% of its initial range; while in the second and third iterations, the input space has been reduced to 10% and 4%, respectively, as seen in Figures 3.10b,c. After several iterations, the final range of both parameters can be detected from the calibration. The wall U-value is calibrated to the range between 1.05 W/m^{2}K and 1.2 W/m^{2}K, and the windows U-value is calibrated to the range between 3.7 W/m^{2}K to 4.8 W/m^{2}K. The final simulated error is reduced to about 6.2% compared with the recorded heat demand.

3.3.3 Multi-parameters calibration

In addition to external wall U-value and windows U-value, several more parameters affect the total energy consumption in building models. In the next case, three additional parameters, the U-value of roof, ground, and internal partition/floor/ceiling, will be calibrated simultaneously for the building model. These five parameters describe the building fabrication and thermal behaviour of all buildings, and they are calibrated together with random values as initially tested points to feed the emulator for the expected value at the untested points. The number of untested points grows exponentially with the number of calibration parameters. To reduce the computational burden of the emulator to calculate the conditional expectation and variance at untested points, the sampling method is used to randomly choose the points from the prior distribution range as untested points.

Based on the calculated value from the emulator, the points with expectation higher than the threshold are retained in the result while the points with lower expectation are reduced from the initial range. Only the points from the remaining range will be considered as the untested points. The points beyond the remaining range are considered as implausible. Due to the high-dimension results of five parameters, the scatter matrix plot is given to compare the parameters in pairs, as seen in Figure 3.11.

In the figure, each row and column show the remaining points of the corresponding parameter regarding the other four parameters. For example, the second plot in the first column indicates the remaining points whose parameters are chosen with the corresponding external wall U-value and windows U-value. The same points with corresponding external wall U-value and roof U-value are given in the third plot in the first column, and so were other plots. The bar charts on the diagonal line show the statistical distribution of each parameter. The first figure in the first column indicates the statistic result of all the remaining points distributed along with the parameters of the external wall. After several iterations in the multi-parameter calibration process, the maximum likelihood distribution of external wall U-value is from 1.05 W/m^{2}K to W/m^{2}K after
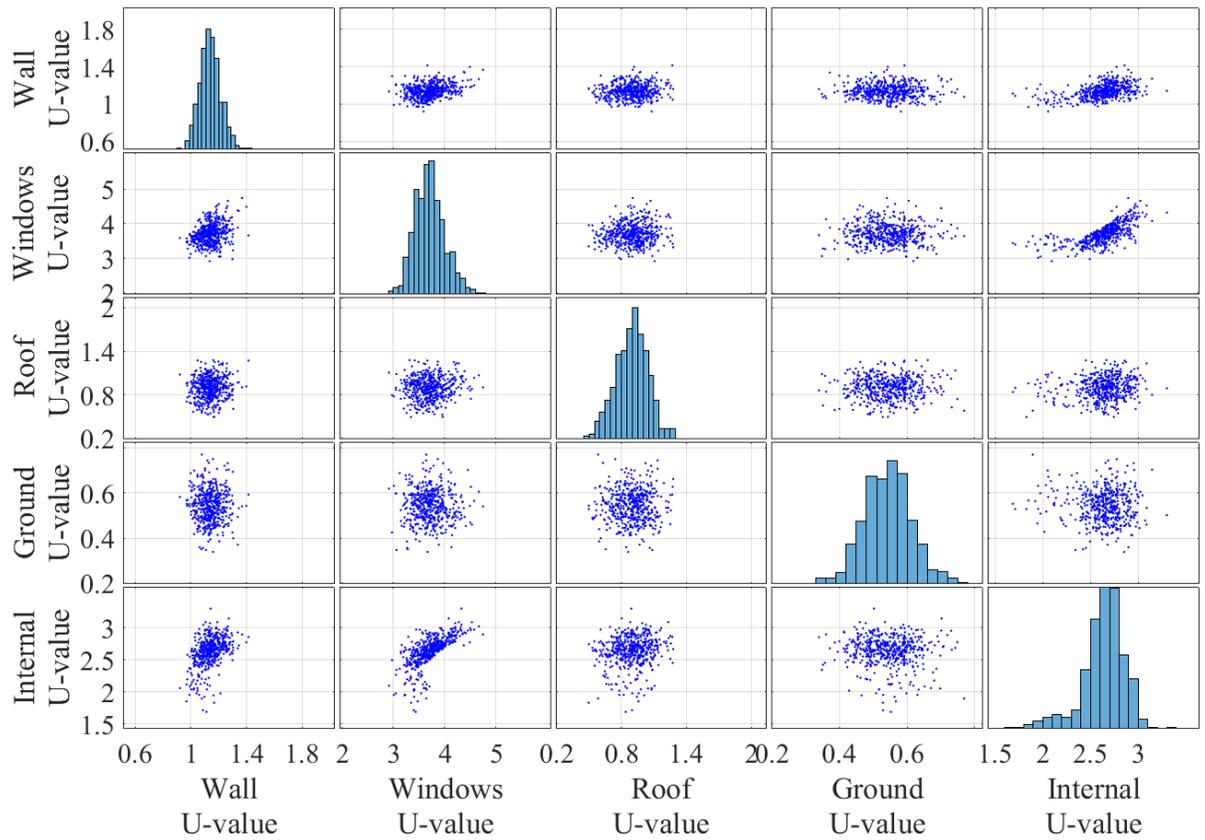
Figure 3.11: The remaining sample points projected to the corresponding parameters of external wall, windows, roof, ground, and internal partition.

the final iteration of multi-parameter calibration. Based on the statistical distribution result of each parameter, the best parameter values can be chosen from the highest expected range from the calibration results. The remaining space of each parameter is the maximum likelihood range to make the model output fit the recorded heat demand. The simulation results show that the monthly and hourly estimation error has been reduced to 4.3% and 12.63%, respectively, compared with the recorded data. The calibrated parameters and their suggested results are shown in Table 3.3.

Table 3.3: Calibrated parameters with initial range and calibrated value.

Parameters	Unit	Initial Range	Calibrated Range	Range to Its Initial (%)	Suggested Value
Wall U-value	W/m^{2}K	0.5-2.5	1.05-1.25	10	1.16
Windows U-value	W/m^{2}K	2-6	3.2-4.2	25	3.8
Roof U-value	W/m^{2}K	0.2-2.2	0.6-1.1	25	0.95
Ground U-value	W/m^{2}K	0.2-0.9	0.4-0.7	43	0.55
Internal U-value	W/m^{2}K	1.5-3.6	2.4-3	29	2.77

To validate the effect of model parameters, the sensitivity analysis has been made to provide the elementary effect of each parameter. The Morris method, which is the most commonly used

global sensitivity analysis method, has been used to test the elementary effect in the case [193]. The basic idea of Morris method is to evaluate the response of model output on the basis of a small change in a single parameter. The mean elementary effect of a single parameter is presented as:

$$\mu_k^* = \frac{1}{U} \sum_{k=1}^{U} \left| \frac{f(x_1, \cdots, x_k + \Delta_k, \cdots, x_n) - f(x_1, \cdots, x_n)}{\Delta_k} \right| \quad (3.24)$$

where μ_i^* is the mean elementary effect of the ith parameter out of n parameters in total, j is the number of tested data out of U tested data in total; Δ is a small change in parameter and $f(\cdot)$ indicates the model output with related parameters.

The results of the sensitivity analysis are shown in Figure 3.12. The x-axis shows five parameters of the walls, windows, roof, ground and internal partition. The y-axis indicates the elementary effect with the unit of kW per change of U value in every 0.1 W/m^{2}K. From the result, the wall U value has the highest elementary effect with about 70 kW increase of heat demand in every increment of 0.1 W/m^{2}K in its U-value. The second most significant parameter for the building thermal resistance is the roof U-value followed by that of windows and internal partitions. The ground U-value is the least significant parameter that affects the building thermal resistance and heat demand.
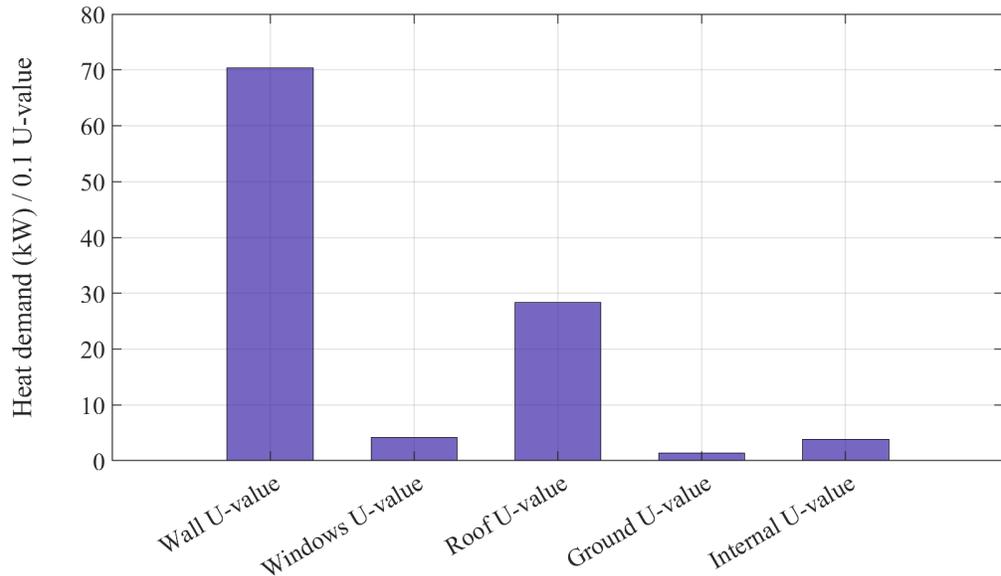
Figure 3.12: Elementary effect sensitivity analysis of each calibrated parameter.

The calibrated result of the IES-VE model is compared with the recorded data and default result before calibration as given in Figure 3.13. The results show that the calibrated monthly estimation error of heat demand has reduced from 16.3% to 4.3% as the best calibration result.

After the most probable parameters chosen after the calibration from the monthly result, the hourly heat demand of each building can be predicted from the model. The result comparison among one parameter calibration, two parameters calibration, and multi parameters calibration is given in Table 3.4. In the one parameter calibration, only the wall thickness has been calibrated
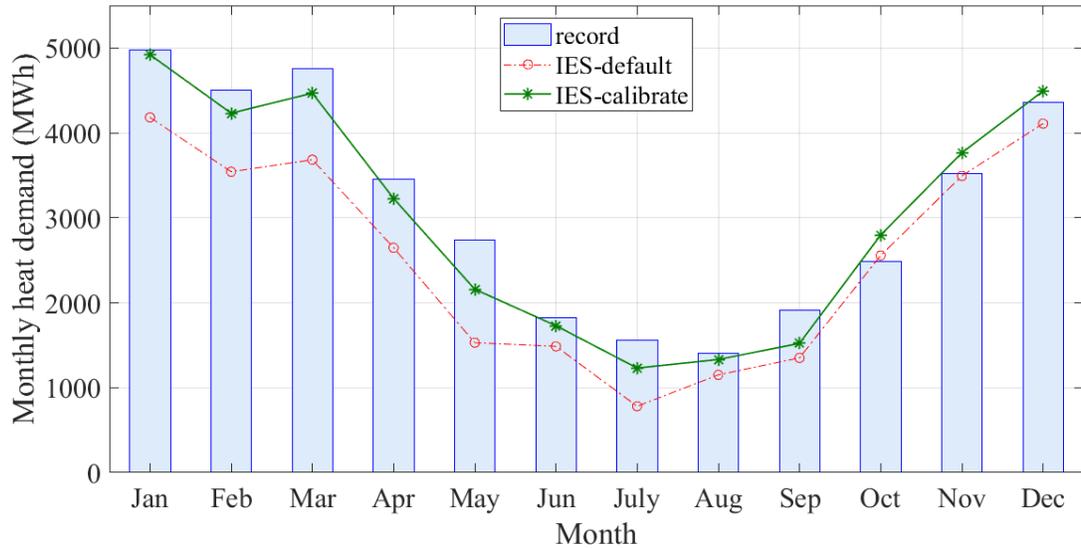
Figure 3.13: Estimation result using IES-VE comparing with recorded monthly heat demand.

and results in the smallest range. The result of that in the two-parameter calibration covers the same range. Because the second parameter has been calibrated together, the final range of the wall thickness is wider than that of the single parameter calibration. The same situation has been found in the multi-parameter calibration as well. The final range of each parameter is wider than that of the single parameter calibration, but the estimation error of model simulation is less in multi-parameter calibrations.

Table 3.4: Estimation result comparison among different number of parameters to be calibrated.

Parameters	Monthly Estimation Error	Hourly Estimation Error
Before calibration	16.3%	18.5%
One parameter calibration	8.02%	15.4%
Two parameter calibration	6.1%	13.2%
Multi parameter calibration	4.3%	12.6%

The average estimation error of hourly heat demand after multi-parameter calibration is around 12.6%. That means the estimation accuracy of using the proposed approach is more than 87% in predicting the hourly heat demand without known the detailed building information. The overall waveform of the predicted heat demand matches the recorded data and can be used for estimation of the whole heating network. In practical, the real building energy consumption not only depends on these parameters and weather conditions but also the human activities and other uncertainties, such as the observation error from the sensor, model uncertainty in simulating the thermal behaviour. In fact, the uncertainties, such as the domestic hot water usage and internal gains, are also contributors of heat demand, but they are hard to predict as it highly depends on the activities of occupants. In the IES-VE simulation, the domestic hot water usage and internal gains have been set as a constant ratio of the total heat demand of each building. The model only
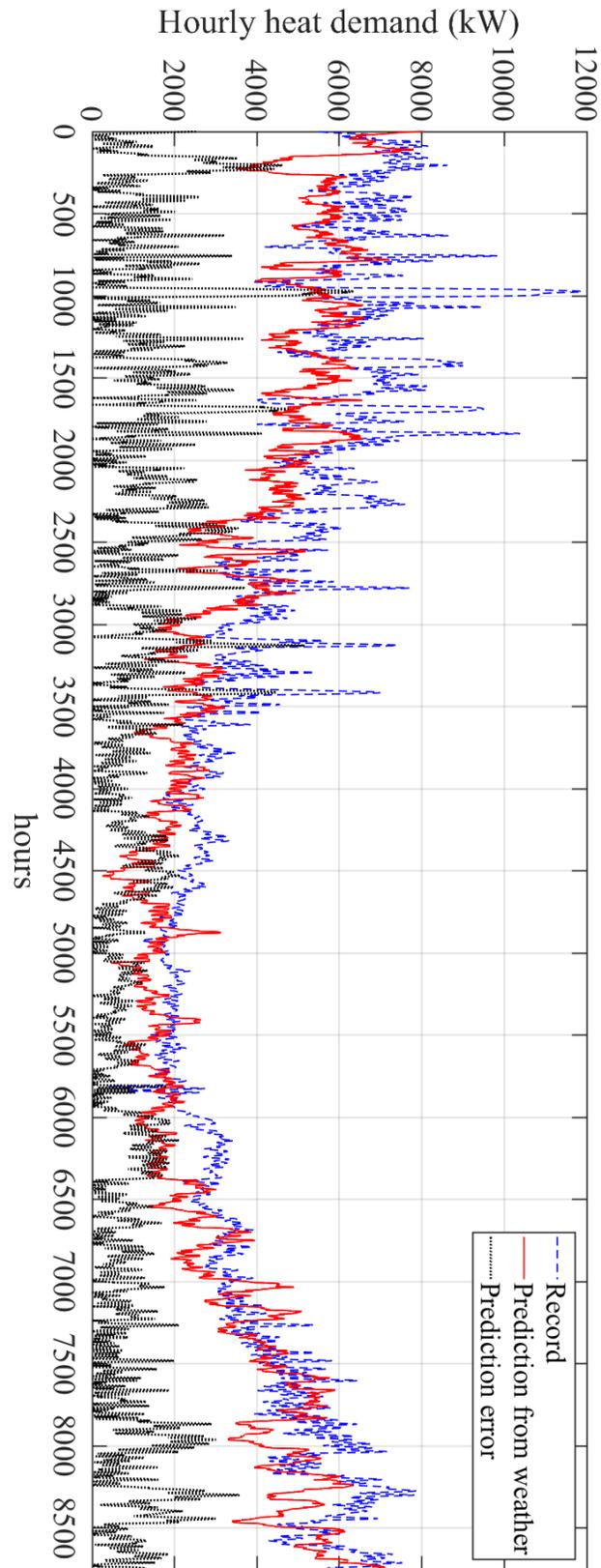
Figure 3.14: Comparison between record and predicted hourly heat demand.

predicts heat demand affected by the weather conditions (which is the dominant factor) in this work. The variation of heat demand caused by those secondary factors such as human activities is beyond the scope of this work, and it can be addressed by building control system.

The hourly heat demand estimation is shown in Figure 3.14. The recorded data of hourly heat demand is shown by the blue line. The red line represents the estmation on the basis of weather profile only, of which the actual human activities within the buildings have not been properly considered in the model. The grey line shows the absolute difference between these two sets of data, namely the estimation error of the model basis on weather profile. In comparison with Figure 3.13, one can find that the hourly estimation has much higher error than the monthly prediction.

There are two points should be highlighted here: (1) The IES-VE model has some default settings of human activities when the type of building is selected. However, such settings are just a constant value (i.e., an extra percentage of heat demand), which are very different from actual user's behaviour that however are dynamic and have not been recorded by the University's energy centre. Hence, the estimation error can be largely attributed to the unknown human activities within the buildings. (2) The model itself does not forecast weather, but uses the hourly weather profile of a whole year embedded in the IES-VE software, which are historical weather data recorded by the weather centre at Glasgow Airport. For this reason, the uncertainty introduced by the weather profile is considered negligible, and thus neglected in this chapter. Furthermore, the hourly weather profile is used, the estimation is can be considered as one hour ahead of time.

Comparing Figure 3.13 with Figure 3.14, with reference to Table 3.4, one can find that the hourly estimation result has much higher error than the monthly estimation result, showing that the accuracy of the model is strongly affected by human activities, especially at short time scale. In general, the more factors being considered, the more accurately the model can predict the heating demand. Since not all input parameters can be accurately measured (e.g., user's behaviour), the calibration methods developed in this chatper becomes necessary and useful.

In summary, the proposed approach uses IES-VE to build the bottom-up model of a district-level heating network, and some unknown parameters are calibrated using the Bayesian-based approach to improve the model's accuracy. The model can predict the hourly heating demand of each building with the error less than 13% for a case study, of which a number of building parameters are unknown.

3.4 Conclusion

This chapter developed an approach to predict the hourly heat demand of a district-level heating network, using a model calibrated by the proposed Bayesian-based calibration method and monthly recorded data. The IES-VE building energy software is used to build the bottom-up

model of buildings in a DHN. Due to the lack of detailed building parameters and heavy computational load of simulating such a large model, a Bayesian history matching based calibration method was proposed to calibrate the model over a shorter time span without the requirement of gradient calculation. In the approach, a statistical method-based emulator has been designed to estimate the conditional expectation and variance of model mismatch at untested points for the calibration iterations. The designed emulator-based calibration method has been compared with a traditional golden search algorithm, and the results show that the Bayesian approach-based emulator performs better with fewer calibration times to find the optimal point. Another case study has been designed to validate the proposed method in calibrating the single parameter, two-parameters, and multi-parameters of building envelope. The results indicate that the method is reliable and efficient to calibrate thermal parameters in building energy models. The multi-parameters calibration gives better performance by calibrating all parameters simultaneously. Finally, it is verified that the proposed approach can predict the hourly heat demand of a DHN without the requirements of detailed building information and accurate recorded data.

Chapter 4

Statistical Modelling of Heat Demand

4.1 Artificial Neural Network Used in Building Heat Prediction

In predicting the energy consumption of a district heating network, using engineering simulation to build a bottom-up model of each building is not efficient because it would require a very large number of hard to get building and occupant activity data, and the simulation would be computationally expensive and time consuming. The motivation of this chapter is to build an ANN model that is able to predict the heat demand of different types of buildings. If the model is only trained with the weather profile without giving the building characteristics, it will only be able to predict the heat demand of a particular building and cannot be transferred to other buildings of the same type. Therefore, the building characteristics is added as the inputs additional to the weather conditions in the ANN training. After the model is well trained, it will be used to predict the heat demand of other buildings by changing the inputs to characteristics of new buildings. For example, if a model is trained to predict the heat demand of an office building, the model should autonomously select its relevant features and train. Then the model can forecast the heat demand of other office buildings by just providing building characteristics, instead of requiring their historical data of heat demand and trained the model for each building.

Thus, the ANN model is expected to predict the heat demand of a district containing dozens of buildings after trained by the data of environment profile as well as building characteristics and heat demand of some sample buildings, as shown in Figure 4.1.

4.1.1 Artificial neural network model

Neural network is a computational model for nonlinear data fitting that typically includes the input layer, hidden layers, and the output layer [75]. There can be one or more hidden layers depending on the complexity of the model and training data. Each layer has several neurons and every neuron is connected to the output of all the neurons in the previous layer through
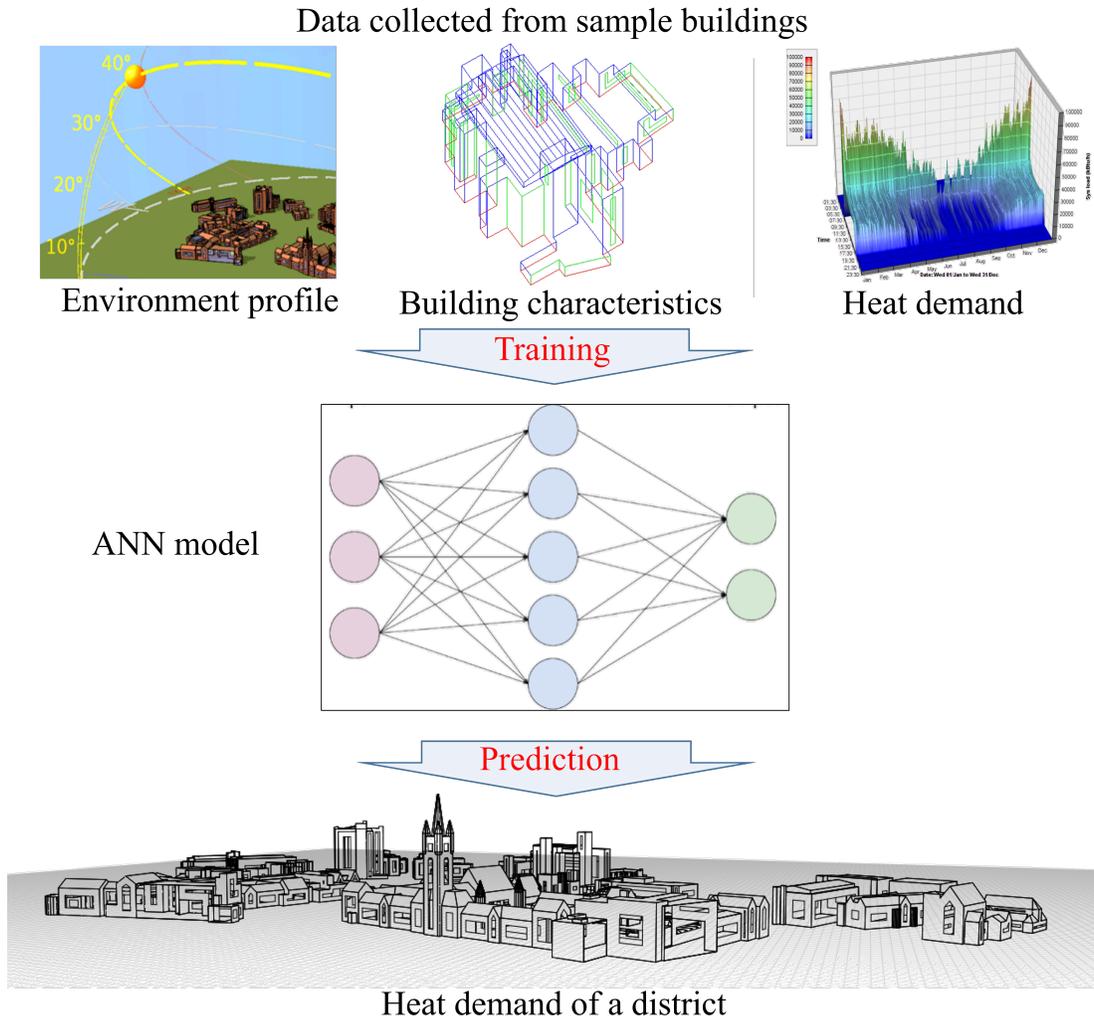
Figure 4.1: Structure of neural network to predict the heat demand of a district.

adaptable synaptic weights [194]. The training of ANN uses a group of input patterns in a data mapping process to produce the dependent variables for the corresponding inputs [195]. In the ANN data mapping process, the neurons in the input layer are multiplied by the weight of corresponding neurons in the hidden layer and then summed up with bias to the neurons in the output layers [77]. The predicted results are compared with the historical data, and their errors are used to update the neuron's weights by suitable adaptation [75].

The process of ANN can be described in mathematical formulas. Define $x_k(k = 1, 2, \ldots, n)$ as the k-th input attribute value which is passed along the links to the other layers. The weighted sum of signals, $\sum$, arriving at the input of the next neuron is subjected to a transfer function, which is most commonly the *sigmoid* function [196] with the following formula

$$f\left(\sum\right) = \frac{1}{1 + e^{-\sum}} \tag{4.1}$$

The j-th hidden neuron $h_j(j = 1, 2, \ldots, p)$ receives the sum of neuron value multiplied by the

weights $w_{kj}^{(2)}$ and bias $b_{kj}^{(2)}$ associated with the link as

$$h_j = f \left(\sum_{k=1}^{n} w_{kj}^{(2)} x_k + b_{kj}^{(2)} \right) \tag{4.2}$$

The output neurons are defined as $y_i (i = 1, 2, \ldots, m)$, which are summed up with their input signals and activation transfer function as

$$y_i = f \left(\sum_{j=1}^{p} w_{ji}^{(1)} f \left(\sum_{k=1}^{n} w_{kj}^{(2)} x_k + b_{kj}^{(2)} \right) + b_{ji}^{(1)} \right) \tag{4.3}$$

where f is the activation function, the sigmoid function used in the chapter; $w_{ji}^{(1)}$, $b_{ji}^{(1)}$, $w_{kj}^{(2)}$ and $b_{kj}^{(2)}$ are the weights and bias linked to the output layer (1) and hidden layer (2), respectively. This is a typical two-layer ANN model with an output layer and one hidden layer.

The error between target vector and predicted outputs from ANN model are used to validate the training performance. The appropriate error function is the mean square error (MSE) defined using the differences between the output vector y_i and the target vector t_i as

$$MSE = \frac{1}{m} \sum_{i=1}^{m} (t_i - y_i)^2 \tag{4.4}$$

The training error is used to update the ANN parameters of weights and bias of each neurons in the hidden and output layers. The training approaches of ANN include general regression, backpropagation (BP), radial basis function and fuzzy inference system. In this chapter, the BP learning algorithm is adopted to a typical two-layer ANN model. The BP algorithm is a supervised iterative training method based on searching the global minimum in the difference between ANN output and target [197]. The errors in the output are propagated back by calculating the derivatives that indicate the amount of 'responsibility' of each neuron using the gradient descent method [196]

$$\delta_i^{(1)} = y_i (1 - y_i) (t_i - y_i) \tag{4.5}$$

$$\delta_j^{(2)} = h_j (1 - h_j) \sum_{i} \delta_i^{(1)} w_{ji} \tag{4.6}$$

where $\delta_i^{(1)}$ and $\delta_j^{(2)}$ indicate the responsibilities of output-layer neurons and hidden-layer neurons, respectively. Then the weights and bias of links can be updated based on the responsibili-

ties [198, 199] as

$$w_{ji}^{(1)} = w_{ji}^{(1)} + \eta \delta_i^{(1)} h_j \tag{4.7}$$

$$w_{kj}^{(2)} = w_{kj}^{(2)} + \eta \delta_j^{(1)} x_k \tag{4.8}$$

$$b_{ji}^{(1)} = b_{ji}^{(1)} + \eta \delta_i^{(1)} \tag{4.9}$$

$$b_{kj}^{(2)} = b_{kj}^{(2)} + \eta \delta_j^{(2)} \tag{4.10}$$

where η is the learning rate of the BP neural network.

Apart from the methods and algorithms of machine learning in data-driven models, the data selection with different features and sizes used for training is also a vital factor of the model performance [75].

4.1.2   Data collection from building energy simulator

The development of a data-driven model will normally consist of data collection and processing as well as model training and testing. The training process of ANN requires a group of datasets from historical data records, which is used as benchmarks to train and test the model's performance. The number of inputs neurons depends on the number of input variables for training the ANN. In the original data, the input data has 26 different variables include both the weather conditions and building characteristics as shown in Table 4.1. The number of hidden neurons is chosen based on the complexity of the target model. The more hidden neurons used, the more complex the model can be achieved, but the longer training time will take. In this chapter, the number of hidden neurons is manually chosen considering both the model complexity and training time. The ANN model is structured with 26 input neurons for ANN training corresponding to the input variables and 20 hidden neurons for internal relation of the ANN model. The data of input variables used for training includes the corresponding hour of the day and months of the year, selected environmental variables, building characteristics and heat demand of all target buildings. The data of building heat demand is used to calculate the training error of the ANN model output. Where the number of people affects the average domestic hot water (DHW) usage, which increases the heat demand. And the heating setpoints indicates the average room temperature to calculate the required heat demand for heating up the room to the setpoint temperature.

Ideally, the most direct way is to measure the end-user heat demand of each building. However, there are many difficulties in measuring the real heat demand of the buildings. Firstly, it is required to install great number of sensors in every parts of the buildings. However, as the case study in University of Glasgow (UofG), many buildings have hundreds of years of age. These buildings do not have building management system (BMS) as modern buildings and it is difficult to reconstruct them. In addition, the collection of the required historical data will take at least

Table 4.1: Variables and units of input variables for ANN training

Types	Variables	Units
Times series	Time of a day	hr
	Days in a week	hr
	Weeks in a year	week
	Month in a year	month
Environment variables	Dry-bulb temperature	°C
	Wet-bulb temperature	°C
	Dew-point temperature	°C
	Daily mean temperature	°C
	Max. adaptive temp.	°C
	Wind speed	m/s
	Wind direction	deg
	Direct radiation	W/m^2
	Diffuse radiation	W/m^2
	Global radiation	W/m^2
	Solar altitude	deg
	Solar azimuth	deg
	Cloud cover	oktas
	Atmospheric pressure	Pa
	External relative humidity	%
	External moisture content	kg/kg
Building characteristics	Floor area	m^2
	Volume of plant	m^3
	Windows area	%
	Plant radiant fraction	0.0~0.1
	Time of heating	hrs/day
	Room heating setpoint	°C
	Number of people	pers
	DHW consumption	I/(h· pers)

several months but ideally several years in recording data from sensors.

Another way is to measure the heat demand from the supplier. In UofG, the buildings are heated via district heating network supplied by the energy centre. But the energy centre can only record the real heat demand of the whole campus but not each building. It makes difficulties in calculating the heat demand of each building from the sum of heat demand of the whole campus. Therefore, the bottom-up building models are built in the IES-VE software referring to all real buildings in campus. The building parameters are then calibrated to allow the sum of heat demand of all building models matching the real heat demand data of whole campus recorded by the energy centre.

IES-VE is an integrated system to build bottom-up models of buildings for thermal analysis and heating load simulation using the Apache engine. After setting the latitude and longitude of the target buildings, the weather profile for the building energy simulation is obtained from the

weather station at the nearest airport then the building thermal behaviour is simulated depending on the variation of environmental parameters, such as change of temperature and solar radiation process as shown in Figure 4.2. For this study it is Glasgow airport which is 9 km from the University of Glasgow. The settings are to make sure the simulated result is as close as possible to the measured heat demand. After the model is well calibrated by the data of whole campus, the simulated heat demand of each building can be obtained. The calibration work and validation result against measured data have been published in previous publications [99, 200]. Then the heat demand of each building simulated from this model is trusted to be close to their real heat demand. In this chapter, to simplify the data collection process, the training data is taken from simulated results generated in the building energy simulation software IES-VE. The data includes the weather profile, building characteristics and simulated heat demand is used for ANN training.
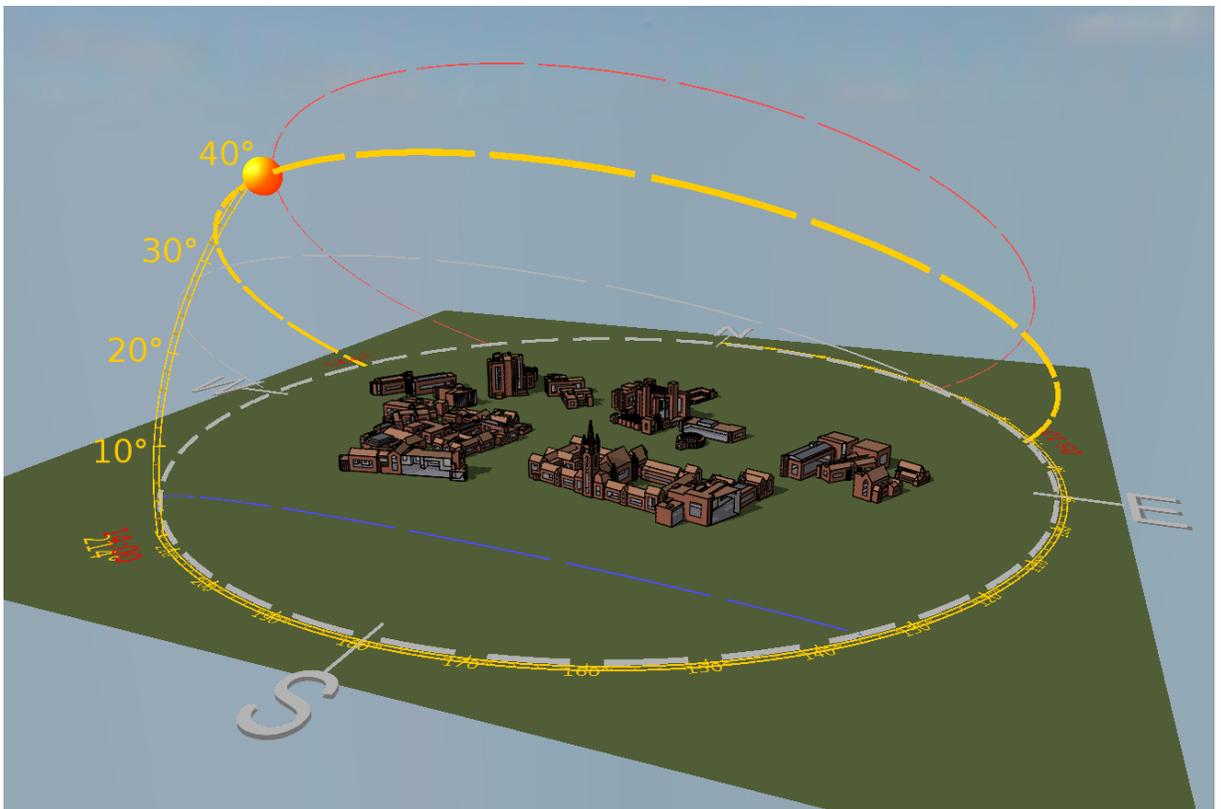
Figure 4.2: Weather profile for building thermal behaviour in IES.

In addition to the weather profile, different building characteristics will also affect the thermal behaviour of the building. The floor area, building volume, windows area and type, and wall thickness and material are the key characteristics of a building. Apart from these, the building heat demand also depends on its type of operational function. For example, the normal working hours of an office building are from 8:00 to 18:00, and its heat demand is different from that of a restaurant opening from 12:00 to 22:00 or a library open for 24 hours. The building fabrication data of every single building is collected from IES as shown in Figure 4.3. The data is collected
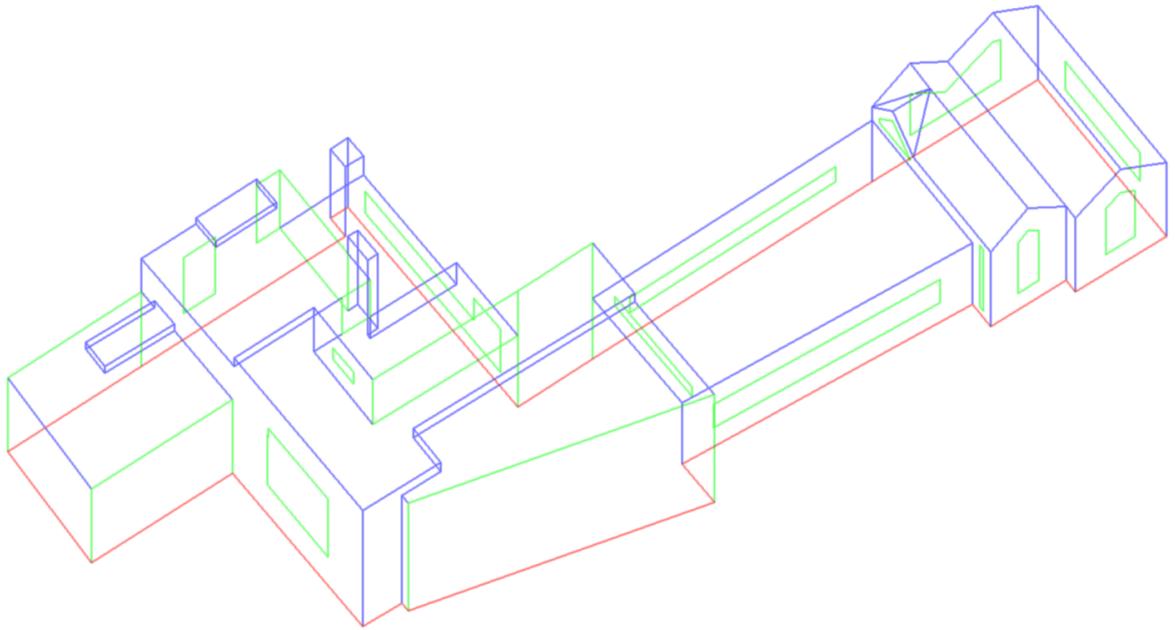
Figure 4.3: Building fabrication information of a sample building.

from the sample building models built in IES-VE which were calibrated to fit the recorded data of real heat energy consumption. The collected data is normalised before it is used to train the ANN model for the prediction of heat demand of the district heating network.

4.2 Development of Feature Selection Approach

4.2.1 Sensitivity analysis of heat demand to input variables

The collected data of weather information and building characteristics are defined as input variables to the ANN model. To validate the effect of model variables, the sensitivity analysis has been designed to provide the elementary effect of each variable. The increase or decrease of each variable will then provide a clear effect on the output. The input and output data used in sensitivity analysis should be exactly the same with that used for ANN training. In the original data, the inputs of sensitivity analysis are the 26 variables including both the environmental condition and building characteristics, while the outputs are the heat demand of sample buildings randomly chosen from the campus.

The commonly used global sensitivity analysis method is the Morris method, which has been used to test the elementary effect in the case. The basic idea of the Morris method is to evaluate the response of the model output on the basis of a small change in a single input variable. The mean elementary effect of a single variable from the complete set of data points of size U is presented as

$$\mu_k^* = \frac{1}{U} \sum \left| \frac{f(x_1, \cdots, x_k + \Delta_k, \cdots, x_n) - f(x_1, \cdots, x_n)}{\Delta_k} \right| \tag{4.11}$$

However, in real engineering applications, the input variables are normally varied simultaneously with time and it is difficult to find the effect of a small change of only a single variable on the target variable. Some research in analysing ecology data has realised the potential weakness of perturbing a single factor at a time [85]. The input x_k with n input variables and output $y = f(x_1, \cdots, x_n)$ at time step (q) comparing with that at time step $(q-1)$ can be presented as

$$x_k^{(q)} - x_k^{(q-1)} = \Delta_k^{(q)} \tag{4.12}$$

$$y^{(q)} - y^{(q-1)} = f\left(x_k^{(q-1)} + \Delta_k^{(q)}\right) - f\left(x_k^{(q-1)}\right) \tag{4.13}$$

Combined with Equation (4.11) with $U = 1$, the output of Equation (4.13) can be represented by the elementary effect μ_k as

$$
\begin{aligned}
y^{(q)} - y^{(q-1)} &= \Delta_1^{(q)} \mu_1 + \Delta_2^{(q)} \mu_2 + \cdots + \Delta_n^{(q)} \mu_n \\
&= \left[\Delta_1^{(q)}, \cdots, \Delta_n^{(q)}\right] \times
\begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix}
\end{aligned}
\tag{4.14}
$$

In the standard Morris method, the mean value is calculated to average the elementary effect of a single variable in fitting the data. To calculate the elementary effect of multiple variables, the problem is changed to finding a pair of μ_k to fit the relationship between changes on each variable to outputs as

$$
\begin{bmatrix}
\Delta_1^{(2)} & \cdots & \Delta_n^{(2)} \\
\vdots & \Delta_k^{(q)} & \vdots \\
\Delta_1^{(U)} & \cdots & \Delta_n^{(U)}
\end{bmatrix}
\times
\begin{bmatrix} \mu_1^* \\ \vdots \\ \mu_n^* \end{bmatrix}
=
\begin{bmatrix}
y(2) - y(1) \\
\vdots \\
y(U) - y(U-1)
\end{bmatrix}
\tag{4.15}
$$

Normally, the simplest way is to calculate the inverse matrix of $\Delta_k^{(q)}$ and its product with the output. However, the inverse matrix can only be calculated if the original matrix is square and non-singular. This is difficult and not always satisfied in data regression problems. The common approach for the linear equation is to find the least-square solution to minimise the unfitted error referring to all used data points for sensitivity analysis. Then the most suitable μ_k^* can be worked out to solve (4.15).

4.2.2   Determine the correlation of input variables

To find the correlation between input variables and outputs, the most common method is regression. Linear regression is a powerful tool that uses mathematical manipulations to transform the relationship between dependent and independent variables into a linear form. Based on this, many procedures were developed to derive the equation of a straight line using the least-squares

criterion for calibration. However, most engineering data is poorly represented by a straight line.

An alternative calibration is to fit polynomials to the data using polynomial regression, where the simplest is quadratic regression [201]. The quadratic regression ensures that the first-order derivative is continuous. The least-squares procedure can easily be extended to fit the data to a 2nd-order polynomial as the quadratic least square regression (QLSR) approach. The QLSR has the advantage to integrate both the convergence property of least squares and the probabilistic property of fuzzy regression to fit a non-linear mapping [202–204].

Define the quadratic polynomial equation presented by the input variables as

$$z = a_0 + a_1 s_i + a_2 s_i^2 \tag{4.16}$$

Assume the fitted value of the kth inputs as x_k can be presented as the sum of polynomial equations of the other $n - 1$ input variables

$$x_k = \sum_{c=1}^{n-1} z_c + \varepsilon \tag{4.17}$$

where ε is the unique information that is independent with any other input variables. Then, the error between the real and fitted value of each input variable can be represented by

$$e_q = x_k^{(q)} - z_c^{(q)} \tag{4.18}$$

To calibrate the value of each parameter in the polynomial equation, the estimated value can be obtained by fitting one variable at a time. The effect of other variables can be added to the error. The accumulated square error of U data points can be presented as

$$E_k = \sum_{q=1}^{U} e_q^2 \tag{4.19}$$

To minimise the accumulated square error by adjusting the parameters, the partial derivative of the accumulated square error with respect to each parameter can be presented as

$$
\begin{cases}
\dfrac{\partial E_k}{\partial a_0} = -2 \sum_{q=1}^{U} \left(x_k^{(q)} - a_0 - a_1 s_k^{(q)} - a_2 s_k^{(q)^2} \right) \\[2mm]
\dfrac{\partial E_k}{\partial a_1} = -2 \sum_{q=1}^{U} \left[\left(x_k^{(q)} - a_0 - a_1 s_k^{(q)} - a_2 s_k^{(q)^2} \right) s_k^{(q)} \right] \\[2mm]
\dfrac{\partial E_k}{\partial a_2} = -2 \sum_{q=1}^{U} \left[\left(x_k^{(q)} - a_0 - a_1 s_k^{(q)} - a_2 s_k^{(q)^2} \right) s_k^{(q)^2} \right]
\end{cases}
\tag{4.20}
$$

After defining the partial derivative as zero, the most suitable parameters a_0^*, a_1^*, and a_2^* can be obtained for the quadratic least square regression. After that, the coefficient of determination,

R^2 shows the quality of the fit of each variable to the target input as

$$R^2 \equiv 1 - \frac{\sum (x_k - z_c)^2}{\sum (x_k - \bar{x}_k)^2} \tag{4.21}$$

where

$$\bar{x}_k = \frac{1}{U} \sum_{q=1}^{U} x_k^{(q)} \tag{4.22}$$

Based on the coefficient of determination, the variable with the highest R^2 value can be chosen to fit the target variable. Thus, one key variable with its optimised parameters can be chosen in each iteration to fit the target input variable. The remaining difference between the target input variable and its fitted value is used in the next iteration. The fitting process using the QLSR is repeated until the R^2 value is lower than the threshold. The low R^2 value indicates that there has not been enough evidence to show that the remainder is determined by other variables. Then the assumed situation in (4.17) has been achieved and the process can be stopped. The lower threshold will cause more iterations in fitting the target variable. However, as the error reduced, the excessive iteration times does not give obvious difference in finding the most related features. Thus, considering both the performance and computational load, the threshold is set at 50% in the following case study to have acceptable fitting result. It validates that the correlation among input variables has been found as the target feature can be presented by other input variables. The result can then be used to remove the duplicated features from input variables to reduce the training data.

