Data-driven Model Predictive Control of Buildings

By

Kui Weng

Supervised by:

Prof. Monjur Mourshed

Prof. Alan Kwan

June 2020

A thesis submitted to Cardiff University for the degree of Doctor of Philosophy

School of Engineering

Cardiff University

Wales, United Kingdom

ACKNOWLEDGEMENTS

LIST OF PUBLICATIONS

- Weng, K., Mourshed, M., 2019. RNN-based forecasting of indoor temperature in a naturally ventilated residential building. IBPSA Building Simulation Conference.

- Chen, X., Weng, K., Meng, F. and Mourshed, M., 2018. Smart Energy Management for Unlocking Demand Response in the Residential Sector. Proceedings, 2(15), 1136.

- Meng, F., Weng, K., Shallal, B., Chen, X. and Mourshed, M., 2018. Forecasting algorithms and optimization strategies for building energy management & demand response. Proceedings, 2(15), 1133.

- Weng, K., Meng, F. and Mourshed, M., 2018. Model-based optimal control of window openings for thermal comfort. Proceedings, 2(15), 1134.

- Weng, K., 2016. Performance of UK dwellings in projected future climates. In Applied Energy conference, ICAE.

ABSTRACT

Buildings account for 30% of the final global energy and 28% of the total carbon emissions in the world. Heating, ventilation and air-conditioning (HVAC) systems can consume up to 60% of the total energy consumption in buildings. Improving the energy efficiency of HVAC systems is important in reducing carbon emissions and mitigating risks associated with global climate change such as overheating of indoor environments. Another benefit of improving the energy efficiency of HVAC systems is to save energy cost for building owners. Many previous studies focused on the design and retrofit for improving building energy efficiency, but few of them looked into how to improve the building operation. As the primary building energy system, Commercial HVAC systems are complex because of the interaction of a large number of sub-systems and uncertainties resulting from the interactions of building mass, thermal inertia, weather and occupancy. The application of Model Predictive Control (MPC) has received significant attention in the last few years from researchers and the industry to the control and management of building energy systems. Despite increasing research on using MPC for improving the energy efficiency of HVAC systems, few of them utilise flexibilities such as time of use (ToU) and killowatt Max (kWmax) control.

This research investigates how the control of building elements (such as windows) and HVAC systems could improve energy efficiency and thermal comfort. This study has been divided into two parts based on three case studies. The first part of the study demonstrates a physics-based case study that assesses the impact of climates on the indoor environment and how the control of window openings for natural ventilation can reduce overheating risk in current and future climates. The results find bedrooms are easier to suffer overheating risks than living rooms but increasing openings for natural ventilation is more effective in reducing overheating hours in bedrooms. By opening 20% of window area for natural ventilation, the results show that 2%, 17% and 45% of the total 108 dwellings' bedrooms are overheated in the 2030s, 2050s and 2080s, compared to living rooms with 30%, 60% and 89%. In the 2030s, increasing the window opening area ratio from 20% to 80% can reduce the number of dwellings with overheating risk in bedrooms from 32 to 14, but find nearly no change in living rooms. However, the passive control of building elements such as windows, blinds and overhangs has limitations in adapting dwellings to climate change. With a maximum window area for opening plus blinds and 2-meter overhangs, it can still not eliminate overheating risk in most UK cities in the 2080s.

After demonstrating the limitations of the control of building elements in future climates, the second part of the study introduces two case studies which turn to study the

optimisation of controls for HVAC systems in a residential and a commercial building. The research goes towards the development of data-driven MPC controllers for the two buildings. A sensor network has been established for building energy metering and environmental monitoring in the residential building to enable remote control of the heating system with the MPC controller. It is found that the MPC controller can improve thermal comfort by allowing more hours with room temperatures within the design comfort band. In the commercial case study building, a data-driven MPC controller has been developed, running optimal control of 9 indoor units per 15 minutes to maintain indoor temperatures within the design comfort band. It proposes a demand response method to minimize energy cost by integrating with ToU and kWmax use cases. The study finds that MPC could take advantage of energy tariffs and flexibility by shifting the loads from high-demand periods to low-demand periods. With the data-driven MPC, it could reduce the peak energy consumption by up to 36% and the peak power by about 15%.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AMR	Automatic Meter Reading
ANN	Artificial Neural Network
ARIMA	Auto-Regressive Integrated Moving Average
ARMA	Autoregressive Moving Average
ARX	Autoregressive Exogenous
ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
BMS	Building Management Systems
BS EN	British adoption of a European standard
CFD	Computational Fluid Dynamics
CI	Computational Intelligence
CIBSE	Chartered Institution of Building Services Engineers
CT	Current Transformer
EIA	Energy Information Administration
EMF	Electromotive Force
EPBD	European Performance of Building Directive
ERT	Encoder Receiver Transmitter
EU	European Union
FEES	Fabric Energy Efficiency Standard
GA	Genetic Algorithm
GHG	Greenhouse Gas
GRNN	General Regression Neural Network
GSM	Global System for Mobile Communications
HEMC	Home Energy Management Controller
HVAC	Heating, Ventilation and Air-conditioning
ICE	Indoor Climate and Energy
IEA	International Energy Agency
IECC	International Energy Conservation Code
IES-VE	Integrated Environmental Solutions-Virtual Environment
IoT	Internet of Things
IQR	Interquartile range
KPI	Key Performance Indicators
kWmax	kilowatt Max
LFS	Labour Force Survey
LM	Levenberg-Marquardt
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MPC	Model Predictive Control
MSE	Mean Squared Error
NRMSE	Normalized Root Mean Square Error
NSGA	Non-dominated Sorting Genetic Algorithm
NTC	Negative Temperature Coefficient
ONS	Office for National Statistics

PCM	Phase Change Material
PHE	Public Health England
PID	Proportional-Integral-Derivative
PRT	Platinum Resistance Thermometer
PV	Photovoltaics
R²	The Coefficient of Determination
RBC	Rule-Based Controllers
RC	Resistive Capacity
RH	Relative Humidity
RMSE	Root mean Square Error
RoSPA	Royal Society for the Prevention of Accidents
RTDs	Resistive Temperature Devices
SVM	Support Vector Machine
SVR	Support Vector Regression
TM	Technical Memorandum
ToU	Time of Use
TRNSYS	TRaNsient SYstem Simulation
UHI	Urban Heat Island
UKCP	UK Climate Projections
VRV	Variable Refrigerant Volume

Chapter 1   Introduction

1.1 Background

Globally, buildings consume about 30% of the final energy and generate approximately 28% of energy-related carbon emissions worldwide (UN Environment & International Energy Agency (IEA), 2017). In Europe, buildings account for about 40% of the total energy (European Commission, 2008), generating 36% of the greenhouse gas (GHG) emission (Grozinger, et al., 2014). The total energy consumption in the building sector reached about 39% in the UK (Grozinger, et al., 2014) and 40% in the USA (US EIA, 2018). As the increasing GHG emission significantly contributes to global climate warming (United Nations, 2015), energy conservation and GHG emission reduction in buildings play a vital role in alleviating the impact of climate change. In Europe, the GHG emission reduction targets are to cut emissions (from 1990 levels) in the EU by at least 20% for 2020 (European Commission, 2012), 40% for 2030 (European Commission, 2014) and achieving a climate-neutral society with net-zero GHG by 2050 (European Union, 2020). Improving energy efficiency is critical to achieving these targets.

The energy consumed by heating, ventilation and air-conditioning (HVAC) accounts for up to 60% of the total energy consumption in buildings (Khan, et al., 2015). Effective and efficient control of the HVAC operation is necessary for energy saving. Conventional controllers, also called classical (Belic, et al., 2015), have been widely used in building energy systems due to the low computational complexity and simplicity of the design (ASHRAE, 2015). A Rule-Based Controllers (RBC) is used for the dynamic control in HVAC sub-systems based on "if-else" logic (Afram & Janabi-Sharifi, 2014), such as On/Off control and Proportional-Integral-Derivative (PID) control (Kwadzogah, et al., 2013). The On/Off controller is the most simple and easy to implement that is commonly used in old building systems without digital control, but it is unable to control moving processes with time delays (Serale, et al., 2018). A process controlled using an on/off controller shows fluctuations from the set-points due to the high thermal inertia of many HVAC processes. While PID controllers are implemented in HVAC systems with digital control and variable frequency drives (ASHRAE, 2015), giving promising results. However, the tuning of the controller parameters is burdensome, as well the performance of the controller reduces if the operating conditions differ from the tuning conditions. Many research discussed optimal tuning and auto-tuning methods for PID controllers, which can be time-consuming and unacceptable due to its intrusive nature in relation to the normal operation of the systems (Afram & Janabi-Sharifi, 2014).

Computational intelligence (CI) methods have been widely used instead of traditional statistical methods for their efficiency and faster convergence in forecasting building

energy demand (Namdar & Berenji, 2014). Unlike statistical methods, CI methods have the versatility to deal with temporal variation (Lee & Tong, 2012). Non-linear methods are often more effective in forecasting intermittent energy demand, typically found in building energy applications. The most widely used methods in building energy are artificial neural network (ANN), genetic algorithm (GA) and support vector machine (SVM) (Azadeh & Tarverdian, 2007). There is a growing trend of using CI methods for energy demand forecasting (Zhao & Magoulès, 2012), energy performance modelling (Foucquier, et al., 2013) and energy system control (Dounis & Caraiscos, 2009) for buildings. Many applications of CI methods on HVAC systems have been introduced in a comprehensive review, including design, control, management, optimization and fault detection and diagnosis (Ahmad, et al., 2016).

The developments in the internet of things (IoT) and embedded sensors now afford us to collect enormous data and information from occupants and buildings at a high resolution in a cost-effective manner. It provides an opportunity in applying CI methods and machine learning techniques in developing data-driven based controllers for HVAC systems in buildings. The research aims to optimise the energy and comfort performance of buildings through sensor network technique and CI methods/machine learning techniques in the data-driven model predictive control (MPC) of building HVAC systems.

1.2 Problem statement

The operational control of HVAC systems in buildings is complex because of its nonlinear characteristics, the interaction of a large number of sub-systems and uncertainties resulting from the interactions of building mass, thermal inertia, weather and occupancy. Conventional control strategies have limitations in dealing with nonlinear HVAC systems and do not cope well with the dynamic nature of occupant behaviour and uncertainties in boundary conditions such as weather – often resulting in underperformance in terms of energy efficiency and occupant comfort.

1.3 Motivation

UK dwellings are mostly naturally ventilated buildings that are vulnerable to climate change-related overheating. The increasing incidence of summertime overheating has been found in the dwellings without air-conditioning in Europe and the UK, especially in existing stocks and new dwellings designed for retaining winter heat (Lomas & Porritt, 2017). Many studies on human bioclimates have found an increasing trend in heat stress and a decreasing trend in cold stress in different climate zones between the 20th and 21st centuries (Li, et al., 2011). Due to the increase in heat stress, summer overheating risk is expected to increase in buildings, especially in naturally ventilated buildings. Overheating and heatwave can increase the heat-related mortality to vulnerable

occupants such as the elderly, chronically ill and infants who are less sensitive to temperature variations and physiologically less capable of regulating their body temperatures (Lomas & Porritt, 2017). Natural ventilation is a crucial sustainable measure in improving indoor air quality and maintaining thermal comfort in buildings (Tong, et al., 2016). It is important to optimise the natural ventilation in UK dwellings to reduce summertime overheating risk, especially in future climates.

Thermal comfort is always of great importance for occupants compared to acoustic and visual comfort and indoor air quality (Frontczak & Wargocki, 2011). The growing demand for better indoor thermal comfort increases the installation of HVAC systems, resulting in a significant rise in energy use in building operation. HVAC consumes a large number of total energy consumption of a building, more than 50% in tropical climates (Chua, et al., 2013). The more energy buildings consume, the more GHG they emit, resulting in escalating climate and environmental crisis. Mitigating climate change and global warming require efforts on improving energy efficiency in both design and operation for building energy systems. However, few studies have looked into how to improve building operation to mitigate climate change while most previous studies focused on design and retrofit (Yang, et al., 2014). It has been an increasing concern in improving the efficiency of building operation, maintaining indoor thermal comfort and mitigating climate change. It is vital to optimising the control of the HVAC system to improve energy efficiency without compromising thermal comfort.

Physics-based models are hard for modelling complex HVAC systems because of a gap between the models and real buildings, tuning efforts from domain experts, the variation of models from building to building, and a requirement of repeating system identification for updating models which change with time (Smarra, et al., 2018). With the development of sensing and communication technologies, a great amount of energy consumption and environmental monitoring data can be collected from the sensors, smart meters and devices in the buildings (Kim, et al., 2019). The unprecedented streaming data provides a great opportunity to use data-driven approaches in building energy management for potential operational efficiency and cost control (Zhou, et al., 2016). Data-driven models outperform physics-based models by reducing computational cost and model complexity in modelling, learning non-linear functions between inputs and outputs, optimizing hyperparameters with learning algorithms and its scalability in repeating the methods with historical data in a different building (Smarra, et al., 2018). A machine learning-based data-driven approach is proposed for modelling the HVAC systems in both residential and commercial buildings. This study is to integrate machine learning-based data-driven models with the optimal control of HVAC operation.

Demand response provides an opportunity for buildings to participate in the energy markets by providing energy flexibility and reliving power imbalance in the grid (Hu, et al., 2019). Temperature dynamics in buildings are relatively slow while HVAC operation can be rapidly changed, providing certain flexibility in exploiting building thermal capacity (Mai & Chung, 2016). MPC controller can take advantage of the flexibility to achieve a cost-saving and comfortable indoor environment. It is a model-based technique that has been widely used to improve the performance of control systems such as the HVAC system. It has been found that the energy consumption of the HVAC systems can be reduced by 7% to more than 50% by adding an MPC controller (Kusiak, et al., 2010; Ferreira, et al., 2012). MPC requires dynamical models of the buildings and energy systems to predict their thermal dynamic or energy needs over a prediction horizon. With the advances of dynamic data-driven models, it readily enables MPC for fast online control (Hu, et al., 2019). The study is to develop a near real-time demand response controller for the HVAC system to shift some load/power away from the peak in response to time-varying energy price.

In summary, the motivation of this study is described in the following four aspects:

1. The importance of ventilation control for sustaining thermal comfort and mitigating overheating risks in the UK dwellings;

2. Optimisation of HVAC operation for enhancing energy efficiency and thermal comfort;

3. Advances in data-driven models and machine learning techniques for optimal control of HVAC systems; and

4. Development of near real-time MPC controllers for HVAC systems for energy efficiency and comfort management.

1.4 Research hypothesis

In light of the problems identified above, this research aims to tackle current limitations concerning building HVAC control and optimisation by adopting the following overarching hypothesis:

Data-driven based model predictive control can allow near real-time optimal control of the HVAC systems through integrating machine learning, the Internet of Things and automated control to enhance the building performance by tackling the uncertainties of changing weather and dynamic system behaviours and integrating demand response.

1.5 Statement of purpose and research questions

The purpose of this study is to optimise HVAC control systems for building energy conservation and comfort management with machine learning and IoT techniques.

Based on the proposed purpose of the study, the research questions are set up as follows:

1. What is the impact of climate change on existing buildings?

2. What are the benefits of using the sensor network and IoT for HVAC controls?

3. How are HVAC models and tools used for simulating the behaviour of HVAC systems while considering their scope and limitations?

4. Can the optimisation of HVAC operational control lead to indoor comfort, energy efficiency and cost reduction?

5. Can the integration of data-driven methods and MPC ease the deployment of energy management and control strategies for control systems in buildings on a wider scale and aid building automation?

1.6 Research objectives

The study aims to optimise HVAC control strategies using machine learning and IoT techniques for building energy efficiency and comfort. In response to the research aim and research questions, the research objectives are to:

1. Investigate the barriers of conventional HVAC controls, considering challenges such as climate change and carbon emission targets;

2. Identify the opportunities of using low-cost sensor network and IoT techniques for building environmental monitoring and metering;

3. Establish data-driven forecasting models using machine learning algorithms for modelling the HVAC systems;

4. Implement model-based predictive control for near real-time control and optimisation of HVAC systems for building energy efficiency and comfort; and

5. Appraise machine learning-based control strategies for HVAC controls.

1.7 Research contribution

The key contributions of this research indicate the significance of the research work in controls and optimisation of HVAC systems, which are expressed as follows.