4.3 Simulation Case Study

4.3.1 Training with partial environmental parameters

Empirically, the temperature is the main drive of the heat demand of buildings. However, the heat demand is not linear related to the ambient temperature. Due to this, the relationship from temperature to the heat demand can be found using nonlinear fitting methods. ANN is theoretically able to fit any type of function with enough training data. Thus, the ANN is trained with a large amount of simulation data to get the relationship between ambient temperature and building heat demand.

The heat demand to ambient temperature, as the only input parameter, is predicted by ANN as in Figure 4.4 (a). The blue points indicate the training data while the black circles indicate the predicted result points. The result verifies that the heat demand is monotonically related to the ambient temperature, the lower temperature causes higher heat demand. However, there are still lots of points distributed around the black line. This presents that heat demand depends on not only the ambient temperature but also much other weather information and human activities.
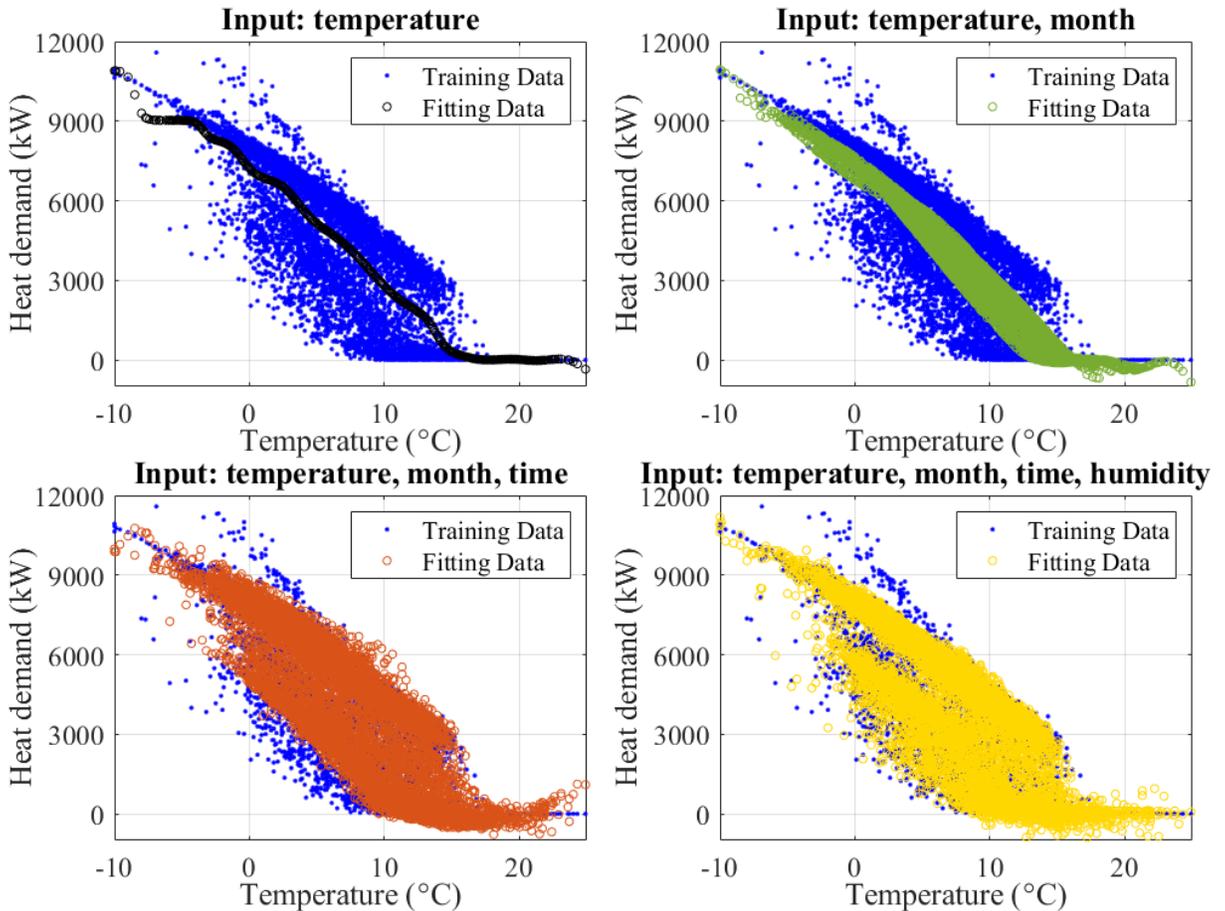
Figure 4.4: Heat demand prediction result of ANN after trained by different input factors.

Apart from the temperature, the heat demand is probably affected by other factors as well, including other environmental factors and time series. Other environmental factors include all weather profiles that can affect heat demand, such as humidity, wind speed, solar radiation. The time series indicates specific behaviour of customer heating as a routine of time, which can include the time of a day and month of a year. In addition, the heat demand in university buildings can depend on the months, such as the summer holiday and examination month. The month in a year indicates the part impact of human activities. The fitting results of temperature as well as other input factors, such as months in a year, time in a day, humidity. Figure 4.4 (b) show the training results of temperature and month in a year to heat demand, Figure 4.4 (c) shows the that of the above variables plus the time of a day, and Figure 4.4 (d) shows that of the above variables plus the humidity. The difference of heat demand predicted by the ANN based statistical model comparing with the engineering model of building energy simulator is shown in Figure 4.5. The bar chart shows the results from the single input variable to the multi-input variable. From the results, the more input variables used to train the ANN, the better fitting results in finding the relationship between inputs and the heat demand and less prediction error achieved.
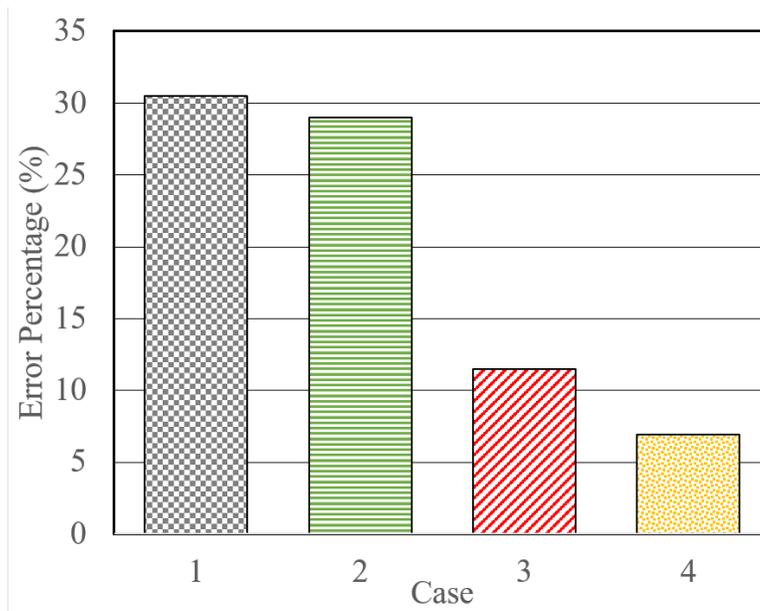
Figure 4.5: Error of ANN heat demand prediction with different training inputs. Case 1 - Input: temperature; Case 2 - Input: temperature & month; Case 3 - Input: temperature, month & time; Case 4 - Input: temperature, month, time & humidity.

4.3.2 Input variables and sensitivity analysis

The data used for ANN training is collected from IES-VE software database containing building characteristics, energy consumption including heat demand, and weather profile, which is recorded by the nearest weather station in an airport after setting the location. In total, the weather profiles include 16 different hourly recorded variables, as given in Table 4.1. A number of input variables are shown in Figure 4.6 for a period of 30 days.

Figure 4.6 (a) shows the five types of temperature information: dry-bulb, wet-bulb, dew-point, daily mean and maximum adaptive temperature. The figures shows that the five temperature variables have similar tendency, i.e. the different temperatures depend on the same weather information. Similar results can also be found in direct radiation, global radiation and diffuse radiation in Figure 4.6 (b). They have similar tendency under most circumstance. Figure 4.6 (c) shows other useful input weather variables including solar altitude, relative humidity and external moisture content for every hour. They have no obvious dependency on other variables but possibly have hidden and non-linear relationships, which will be discussed later. The ANN training aims to use these weather-based variables along with building characteristics as inputs and corresponding hourly heat demand as outputs.

Due to the multiplicity of the input variables, ANN training could take an extremely long time, furthermore, it reduces the efficiency in finding an effective input-output relationship in the ANN model. The sensitivity analysis (SA) introduced in Section 4.2.2 aims at analysing the sensitivity of each input variable to the output. The result of SA of input variables is shown in Figure 4.7. From the result, the wet-bulb temperature has the highest sensitivity to the building
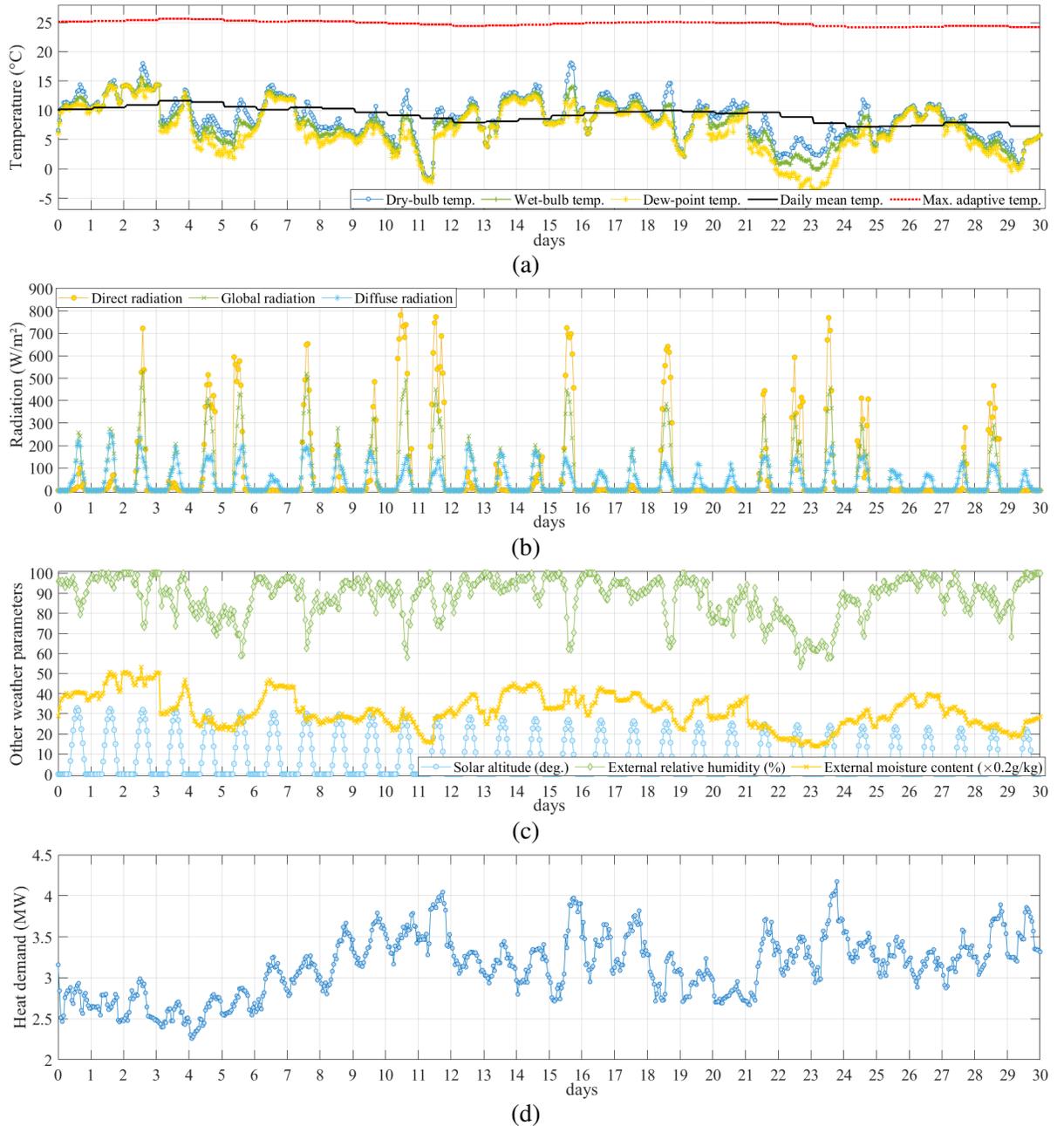
Figure 4.6: A sample of input variables used in ANN training: (a) dry-bulb, wet-bulb, external dew-point, daily mean and max adaptive temperatures, (b) direct, global and diffuse radiations, (c) solar altitude, external relative humidity, and moisture content, (d) recorded heat demand.

heat demand. The input variables include both the weather-based variables and the building characteristics. From the SA result, the top 5 highly sensitive variables are all temperature related. And the building characteristics with the highest sensitivity for heat demand prediction is the building volume. The wind direction is the variable with the lowest sensitivity to heat demand. With the result of sensitivity analysis, the number of weather input variables can be reduced via choosing the inputs with the highest sensitivity for training.

4.3.3   Analysis of correlation among input variables

In the sensitivity analysis, the five input variables with the highest influence on the output heat demand are wet-bulb, dry-bulb, dew-point, daily mean and maximum adaptive temperatures. Even though the five variables represent different types of temperature, they contain key information that affects all temperature-based variables. Thus, if the internal relationship between different input variables can be found, the number of input variables can be further reduced.

As the wet-bulb temperature has the highest influence in the sensitivity analysis to the output heat demand, it is tested and used as an example to show the fitting result. The QLSR method proposed in Section 4.2.1 is used to find the duplicated features in the target input variable. The fitting result in Figure 4.8 show the wet-bulb temperature on the y-axis and the value of the correlated variables on the x-axis. Figure 4.8 shows the fitting result in the first iteration. From which, it shows the relationship between the wet-bulb temperature with the other fifteen weather-based variables. To make it clearer in comparison, the coefficient of determination, which is also called R-squared (R^2), is used to determine the dependency between variables, as shown in Figure 4.9 (a).

The variable with the highest R^2 value is the dry-bulb temperature which has around 97% determination with wet-bulb temperature. After subtracting the determined part of dry-bulb temperature from the wet-bulb temperature, the remaining part is used as the next target to run the second iteration in QLSR. As shown in Figure 4.9 (b) to (d), the next determination variables are the relative humidity and dew-point temperature. In the 4th iteration, the R^2 value has dropped below 35% and the error is less than 0.1% of the nominal temperature range. Then the iteration stops.

The fitting result from QLSR shows that the wet-bulb temperature depends on three other features, the dry-bulb temperature, relative humidity and dew-point temperature, whose parameters of weights also provided from the QLSR fitting. After that, the wet-bulb temperature can be fitted by other three and the fitting weights based on equation (4.16). The fitted wet-bulb temperature is compared with its real value as demostrated in Figure 4.10. In the figure, the solid black line indicates the real wet-bulb temperature and the read dots indicate the fitted one that is calculated by other features. The result shows that the average fitting error, which is given in the bottom figure with the blue dots, is less than 0.5%. It also means thta the three fitted variables have more than 99.5% information of wet-bulb temperature. Therefore, it verifies that

the wet-bulb temperature can be fully described by the other three features.
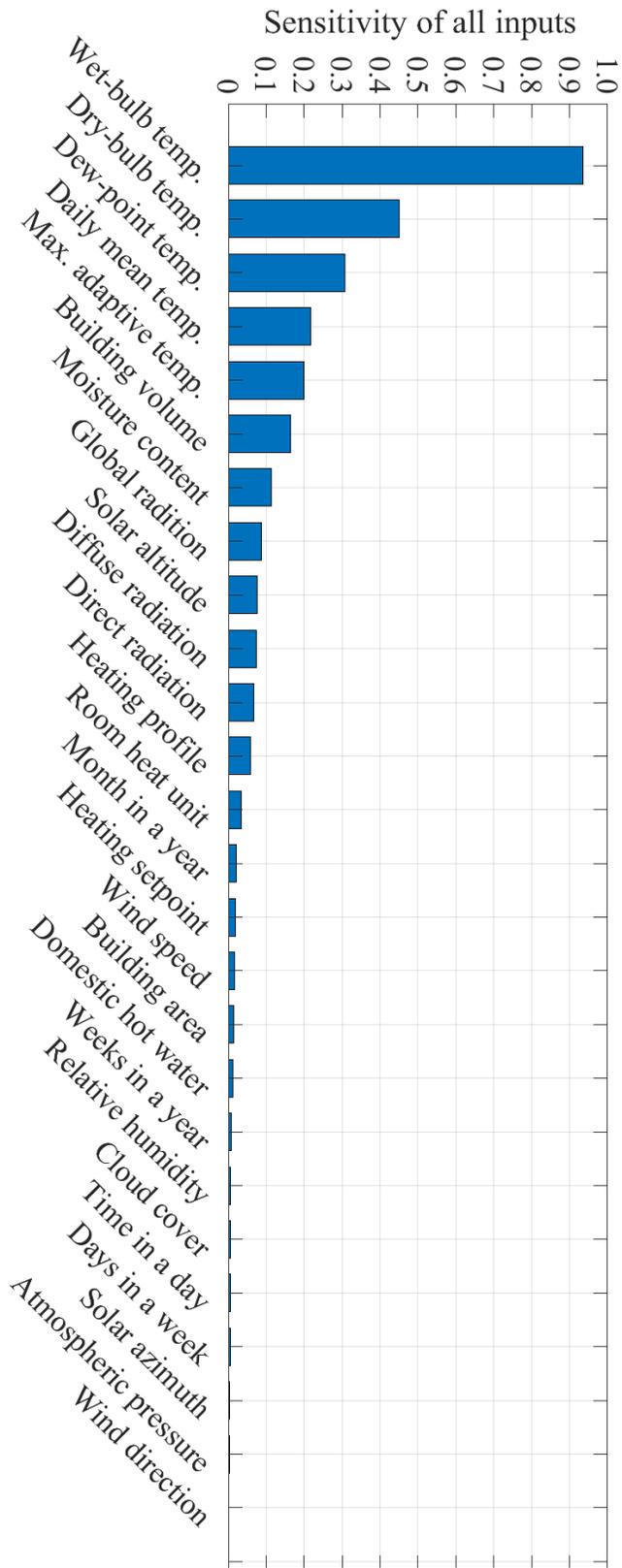
Figure 4.7: Sensitivity analysis result of all input variables
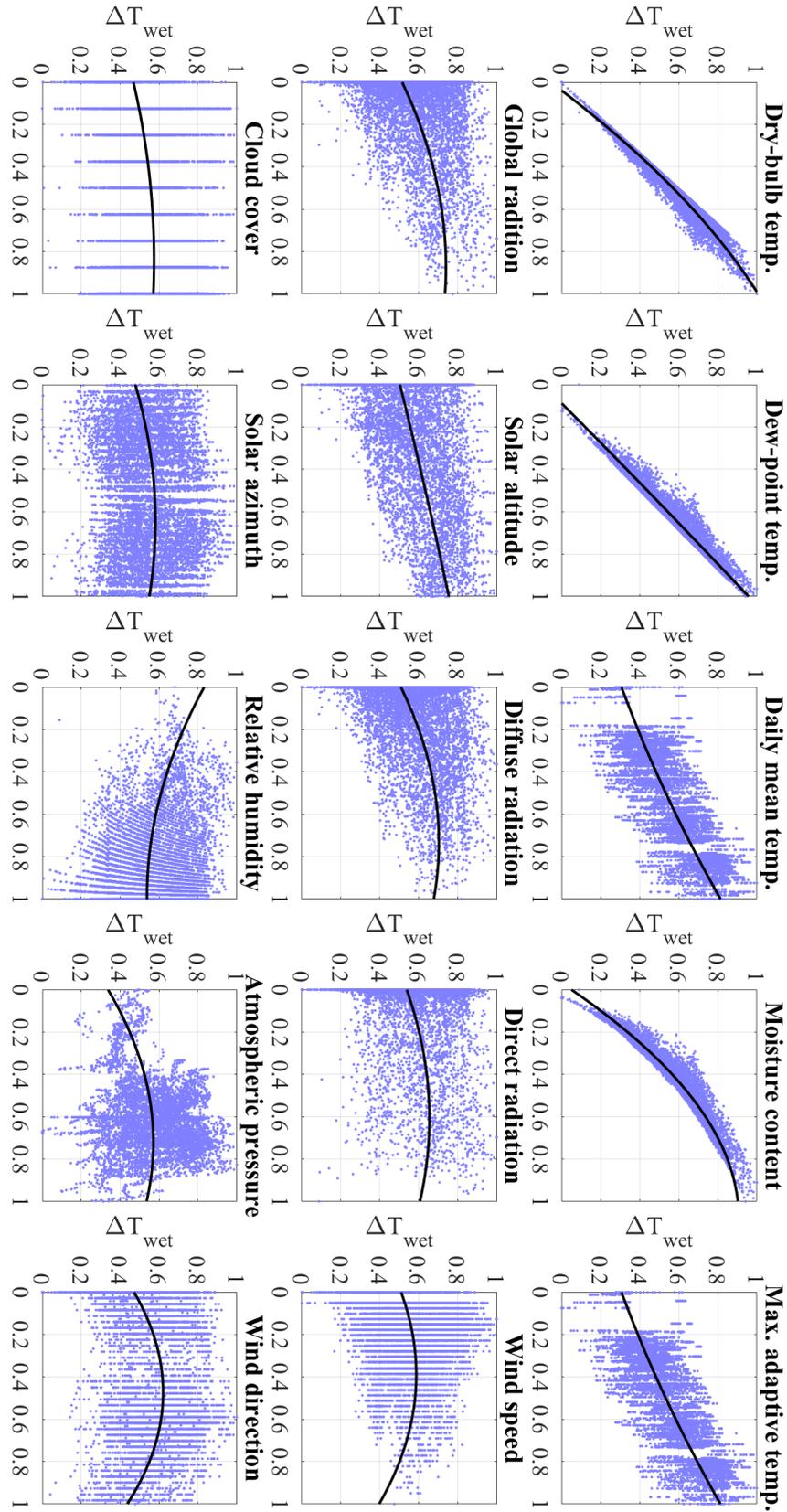
Figure 4.8: Quadratic least square regression fitting curve of the wet-bulb temperature using other weather-based input variables (1st iteration).
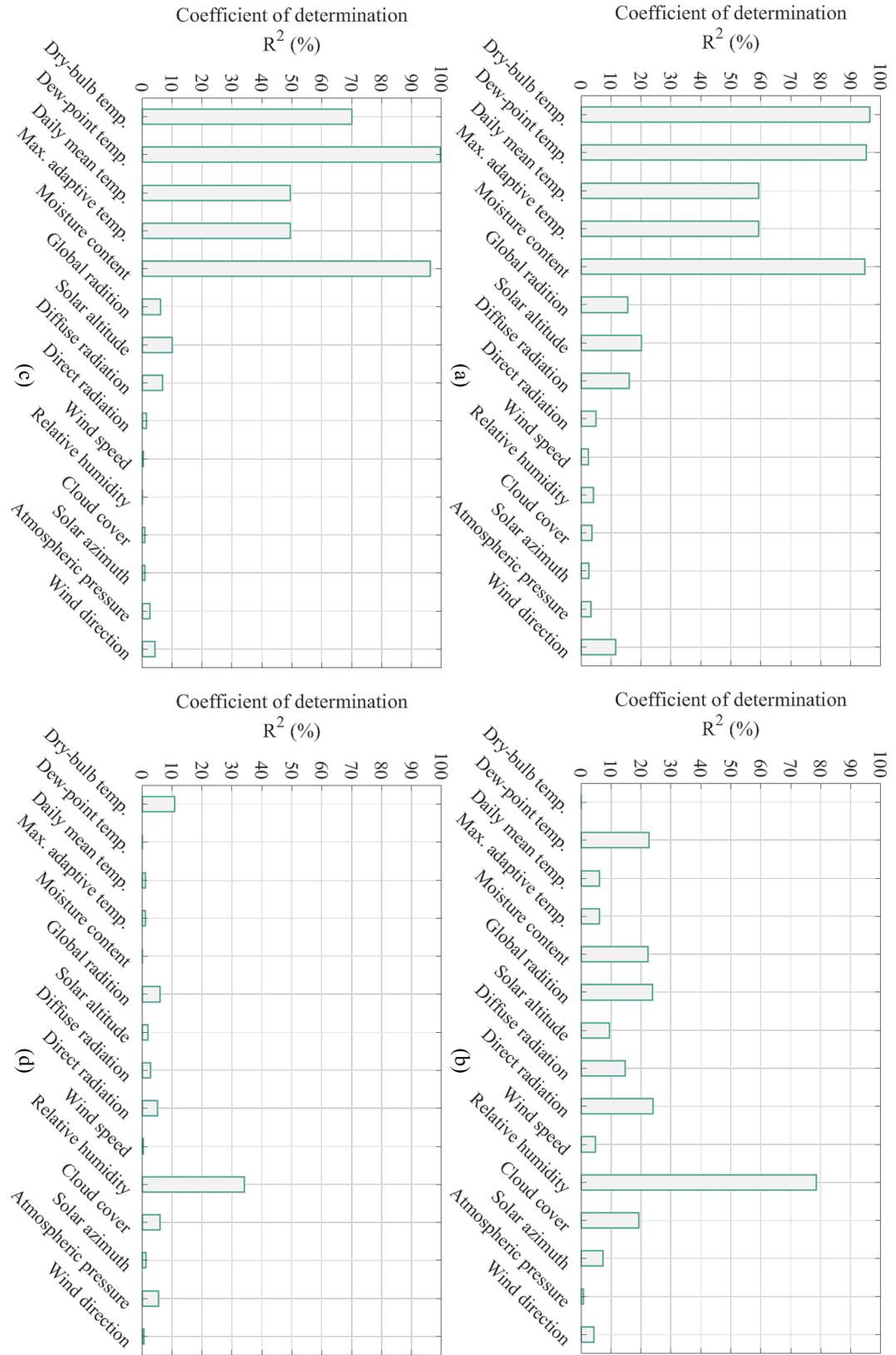
Figure 4.9: Coefficient of determination of other weather-based input variables to Wet-bulb temperature in each iteration: (a) 1st iteration, (b) 2nd iteration, (c) 3rd iteration and (d) 4th iteration.

If the wet-bulb temperature and other three features are all chosen as the input variables for the ANN training, the output heat demand will be found related to both the wet-bulb temperature and the other three feature. It causes the repeated training to the duplicated features. Thus, in order to improve the effectiveness of ANN training, the duplicated feature of wet-bulb temperature can be removed from the training inputs and the number of variables are then reduced for a faster training speed.

Similar with the QLSR approach used on the wet-bulb temperature fitting, other variables are also tested for duplicated features before they are used as inputs of ANN training. In the traditional sensitivity analysis result given in Figure 4.7, the cut-off criteria is set as 0.02 to allow 15 variables out of 26 used for the ANN training. After running the same approach on the 15 variables, the result indicates that 3 features out of the 15 can be further removed to reduce the training load of ANN. They are the wet-bulb temperature, the maximum adaptive temperature and the moisture content. Using the QLSR approach, they are found containing duplicated features with other variables and thus can be removed from the training inputs. All other 12 variables, which are known as the result of sensitivity analysis with reduced features, are chosen for the ANN training.

However, it has to say that the number of variables with duplicated features depends on collected data and different case studies. In other cases that using different data sources, the removed variables could be found no longer containing duplicated feature with others. Thus, it is necessary to run the proposed sensitivity analysis with reduced features approach for each data source and case study.
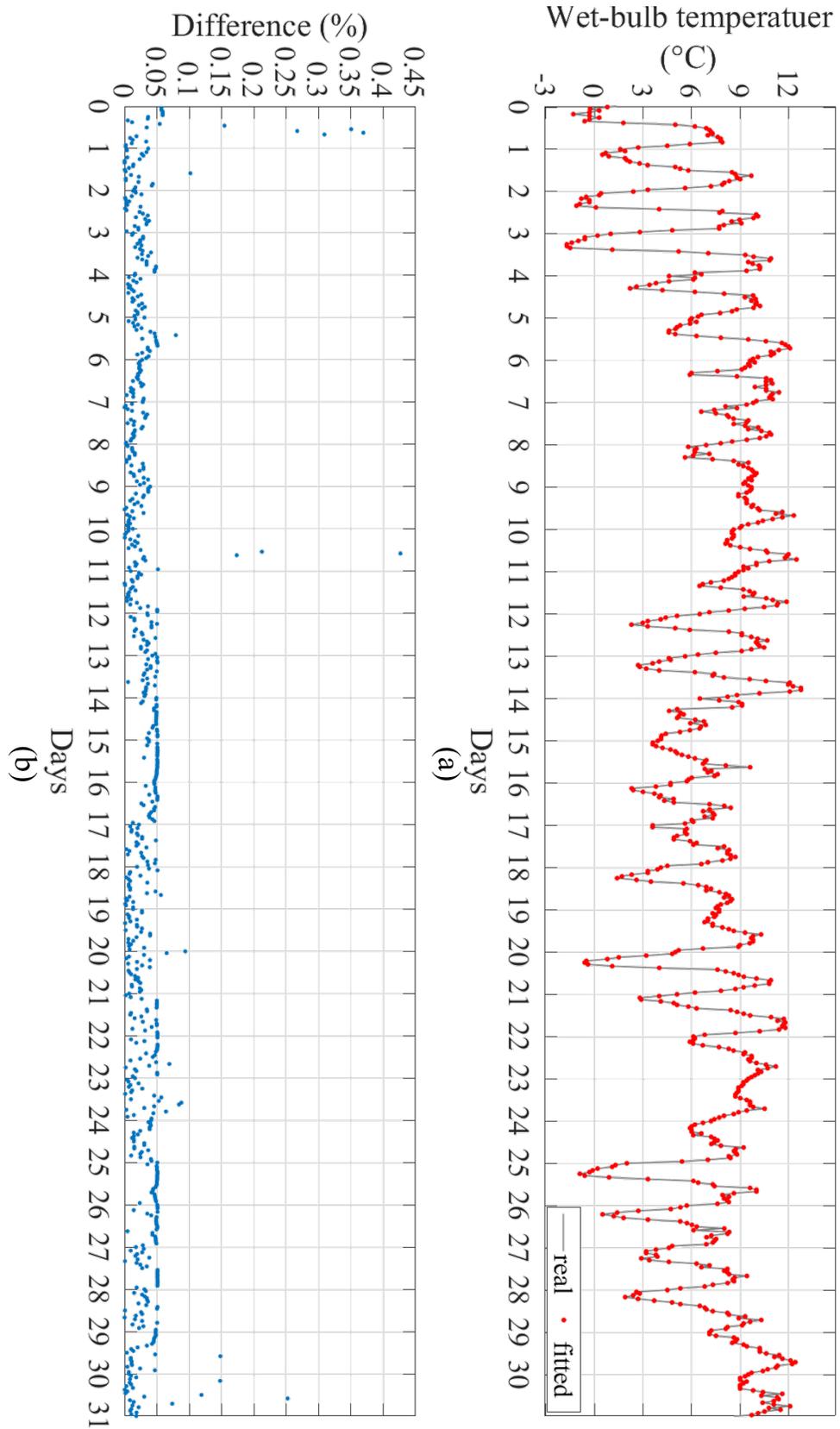
Figure 4.10: Fitting result of real wet-bulb temperature using searched input variables: (a) comparison between real and fitted wet-bulb temperature, (b) difference between real and fitted temperature.

4.3.4 ANN training and prediction performance comparison

To verify the effectiveness of the proposed approach, this section gives the comparison among the ANN training result of using all 26 input variables (All), the top 15 input variables from sensitivity analysis (SA), and 12 input variables chosen from the sensitivity analysis with reduced features (SARF).

The ANN is built and trained in Matlab with the built-in neural network toolbox. The parameter of ANN models is set using 1 hidden layer with 20 hidden nodes. The models are from the building models from the University of Glasgow, which has 36 different buildings with different heating types, and the total recorded data of 12 months' weather profile. The training of ANN uses part of the recorded data includes weather profile of 4 months from January to April, building characteristics of 10 randomly chosen sample buildings and their heat demand. As the weather profile and heat demand of the first 4 months and 10 sample buildings are used to train the ANN, the remained 8 months weather profile is used to test the training performance by predicting the heat demand of the remained 26 buildings.

Furthermore, the ANN training is using random initial weights and bias value. In the ANN training, the stopping criteria is that the gradient reduction of training error using the initial parameters is low enough, which means the training error is difficult to be further reduced. The training performance is mostly dependent on the initial parameters of weights and bias. Therefore, only one set of training result is not enough to show and compare the performance for randomly chosen initial weights. The simulation test uses the MATLAB neural network toolbox for ANN model training. In order to make a fair comparison and to give a convincing conclusion, the ANN training and heat demand prediction should be repeated many times so that the statistic results can be compared to show the average training performance. In the case study, each method for ANN training has been repeated one thousand times and the statistical results of all methods are compared. The probability density functions of statistical prediction error of heat demand and training time of ANN model for the one thousand repeated tests are shown in Figure 4.11.
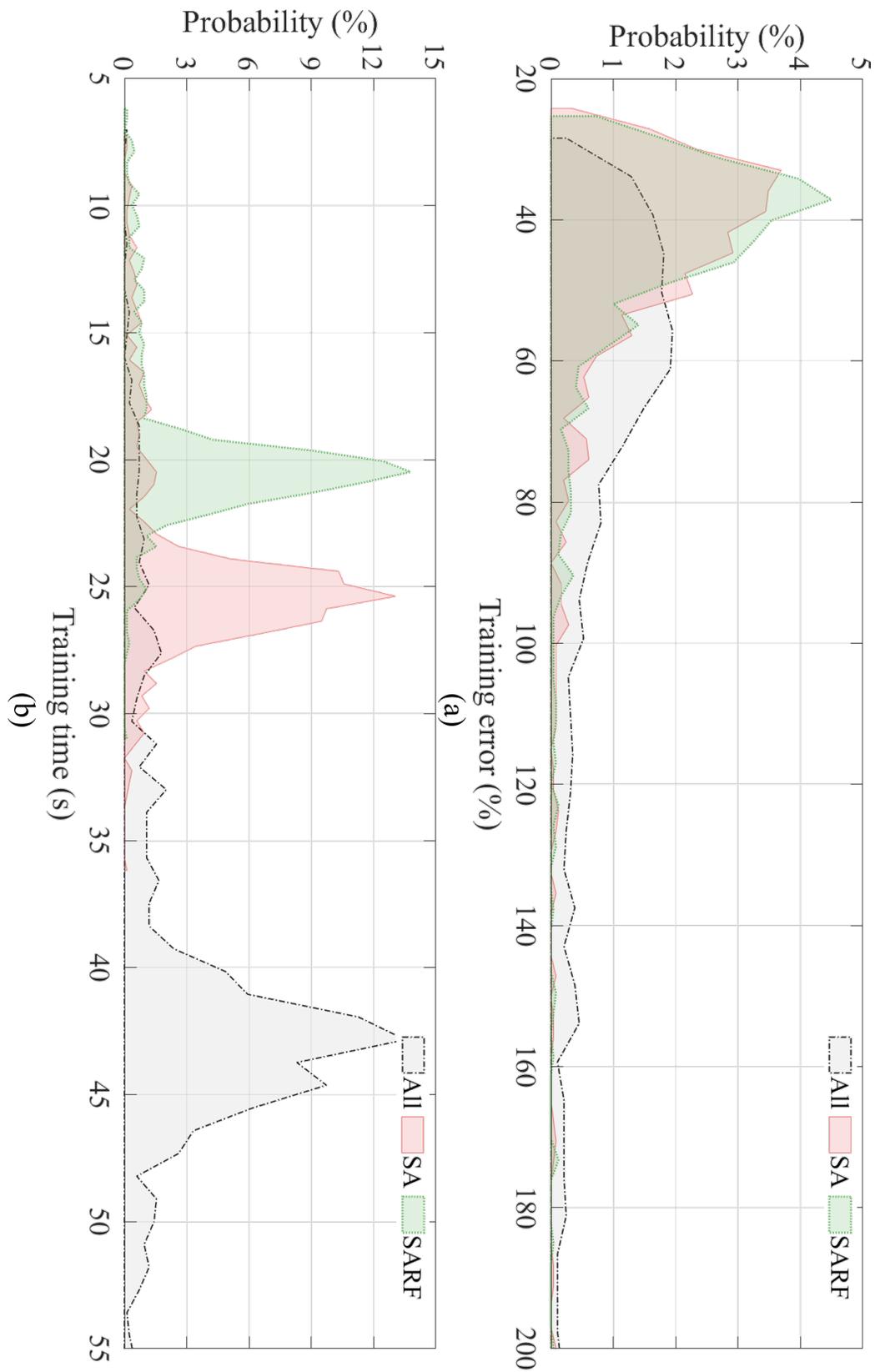
Figure 4.11: Statistic ANN training performance and result comparison: (a) probability of training error based on statistics, (b) probability of training time based on statistics.

Due to the randomness of the ANN training process, the results show some variability and, thus, the medians of the statistical results are given and compared. In the result of Figure 4.11 (a), the median of prediction error using the original method is about approximate 60% prediction error while that of SA and SARF are around 38% prediction error. And in Figure 4.11 (b), it is clear that the median of ANN training time with all variables is about 43s, the median of training time of SA is 26s and that of SARF is 21s, around 20% less training time than SA and 50% less than original all inputs. In addition to the median of prediction error and training time, the variance of SARF is obviously smaller than the variance of the case with all input variables and similar to the variance of SA. This verifies that SARF will reduce the uncertainty in ANN training and training time while achieving better performance.

In addition, the result in Figure 4.11 shows that the ANN training performance using the Levenberg-Marquardt (LM) training function, which is the most widely used training algorithm. However, there are many other training algorithms, including BFGS Quasi-Newton (BFG), resilient backpropagation (RP), scaled conjugate gradient (SCG), conjugate gradient with Powell/Beale restarts (CGB), Fletcher-Powell conjugate gradient (CGF), Polak-Ribiére conjugate gradient (CGP), one step secant (OSS), and variable learning rate backpropagation (GDX). The next step is to verify the performance of the developed SARP method in other training algorithms. Table 4.2, 4.3 and 4.4 show the prediction error of all training functions using the original 26 input variables, 15 input variables from SA, and 12 input variables from SARF. The indices for performance comparison are chosen as the mean, min and max prediction error its standard deviation (STD) as well as the mean, min and max prediction time and its STD.

To make the comparison clearer, the mean prediction error, error STD, mean training time, and time STD are drawn in bar charts as shown in Figure 4.12. The results show the prediction error and time of different training functions in ANN model training. In comparison among different training functions, the BFG, SCG, CGB, CGG, and CGP have the relatively better prediction result while the RP and GDX require the least training time. For all the training function used in the above ANN training, the SARF uses the least training time to perform a relatively better performance with the least prediction error. It verifies that the proposed SARF method can improve the ANN training efficiency and reduce the training time to obtain the same or even better performance in the heat demand prediction.
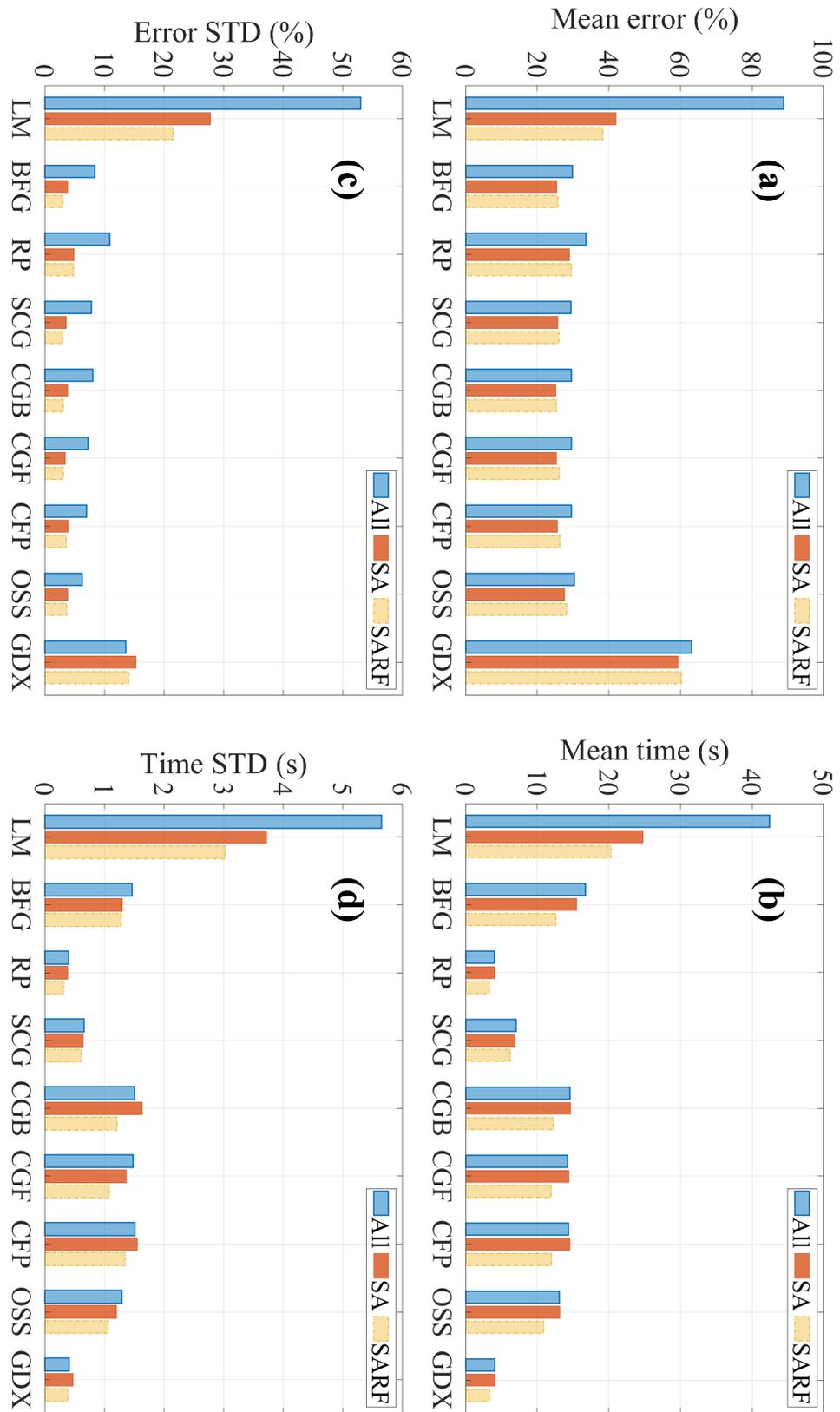
Figure 4.12: Comparison of using SARF with traditional SA method in different training functions. (a) Mean prediction error, (b) mean training time, (c) standard deviation of prediction error, (d) standard deviation of training time.

Table 4.2: ANN test result in heat demand prediction using original 26 input variables

Algorithm	Mean Error (%)	Min. Error (%)	Max. Error (%)	Std. (%)	Mean Time (s)	Min. Time (s)	Max. Time (s)	Std. (s)
LM	88.81	24.56	199.96	52.99	42.47	33.46	58.63	5.93
BFG	29.90	20.56	121.61	8.38	16.77	13.47	23.38	1.54
RP	33.66	21.58	112.60	10.94	4.05	3.10	5.81	0.43
SCG	29.41	21.34	93.62	7.81	7.06	5.23	8.90	0.69
CGB	29.52	20.25	83.62	8.07	14.61	7.57	21.86	1.58
CGF	29.53	21.43	112.18	7.29	14.23	10.84	18.27	1.55
CGP	29.59	19.99	86.54	7.06	14.37	0.69	19.29	1.59
OSS	30.39	21.93	89.76	6.31	13.12	9.70	16.36	1.36
GDX	63.18	38.80	171.49	13.58	4.06	3.21	5.77	0.43

Table 4.3: ANN test result in heat demand prediction using 15 input variables from SA

Algorithm	Mean Error (%)	Min. Error (%)	Max. Error (%)	Std. (%)	Mean Time (s)	Min. Time (s)	Max. Time (s)	Std. (s)
LM	42.02	20.21	195.74	27.84	24.81	27.17	57.70	5.09
BFG	25.46	18.06	46.96	3.88	15.53	12.05	19.80	1.78
RP	29.09	20.84	58.42	4.89	4.05	2.99	5.74	0.52
SCG	25.72	19.29	55.29	3.64	6.96	5.50	11.67	0.89
CGB	25.29	18.98	54.58	3.84	14.67	0.51	22.86	2.24
CGF	25.40	18.98	43.80	3.42	14.44	10.83	21.45	1.87
CGP	25.66	19.18	49.98	3.96	14.55	11.25	22.73	2.12
OSS	27.74	20.66	55.91	3.88	13.15	10.42	19.23	1.64
GDX	59.32	33.85	153.99	15.28	4.12	3.08	8.18	0.65

Table 4.4: ANN test result in heat demand prediction using 12 input variables from SARF

Algorithm	Mean Error (%)	Min. Error (%)	Max. Error (%)	Std. (%)	Mean Time (s)	Min. Time (s)	Max. Time (s)	Std. (s)
LM	38.31	20.84	209.48	21.53	20.80	3.95	38.34	7.71
BFG	25.72	17.48	38.74	3.08	12.65	9.15	19.87	1.75
RP	29.60	22.08	54.95	4.86	3.35	2.69	4.83	0.44
SCG	26.02	18.72	35.00	3.07	6.29	4.98	9.70	0.84
CGB	25.43	19.03	41.73	3.16	12.24	7.86	17.68	1.66
CGF	26.13	19.06	38.25	3.12	11.97	9.69	17.53	1.48
CGP	26.37	19.87	48.80	3.62	12.02	9.22	18.94	1.84
OSS	28.27	20.96	43.46	3.68	10.95	8.92	17.13	1.45
GDX	60.31	39.64	119.06	14.07	3.34	1.37	5.23	0.53

4.4 Conclusion

With the development of ANN technology, the sensitivity analysis is necessary to rank the importance of input variables due to a large amount of training data. In predicting the district heat demand using weather information, it is found that most input variables contain duplicated features which is not required to train the ANN model. This chapter proposed a method with the ability to remove both the variables with low importance and the variables that have high importance but contain duplicated features. The proposed approach analysed the correlation among input variables via detecting the coefficient of determination of each variable with others

referring to the fitting error of quadratic least square regression. The approach is validated in a case study of predicting heat demand in a district using an ANN model that is trained by historic data from several sample buildings. The traditional sensitivity analysis method ranked the input variables based on their influence on the heat demand. It was shown that the 15 most important features can be used to predict the district heat demand with the same or even better performance than the complete set of features. The proposed method further removed 3 important variables that are determined by other variables via analysing the determination of each variable. The results show that the proposed method can reduce training time by around 20% while achieving the same training and prediction performance compared with the traditional sensitivity analysis method. With the developed sensitivity and correlation analysis approach, the training data is simplified and the efficiency of training an ANN model can then be improved.

Chapter 5

Statistical Modelling of Electricity Demand

5.1 Data Processing and Modelling of Electricity Demand

5.1.1 Electricity demand data processing

In University of Glasgow, the electricity power is supplied by both the power grid and CHP from energy centre, as shown in Figure 5.1. The CHP consumes gas and generates heat and electricity for the buildings in campus.
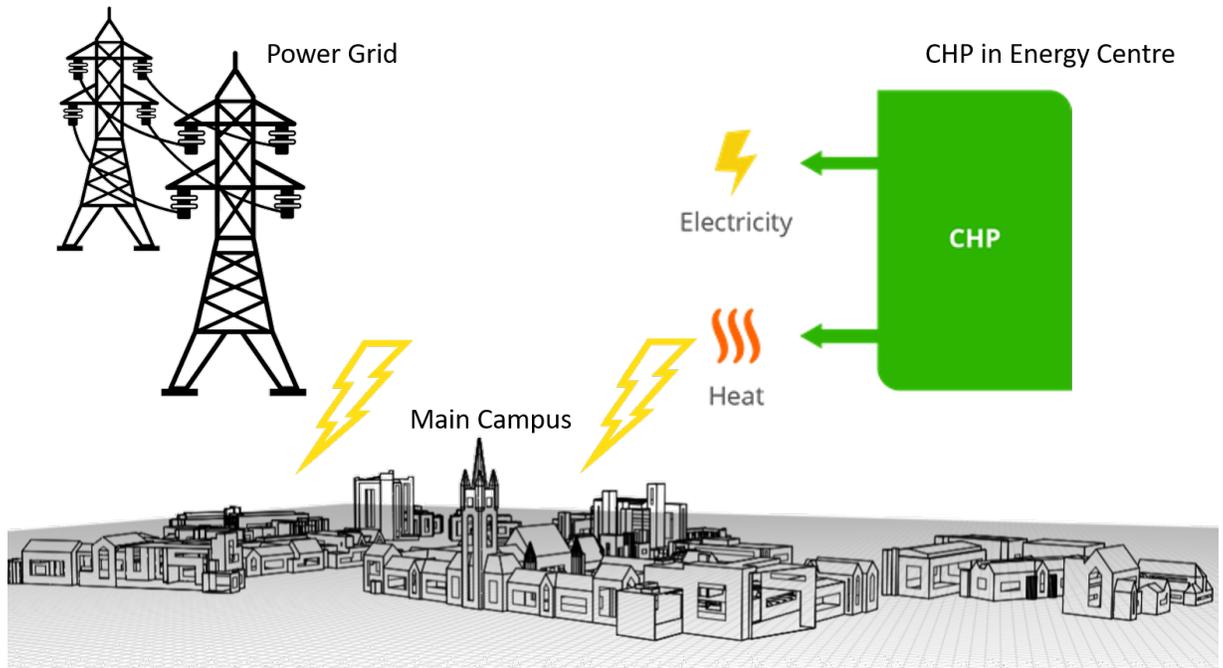
Figure 5.1: Schematic diagram of campus electricity from CHP and grid.

When the electricity demand of campus is less than the peak power output of CHP, all the power is supplied by the CHP and no power is from grid. If the electricity demand is higher than the peak power output of CHP, or in the summer, no heat demand is required from the CHP, the power is supplied from grid. The actual power consumption of main campus is the sum of

power from grid and power from CHP, as shown in Figure 5.2. The blue line shows the power from CHP, red line shows the power from grid.
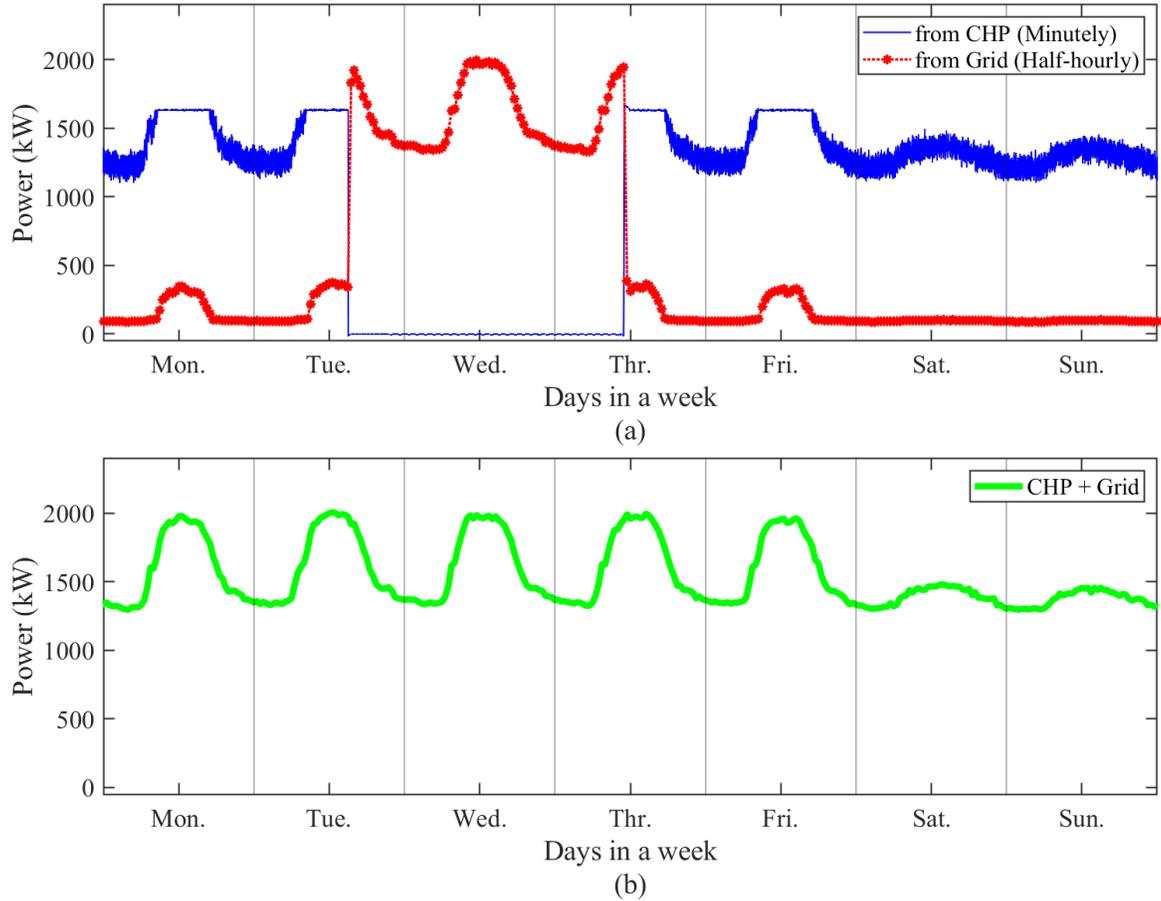
Figure 5.2: Campus electricity power demand from CHP and grid in a week.

However, in the data of power consumption, the CHP power output is recorded in every minute, while the grid power flow is recorded in every half-hour. That could cause the mismatch between data from CHP and data from grid. Figure 5.3 shows the example of one day. The blue line shows the power from grid, the red line shows the power from CHP and the green line shows the total power. In some period, the result shows obvious oscillation, which makes the result not reliable. In order to avoid the power oscillation and make the power demand more reliable, the data processing by moving the time series of CHP to compensate the time difference and match that of the power from grid. Before processing, the original data of grid and CHP has some mismatch on time points. After moving the data forward or backward to match their time points, the final result of the total power consumption has less oscillation, which is more reliable.

5.1.2 Problems in electricity demand data fitting

In matching the electric power to weather conditions, a problem is found that the relationship of power consumption and temperature are in different sign of proportional in short-term (hourly
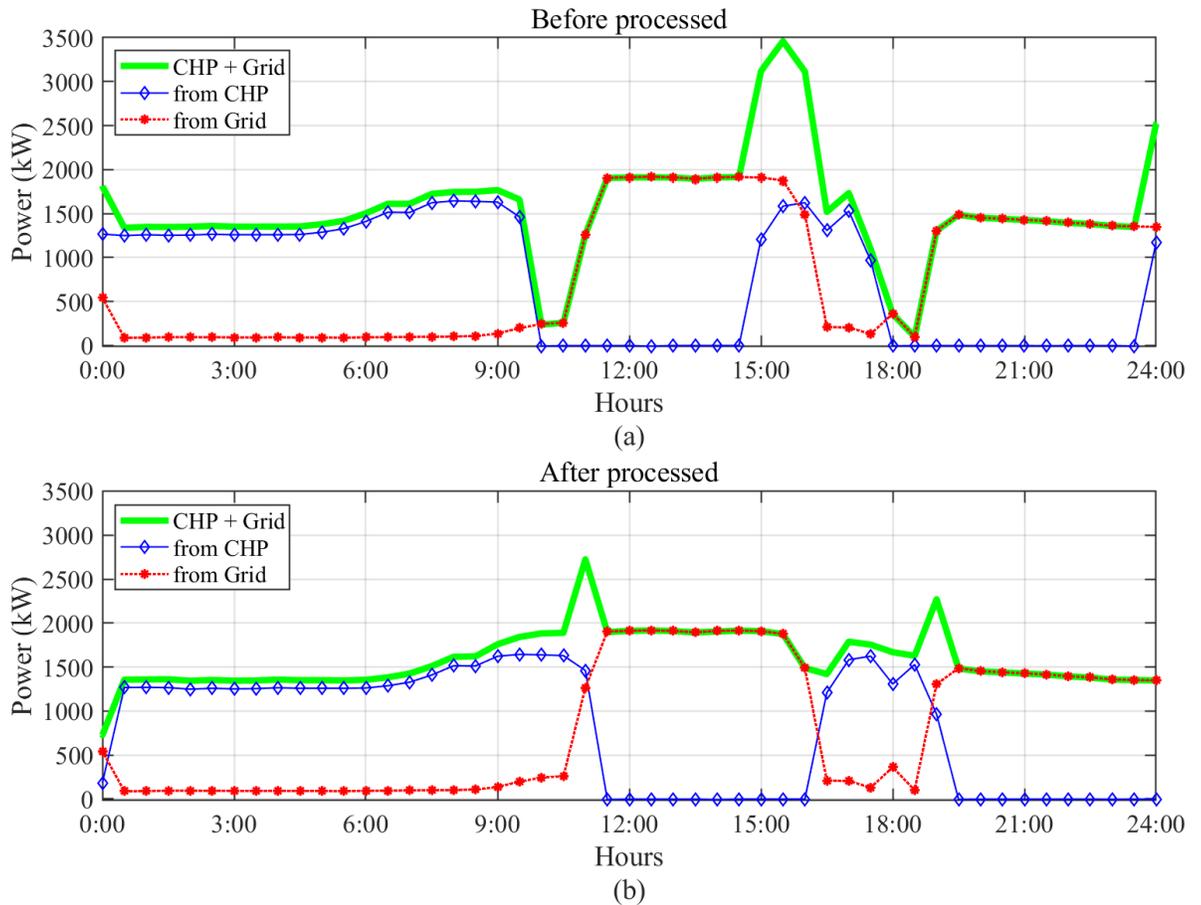
Figure 5.3: Data processing of electricity power consumption from grid and CHP

variation) and long-term (seasonal variation) time horizons. As in the example of the Maths &
Statisc School building, the comparison between ambient temperature and electricity demand is
given at the top of Figure 5.4. The total data of 1.5 years on the left shows that the relationship
between ambient temperature and electricity demand is in obvious inverse proportional (nega-
tive correlation). If enlarge the comparison in days, their relationship is in direct proportional
(positive correlation), as shown in the figure on the right. The lower air temperature (blue curve),
the higher electricity demand (red curve). But if the comparison is enlarged into single days, as
shown in the figure on the right, their relationship is positive correlation. The electricity demand
is increasing and decreasing following the temperature.

The main reason of this situation can be explained as the impact of human activities of oc-
cupants in the building. For example, during working time, occupants are using electric devices
with the demand in an approximate Gaussian distribution. Thus, the power demand at noon is
the highest while the power demand at night is the lowest. This tendency of power demand in a
day matches the tendency of temperature, but it cannot say that the power demand is determined
by the temperature.

In a long-term period or seasonal variation, the HVAC (heating, ventilation, and air condi-
tioning) system is managed to heat the building continuously. Thus, the power demand is in
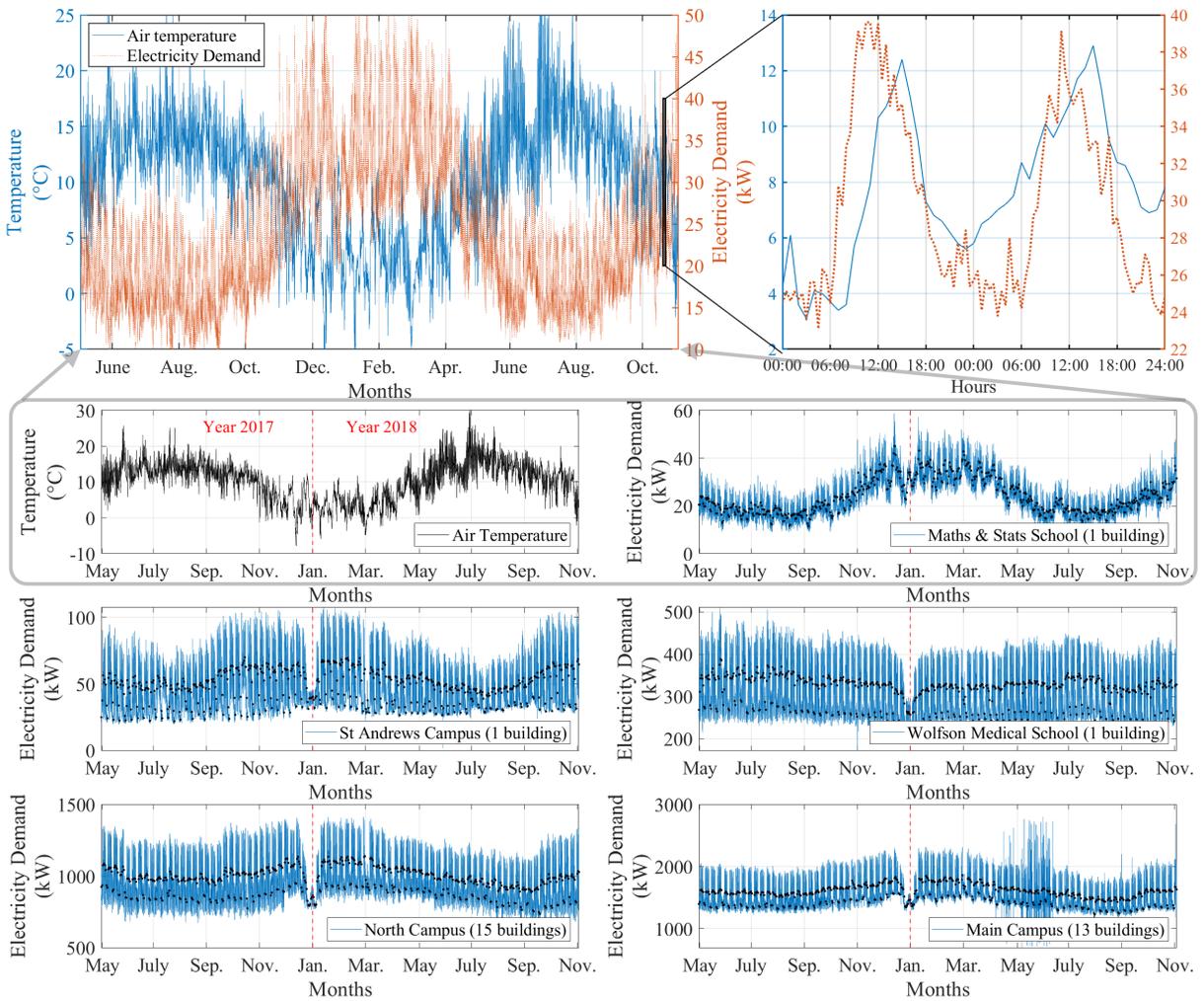
Figure 5.4: Half-hourly power demand of different campuses comparing with temperature.

inverse correlation with the temperature, or the lower temperature causes the higher power consumption for space heating. It shows obvious inverse correlation between electricity demand and temperature.

This is a typical issue in office/education buildings heated by electricity. It is more obvious in cold areas, such as in Scotland, that the heating load is much higher than the cooling load. In the opposite condition where the cooling load is much higher, the inverse correlation between electricity demand and temperature in the long term is not as obvious as the issue shown above. This chapter focuses only on fitting the data of electricity demand of buildings where the heating demand is much higher than cooling. Therefore, in a longer time period, the temperature and electricity demand are in inverse correlation, while in a shorter time period, the temperature and electricity demand are in positive correlation. That causes difficulties in data fitting using traditional statistical methods of fitting the power demand with temperature directly.

The main reason of this situation can be explained as the impact of human activity of occupants in the building. For example, during working time, occupants are using the electric devices with the demand in an approximate Gaussian distribution. Thus, the power demand at noon is
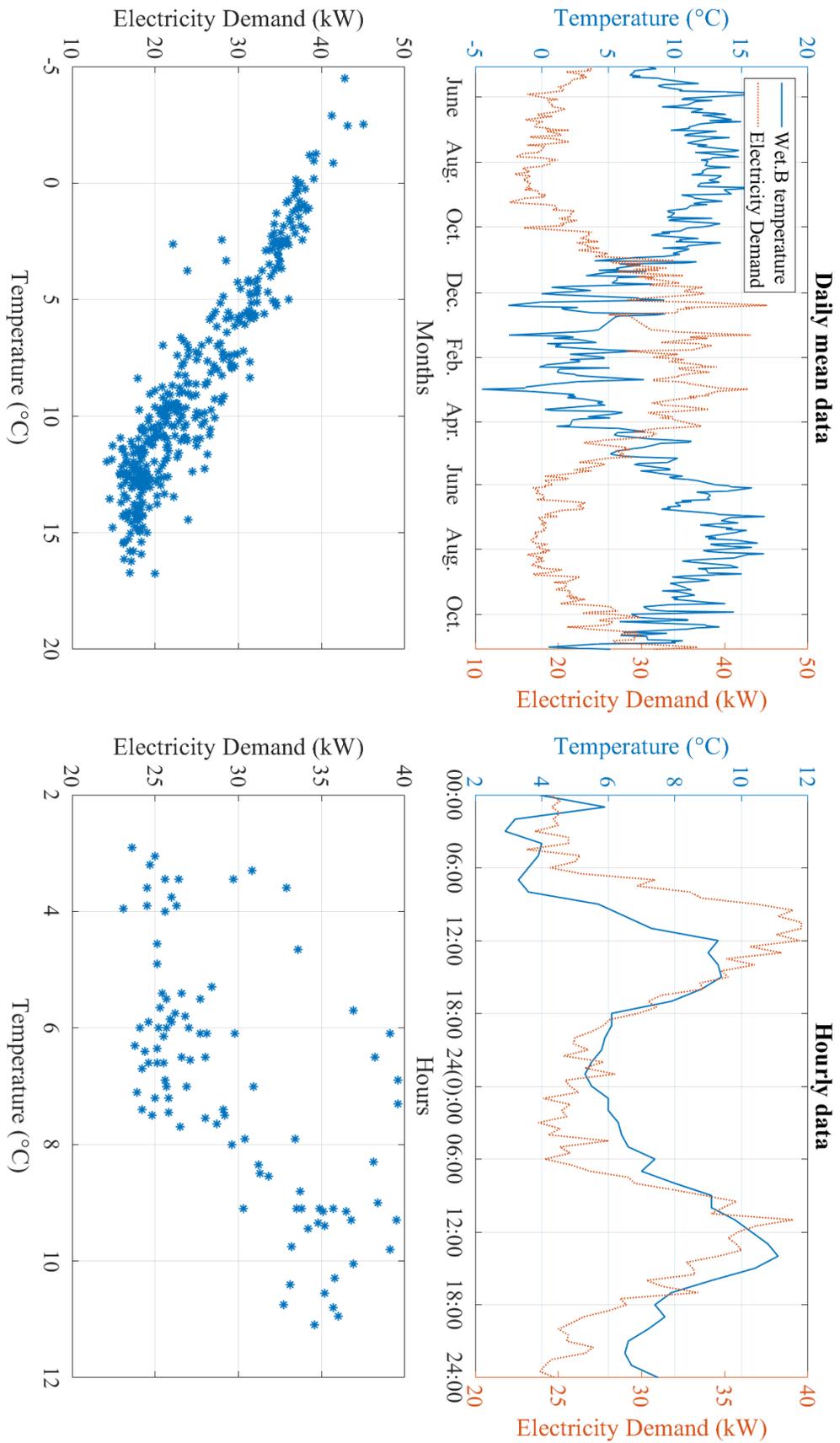
Figure 5.5: Comparison between daily mean power demand and half-hourly demand.

the highest while the power demand at night is the lowest. This tendency of power demand in a day matches the tendency of temperature. But it cannot say that the power demand is determined by the tempertaure in a direct proportion.

In a long-term time horizon, the building management system leads a higher ratio in power consumption of electricity, such as the HVAC (heating, ventilation, and air conditioning) system. Thus, the power demand is inversely proportional with the temperature, or the lower temperature causes the higher power consumption for space heating. It shows obvious inverse porportional relationship between electricity demand and temperature. So in a longer time horizon, the temperature and electricity demand are in negative relationship, while in shorter time horizon, the temperature and electricity demand are in positive relationship. That causes difficulties in data fitting using statistical method.

To see the relationship between temperature and electricity demand, the daily mean data can be found with a more obvious inverse proportional relationship. The daily mean power demand can be fitted easily with linear regression. But the hourly power demand can be found with a relatively direct porportional relationship, as shown in Figure 5.5. This difference causes the difficulty in data fitting between power consumption of electricity and weather conditions, such as temperature.

5.2   Development of Electricity Demand Prediction Approach

5.2.1   Working hour splitting based regression approach (Approach 1)

The approach is designed to split the building baseload power consumption and occupants' activities by time series. In order to achieve this, it is assumed that the occupants only consume power at particular period of time while building baseload power is consumed 24 hours continuously. In the time period when the occupancy rate is low or zero, the power demand is mainly the baseload of the building; and in the time period when the occupancy rate is high, the power demand data includes both the building baseload and occupants' activities. It is difficult to separate them from the recorded data. Therefore, the approach uses the power demand data of the low occupancy time period to find the dependence on environmental conditions to adapt to the building baseload, which is independent with occupants' activities. After that, the building baseload in the remaining time period can be predicted by the weather variables of the same period and the fitted dependence above. It is assumed that the difference between the actual power demand and the building baseload during this period is caused by the influence of the occupants' activities.

In the case study, the target buildings at university campus are used for teaching and officing. Thus, the normal working hours are between 9:00 and 17:00, depending on the work schedule. However, the actual working hours are flexible for employees, depending on their preference
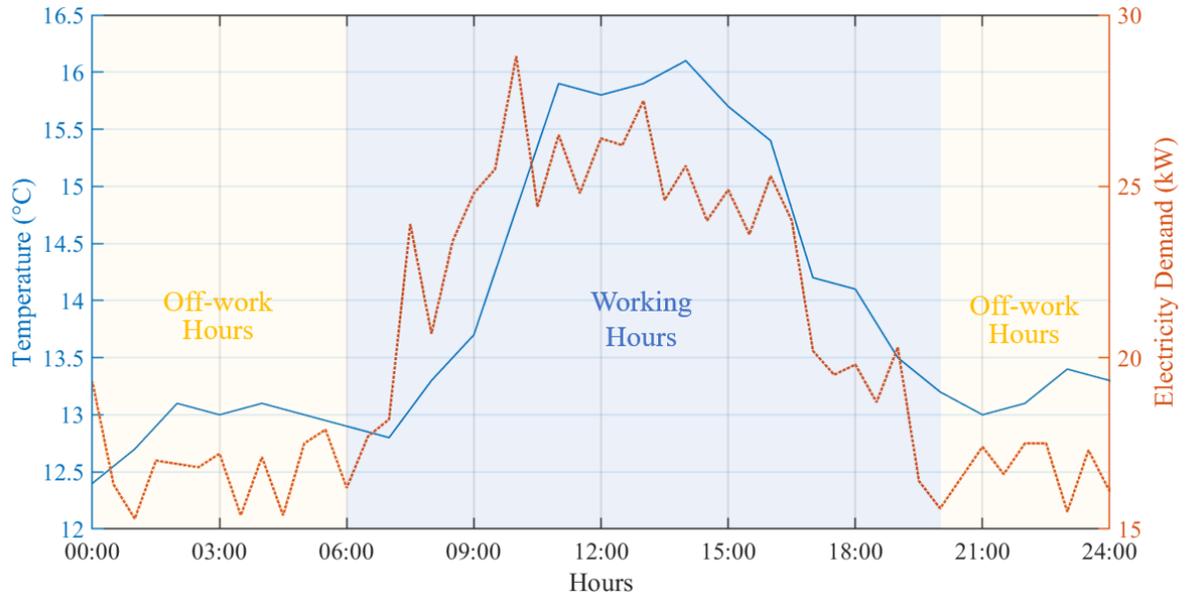
Figure 5.6: Spliting the time into working hours and off-work hours of every 24 hours.

on working time. After comparing between the electricity demand data of normal working days and public holidays, it is found that the impact period of occupant's power consumption is several hours wider than the normal working hours. Therefore, in order to further reduce the uncertainty of occupants' activities, a three hours redundancy period has been added before and after the normal working hours to eliminate possible occupants' activities during non-working hours. The data between 20:00 every day and 6:00 the next day is defined as the non-working hours and used to fit weather variables to electricity consumption without occupants' activities, as shown in Figure 5.6.

In the approach, the non-working hours power demand data is used to fit the most significant weather variables as the building baseload power demand. To choose the most significant weather variable for fitting the non-working hours power demand, the sensitivity analysis technique is used to find which weather variable has the highest sensitivity to the power demand. The sensitivity analysis uses the coefficient of determination, R_k^2, of the kth model input variable as the index of showing the quality of each weather variable to the electricity demand as

$$R_K^2 = 1 - \frac{\sum_i \left(y_i - x_{k,i}\right)^2}{\sum_i \left(x_{k,i} - \bar{x}_k\right)^2} \tag{5.1}$$

where $x_{k,i}$ and y_i indicate the ith sample points data of the kth model input and model output, respectively; $\bar{x}_k$ indicates the mean of model input.

After the weather variable with highest sensitive to power demand is found, their relationship is found using the simplest statistical method (i.e., the linear regression method) to predict the building baseload power demand. The difference between the recorded real power demand and fitted baseload power is then known as the power demand determined by occupants' activities.
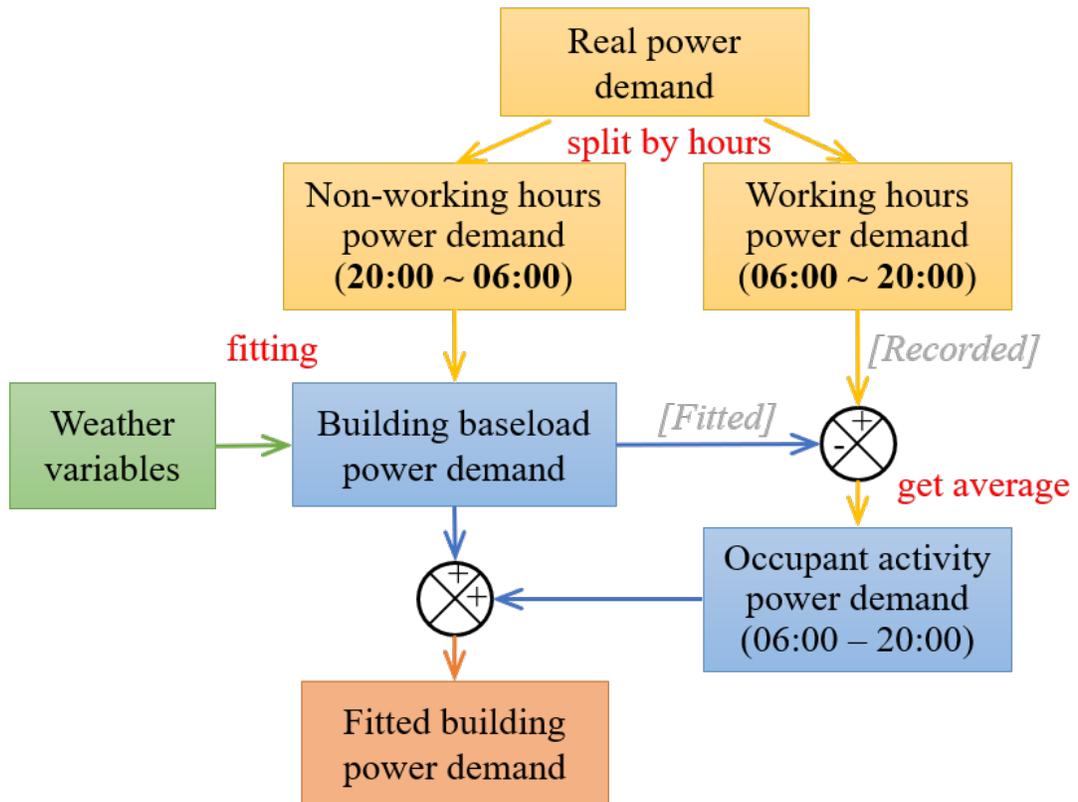
Figure 5.7: Flow chart of electricity power demand fitting approach.

It is assumed that the power demand of occupants' activities in each hour satisfies a normal distribution. Therefore, at the same hour on different dates, the mean value of the difference between the recorded power demand and the fitted building baseload is used to predict the most possible power demand caused by occupants' activities. After getting the average of human activities power demand caused by building occupants, the sum of power demand caused by fitted building baseload and occupant activities are known as the fitted building power demand. The whole process is produced as the flowchart in Figure 5.7.

5.2.2 ANN with fuzzy hours splitting approach (Approach 2)

The design above is the basic method to verify the approach of splitting the data by time of working hours. However, the linear regression method to fit the electricity demand and ambient temperature cannot fully capture the nonlinear dynamics of multi weather variables. In addition, the average occupant power demand ignored a lot of information that causes the power variation of different human activities. In order to improve the power demand prediction approach, an approach is developed with artificial intelligent technologies.

In the original design, the splitting method is based on a fixed time from normal working hours. However, this splitting approach is too arbitrary and does not consider the possible changes of different conditions. In the new approach, the full-occupancy hours are defined from
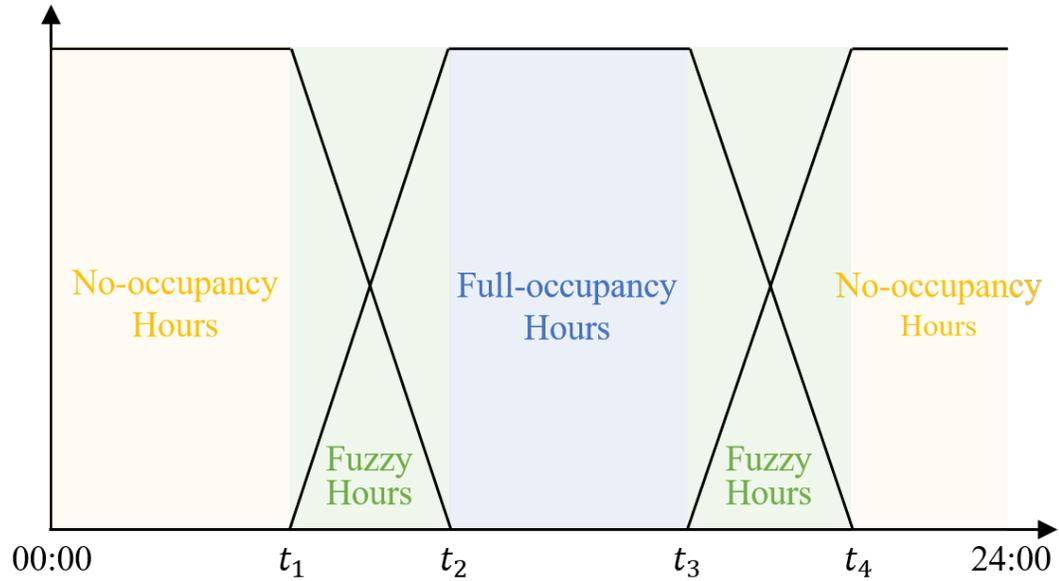
Figure 5.8: Spliting the time into working hours and off-work hours by fuzzy logic.

t_2 to t_3 while the no-occupancy hours are defined from t_4 to t_1. The area between full-occupancy hours and no-occupancy hours are defined as the fuzzy area that is not quite clear to belong to the full-occupancy or no-occupancy, as shown in 5.8.

The power demand can be split into three sections by time as

$$P_{\text{building}}(t) = \begin{cases} P_{no} & t_4 \leq t < t_1 \\ P_{fuzzy} & t_1 \leq t < t_2 \& t_3 \leq t < t_4 \\ P_{full} & t_2 \leq t < t_3 \end{cases} \tag{5.2}$$

Assume the membership function of the fuzzy area is $f(t)$, which indicates the occupancy rate. The power demand of the fuzzy hours $P_{fuzzy}(t)$ can be obtained from the power demand of no-occupancy power, the full-occupancy power and the membership function as

$$P_{\text{fuzzy}}(t) = P_{\text{no}}(t) \cdot (1 - f(t)) + P_{\text{full}}(t) \cdot f(t) \tag{5.3}$$

where $P_{\text{no}}(t)$ is the power demand of time during no-occupancy hours, which also indicates the building baseload power demand without the impact of the occupants. $P_{\text{full}}(t)$ is the power demand of time during full-occupancy hours, which includes the power demand at peak load caused by full-occupancy activities.

The membership functions of fuzzy area can be designed in different shapes. The general classification of membership functions can be piecewise linear function, Gaussian function, bell-shaped function, sigmoid function, and polynominal based function [205].

The piecewise linear function constitute the simplest type of membership functions by two

parameters c_1 and c_2 with changing of variable x as

$$f(x;c_1,c_2) = \begin{cases} 0, & x \leq c_1 \\ \dfrac{x-c_1}{c_2-c_1}, & c_1 < x < c_2 \\ 1, & x \geq c_2 \end{cases}$$ (5.4)

where the parameters c_1 and c_2 locate at the bottom and top of the curve, respectively.

The Gaussian membership function is given by the expression as

$$f(x;\sigma,c_2) = \begin{cases} e^{\frac{-(x-c_2)^2}{2\sigma^2}}, & x < c_2 \\ 1, & x \geq c_2 \end{cases}$$ (5.5)

where the σ indicates the width of the membership function curve.

The bell shaped membership function is given by the expression as

$$f(x;a,b,c_2) = \begin{cases} \dfrac{1}{1+\left|\frac{x-c_2}{a}\right|^{2b}}, & x < c_2 \\ 1, & x \geq c_2 \end{cases}$$ (5.6)

where parameter a indicates the width of the membership function, b is a positive that indicates the orders of the curve.

The sigmoidal function is generally expressed as

$$f(x;a,c_1,c_2) = \frac{1}{1+e^{-a(x-\frac{c_1+c_2}{2})}}$$ (5.7)

where parameter a indicates the speed and shape of the curve depending on its value and sign.

The curves of linear, Gaussian, Bell-shaped and Sigmoidal membership functions from Eq. (5.4) to (5.7) are shown in Figure 5.9.

In the fitting approach between weather variables and power demand, the working hour splitting based regression approach uses linear regression to find the relationship between non-working hours power demand and the most sensitive weather variable. That approach assumes that the demand of building baseload is in linear relation to a single weather variable. The impact of other less sensitive weather variables and other less dominate nonlinear relationships are ignored. Therefore, in the new approach, the artificial neural network (ANN) method is used to find the nonlinear relationship from multiple weather variables to power demand of no-occupancy hours and full-occupancy hours, respectively. Similar with the regression approach, the sensitivity analysis referring to the coefficient of determination is required to rank the weather variables based on their sensitivity. But different with the regression approach above that only choose one most significant variable, the ANN based approach chooses more weather variables for fitting their nonlinear relationships with power demand. Based on the result of sen-
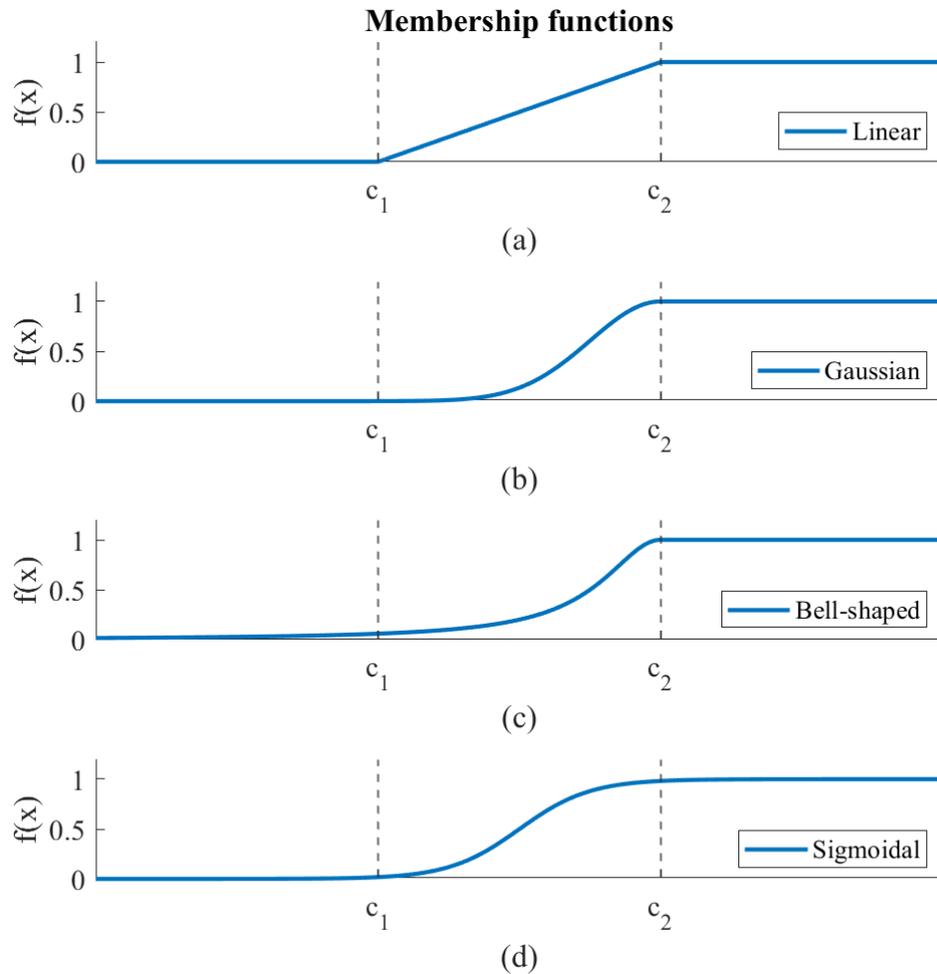
Figure 5.9: Commonly used membership functions of fuzzy logic.

sitivity analysis, the number of input neurons in ANN can be defined according to the number of weather variables with high sensitivity to the power demand.

The process of ANN can be described in mathematical formulas. Define x_k $(k = 1, 2, \cdots, n)$ as the k-th input attribute value, which is passed along the links to the other layers. The weighted sum of signals, Σ, arriving at the input of the next neuron is subjected to a transfer function, which is the most commonly used 'sigmoid' function as

$$f(\Sigma) = \frac{1}{1 + e^{-\Sigma}} \qquad (5.8)$$

The jth hidden neuron h_j $(j = 1, 2, \cdots, p)$ receives the sum of neuron value multiplied by the weights $w_{kj}^{(2)}$ and bias $b_{kj}^{(2)}$ associated with the link as

$$h_j = f\left(\sum_{k=1}^{n} w_{kj}^{(2)} x_k + b_{kj}^{(2)}\right) \qquad (5.9)$$

The output neurons are defined as y_i $(i = 1, 2, \cdots, m)$, which are summed up with their input

signals and activation transfer function as

$$y_i = f \left(\sum_{j=1}^{p} w_{ji}^{(1)} f \left(\sum_{k=1}^{n} w_{kj}^{(2)} x_k + b_{kj}^{(2)} \right) + b_{ji}^{(1)} \right) \tag{5.10}$$

where $f(\cdot)$ is the activation function, the sigmoid function used in the chapter; $w_{ji}^{(1)}$, $b_{ji}^{(1)}$, $w_{kj}^{(2)}$ and $b_{kj}^{(2)}$ are the weights and bias linked to the output layer (1) and hidden layer (2), respectively. This is a typical two-layer ANN model with an output layer and one hidden layer.