The core contribution of this work is the development and implementation of a near real-time optimal control strategy for HVAC systems, which allows the optimisation and

controls completed within 15 minutes. The near real-time optimal control enables HVAC systems to cope well with uncertainties and shift load response to demand and energy price. Due to its short response time and capability in predicting uncertainties and peak demands, it can be embedded in a simple controller or building energy management systems to improve energy efficiency and manage indoor comfort in real buildings.

Another core contribution is the implementation of machine learning techniques in data-driven models for real HVAC systems in a residential building and a commercial building. It allows us to compare the performance of machine learning-based models with conventional models such as physics-based simulation models in two different kinds of buildings. The state-of-the-art review of machine learning algorithms in modelling HVAC systems makes an additional contribution, providing a better understanding of computational intelligence in HVAC and building controls.

The third core contribution is the establishment of a sensor network monitoring system in a residential building. It forms an IoT based monitoring system for measuring high-resolution data and a test environment to optimise controls through wireless communication.

Apart from the systems, this study provided an improved understanding of the performance of dwellings and overheating risks in present-day and future climates. This allows us to establish the need for automated control of building elements (e.g. fabric) and systems. While current practice relates to the automation of building systems, this research highlights the need for automating the control of building systems and elements together.

1.8 Thesis structure

This thesis consists of eight chapters, and the content of each chapter is briefly introduced in this section.

Chapter 1 provides the broader context of this thesis by introducing the background, problem statement, motivation, hypothesis, objectives, and contribution of the research.

Chapter 2 undertakes a state-of-art review on the widely used modelling methods of HVAC systems, energy and environmental monitoring techniques, machine learning in data-driven models and the optimisation through the MPC for HVAC operational control.

Chapter 3 presents the methodologies for developing the data-driven MPC for the operation of the HVAC systems. The first section outlines data acquisition and storage comprised in the monitoring system. The second investigates the forecasting models for HVAC systems. Finally, the third section explores the use of MPC for near real-time optimal control of HVAC systems.

Chapter 4 aims to investigate the impact of climate change on the thermal comfort of occupants. It establishes a study in investigating how future climates influence the overheating risks in the UK dwellings. It emphasises the requirement of the control of building elements for thermal comfort management.

Chapter 5 introduces the case studies in part II of the study and the analysis of the data from the case study buildings.

Chapter 6 is to develop the data-driven models for the HVAC systems in two experiment buildings. It evaluates the performance of the widely used machine learning algorithms in the day-ahead forecasts.

Chapter 7 develops the data-driven based MPC for near real-time optimal control of the HVAC systems in both case study buildings. It also introduces how optimisation of the controls of HVAC systems is achieved by the MPC.

Chapter 8 concludes the work completed in previous chapters and outlines the key contributions of the study.

Chapter 2   Literature review

This chapter provides a comprehensive literature review in five sections, including the links between climate change and overheating, HVAC forecasting models, energy metering and environmental monitoring, machine learning and MPC for HVAC operation. Each section is linked to another, aiming to answer the research questions and test the hypothesis presented in Chapter 1. Section 2.1 begins with identifying the impact of climate on naturally ventilated residential buildings. It provides insight into high overheating risk for buildings in future climates and the importance of control in building HVAC systems and elements to mitigate the impact of climate change. Section 2.2 introduces the most widely used methods for modelling and forecasting of HVAC systems and brings out the advantages of data-driven approaches for near real-time operational control of HVAC systems. Section 2.3 reviews building energy metering and environmental monitoring, providing high-resolution data for applying data-driven methods using data from the formed monitoring system. Section 2.4 investigates the machine learning techniques in data-driven models. Section 2.5 identifies the MPC in the near real-time control of HVAC systems.

2.1 Climate change on buildings

2.1.1   Overheating risk

High summertime temperatures lead to a rise in indoor overheating risks, resulting in thermal discomfort and heat-related health issues to the occupants (Santamouris & Kolokotsa, 2015). Indoor temperatures were found to reach 30°C in living rooms and 36°C in bedrooms in Manchester homes (37.9°C in a London flat) when the daily average outdoor temperature was above 20°C (Lomas & Kane, 2013). The UK Climate Projections 2009 reported (UKCP09) that average summer temperature would increase by 1.3 - 4.6°C in London by the 2050s under a high emission scenario (UKCP, 2009). In 2013 a heatwave hit the UK, particularly in the South and West UK, with daily maximum temperatures exceeding 30°C for seven consecutive days from 13 to 19 July and somewhere in England exceeding 28°C for 19 consecutive days from 6 to 24 July (Met office, 2013). The impact of heatwave on mortality was considerably less in 2013 (301 deaths) than in 2006 (2323 deaths) and 2003 (2234 deaths) (Green, et al., 2016). Although heatwaves are unusual extreme weather scenarios according to current climate projection, they will occur every two to three years by the 2050s (Greater London Authority, 2008). Life-threatening heat-related illness, such as heatstroke and sunstroke increased significantly in the 2013 heatwave (Elliot, et al., 2014). Especially for the

elderly, more evidence supported that the comfortable temperature they feel is actually not healthy for them.

Previous guidance and criteria have defined overheating risk for UK dwellings. According to the Chartered Institution of Building Services Engineers (CIBSE) Guide A in 2006, overheating risk should be considered when the hours of indoor operative temperature exceeding 26°C for bedrooms (28°C for living rooms) is over 1% of annual occupied hours (CIBSE, 2006). In 2013, the CIBSE updated its Technical Memorandum (TM) for overheating criteria TM52 (Nicol, 2013), which applied an adaptive comfort approach based on European and British adaptive thermal comfort standard BS EN15251 (2007). TM52 considered human's adapting ability to temperature change and redefined overheating threshold temperature relating to the running mean outdoor temperature. The hours of indoor operative temperature exceeding the overheating threshold should not exceed 3% of annual occupied hours. In 2017, CIBSE's newest overheating criteria TM59 (CIBSE, 2017) updated overheating criteria for bedrooms, stating that the total bedroom occupied hours with temperature exceeding 26°C should not be over 1% of annual occupied hours (1% of annual hours between 10 pm and 7 am is 33 hours). This overheating criterion was established considering occupant thermal comfort, defining the overheating that occurred at a point or range when occupants experienced thermal discomfort.

2.1.2   UK housing stock

The 2016-2017 housing stock survey showed that there were around 27.2 million dwellings in the UK. As shown in Figure 2-1a, there is more housing stock in England, especially in South East England and London. The majority of dwelling types are houses, including detached (17%), semi-detached (25%), and terraced houses (28%). Bungalows and Low-rise purpose-built flats take account for 9% and 15% of the building stock, respectively. High-rise and converted flats are the least, accounting for 2% and 4%, shown in Figure 2-1b. Most of the dwellings are in England (around 83%) and Scotland (around 9%), according to 2017's households' survey. Figure 2-1c showed a steady increase in the UK housing stock from 1996 to 2017. The UK housing stock is generally old. The English housing survey 2016-2017 (shown in Figure 2-1d) revealed that 21% of the houses were built before 1919 in England, and more than 75% of the English dwellings were built before the 1980s (MHCLG, 2018). Figure 1e presented the usual insulation measurements of dwellings, including cavity, wall insulation, double glazing, and loft insulation. Most of the dwellings have cavity walls and the entire house double glazing, about 93% and 83%. There has been an increasing need for cavity walls and insulation to comply with building regulations since 1990. Dwellings are becoming more air tightened to meet new building regulations and carbon emission reduction

targets, such as the UK's 'Zero Carbon' homes target by 2016. However, a high level of insulation could improve energy efficiency by reducing heat loss. But at the same time, it could also increase the airtightness of the dwellings resulting in lower air infiltration and higher overheating risk in the summer (McLeod, et al., 2013).
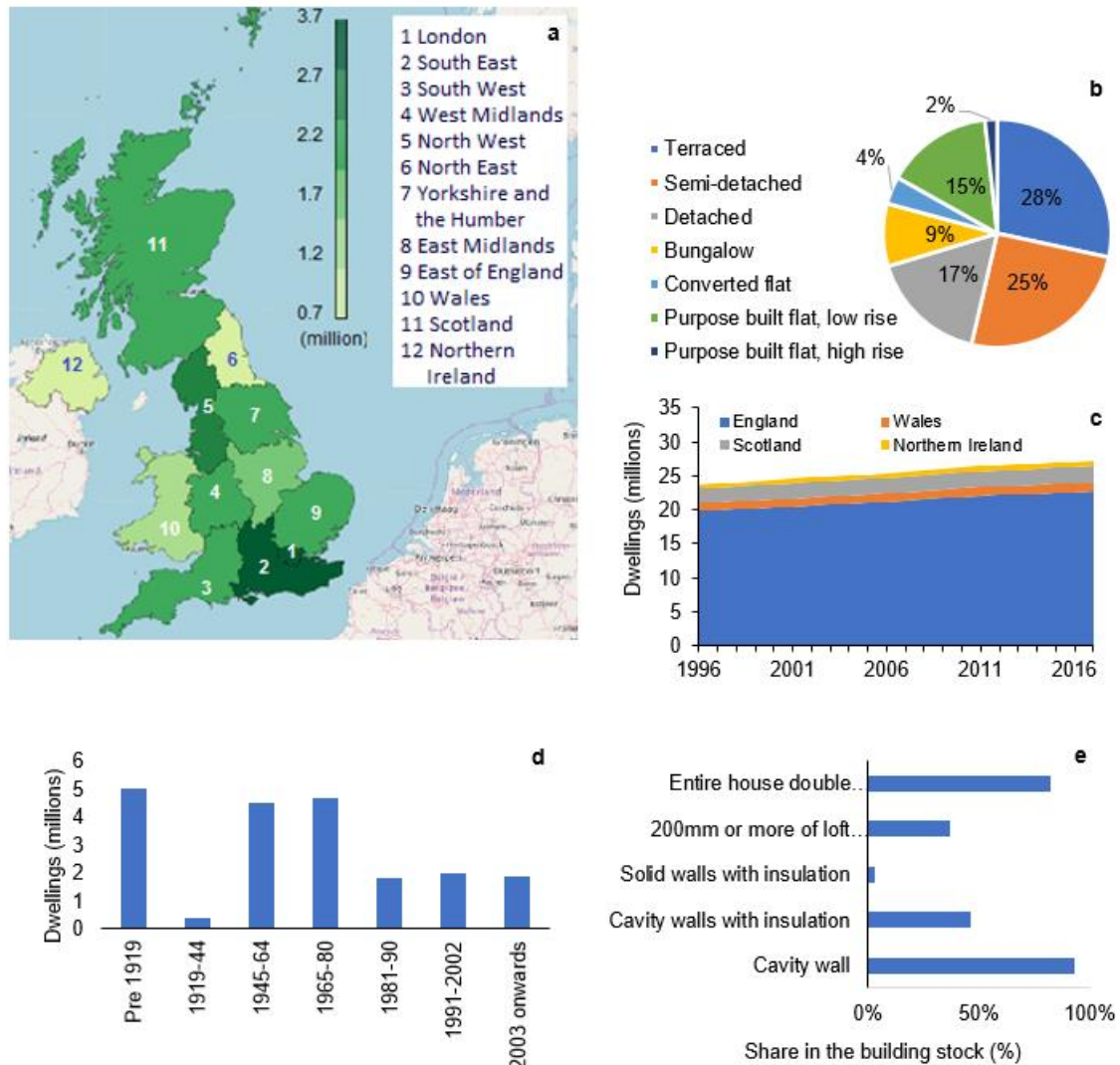
Figure 2-1 The UK housing stock. (a) The number of homes by administrative regions. (b) Percentage of English dwellings by building types. (c) Building stock by country between 1996 and 2017. (d) English dwellings by age. (e) English dwellings by insulation measures. Source of data: (a) and (c) Labour Force Survey (LFS), Office for National Statistics (ONS, 2018); and (b), (d) and (e) English Housing Survey headline report 2016 to 2017 (MHCLG, 2018).

Non-air-conditioned buildings are more vulnerable to climate change (Lomas & Porritt, 2017). Dwellings in the UK are mostly built with heating systems and natural ventilation. They are projected to experience high indoor summertime temperatures as extreme weather events such as heatwaves are becoming more common and lasting longer due to anthropogenic climate change (Mavrogianni, et al., 2012). Gupta and Gregg used Integrated Environmental Solutions-Virtual Environment (IES-VE) for modelling main

types of the UK dwellings based on UKCP09 probabilistic weather data within a series of scenarios to examine the impact of several passive adaptation measures on overheating hours. They suggested that flats and mid-terraced houses had much more risk of overheating in future climatic scenarios than detached or semi-detached houses, and user-controlled shading was the most effective passive cooling measures (Gupta & Gregg, 2012).

2.1.3 Natural ventilation and controls

Natural ventilation is vital as most UK dwellings are built with natural ventilation which is usually the only cooling measure in hot summer. A BRE study of ventilation and air quality of dwellings has been implemented since 1995 when building regulations were revised. The further studies were carried out for pilot study, main study and peak level study in 2001, 2002 and 2003 to measure tightness, temperature, humidity, indoor air ventilation, and pollutant concentrations in homes during occupancy (Dimitroulopoulou, et al., 2005). A study of ventilation in European dwellings pointed out that 30% of UK homes in summer had an inadequate ventilation rate of fewer than 0.5 air change per hour (ach^{-1}) which was recommended to decrease the growth of mould and 0.8 ach^{-1} for house dust mites (Dimitroulopoulou, 2012).

Night-time ventilation is highly recommended for cooling the buildings in the summer. The basic concept of this strategy is to cool the structure overnight during the occupied hours and form a heat sink to guarantee thermal comfort without mechanical cooling during the day (Kolokotroni, et al., 2001). Givoni suggested that night-time ventilation is best suited to arid regions where night-time temperature is below 20°C, and the daytime temperature is around 30°C to 36°C which natural ventilation is not able to achieve indoor comfort (Givoni, 1994). Artmann et al. proposed a degree-hours based method using the climatic data of 259 European stations to analyse the effect of night ventilation for passive cooling and suggested that night ventilation (assumed constant airflow rate of 6 ach^{-1}) has a high potential for passively cooling the commercial buildings in Europe (Artmann, et al., 2007). McLeod et al. compared a naturally ventilated Fabric Energy Efficiency Standard (FEES) dwelling with mechanically cooled Passivhaus dwellings under UKCP09 projected weather in central London and found that both of them suffered similarly high overheating risks by 2050 (McLeod, et al., 2013). However, the FEES dwellings used purge ventilation which restricted window opening with maximum opening angles of 10°, which limited the natural ventilation rate for cooling. Such an opening angle is for using the window restrictors in UK new-build social housing to comply with the 2014 guidance from the Royal Society for the Prevention of Accidents (RoSPA) to prevent children falling from the windows (RoSPA, 2014). In addition, the

Urban Heat Island (UHI) effect from 11 pm to 8 am in central London also affected the cooling (McLeod, et al., 2013).

2.2 HVAC system modelling and forecasting

2.2.1 Physics-based model

It is also known as a white-box model, requiring the understanding of the underlying principles and process physics. The physics-based models is generally created and simulated on simulation tools. The model has good accuracy in forecasting but requires expert knowledge for selecting structure and parameters. Simulation tools such as EnergyPlus (Zhao, et al., 2015) are usually used in the design stage. The optimisation of models in the design stage does not need to consider the complexity and computational time. However, the MPC for real-time control has to consider the time consumption and accuracy. Due to the complexity of the simulation model, it is difficult to be used in MPC for real-time control. To overcome this problem, Ascione et al. (Ascione, et al., 2016) implemented a simulation-based MPC for the control of the HVAC system, suggesting a minimum run period of 10 days covered by EnergyPlus simulation to reduce the computational time in the optimisation stage as several runs of the simulation were required in this stage.

The physics-based method uses building simulation engines (e.g. EnergyPlus), with their multi-domain (thermal, lighting, network airflow, etc.) modelling capabilities, to predict energy consumption (Mourshed, et al., 2003). The simulation of buildings based on physical models can reduce uncertainties of evaluation and provide fine spatial and temporal resolution. However, the use of engines requires computation time for simulation, making them unsuitable for online or near-online applications (Ahmad, et al., 2016). This method can monitor the dynamic behaviour of the thermal storage system (e.g. heat pump) for load control. Arteconi et al. designed a physical model for the heat pump to shift electricity peak load by deciding when to supply hot water (Arteconi, et al., 2013). The system can be switched on and off for shifting load but had the interval of 6 minutes between each switching to avoid excessively rapid cycling of the pump. A similar method was used by Lee and Braun for simulation chiller power, and the computation time is around 15 minutes to 1 hour as time is required for data generation and simulation (Lee & Braun, 2008).