The training error is used to update the ANN parameters of weights and bias of each neurons in the hidden and output layers. The training based on backpropagation (BP) learning algorithm is adopted to a typical two-layer ANN model to search for the global optimum as

$$w_{ji}^{(1)} = w_{ji}^{(1)} + \eta \delta_i^{(1)} h_j \tag{5.11}$$

$$w_{kj}^{(2)} = w_{kj}^{(2)} + \eta \delta_j^{(2)} x_k \tag{5.12}$$

$$b_{ji}^{(1)} = b_{ji}^{(1)} + \eta \delta_i^{(1)} \tag{5.13}$$

$$b_{kj}^{(2)} = b_{kj}^{(2)} + \eta \delta_j^{(2)} \tag{5.14}$$

where η is the learning rate of the BP based ANN.

The original design assumes that the power demand of occupants' activities only related with the working time and calculates the average occupants' power consumption by subtracting the fitted building baseload from the recorded full power demand. However, the occupant activities can somehow be affected by different weather variables. The new approach uses the ANN to find the occupancy rate in fuzzy hours, in which the number of occupants is varying from no occupants to full occupants. Thus, the occupancy rate of fuzzy hours is described by membership function between no-occupancy and full-occupancy and is fitted with the weather variables in ANN. Assuming the fuzzy hours are between t_1 and t_2, the fitted fuzzy area membership function can be obtained by

$$\overline{f}(t)\big|_{t_1}^{t_2} = \min \left(\max \left(\frac{P_{\text{fuzzy}}(t) - \overline{P}_{\text{no}}(t)}{\overline{P}_{\text{full}}(t) - \overline{P}_{\text{no}}(t)}, 0 \right), 1 \right) \tag{5.15}$$

where $P_{\text{fuzzy}}(t)$ is the real power demand of fuzzy hours for training, $\overline{P}_{\text{no}}(t)$ indicates the predicted baseload power demand in fuzzy hours using the ANN trained by no-occupancy hours data. $\overline{P}_{\text{full}}(t)$ indicates the predicted full occupants power demand in fuzzy hours. The power demand data of no-occupancy hours and full-occupancy hours is used to train the ANN model. After it is well trained, the model can be used to calculate the no-occupancy power demand and full-occupancy power demand. Therefore, in the fuzzy hours, its no-occupancy power demand and full-occupancy power demand can be predicted and used as the upper and lower limit. Its real value can be obtained using both the predicted limits and the fitted membership function

of occupancy rate, $\overline{f}(t)\big|_{t_1}^{t_2}$, in fuzzy hours between t_1 and t_2. The fitted membership function shows the occupancy rate between 0 and 1. If the membership value is 0, no occupant activity affects the power demand, and the building power demand only includes the baseload power demand. If the membership value is 1, the power demand is impacted by full occupants. The membership value between 0 and 1 indicates the ratio of current occupants to the full occupants. Different with Approach 1 that assumes the occupancy rate is the same in all weekdays or weekends, the fitted occupancy rate in Approach 2 depends on both the weather variables and time horizon. Thus, Approach 2 covers the uncertainties of occupancy rate caused by weather and better predicts the real occupancy rate of target buildings.

In the ANN design, it combines three ANN together. The baseload power demand is fitted using m nodes and trained with the data of no-occupancy hours. The full-occupancy power demand is fitted using additional n_2 nodes and trained with the data of working hours as the full-occupancy power demand is calculated by both the n_1 nodes for baseload and n_2 nodes for occupants. The membership function of occupancy rate in fuzzy hours for training is calculated using Eq. (5.15) with the real power demand in fuzzy hours and the fitted full-occupancy power demand from $n_1 + n_2$ nodes. The membership function is fitted to time series and weather variables using another n_3 nodes. In summary, the ANN is developed to use the weather variables and time series as inputs and the power demand data split into no-occupancy hours, full-occupancy hours and fuzzy hours, as shown in Figure 5.10.

The power demand of the fuzzy area ($t_1 \sim t_2$ and $t_3 \sim t_4$) is calculated according to the trained membership function of occupancy rate. Then the predicted power demand from 00:00 to 24:00 is compared with the recorded result to validate the approach.

The final artificial intelligent based approach is developed as the flowchart shown in Figure 5.11. The real power demand data is split the full-occupancy hours ($t_2 \sim t_3$) and no-occupancy hours ($t_4 \sim t_1$) to fit with weather variables, respectively. The power demand data of fuzzy hours ($t_1 \sim t_2$ and $t_3 \sim t_4$) is then used to calculate the occupancy rate using Eq. (5.15) and fitted with weather variables as well. At last, the fitted building power demand is calculated from the predicted no-occupancy power, full-occupancy power and predicted occupancy rate from 00:00 to 24:00 from the ANN for validation.

5.3 Simulation Result of Electricity Demand Prediction

In the case study of electricity consumption of the University of Glasgow, most buildings do not have individual power meters to record their electricity usage. The buildings with individual power meter to record their electricity demand are the Maths & Stats school, the St Andrews building, and the Wolfson medical school. Other buildings are included in the two campuses, the North campus and Main campus, as shown in Figure 5.12. The Maths & Stats school building is used as the target building to develop the approach for electricity demand prediction.
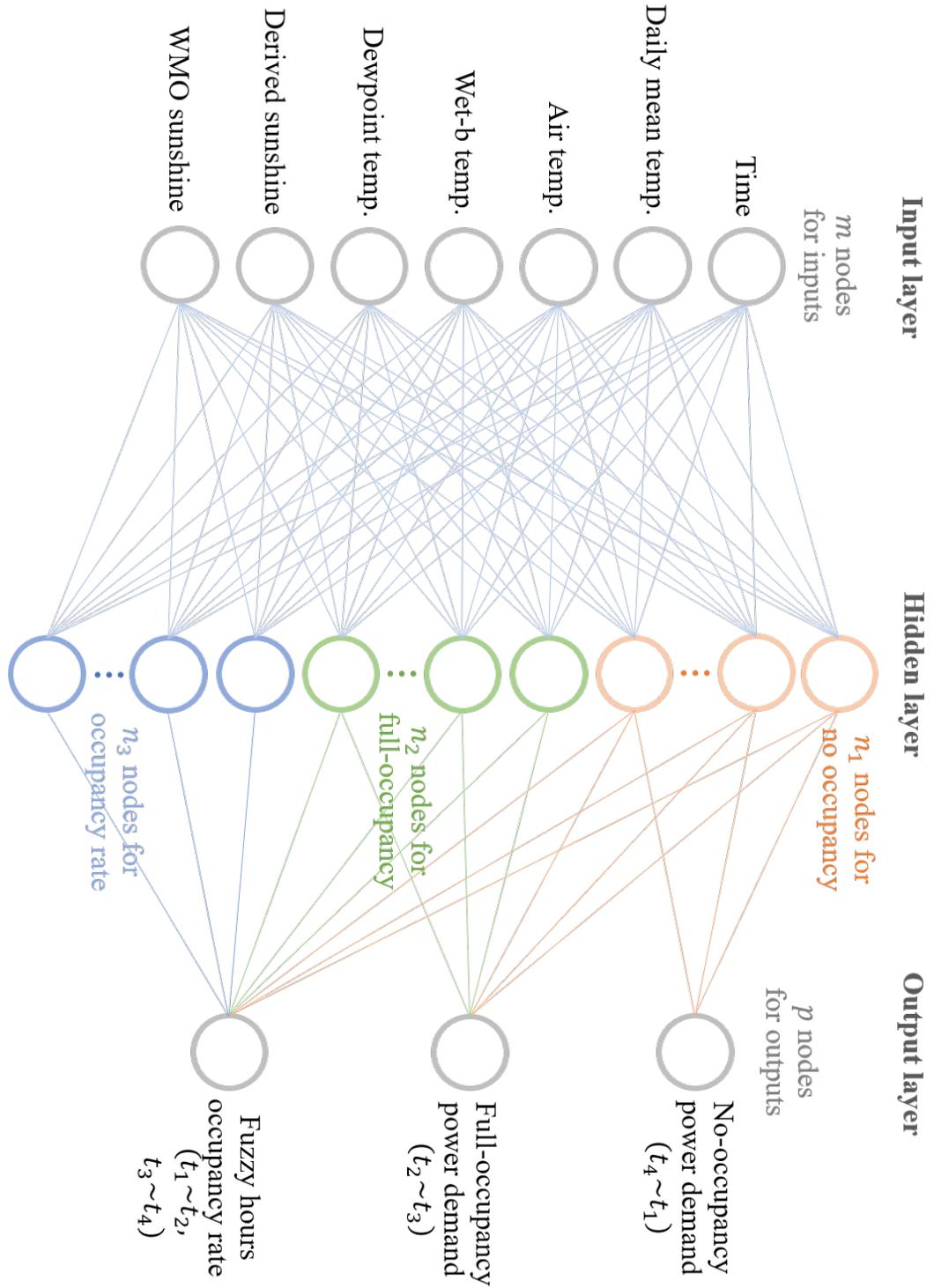
Figure 5.10: Artificial neural network for fitting the weather conditions and off-work hours and working hours power demand.
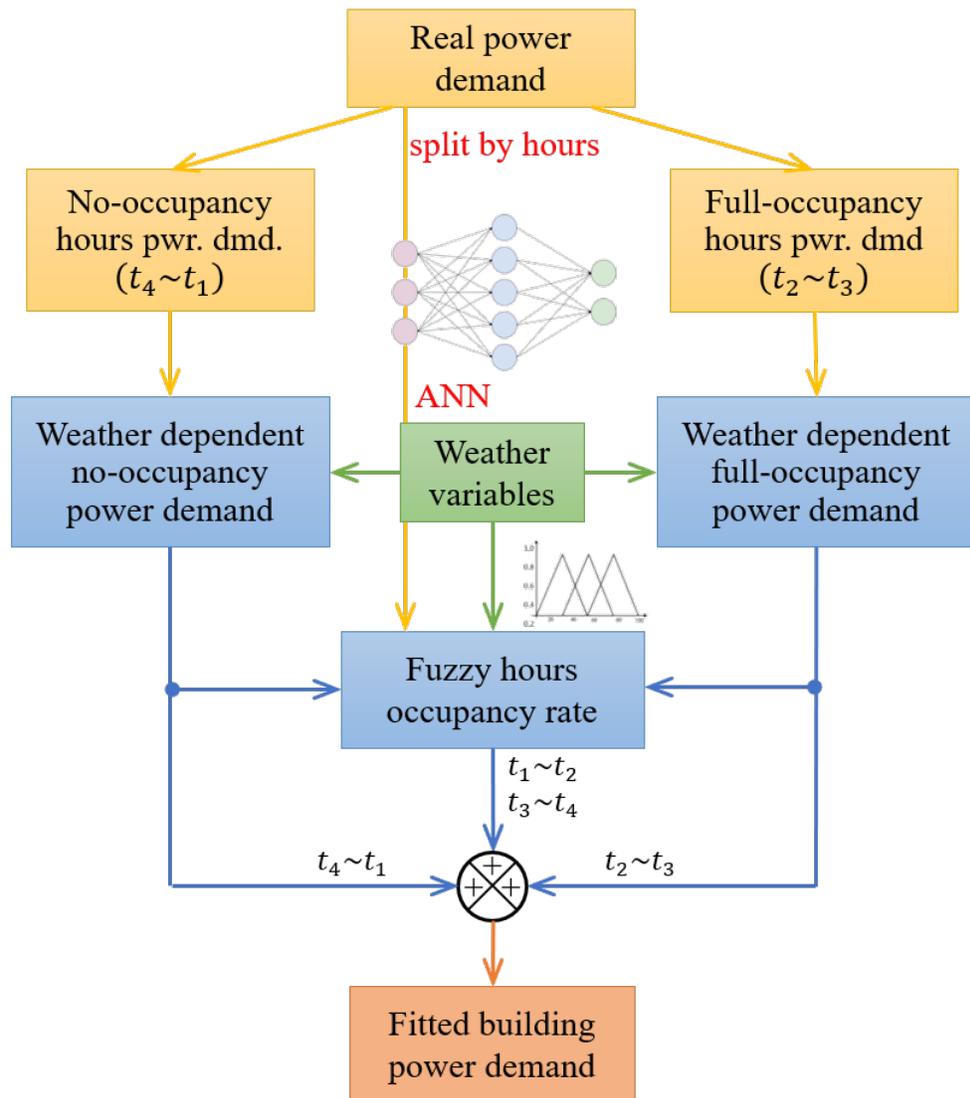
Figure 5.11: Embed the developed electricity power demand fitting approach with artificial intelligent techniques.
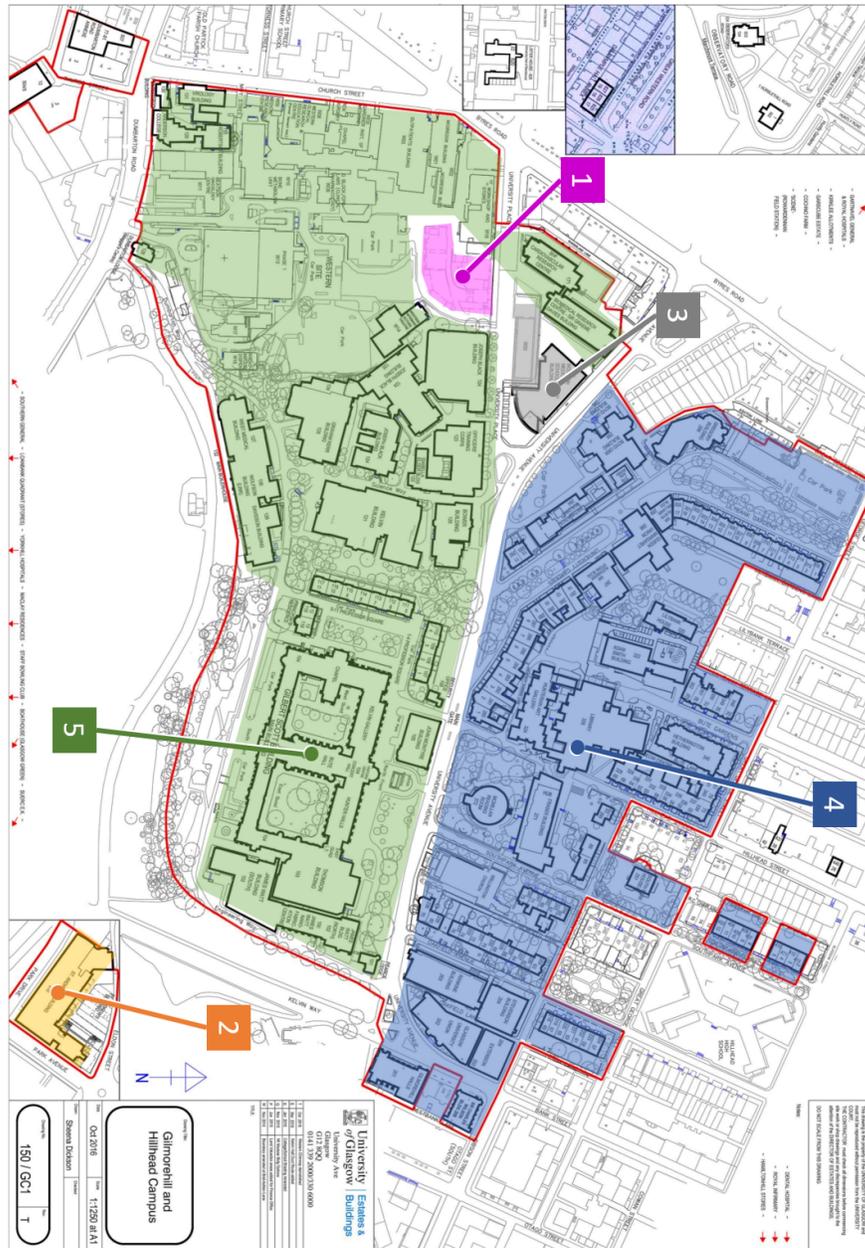
1 Math & Stats School

2 St Andrews Campus

3 Wolfson Medical School

4 North Campus

5 Main Campus

Figure 5.12: Campuses of the University of Glasgow with individual electricity demand data.

5.3.1   Fitting of electricity demand

The data used in the case study included 13 weather variables recorded by the local weather station and the electricity consumption of whole campus recorded by the university energy centre. The weather variables used for fitting the whole hours power demand include the dry-bulb or air temperature, the wet-bulb temperature, the dew-point temperature, the daily mean temperature, the derived sunshine, the WMO sunshine (measured by World Meteorological Organization), the wind speed, the wind direction, the relative humidity, the station pressure, the mean sea-level pressure, the visibility, and the cloud base height.

The weather conditions used for fitting the whole hours power demand include the dry-bulb temperature, the wet-bulb temperature, the dew-point temperature, the daily mean temperature, the maximum adaptive temperature, the global radiation, the direct radiation, the solar altitude, the moisture content and relative humidity. The comparison of each weather condition with the building total power demand in a week is shwon in Figure 5.13. In the figure, the blue curve shows the power demand of electricity while the red curve shows the corresponding weather condition. In the comparison, one can find that the weather condition with the highest sensitivity to the power demand is the maximum adaptive temperature and daily mean temperature. These two weather conditions are both the constant index for a single day. That means the short-term data and long-term data are in opposite proportional relationships and it makes the hourly weather information useless in the whole time range fitting method.

In the sensitivity analysis, each weather variable is analysed with the electricity demand with R^2. The data is recorded in every hour for a total of 549 days from 1st May 2017 to 31st October 2018. Therefore, the full hours power demand including the total of 13176 hours data points and the non-working hours power demand including just 4941 hours data point are used in the sensitivity analysis. The comparison between analysing full hours power demand and non-working hours power demand is shown in Figure 5.14. In the result, the blue bar indicates the sensitivity of each weather variable to non-working hours power demand, where the air temperature is higher than other weather variables. The red bar indicates the sensitivity of each weather variable to 24 hours power demand, where the daily mean temperature is obviously higher than other weather variables and the sensitivity of all weather variables is less than 0.5 in R^2. This verifies that the effectiveness of the proposed approach that the non-working hours power demand has higher sensitivity to the electricity demand and the impact of occupant's activities in working hours can reduce the sensitivity and causes more uncertainties.

The relationship between weather condition and off-work hours power demand is fitted in linear regression. As in the sensitivity analysis, the wet-bulb temperature has the highest sensitivity to the off-work hours power demand. Thus, the linear regression method is used to find the linear proportion between wet-bulb temperature and off-work hours power demand.

In the results of sensitivity analysis, the air temperature has the highest sensitivity to the non-working hours power demand. Thus, the linear regression method is used to find the linear
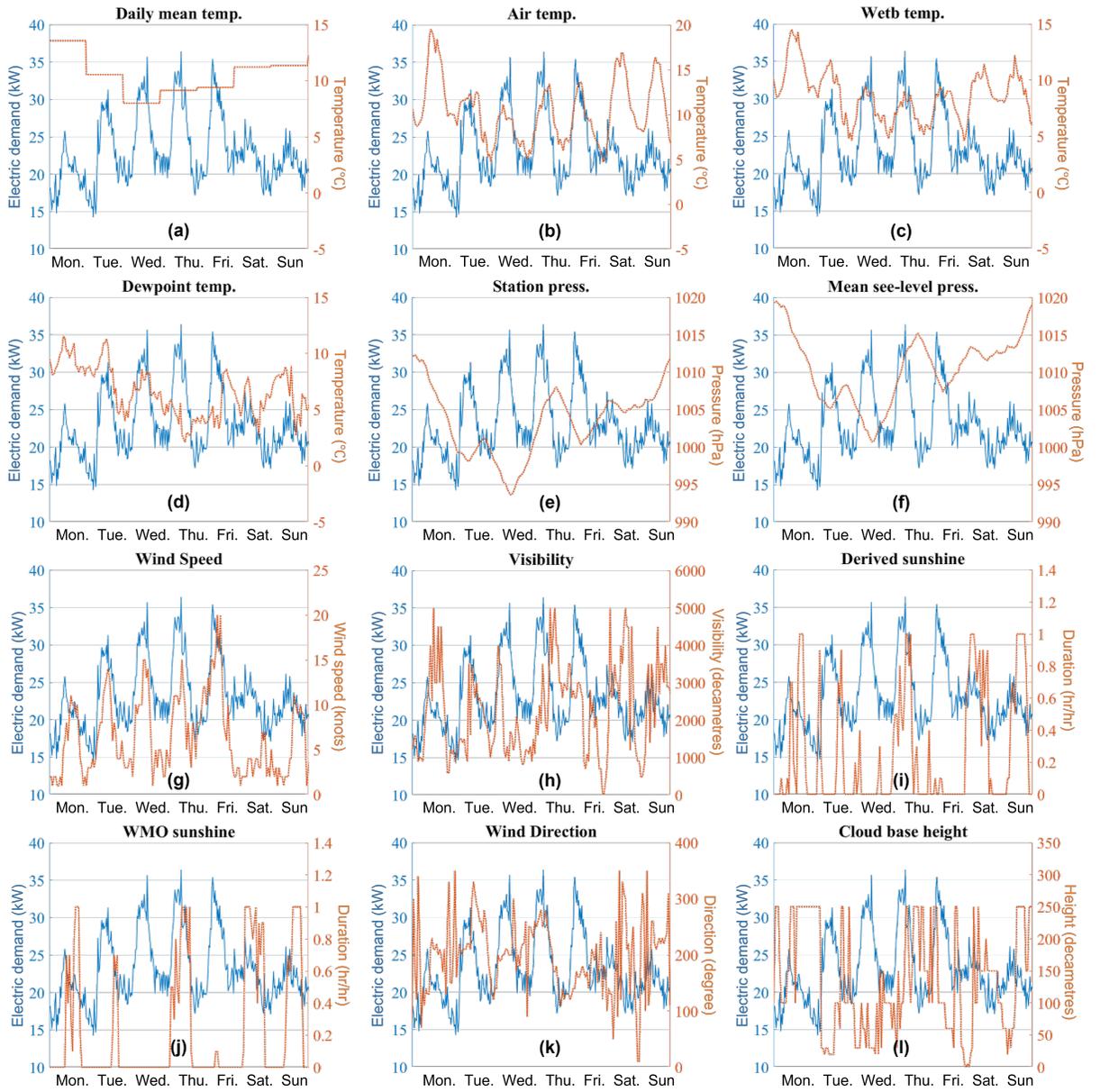
Figure 5.13: Example of electricity consumption to weather conditions in one week.
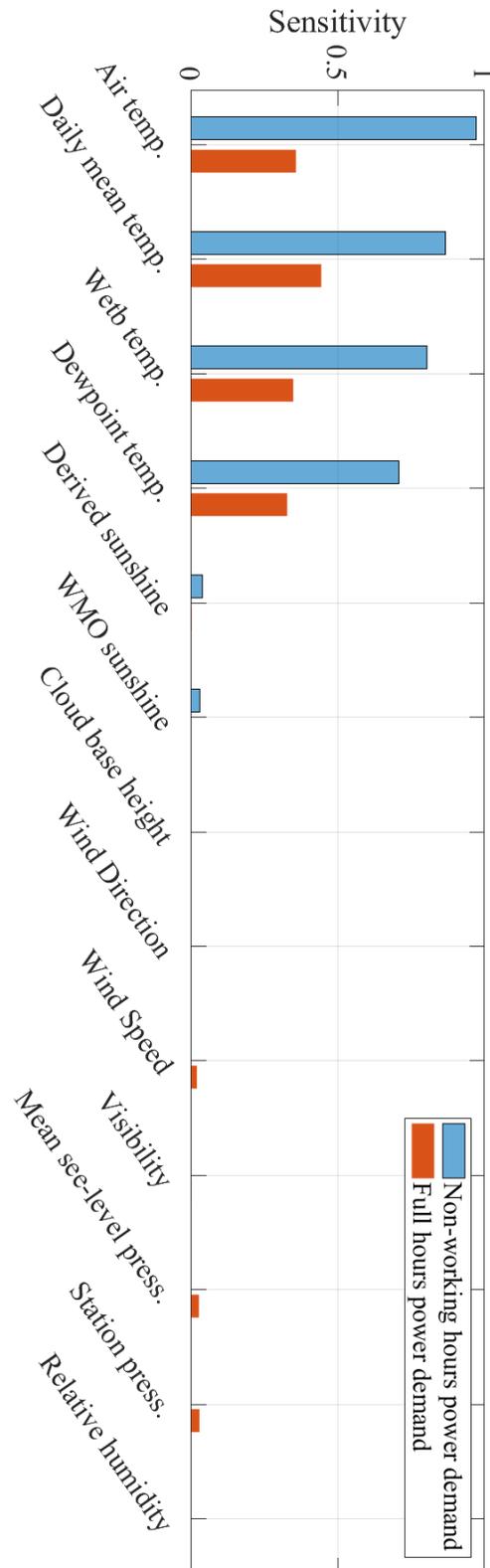
Figure 5.14: Sensitivity analysis of each weather condition to off-work hour power demand.

proportion between air temperature and non-working hours power demand. The fitted power demand with linear regression is defined as the building baseload power demand. The difference between real power and fitted baseload power is known as the occupant's activities and its average power demand can be obtained from the boxplot shown in Figure 5.15. The occupant's activities are different between workdays and weekends/holidays. Therefore, the fitting of power difference is separated by workdays and weekends/holidays and getting their mean hourly power demand of occupants, respectively.

In the result of mean value, the peak power demand caused by occupants' activities in weekdays is about 15kW while that on weekends is only approximate 6kW. In this test, the power demand caused by human activities of occupants are assuming to be the same every workday or weekends. In the box plot, the red central segment indicates the median, the top and bottom edges of each box indicate the 1/4 and 3/4 percentiles, and the '+' symbol indicates the outliers. The uncertainties are obtained as the range between upper and lower limits to the median by ignoring the outliers. In the result, the average power demand caused by human activities of occupants has the uncertainties of up to ± 10 kW in both working days and non-working days. For example, at the time of 12:00, the occupants' power demand in weekdays is 15 ± 10 kW while that in weekends is 5 ± 10 kW. The negative power demand just presents the difference between the average fitted power demand and real value. This regression approach mainly considers the varying of occupancy with time, and its relation to weather variables is not considered in the regression for keeping the simplification of this approach. The varying of occupancy behaviour caused by weather variables is considered as uncertainties in this approach. But this approach has a limitation that the relationship between occupancy behaviour and weather variables is not considered and fitted.

After the fitting of both building baseload power demand and occupants power demand, their sum is known as the total power demand of the target building and it is then compared with the real power demand. Figure 5.16 shows the results of the comparison between real power demand (blue solid line) and fitted power demand (red dashed line). Their difference is known as the fitting error as shown below. In the fitting results, the fitted power demand matches well with the recorded real power demand. The average fitting error is around 18%.

The fitted power demand is the sum of weather dependent baseline power and the power caused by the occupants' activities. The example of fitted power demand in one week is shown on the right of Figure 5.16. The blue dashed line shows the temperature. The blue shadow shows the fitted baseload power, which is in negative correlation to temperature. The yellow shadow shows the fitted power caused by occupants' activities. The red line shows the final fitted power demand while the black pointed line shows the real power demand for comparison.

However, due to the power of the occupants' activities is fitted using the mean value, it causes more fitting error of the unpredictable human activities. Therefore, using ANN to find the relationship between the power demand caused by occupants and the weather variable in the
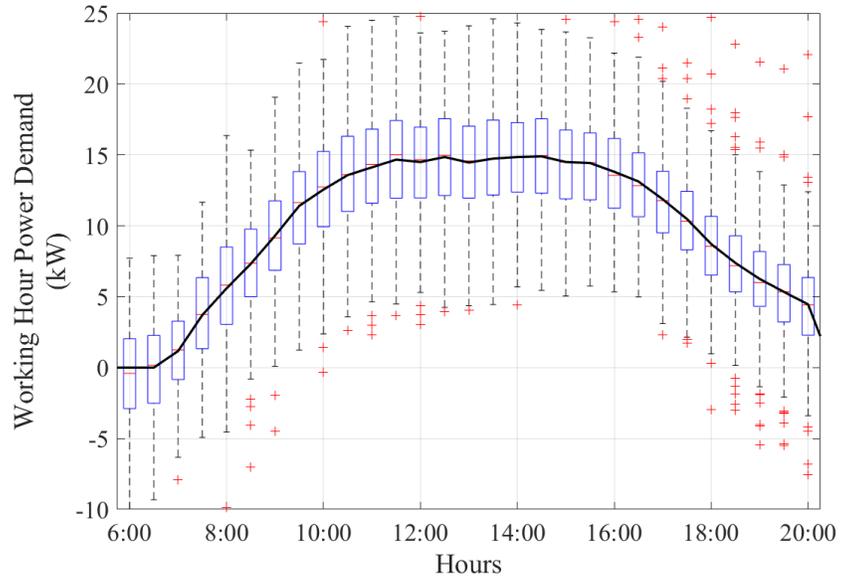
(a)
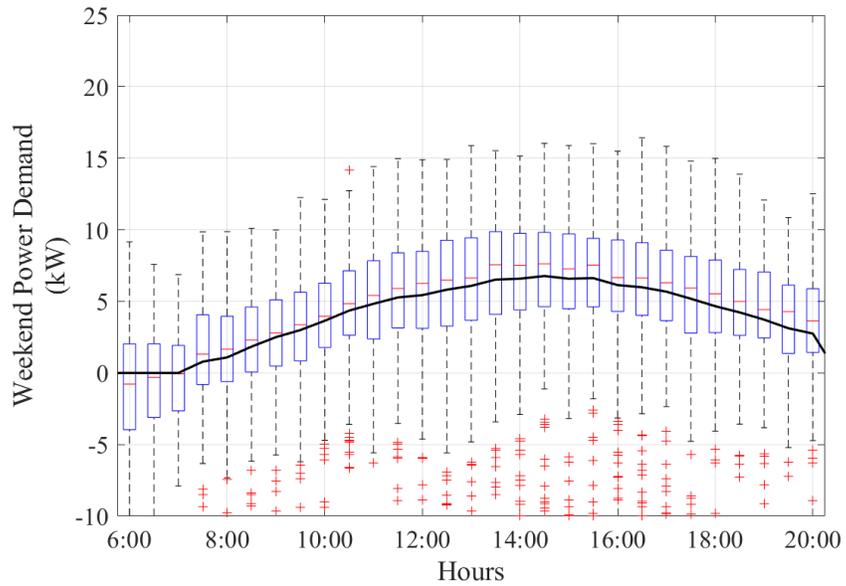
(b)

Figure 5.15: Statistical mean value of working hours power demand caused by occupant activities in working days and weekends/holidays.
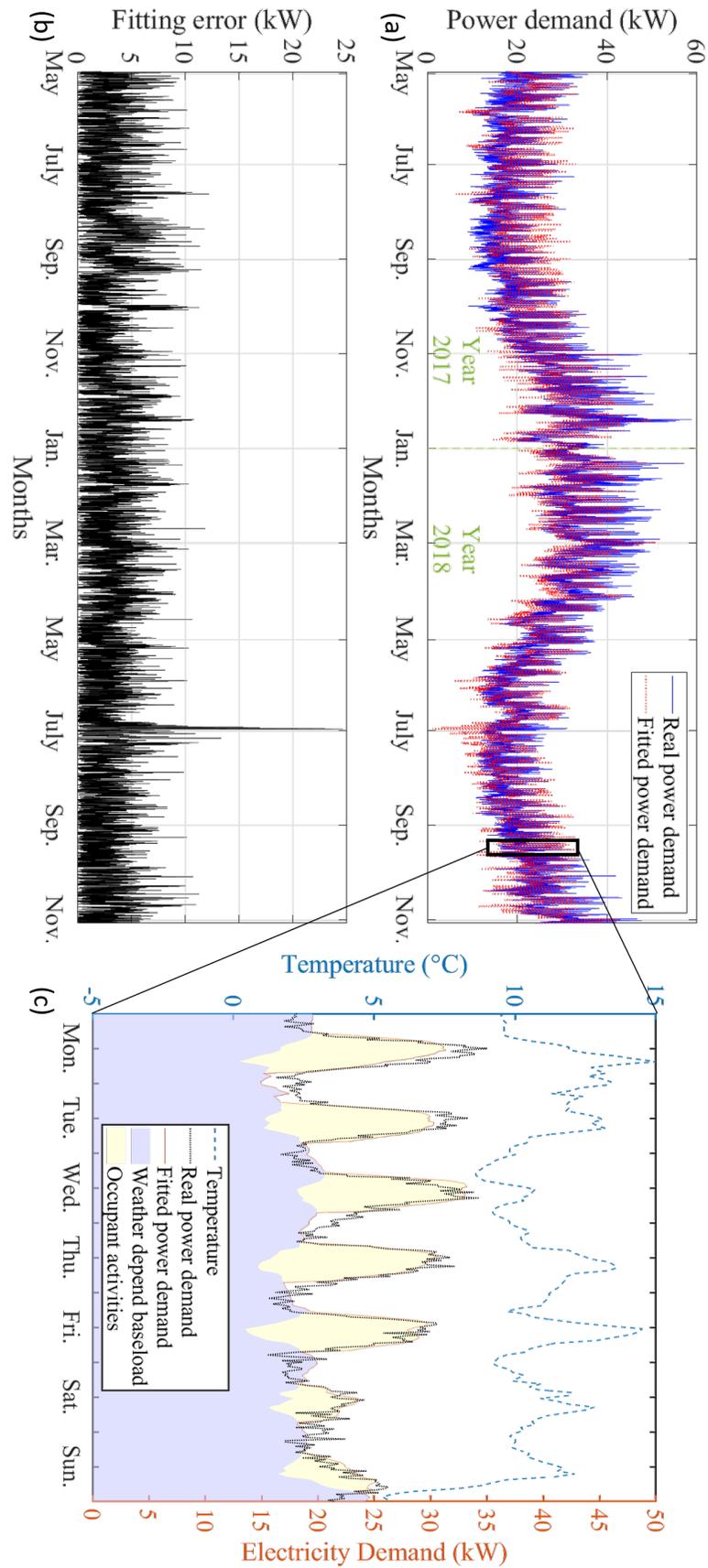
Figure 5.16: Prediction result of fitted power demand comparing with real power demand.
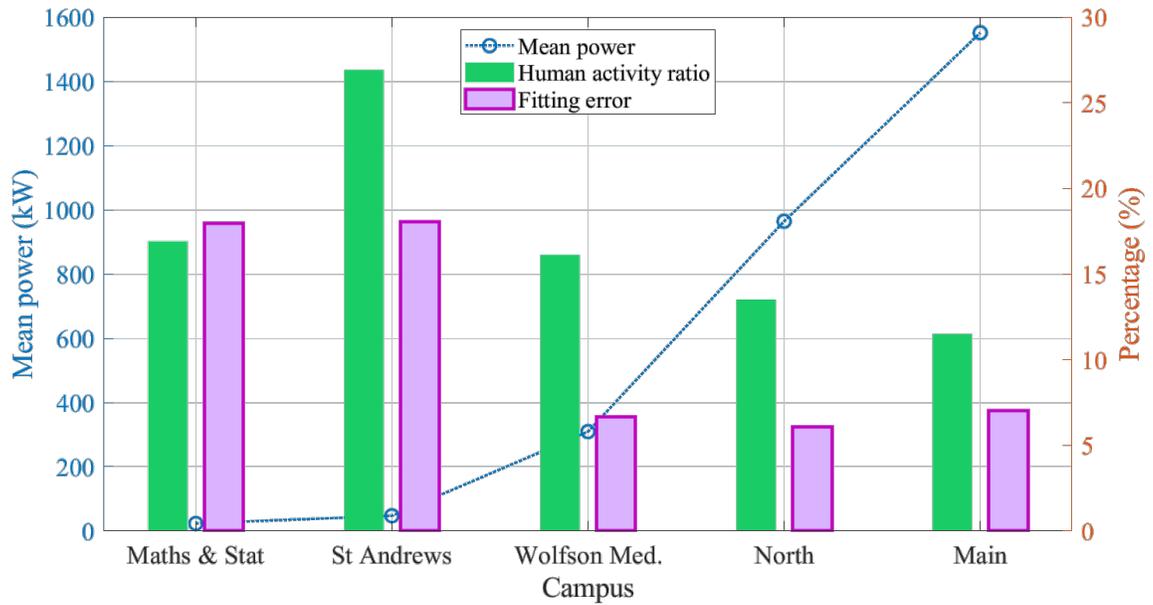
Figure 5.17: Ratio of human activity impact to total power demand and fitting error of each campus.

new design is theoretically one solution to improve the prediction accuracy.

In the University of Glasgow campus, the data of electricity power consumption is recorded with 5 different campuses, including the Maths & Statistics school (1 building), the St. Andrews campus (1 building), the Wolfson Medical School (1 building), the North campus (15 buildings), and the Main campus (13 buildings). Each campus has different capacity of rated powe demand and occupant behaviour.

The campus power data shows different relationship with temperature. In order to compare these information, a bar chart is shown in Figure 5.17 that indicates the ratio of human activity impact to total power demand and fitting error of each campus. The human activity ratio is calculated with the fitted power demand caused by occupants to the total power demand including both the building baseload power and occupants power demand.

The ratio of weather dependent power to the total power consumption in building are different. The blue circle in the bar chart shows their mean power, which indicates the size of the buildings. The mean power demand of each campus is ranking with an increasing order. The green bar shows the ratio of fitted power demand caused by occupants to its total power consumption including both the fitted building baseload power and occupants power of each campus. The purple bar shows the average fitting error. From the result, it can approximately find that, the larger campus normally has less impact of human activity, and the less ratio of human activity in power consumption results in less fitting error occurred.

5.3.2 Fitting result of AI based approach

In this section, the artificial intelligent based fitting approach is tested in fitting the power demand of target building/campus. As in the design, the ANN method is used to find the relationship between weather variables and power demand of no-occupancy and full-occupancy activities. In the previous case, the statistical results of power demand caused by occupancy show that the probability density function of occupancy within a day is approximately a bell-shaped curve. Therefore, in order to minimise the impact of the occupancy rate varying, an interval where the occupancy rate is as stable as possible should be selected near the peak and valley values of the occupancy rate. This will result in a narrower time interval. However, if the time interval is selected as narrow as possible, it will result in less data available for ANN training, which will reduce the prediction accuracy. Therefore, it is necessary to consider the trade-off between narrower time intervals and more data for ANN training in defining the full-occupancy and no-occupancy time period. In the ANN training, the full-occupancy hours are set as from t_2=11:00 to t_3=15:00. The no-occupancy hours are set as from t_4=00:00 to t_1=04:00. The switching between no-occupancy hours and full-occupancy hours is using the predicted occupancy rate from the ANN model.

The fitted membership function of occupancy rate is related to the weather variables and time horizon using the trained ANN. Figure 5.18(a) shows fitted membership function for one week as an example. The result shows that the occupant's membership is the highest at noon and is the lowest at night when it is normally close to zero. The result of predicted power fitted by ANN is shown in Figure 5.18(b). The blue shadow on the bottom shows the fitted no-occupancy power demand, which indicates the building baseload. The green dashed line shows the fitted full-occupancy power demand, which indicates the peak load. The yellow shadow in the middle shows the fitted power demand caused by occupants' activities from the predicted occupancy rate in Figure 5.18(a). Combining the baseload power demand, the occupants' activities related power demand and occupancy membership function fitted from ANN, the final fitted power demand is shown as the solid red line shown in Figure 5.18(b). Comparing it with the real power demand shown as the dotted black line, the fitted power demand tracks the real power demand and is able to predict the future electricity demand.

In order to validate the effectiveness of the proposed approaches, the comparison among the linear regression-based approach (Approach 1), the ANN based approach (Approach 2) and the conventional ANN fitting approach in one week is shown in Figure 5.19. The result of conventional ANN is shown with the dashed green line. Due to the issues mentioned in Section 5.1.2, the conventional ANN cannot fit the data because the long-term and short-term time horizons are negative and positive correlation with weather variables. Thus, the conventional ANN cannot find the best relationship between weather variables and target power demand. With splitting the data by different time periods of no-occupancy hours and full-occupancy hours, the proposed approaches can have better prediction performance and less absolute prediction error.
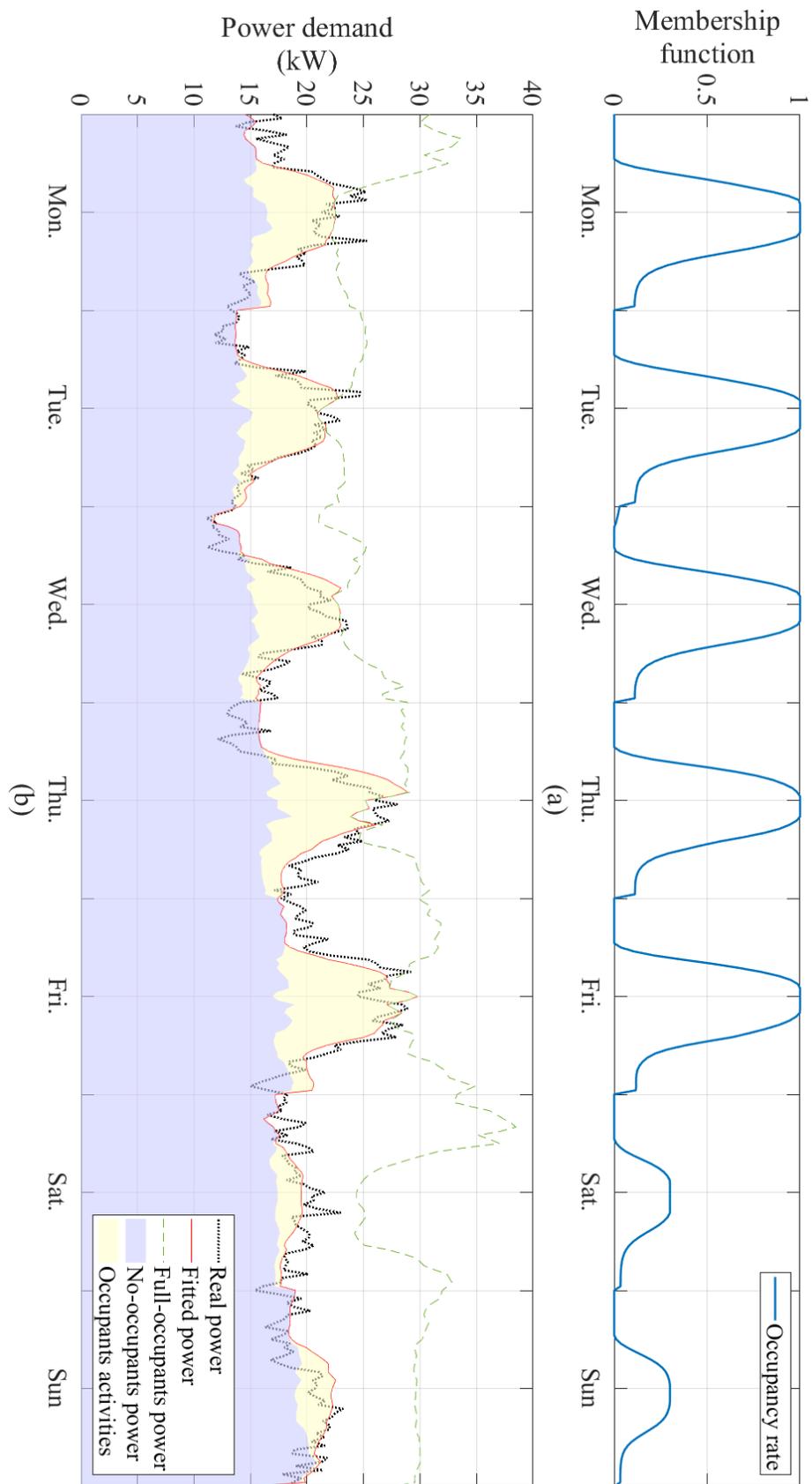
Figure 5.18: Prediction result of proposed ANN based approach (approach 2) in one week as an example. (a) Fitted membership function of occupancy rate, (b) fitted power demand.
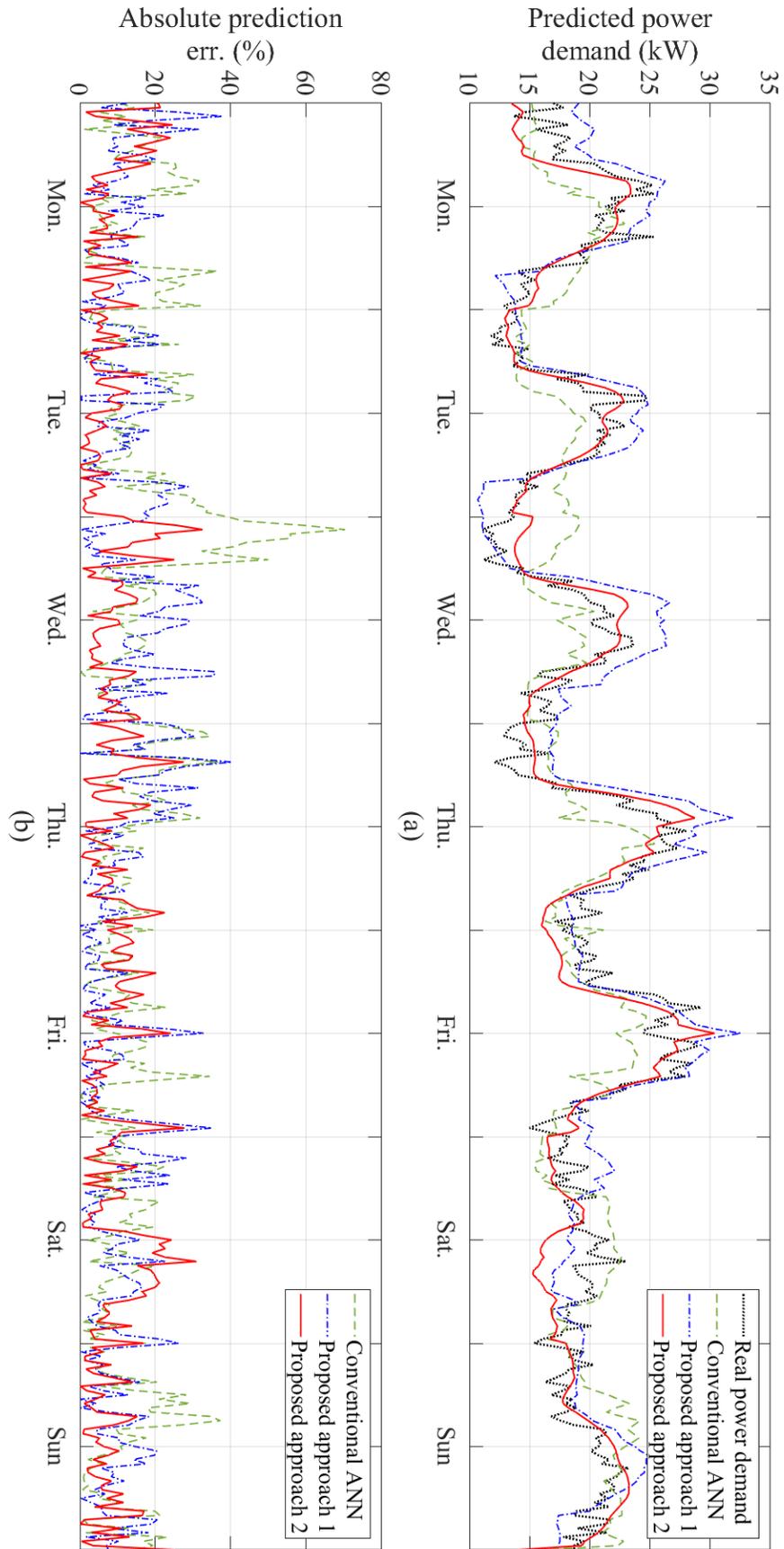
Figure 5.19: Fitting result comparison among the ANN to full power, the baseload & human activity approach, and the ANN and fuzzy based approach.