The physics-based simulation method uses time-series weather data and building physical parameters in the dynamic simulation. Building parameters such as construction details are used in building simulation models for energy consumption. The simulation is based on physical modelling of the heat transfer and can be written via the energy conservation law as follows:

$$Q_{\text{in}} + Q_{\text{source}} = Q_{\text{out}} + Q_{\text{store}} \qquad (4.2)$$

Q_{in} is the heat flux entering the system. Q_{source} is the heat flux of an eventual heat source. Q_{out} represents the heat flux leaving the system and Q_{store} represents the heat flux stored.
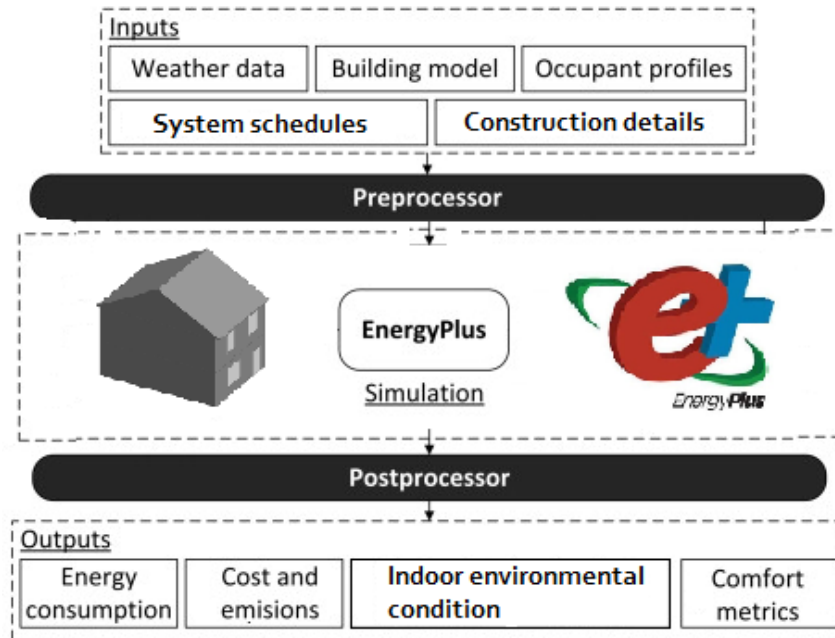
Figure 2-2 The schematic of a typical EnergyPlus simulation model.

The heat transfer equation takes account of the conduction through walls, the convection, the short and longwave radiation and the ventilation (Foucquier, et al., 2013). Physics-based modelling uses a multi-zone approach that each zone is considered to be homogeneous for the entire building. Based on this approach, there are some commonly used simulation software such as EnergyPlus (EnergyPlus, 2020), ESP-r (ESP-r, 2020), TRANSYS (TRNSYS, 2019), and e-QUEST (eQUEST, 2018). Figure 4-2 presents a schematic of the physics-based simulation model using EnergyPlus. The inputs and outputs have been illustrated in Figure 2-2.

Typically, MPC will need the model to be computationally efficient on real-time HVAC control and optimisation. In order to reduce the model complexity, the reduced-order model, such as a simplified mathematical model, is usually used to describe the thermal network and system behaviour (Salakij, et al., 2016). The simplified model will still need an understanding of the physical correlation among parameters for reflecting the interaction between the sub-systems. It has been a major challenge to develop an accurate simplified mathematical model for predicting the nonlinearity of the HVAC system. An alternative is to use a state-space model. It is generally developed based on a thermal network such as the resistance-capacitance (RC) method. This method uses an electrical network to model the heat transfer between HVAC sub-systems, while

electrical resistor, capacitor, current and voltage represent thermal resistivity, capacitance, heat transfer and temperature respectively. (Afram & Janabi-Sharifi, 2014). The state-space model has been widely used for creating control-oriented models in MPC (Hu & Karava, 2014). Based on this method for predicting building systems, the optimisation process can be linear (Braun & Chaturvedi, 2002) and nonlinear (Goyal & Barooah, 2012).

2.2.2 Grey box model

The Grey box model requires both knowledge and data of physical phenomenon to complete a model. It is generally a non-linear model using the physics-based models as the model structure and measured data as the model parameters. Therefore, this approach provides physical meaning and better generalisation capability to a data-driven model. The estimation algorithms for selecting the particular parameters decide the most common grey box approaches, including simplex search, nonlinear least squares, and genetic algorithms (GA) (Afram & Janabi-Sharifi, 2014).

Balan et al. (Balan, et al., 2011) introduced a grey box model based on a thermal model and genetic algorithm. The thermal model contains parameters, including outdoor temperature, indoor temperature, and energy consumption. The approach is to develop a lumped parameter model as the thermal model to represent heat transfer between the zone air and wall using the concept of resistive capacity (RC) circuit. Wu and Sun (Wu & Sun, 2012) developed a grey model based on thermodynamic equations and autoregressive moving average (ARMA) algorithm, predicting the room temperature accurately up to ten weeks. However, the equation requires a lot of heat- and energy-related parameters, including temperature parameters (room, discharge air, wall surface, inside window surface), and airflow parameters (discharge air density, the heat capacity of air, heat gain from the solar flux and internal sources such as human and electrical devices).

2.2.3 Data-driven model

A data-driven model, or black-box model, fits the function model using the real data as inputs and outputs. Compared with the physics-based model, it does not require expert knowledge of system physics. It can model the internal relationship among parameters requiring only the input-output data from the system or building environment. Therefore, it has its advantage of the ease of model development (Afram, et al., 2017).

The data-driven model is used in the MPC to identify complex systems such as HVAC systems considering energy, cost and thermal comfort. Various data-driven models can be used as the predictive model for building HVAC systems, including statistical models (e.g., Autoregressive exogenous (ARX) (Ma, et al., 2012)), Artificial Neural Network

(ANN) (Afram, et al., 2017), Random Forest (RF) (Afram & Janabi-Sharifi, 2014). Although the statistical models, such as ARX, show their ability to linearise the model, there is still a challenge to forecast accurately (Li, et al., 2015).

2.3 Energy metering and environmental monitoring

Energy metering can be useful in building energy inefficiencies identification, energy benchmarking, load planning, and demand management for balancing system reliability and price volatility (Genet & Schubert, 2011). A comprehensive building energy metering and environmental monitoring provide valuable information about building performance and play an important role in data-driven measures on energy efficiency (Ahmad, et al., 2016). Therefore, it can involve all stakeholders, including tenants, property owners and energy managers, to implement the measures for energy efficiency (Genet & Schubert, 2011).

The updating energy-related regulations and legislations have more requirements in building energy efficiency, which motivate the development of building metering, environmental monitoring and automated controls (Ahmad, et al., 2016). European Performance of Building Directive (EPBD) requires European members to release building regulations to encourage intelligent metering and building automation in new and renovated (European Commission, 2012).In the USA, metering is required for federal buildings in the energy Policy Act of 2005. The requirement for metering and control was enhanced by the International Energy Conservation Code (IECC), which has been adopted in most US states (International Code Council, 2016). In the UK, energy suppliers were required to roll out smart meters to comply with energy regulations EnergyAct 2008 and Energy Act 2011, which aimed at reducing energy use, improving the environment and saving energy cost (DBEIS&Ofgem, 2013). Government guidance such as Smart meters: a guide has mentioned that smart meters will be implemented across the country by the end of 2020 (DBEIS, 2018).

2.3.1   Electricity

Electricity meters measure energy consumption by sensing the voltage and current. The traditional electricity meters were electromechanical meters containing a metal disc to rotate at speed in proportion to the power through the meters. The inside disc can be influenced by the induction coils producing magnetic flux proportional to the voltage and the current (Edison Electric Institute, 2014). The old mechanical meters must measure energy usage by reading the dial directly. Lagging is an issue for those mechanical meters. According to the side-by-side testing by Oncor, 5% of the mechanical meters have a lagging problem, resulting in an inaccuracy of more than 2% for electricity metering (although 2% is still accurate for mechanical meters). Due to no storage and

communication of data, meter readers were sent to come to read the measurement. Fraud is another issue as meters could be disabled by customers when they are not around.

Morden electricity meters can generate electronic or pulse outputs using encoders or other registering mechanisms (Parker, et al., 2015). One example is the encoder receiver transmitter (ERT) developed by Itronwhich allows data to be transmitted from the utility meters to a utility vehicle through a packet radio protocol. Although electromechanical meters have been used in residential and commercial buildings for many years, electronic meters are replacing them with many advantages (Sun, et al., 2015). Electronic meters have no moving parts and are good at displaying, storing and transmitting data with higher accuracy. Morden electronic meters can measure various types of data, such as maximum power demand and support time-of-day billing. They can track the direction of energy if solar panels or wind turbines have been installed in a residential or commercial building in a power network. With wireless technology such as ERT for remote metering, automatic meter reading (AMR) can automatically collect energy consumption and transfer the data to a data centre for billing or analysis. Although modern meters, such as AMR meters, have many advantages, some customers still want to keep their old analogue meters. The prime reasons are the dirty electricity (voltage spikes, surges and high-frequency variations forming a harmful electromagnetic field in the buildings) and overexposure to electromotive force(EMF) radiation (Havas, 2006). Dirty electricity and the growing exposure to radiofrequency radiation have been recognized to cause public health epidemic (Kleiber, 2017).

Table 2-1 Characteristics of domestic and commercial electronic electricity meters.

Parameter	Use				
	Domestic		Commercial		
Phase	1	3	1	3	3
Connection	Direct	Direct	Direct	Direct	CT
Maximum voltage (V)	290	400	275	415	1500
Maximum current (A)	100	125	100	100	100
Minimum error (%)	1	1	1	1	0.2
Maximum error (%)	2	2	2	2	1
Maximum number of tariff rates	8	4	2	8	16
Max power dissipation in current circuit per phase (W)	0.5	0.5	0.8	0.5	0.5
Price ($)	20 - 150	75 - 245	25 - 208	68 - 125	68 - 3500

AMR has been enhanced by the development of communication technologies such as ZigBee, radio frequency and global system for mobile communications (GSM) (Rawat, et al., 2016). It has been used by some business users in the UK, helping them monitor energy use and control the costs. The meter contains a remote reading device either connected to or embedded in the body. The signal sent by AMR meters can be picked by a 'collector' such as a meter reader's vehicle. Through AMR, no manual reading is required as the energy supplier can remotely get the readings and calculate the monthly energy bills for the users. Unlike AMR meters, smart meters are normally two-way meters connected to the smart grid while AMR meter is capable of one-way wireless communication from the energy supplier side. Smart meters enable users from domestic and business to monitor their energy use in near real-time and subsequently change the energy use behaviour. They are operated through a centralised data communication company. Besides, the automatic outage reporting function is embedded in the smart meters (Jiang, et al., 2016). Due to the capability of two-way communication, it is considered to be worse than AMR meters as more frequent emissions and dirty electricity pose a long-term risk on human's health. However, the distance to the meter makes a significant difference in the emissions received by the occupants (Pope, 2019). It is recommended to sleep as far from the meter as possible and stay away from it as much as possible.

Table 2-1 illustrates the parameters of electronic electricity meters in domestic and commercial buildings. The three-phase meter is significantly more expensive than the single-phase meter, especially for the current transformer (CT) connected meter which transforms high current (over 100Amps) to low current (e.g.5Amps). Those CT meters can cost more than three thousand pounds. However, they can provide higher accuracy (e.g. 2%), while most meters can only reach 1% or 2%. Both domestic and commercial electricity meter can have dual or multi tariff rates. The maximum power dissipation in the current circuit per phase is around 0.5 to 0.8 W.

2.3.2 Gas

Gas consumption is measured with gas meters, which should comply with the local codes. There are different types of gas meters on the market, while diaphragm, rotary and turbine are the most widely used. The diaphragm gas meters operate through the two or multiple chambers formed by the moving diaphragms. The gas flow is measured by counting the rotary motion converted from the linear motion of the diaphragms. Rotary gas meters can provide higher measuring precision over the higher gas flow rate. This type of gas meter uses two aligned spinning motors to allow a certain amount of gas to pass through each complete turn. The rotation of motors is related to the gas flow rate and recorded using mechanical counters or electrical pulses. Turbine gas meters

measure the gas consumption by measuring the gas velocity through the meters. They have been often used in the industrial applications where the continuity of the flow is crucial and high accuracy and reliability are needed, for instance, custody transfer metering (Cascetta & Rotondo, 2015).

All three types of gas meters use dynamic mechanisms to measure the flow rate. Although these meters have been widely supplied and installed by the energy suppliers due to their simplicity and low cost, they have disadvantages such as high-pressure losses, wear and tear, mechanical outputs and inability to show an instantaneous gas flow rate. In recent years, static gas flow meters have been encouraged because of the accuracy, safety and remote-reading. Static ultrasonic flow meters are well known for its non-intrusiveness and performance characteristics, including data-recording, auto-diagnostic and integration with electronic sensors. (Buonanno, 2000). Table 2-2 presents the parameters of four different types of gas meters. The technical data were collected from manufacturers for comparison. Diaphragm meters have been widely used in domestic and commercial areas as they can measure a considerably narrow range in flow rate. Rotary and Turbine meters are designed with a higher flow rate which allows them for fiscal metering. Ultrasonic can measure a broader range of flow rate. For example, a domestic one can measure a small flow rate from 0.04 to 6 m3/h. However, the accuracy can be improved if the size of the tube is larger than 150 mm from the lowest of 5% to 1%.

Table 2-2 Characteristics of the most common gas meter types.

Parameter	Meter type			
	Diaphragm	Rotary	Turbine	Ultrasonic
Min flow rate (m^3/h)	0.001	0.25	2.5	Speed dependent
Max flow rate (m^3/h)	100	650	10000	Depends on speed
Min temperature (°C)	-25	-20	-25	-10
Max temperature (°C)	65	60	70	40
Accuracy at maximum flow rate (± %)	1 - 1.5	1 - 1.5	1 - 1.5	1.5 - 2
Min cyclic volume (l)	1.2	-	-	-
Max cyclic volume (l)	65	-	-	-
Use	Domestic/ Commercial	Fiscal	Fiscal	Domestic/ Commercial/ Fiscal

2.3.3 Temperature

Air temperature can be defined as 'the thermal of the air surrounding the human body, which represents the aspect of the surroundings and determines the heat flow between the body and air (Parsons, 2014). In the building, the indoor air temperature is important for the occupants regarding human's thermal comfort, health and wellbeing (Teli, et al., 2018). The room temperature is suggested to be a modifiable risk factor in hypertension-related morbidity and mortality as it could influence people's blood pressure, especially for those who do not take physical activity regularly (Zhao, et al., 2019).

Various types of sensors can be used to measure temperature by sensing the change in physical properties. Table 2-3 illustrates the most common types of temperature sensors in the industry. The most common types of temperature sensors are thermocouples, resistive temperature devices (RTDs) and thermistors (Tong, 2001). The measurement of temperature is based on different kinds of phenomena, including thermoelectricity, thermal expansion, the temperature-dependent electrical resistivity of the conductor, and spectral characteristics and fluorescence (Childs, et al., 2000). Most of the temperature sensors rely on mechanical and electrical principles. Mechanical sensors are easily installed and maintained without a power supply. One example is the room thermostat using bimetallic thermometers to measure temperature through different thermal expansion rate of two different metals (Venkateshan, 2015). Another example is the fluid-expansion thermal thermometers which use mercury or organic-liquid. However, it is hard to record or transmit data, and they are not as accurate as of the thermocouples or RTDs (Omega, 2018).

Electrical sensors can convert the measured data to digitally encoded signals which allow direct communication for control and management. Thermocouples temperature sensors are based on thermoelectric effect with conversion between the temperature and voltage. They have a wide measuring range (e.g. type K thermocouple's measuring range from -200°C to +1250°C) and reasonable accuracy (within 1°C to 2°C traditionally) (Duff & Towey, 2010). RTDs measures temperature based on the resistance changes of a conductor corresponding to the temperature change. This type of sensors is accurate, reliable and stable but has some disadvantages such as large size, high time response and cost (Duff & Towey, 2010). One of the RTDs used in the building sector is the platinum resistance thermometer (PRT is an RTDs using platinum as the conductor). Instead of using pure metal, thermistor uses semiconductor such as metallic oxides with lower cost in a smaller size. However, one of the problems is a drift in resistance value over time (Micro-Chip Technologies, 2010). Morden thermistors have improved its accuracy from the best tolerance of 5% in the past to an acceptable error of 0.18°C (Tong, 2001).

Most of the temperature sensors can measure temperature between -20°C and 70°C. However, the accuracy will be decreased if the temperature is not within the range between 0°C and 50°C. Due to the drop of the sensing accuracy in extreme temperature, those sensors are not suitable for outdoor temperature metering but are still good for indoor use. Most of those sensors are negative temperature coefficient (NTC) thermistors measuring similar temperature range and accuracy. One of them has thin-film RTD sensors inside, which allows it to measure a wider temperature range (e.g. -200°C to 800°C) with high accuracy and resolution. However, it is more expensive than the common temperature loggers with internal NTC thermistors. Radio Frequency is still the most common communication methods for wireless temperature logger. However, Wi-Fi, ZigBee and Bluetooth are becoming popular. Especially, ZigBee is often used to form the mesh network in the buildings. Bluetooth is often used to pair with cell phones through applications. Therefore, temperature loggers with Bluetooth normally do not have a central gateway to allow sensors sending data.

Table 2-3 The most common temperature sensors in the industry.

Parameter	Meter type			
	Thermocouple	**RTD**	**Thermistor**	**Integrated Silicon**
Temperature range (°C)	−270 to 1800	−250 to 900	−100 to 450	−55 - 150
Accuracy (± °C)	0.5	0.01	0.1	1
Responsiveness in stirred oil (Sec)	<= 1	1 - 10	1 - 5	4 - 60
Excitation	None Required	Current Source	Voltage Source	Typically supply voltage
Form of Output	Voltage	Resistance	Resistance	Voltage, current, or digital
Price ($)	1 - 50	25 - 1000	2 – 10	1 - 10

2.3.4　Relative humidity

Relative humidity (RH) is the ratio of actual water vapour density to the saturated water vapour density at the same temperature. The RH value is increased when air is cooled as cold air can hold less moisture. RH is important for thermal comfort and health of occupants in the buildings (Vellei, et al., 2017). RH sensing is usually based on measuring wet-dry bulb temperature and monitoring the change of material properties, such as their physical or electrical parameters. Various methods have been developed to measure humidity, from the simplest methods based on the expansion and contraction of the materials (e.g. human hair) to complex methods by using miniaturized chips. (Tan, et al., 2005). The whirling hygrometer is one of the most often used instruments for measuring relative humidity. By rotating two thermometers, it measures the wet- and dry-bulb temperature for calculating or searching the RH value in the tables (Parsons, 2014). Based on the change in electrical capacitance and resistance to humidity, the electronic humidity sensors are mainly classed into capacitive humidity sensors and resistive sensors. In capacitive humidity sensors, they measure the RH by detecting the dielectric change of the conductor, which is proportional to RH. Resistive sensors measure the RH based on the resistance change of the hygroscopic medium.

Both types of sensors are low cost and power consumption, covering a wide range of humidity with good repeatability. However, they suffer from cross-sensitivities and temperature dependency on some chemical species (Blank, et al., 2016). In recent years, many kinds of materials have been used in water or humidity sensing based on the change in electrical parameters of materials, including electrolytes, organic polymers, ceramics and composite materials (Erol, et al., 2011). Polymers and porous ceramic are widely used in commercial humidity sensors (Arshak & Twomey, 2002). Compared with electrical humidity sensors, optical fibre-based humidity sensors have serval advantages, including fast time response, no electrical contacts and electromagnetic immunity, allowing them to work in flammable environments. The disadvantages are the high cost, the complexity of the system and precise installation (Yeo, et al., 2008).

Table 2-4 presents some RH loggers found on the market. Most humidity sensors are capable of measuring RH between 0 to 100%. However, most of them can measure the highest accuracy of RH between 25% and 75%. Out of this range, the error will increase from around 3% to 10%. Same as temperature logger, the RH loggers often use radio for wireless communication. Other than radio, ZigBee has been increasingly used in wireless RH logger.

Table 2-4 A summary of capabilities of RH sensors on the market.
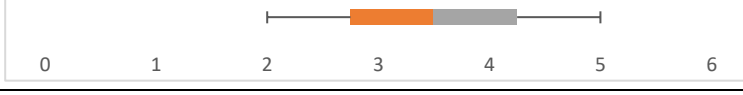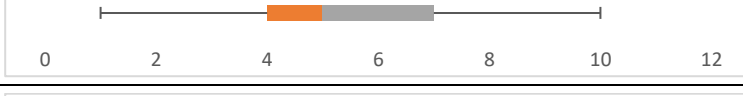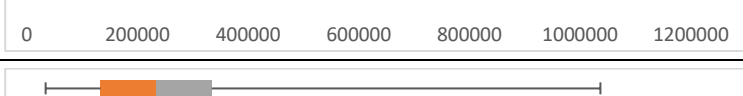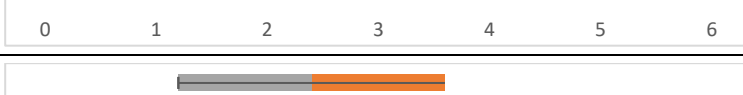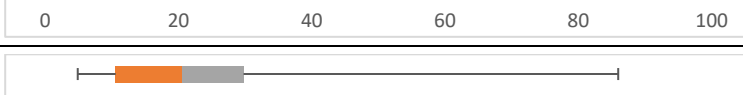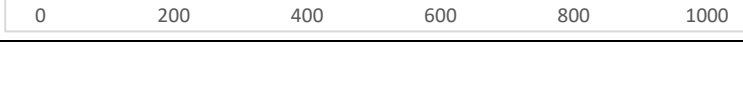
Parameter	Min	Max	Box chart
Min RH (%)	0	15	
Max RH (%)	90	100	
Resolution (%)	0.001	1	
Response time (s)	11	720	
Max error (%)[limited]	1	5	
Max error (%) [full]	1.5	10	
Sample rate from (minute)	0.02	11	
Sample rate to(minute)	3	1440	
Max transmission range (m)	0	152	
Max battery life (month)	3	36	
Max readings	4000	3200000	
Price ($)	15	678	

2.3.5 Daylight

As a passive strategy, daylighting is important to occupants' vitality, performance and visual comfort in the buildings. The design of modern lighting systems needs to take into account the visual comfort of people. As a natural source, good design of utilizing daylight can also increase the efficiency of space lighting and reduce the energy consumption of lighting (Nasrollahi & Shokri, 2016). The light sensor is often used to detect daylight or light illuminance level, which can be used to control the light switching on or off, or dimming by setting the threshold of daylight. It is typically placed on the lighting device to adjust the artificial lighting to achieve a design illuminance. Combining light sensors for local illumination can efficiently save the lighting energy and tune the illumination. The lighting system can also be used in conjunction with occupancy sensors

to turn off or reduce lighting requirements in unmanned areas (Pandharipande & Caiced, 2015). In Europe, the standard about illumination such as EN 12464-1 recommended a minimum illuminance values of 500 lux (occupied) and 300 lux (unoccupied) in the office building (European Committee for Standardization, 2002).

Table 2-5 Characteristics of the common light sensor on the market.

Parameters	Min	Max	Box chart
Min. illuminance (lux)	0	322	
Max. illuminance (lux)	220	400000	
Highest accuracy (±%)	2	5	
Lowest accuracy (±%)	1	10	
Highest accuracy (± lux)	5	30	
Lowest accuracy (± lux)	100000	100000	
Min. resolution (lux)	0.01	2	
Min. sampling Interval (second)	0.4	60	
Record capacity (readings)	3000	1000000	
Battery life (year)	0.0057	5	
Min. working temperature (°C)	-40	0	
Max. working temperature	30	85	
Price ($)	56	867	

2.4 Machine learning techniques

Machine learning algorithms have been widely used for creating data-driven (or black-box) models to explain the interactions among the measured input and output variables.

2.4.1   Artifice Neural Network

Artifice Neural Network (ANN) is a non-linear statistical method to determine the complex relationship between the input and output without prior knowledge (Haykin, 1994). This method is usually used to trace previous load patterns for predicting future loads and can adapt to future changes with new data (Park, et al., 1991). It is capable of forecasting short-, medium-, and long-term energy consumption (Deb, et al., 2017). Compared with physics-based methods in forecasting HVAC energy demand, ANN showed less error with a simpler model which can use only temperature as input (Neto & Fiorelli, 2008).

Among the CI/machine learning methods, ANN is the most widely used artificial intelligence algorithm for predicting building energy demand (Zhao & Magoulès, 2012). It can be used for forecasting demands for various building energy applications, including singe application (e.g. heating) (Díaz, et al., 2001), multiple applications (e.g. cooling and lighting) (Aydinalp, et al., 2002), and building system with solar energy system (Kalogirou, 2006). The forecasting by ANN can rely on the single input variable (e.g. external temperature) (Ben-Nakhi & Mahmoud, 2004), or multiple variables (e.g. occupancy and weather) (Kalogirou & Bojic, 2000). Both short-term (Gonzalez & Zamarreno, 2005) and long-term (Ekonomou, 2010) forecasting are satisfied with ANN. For industrial buildings with high fluctuating energy consumption, ANN can also be used to predict energy demand (Azadeh, et al., 2008). Neural Networks based model is one of the most popular machine learning method for predicting energy consumption in buildings. An ANN model has been proposed for predicting the Predicted Mean Vote (PMV) index, which is used in MPC of the HVAC system to achieve thermal comfort and energy saving (Ferreira, et al., 2012).

2.4.2   Random Forest

Random Forest (RF) is one of the algorithms in the regression tree (RT) family. RF model is improved from bagging regression tree (Prasad, et al., 2006), consisting of an ensemble of decision trees. Each tree can produce a response and presented as a set of predictors which uses the conditional mean of the observations on the resulting leaf. The forest can form an overall better model by averaging the variance from trees (Voyant, et al., 2017). RF is an attractive tool for predicting the energy consumption of HVAC systems because it can (i) involve the interaction between predictors; (ii) learn simple and complex problems based on its ensemble learning theory; (iii) save time without much fine-tuning of parameters compared with other machine learning

techniques such as ANN (Ahmad, 2017). Compared with a single classifier or regression tree, RF methods offer a significant improvement in regression (Siroky, 2009). Compared with ANN, it is faster to train and tune. The training time for RF is 17.8 seconds for one job, while it is 11.1 minutes for ANN (Ahmad, 2017).

RF was proposed to model the nonlinear function of the HVAC system regarding the dynamics of the room temperature and power consumption, which shows high accuracy with Normalized Root Mean Square Error (NRMSE) of more than 92% for power and more than 96% for room temperature (Smarra, et al., 2018). An RF model was developed to infer the interaction between internal temperatures, external temperatures and loads of the HVAC system. Based on it, a next 24-h energy optimizer for the HVAC system was proposed to dynamically adjust the On/Off of the HVAC and the operational schedule of the mechanical ventilation, achieving 48% and 39% of energy reduction for heating and cooling (Manjarres, et al., 2017).

2.4.3 SVM

SVM is one of the most robust and accurate data mining algorithms (Wu, et al., 2008), and is recognized as a new neural network algorithm for forecasting (Dong, et al., 2005). Compare with ANN using training data to minimize classification error, SVM can find the unique global optimum by minimizing overall error while ANNs may only find a local optimum (Mitchell, et al., 2017). SVM was firstly used by Dong et al. to estimate building load, and the forecasting result showed less error with SVM than that with neural network and genetic programming (Dong, et al., 2005). A review by Zhao and Magoulès showed SVM and ANN performed better, showing more promising results than simulation and statistical methods based on historical data analysis (Zhao & Magoulès, 2010). This method is reported to have remarkable improving accuracy in forecasting energy demand with clustering techniques (e.g. K-means algorithm) for weather data (Chen, et al., 2004) and energy consumption data (Alvarez, et al., 2011). However, one of the drawbacks is that the training process would become extremely slow when training data is very large (e.g. multiple buildings' heating load), and parallel SVM was recommended to improve training speed (Zhao & Magoulès, 2010).

SVM showed better accuracy than ANN in short-term (hourly (Fattaheian-Dehkordi, et al., 2014) and every 5-min (Setiawan, et al., 2009)) electricity demand forecasting. SVM method was used by Fu et al. for hourly lighting, heating, cooling and total energy forecasting, and it outperformed ARIMAX (ARIMA with explanatory variable), Decision Tree, and ANN (Fu, et al., 2015). Pai and Hong found SVM model performed better than ARIMA and general regression neural network (GRNN) model in forecasting electricity load (Pai & Hong, 2005). SVM is also a supervised learning method similar to ANN but

works differently by first classifying non-linear data. It uses the structural minimization principle to achieve a global optimum (Hong, 2013). It is highly effective to solve non-linear problems and is often used to forecast energy consumption with high accuracy (Zhao & Magoulès, 2012).

2.5 Model predictive control

MPC is an acknowledged method for constrained control, which has recently gained significant attention from researchers and industry in controlling and managing the building energy system (Goyal, et al., 2013), appliances, building interaction with a smart grid, and on-site renewable energy system generation (Serale, et al., 2018), as MPC has the ability to exploit disturbances' predictions and prior constraints to predict the energy demand the building. As well, it provides great potentials for the achievement of energy-efficient control, system dynamics and solidity along with time delays (Shaikh, et al., 2014). MPC is considered a promising method to reduce energy use while saving cost, as it can integrate constraint handling, disturbance rejection with energy-saving and dynamic control strategies into controller formulation. It has been adopted for optimising the control of building systems such as cooling, heating, ventilation systems, water heating, thermal storages and window control (Serale, et al., 2018).

Demand strategies are generally applied in building energy consuming or generating systems such as HVAC, lighting and photovoltaic systems. End-customers lacks knowledge and information about efficiently scheduling the operation of building energy systems in response to electricity pricing (Shan, et al., 2016). A home energy management controller (HEMC) was developed by Ahmed et al. (Ahmed, et al., 2016), which is trained by a feed-forward neural network type and Levenberg-Marquardt (LM) algorithm, can reduce energy usage by switching the home appliances through the signals of ON or OFF considering inputs such as room temperature, the temperature of electric water heater, DR signal and total power consumption. A total energy saving of over 3 % per 5 hours can be achieved.

Evolutionary computation method such as Genetic Algorithm (GA) is often used as a search tool to find the optimal weights of the fitness function. It is generally used to optimise other forecasting algorithms rather than using it directly for energy demand forecasting (Ceylan & Ozturk, 2004). GA is often used for the determination of simple energy demand forecasting models, and a large amount of data is normally required for evolution (Foucquier, et al., 2013). It integrates with forecasting models to improve accuracy (Holland, 1992). Togun and Baysec pointed out that the GA-based model performs as well as ANN for forecasting energy demand (Togun & Baysec, 2010). GA was used to optimise statistical models (logarithmic linear, exponential and quadratic) to

predict electricity energy demand (Azadeh, et al., 2007). The result showed GA could improve the accuracy of these regression models for energy forecasting (Azadeh, et al., 2007), and it showed better accuracy than time-series and simulation-based GA models (Azadeh & Tarverdian, 2007). The model combining GA and BP network designed by Yugui can outperform others in training and operation time (Yugui, 2013).

GA can be used as an optimization algorithm to find the optimal global solution. Therefore, it is a popular method for solving optimisation problems (Shariatzadeh, et al., 2016). It has been proved to be an effective method employed in the power system to find an optimal load shedding strategy (Luan, et al., 2002). GA can deal with non-linear objective problems with discrete decision variables. It can also be used to solve multi-objective optimisation problem with acceptable computational cost (Shariatzadeh, et al., 2016).

2.6 Summary

Summertime overheating risk in buildings is of increasing concern. High temperatures have been found in the UK dwellings, exceeding 36°C in bedrooms in Manchester homes and reaching 37.9°C in a London flat. The projection of future climates proposed a rise in average summer temperatures by 1.3°C to 4.6°C in London by the 2050s. Under climate change and global warming, the more frequent heatwave is expected to be a life-threatening risk for people, especially for vulnerable people such as the elderly. The literature looked into the UK housing stock in 2017, discovering the most homes are locating in the South of the UK, where the weather is warmer than North of the UK. Therefore, a large group of people are going to experience overheating risk in the future. Another finding is that most of the dwellings are built with natural ventilation and heating system, lack of mechanical cooling to adapt to climate change and increasing overheating risk. Potential cooling can be achieved by natural ventilation. However, the design of inadequate ventilation, the design of air-tightening new homes, the limit of the window opening in urban cities and the heat island effect are influencing the cooling effect of natural ventilation. Research turns to investigate if the control of natural ventilation can reduce the overheating risk in the current and future climates and how to optimize the building operation if the HVAC systems are installed for better thermal comfort and mitigation of climate change.