Using Approach 1, the non-working hour power demand is fitted with the linear regression of air temperature and working hour occupancy power demand is obtained from its average value. This test validates the effectiveness of splitting the data by working and non-working hour time periods. However, the linear regression cannot fully use the information of weather variable. In the ANN with fuzzy hours splitting approach (Approach 2), the no-occupancy hours power demand, the full-occupancy hours power demand and occupancy rate is fitted with weather variables by the ANN approach. The result shows that Approach 2 has less absolute prediction error than the linear regression-based approach and the conventional ANN.

In neural network technology, data is usually classified into training set and test set, usually in the ratio of 70:30, to guarantee the model accuracy. This case study contains the electricity demand data of 18 months in total, from May 2017 to October 2018. Therefore, the data of the first 12 months is used to train the ANN, and the data of the next 6 months is used for testing. In addition, the power demand data from December to February has the highest power demand in a year. As the training data includes this period, the power demand of other months can be guaranteed to be within the boundary of the model.

The power demand of the target building of Maths & Stat School for training and the predicted power demand is shown in Figure 5.20(a). In addition to the Maths & Stat School, other university campuses have different dependency of power demand to weather variables and each campus has different capacity of rated power demand and occupant behaviour.

As described in Section 5.1.2, the Maths & Stat School, St Andrews Campus and Wolfson Medical School have their own individual power meter to record the electricity consumption. The north campus and main campus only recorded the total power consumption of tens of buildings. In addition, the university has its own district heating system for most old buildings. As the Maths & Statistic School is newly built, it is not included into the district heating system. Its space heating is fully supplied by the electrifying heating and, therefore, its electricity demand shows more relating to the weather conditions. Other buildings have different percentage of electrifying heating depending on occupants' behaviour. The weather conditions have less influence on these buildings comparing with the Maths & Statistic School. Therefore, to validate the universality of the ANN based approach and its robustness to different data, the prediction results of other buildings or campuses using the same prediction approach are given in Figure 5.20(b)-(e).

In order to compare their prediction performance of all methods numerically, the performance index is choosing the root-mean-square error (RMSE) between predicted power demand and real power demand of the last 6 months. The RMSE of all five campuses predicted by the conventional ANN approach and two proposed approaches are compared in a bar chart as shown in Figure 5.21.
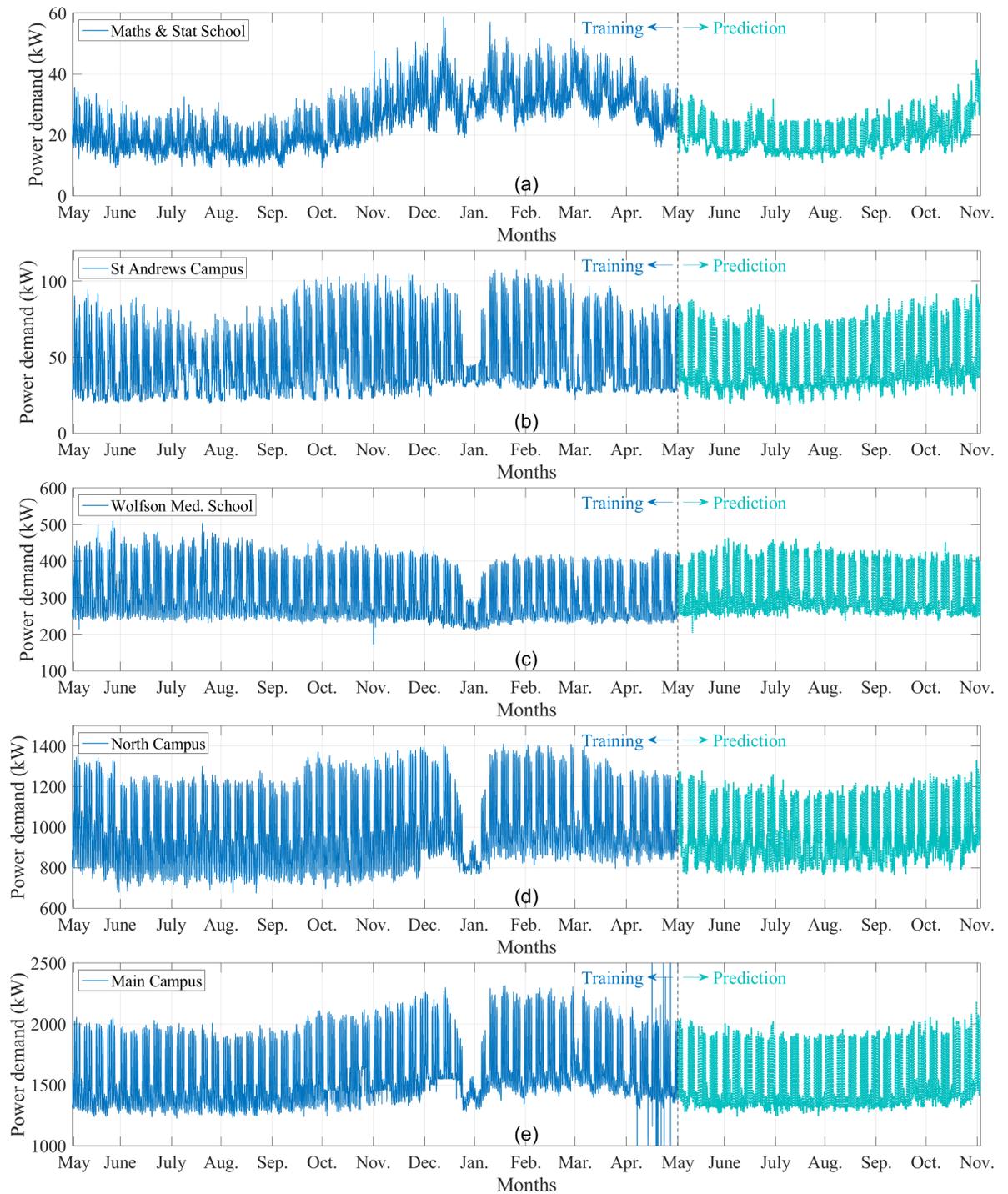
Figure 5.20: Prediction of power demand of different campuses using the proposed approach.
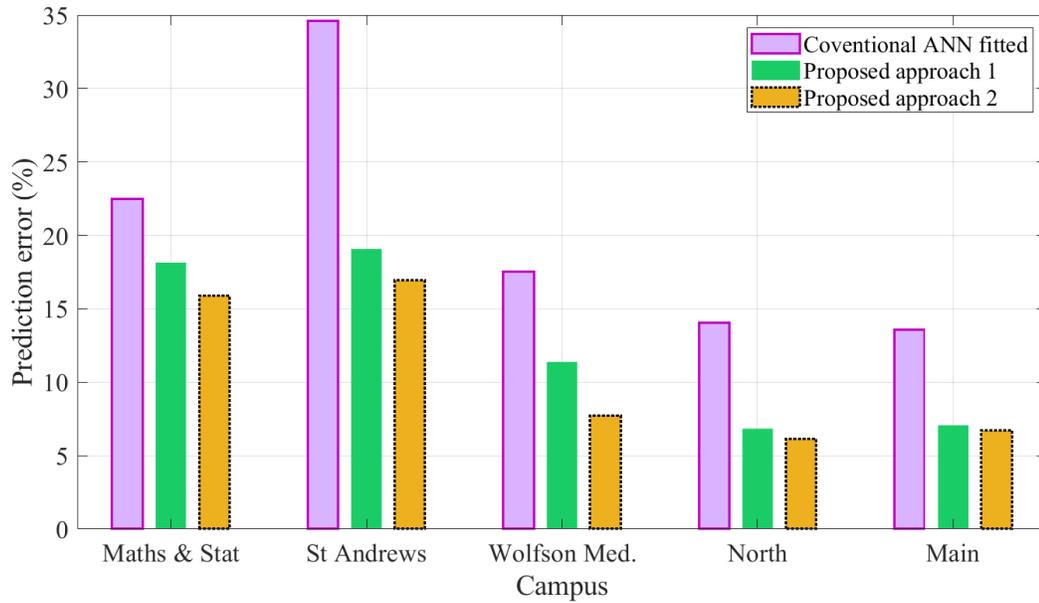
Figure 5.21: Prediction error of different fitting approaches to each campus.

From the result, the proposed approach that uses the ANN to fit the power demand data of working and non-working hours separately can get the best prediction performance with less RMSE. The proposed Approach 2 can reduce RMSE by 5% to 30% compared with the proposed Approach 1, and it reduces RMSE by 30% to 55% compared with the conventional ANN. In average, the proposed working hour splitting based regression approach and ANN with fuzzy hours splitting approach can reduce the RMSE prediction error by 35% and 42%, respectively.

The author reviewed publications on the same research topic in recent years, and the average prediction error of office building energy consumption prediction was between 5% and 15% [67, 108, 117, 206, 207]. For a more accurate prediction model, the model can be improved by installing a large number of meters and sensors to obtain more measurement data, and the prediction error can be less than 5% [103]. In our work, due to the lack of detailed data and uncertain occupancy rate in the model, the estimation results are reasonable between 5% and 20%, and the accuracy is acceptable.

Thus, the proposed approaches can be used to predict the power demand, whose long-term and short-term data are in negative and positive correlation, respectively. In addition, the proposed approach with ANN and fuzzy technologies can be used as a 'grey-box' to include the knowledge of physical process to explain the effect of occupancy activities.

5.4   Conclusion

The electricity demand of office buildings seems to be in negative and positive correlation with weather variables in long-term and short-term time horizon, respectively, as a result the conventional ANN approach cannot accurately capture the relationships between them. In this chapter,

two electricity demand prediction approaches have been proposed to solve this issue. The initial proposed approach splits the power demand data by working hours and non-working hours to avoid the impact of occupants' activities to building power demand. Using this method, the linear regression approach is used to fit the building baseload power to a weather variable using the non-working hours data and find the average occupants power demand using the data of working hours. To fit the power demand with more weather variables, the proposed approach is further developed to use ANN to fit the non-working hours data and working hours data and the membership function of fuzzy hours between them. With the second proposed approach, more weather variables can be considered in the model to predict the power demand more accurately. In the simulation results, both approaches have been validated to show less RMSE value than the conventional ANN approach in predicting the power demand. In addition, the ANN with fuzzy hours splitting approach has the best performance among the three approaches and reduces RMSE by 5% to 30% compared with the working hour splitting based regression approach and reduces RMSE by 30% to 55% compared with the conventional ANN. Therefore, both approaches are able to solve the issue that the input and output fitting data are in negative and positive correlation in long-term and short-term time horizon, respectively. The proposed approaches can achieve good performance with RMS prediction error as low as 6% in building power demand prediction. In future works, the working hour splitting approach and fuzzy hour approach will be applied to other models, such as deep learning and stochastic models, to validate the result.

Chapter 6

Integrated Heat and Power Network

6.1   Integrated Energy Network in University Campus

In the previous chapters, the heat demand and the electricity demand are analysed independently. But in practice, heat demand and electricity demand can potentially be combined and analysed at the same time to achieve better performance in improving energy efficiency and decarbonisation. In recent years, the integrated energy network (IEN) has developed rapidly in terms of energy efficiency improvement, carbon dioxide emission reduction, and renewable energy integration.

In the University of Glasgow, the electricity and heat is supplied by the energy center, which includes a 3.4MW combined heat and power (CHP) gas engine, whose model is GE Jenbacher 620, and five gas boilers with a total heat production capacity of 42.4MW. Both the CHP and boiler consume gas from the gas network, as shown in Figure 6.1.
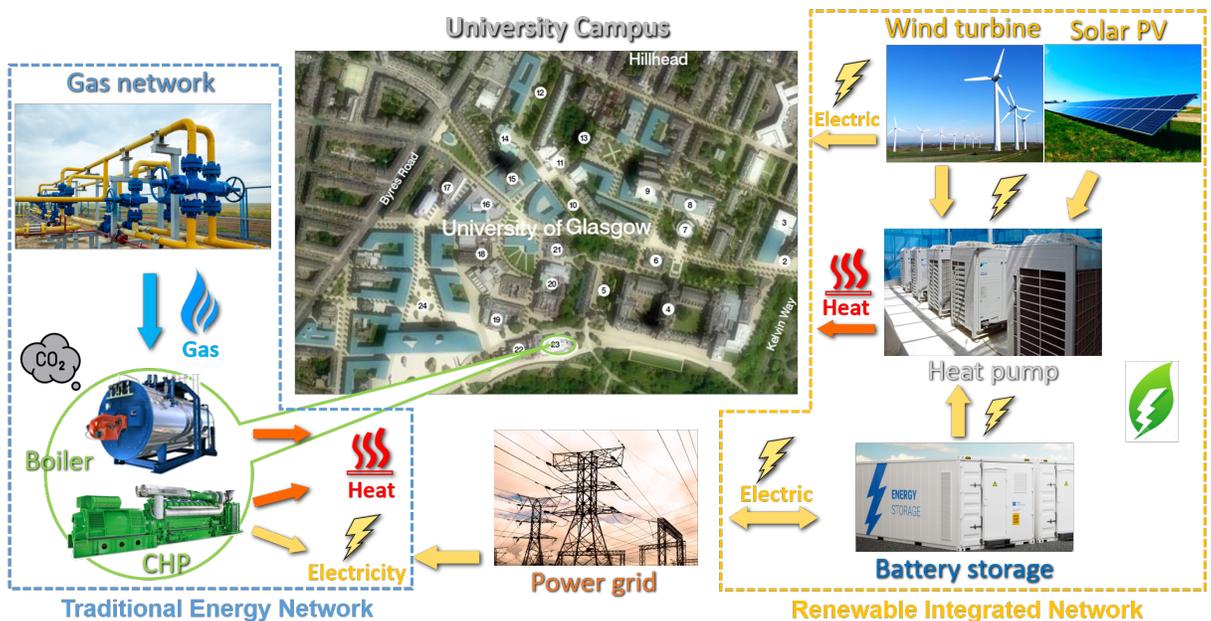
Figure 6.1: Integrated heat and power network of the University of Glasgow including renewables power generation.

The heat generated from both boiler and CHP is supplied to buildings across the whole campus while the electricity generated from the CHP is only supplied to buildings in main campus, as defined in Figure 5.12. The CHP generates both heat and electricity, but its operational control method is electricically lead rather than heat lead. This means that CHP gives priority to electricity demand. When the electricity generated by CHP is lower than the electricity demand of the campus, the remained electricity demand is supplied by the grid. When the heat generated by CHP is lower than the heat demand of the campus, the remained heat demand is supplied by the gas boiler. In summer, due to the high temperature, the campus does not require a large amount of heat supply, and the CHP will be shut down for couple of months to save energy and carry out annual maintenance. And the electricity is mainly supplied by power grid.

The demand of heat and electricity for university campus has been modelled and predicted in Chapter 3, 4 and 5. The result of hourly heat and electricity demand of University of Glasgow in a whole year of 2018 comparing with local tempertaure is shown in Figure 6.2. It can be seen from the results that the power demand has a small correlation with temperature [208], while the heat demand has an obvious correlation with temperature changes. In summer, heat demand is less than half of electricity demand, while in winter, heat demand can reach more than twice the electricity demand.
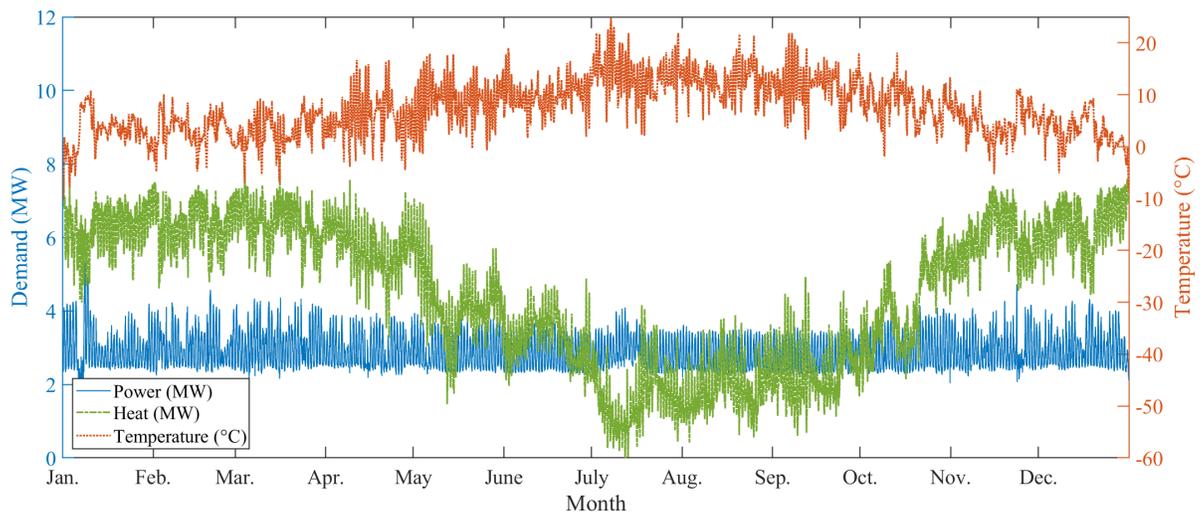
Figure 6.2: Demand of heat and electricity power of University of Glasgow in a whole year of 2018 in comparison with local temperature.

The operation strategy of the traditional gas boiler/CHP based IEN of the University of Glasgow Energy Center will effectively control the monthly cost. The monthly gas and electricity costs are shown in Figure 6.3. From the result, it can be found that the cost for electricity is mainly in summer, which takes more than half of the monthly cost. In other seasons, the main cost is from the gas consumption. The approximate amount of CO_2 emissions can also be estimated from the amount of gas consumption from the result. However, if the current gas boiler/cogeneration heating system is replaced with an electrified heating system, it will signifi-
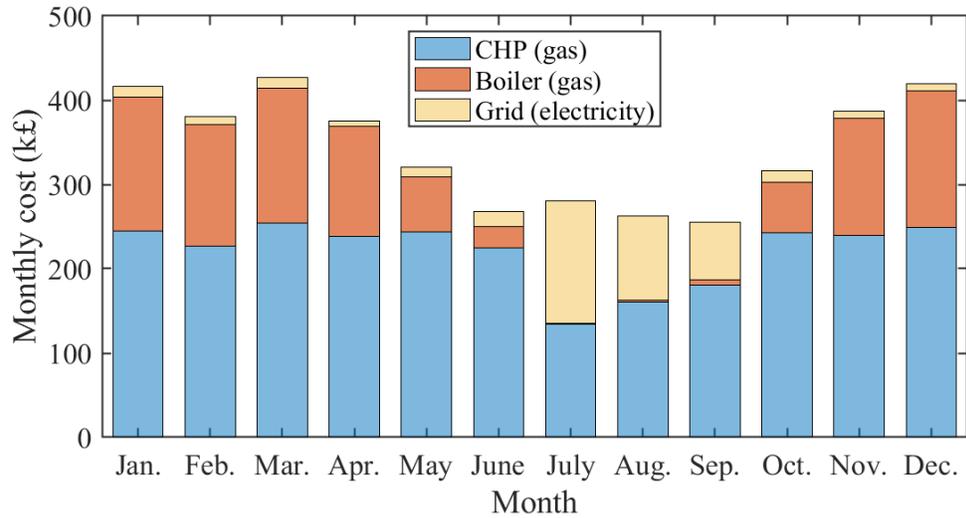
Figure 6.3: Monthly cost of energy demand with both heat and power in traditional energy network.

cantly increase the monthly electricity demand and cost.

Since grid-powered electrified heating system may significantly increase the power demand of the grid, an alternative way is using the local renewable resources to generate power to reduce the load of power grid, as the renewable integrated network shown in Figure 6.1. In addition, the battery storage can be used to store power when the power demand is low, and discharge when the power demand is high in order to reduce the peak load to the grid. The renewable power generation is based on both the installation capacity and local renewable energy resources. Taking the location of the University of Glasgow as an example, the monthly power generation of wind turbines and solar PV installed capacity of 1MW is shown in the Figure 6.4. The result shows that the local wind resources (blue bars) are better in winter and local solar resources (red bars) are better in summer.
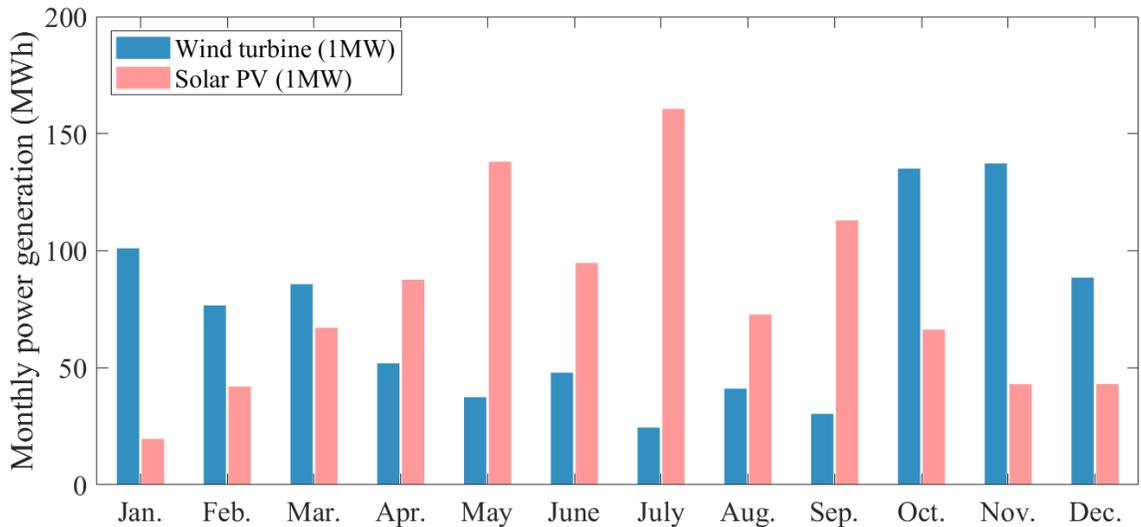
Figure 6.4: Monthly power generation of 1MW wind turbine and solar PV.

6.2 Renewable Heating System Modelling

Traditional heating systems use combined heat and power (CHP) and gas boilers to consume gas to generate heat for space heating of buildings. The use of gas can cause a large amount of CO_2 emissions. In order to reduce greenhouse gas emissions, one solution is to use electrically driven heat pumps to heat buildings. However, this will increase the cost in electricity and have an impact on the electricity grid, especially in winter when the heat demand is relatively high. This can lead to extra load to the power grid and affect its stability. To solve this problem, the usage of local renewable resources is considered as a substitute. Renewable resources mainly include wind and solar energy that both depend on weather conditions. Thus, the heating system requires battery storage systems and thermal energy storage to compensate the imbalance between generation and demand for electrical energy and thermal energy, respectively.

6.2.1 Wind turbine

The role of wind turbines (WT) is to convert wind energy, an environmentally friendly energy source, into a form of electrical energy. It is one of the most promising renewable energy sources and its global installed capacity is increasing year by year [209]. According to the direction of their rotation axis, WTs can be classified into horizontal axis wind turbines (HAWT) and vertical axis wind turbines (VAWT). HAWT extracts wind energy on the horizontal axis and the blades rotate due to the lift provided by the aerodynamic force from wind. Because of it has higher efficiency than VAWT, the HAWT is more popular in research. However, HAWT needs to always point to the wind direction to work effectively. For unpredictable wind directions, the HAWT needs additional mechanisms to ensure that the blades are always facing the wind direction. The VAWT has its blades perpendicular to the ground and rotating around the vertical axis. VAWT can receive wind power from any direction. Compared with HAWT, VAWT has several advantages, including no need to face to the wind flow, the ability of generating electricity at low wind speed, no need for additional control on yaw and pitch, less maintenance costs, and less noise [210]. However, VAWT has much lower efficiency at high wind speeds than HAWT, its dynamic stability is relatively poor due to its light weight, and it is vulnerable to backtracking wind. For the above reasons, VAWT is often installed in slow and turbulent wind environments, such as roof top, for low-power generation [211], while HAWT is often used for high-power generation in isolated and remote areas, such as on-shore or off-shore wind farms [212]. In recent years, some researchers have combined the two types of WT together to improve its operational capabilities [213].

Although HAWT and VAWT have lots of difference in structure, control strategy and power efficiency, the methods they use to capture power from wind energy are the same, and thus can be expressed by the same equation. The available power output of a WT is proportional to the cube of the wind speed. If the wind speed is too low, the output power of the WT can be ignored.

When the wind speed increases, the output power of the WT increases rapidly. Therefore, in order to ensure the normal operation of a WT, three threshold wind speeds usually need to be considered, including the cut-in speed, the rated speed and the cut-out speed [214]. When the wind speed is lower than the cut-in speed, the WT stops to avoid unnecessary mechanical wear and energy consumption; when the wind speed is higher than the cut-in speed and lower than the rated speed, the WT controls its rotating speed to track the maximum power point referring to the current wind speed; when the wind speed is higher than the rated speed and lower than the cut-out speed, the WT adjusts its pitch angle to limit the captured wind power not to exceed the limit of the generator; when the wind speed exceeds the cut-out speed, the WT shuts down to protect the rotor and the generator from the damage of strong wind. The example of local wind speed in a year is shown in Figure 6.5 (a).
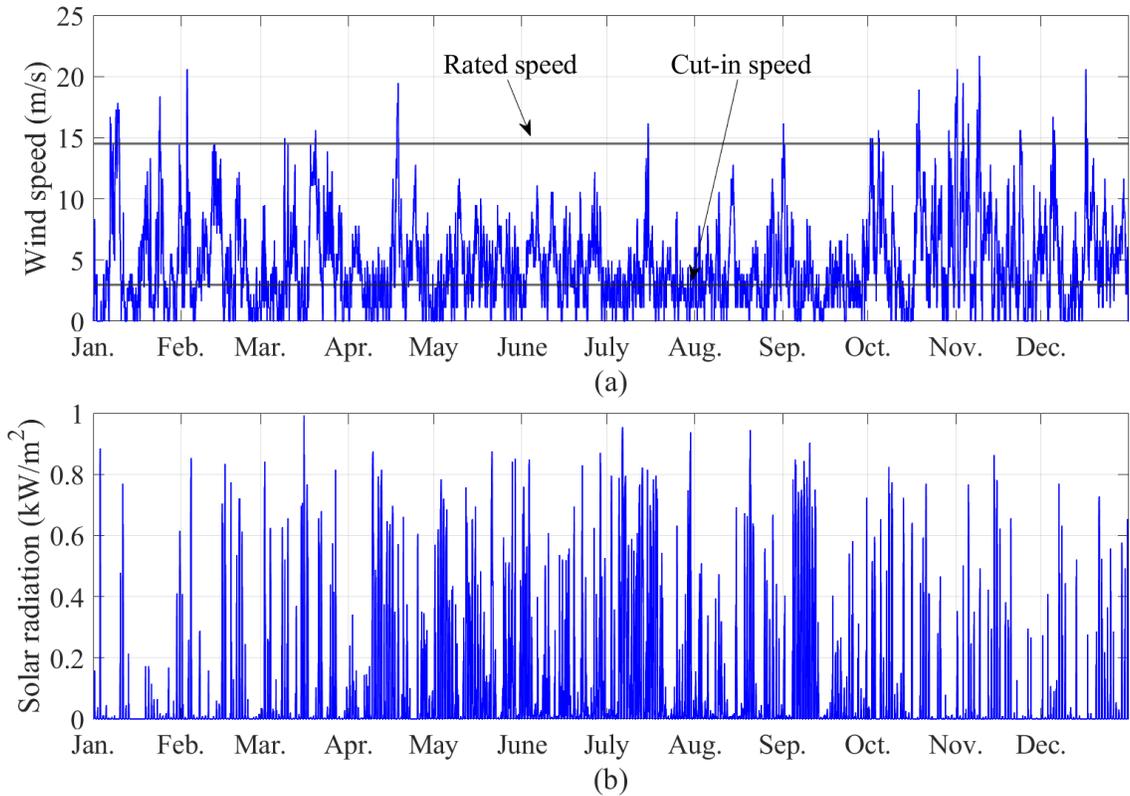
Figure 6.5: Power generation from wind turbine depending on wind speed.

The power output of WT captured from wind is presented as

$$P_{\mathrm{WT}}(h) = \begin{cases} 0 & V_{\mathrm{ws}}(h) < V_{\mathrm{cut-in}} \text{ or } V_{\mathrm{ws}}(h) \geq V_{\mathrm{cut-out}} \\ \frac{1}{2}\rho A_{\mathrm{r}} C_{\mathrm{p}} V_{\mathrm{ws}}^3(h) & V_{\mathrm{cut-in}} \leq V_{\mathrm{ws}}(h) < V_{\mathrm{rate}} \\ P_{\mathrm{WT,rate}} & V_{\mathrm{rate}} \leq V_{\mathrm{ws}}(h) < V_{\mathrm{cut-out}} \end{cases} \qquad (6.1)$$

where ρ is the air density, A_{r} is the swept area of WT that depends on the turbine radius or length of blades, $V_{\mathrm{ws}}(h)$ is the hourly mean wind speed, and C_{p} is the power conversion coefficient that is determined by the aerodynamic structure of the WT. $V_{\mathrm{cut-in}}$, V_{rate} and $V_{\mathrm{cut-out}}$ denote the cut-in

wind speed, rated wind speed and cut-out wind speed.

When designing a renewable energy system, it is necessary to determine the suitable capacity of the WT according to the rated power of the WT and the local wind resources. During the operation, it is necessary to consider whether the current wind speed has reached the cut-in speed, rated speed and cut-out speed, and calculate the power that the WT can generate according to the wind speed.

6.2.2 Solar photovoltaic

Solar energy is another well-known clean energy. Since most parts of the world have sufficient solar irradiance, solar power generation technology has attracted a large number of researchers. Photovoltaic (PV) is a device that uses the photoelectric effect to directly and conveniently convert solar energy into electricity [156]. In recent years, with the development of PV technology and the rapid decline of PV costs, large-scale grid-connected solar PV power stations have been built all over the world [215].

However, PV systems are highly dependent on solar radiation and weather characteristics. The example of local solar radiation in a year is shown in Figure 6.5 (b). Due to the variability of solar radiation and ambient temperature, the power output of the PV system is uncertain and random [215]. The uncertainty is from the alternation of day and night, effect of seasonal changes, and random cloud movement [216]. Compared with the distributed PV system, the power output of the relatively concentrated and single format PV power generation system in a small area can change very quickly. An isolated community can be completely or partially obscured by fast-moving clouds in the span of a few seconds to a few minutes [217]. The movement of the cloud will affect the performance of the PV system, so prediction must be made to avoid undesirable technical problems and cost losses [159]. Predicting the power output of a PV system is a challenging task.

For a photovoltaic array composed of PV panels, the maximum power output can be expressed as [218]

$$P_{\text{PV}}(h) = \eta_{\text{pv}} A_{\text{pv}} S_{\text{T}}(h) \cdot [1 - 0.005 (T_{\text{air}}(h) - 25)] \tag{6.2}$$

where η_{pv} is the photoelectric conversion efficiency of the PV array, A_{pv} is the total area of the PV array, S_{T} is the solar radiation incident on the panel, and T_{air} is the air temperature of the ambient in Celsius degree.

The power output of PV system mainly depends on the amount of solar radiation and the operating temperature of the installation location. In addition, as the equipment ages, the photoelectric conversion efficiency and power conversion efficiency will gradually decrease. But their changes in the system life cycle are very small, so they can be ignored. Therefore, as long as the solar radiation and panel temperature are known, the power output of the PV system can be predicted.

6.2.3 Battery storage system

As the main renewable energy sources (RESs), solar and wind energy are intermittent and fluctuating weather resources. One of the possible solutions is to use a hybrid renewable energy system to integrate various renewable energy sources in the best combination [160]. A good example is the complementarity of solar and wind energy [219]. However, this approach still has large uncertainties. If the distributed resources cannot provide sufficient flexibility in power generation, network operators will have to limit the penetration rate of renewable energy to a certain level, which will adversely affect the use of RESs [159]. For example, the power generation of the PV system usually exceeds the residential electricity load during the day but cannot meet the demand in the late afternoon and evening when the residential load tends to increase [155]. Therefore, in order to highly integrate variable wind and solar energy and reduce the impact of weather changes on renewable power generation, energy storage systems have been widely accepted as one of the potential solutions [220–222]. The main function of energy storage system (ESS) is to compensate the imbalance between power generation and demand, transfer energy from periods of high power generation to periods of low power generation, so that the system can maintain full functionality under various operating conditions [157].

ESS includes a variety of technologies, such as pumped water storage (PWS), compressed air energy storage (CAES), hydrogen storage, flywheels, supercapacitors, and batteries [124]. The batteries have higher energy density than supercapacitors, higher charging rate than pumped water storage and less power loss than flywheels. Therefore, the battery storage system (BSS) is the most suitable storage system to store the electricity energy for such localised small-scale applications due to its advantage of higher charging rate, higher energy density and shorter response time. The response time of BSS is between milliseconds and seconds, while that of PWS and CAES are from a few seconds to minutes [223]. Due to their durability, low maintenance and low social environmental impact, it is expected that the development and use of BSS will increase significantly in the next decades [124]. Among batteries, the most used are lithium-ion batteries (LIB) and vanadium redox flow batteries (VRB). The advantages of LIB are its high energy density, high efficiency, long lifecycle, and more environment friendly characteristics. However, due to these advantages of LIB, they are also widely used in a variety of other applications, including small electronic products and electric vehicles. VRB is a good substitute for LIB because of its safety, long service life, better scalability and high recyclability [124]. However, the disadvantage of VRB is its low energy density [223]. Therefore, a VRB requires more floor space than a LIB at the same capacity.

In the charging and discharging behaviour of a battery, the state of charge (SOC) is often used to reflect the ratio of the battery's remaining capacity to its rated capacity [224]. To protect battery lifecycle, it is necessary to limit the SOC within a certain range. During battery charging,

the SOC can be given as [225]

$$\text{SOC}_{\text{BSS}}(h) = \text{SOC}_{\text{BSS}}(h-1) \cdot (1-\sigma) + \frac{P_{\text{BSS}}(h) \cdot \Delta h}{E_{\text{BSS,cap}}} \cdot \eta_{\text{ch}} \tag{6.3}$$

and during battery discharging, the SOC can be given as

$$\text{SOC}_{\text{BSS}}(h) = \text{SOC}_{\text{BSS}}(h-1) \cdot (1-\sigma) + \frac{P_{\text{BSS}}(h) \cdot \Delta h}{E_{\text{BSS,cap}}} / \eta_{\text{dch}} \tag{6.4}$$

where σ denotes the self-discharge rate on an hourly basis, P_{BSS} denotes the power of BSS generated or absorbed, $E_{\text{BSS,cap}}$ denotes the energy capacity of BSS, η_{ch} and η_{dch} denote the battery charge and discharge efficiency, respectively. Δh represents the time step, in this article, the time step is calculated in hours.

6.2.4 Heat pump

Heat pumps are a group of energy systems that can extract heat energy from lower temperature sources such as ambient air, soil, water in lakes and rivers, and then upgrade and deliver it at a higher temperature for heating applications [226]. After the liquid absorbs heat and converts it into vapor, input external energy to compress it and convert it into hot vapor. The hot vapor is sent to another heat exchanger, the condenser. Here, the hot steam releases the heat obtained from the heat source in the evaporator and is condensed into hot liquid in the process [226]. The heat released by the hot steam is used to heat the indoor space. If the heat source is outside air, the system is called an air source heat pump (ASHP). If the heat source is the ground, soil or groundwater, it is called a geothermal heat pump (GHP). GHP is one of the fastest growing renewable energy sources in the world [153]. Compared with traditional heating fuels, such as natural gas, heating oil and propane, the effectiveness of operational cost of GHPs is directly related to the cost of electricity that drives the HP [150]. But the installation costs of GHPs can be an order of magnitude higher compared to a gas boiler.

Heat energy delivered by the HP normally exceeds its electricity consumption. Coefficient of performance (COP) defined as the ratio of heat output to input power is used to describe the energy performance of heat pumps [151]. The COP of heat pumps can reach 3-5 or more depending on temperature conditions, which means the heat pump can provide 3-5kW of heat by consuming 1kW of electricity. Furthermore, the smaller temperature difference between the heat source and the radiator, the greater the COP of the HP system. In cold weather, the COP of an ASHP is usually about 2 or lower. Therefore, GHP have higher efficiency than ASHP because the underground temperature is closer to the required indoor conditions than the outdoor air temperature in winter and the GHP usually use water which reduces the size of the evaporator heat exchanger due to the higher energy density compared to air. However, as GHP requires higher installation costs and larger spaces, ASHP is more suitable for heating smaller spaces.

In HP applications, its COP can be affected by different variables, such as external temperature, water supply temperature, inlet water temperature and load factor. Simplifications can be made to reduce this complexity. The COP can be calculated as [149]:

$$\text{COP}_{\text{HP}} = \frac{\dot{Q}_{\text{HP}}}{P_{\text{HP,comp}}} \tag{6.5}$$

where $\dot{Q}_{\text{HP}}$ is the rate of heat transfer in condenser of HP, $P_{\text{HP,comp}}$ is the input power to the compressor.

The rate of heat transfer from the condenser is obtained as

$$\dot{Q}_{\text{HP}} = \dot{m}_{\text{ref}} \cdot (T_{\text{HP}} - T_{\text{return}}) \cdot C \tag{6.6}$$

where $\dot{m}_{\text{ref}}$ is the mass flow rate and C is the specific heat constant in the unit of $kJ/kg \cdot K$.

In this study, the air source heat pump is considered. The COP of the HP is modelled as a function of temperature lift, or the temperature lift is known as the difference between the supply water temperature T_{wat} and the external air temperature T_{temp} [141].

$$\text{COP}_{\text{HP}} = a \cdot (T_{\text{wat}} - T_{\text{temp}}) + b \tag{6.7}$$

where a and b are the parameters from the linear regression fits of manufacturer's data [141]. However, the linear regression is just an initial fitting for the easy implementation of modelling. A nonlinear model can be more accurate. Since this chapter aims to provide a feasible solution for investment of heat pump rather than operational control, the linear regression model is sufficient for estimating the COP of HP.

As a case study, Figure 6.6 shows the monthly power demand of electricity consumption and electric-driven heat pump versus ambient temperature. It can be seen from the figure that the fluctuation of electricity consumption has less dependency with temperature, while the power required by electric-driven heat pumps is highly dependent on temperature, that is, higher in winter and lower in summer. In winter, the power demand of the electric-driven heat pump is equivalent to the total electricity consumption, which means the power demand from grid will be doubled due to the use of heat pump. This verifies that the usage of electric-driven heat pump will significantly increase the peak power demand in winter and will bring further challenges to the national grid in terms of creating additional power generation capacity and balancing power generation and demand.

6.2.5 Thermal energy storage

The outdoor temperature changes between summer and winter during the year can produce huge changes of heat demand [227]. In this case, a lot of heat will be wasted in the summer when
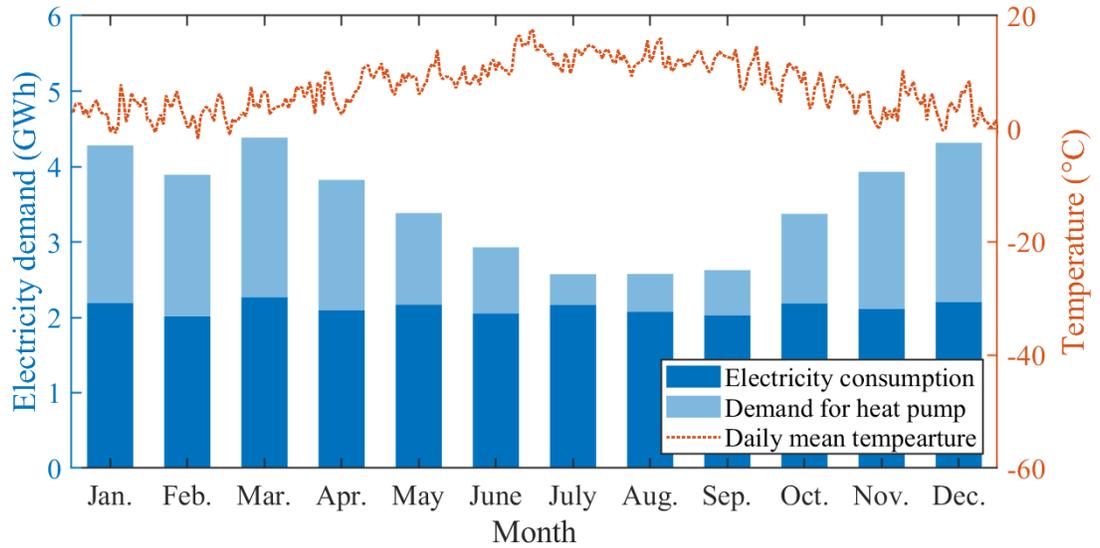
Figure 6.6: Monthly electricity consumption and electricity demand for heat pump.

the heating demand is minimal, and there will be insufficient heat in the winter. In addition, the mismatch between energy demand and supply is often observed. Among several storage technologies, thermal energy storage (TES) seems to be one of the most promising technologies that can compensate for the intermittent heat generation of new energy sources [140]. The advantages of using TES in energy systems include improved overall efficiency, better reliability, bring better economy, reduced investment and operating costs, and less pollution and CO_2 emission to the environment [228].