Data-driven models are suitable for developing MPC for HVAC systems compared to physic-based and hybrid models. Both physics-based and data-driven models can be types of linear/non-linear, explicit/implicit and static/dynamic models. However, physics-based models are usually continuous and deterministic, while data-driven models are stochastic models which are generally discrete and deterministic (Afram &

Janabi-Sharifi, 2014). The grey box model (or hybrid model) is a combination of the previous two models, requiring specific knowledge of the system and data to form a semi-physical model. Therefore, the difficulty in founding the theoretical structure of the grey model may prevent non-expert users from using this method. Besides, this method still requires data to create the model, and some data such as wall surface temperature is hard to measure through the sensor.

A sensor network is an option to collect time-series data for data-driven approaches for optimisation of the HVAC operation. The sensor network leverages data collected from sensors to embed intelligence in the control of the building and its systems, including but not limited to lighting, heating, ventilation, air conditioning, security and circulation. The integration of data analytics with building management systems (BMS) for intelligent control in a smart building sets itself apart from a conventional building. Therefore, the collection and availability of sensor data from a building monitoring system play a vital role in enabling intelligence in smart buildings. The summary of the reviews for sensors (including electricity, gas, relative humidity and light) is listed below.

1. Conventional mechanical electricity meter has an issue in lagging, resulting in a metering error of more than 2%. Morden electronic ones are more accurate and allow for AMR. However, they are reported to have dirty electricity, which can be harmful to the health of users.

2. Most gas meters use dynamic mechanisms to measure the flow rate, but ultrasonic gas meters provide a non-intrusive way for measuring gas consumption. The accuracy can be affected by the flow rate of the gas and the size of the tube. Ultrasonic meters used in larger size of tube brings higher accuracy, therefore, are widely used for commercial purpose.

3. Most temperature sensors on the market can measure the temperature ranging from -20°C to 70°C, such as NTC thermistors. However, the accuracy can be dramatically affected if the temperature is out of the range between 0°C and 50°C. Expensive temperature sensors such as thin-film RTD sensor can measure a wider range, from -200°C to 800°C.

4. Mosy RH sensors can measure the RH between 0 and 100%, promising an acceptable error from 1% to 5%. However, the accuracy decreases significantly (error increases from 3% to 10%) when surrounding RH is out of the range between 25% and 75%.

5. The light sensors can be used to detect the illuminance level. They can be integrated with artificial lighting to provide lighting control. On/Off and dimming are the main strategies to control artificial lighting, offering an opportunity for

energy saving by maximizing the use of daylight. The recommended minimum illuminance level is 500 lux for occupied space.

This chapter also reviewed machine learning algorithms, including ANN, RF and SVM. ANN is the most widely used machine learning algorithms for forecasting temperature and energy in buildings. RF is an ensemble-based machine learning algorithm, gaining popularity for its feasibility and effectively for predicting hourly electricity consumption. It also outperforms ANN because it is easy to train and tune. SVM is more accurate than ANN in short-term (such as 5-minute timestep) forecasting of electricity demand. Section 2.5 introduced GA, which is capable of finding the global optimal in the optimisation problem.

Chapter 3   Research methodology

This chapter introduces the research methodology used to conduct the research study and achieve the objectives proposed in Chapter 1. The methodology section is divided into two parts. Part I is to design a case study for studying the relationship between the buildings and climates. Part II is to design case studies for developing data-driven based model predictive control (MPC) for HVAC systems of buildings. Section 3.1 introduces the case study method used to investigate the impact of climate change on UK dwellings and the effect of natural ventilation control strategies for reducing overheating risk. Section 3.2 demonstrates that case studies for MPC development and validation. Section 3.3 illustrates the key machine learning algorithms used in developing data-driven models of HVAC systems. Section 3.4 identifies the MPC and optimisation algorithms for optimizing the operational control of the HVAC systems. Section 3.5 outlines the key performance indicators for error analysis for evaluating the forecasting performance and validating the performance of the near real-time MPC.

3.1 Part I: Climate change and the need for building controls

3.1.1   Why EnergyPlus

EnergyPlus, IDA Indoor Climate and Energy (ICE) and TRaNsient SYstem Simulation (TRNSYS) were the most popular building performance simulation tools considering the requirements for accuracy, flexibility, and high-speed dynamic simulation (Mazzeo, et al., 2020). Mazzeo et at. compared them in simulating building thermal behaviour in a solar box. Their study illustrated that EnergyPlus and IDA ICE have higher accuracy than TRNSYS in predicting indoor temperatures, but EnergyPlus presents the highest R2 in predicting glass internal surface temperatures. Although IDA ICE performs better in two of three testing months, EnergyPlus shows the minimum RMSE value in predicting indoor temperatures in another month. However, IDA ICE does not consider the directionality effects of direct solar radiation. Compared with IDA ICE, EnergyPlus and TRNSYS are more sophisticated in modelling the direct solar radiation. EnergyPlus provides more accurate phase change material (PCM) models than TRNSYS's, resulting in a better method for predicting the thermal behaviour of a PCM in external walls (Mazzeo, et al., 2020).

Same as EnergyPlus, some building modelling tools such as eQUEST is also based on DOE-2. However, eQUEST is only built upon DOE-2, which has a limitation in modelling emerging technologies, while EnergyPlus is built on DOE-2 and BLAST that also combines heat balance of BLAST with a generic HVAC system. eQUEST is often used during the design phase due to its ease of use and quick in producing results, but

EnergyPlus can be used in modelling complex systems and produce more accurate results (Rallapalli, 2010).

Computational Fluid Dynamics (CFD) models have been widely used in simulating natural ventilation. Compared with EnergyPlus, CFD is computationally expensive as it divides the modelled space into fine meshes to solve flows in and around the space. In addition, CFD is not suitable for modelling conjugate heat transfer between solid and fluid. The response times to thermal energy for walls and air are hours and seconds, making it stiff (Zhang, et al., 2013).

Finally, EnergyPlus is open source and free to download. It is also the core of some interfaces, including OpenStudio [1] and DesignBuilder[2] , which are user-friendly platforms to support building energy modelling.

3.1.2 Building simulation

The building simulation program EnergyPlus is used to simulate summer indoor operative temperatures in UK dwellings, and CIBSE TM 59 is applied to assess the overheating risk. The most common types of UK dwellings have been modelled, including detached, semi-detached, terraced houses and 4-floor and 10-floor residential flats. The buildings sit in the six most populated and latitude-varying cities (London, Cardiff, Birmingham, Manchester, Edinburgh and Belfast) to represent UK dwellings and consider the change of weather on locations. Four climate scenarios (the current, 2030, 2050 and 2080) are used to show the impact of climate change on UK dwellings, and all scenarios are based on very likely (90% probability) medium emissions from the UKCP09 database.

The effect of natural ventilation on indoor temperature is studied by opening window area of 0%, 20%, 40%, 60% and 80%, while 0% stands for a closed window and 80% represents the maximum window opening regardless of the window frames. Opening windows during a different time in a day can significantly influence the effect of natural ventilation on indoor thermal condition. There are three simple window opening schedules based on opening hours: day-time (8 am to 6 pm), night-time (6 pm to 8 am), and whole-day (0 am to 12 pm). Besides, shades such as blinds and external overhangs are applied to investigate the effect of mitigation strategies on overheating risk.

3.1.3 Building physics

There are 18 types of dwellings of different build years from the 1900s to 2010s in this study, shown in Table 3-1. Most common types of UK residential buildings are tested,

---

[1] https://www.openstudio.net/
[2] https://designbuilder.co.uk/

including detached houses, semi-detached houses, middle terraced houses, low-rise (4-floor) and high-rise (10-floor) flats. Early houses were built around the 1900s, and early flats were built around the 1950s in the UK. Dwellings built in the same period were assumed to use the same materials for the construction. The construction methods and materials of the building facade are different for houses and flats.

Table 3-2 and Table 3-3 show the details of the construction for the external walls of UK houses and flats built in different periods. As different size and orientation of windows can cause different solar heat gain for a room, resulting in a significant change in indoor temperature. Therefore, the size and orientation of windows are assumed to be the same for all the dwellings built in the same period. As occupants spend most of their time in the buildings, especially in the bedrooms and living rooms, operative temperatures in both types of rooms are studied to assess the summer overheating risk in the UK dwellings.

Table 3-1 Dwelling models in types and build years.
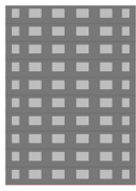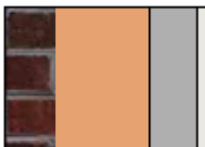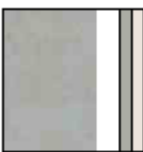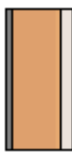
Build year	Building type/Code				
	Detached house	Semi-detached house	Terraced house	Low-rise flat	High-rise flat
1900s	D1900	S1900	T1900	-	-
1950s	D1950	S1950	T1950	L1950	H1950
2000s	D2000	S2000	T2000	L2000	H2000
2010s	D2010	S2010	T2010	L2010	H2010

Table 3-2 Construction of external walls for houses and low-rise flats.

Build year	1900s	1950s	2000s	2010s
Section				
Construction type	Solid wall	Uninsulated cavity brick/block wall	Min wool insulated brick/block cavity wall	XPS polystyrene Insulated brick/block wall
Brick layer (mm)	225	100	105	105
Air gap (mm)	0	50	25	0
Concrete layer (mm)	0	100	100	105
Insulation (mm)	0	0	50	200
Plaster/render (mm)	13	15	13	15
U values (W/m²K)	2.062	1.487	0.350	0.250
Infiltration at 50 pa (m³/ m²hr)	12	16	10	5

Note: The values of parameters are from DesignBuilder

Table 3-3 Construction of external walls for high-rise flats.

Build year	1950s	2000s	2010s
Section			
Construction type	Uninsulated lightweight wall	XPS polystyrene insulated lightweight wall	XPS polystyrene Insulated lightweight wall
Concrete layer (mm)	200	0	0
Air gap (mm)	25	0	0
Steel/ Metallic cladding (mm)	1	10	10
Insulation (mm)	0	89.7	121.9
Plaster/render (mm)	10	13	15
U values (W/m2K)	0.708	0.347	0.263
Infiltration at 50 pa (m3/ m2hr)	16	10	5

Note: The values of parameters are from DesignBuilder

3.1.4 Locations and climates

The UK has a wide range in its latitude, from 49° to 61° N, resulting in various regional climates. For example, the summer is cool in the North West and North East but is warm in South East and South West. Due to the enormous variations in regional climates, the six most populated cities at different latitudes are chosen for the locations of the dwellings in the UK: London, Cardiff, Birmingham, Manchester, Edinburgh and Belfast.

Climate changes significantly over time, and the temperature is likely to increase in future. The regional weather data for these cities was downloaded from the UK climate projections (UKCP) database, based on the medium emission scenario (three emission scenarios: low, medium and high) at the 90% probability level. Climates in periods: the current, 2030s, 2050s and 2080s, are adopted as the weather inputs for the dwellings to investigate the performance of the buildings in the present time and future.

3.1.5 Heat gains

The room temperature can increase through gaining heat from surrounding people, lighting or other electrical equipment. Each model has considered heat gain from people, lighting and other electrical equipment. As shown in Figure 3-1, all of the three elements can contribute heat to the indoor environment. Electrical equipment such as lights is assumed to be switched off in the bedrooms during the night. People become the main contributor to heat gain. While in the morning and evening, lights and other electrical equipment consume a significant amount of energy during the occupancy.
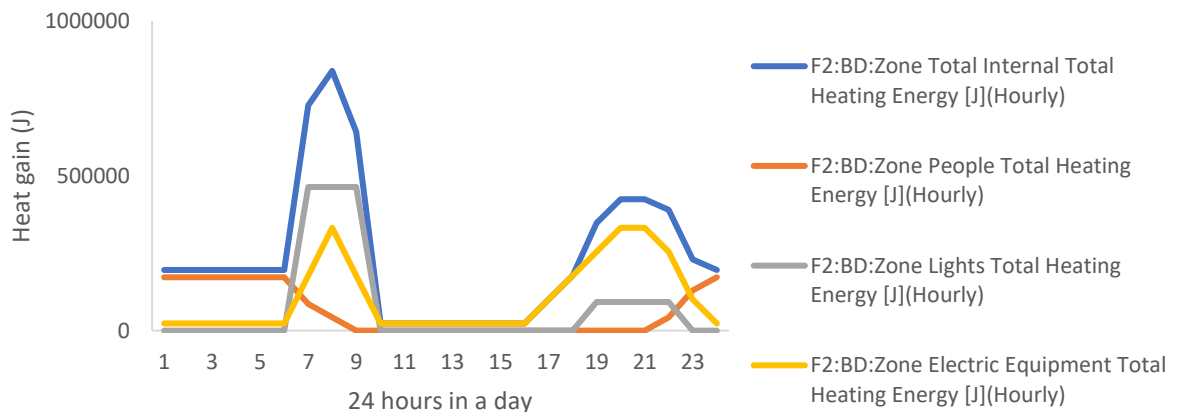
Figure 3-1 The 24-hour bedroom heat gain pattern for a 1900s detached dwelling. The average 24-hour heat gain pattern for an example detached dwelling presents the hours of bedroom heat gains generated in a day.

3.1.6 Opening and shading strategies

Across the UK, most dwellings use natural ventilation through the window opening for cooling in the summer. This study aims to find out whether the current dwellings can still provide thermal comfort in the present and future climates. Different window schedules (daytime, night-time and all-day) and opening areas (0%, 20%, 40%, 60% and 80%) are applied to assess the overheating risks of dwellings in the UK dwellings in the current and future climates. Shading strategies such as blinds and overhangs installation are applied as extra measures to reduce overheating hours. The overheating assessment of the UK dwellings is carried out in the following scenarios.

- All-day closed window

- Daytime (8 am to 6 pm) opening

- Night-time (6 pm to 8 am) opening

- All-day window opening

- Night-time window open plus daytime blinds

- Night-time window open plus daytime blinds and 0.25 m overhangs

- Night-time window open plus daytime blinds and 0.5 m overhangs

Scenario 1 aims to investigate the effects of non-ventilation on overheating hours. Scenario 2 evaluates the effect of daytime ventilation. In this scenario, the windows are assumed to be opened during the day but closed during the night for the occupants' sleeping. Scenario 3 models the effect of night-time ventilation. The occupants are assumed to leave windows opened for cooling when they stayed or slept in the bedrooms during the night but close windows for safety during the day. Scenario 4 examines the maximum ventilation for all-day cooling. Scenarios 5 applies blinds during the day to reduce solar heat gain and leaves the window opened during the night. Scenario 6 and 7 add 0.25 m and 0.5 m overhangs above the windows for additional shading.

3.1.7 Overheating criteria

CIBSE TM59 overheating assessment method is used to evaluate the overheating risk for UK dwellings. There are two criteria (Criteria I and II) for the assessment of overheating risk for living rooms and bedrooms, respectively. Criteria II for bedrooms is more strict as it considers the temperature of 26°C as the threshold of thermal comfort during the sleeping time. Indoor temperature exceeding 26 °C is considered as high temperature for bedrooms at night, leading to sleep deprivation.
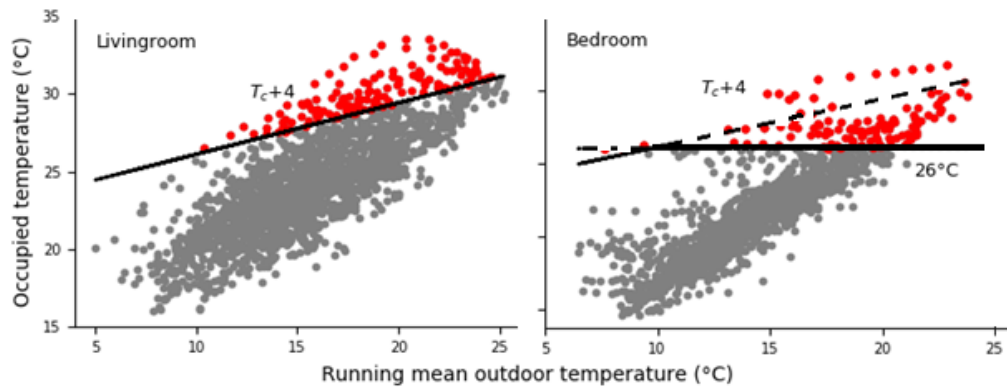
Figure 3-2 Overheating assessment for the living room and bedroom.