In recent decades, the TES technology has been widely studied. Considering the required storage time, the TES system can be divided into short-term storage and long-term storage. The short-term storage TES has a charging and discharging cycle of hours to several days, which is called diurnal thermal energy storage. The long-term storage TES can store heat for several months and is called seasonal TES [140, 229, 230]. The principle of seasonal TES is to store heat energy through heating equipment in summer, and then discharge it for space heating in winter [231]. The use of TES can significantly increase the flexibility and self-consumption of renewable energy for end users [232]. Therefore, this concept makes a significant contribution to the efficient use of renewable energy in district heating systems and the decarbonisation of the building sector.

The most traditional TES design is provided in the form of a water tank. When there is excess thermal energy on the supply side, these tanks are charged (heated up), and when the demand for thermal energy exceeds the supply, these tanks are discharged [233]. In the thermal system, the charging and discharging process of the storage tank can be carried out by direct heat exchange. The energy stored in the TES at time h is expressed as [141]

$$Q_{\text{TES,sto}}(h) = Q_{\text{TES,sto}}(h-1) + \left[\dot{Q}_{\text{TES,ch}}(h) - \dot{Q}_{\text{TES,dch}}(h) - \dot{Q}_{\text{loss}} \right] \cdot \Delta h \qquad (6.8)$$

where $Q_{TES,sto}$ is the energy stored in TES, $\dot{Q}_{TES,ch}$ and $\dot{Q}_{TES,dch}$ denote the charge the discharge power, $\dot{Q}_{loss}$ denotes the standing losses, which normally can be found from manufacturer's datasheet.

Among the performance indicators used to evaluate TES, the most common is TES efficiency, which is the ratio of the total heat energy recovered from the heat accumulator at the discharge temperature to the total heat input at the charging temperature as [140]

$$\eta_{TES} = \frac{Q_{TES,dch}}{Q_{TES,ch}} \tag{6.9}$$

6.3   Optimal Sizing of Renewable Heating System

6.3.1   Optimisation problem

The purpose of this study is to find the most effective investment plan to reduce CO_2 emissions in building heating under a given budget. Therefore, the optimisation problem is transferred to minimising the total CO_2 emission in the period between 2020 and 2050. The objective function is expressed as

$$\min \sum_{y_i=2020}^{2050} S_{CO2}(y_i) \tag{6.10}$$

where $S_{CO2}(y_i)$ denotes the annual CO_2 emission in year y_i, which is expressed as the product of specific CO_2 emission factor for nature gas, g_{gas}, in the unit of g/kWh, and heat energy produced by gas boiler, Q_{gas}, in the unit of kWh. The emission factor for nature gas indicates the average CO_2 emission when using gas boiler to consum gas and generate heat. Therefore, it already considered the efficiency of the conversion technology.

In this work, the RHS uses the electricity from local renewable energy. Thus, it is assumed that electricity power generation has no CO_2 emissions in the RHS, and the CO_2 emissions are entirely generated by using the natural gas in the traditional heating system. This assumption can intuitively indicate the reduction in the use of traditional heating system and nature gas as well as the total CO_2 emission during operation. The factor of $S_{CO2}(y_i)$ can be used as an indicator to quantify the decarbonisation performance, which is defined as follows

$$S_{CO2}(y_i) = g_{gas} \cdot Q_{gas}(y_i) \tag{6.11}$$

This equation only considers the CO_2 generated from gas and assumed that the electricity generation has no CO_2 emission in the RHS. This assumption is simple and directly react the CO_2 in operation. And use this factor as an index to justify the performance of decarbonisation.

And Q_{gas} can be presented as

$$Q_{\text{gas}}(y_i) = \sum_{h=1}^{8760} \left[\dot{Q}_{\text{heat}}(h) - \dot{Q}_{\text{HP}}(h) + \dot{Q}_{\text{TES,ch}}(h) - \dot{Q}_{\text{TES,dch}}(h) \right] \qquad (6.12)$$

where $\dot{Q}_{\text{heat}}(h)$ is the hourly heat demand of buildings and $\dot{Q}_{\text{HP}}(h)$ is the hourly heat produced by heat pump, h is the time factor of hours in a whole year, which in total 365 days times 24 hours. The $\dot{Q}_{\text{HP}}(h)$ is determined by both the electricity generated from PV, WT and BSS and its COP as

$$\dot{Q}_{\text{HP}}(h) = \text{COP}_{\text{HP}}(T_{\text{air}}(h)) \cdot (P_{\text{PV}}(h) + P_{\text{WT}}(h) - P_{\text{BSS}}(h)) \qquad (6.13)$$

where P_{BSS} indicates the power storing into the BSS, thus the positive value shows the charging process whie negative value shows the discharging process.

In practice, the output of power generator and heating equipment cannot exceed its rated value, and the energy stored by storage equipment cannot exceed its specified capacity. The constraints on operations of all types of components are given as

$$0 \leq P_{\text{PV}}(h) \leq P_{\text{PV,rate}}(y_i) \qquad (6.14)$$

$$0 \leq P_{\text{WT}}(h) \leq P_{\text{WT,rate}}(y_i) \qquad (6.15)$$

$$0 \leq \dot{Q}_{\text{HP}}(h) \leq \dot{Q}_{\text{HP,rate}}(y_i) \qquad (6.16)$$

$$\text{SOC}_{\text{min}} < \text{SOC}_{\text{BSS}}(h) < \text{SOC}_{\text{max}} \qquad (6.17)$$

$$0 \leq Q_{\text{TES,sto}}(h) \leq Q_{\text{TES,cap}}(y_i) \qquad (6.18)$$

where h indicates the operational time step in hour, and y_i indicates the ith year of investment.

In addition, the service lifetime of devices is not considered in this planning scheme. Thus, the annual installed capacity only uses positive values, which means that the total capacity of each component cannot be lower than the previous year. This assumption is not accurate, but will simplify the optimisation process. The constraints on investment planning of all types of components are given as

$$P_{\text{PV,rate}}(y_i) \geq P_{\text{PV,rate}}(y_i - 1) \qquad (6.19)$$

$$P_{\text{WT,rate}}(y_i) \geq P_{\text{WT,rate}}(y_i - 1) \qquad (6.20)$$

$$\dot{Q}_{\text{HP,rate}}(y_i) \geq \dot{Q}_{\text{HP,rate}}(y_i - 1) \qquad (6.21)$$

$$E_{\text{BSS,cap}}(y_i) \geq E_{\text{BSS,cap}}(y_i - 1) \qquad (6.22)$$

$$Q_{\text{TES,cap}}(y_i) \geq Q_{\text{TES,cap}}(y_i - 1) \qquad (6.23)$$

In addition to the above technical constraints, another constraint is the financial budget. In this study, it is assumed that the investment plan has an Annual Investment Cost (AIC) limit, which is only used for investment and not includes the operational cost. The investment cost

in each year can be used to increase the capacity of any component, but the total cost cannot exceed AIC as

$$
\begin{aligned}
\text{AIC} \geq & \Delta P_{\text{PV,rate}}(y_i) \cdot U_{\text{PV}}(y_i) + \Delta P_{\text{WT,rate}}(y_i) \cdot U_{\text{WT}}(y_i) + \Delta \dot{Q}_{\text{HP,rate}}(y_i) \cdot U_{\text{HP}}(y_i) \\
& + \Delta E_{\text{BSS,cap}}(y_i) \cdot U_{\text{BSS}}(y_i) + \Delta Q_{\text{TES,cap}}(y_i) \cdot U_{\text{TES}}(y_i)
\end{aligned}
\tag{6.24}
$$

where $\Delta P_{\text{PV,cap}}$, $\Delta P_{\text{WT,cap}}$ and $\Delta \dot{Q}_{\text{HP,rate}}$ denote the increased rated power of PV, WT and HP and capacity of BSS and TES of each year with the unit of kW. $\Delta E_{\text{BSS,cap}}$ and $\Delta Q_{\text{TES,cap}}$ denote the capacity of BSS and TES of each year with the unit of kWh. $U_{\text{PV}}(y_i)$, $U_{\text{WT}}(y_i)$, $U_{\text{HP}}(y_i)$, $U_{\text{BSS}}(y_i)$ and $U_{\text{TES}}(y_i)$ denote the unit cost of each component in year y_i with the unit of $£/kW$ or $£/kWh$.

The forecasted component unit price in future years is obtained based on different financial assumptions in recently published papers [234–236]. In this study, the financial forecast for the unit cost of components for 2020-2050 is in British pounds, as shown in Figure 6.7. The prediction shows that the unit prices of PV, WT, HP, BSS and TES will continue to fall in the next few decades, while the unit price of natural gas and CO_2 will increase rapidly in the coming decades.

The price of GHG is based on a probabilistic assumption of GHG for possible future climate damage, used to balance consumption today against unknown damages in the distant future [237]. The CO_2 price used in this paper is the price path in the literature based on probabilistic assumptions of climate damage [234].
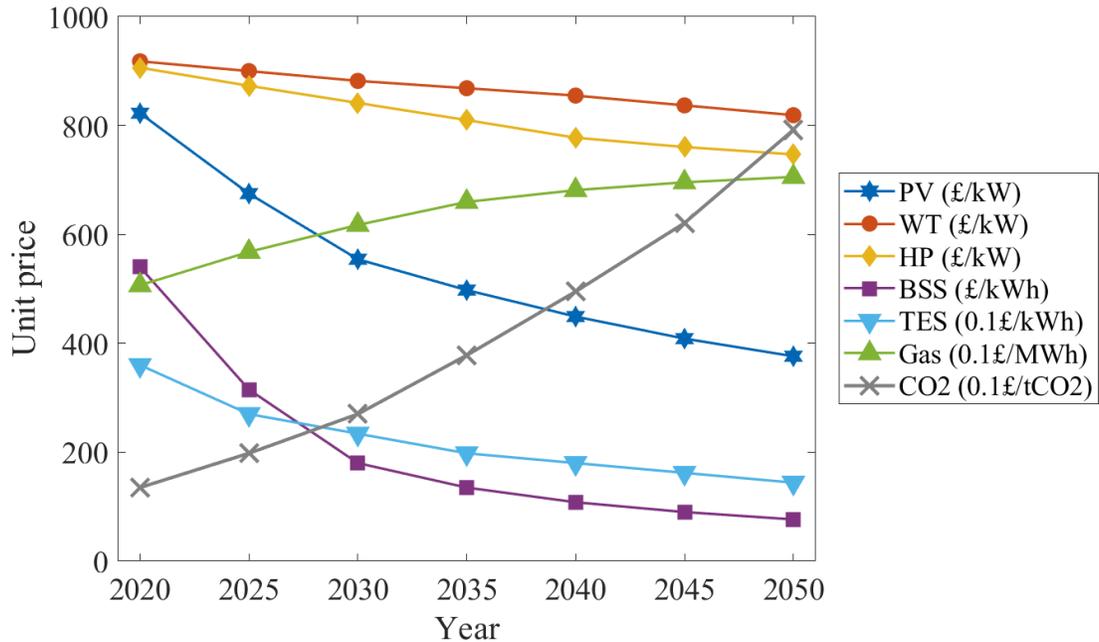
Figure 6.7: Financial assumptions for components of future years.

6.3.2 Optimisation approach

From the literature in the past decades, there are four main optimisation approaches that are commonly used, including the direct search, calculus-based optimisation, genetic algorithm (GA), and particle swarm optimisation (PSO). Direct search is a straightforward optimisation method that does not consider the derivatives of inputs. Thus, the direct search method can be applied in optimising many nonlinear functions that have less dependence on its derivatives [176]. The most commonly used Calculus-based Optimisation method is to set the gradient of the objective function to zero [40]. Another calculus-based optimisation method is Newton's method, which is similar to the Steepest Descent Method that uses an iterative process from an initial guessed starting point to finally converge to the optimum point. The GA is a population-based algorithm to search for the global-optimum solution. The iterative process of GA will converge to better solutions based on the breeding of the parents with higher performance [40, 177]. However, the GA has shortages that it requires a large amount of non-optimal data as the requirement of global search [40]. The PSO is a population-based metaheuristic algorithm that attempts to find the global optimal solution of the optimisation problem by simulating the social gathering behaviour of animals [40, 238]. This chapter chooses the PSO method to find the optimal sizing of RHS due to its strength in finding the global optimal solution and easy to implement.

PSO method uses a large number of the swarm to search the optimum point globally and share the information among all swarm for the next search step. In the PSO algorithm, each feasible solution is called a particle, which is specified by a vector containing problem variables [122]. Particles can remember their previous optimal position and share it with others [239]. The motion of each particle is composed of two randomly weighted influences and an initial random velocity [239]. The PSO algorithm simulates the sociality of particle, combines the best position of the entire swarm and the trend of its own movement, effectively avoids the particles and swarm falling into the local optimal solutions [240].

In the PSO algorithm, the state of each particle is represented by its position x_j and velocity v_j. The velocity update equation of the PSO algorithm has three key parameters, including the inertia constant w, the acceleration constant c_1 that controls the direction of the particles toward its best position in history, and the acceleration constant c_2 that attracts the particles to the best position of the swarm. The formula for updating the velocity and position of each particle in the space can be expressed as [122, 241].

$$v_j^{k+1} = w^k \cdot v_j^k + c_1 \cdot r_1^k \cdot \left(p_{best,j}^k - x_j^k \right) + c_2 \cdot r_2^k \cdot \left(g_{best}^k - x_j^k \right) x_j^{k+1} = x_j^k + v_j^{k+1} \qquad (6.25)$$

where v_j^k and x_j^k are the velocity and position of the jth particle in the kth iteration, $p_{best,j}^k$ is the best position achieved by the jth particle and g_{best}^k is the best position of the swarm, r_1 and r_2 are the random factors between 0 and 1.

6.4 Case Study of Heat Pump Powered by Grid

Heat demand is the most important input parameter for any heating system optimisation. It is essential to obtain good performance of the real system. A heat demand model needs to be established to predict the required heat through weather conditions such as ambient temperature. Modelling and simulation of energy systems are usually implemented using engineering models or data-driven models. The simulation test uses energy data from the Glasgow University campus as a case study. In previous chapters, both the engineering modelling approach [99] and data-driven approach [100] have been used to predict the heat demand of campus buildings. The heat demand of the campus and the corresponding ambient temperature data are shown in Figure 6.8.
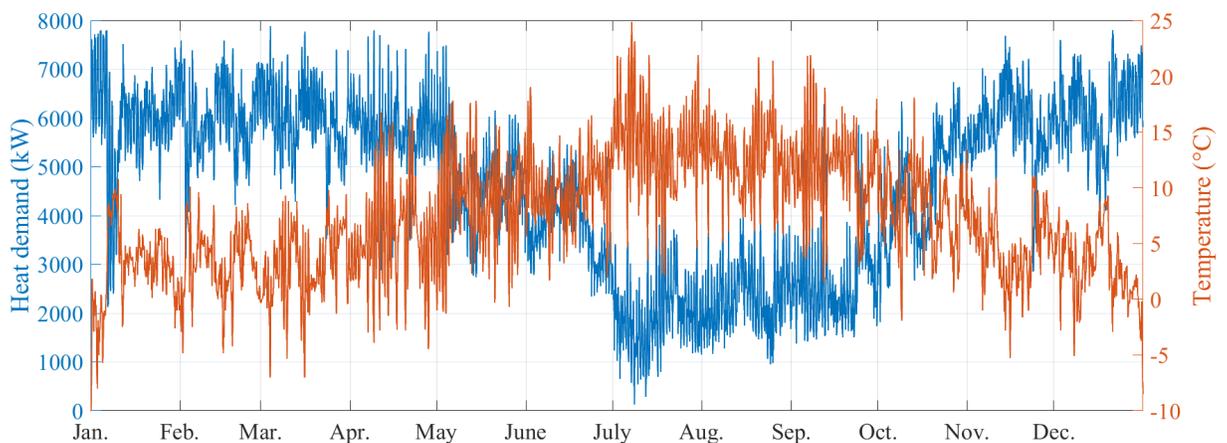
Figure 6.8: Heat demand of university campus versus ambient temperature.

The hourly temperature obtained from the nearest weather station shows that the local annual temperature fluctuates between -10°C and 25°C. The heat demand data is recorded at the University Energy Centre, given as a blue line, varying from 0 kW to 8000 kW. From the data, the temperature is lower in winter and higher in summer, while the heat demand is the opposite.

6.4.1 Size of heat pump

In the grid driven RHS, the core equipment is the heat pump, which can significantly reduce the use of natural gas and carbon dioxide emissions in building space heating. In this case, it is assumed that all the heat demand is provided by HP, and its input electric power is completely supplied by the grid. The operating costs of HP refer to electricity tariffs of the grid. In this case, two types of electricity prices are considered: standard and economy. The standard electricity price gives a constant unit price of 0.144 GBP/kWh throughout the day. The economic electricity price considers the off-peak hours from 0:00 to 07:00 every day, and the unit price in this period is 0.0765 GBP/kWh. Other durations are considered as peak hours, whose unit price is 0.1747 GBP/kWh [141]. The electricity tariffs of Standard type and Economy type are shown in Figure

6.9.
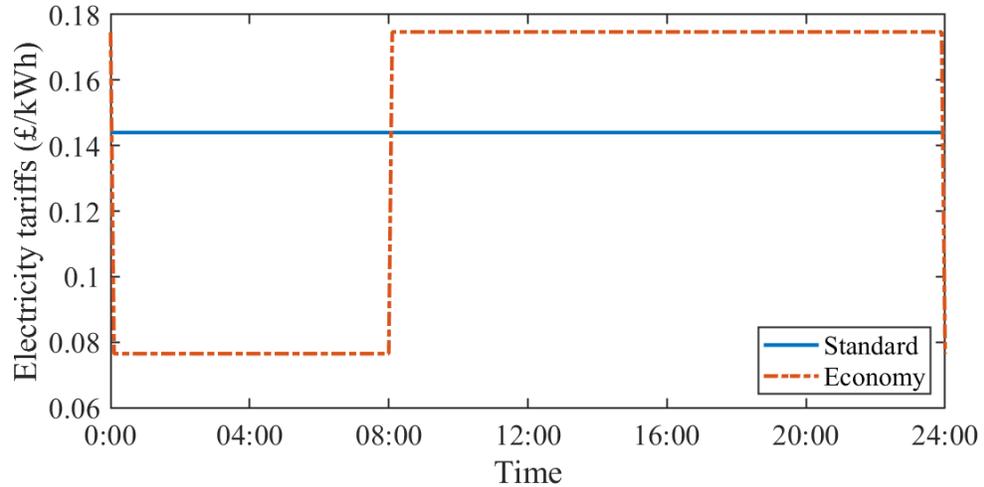
Figure 6.9: The different electricity tariffs, standard and economy with off-peak duration.

Using Economic tariff, BSS can be used to store electricity during off-peak hours and release the stored electricity during peak hours to assist HP. Under this strategy, the higher the BSS capacity, the greater the reduction in electric operational cost (EOC). However, if the BSS capacity is large enough, the stored energy may not be fully utilised. Therefore, in this case, it is necessary to find the BSS capacity that best suits the heat demand. Figure 6.10 shows the results of reducing EOC by increasing BSS capacity under Economic tariff. From the results, it can be found that when the BSS capacity reaches about 50MWh, the lowest power operating cost can be achieved.
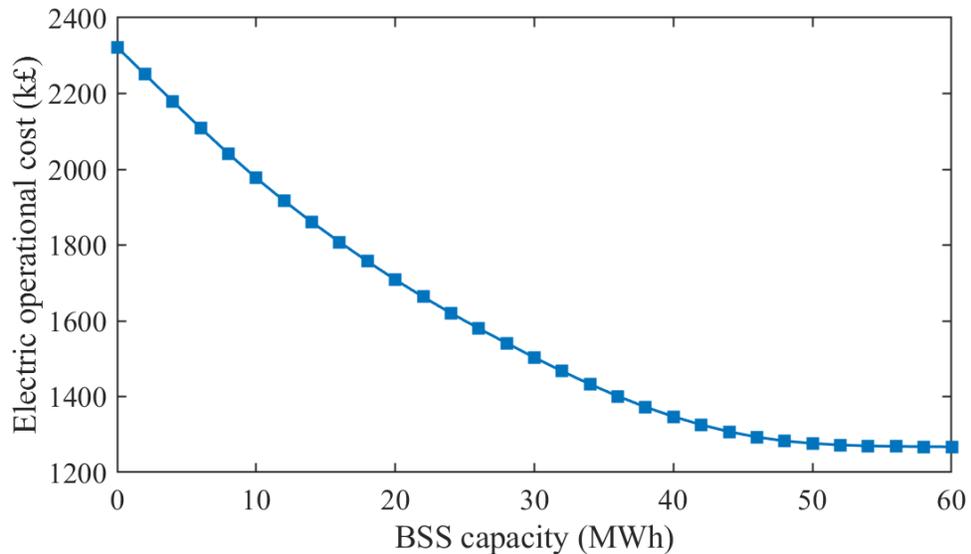
Figure 6.10: Reduce electric operational cost by increasing BSS capacity under Economic tariff

Figure 6.11 shows the monthly EOC of using HP to heat all buildings under the three different tariff strategies. It can be found that the EOC under the Economy tariff is slightly lower than
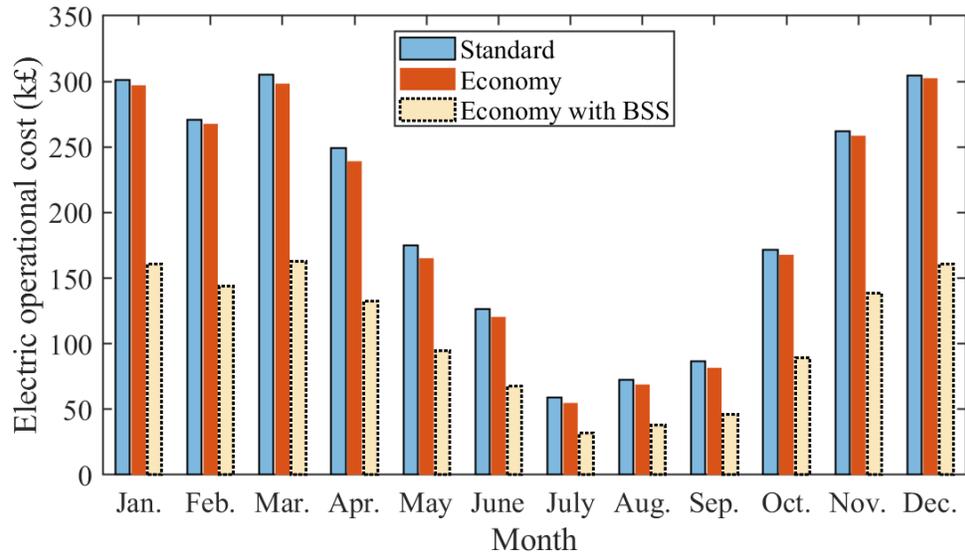
Figure 6.11: Monthly electric operational cost of heat pump refers to different types of electricity tariffs.

the Standard tariff, but not obvious. However, after using BSS, the EOC can be significantly reduced, as shown in Figure 6.11. In quantitative comparison, Economic tariff can reduce the EOC by 3.4% compared to Standard tariff, while Economical tariff with BSS can reduce the EOC by about 46.8%.

To show in details how the size of HP affecting the monthly electricity demand and CO_2 emission, Figure 6.12 shows the surface of montly required electricity from grid and monthly CO_2 emission under to different size of HP. For example, in the original CHP/Boiler heating system that has no HP applied, the required electricity from grid is less and mainly in summer, similar with the result shown in Figure 6.3. But the CO_2 emission of the original heating system with no HP applied is much higher, especially in winter time. This is because in the original heating system, the electricity demand is mostly supplied by the CHP together with the heat generation. With the increased size of HP, the required electricity from grid is increased but the CO_2 emission reduced accordingly, especially in winter time that has large amount of heat demand.

In order to compare them directly, the yearly amount of electric demand from grid and CO_2 emission with the increased size of HP are shown in Figure 6.13. When the HP size increased to about 7 MW, the CO_2 emission can be reduced to zero and the electricity demand from grid will increase to about 40GWh in the whole year.

6.4.2   Operational cost under different electricity tariffs

The electricity tariffs include three types, the Standard tariff, the Economy tariff with 7 hours off-peak duration (E7), and the Economy tariff with 10 hours off-peak duration (E10). The off-peak hours for E7 are from 00:00 to 07:00, while the off-peak hours for E10 are between 00:00

(a)

(b)
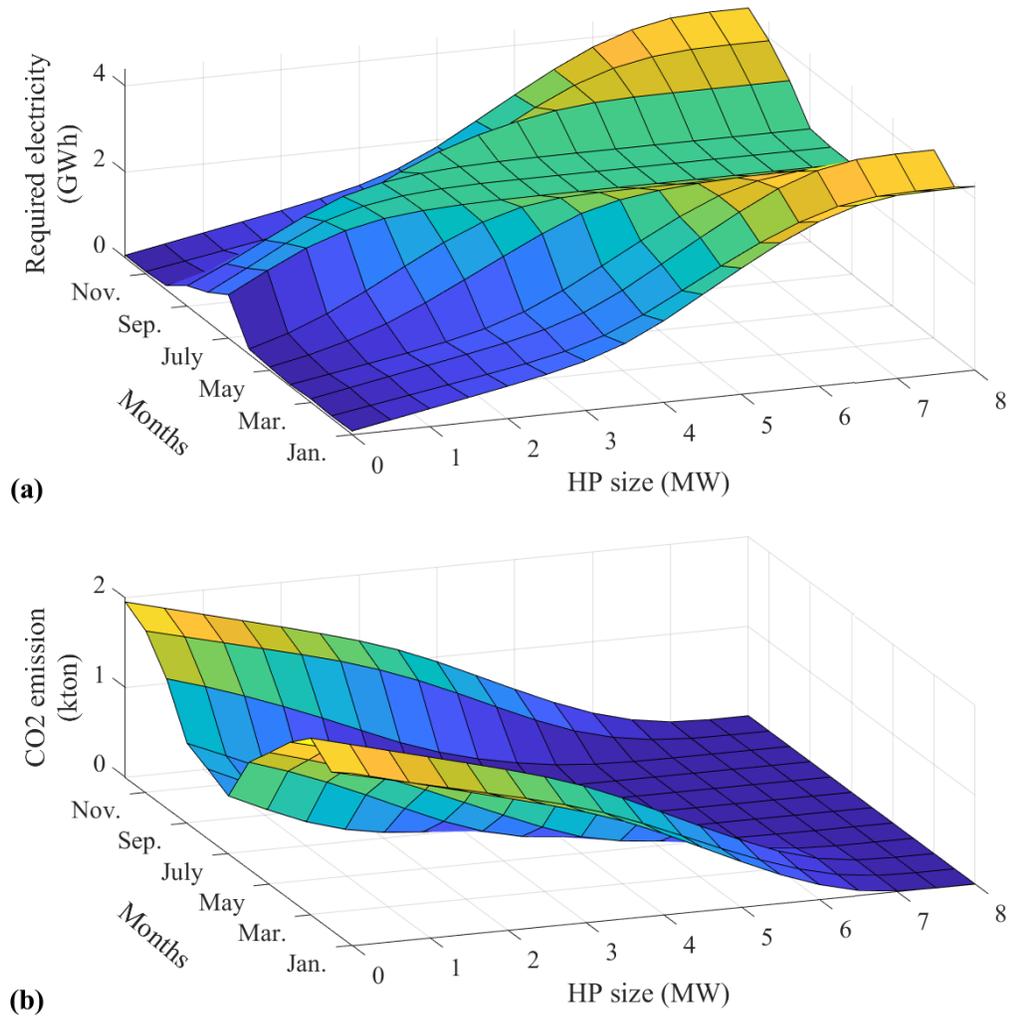
Figure 6.12: Monthly additional electricity demand from grid and CO_2 emission under different size of HP.
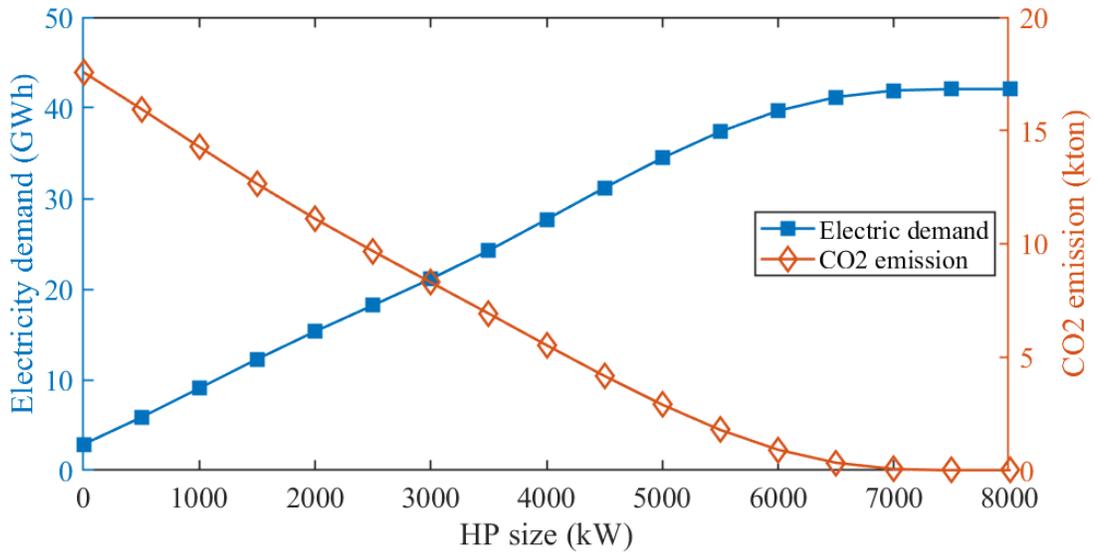
Figure 6.13: Total electric demand and CO_2 emission with the increased size of HP.

to 05:00, 13:00 to 16:00 and 20:00 to 22:00. The standard electricity price gives a constant unit price of 0.144 GBP/kWh throughout the day. In the E7 tariff, the price at off-peak hours is 0.0765 GBP/kWh while the price of on-peak hours is 0.1747 GBP/kWh. In the E10 tariff, the price at off-peak hours is 0.071 GBP/kWh while the price of on-peak hours is 0.1744 GBP/kWh [141]. The electricity tariffs of Standard, E7 and E10 are shown in Figure 6.14(a).

The electricity operational cost of power from grid of different tariffs are shown in the bar chart in Figure 6.14(b). Due to different off-peak hours and price in E7 and E10, the tariff E10 can get cheaper operational cost of electricity from grid. From the result, E7 and E10 can reduce the operational cost by approximate 1% and 15%, respectively.

In addition to different tariffs, the use of local renewable resources and energy storage can further reduce the operational cost. The capacity of wind turbine and solar PV are 3.5 MW of each component and the energy storage is chosen as 7 MWh. Using Economic tariff, BSS can be used to store electricity during off-peak hours and release the stored electricity during peak hours to assist HP. The operation strategy of the storage is controlled to be fully charged during off-peak hours and discharged during on-peak hours. The results of electricity operational cost from grid under different tariffs are by using renewables and storage are shown in Figure 6.14(c). It can be found that the E7 and E10 tariffs reduced the operational cost by 7% and 30%, respectively, comparing with the Standard tariff.

The installation of renewable power generation and energy storage devices can bring additional cost of the whole system. However, due to the unit price of gas and CO_2 increase year by year as the prediction shown in Figure 6.15(a), the total save on operational cost will increase within future years. Figure 6.15(b) shows the save of operational cost of E7 and E10 tariffs from 2020 to 2050 based on the prediction of unit cost of gas and CO_2. It can be found that the E10 can save more than 4.5 million pounds by 2050 due to the increased unit price of gas and CO_2. This could significantly reduce the average investment cost of renewable power generation and energy storage devices and encourage the installation of heat pump based renewable heating system in terms of economic benefits.

6.5 Case Study of Grid Independent Renewable Heating System

In the previous simulation case, it can be found that the use of HP will bring additional power load to the grid. This shows that although the use of HP can greatly reduce CO_2 emissions, there is a potential risk of overloading or even fault occurring in the power grid. Therefore, the most promising alternative is to use local renewable resources to generate electricity to drive HP to heat buildings. This will reduce the impact of the large amount of electrical energy required by HP on the grid. This section considers using local renewable energy to provide power for heat pumps, which is known as the grid independent renewable heating system (RHS) including
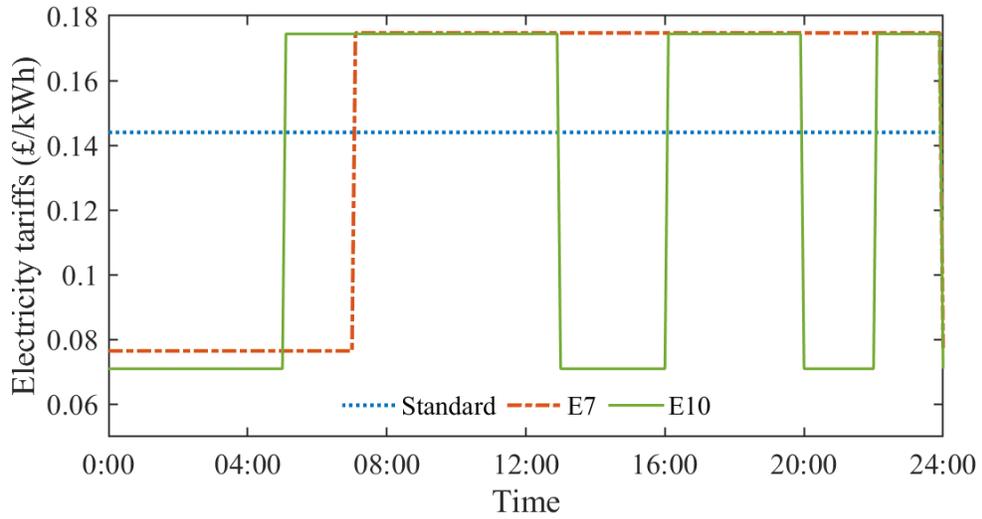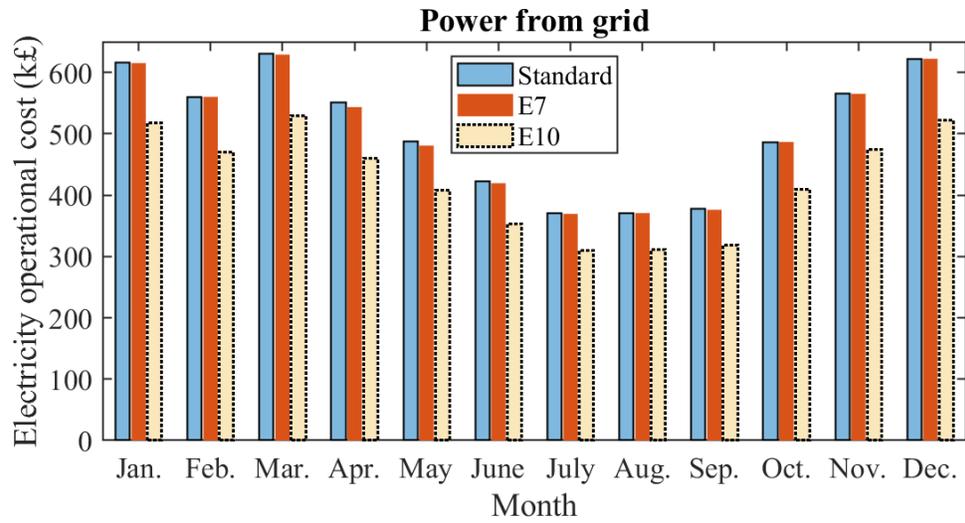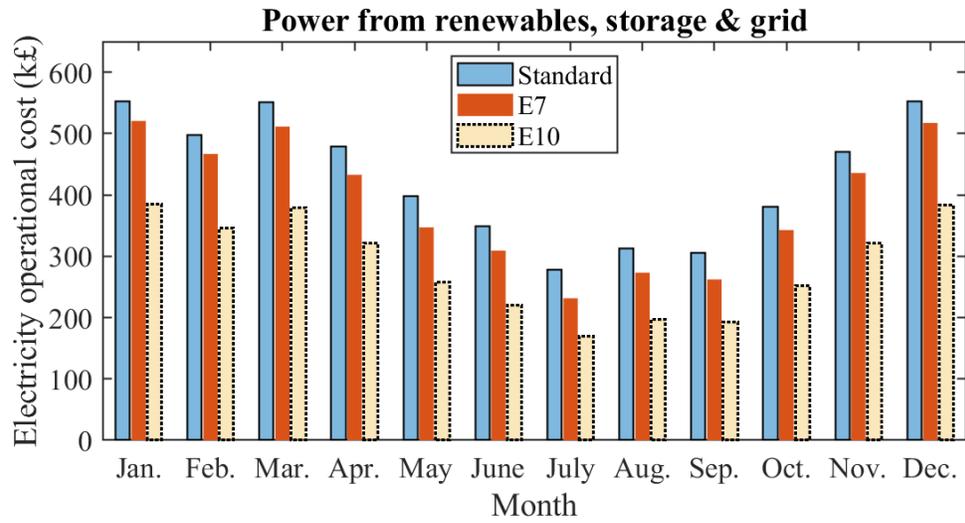
Figure 6.14: Electricity tariffs and operational cost based on different profile. (a) Electricity tariffs, (b) operational cost of electricity all from grid, (c) operational cost of electricity with renewable energy and battery storage system.
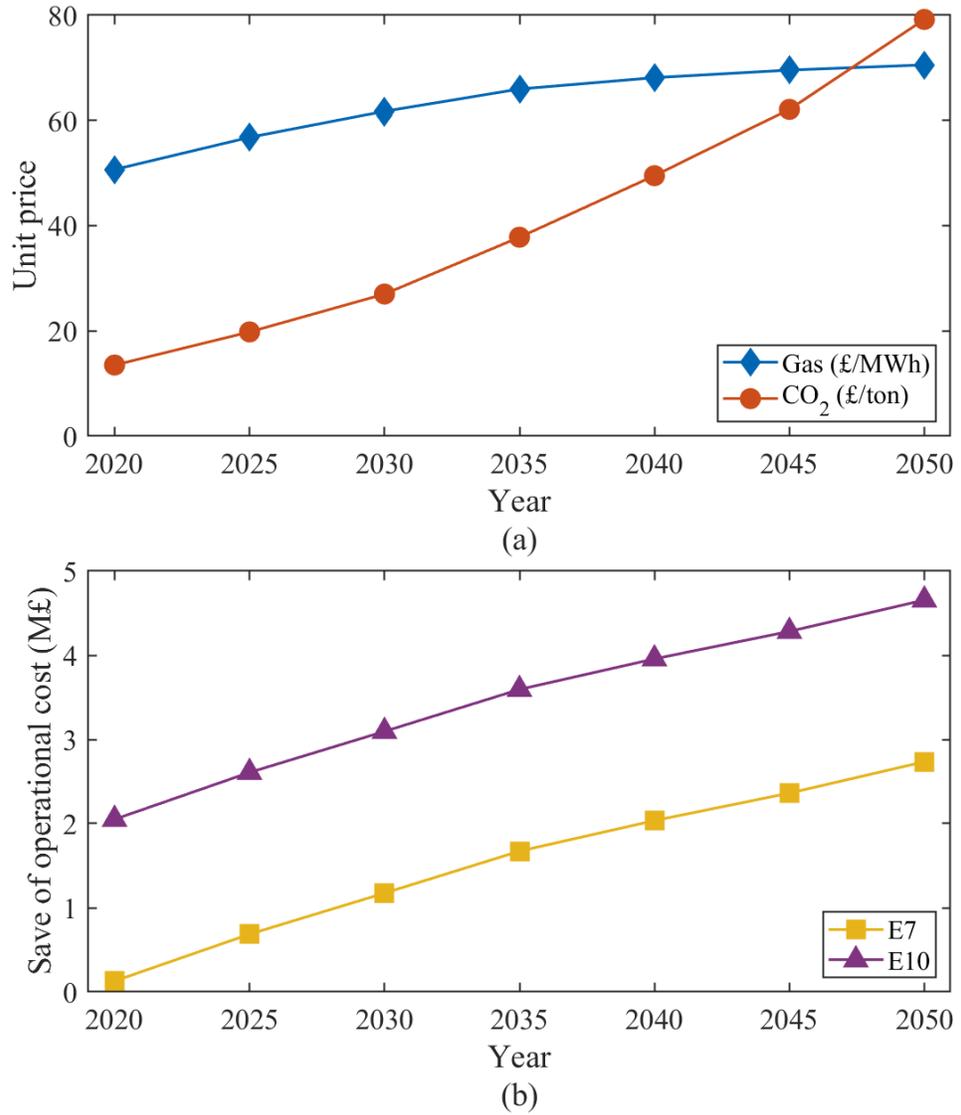
Figure 6.15: Unit price of gas and CO_2 in future years till 2050 and predicted financial save of operational cost by using renewable integrated energy network comparing with the traditional energy network.

photovoltaic, wind turbine, battery storage system and thermal energy storage [242]. The district heating network of the traditional heating system and the proposed RHS is shown in Figure 6.16.
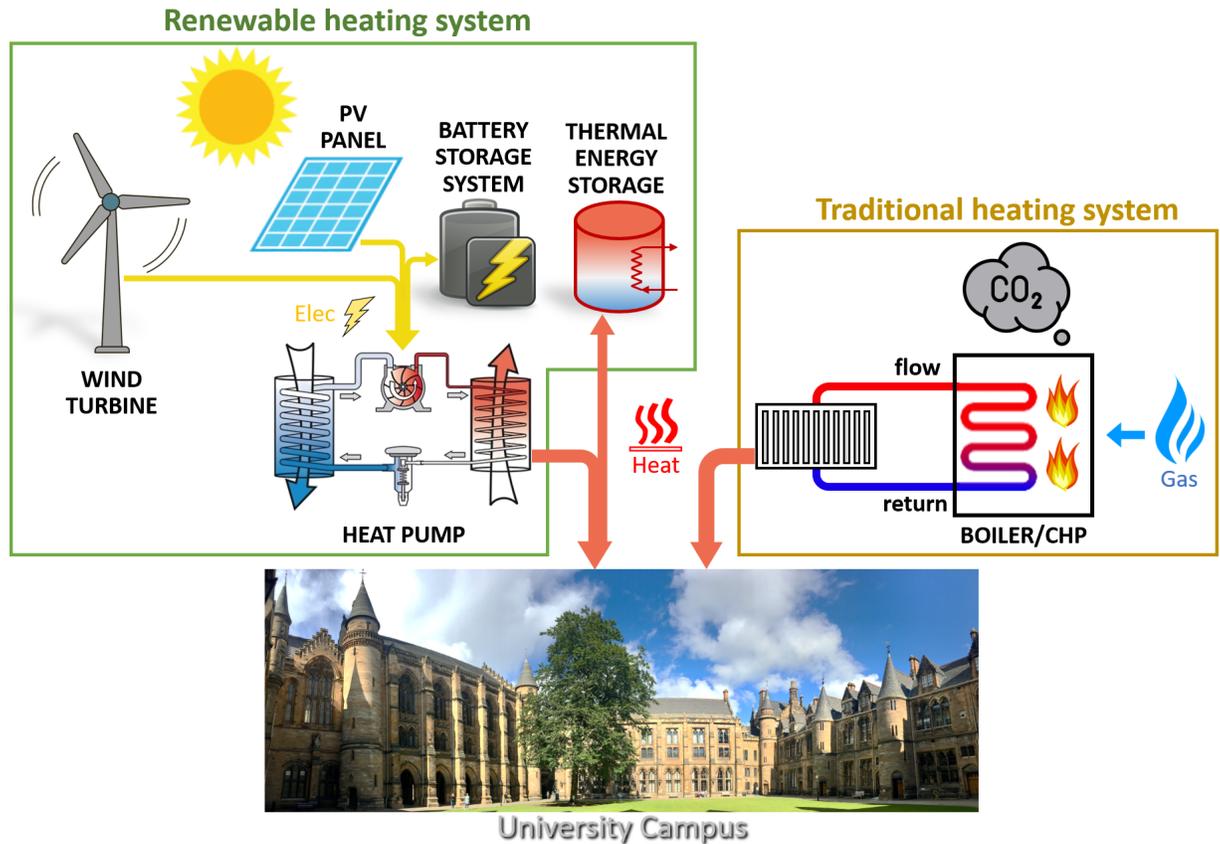
Figure 6.16: Grid independent heating system of university campus.