Figure 3-2 illustrates an example of using the two criteria, implementing them to evaluate the overheating hours for the living rooms and bedrooms in a semi-detached house in Birmingham in the 2030s. The scatter plots show the indoor temperature is positively correlated to outdoor temperature for both room types, with most of the points falling in a stripe shape (highly correlated when all points in a straight line). The bedroom's temperature is more correlated to outdoor temperature as the points form a narrow stripe shape.

Criterion I is to assess overheating risk for living rooms, kitchens and bedrooms. This criterion defines overheating risk based on the adaptive thermal comfort standard BS EN 15251. The number of the occupied hours for the indoor temperature exceeding the maximum allowable temperature by at least 1 K ($\Delta T \geq 1$) shall be no more than 3% of occupied hours from 1st May to 30th September (3% of occupied hours in summer between 10 am and 10 pm is 59 hours). The ΔT is the difference between the actual indoor operative temperature (T_{op}) and the maximum allowable temperature (T_{max}).

$$\Delta T = T_{op} - T_{max} \tag{1}$$

T_{max}, according to BS EN 15251, is equal to ($T_c + 3$) for the existing buildings.

Criteria II is an additional criterion for assessing the overheating risk in bedrooms only. It is more strict as it also considers the temperature of 26°C as the threshold of thermal comfort during sleeping time. The overheating risk is defined if the number of hours for a bedroom occupied temperature exceeding 26°C more than 1% of annual occupied hours (33 hours). The TM59 overheating criteria require the building to pass both Criteria I and II for relevant rooms to prevent overheating risk.

3.1.8 Adaptive thermal comfort standard BS EN15251

BS EN15251 is an adaptive thermal comfort standard designed to set up limits for the indoor thermal comfort of the occupants. The neutral comfort temperature (T_c) is calculated from the running mean outdoor temperature (T_{rm}), while T_{rm} is calculated from the previous 7-day daily mean outdoor temperature (T_{od}).

$$T_c = 0.33T_{rm} + 18.8$$

<div align="right">(2)</div>

$$T_{rm} = (1 - 0.8) * \{T_{od-1} + 0.8 * T_{od-2} + 0.8^2 * T_{od-3} + 0.8^3 * T_{od-4}$$
$$+ 0.8^4 * T_{od-5} + 0.8^5 * T_{od-6} + 0.8^6 * T_{od-7}\}$$

<div align="right">(3)</div>

The adaptive thermal comfort criteria suggest acceptable comfort temperatures range of 2°C and 3°C from neutral comfort temperature (T_c) for 90% and 80% acceptability limits. The levels of acceptability or categories indicate indoor temperature should be within the suggested acceptable range to achieve different levels of expectation for specified occupants or buildings. In existing dwellings, $T_c + 3$ is for the upper limit of the comfort range.

3.1.9 Heatwave

Summer heatwave days can be calculated using the weather data from climate projection and the definition of a heatwave from the Met Office. More than three consecutive days with daily maximum temperature over temperature thresholds is considered a heatwave. The temperature threshold varies from 25 to 28 in different locations in the UK: 25 °C in parts of Scotland, North and South-West England, Wales, Northern Ireland; 26 °C in Dorset and the area from Lincolnshire to Cheshire; 27 °C in East Anglia, parts of the middle land and much of the Home Countries; and 28 °C in London.

3.1.10 Data processing

Previous sections have explained why EnergyPlus has been adopted for modelling the dwellings. This section is to introduce the main steps and tools used on how to create 3D models, modify inputs and plot outputs. Figure 3-3 shows the data flow from model development to the plotting of the results. DesignBuilder is used to create 3D models for dwellings. As a user-friendly graphical interface based on EnergyPlus, the models can be exported into input files for EnergyPlus to recognize, edit and run for simulation. There is massive simulation work needed for this study (18 types of dwellings, 6 locations, 4 weather files from the 1990s to the 2080s, 5 openings from 0% to 80%, 3 opening strategies and 3 shading strategies). The programming tool Python is used to automate the process. It can connect with EnergyPlus to edit the input files, run the simulation, and save the results on the local computer. The two most useful packages in Python are used to deal with data: Pandas and Matplotlib. Pandas is used to open and edit EnergyPlus input files, while Matplotlib is used to generate plots for visualization.
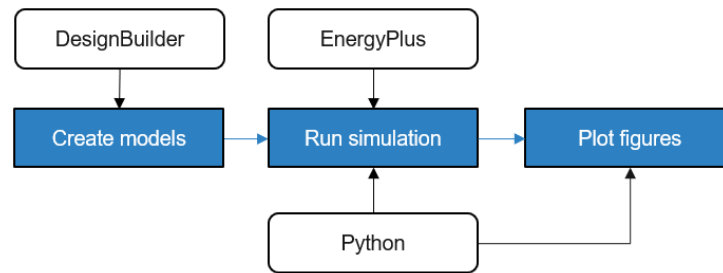
Figure 3-3 Flow chart of the main steps and used tools.

3.2 Part II: Case studies for MPC

3.2.1 Residential building

One of the case study buildings is a mid-terraced house locating in Cardiff, shown in Figure 3-4. The sensing and controllable devices have been installed in the building. A Photovoltaics (PV) system has been installed in the roof of the buildings. Besides, sensors and other devices have been installed to help the building become 'smart'.
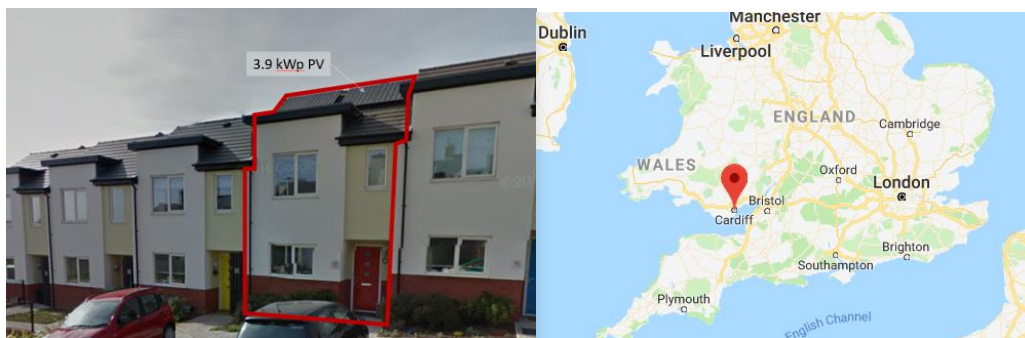
Figure 3-4 The overview and geometry of the residential building in Cardiff site.

Control systems

The test house uses a Nest thermostat[3] as the heating controller for maintaining indoor air temperature within the range of occupants' preference. Nest operates on two predominant modes: operating schedule and home/away. The occupants use the Nest thermostat to set the operating schedule for the conventional heating system. Nest can also detect the presence of occupants using the on-device presence sensor and occupants' location with a detached sensor. This helps to determine if the boiler should be turned off while occupants are away from home. Nest is marketed as a learning thermostat as it can automatically learn occupants' daily routines and create an operational schedule to match occupants' preferences. The occupants can also manually change the schedule remotely using a phone or PC via the internet to adjust the room temperature.

---

[3] https://nest.com/uk/thermostats/nest-learning-thermostat/overview/

Weather station

Real-time outdoor environmental conditions and solar radiation are measured using the Davis Vantage Pro[4] weather station. Outdoor dry-blub and dew point temperatures, wind speed and direction, relative humidity, and global horizontal solar radiation are measured on a 5-minute interval. The data are collected via a Davis gateway, stored in the local database and uploaded to Weather Underground so that the station can be accessed online[5]. Like indoor environmental monitoring system, the data from the weather station are not currently utilised in the BMS but rather used for validating forecasting algorithms.

Monitoring system

In the residential building, the building monitoring system has been developed by installing sensors (including data loggers and data receivers) to detect environmental data (including temperature, humidity, etc.) and energy meters (generation and consumption). The temperature and humidity sensors have been deployed in the living room, kitchen, hallway and bedrooms. In addition, the meter sensors have been installed in the electricity meter, gas meter and PV system. All the data will be wirelessly transmitted to the database on a computer to report historical environmental data and energy consumption. Moreover, the weather station has been installed to receive the weather data (e.g. outdoor temperature). All the data has been stored for analysis. Other than these sensors, more smart devices (smart as they have simple optimiser or can be connected to the internet for control), such as Nest thermostat, have been installed to simply control heating system bases on set point and room temperature. The devices and sensors installed in the house have been illustrated in Figure 3-5.
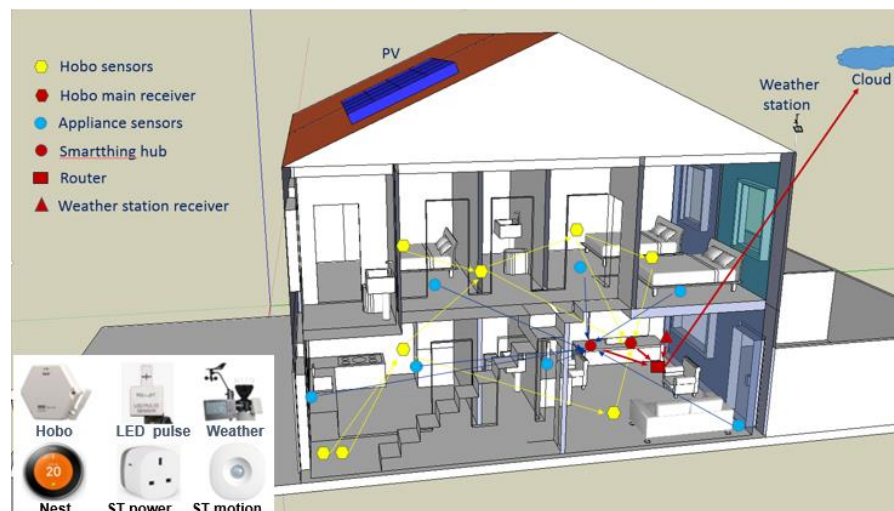
Figure 3-5 Devices and sensors installed in the house.

---

[4] https://www.davisnet.com/solution/vantage-pro2/
[5] https://www.wunderground.com/personal-weather-station/dashboard?ID=IPENARTH2
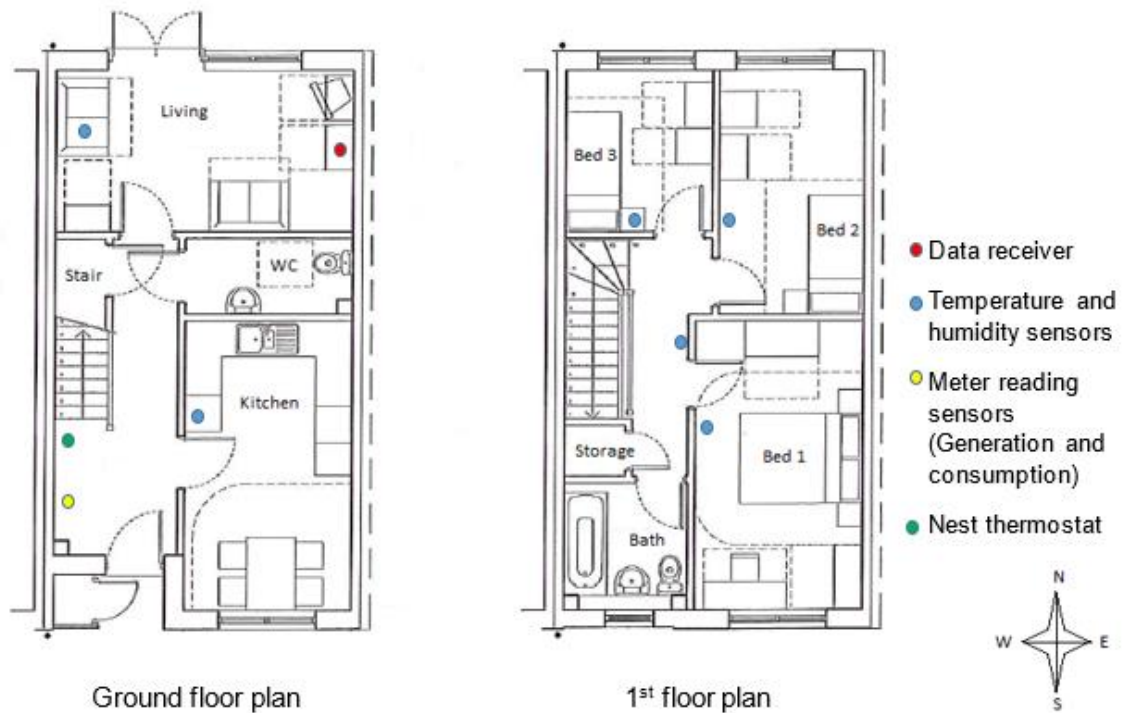
Figure 3-6 Sensor locations in the house.

The building environmental monitoring system uses Onset Hobo ZW-series wireless[6] nodes. The nodes are set up as an independent wireless mesh network, comprising a gateway (ground floor) and a router (first floor). Each node has stand-by battery power for fail-safe operation in the event of a power failure. Temperature and humidity are monitored in each of the five rooms at a 1-minute interval. In addition, the import of electricity and the export of excess solar PV generation are monitored using ZW series pulse meters at a 1-minute interval. The data from environmental sensors and electricity meters are collated and stored on an Intel-Atom PC running 24 hours a day. The Hobo monitoring network is independent of the building management system, and its key purpose is to validate the ongoing and future research for assessing how indoor environmental conditions are affected by system operation.

The test house has a sensor network monitoring system that collects data from various sensing and monitoring systems as well as a heating controller (Nest thermostat) that sends control signals to the boiler. Key systems in the test house are described below for a better understanding of the systems and components in the house. The locations of key sensors and controllers are illustrated in Figure 3-6.

---

[6] http://www.onsetcomp.com/products/data-loggers/zw-indoor-wireless-hobo-data-nodes
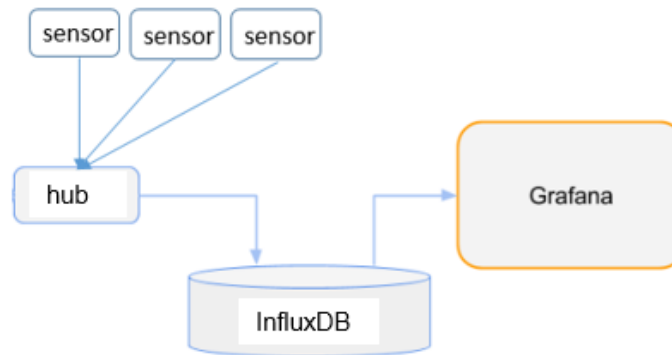
Figure 3-7 InfluxDB[7] database for aggregating sensor monitoring data.

Data collection

Figure 3-7 illustrates the data flow in the energy metering and environmental monitoring system in the residential building. As the hub/data receiver collects data from different sensors, it is necessary to aggregate all the data in one place. InfluxDB, an open-source time-series database, is adopted as a local database to store all the data collected from different sensors. The database is designed for storing time-series data. It is used to collect measurement of the indoor environment (such as indoor temperature), energy consumption and generation. It is installed in the low-cost Raspberry Pi and backed up in a mini Lenovo PC. As shown in Figure 3-7, data is collected by the sensors, sent to hub/data receiver, saved in the InfluxDB database and visualized in Grafana. Grafana is an open-source metrics dashboard and graph editor supporting InfluxDB for near real-time data visualization. The collected data in the database is used as input for training the model for the heating system of the Cardiff residential building.

3.2.2 Commercial building

Another case study building is an office building in Barcelona, Spain. It has six stories and a mezzanine level in the first two floors. The HVAC system of the building is a VRF system consisting of 10 outdoor and 78 indoor units. The exterior view of the building and the layout of the 4th floor with the indoor units are shown in Figure 3-8.

---

[7] www.influxdata.com
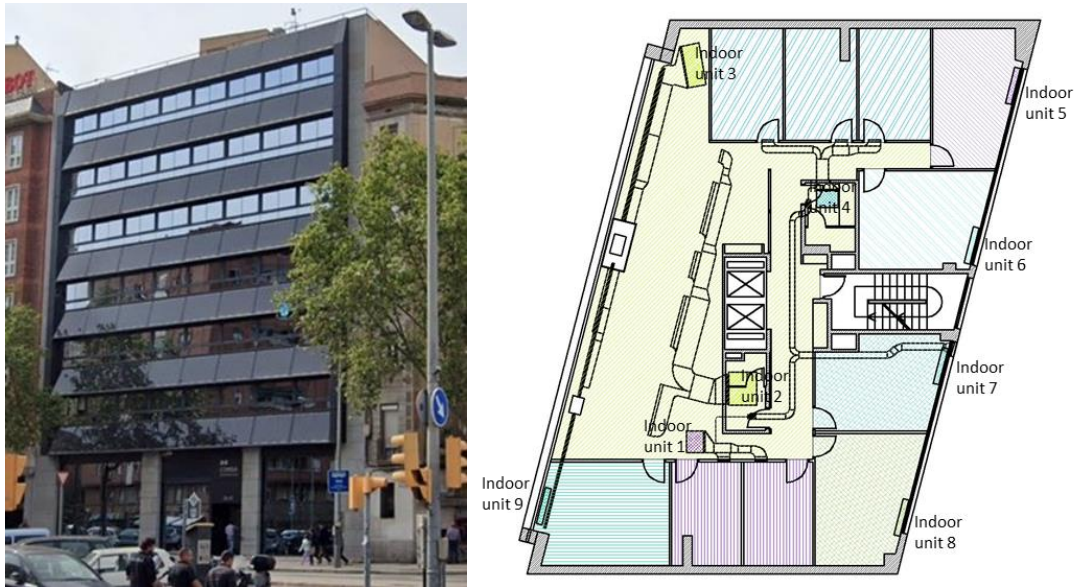
Figure 3-8 The exterior view of the commercial building and the layout of indoor HVAC units. It shows 9 indoor HVAC units with their corresponding supply ducts and served zones in the 4th floor.

HVAC System

The office building has a variable refrigerant volume (VRV) technology-based HVAC system. This means there are no chillers and no boilers. The energy used in the system is electrical energy. The indoor units are located across the building, deliver the thermal energy to the air to condition the space occupied by people, while the external units are located on the rooftop. These units convey thermal energy to or from the indoor units through a refrigerant fluid.

Indoor unit

Each one of the indoor units has its own autonomous built-in controller. This controller communicates with a local panel to provide the building user with the ability to adjust working parameters such as temperature set point. It also communicates with the corresponding outdoor unit so that it can provide the needed thermal energy through the refrigerant fluid to all its associated indoor units. At the office building, all communications happen through two separate Mitsubishi M-Net proprietary Fieldbuses. The first one connects indoor units and control panels on floors 0, mezzanine, 1, and 2 with the corresponding outdoor units on the rooftop. The second one connects indoor units and control panels on floors 3, 4, 5, and 6 with the corresponding outdoor units on the rooftop. In order to provide the HVAC system with the capability of centralized control, two Mitsubishi GB-50 units exist (one for each Fieldbus), which allow supervision and control from a PC with a web browser. With the corresponding outdoor unit so that it can provide the needed thermal energy through the refrigerant fluid to all its associated indoor units. Table 3-4 shows the monitored parameters in the commercial building, including

paramters of indoor units (such as temperature set-points), weather parameters, indoor temperature and humidity, and energy and power meters.

Table 3-4 Monitored energy, environmental, weather and system parameters at the commercial building.

Category	Parameter	Scope
HVAC indoor unit	Mode [cooling/heating] Operation status [on/off] Temperature set-point Unit temperature	Unit
Weather	Humidity outdoor Solar radiation Temperature outdoor	Site
Indoor environment	Humidity indoor Temperature indoor	Floor
Energy	Energy active global Power active global Energy active HVAC	Floor

Data collection
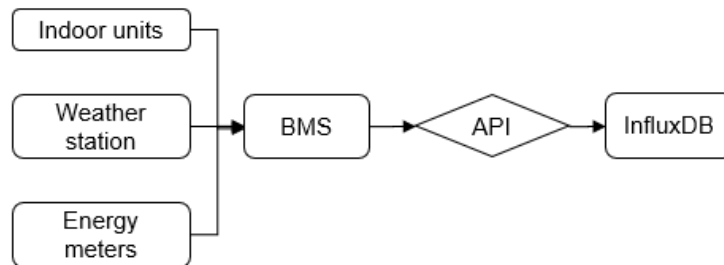
Figure 3-9 Data flow from commercial building to local database

Figure 3-9 shows data collection from the commercial building to the local database. The data flow can be divided into two main parts. The first part is the data collection inside the building. The commercial building has its own building monitoring system (BMS), which is based on Mitsubishi commercial products. Mitsubishi M-Net proprietary Fieldbuses are used to monitor and collect data from indoor units, outdoor units, weather station and energy meters. All the data are collected and saved in local BMS. Inside the BMS, the gateway and API are set up to share the data. The second part of the data collection is to use API and fetch data via the internet. High-resolution data per minute are monitored and collected inside the building, and the Python code is developed to fetch these data via the open API. The most important Python package used to fetch the data is the InfluxDB client, which saves data at the local InfluxDB database

3.3 Model predictive control

3.3.1 Data processing

Data processing includes data collection and data clean. Data collection for the residential and commercial building has been introduced in section 3.2.1 and section 3.2.2, illustrating the devices, tools and techniques used in monitoring and collecting the data. This section introduces the methods in data clean.

The data stored in the database are the raw data collecting from the buildings. These raw data can not be used as inputs for training data-driven models. The data needs to be cleaned before applying them into model training. The methods for cleaning the data from both buildings are nearly same. For example, the most important step in cleaning data is to remove the outliers.

Outliers can be extremely high or low values in the data. Such outliers may come from errors of sensors and should be removed to improve the accuracy of models. Box plot can be used to display data and detect outliers. It uses the interquartile range (IQR) to calculate the difference of data between 25th (Q1) and 75th (Q3) percentiles (IQR = Q3 – Q1). The IQR represents the middle half of the data. Values far away from this IQR can be treated as outliers. Figure 3-10 introduces the rule of 1.5 IQR, which is often used to detect the outliers by finding out if the value is with the range of (Q1-1.5 IQR) and (Q3 + 1.5 IQR). Another method is to find the outliers by simply plotting the data. By plotting data, there may be a few extreme values. Deleting them is a simple method to remove the outliers.

Other than outlier removal, data resampling is also the main task in data cleaning. As original data can be less or equal to per minutes, it needs to be changed into per 15 minutes for use in models. Python package Pandas has been used to read excel files downloaded from InfluxDB and resample the data.
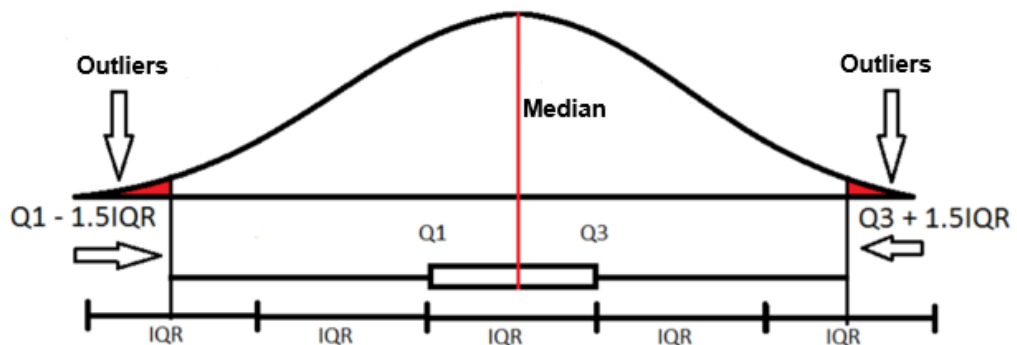
Figure 3-10 The rule of 1.5 IQR for outlier detection

3.3.2 Machine learning

Machine learning algorithms can reveal the non-linear relationship between the input and output variables. In the study, the data-driven approach requires machine learning algorithms for forecasting indoor temperatures, energy demand and power. Machine learning algorithms ANN, RF and SVM are used for developing data-driven forecasting models. Sciket-Learn is a free Python package for machine learning, providing various machine learning algorithms including ANN, RF and SVM.

Model training and testing are two important stages in the development of machine learning-based models. After the data has been cleaned as inputs for training models, data are divided into 70%, 20% and 10% for training, validation and testing stages. The reason to separate them is to prevent the models from overfitting to accurately evaluate the models. Sciket-Learn also has functions in splitting data from training and testing.

In the training stage, the decision of parameters for models is import as different parameters can directly influence the accuracy. Optimisation algorithms are used to optimise the choice of parameters. Keras is a free and open-source Python library for deep learning models. It can be used to optimise the parameters in the training stage. After training and validation, models need to be saved and reloaded for future use, such as testing the accuracy of models with new input data for prediction. Pickle is a Python library used for saving and calling data-driven models.

Table 3-5 presents the main Python tools used in this study. It includes Python libraries not only for data processing and model training but also for API calls, visualisation and optimisation.

Table 3-5 Main Python tools used in the study.

Python library	Description
Pandas	Deal with data, including data processing and analysis
Scikit-Learn	Provide Machine learning algorithms
Keras	Optimize the parameters in the training stage; Also provide machine learning algorithms such as ANN
Influxdb	Access to influxDB database, upload and download data
Platypus	Provide optimization algorithm for MPC
Json	Read json file (monitored data from buildings) from API
Request	Call API via internet
Matplotlib	Plot figures for visualization
Pickle	Save and load data-driven models

3.3.3 Optimization and control
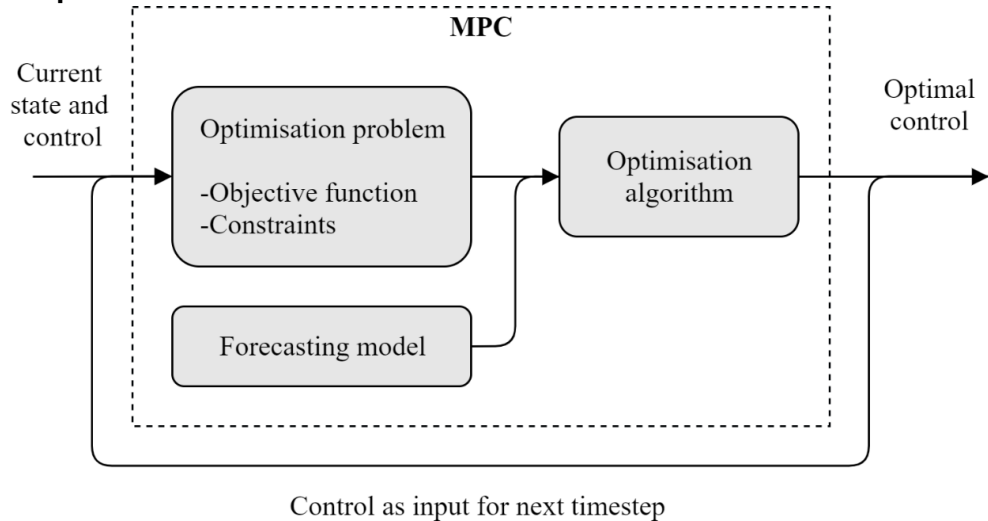
Figure 3-11 A framework of MPC optimisation process.

The forecasting model is used to describe the behaviour of the HVAC system, which uses historical inputs to forecast next time step outputs such as energy consumption. The inputs involve three main types: manipulated, measured and controlled. The manipulated variables are the control signals, such as the next time step temperature set-point given to the indoor unit. The measured variables are from the measurement, including last time step state (e.g., energy or indoor temperature) and outdoor weather parameters. The controlled variables are next time step energy consumption, power and indoor temperature monitored by each indoor unit thermostat. With the forecasting model of the system, it is able to find the impact of the indoor unit control signals on the future change of the energy consumption, power and indoor thermal condition, considering the disturbance of the outdoor weather. Figure 3-11 illustrates a simple structure for the optimisation process of the MPC. Based on the model, the MPC aims to find the optimal solution (e.g., next time step control signals) to achieve maximum energy reduction but satisfy the constrained power and comfort temperature.

Optimisation of the control on indoor units for future time steps is built based on the system forecasting model, objective and constraints. The Non-dominated Sorting Genetic Algorithm II (NSGA-II) has been used as the optimisation method for optimizing the control of the indoor units based on the data-driven model. It is an evolutionary algorithm generating a group of equally optimal solutions, which is called Pareto non-dominated front. The selection of the optimal solution of the Pareto front is depended on the objectives and constraints.

Figure 3-12 illustrates the flow chart of the MPC. The GA algorithm initializes the population based on the identified type and range of the manipulated inputs. With the system model, the output variables such as next time step energy are forecasted. The

forecasted results at time step k can be used as the inputs for time step k+1. Repeatedly, the model can forecast the next N (N = 96 for the day ahead forecasting) time step system states. The sorting process is based on the non-domination criteria of the population. The objective of the NSGA-II algorithm is to improve the adaptive fit of a group of optimal solutions to a Pareto front.
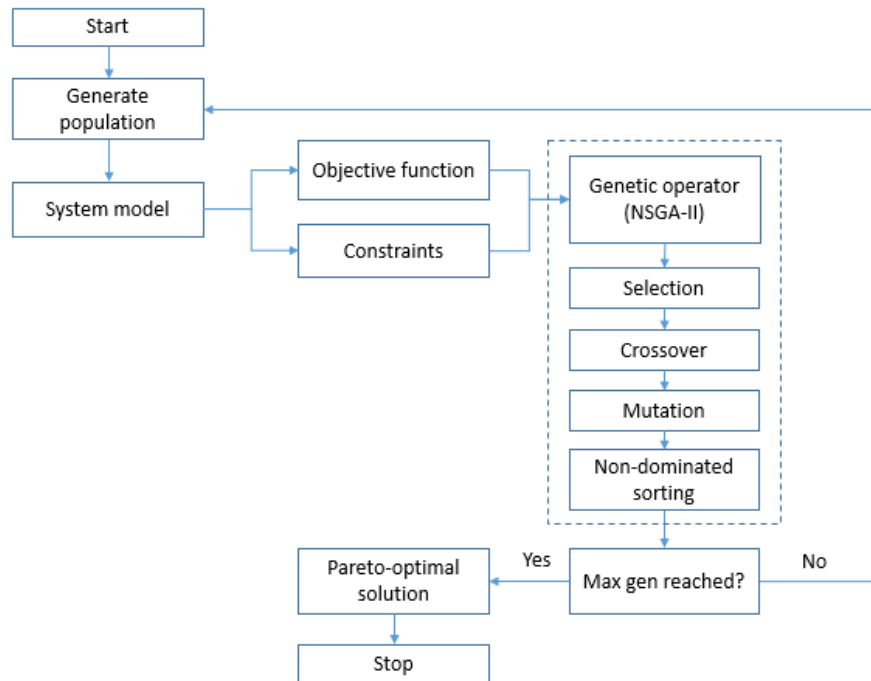
Figure 3-12 MPC flow chart. The chart illustrates the main steps in the MPC using data-driven system model and NSGA-II optimisation method.

3.4 Error analysis

The accuracy of the forecasting model is the key indicator of model performance. There are various methods to evaluate the forecasting model accuracy (Smith & Sincich, 1992). Some of these methods are the mean absolute error (MAE), mean squared error (MSE), root mean square error (RMSE), the mean absolute percentage error (MAPE) and the coefficient of determination (R^2). The forecasting results are assessed by using these widely used accuracy measures as the key performance indicators (KPI)s.

Table 3-6 lists the five most common metrics for measuring the accuracy for the forecast through measuring the errors. They are used as KPIs for evaluating the forecasting results. In Table 3-6, A_t is the actual value at time t; $\overline{A}$ is the average of the actual values; F_t is the forecasted value at time t; n is the total data points and t represents time. KPI is to measure the forecasting error; thus, the smaller error (from number 1 to 4) is the better. R^2 (number 5 of KPIs) value is used to represent the goodness of fit of a model, with values ranging from 0 to 1, while 1 shows the predictions fit the data perfectly.

Table 3-6 KPIs for measuring the load/demand and indoor temperature forecasts.