6.5.1 Gas consumption and required size of storage

In this case, it is assumed that the heating system is independent of the grid, and only uses renewable energy for power supply to the HP. The heating system combines the RHS based on local renewable resources and the traditional heating system using gas boiler. When renewable resources can generate enough power, the heat demand is completely supplied by HP, and the unconsumed power is stored in the BSS. When renewable resources driven HP cannot provide enough heat for the building, traditional gas boilers will be used to compensate for the remaining heat demand by consuming natural gas. The parameters of each component in the RHS are given in Table 6.1.

Figure 6.17 shows an example of monthly heat production when renewable generation capacity is 10 MW, where the rated power of PV and WT are both 5MW. Since this case aims to consider the role of the BSS rather than its actual size, the capacity of the BSS is assumed to be large enough to store all unused energy from renewable energy sources. The blue part, the green part, and the yellow part in the bar chart represent the heat generated by the HP driven by power from PV, WT, and BSS, respectively. It can be seen from the results that, due to the large changes

Table 6.1: Prediction result comparison among different number of parameters to be calibrated.

Parameters	Symbol	Value
WT – cut-in wind speed	V_{cut-in}	3.5 m/s
WT – rated wind speed	V_{rate}	14.5 m/s
WT – cut-out wind speed	$V_{cut-out}$	25 m/s
WT – air density	ρ	1.225 kg/m^3
WT – swept area	A_r	1735 m^2
WT – rated power	$P_{WT,rate}$	750 kW
PV – photoelectric conversion efficiency	η_{pv}	0.9
PV – area of unit array	A_{pv}	500 m^2
PV – rated power	$P_{PV,rate}$	350 W
BSS – self-discharge rate	σ	0.00005
BSS – charge efficiency	η_{ch}	0.95
BSS – discharge efficiency	η_{dch}	0.95
HP – linear regression parameter of COP	a	-0.066
HP – linear regression parameter of COP	b	5.7
HP – supply water temperature	T_{wat}	60 °C
TES – efficiency	η_{TES}	0.9
CO$_2$ – natural gas GHG emission	g_{gas}	241 g/kWh

in the illumination time and the angle of solar radiation in high latitude areas such as Scotland, PV power generation is higher in summer and lower in winter. WT has the opposite trend as it will produce more power in winter and less power in summer. From the results of each month, the heat demand of the building in July, August and September is less, so the heat demand can be fully supplied by HP. In addition, due to the low heat demand, the unconsumed power generated by renewable resources is stored in the BSS. Therefore, in October and November, most of the electricity is provided by the BSS that has stored electricity in the past few months. From December to April, the heat demand is high, but the power from renewable energy and BSS are low, so a large amount of heat is provided by gas boilers for building heating.

It can be seen from the above results that the natural gas consumption of gas boilers is related to the installed capacity of renewable energy. Therefore, the relationship between natural gas consumption and renewable energy installed capacity is shown in Figure 6.18(a). Without the use of BSS, the gas consumption decreases slowly as the installed capacity of PV or WT increases, as shown by the dashed lines. Under the same installed capacity, WT-driven HP performs better than PV-driven HP in reduction of gas consumption. The reason is that WT's power generation is large in winter and small in summer, which is more in line with building heat demand. However, PV power generation has a small amount of power generation in winter, so the power generation of PV in winter is not enough to support HP consumption. However, after using BSS with installed capacity in Figure 6.18(b), the time distribution of renewable resources has less impact, and the total power generation is more important. As the installed capacity of renewable energy increases, the consumption of natural gas decreases faster since unconsumed
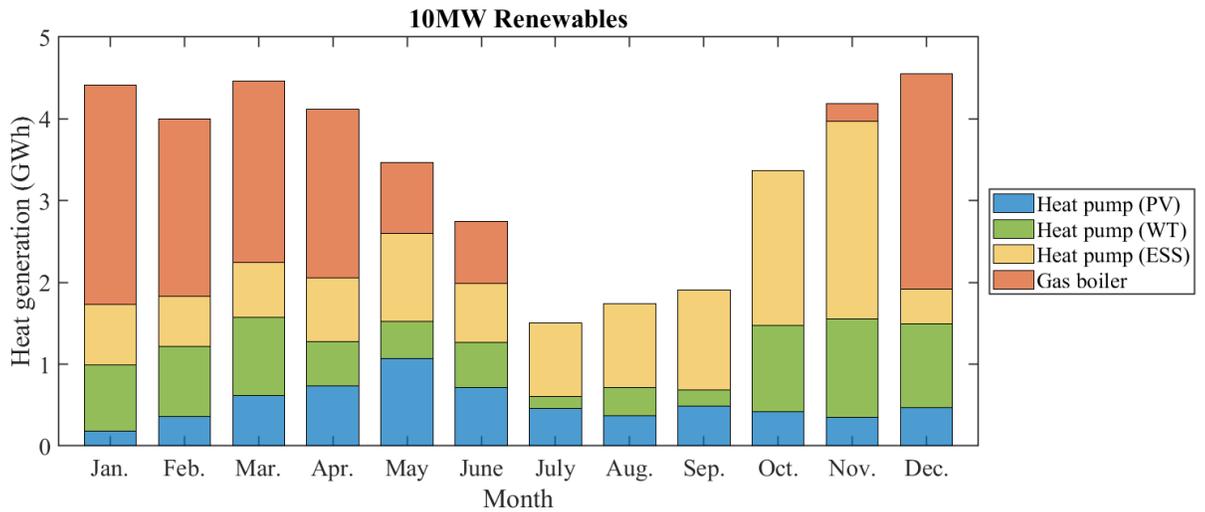
Figure 6.17: Example of monthly heat generation from renewables and gas boiler.

electricity is stored in the BSS and used to assist the renewable energy generation. With the same installed capacity, PV can generate more electricity than WT, and thus results in less gas consumption. In this case, HP driven by PV can reduce gas consumption faster than HP driven by WT. However, due to the time mismatch between heat demand and PV power generation, the BSS capacity required for PV is much higher than that of WT. The required capacity of BSS corresponding to different renewable energy installation capacity is shown in Figure 6.18(b). When the PV installed capacity reaches 15 MW, the gas consumption can be reduced to zero. The corresponding required BSS is nearly 5 GWh, which is a very high capacity, while the same installation capacity of WT requires the BSS for just 1 GWh, which is still a massive capacity. Therefore, this result only considers the technical requirements and does not consider the financial feasibility.

The unit of gas flow is in m^3. The convertion of gas units m^3 into kWh can be calculated by mutiplied by volume correction factor (1.02264) and calorific value (39.39355) and divided by conversion factor of kilo per hour. The final conversion value of gas unit from m^3 to kWh is 11.19 kWh/m^3. The units of measured gas flow is in kNm^3, which is 'Kilo Normal Meter Cubed'. The Normal refers to normal conditions of 0 ∘C and 1 standard atmosphere, which equals to 101.325 kPa.

The results in Figure 6.18 are for cases with only a single renewable generation technology, i.e., either PV or WT but not at the same time. The result shows that PV driven HP is better in terms of total power generation but requires more BSS for energy storage. For this reason, another method is to install PV and WT at the same time to combine their advantages. The gas consumption

The gas consumption and the required BSS corresponding to the PV and WT at different rated powers are shown in the 2D surfaces of Figure 6.19(a) and (b). When the rated power of PV and WT is greater than 7MW, only 2.5 GWh of BSS can reduce natural gas consumption to
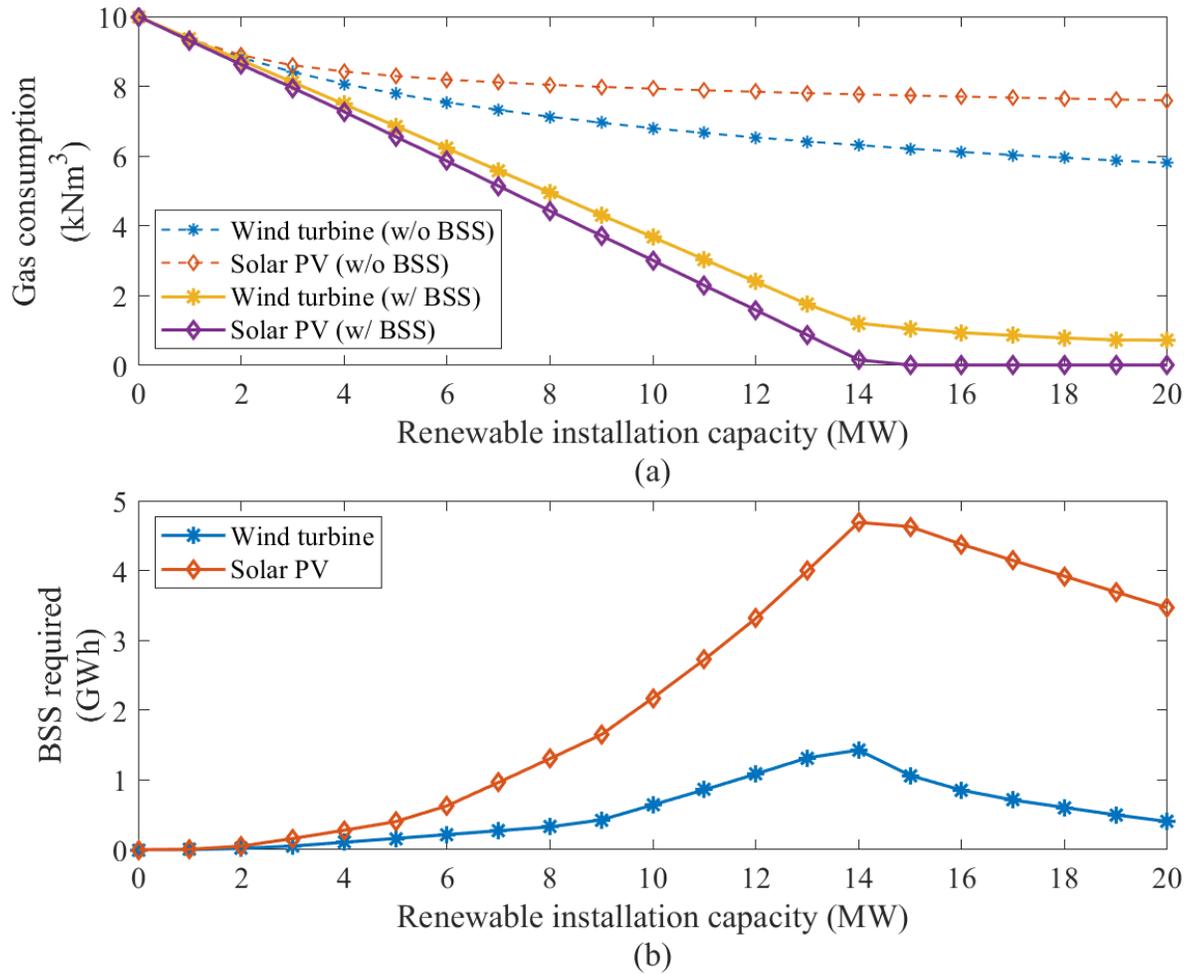
Figure 6.18: Boiler gas consumption and required BSS capacity according to the installation capacity of wind energy or solar energy.

zero, which is only half the installed capacity of PV alone.

6.5.2   Optimal sizing considering CO_2 emission and financial cost

The previous case only considered the technical requirements for the selection of components in RHS. This section considers the financial cost in reducing CO_2 emissions and optimises the investment capacity of each component for every year. The financial assumptions for the unit cost of each component in future years use the predictions given in Figure 6.7. Due to the limited AIC, more investment on one component means less investment on other components. Thus, its sizing needs to be optimised to find the optimal scheme to achieve the minimum CO_2 emission. The variables for optimisation include the capacity of PV, WT, BSS, HP, TES considering their constraints described in Section 6.3.1. For this multi-parameter optimisation problem, the optimisation method uses the PSO method described in Section 6.3.2 to find the most suitable component size to minimise the total CO_2 emissions. Since the financial assumption for the unit cost of each component in future years predicted in Figure 6.7 is in every five years, the optimi-
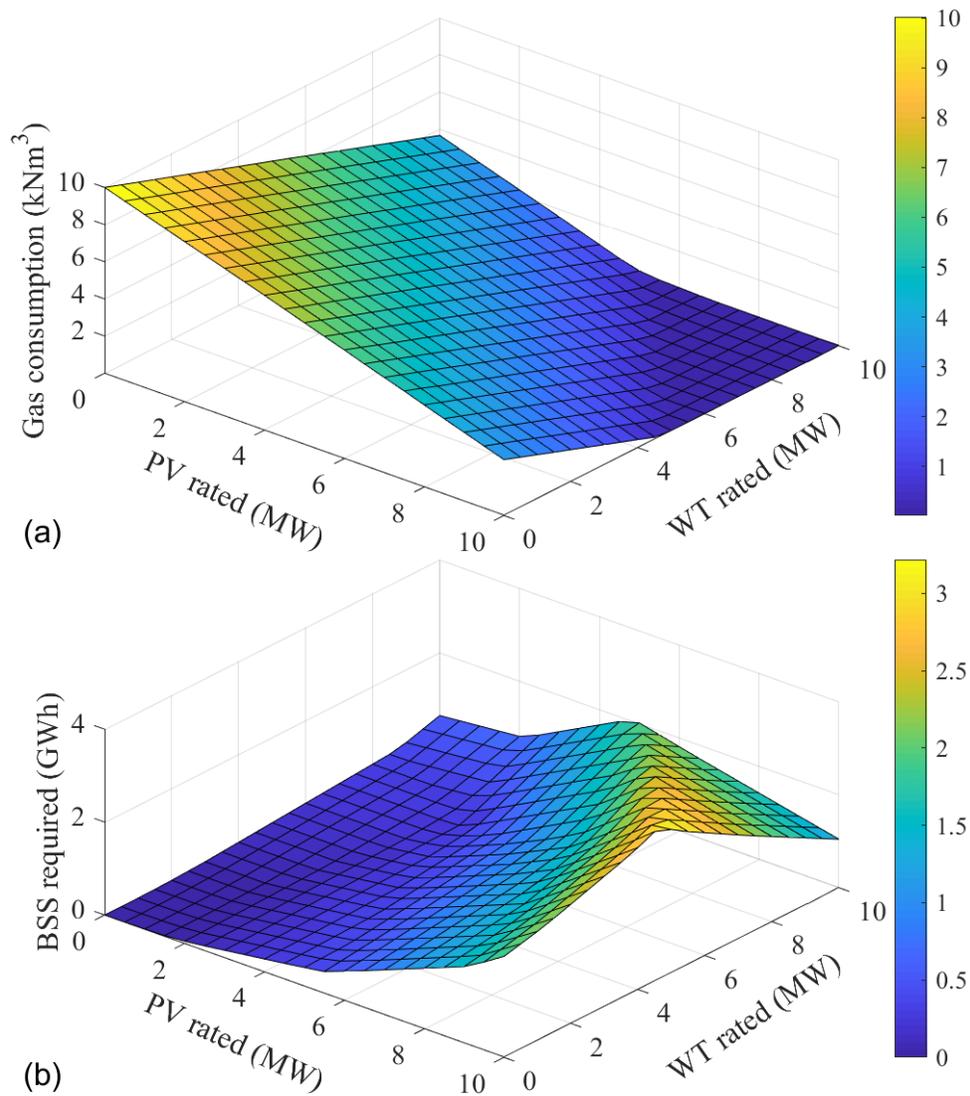
Figure 6.19: 2D surface of required ESS and boiler gas consumption according to rated PV and WT.

sation purpose is the total installed amount of each component in every five years. The indicator for the optimisation is the total amount of CO_2 emissions in five years. After that, the optimal size of each period is based on the installation amount of each component at the previous period and re-run the optimisation approach according to AIC. Therefore, the optimisation results of every five years are based on the optimal solution for the current period. It is ensured that the size of the components in the RHS will increase, and the CO_2 emissions will decrease year by year, and finally achieve the optimal result.
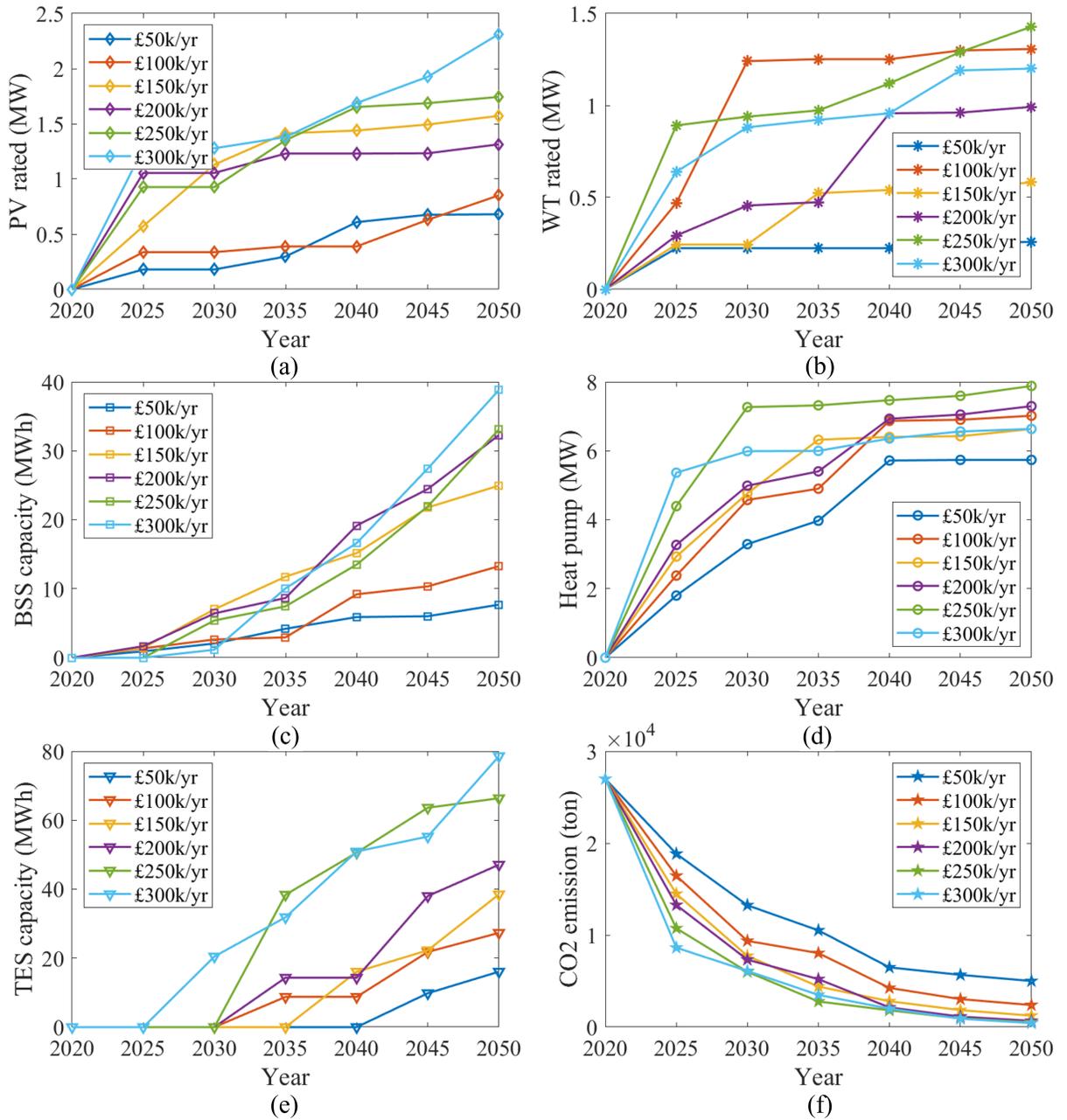
Figure 6.20: Optimisation result of installing renewable energy, heat pump and BSS, and the reduction of CO_2 emission.

The optimisation result is shown in Figure 6.20, which shows the annual investment capacity of each component in the RHS based on different investment budgets. For example, if the university plans to spend 100,000 pounds per year, as shown by the blue line, the optimal capacity of PV, WT, HP, BSS, and TES for investment according to the optimisation results are given in Figure 6.20(a)-(e). Using the optimal scheme, the least CO_2 emissions can be obtained, as shown in Figure 6.20(f). In the blue line, the CO_2 emission will still not be reduced to zero by 2050 but will be reduced from 2.7 tons/year to 0.7 tons/year. If the university plans to use the grid independent RHS to completely eliminate CO_2 emissions by 2050, it will need at least

400,000 pounds per year. The optimisation results give a preliminary idea of how to construct the RHS. From the results, in the first 10 years, HP, as the main component of RHS, has the fastest growth in its investment capacity. At least after 2030, the investment of TES will have an effective impact on reducing CO_2 emissions. The investment capacity of PV, WT and BSS is increased according to the local renewable energy resources and the unit cost of each component to achieve the optimal application of renewable energy.

Compared with different AIC, the more investment, the less CO_2 emissions are obtained, as shown in Figure 6.21(a). Numerically, if RHS is not invested, the traditional heating system will generate at least 8×10^5 tons of CO_2 from 2020 to 2050. The higher the AIC of RHS, the less total CO_2 emissions are obtained. If the AIC reaches £ 300,000 per year, the total CO_2 emissions from 2020 to 2050 can be reduced to 1.6×10^5 tons, which is only 20% of continuous using the traditional heating system.
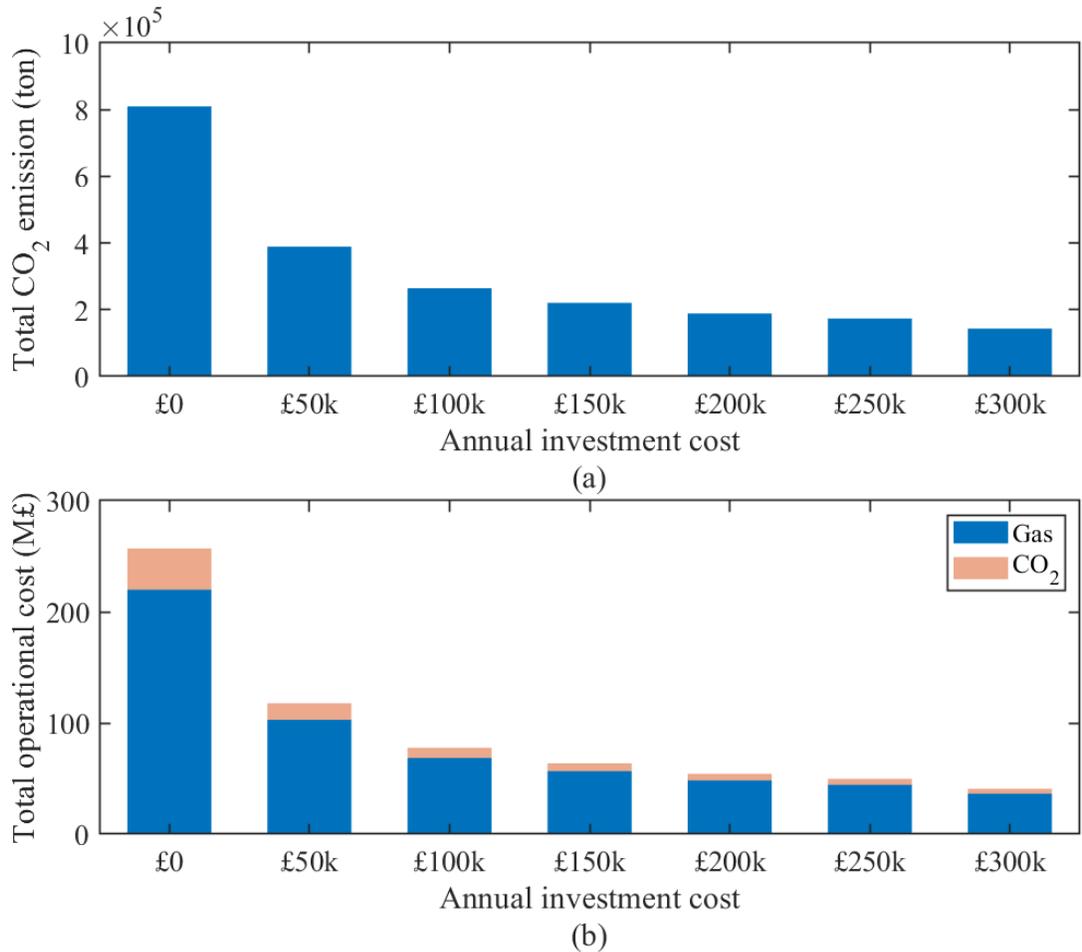
Figure 6.21: Total CO_2 emission and total operational cost versus different annual investment cost until 2050.

The estimated CO_2 emissions are based on the output of a district heating network model, which has been calibrated based on data from the past few years. The actual CO_2 emissions of the entire university are much more complicated and include commuting and business travel.

Therefore, the estimated CO_2 emissions in this article are only for preliminary verification of the feasibility and effectiveness of RHS. In addition, the investment cost shown in this chapter only considers the predicted unit price of devices. This makes the price and investment amount of the RHS system look much lower than expected. The actual cost will also include other cost, such as the installation fee, maintenance fee, ground rent, etc. These will greatly increase the complexity of prediction and optimisation, so this chapter does not consider these factors.

In addition to the total CO_2 emissions, the operating cost of natural gas and CO_2 is another cost of the heating system, as shown in Figure 6.21(b). If a traditional heating system is used, the total operating cost of natural gas and CO_2 is approximately 260 million pounds. And if the AIC of RHS reaches £ 300k per annum for 30 years from 2020 to 2050, the total operating cost will be reduced to 36 million pounds, which is just 14% of the operating cost of the traditional heating system. This shows that the investment of RHS can reduce operating costs as well, which is a further benefit and profit for this investment.

6.6 Conclusion

This chapter proposed a complete RHS framework, considering PV, WT, BSS, HP and TES. In the RHS, it is analysed the reduction of gas consumption and the required BSS under the corresponding installed capacity of renewable components from their technical requirements. Meanwhile, in the context of British tariffs and government incentives, the size of each component in the system is optimised through the PSO algorithm according to different AICs. The results verify that the optimal size of RHS provided by this approach can minimise CO_2 emissions and reduce the operating cost of natural gas. This provides a preliminary feasibility plan for how to invest in RHS to replace the traditional heating system in the university campus. In the future work, we will design and analyse the operation scheme and control strategy of the RHS system as well as more detailed HP models. This aims to increase the efficiency and reduce the loss of the RHS system in order to further reduce the carbon emissions and help to alleviate the global warming issue.

Chapter 7

Conclusion and Future Works

7.1 Conclusion

In this thesis, the integrated urban energy system is modelled and optimised in multi processes. As a target system, the heat and power demand of University of Glasgow is considered as a case study. The building heat and power demand prediction is essential for building energy consumption management and GHG emission estimation, as well as the integration of renewable energy resources.

In the building heat demand modelling, the engineering approach is demonstrated in Chapter 3 to predict the hourly heat demand of a district-level heating network. The developed model is calibrated by the proposed Bayesian-based calibration method to fit the monthly recorded data of heat demand in the university campus. The model development is on a platform of the IES-VE building energy software, which is used to build the bottom-up model of buildings in a district heating network. Due to the lack of detailed building parameters and heavy computational load of simulating such a large model, a Bayesian history matching based calibration method was proposed to calibrate the model over a shorter time span without the requirement of gradient calculation. In the approach, a statistical method-based emulator has been designed to estimate the conditional expectation and variance of model mismatch at untested points for the calibration iterations. The designed emulator-based calibration method has been compared with a traditional golden search algorithm, and the results show that the Bayesian approach-based emulator performs better with fewer calibration times to find the optimal point. Another case study has been designed to validate the proposed method in calibrating the single parameter, two-parameters, and multi-parameters of building envelope. The results indicate that the method is reliable and efficient to calibrate thermal parameters in building energy models. The multi-parameters calibration gives better performance by calibrating all parameters simultaneously. Finally, it is verified that the proposed approach can predict the hourly heat demand of a DHN without the requirements of detailed building information and accurate recorded data.

The statistical approach of building heat demand modelling is demonstrated in Chapter 4,

which used the ANN technology to predict the heat demand according to weather conditions. In predicting the district heat demand using weather information, it is found that most input variables contain duplicated features which is not required to train the ANN model. In the developed ANN technology, a new sensitivity analysis approach is proposed to remove both the variables with low importance and the variables that have high importance but contain duplicated features. The proposed approach analysed the correlation among input variables via detecting the coefficient of determination of each variable with others referring to the fitting error of quadratic least square regression. The approach is validated in a case study of predicting heat demand in a district using an ANN model that is trained by historic data from several sample buildings. The traditional sensitivity analysis method ranked the input variables based on their influence on the heat demand. It was shown that the 15 most important features can be used to predict the district heat demand with the same or even better performance than the complete set of features. The proposed method further removed 3 important variables that are determined by other variables via analysing the determination of each variable. The results show that the proposed method can reduce training time by around 20% while achieving the same training and prediction performance compared with the traditional sensitivity analysis method. With the developed sensitivity and correlation analysis approach, the training data is simplified and the efficiency of training an ANN model can then be improved.

In the building electricity demand modelling, the statistical approach is demonstrated in Chapter 5 to predict the electricity demand of each campuses and their hourly occupancy rate. In this study, it is found that the electricity demand of office buildings is in negative and positive correlation with weather variables in long-term and short-term time horizon, respectively, as a result the conventional ANN approach cannot accurately capture the relationships between them. This can be a typical issue in office/education buildings heated by electricity, especially for buildings in cold areas, such as in Scotland. In order to solve this issue, two electricity demand prediction approaches have been developed. The initial proposed approach splits the power demand data by working hours and non-working hours to avoid the impact of occupants' activities to building power demand. Using this method, the linear regression approach is used to fit the building baseload power to a weather variable using the non-working hours data and find the average occupants power demand using the data of working hours. To fit the power demand with more weather variables, the proposed approach is further developed to use ANN to fit the non-working hours data and working hours data and the membership function of fuzzy hours between them. With the second proposed approach, more weather variables can be considered in the model to predict the power demand more accurately. In the simulation results, both approaches have been validated to show less RMSE value than the conventional ANN approach in predicting the power demand. In addition, the ANN with fuzzy hours splitting approach has the best performance among the three approaches and reduces RMSE by 5% to 30% compared with the working hour splitting based regression approach and reduces RMSE by 30% to 55%

compared with the conventional ANN. Therefore, both approaches are able to solve the issue that the input and output fitting data are in negative and positive correlation in long-term and short-term time horizon, respectively. The proposed approaches can achieve good performance with RMS prediction error as low as 6% in building power demand prediction.

The building space heating and electricity energy consumption are integrated into a virtual integrated energy network (IEN) in Chapter 6. In the virtual IEN, a renewable heating system (RHS) framework is proposed, considering the electric heat pump, renewable resources, such as photovoltaic and wind turbine, as well as the energy storage systems, such as battery storage system and thermal energy storage. In the RHS, it is analysed the reduction of gas consumption and the required battery storage under the corresponding installed capacity of renewable components from their technical requirements. Meanwhile, in the context of British tariffs and government incentives, the size of each component in the system is optimised through the PSO algorithm according to annual investment costs and total CO_2 emissions. The results verify that the optimal size of RHS provided by this approach can minimise CO_2 emissions and reduce the operating cost of natural gas. This provides a preliminary feasibility plan for how to invest in RHS to replace the traditional heating system in the university campus.

In summary, the major contributions of this thesis can be summarised as follows:

- Developed a prediction approach using a building energy model whose parameters are calibrated by Bayesian-based calibration method to match the recorded data of monthly heat demand. In the proposed calibration approach, an emulator is established to evaluate the untested values of thermal parameters using Bayesian method, and then use the evaluation results to search for the most suitable parameters value. The designed approach greatly accelerates the calibration speed. The method is used to calibrate a single parameter and multiple parameters of the building thermal energy models for a district heating network. After it has been verified with measured data, the developed calibration method is used to calibrate parameters of building energy models. The output of the calibrated model can predict the hourly building heat demand in district heating networks.

- Developed a sensitivity analysis approach in ANN to analyse the correlation between input variables and to detect the variables that have high importance but contain duplicated features. The proposed approach is validated in a case study that predicts the heat demand of a district heating network containing tens of buildings at a university campus. The proposed approach can detect and remove several unnecessary input variables and helps the ANN model to reduce its training time while maintaining the prediction accuracy. It indicates that the approach can be applied for analysing large number of input variables to help improving the training efficiency of ANN in district heat demand prediction and other applications.

- Proposes two occupancy rate estimation based approaches for fitting and predicting the

electricity demand of office buildings. The first proposed approach splits the electricity demand data into fixed time periods, containing working hours and non-working hours, to reduce the impact of occupants' activities. After finding the most sensitive weather variable to non-working hour electricity demand, the building baseload and occupant activities can be predicted separately. The second proposed approach uses the ANN and fuzzy logic techniques to fit the building baseload, peak load, and occupancy rate with multi-variables of weather variables. In this approach, the power demand data is split into a narrower time range as no-occupancy hours, full-occupancy hours, and fuzzy hours between them, in which the occupancy rate is varying depending on the time and weather variables. The proposed approaches can provide more information for building energy management, including the predicted baseload, peak load, and occupancy rate, without requiring additional building parameters.

- Investigated a complete renewable heating system framework and sizing the components to decarbonise building heating in an integrated energy network. Since grid-powered heat pumps may significantly increase the power demand of the grid, the framework considers using local renewable energy to provide power for heat pumps, which is known as the grid independent renewable heating system including photovoltaic, wind turbine, battery storage system and thermal energy storage. The relationship between the reduction of gas consumption and the requirement of BSS under the corresponding installation capacity of renewable components is analysed with their technical requirements. Then, according to different investment plans, this work uses the particle swarm optimisation algorithm for optimal sizing of each component in the RHS to find a solution to minimise CO_2 emissions. The RHS with optimal sizing can minimise CO_2 emissions and reduce the operational cost of natural gas. This work provides a feasible solution of how to invest the RHS to replace the existing heating system based on gas boilers and CHPs.

7.2   Future Works

- In the statistical modelling of heat demand and power demand, I used the ANN method based on the back propagation algorithm, which is the most basic method in artificial intelligence algorithms, for modelling and data training. In my future works, I will focus on more advanced artificial intelligence methods and deep learning algorithms applied to the prediction of heat demand and electricity demand to improve the accuracy of demand forecasting.

- In the optimal sizing of grid independent RHS, the PSO algorithm is used for the annual installation capacity of each component to minimise the CO_2 emission. In my future works, different optimisation algorithms, such as genetic algorithm, will be analysed, compared and evaluated to choose the most suitable optimisation methods for the topic.

- In the integrated heat and power network, this work provides a preliminary feasibility plan for how to invest in RHS to replace the traditional heating system. But the plan is basic, ignoring many detailed information, such as the influence of operation schemes and the nonlinear behaviour of the HP model. In my future works, I will design and analyse the operation scheme and control strategy of the RHS system as well as more detailed HP models. This aims to increase the efficiency and reduce the loss of the RHS system in a more practical situation to further reduce the carbon emissions and help to alleviate the global warming issue.

- In terms of model validation, I used the heat and electricity demand of the University of Glasgow as a case study in my work. However, the analysis and modelling of one system may have limitations as the types of buildings are all for office use. In my future works, I will try to extend the demand prediction methods to a wider range of applications, such as the demand for heat and electricity throughout the UK, in order to include different building types and profiles of heat and electricity demand.

Bibliography

[1] Danny Dorling. World population prospects at the un: our numbers are not our problem? *The Struggle for Social Sustainability: Moral Conflicts in Global Social Policy*, page 129, 2021.

[2] Nicholas P Webb, Nadine A Marshall, Lindsay C Stringer, Mark S Reed, Adrian Chappell, and Jeffrey E Herrick. Land degradation and climate change: building climate resilience in agriculture. *Frontiers in Ecology and the Environment*, 15(8):450–459, 2017.

[3] Frederica P Perera. Multiple threats to child health from fossil fuel combustion: impacts of air pollution and climate change. *Environmental health perspectives*, 125(2):141–148, 2017.

[4] Haleh Moghaddasi, Charles Culp, Jorge Vanegas, and Mehrdad Ehsani. Net zero energy buildings: Variations, clarifications, and requirements in response to the paris agreement. *Energies*, 14(13):3760, 2021.

[5] UNFCCC. Adoption of the Paris agreement: decision 1/CP/17. `http://unfccc.int/resource/docs/2015/cop21/eng/l09r01.pdf`, December 2015.

[6] Heleen L van Soest, Michel GJ den Elzen, and Detlef P van Vuuren. Net-zero emission targets for major emitting countries consistent with the paris agreement. *Nature communications*, 12(1):1–9, 2021.

[7] IEA. Renewables 2019. `www.iea.org/reports/renewables-2019`, 2019.

[8] Peter C Slorach and Laurence Stamford. Net zero in the heating sector: Technological options and environmental sustainability from now to 2050. *Energy Conversion and Management*, 230:113838, 2021.

[9] IEA. Heating. `www.iea.org/reports/heating`, 2020.

[10] Climate Change Act. Climate Change Act 2008 (2050 Target Amendment) Order 2019 (SI 2019/1056). `http://www.legislation.gov.uk/uksi/2019/1056/made`, 2019.

[11] BEIS. Digest of UK Energy Statistics (DUKES) 2019: Department for Business, Energy & Industrial Strategy (BEIS). `www.gov.uk/government/statistics/digest-of-uk-energy-statistics-dukes-2019`, 2019.

[12] Richard Lowes and Bridget Woodman. Disruptive and uncertain: Policy makers' perceptions on uk heat decarbonisation. *Energy policy*, 142:111494, 2020.

[13] Vicky Cheng and Koen Steemers. Modelling domestic energy consumption at district scale: A tool to support national and local energy policies. *Environmental Modelling & Software*, 26(10):1186–1198, 2011.

[14] Oliver Broad, Graeme Hawker, and Paul E Dodds. Decarbonising the uk residential sector: The dependence of national abatement on flexible and local views of the future. *Energy Policy*, 140:111321, 2020.

[15] BEIS. Energy consumption in the uk (ecuk), end use tables. `https://www.gov.uk/government/statistics/energy-consumption-in-the-uk.`, 2019.

[16] Haodong Lin, Julie Clavreul, Camille Jeandaux, Jenny Crawley, and Isabela Butnar. Environmental life cycle assessment of heating systems in the uk: Comparative assessment of hybrid heat pumps vs. condensing gas boilers. *Energy and Buildings*, 240:110865, 2021.

[17] Nick Eyre and Pranab Baruah. Uncertainties in future energy demand in uk residential heating. *Energy Policy*, 87:641–653, 2015.

[18] The Scottish Government. The future of energy in scotland: Scottish energy strategy. `https://www.gov.scot/publications/scottish-energy-strategy-future-energy-scotland-9781788515276/documents/`, December 2017.

[19] Tina Fawcett, Nick Eyre, and Russell Layberry. Heat pumps and global residential heating. *Proceedings, ECEEE summer study, Paper*, pages 6–321, 2015.

[20] Martin Geidl. *Integrated modeling and optimization of multi-carrier energy systems*. PhD thesis, ETH Zurich, 2007.

[21] Thomas Nuytten, Bert Claessens, Kristof Paredis, Johan Van Bael, and Daan Six. Flexibility of a combined heat and power system with thermal energy storage for district heating. *Applied energy*, 104:583–591, 2013.

[22] Johan Kensby, Anders Trüschel, and Jan-Olof Dalenbäck. Potential of residential buildings as thermal energy storage in district heating systems–Results from a pilot test. *Applied Energy*, 137:773–781, 2015.

[23] Jinfu Zheng, Zhigang Zhou, Jianing Zhao, and Jinda Wang. Integrated heat and power dispatch truly utilizing thermal inertia of district heating network for wind power integration. *Applied energy*, 211:865–874, 2018.

[24] Zhaoguang Pan, Qinglai Guo, and Hongbin Sun. Feasible region method based integrated heat and electricity dispatch considering building thermal inertia. *Applied Energy*, 192:395–407, 2017.

[25] Florian Kienzle, Patrick Favre-Perrod, Michèle Arnold, and Göran Andersson. Multi-energy delivery infrastructures for the future. In *2008 First international conference on infrastructure systems and services: building networks for a brighter future (INFRA)*, pages 1–5. IEEE, 2008.

[26] Thilo Krause, Göran Andersson, Klaus Fröhlich, and Alfredo Vaccaro. Multiple-energy carriers: modeling of production, delivery, and consumption. *Proceedings of the IEEE*, 99(1):15–27, 2010.

[27] Mark OMalley and Ben Kroposki. Energy comes together: the integration of all systems [guest editorial]. *IEEE Power and Energy Magazine*, 11(5):18–23, 2013.

[28] Pierluigi Mancarella. MES (multi-energy systems): An overview of concepts and evaluation models. *Energy*, 65:1–17, 2014.

[29] Reg Platt and Andrew Pendleton. Green streets, strong communities. *Institute for Public Policy Research*, 2011.

[30] Yeonsook Heo, Ruchi Choudhary, and GA Augenbroe. Calibration of building energy models for retrofit analysis under uncertainty. *Energy and Buildings*, 47:550–560, 2012.

[31] PF Lyons, NS Wade, T Jiang, PC Taylor, F Hashiesh, M Michel, and D Miller. Design and analysis of electrical energy storage demonstration projects on UK distribution networks. *Applied Energy*, 137:677–691, 2015.

[32] Mattia De Rosa, Vincenzo Bianco, Federico Scarpa, and Luca A Tagliafico. Heating and cooling building energy demand evaluation; a simplified model and a modified degree days approach. *Applied energy*, 128:217–229, 2014.

[33] Jihoon Jang, Jumi Baek, and Seung-Bok Leigh. Prediction of optimum heating timing based on artificial neural network by utilizing BEMS data. *Journal of Building Engineering*, 22:66–74, 2019.

[34] Jonathan Reynolds, Muhammad Waseem Ahmad, Yacine Rezgui, and Jean-Laurent Hippolyte. Operational supply and demand optimisation of a multi-vector district energy

system using artificial neural networks and a genetic algorithm. *Applied energy*, 235:699–713, 2019.

[35] Henrik Lund, Anders N Andersen, Poul Alberg Østergaard, Brian Vad Mathiesen, and David Connolly. From electricity smart grids to smart energy systems–a market operation based approach and understanding. *Energy*, 42(1):96–102, 2012.

[36] Fan Wu, Qinglai Guo, Hongbin Sun, and Zhaoguang Pan. Research on the optimization of combined heat and power microgrids with renewable energy. In *2014 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, pages 1–5. IEEE, 2014.

[37] H Tarik Ozkahraman and Ali Bolatturk. The use of tuff stone cladding in buildings for energy conservation. *Construction and Building Materials*, 20(7):435–440, 2006.

[38] Y Cøengel. Heat and mass transfer: a practical approach, 2007.

[39] Cheng Fan, Jiayuan Wang, Wenjie Gang, and Shenghan Li. Assessment of deep recurrent neural network-based strategies for short-term building energy predictions. *Applied Energy*, 236:700–710, 2019.

[40] Farshad Kheiri. A review on optimization methods applied in energy-efficient building geometry and envelope design. *Renewable and Sustainable Energy Reviews*, 92:897–920, 2018.

[41] Nelson Fumo. A review on the basics of building energy estimation. *Renewable and Sustainable Energy Reviews*, 31:53–60, 2014.

[42] Peder Bacher and Henrik Madsen. Identifying suitable models for the heat dynamics of buildings. *Energy and Buildings*, 43(7):1511–1522, 2011.

[43] Yeonsook Heo, Godfried Augenbroe, Diane Graziano, Ralph T Muehleisen, and Leah Guzowski. Scalable methodology for large scale building energy improvement: Relevance of calibration in model-based retrofit analysis. *Building and Environment*, 87:342–350, 2015.

[44] ASHRAE Handbook-Fundamentals. Energy estimating and modeling methods. 2009.

[45] NC Balaji, Monto Mani, and BV Venkatarama Reddy. Dynamic thermal performance of conventional and alternative building wall envelopes. *Journal of building engineering*, 21:373–395, 2019.

[46] T Agami Reddy. Literature review on calibration of building energy simulation programs: uses, problems, procedures, uncertainty, and tools. *ASHRAE transactions*, 112:226, 2006.

[47] Mohammad Royapoor and Tony Roskilly. Building model calibration using energy and environmental data. *Energy and Buildings*, 94:109–120, 2015.

[48] Paul Raftery, Marcus Keane, and James O'Donnell. Calibrating whole building energy models: An evidence-based methodology. *Energy and Buildings*, 43(9):2356–2364, 2011.

[49] Paul Raftery, Marcus Keane, and Andrea Costa. Calibrating whole building energy models: Detailed case study using hourly measured data. *Energy and buildings*, 43(12):3666–3679, 2011.

[50] Veronica I Soebarto. Calibration of hourly energy simulations using hourly monitored data and monthly utility records for two case study buildings. In *Proceedings of Building Simulation*, volume 97, pages 411–419, 1997.

[51] Stéphane Bertagnolio. *Evidence-based model calibration for efficient building energy services*. PhD thesis, Université de Liége, Liége, Belgium, 2012.

[52] Zheng Yang and Burcin Becerik-Gerber. A model calibration framework for simultaneous multi-level building energy simulation. *Applied Energy*, 149:415–431, 2015.

[53] Germán Ramos Ruiz, Carlos Fernández Bandera, Tomás Gómez-Acebo Temes, and Ana Sánchez-Ostiz Gutierrez. Genetic algorithm for building envelope calibration. *Applied energy*, 168:691–705, 2016.

[54] Christian Struck, Jan Hensen, and Petr Kotek. On the application of uncertainty and sensitivity analysis with abstract building performance simulation tools. *Journal of building physics*, 33(1):5–27, 2009.

[55] Bryan Eisenhower, Zheng O'Neill, Vladimir A Fonoberov, and Igor Mezić. Uncertainty and sensitivity decomposition of building energy models. *Journal of Building Performance Simulation*, 5(3):171–184, 2012.

[56] A Tindale. Third-order lumped-parameter simulation method. *Building Services Engineering Research and Technology*, 14(3):87–97, 1993.

[57] CP Underwood. An improved lumped parameter method for building thermal modelling. *Energy and Buildings*, 79:191–201, 2014.

[58] Mehmet Aksoezen, Magdalena Daniel, Uta Hassler, and Niklaus Kohler. Building age as an indicator for energy consumption. *Energy and Buildings*, 87:74–86, 2015.

[59] Rosanne Walker and Sara Pavía. Thermal performance of a selection of insulation materials suitable for historic buildings. *Building and environment*, 94:155–165, 2015.

[60] Giorgio Mustafaraj, Dashamir Marini, Andrea Costa, and Marcus Keane. Model calibration for building energy efficiency simulation. *Applied Energy*, 130:72–85, 2014.

[61] S Yang, TJ Pilet, and JC Ordonez. Volume element model for 3D dynamic building thermal modeling and simulation. *Energy*, 148:642–661, 2018.

[62] Jun Yuan, Victor Nian, and Bin Su. A meta model based Bayesian approach for building energy models calibration. *Energy Procedia*, 143:161–166, 2017.

[63] Jun Yuan, Victor Nian, Bin Su, and Qun Meng. A simultaneous calibration and parameter ranking method for building energy models. *Applied energy*, 206:657–666, 2017.

[64] Hai-xiang Zhao and Frédéric Magoulès. A review on the prediction of building energy consumption. *Renewable and Sustainable Energy Reviews*, 16(6):3586–3592, 2012.

[65] AS Ahmad, MY Hassan, MP Abdullah, HA Rahman, F Hussin, H Abdullah, and R Saidur. A review on applications of ANN and SVM for building electrical energy consumption forecasting. *Renewable and Sustainable Energy Reviews*, 33:102–109, 2014.

[66] Young Tae Chae, Raya Horesh, Youngdeok Hwang, and Young M Lee. Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings. *Energy and Buildings*, 111:184–194, 2016.

[67] Yixuan Wei, Liang Xia, Song Pan, Jinshun Wu, Xingxing Zhang, Mengjie Han, Weiya Zhang, Jingchao Xie, and Qingping Li. Prediction of occupancy level and energy consumption in office building using blind system identification and neural networks. *Applied Energy*, 240:276–294, 2019.

[68] Aurélie Foucquier, Sylvain Robert, Frédéric Suard, Louis Stéphan, and Arnaud Jay. State of the art in building modelling and energy performances prediction: A review. *Renewable and Sustainable Energy Reviews*, 23:272–288, 2013.

[69] David Connolly, Henrik Lund, Brian Vad Mathiesen, and Martin Leahy. A review of computer tools for analysing the integration of renewable energy into various energy systems. *Applied energy*, 87(4):1059–1082, 2010.

[70] Stephanie Hay and Anna Ferguson. A review of power system modelling platforms and capabilities. *IET Special Interest Publication for the Council for Science and Technology*, pages 3–13, 2015.

[71] Betul Bektas Ekici and U Teoman Aksoy. Prediction of building energy consumption by using artificial neural networks. *Advances in Engineering Software*, 40(5):356–362, 2009.

[72] F Martellotta, U Ayr, P Stefanizzi, A Sacchetti, and G Riganti. On the use of artificial neural networks to model household energy consumptions. *Energy Procedia*, 126:250–257, 2017.

[73] Andrew Hunter, Lee Kennedy, Jenny Henry, and Ian Ferguson. Application of neural networks and sensitivity analysis to improved prediction of trauma survival. *Computer methods and programs in biomedicine*, 62(1):11–19, 2000.

[74] Simon Haykin. *Neural networks: a comprehensive foundation*. Prentice-Hall, Inc., 2007.

[75] Kadir Amasyali and Nora M El-Gohary. A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews*, 81:1192–1205, 2018.

[76] Hao Bai, Chen Liu, Elena Breaz, and Fei Gao. Artificial neural network aided real-time simulation of electric traction system. *Energy and AI*, page 100010, 2020.

[77] MA Rafe Biswas, Melvin D Robinson, and Nelson Fumo. Prediction of residential building energy consumption: A neural network approach. *Energy*, 117:84–92, 2016.

[78] Ali HA Al-Waeli, K Sopian, Jabar H Yousif, Hussein A Kazem, John Boland, and Miqdam T Chaichan. Artificial neural network modeling and analysis of photovoltaic/thermal system based on the experimental study. *Energy Conversion and Management*, 186:368–379, 2019.

[79] O Pastor-Bárcenas, Emilio Soria-Olivas, José D Martín-Guerrero, Gustavo Camps-Valls, José L Carrasco-Rodríguez, and Secundino del Valle-Tascón. Unbiased sensitivity analysis and pruning techniques in neural networks for surface ozone modelling. *Ecological Modelling*, 182(2):149–158, 2005.

[80] Younes Noorollahi, Mohammad Ali Jokar, and Ahmad Kalhor. Using artificial neural networks for temporal and spatial wind speed forecasting in Iran. *Energy Conversion and Management*, 115:17–25, 2016.

[81] Emanuel Federico Alsina, Marco Bortolini, Mauro Gamberi, and Alberto Regattieri. Artificial neural network optimisation for monthly average daily global solar radiation prediction. *Energy conversion and management*, 120:320–329, 2016.

[82] Chirag Deb, Lee Siew Eang, Junjing Yang, and Mattheos Santamouris. Forecasting diurnal cooling energy load for institutional buildings using artificial neural networks. *Energy and Buildings*, 121:284–297, 2016.

[83] Sonia Leva, Alberto Dolara, Francesco Grimaccia, Marco Mussetta, and Emanuele Ogliari. Analysis and validation of 24 hours ahead neural network forecasting of photovoltaic output power. *Mathematics and computers in simulation*, 131:88–100, 2017.

[84] Alberto Dolara, Francesco Grimaccia, Sonia Leva, Marco Mussetta, and Emanuele Ogliari. A physical hybrid artificial neural network for short term forecasting of PV plant power output. *Energies*, 8(2):1138–1153, 2015.

[85] Muriel Gevrey, Ioannis Dimopoulos, and Sovan Lek. Two-way interaction of input variables in the sensitivity analysis of neural network models. *Ecological modelling*, 195(1-2):43–50, 2006.

[86] AH Sung. Ranking importance of input parameters of neural networks. *Expert systems with Applications*, 15(3-4):405–411, 1998.

[87] Jacek M Zurada, Aleksander Malinowski, and Ian Cloete. Sensitivity analysis for minimization of input data dimension for feedforward neural network. In *Proceedings of IEEE International Symposium on Circuits and Systems-ISCAS'94*, volume 6, pages 447–450. IEEE, 1994.

[88] L Srivastava, SN Singh, and J Sharma. Comparison of feature selection techniques for ANN-based voltage estimation. *Electric Power Systems Research*, 53(3):187–195, 2000.

[89] Sven F Crone and Nikolaos Kourentzes. Feature selection for time series prediction–a combined filter and wrapper approach for neural networks. *Neurocomputing*, 73(10-12):1923–1936, 2010.

[90] Zihao Li, Daniel Friedrich, and Gareth P Harrison. Demand forecasting for a mixed-use building using agent-schedule information with a data-driven model. *Energies*, 13(4):780, 2020.

[91] Sally M Blower and Hadi Dowlatabadi. Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example. *International Statistical Review/Revue Internationale de Statistique*, pages 229–243, 1994.

[92] Anthony TC Goh. Back-propagation neural networks for modeling complex systems. *Artificial intelligence in engineering*, 9(3):143–151, 1995.

[93] Sovan Lek, Marc Delacoste, Philippe Baran, Ioannis Dimopoulos, Jacques Lauga, and Stéphane Aulagnier. Application of neural networks to modelling nonlinear relationships in ecology. *Ecological modelling*, 90(1):39–52, 1996.

[94] Julian D Olden and Donald A Jackson. Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks. *Ecological modelling*, 154(1-2):135–150, 2002.

[95] Michele Scardi. Artificial neural networks as empirical models for estimating phytoplankton production. *Marine Ecology Progress Series*, 139:289–299, 1996.

[96] Yannis Dimopoulos, Paul Bourret, and Sovan Lek. Use of some sensitivity criteria for choosing networks with good generalization ability. *Neural Processing Letters*, 2(6):1–4, 1995.

[97] B Henderson-Sellers and A Henderson-Sellers. Sensitivity evaluation of environmental models using fractional factorial experimentation. *Ecological Modelling*, 86(2-3):291–295, 1996.

[98] Diane L Beres and Douglas M Hawkins. Plackett–Burman technique for sensitivity analysis of many-parametered models. *Ecological modelling*, 141(1-3):171–183, 2001.

[99] Si Chen, Daniel Friedrich, Zhibin Yu, and James Yu. District heating network demand prediction using a physics-based energy model with a Bayesian approach for parameter calibration. *Energies*, 12(18):3408, 2019.

[100] Si Chen, Yaxing Ren, Daniel Friedrich, Zhibin Yu, and James Yu. Sensitivity analysis to reduce duplicated features in ANN training for district heat demand prediction. *Energy and AI*, 2:100028, 2020.

[101] Linda Pedersen, Jacob Stang, and Rolf Ulseth. Load prediction method for heat and electricity demand in buildings for the purpose of planning for mixed energy distribution systems. *Energy and Buildings*, 40(7):1124–1134, 2008.

[102] Matteo De Felice, Andrea Alessandri, and Paolo M Ruti. Electricity demand forecasting over Italy: Potential benefits using numerical weather prediction models. *Electric Power Systems Research*, 104:71–79, 2013.

[103] Yan Ding, Qiaochu Wang, Zhaoxia Wang, Shuxue Han, and Neng Zhu. An occupancy-based model for building electricity consumption prediction: A case study of three campus buildings in Tianjin. *Energy and Buildings*, 202:109412, 2019.

[104] Urs Wilke, Frédéric Haldi, Jean-Louis Scartezzini, and Darren Robinson. A bottom-up stochastic model to predict building occupants' time-dependent activities. *Building and Environment*, 60:254–264, 2013.

[105] Guy R Newsham and Benjamin J Birt. Building-level occupancy data to improve ARIMA-based electricity use forecasts. In *Proceedings of the 2nd ACM workshop on embedded sensing systems for energy-efficiency in building*, pages 13–18, 2010.

[106] XJ Luo, Lukumon O Oyedele, Anuoluwapo O Ajayi, Olugbenga O Akinade, Juan Manuel Davila Delgado, Hakeem A Owolabi, and Ashraf Ahmed. Genetic algorithm-determined deep feedforward neural network architecture for predicting electricity consumption in real buildings. *Energy and AI*, 2:100015, 2020.

[107] Seokho Kim, Yujin Song, Yoondong Sung, and Donghyun Seo. Development of a consecutive occupancy estimation framework for improving the energy demand prediction performance of building energy modeling tools. *Energies*, 12(3):433, 2019.

[108] Jihui Yuan, Craig Farnham, Chikako Azuma, and Kazuo Emura. Predictive artificial neural network models to forecast the seasonal hourly electricity consumption for a university campus. *Sustainable Cities and Society*, 42:82–92, 2018.

[109] Huiling Cai, Shoupeng Shen, Qingcheng Lin, Xuefeng Li, and Hui Xiao. Predicting the energy consumption of residential buildings for regional electricity supply-side and demand-side management. *IEEE Access*, 7:30386–30397, 2019.

[110] Aaron Zeng, Sheng Liu, and Yao Yu. Comparative study of data driven methods in building electricity use prediction. *Energy and Buildings*, 194:289–300, 2019.

[111] Joaquim Massana, Carles Pous, Llorenç Burgas, Joaquim Melendez, and Joan Colomer. Short-term load forecasting for non-residential buildings contrasting artificial occupancy attributes. *Energy and Buildings*, 130:519–531, 2016.

[112] SSAK Javeed Nizami and Ahmed Z Al-Garni. Forecasting electric energy consumption using neural networks. *Energy policy*, 23(12):1097–1104, 1995.

[113] Nikolaos G Paterakis, Elena Mocanu, Madeleine Gibescu, Bart Stappers, and Walter van Alst. Deep learning versus traditional machine learning methods for aggregated energy demand prediction. In *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, pages 1–6. IEEE, 2017.

[114] Jawad Ahmad, Hadi Larijani, Rohinton Emmanuel, Mike Mannion, Abbas Javed, and Mark Phillipson. Energy demand prediction through novel random neural network predictor for large non-domestic buildings. In *2017 Annual IEEE International Systems Conference (SysCon)*, pages 1–6. IEEE, 2017.

[115] Aowabin Rahman, Vivek Srikumar, and Amanda D Smith. Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks. *Applied energy*, 212:372–385, 2018.

[116] Younghoon Kwak, Donghyun Seo, Cheolyong Jang, and Jung-Ho Huh. Feasibility study on a novel methodology for short-term real-time energy demand prediction using weather forecasting data. *Energy and Buildings*, 57:250–260, 2013.

[117] Aaron Zeng, Hodde Ho, and Yao Yu. Prediction of building electricity usage using Gaussian process regression. *Journal of Building Engineering*, 28:101054, 2020.

[118] Orhan Ekren and Banu Yetkin Ekren. Size optimization of a PV/wind hybrid energy conversion system with battery storage using response surface methodology. *Applied energy*, 85(11):1086–1101, 2008.

[119] Ricardo Luna-Rubio, Mario Trejo-Perea, D Vargas-Vázquez, and GJ Ríos-Moreno. Optimal sizing of renewable hybrids energy systems: A review of methodologies. *Solar energy*, 86(4):1077–1088, 2012.

[120] Ali M Eltamaly, Mohamed A Mohamed, and Abdulrahman I Alolah. A novel smart grid theory for optimal sizing of hybrid renewable energy systems. *Solar Energy*, 124:26–38, 2016.

[121] Kasereka Ndwali, Jackson G Njiri, and Evan M Wanjiru. Multi-objective optimal sizing of grid connected photovoltaic batteryless system minimizing the total life cycle cost and the grid energy. *Renewable Energy*, 148:1256–1265, 2020.

[122] Akbar Maleki, Morteza Gholipour Khajeh, and Mehran Ameri. Optimal sizing of a grid independent hybrid renewable energy system incorporating resource uncertainty, and load uncertainty. *International Journal of Electrical Power & Energy Systems*, 83:514–524, 2016.

[123] EU An. Strategy on heating and cooling. *European Commission COM*, 51, 2016.

[124] Lígia da Silva Lima, Mattijs Quartier, Astrid Buchmayr, David Sanjuan-Delmás, Hannes Laget, Dominique Corbisier, Jan Mertens, and Jo Dewulf. Life cycle assessment of lithium-ion batteries and vanadium redox flow batteries-based renewable energy storage systems. *Sustainable Energy Technologies and Assessments*, 46:101286, 2021.

[125] Seungwon An, Qing Li, and Thomas W Gedra. Natural gas and electricity optimal power flow. In *2003 IEEE PES Transmission and Distribution Conference and Exposition (IEEE Cat. No. 03CH37495)*, volume 1, pages 138–143. IEEE, 2003.

[126] Martin Geidl and Göran Andersson. Optimal power flow of multiple energy carriers. *IEEE Transactions on power systems*, 22(1):145–155, 2007.

[127] Clodomiro Unsihuay, JW Marangon Lima, and AC Zambroni De Souza. Modeling the integrated natural gas and electricity optimal power flow. In *2007 IEEE Power Engineering Society General Meeting*, pages 1–7. IEEE, 2007.

[128] Clodomiro Unsihuay-Vila, Jose Wanderley Marangon-Lima, AC Zambroni De Souza, Ignacio J Perez-Arriaga, and Pedro P Balestrassi. A model to long-term, multiarea, multistage, and integrated expansion planning of electricity and natural gas systems. *IEEE Transactions on Power Systems*, 25(2):1154–1168, 2010.

[129] Alberto Martinez-Mares and Claudio R Fuerte-Esquivel. A unified gas and power flow analysis in natural gas and electricity coupled networks. *IEEE Transactions on Power Systems*, 27(4):2156–2166, 2012.

[130] Getnet Tadesse Ayele, Pierrick Haurant, Björn Laumert, and Bruno Lacarrière. An extended energy hub approach for load flow analysis of highly coupled district energy networks: Illustration with electricity and heating. *Applied energy*, 212:850–867, 2018.

[131] Marc Rees, Jianzhong Wu, and Bishoy Awad. Steady state flow analysis for integrated urban heat and power distribution networks. In *2009 44th International Universities Power Engineering Conference (UPEC)*, pages 1–5. IEEE, 2009.

[132] Xuezhi Liu, Jianzhong Wu, Nick Jenkins, and Audrius Bagdanavicius. Combined analysis of electricity and heat networks. *Applied Energy*, 162:1238–1250, 2016.

[133] Xuezhi Liu and Pierluigi Mancarella. Modelling, assessment and Sankey diagrams of integrated electricity-heat-gas networks in multi-vector district energy systems. *Applied Energy*, 167:336–352, 2016.

[134] Xinyu Chen, Chongqing Kang, Mark O'Malley, Qing Xia, Jianhua Bai, Chun Liu, Rongfu Sun, Weizhou Wang, and Hui Li. Increasing the flexibility of combined heat and power for wind power integration in China: Modeling and implications. *IEEE Transactions on Power Systems*, 30(4):1848–1857, 2014.

[135] Xinyu Chen, Michael B McElroy, and Chongqing Kang. Integrated energy systems for higher wind penetration in China: Formulation, implementation, and impacts. *IEEE Transactions on Power Systems*, 33(2):1309–1319, 2017.

[136] Zhigang Li, Wenchuan Wu, Mohammad Shahidehpour, Jianhui Wang, and Boming Zhang. Combined heat and power dispatch considering pipeline energy storage of district heating network. *IEEE Transactions on Sustainable Energy*, 7(1):12–22, 2015.

[137] Zhigang Li, Wenchuan Wu, Jianhui Wang, Boming Zhang, and Taiyi Zheng. Transmission-constrained unit commitment considering combined electricity and district heating networks. *IEEE Transactions on Sustainable Energy*, 7(2):480–492, 2015.

[138] Youcai Liang, Mohammed Al-Tameemi, and Zhibin Yu. Investigation of a gas-fuelled water heater based on combined power and heat pump cycles. *Applied energy*, 212:1476–1488, 2018.

[139] Victor Tulus, Dieter Boer, Luisa F Cabeza, Laureano Jiménez, and Gonzalo Guillén-Gosálbez. Enhanced thermal energy supply via central solar heating plants with seasonal storage: A multi-objective optimization approach. *Applied Energy*, 181:549–561, 2016.

[140] Abdulrahman Dahash, Fabian Ochs, Michele Bianchi Janetti, and Wolfgang Streicher. Advances in seasonal thermal energy storage for solar district heating applications: A critical review on large-scale hot-water tank and pit thermal energy storage systems. *Applied Energy*, 239:296–315, 2019.

[141] Renaldi Renaldi, Arisitides Kiprakis, and Daniel Friedrich. An optimisation framework for thermal energy storage integration in a residential heat pump heating system. *Applied energy*, 186:520–529, 2017.

[142] Energy Department for Business and Industrial Strategy. 2017 UK greenhouse gas emissions, final figures. 2019.

[143] Michael-Allan Millar, Zhibin Yu, Neil Burnside, Greg Jones, and Bruce Elrick. Identification of key performance indicators and complimentary load profiles for 5th generation district energy networks. *Applied Energy*, 291:116672, 2021.

[144] Ottmar Edenhofer, Ramón Pichs-Madruga, Youba Sokona, Kristin Seyboth, Susanne Kadner, Timm Zwickel, Patrick Eickemeier, Gerrit Hansen, Steffen Schlömer, Christoph von Stechow, et al. *Renewable energy sources and climate change mitigation: Special report of the intergovernmental panel on climate change*. Cambridge University Press, 2011.

[145] Mohammed Ridha Jawad Al-Tameemi, Youcai Liang, and Zhibin Yu. Design strategies and control methods for a thermally driven heat pump system based on combined cycles. *Frontiers in Energy Research*, 7:131, 2019.

[146] YW Kim, J Ramousse, G Fraisse, P Dalicieux, and P Baranek. Optimal sizing of a thermoelectric heat pump (THP) for heating energy-efficient buildings. *Energy and buildings*, 70:106–116, 2014.

[147] Yunus A Cengel and Michael A Boles. *Thermodynamics: An Engineering Approach 6th Editon (SI Units)*. The McGraw-Hill Companies, Inc., New York, 2007.

[148] Valentin Trillat-Berdal, Bernard Souyri, and Gilbert Achard. Coupling of geothermal heat pumps with thermal solar collectors. *Applied Thermal Engineering*, 27(10):1750–1755, 2007.

[149] Onder Ozgener and Arif Hepbasli. Modeling and performance evaluation of ground source (geothermal) heat pump systems. *Energy and Buildings*, 39(1):66–75, 2007.

[150] Stuart J Self, Bale V Reddy, and Marc A Rosen. Geothermal heat pump systems: Status review and comparison with other heating options. *Applied energy*, 101:341–348, 2013.

[151] Alessandro Casasso and Rajandrea Sethi. Efficiency of closed loop geothermal heat pumps: A sensitivity analysis. *Renewable Energy*, 62:737–746, 2014.

[152] Walter Grassi. *Heat pumps: fundamentals and applications*. Springer, 2017.

[153] Halit Arat and Oguz Arslan. Optimization of district heating system aided by geothermal heat pump: A novel multistage with multilevel ANN modelling. *Applied Thermal Engineering*, 111:608–623, 2017.

[154] Hongyang Li, Sham Rane, and Zhibin Yu. Investigation of the performance and flow characteristics of two-phase reaction turbines in total flow geothermal systems. *Renewable Energy*, 175:345–372, 2021.

[155] Eric O'Shaughnessy, Dylan Cutler, Kristen Ardani, and Robert Margolis. Solar plus: Optimization of distributed solar PV through battery storage and dispatchable load in residential buildings. *Applied Energy*, 213:11–21, 2018.

[156] Derong Liu, Yancai Xu, Qinglai Wei, and Xinliang Liu. Residential energy scheduling for variable weather solar energy based on adaptive dynamic programming. *IEEE/CAA Journal of Automatica Sinica*, 5(1):36–46, 2017.

[157] Tomislav Dragičević, Hrvoje Pandžić, Davor Škrlec, Igor Kuzle, Josep M Guerrero, and Daniel S Kirschen. Capacity optimization of renewable energy sources and battery storage in an autonomous telecommunication facility. *IEEE Transactions on Sustainable Energy*, 5(4):1367–1378, 2014.

[158] Rahmat Khezri, Amin Mahmoudi, and Mohammed H Haque. Optimal capacity of solar PV and battery storage for Australian grid-connected households. *IEEE Transactions on Industry Applications*, 56(5):5319–5329, 2020.

[159] Taskin Jamal, Craig Carter, Thomas Schmidt, GM Shafiullah, Martina Calais, and Tania Urmee. An energy flow simulation tool for incorporating short-term PV forecasting in a diesel-PV-battery off-grid power supply system. *Applied Energy*, 254:113718, 2019.

[160] Yuqing Yang, Stephen Bremner, Chris Menictas, and Merlinde Kay. Battery energy storage system size determination in renewable energy systems: A review. *Renewable and Sustainable Energy Reviews*, 91:109–125, 2018.

[161] Safat Bin Wali, MA Hannan, MS Reza, Pin Jern Ker, RA Begum, MS Abd Rahman, and M Mansor. Battery storage systems integrated renewable energy sources: A biblio metric analysis towards future directions. *Journal of Energy Storage*, 35:102296, 2021.

[162] Soheil Kavian, Cyrus Aghanajafi, Hassan Jafari Mosleh, Arash Nazari, and Ashkan Nazari. Exergy, economic and environmental evaluation of an optimized hybrid photovoltaic-geothermal heat pump system. *Applied Energy*, 276:115469, 2020.

[163] Bo Yang, Junting Wang, Yixuan Chen, Danyang Li, Chunyuan Zeng, Yijun Chen, Zhengxun Guo, Hongchun Shu, Xiaoshun Zhang, Tao Yu, et al. Optimal sizing and placement of energy storage system in power grids: a state-of-the-art one-stop handbook. *Journal of Energy Storage*, 32:101814, 2020.

[164] Bo Yang, Lei Yu, Yixuan Chen, Haoyin Ye, Ruining Shao, Hongchun Shu, Tao Yu, Xiaoshun Zhang, and Liming Sun. Modelling, applications, and evaluations of optimal sizing and placement of distributed generations: A critical state-of-the-art survey. *International Journal of Energy Research*, 45(3):3615–3642, 2021.

[165] Modassar Chaudry, Nick Jenkins, and Goran Strbac. Multi-time period combined gas and electricity network optimisation. *Electric power systems Research*, 78(7):1265–1279, 2008.

[166] Michele Arnold, Rudy R Negenborn, Goran Andersson, and Bart De Schutter. Model-based predictive control applied to multi-carrier energy systems. In *2009 IEEE Power & Energy Society General Meeting*, pages 1–8. IEEE, 2009.

[167] Bieshoy Awad, Modassar Chaudry, Jianzhong Wu, and Nick Jenkins. Integrated optimal power flow for electric power and heat in a microgrid. In *CIRED 2009-20th International Conference and Exhibition on Electricity Distribution-Part 1*, pages 1–4. IET, 2009.

[168] MT Rees, Jianzhong Wu, Nicholas Jenkins, and Muditha Abeysekera. Carbon constrained design of energy infrastructure for new build schemes. *Applied energy*, 113:1220–1234, 2014.

[169] N Awad, Jianzhong Wu, Janaka Ekanayake, and Nicholas Jenkins. Integrated operation of an energy MicroGrid with islanded electricity network. 2011.

[170] Xianjun Zhang, George G Karady, Kalyan R Piratla, and Samuel T Ariaratnam. Network capacity assessment of combined heat and power-based distributed generation in urban energy infrastructures. *IEEE Transactions on Smart Grid*, 4(4):2131–2138, 2013.

[171] Xianjun Zhang, George G Karady, and Samuel T Ariaratnam. Optimal allocation of CHP-based distributed generation on urban energy distribution networks. *IEEE Transactions on Sustainable Energy*, 5(1):246–253, 2013.

[172] JH Zheng, QH Wu, and ZX Jing. Coordinated scheduling strategy to optimize conflicting benefits for daily operation of integrated electricity and gas networks. *Applied energy*, 192:370–381, 2017.

[173] Yibo Jiang, Jian Xu, Yuanzhang Sun, Congying Wei, Jing Wang, Siyang Liao, Deping Ke, Xiong Li, Jun Yang, and Xiaotao Peng. Coordinated operation of gas-electricity integrated distribution system with multi-CCHP and distributed renewable energy sources. *Applied energy*, 211:237–248, 2018.

[174] Yulong Yang, Kai Wu, Xu Yan, Jianchao Gao, and Hongyu Long. The large-scale wind power integration using the integrated heating load and heating storage control. In *2015 IEEE Eindhoven PowerTech*, pages 1–6. IEEE, 2015.

[175] Wei Gu, Jun Wang, Shuai Lu, Zhao Luo, and Chenyu Wu. Optimal operation for integrated energy system considering thermal inertia of district heating network and buildings. *Applied energy*, 199:234–246, 2017.

[176] Tamara G Kolda, Robert Michael Lewis, and Virginia Torczon. Optimization by direct search: New perspectives on some classical and modern methods. *SIAM review*, 45(3):385–482, 2003.

[177] Emad Elbeltagi, Tarek Hegazy, and Donald Grierson. Comparison among five evolutionary-based optimization algorithms. *Advanced engineering informatics*, 19(1):43–53, 2005.

[178] NC Balaji, Monto Mani, and BV Reddy. Discerning heat transfer in building materials. *Energy Procedia*, 54(0):654–68, 2014.

[179] Kévin Beck, M Al-Mukhtar, Olivier Rozenbaum, and Michel Rautureau. Characterization, water transfer properties and deterioration in tuffeau: building material in the loire valley—France. *Building and environment*, 38(9-10):1151–1162, 2003.

[180] L Evangelisti, C Guattari, P Gori, and F Bianchi. Heat transfer study of external convective and radiative coefficients for building applications. *Energy and Buildings*, 151:429–438, 2017.

[181] Hao Gao, Christian Koch, and Yupeng Wu. Building information modelling based building energy modelling: A review. *Applied energy*, 238:320–343, 2019.

[182] Sheena Dickson. Gilmorehill and Hillhead campus map. `https://www.gov.uk/government/statistics/energy-consumption-in-the-uk.`

[183] Paul Baker. U-values and traditional buildings: in situ measurements and their comparisons to calculated values. *Historic Scotland Technical Paper*, 10:70, 2011.

[184] Vladislav Masatin, Eduard Latõšev, and Anna Volkova. Evaluation factor for district heating network heat loss with respect to network geometry. *Energy Procedia*, 95:279–285, 2016.

[185] Tania Ullah and William M Healy. The effect of piping heat losses on the efficiency of solar thermal water heating in a net-zero energy home. In *Proceedings of the 2015 ASHRAE Winter Conference, Chicago, IL, USA*, pages 24–28, 2015.

[186] Ding Yan, Tian Zhe, Wu Yong, and Zhu Neng. Achievements and suggestions of heat metering and energy efficiency retrofit for existing residential buildings in northern heating regions of China. *Energy Policy*, 39(9):4675–4682, 2011.

[187] Mattias Vesterlund, Johan Sandberg, Bo Lindblom, and Jan Dahl. Evaluation of losses in district heating system, a case study. In *International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems: 16/07/2013-19/07/2013*, 2013.

[188] Maria J Bayarri, James O Berger, Rui Paulo, Jerry Sacks, John A Cafeo, James Cavendish, Chin-Hsu Lin, and Jian Tu. A framework for validation of computer models. *Technometrics*, 49(2):138–154, 2007.

[189] Marc C Kennedy and Anthony O'Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464, 2001.

[190] Christian Robert and George Casella. *Monte Carlo statistical methods*. Springer Science & Business Media, 2013.

[191] Ioannis Andrianakis, Ian R Vernon, Nicky McCreesh, Trevelyan J McKinley, Jeremy E Oakley, Rebecca N Nsubuga, Michael Goldstein, and Richard G White. Bayesian history matching of complex infectious disease models using emulation: a tutorial and a case study on HIV in Uganda. *PLoS Comput Biol*, 11(1):e1003968, 2015.

[192] Carl Edward Rasmussen. Gaussian processes in machine learning. In *Summer School on Machine Learning*, pages 63–71. Springer, 2003.

[193] Francesca Campolongo, Jessica Cariboni, and Andrea Saltelli. An effective screening design for sensitivity analysis of large models. *Environmental modelling & software*, 22(10):1509–1518, 2007.

[194] Rajesh Kumar, RK Aggarwal, and JD Sharma. Energy analysis of a building using artificial neural network: A review. *Energy and Buildings*, 65:352–358, 2013.

[195] Frédéric Magoulès and Hai-Xiang Zhao. *Data mining and machine learning in building energy analysis*. Wiley Online Library, 2016.

[196] Miroslav Kubat. *An introduction to machine learning*. Springer, 2017.

[197] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. *nature*, 323(6088):533–536, 1986.

[198] Nii O Attoh-Okine. Analysis of learning rate and momentum term in backpropagation neural network algorithm trained to predict pavement performance. *Advances in Engineering Software*, 30(4):291–302, 1999.

[199] Zhaoyang Ye and Moon Keun Kim. Predicting electricity consumption in a building using an optimized back-propagation and Levenberg–Marquardt back-propagation neural network: Case study of a shopping mall in China. *Sustainable Cities and Society*, 42:176–183, 2018.

[200] Si Chen, Yaxing Ren, Daniel Friedrich, Zhibin Yu, and James Yu. District heat demand prediction using artificial neural network with data of sample building. In *International Conference on Applied Energy (ICAE 2019). August 12-15, Västerås, Sweden*, 2019.

[201] Huidong Gu, Guowen Liu, Jian Wang, Anne-Francoise Aubry, and Mark E Arnold. Selecting the correct weighting factors for linear and quadratic calibration curves with least-squares regression algorithm in bioanalytical LC-MS/MS assays and impacts of using incorrect weighting factors on curve stability, data quality, and assay performance. *Analytical Chemistry*, 86(18):8959–8966, 2014.

[202] Hideo Tanaka and Haekwan Lee. Interval regression analysis by quadratic programming approach. *IEEE Transactions on Fuzzy Systems*, 6(4):473–481, 1998.

[203] G Baffi, EB Martin, and AJ Morris. Non-linear projection to latent structures revisited: the quadratic PLS algorithm. *Computers & Chemical Engineering*, 23(3):395–411, 1999.

[204] Giuseppe Baffi, EB Martin, and AJ Morris. Non-linear projection to latent structures revisited (the neural network PLS algorithm). *Computers & Chemical Engineering*, 23(9):1293–1307, 1999.

[205] Jin Zhao and Bimal K Bose. Evaluation of membership functions for fuzzy logic controlled induction motor drive. In *IEEE 2002 28th Annual Conference of the Industrial Electronics Society. IECON 02*, volume 1, pages 229–234. IEEE, 2002.

[206] MC Leung, CF Norman, Loi Lei Lai, and Tin Tai Chow. The use of occupancy space electrical power demand in building cooling load prediction. *Energy and Buildings*, 55:151–163, 2012.

[207] Jatin Bedi and Durga Toshniwal. Deep learning framework to forecast electricity demand. *Applied energy*, 238:1312–1326, 2019.

[208] Si Chen, Yaxing Ren, Daniel Friedrich, Zhibin Yu, and James Yu. Prediction of office building electricity demand using artificial neural network by splitting the time horizon for different occupancy rates. *Energy and AI*, page 100093, 2021.

[209] Zhengru Ren, Amrit Shankar Verma, Ye Li, Julie JE Teuwen, and Zhiyu Jiang. Offshore wind turbine operations and maintenance: A state-of-the-art review. *Renewable and Sustainable Energy Reviews*, 144:110886, 2021.

[210] Sandra Eriksson, Hans Bernhoff, and Mats Leijon. Evaluation of different turbine concepts for wind power. *renewable and sustainable energy reviews*, 12(5):1419–1434, 2008.

[211] Muhammad Mahmood Aslam Bhutta, Nasir Hayat, Ahmed Uzair Farooq, Zain Ali, Sh Rehan Jamil, and Zahid Hussain. Vertical axis wind turbine–A review of various configurations and design techniques. *Renewable and Sustainable Energy Reviews*, 16(4):1926–1939, 2012.

[212] M Johari, Muhd Jalil, and Mohammad Faizal Mohd Shariff. Comparison of horizontal axis wind turbine (HAWT) and vertical axis wind turbine (VAWT). *International Journal of Engineering and Technology*, 7(4.13):74–80, 2018.

[213] Bala Govind. Increasing the operational capability of a horizontal axis wind turbine by its integration with a vertical axis wind turbine. *Applied Energy*, 199:479–494, 2017.

[214] Yaxing Ren, Liuying Li, Joseph Brindley, and Lin Jiang. Nonlinear PI control for variable pitch wind turbine. *Control Engineering Practice*, 50:84–94, 2016.

[215] Paras Mandal, Surya Teja Swarroop Madhira, Julian Meng, Ricardo L Pineda, et al. Forecasting power output of solar photovoltaic system using wavelet transform and artificial intelligence techniques. *Procedia Computer Science*, 12:332–337, 2012.

[216] Taskin Jamal, Tania Urmee, Martina Calais, GM Shafiullah, and Craig Carter. Technical challenges of PV deployment into remote Australian electricity networks: A review. *Renewable and Sustainable Energy Reviews*, 77:1309–1325, 2017.

[217] Thomas Schmidt, Martina Calais, Eric Roy, Ashton Burton, Detlev Heinemann, Thilo Kilper, and Craig Carter. Short-term solar forecasting based on sky images to enable higher PV generation in remote electricity networks. *Renewable Energy and Environmental Sustainability*, 2:23, 2017.

[218] Cai Tao, Duan Shanxu, and Chen Changsong. Forecasting power output for grid-connected photovoltaic power system without using solar radiation measurement. In *The*

2nd International Symposium on Power Electronics for Distributed Generation Systems, pages 773–777. IEEE, 2010.

[219] Abhnil A Prasad, Robert A Taylor, and Merlinde Kay. Assessment of solar and wind resource synergy in Australia. *Applied Energy*, 190:354–367, 2017.

[220] Guilherme de Oliveira e Silva and Patrick Hendrick. Photovoltaic self-sufficiency of Belgian households using lithium-ion batteries, and its impact on the grid. *Applied energy*, 195:786–799, 2017.

[221] Oytun Babacan, Elizabeth L Ratnam, Vahid R Disfani, and Jan Kleissl. Distributed energy storage system scheduling considering tariff structure, energy arbitrage and solar PV penetration. *Applied Energy*, 205:1384–1393, 2017.

[222] Yaxing Ren, Saqib Jamshed Rind, and Lin Jiang. A coordinated control strategy for battery/supercapacitor hybrid energy storage system to eliminate unbalanced voltage in a standalone AC microgrid. *Journal of Intelligent Manufacturing and Special Equipment*, 2020.

[223] Anya Castillo and Dennice F Gayme. Grid-scale energy storage applications in renewable energy integration: A survey. *Energy Conversion and Management*, 87:885–894, 2014.

[224] Junjie Yang, Juan Liu, Zilu Fang, and Weiting Liu. Electricity scheduling strategy for home energy management system with renewable energy and battery storage: a case study. *IET Renewable Power Generation*, 12(6):639–648, 2018.

[225] Weiping Zhang, Akbar Maleki, Marc A Rosen, and Jingqing Liu. Optimization with a simulated annealing algorithm of a hybrid system for renewable energy including battery and hydrogen storage. *Energy*, 163:191–207, 2018.

[226] George R Watzlaf and Terry E Ackman. Underground mine water for heating and cooling using geothermal heat pump systems. *Mine water and the environment*, 25(1):1–14, 2006.

[227] Luyi Xu, J Ignacio Torrens, Fang Guo, Xudong Yang, and Jan LM Hensen. Application of large underground seasonal thermal energy storage in district heating system: A model-based energy performance assessment of a pilot system in chifeng, china. *Applied Thermal Engineering*, 137:319–328, 2018.

[228] Ioan Sarbu and Calin Sebarchievici. A comprehensive review of thermal energy storage. *Sustainability*, 10(1):191, 2018.

[229] Renaldi Renaldi and Daniel Friedrich. Techno-economic analysis of a solar district heating system with seasonal thermal storage in the UK. *Applied Energy*, 236:388–400, 2019.

[230] Serguey A Maximov, Sajid Mehmood, and Daniel Friedrich. Multi-objective optimisation of a solar district heating network with seasonal storage for conditions in cities of southern Chile. *Sustainable Cities and Society*, page 103087, 2021.

[231] Kamil Kubiński and Łukasz Szabłowski. Dynamic model of solar heating plant with seasonal thermal energy storage. *Renewable Energy*, 145:2025–2033, 2020.

[232] R Waser, Faisal Ghani, S Maranda, TS O'Donovan, P Schuetz, M Zaglio, and J Worlitschek. Fast and experimentally validated model of a latent thermal energy storage device for system level simulations. *Applied Energy*, 231:116–126, 2018.

[233] MY Abdelsalam, HM Teamah, MF Lightstone, and JS Cotton. Hybrid thermal energy storage with phase change materials for solar domestic hot water applications: Direct versus indirect heat exchange systems. *Renewable Energy*, 147:77–88, 2020.

[234] Arthur Rinaldi, Martin Christoph Soini, Kai Streicher, Martin K Patel, and David Parra. Decarbonising heat with optimal PV and storage investments: A detailed sector coupling modelling framework with flexible heat pump operation. *Applied Energy*, 282:116110, 2021.

[235] Xiufeng Yue, Neha Patankar, Joseph Decarolis, Alessandro Chiodi, Fionn Rogan, JP Deane, and Brian O'Gallachoir. Least cost energy system pathways towards 100% renewable energy in Ireland by 2050. *Energy*, 207:118264, 2020.

[236] Dominik Keiner, Manish Ram, Larissa De Souza Noel Simas Barbosa, Dmitrii Bogdanov, and Christian Breyer. Cost optimal self-consumption of PV prosumers with stationary batteries, heat pumps, thermal energy storage and electric vehicles across the world up to 2050. *Solar Energy*, 185:406–423, 2019.

[237] Kent D Daniel, Robert B Litterman, and Gernot Wagner. Declining co2 price paths. *Proceedings of the National Academy of Sciences*, 116(42):20886–20891, 2019.

[238] Russell Eberhart and James Kennedy. A new optimizer using particle swarm theory. In *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, pages 39–43. Ieee, 1995.

[239] Yuhui Shi et al. Particle swarm optimization: developments, applications and resources. In *Proceedings of the 2001 congress on evolutionary computation (IEEE Cat. No. 01TH8546)*, volume 1, pages 81–86. IEEE, 2001.

[240] Ali Baniasadi, Daryoush Habibi, Waleed Al-Saedi, Mohammad AS Masoum, Choton K Das, and Navid Mousavi. Optimal sizing design and operation of electrical and thermal energy storage systems in smart buildings. *Journal of Energy Storage*, 28:101186, 2020.

[241] C Mekontso, Abdulkarim Abubakar, S Madugu, O Ibrahim, and YA Adediran. Review of optimization techniques for sizing renewable energy systems. *Computer Engineering and Applications Journal*, 8(1):13–30, 2019.

[242] Si Chen, Daniel Friedrich, and Zhibin Yu. Optimal sizing of a grid independent renewable heating system for building decarbonisation. *Frontiers in Energy Research*, page 455, 2021.