KPI	Formula/Equation	Number
Mean Absolute Error (MAE)	$$\mathrm{MAE} = \frac{\sum_{t=1}^{n} |A_t - F_t|}{n}$$	1
Mean Squared Error (MSE)	$$\mathrm{MSE} = \frac{\sum_{t=1}^{n}(A_t - F_t)^2}{n}$$	2
Root Mean Square Error (RMSE)	$$\mathrm{RMSE} = \sqrt{\frac{\sum_{t=1}^{n}(A_t - F_t)^2}{n}}$$	3
Mean Absolute Percentage Error (MAPE)	$$\mathrm{MAPE} = \frac{100\%}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$$	4
The coefficient of determination (R^2)	$$R^2 = 1 - \frac{\sum_{t=1}^{n}(A_t - F_t)^2}{\sum_{t=1}^{n}\left(A_t - \overline{A}\right)^2}$$	5

Chapter 4   Need for building controls

This chapter investigates the impact of climate change on buildings to examine if natural ventilation can mitigate overheating risk in future climates. It allows the study to identify the limitation of natural ventilation, the increasing needs of HVAC installation and the importance of operational control for energy systems or controllable elements in buildings for improving the performance of energy and comfort.

4.1 Introduction

Although HVAC systems have been widely used in buildings to cope with weather or achieve better thermal comfort, some buildings do not apply HVAC systems for direct heating or cooling but rely on the control of building elements such as windows. Buildings such as most of the UK dwellings use natural ventilation for cooling in summer because of infrequent high temperatures in the mild climate of the UK. Therefore, the HVAC systems in existing UK dwellings are often designed with heating systems only to meet the heating demand in cold winter. In summer, occupants open windows for natural ventilation cooling.

Increasing summertime ambient temperatures and resulting indoor overheating in dwellings adversely affect occupant thermal comfort; hence productivity, sleep, and health and safety. Might passive strategies such as natural ventilation and shading be sufficient to prevent summertime overheating at present and in the future without resorting to expensive building refurbishments. We investigated the effectiveness of passive measures in mitigating overheating risks in representative UK dwellings under multiple climate change projections.

4.2 Increasing summer heatwave days

Monthly heatwave days have been calculated from May to September in four climate scenarios from the present to the 2080s, shown in Figure 4-1a. It illustrates summer heatwave becomes more and more intense across the UK in the future. Most of the heatwave days fall in July and August since the 2030s. No significant heatwave is found in the present days for all cities, but heatwave days increase dramatically in future periods in southern cities. Especially in London, heatwave days can increase to 10 in the 2030s, 26 in the 2050s and 58 (63% of the summer period) in the 2080s. Edinburgh and Belfast are cities in the North, have much fewer heatwave days than the other cities. Both have less than 10 heatwave days in the 2080s, taking up 15% of the summer. Figure 4-1b presents the mean of summer daily maximum indoor temperature in dwellings' types and age bands from the present to 2080s. Maximum indoor temperature is usually used to compare with the threshold to evaluate the impact of heat on health

(Public Health England (PHE), 2014), and 26 °C was regarded as the most suitable threshold for at-risk groups (Tham, et al., 2020). It is found dwellings in London have the highest indoor temperature. The indoor temperature also reflects the higher outdoor temperature and more heatwave days in London. Terraced houses built in the 1900s have much higher indoor temperature than other types of houses and flats. The figure also shows the maximum outdoor temperature in a heatwave can be more than 30°C, leading to a life-threatening heat-related risk to vulnerable people such as the elderly and infants. Longer duration of the heatwave will result in a consistently increased attributable number of deaths (Cheng, et al., 2018). People will be more vulnerable to heat strain as they are not able to physiologically adapt to a sudden change of weather (Parsons, 2009).
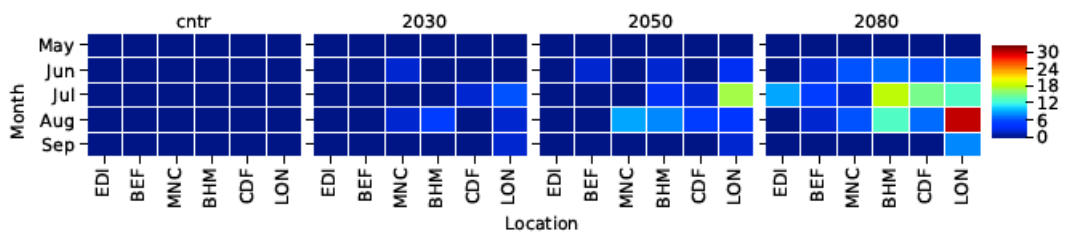
(a)
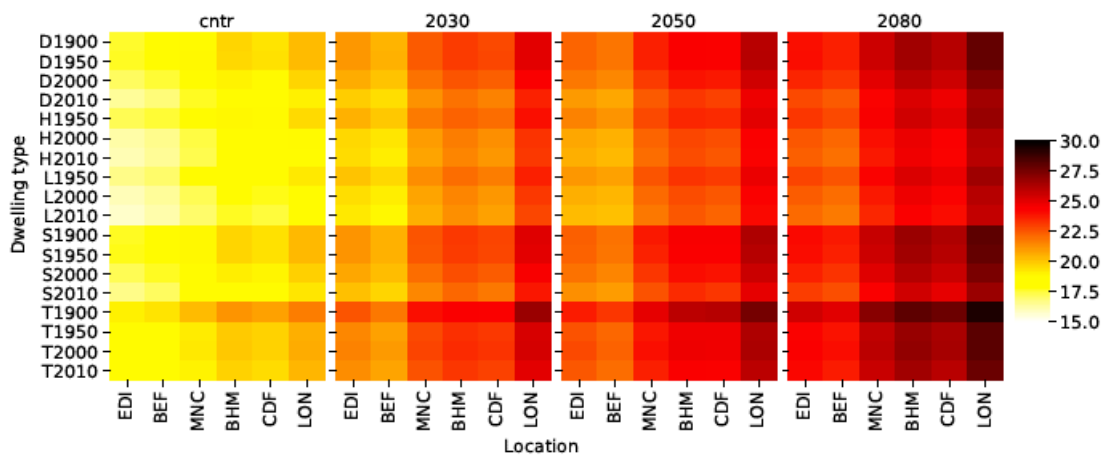
(b)

Figure 4-1 (a) Monthly summer heatwave days from May to September. (b) Summer indoor average daily maximum indoor operative temperatures. Summer days with temperature exceeding 30°C in the daytime and 15°C at night in at least 2 consecutive days are considered as heatwave days. Summer indoor average daily temperatures are represented in colormap for the most often seen UK dwellings under climates from the 1990s to 2080s. All dwellings are applied with night ventilation with 20% of window area and blinds.

4.3 The impact of locations and climates

In the UK, dwellings in the lower latitude cities have a higher overheating risk. Figure 4-2 and Figure 4-3 illustrate the overheating hours in living rooms and bedrooms for dwellings in various locations and periods. The results show the locations of the dwellings have a significant impact on the overheating hours. Dwellings in Southern cities such as London have more overheating hours than those in the North such as Belfast and Edinburgh. Besides, the overheating hours of dwellings increase as the periods increase from the present to the 2080s. Early built dwellings in the 1900s and 1950s have a low overheating risk in the high latitude cities under climate change. In Belfast, terraced houses built in the 1900s and 1950s has overheating hours less than 20 until 2050s, well below the threshold of overheating (59 for living rooms). However, the number will reach 400 in London in the same period.
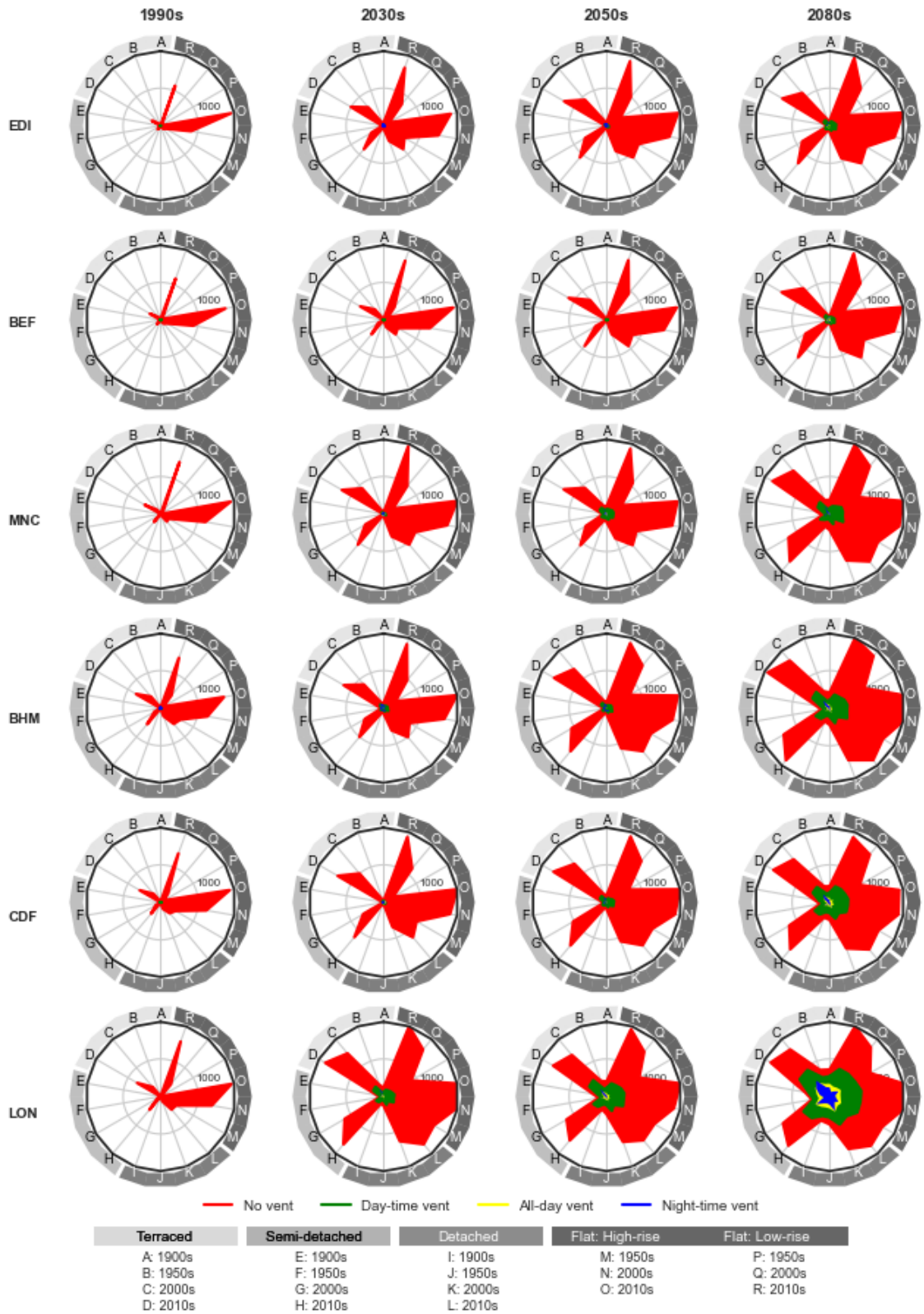
Figure 4-2 Overheating hours in living rooms with different opening duration (Scenario 1, 2, 3 and 4) between the 1990s and 2080s. Each line in the radar plot represents the number of overheating hours for one type of dwellings. The inner circle (1000 hours) and outer circle account for approximately 50% and 100% of the living room occupied hours in summer.
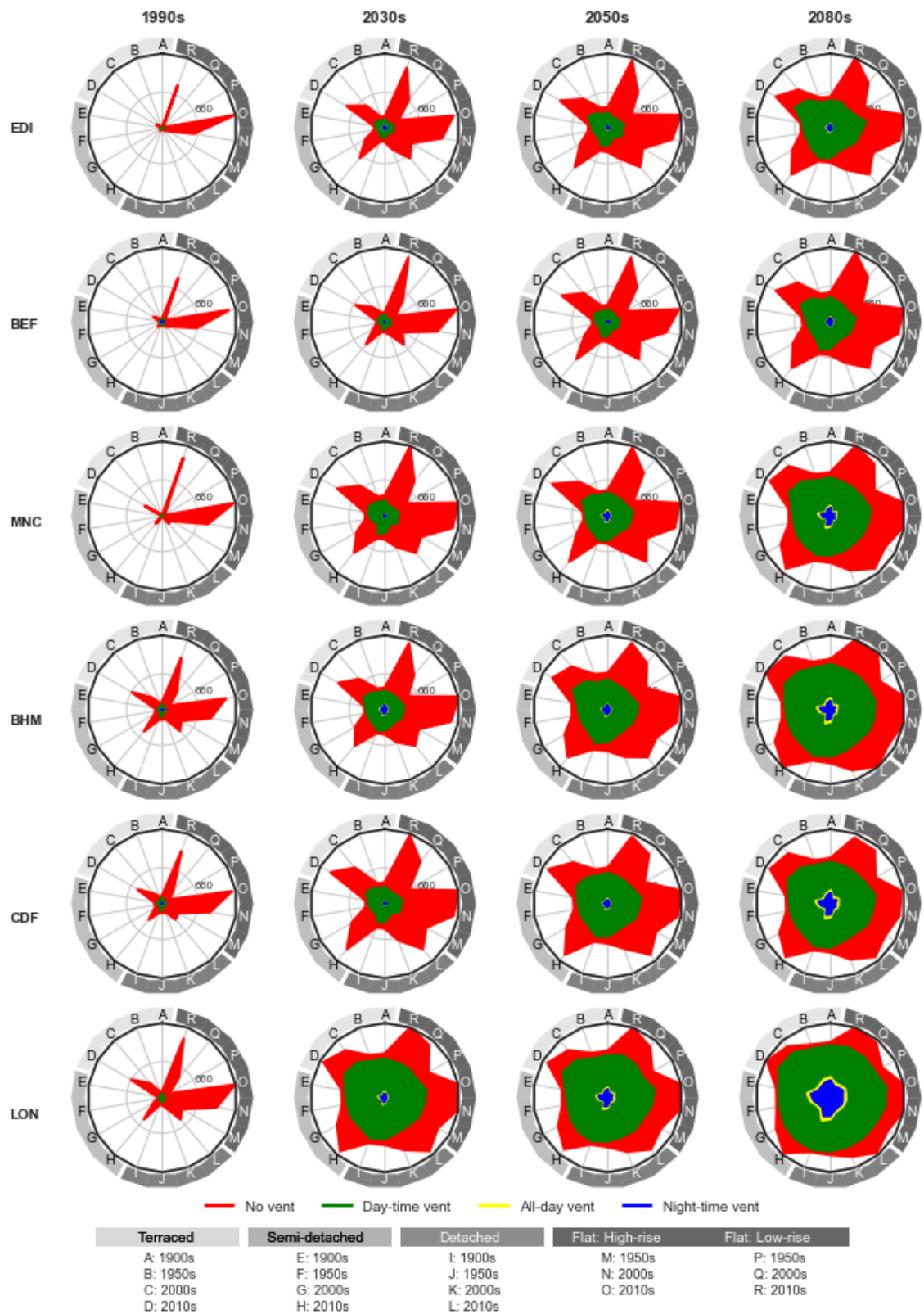
Figure 4-3 Overheating hours in bedrooms with different opening duration (Scenario 1, 2, 3 and 4) between the 1990s and 2080s. Each line in the radar plot represents the number of overheating hours for one type of dwellings. The inner circle (660 hours) and outer circle account for approximately 50% and 100% of a bedroom occupied hours in summer.

4.4 The overheating risk without natural ventilation

Without natural ventilation, dwellings can be easily overheated in summer. Figure 4-2 shows the overheating hours in living rooms in various types of dwellings with opening windows during different periods. When the windows are closed with no natural ventilation, newly built high-rise flats (O) experience more overheating risk, with almost all the hours overheated under the current climates. Followed by the 2000s' high-rise flats (Q) and 2010s' low-rise flats (R), both have overheating hours of more than 1000 hours which are about half of the total summer hours. In future climates from the 2030s to the 2080s, the newly built houses in the 2000s and 2010s also have significant overheating risks as flats. However, houses built in the 1900s and 1950s have the least increment in overheating hours.

More types of dwellings suffer severe overheating risk in the bedrooms. Figure 4-3 shows the overheating hours in bedrooms in various types of dwellings with different opening strategies. Compared to the living rooms, the overheating hours in the bedrooms of the old buildings (the 1900s and 1950s) increase more significantly than in living rooms, shown in Figure 2. Without ventilation, the bedrooms in the recently built dwellings (the 2000s and 2010s) behaved similarly as living rooms. For dwellings built in the 1900s and 1950s, the largest growth of the overheating hours was found between the present and 2030s. Bedrooms in the 1900s Detached and 1950s low-rise flats have a rise of over 700 hours of overheating hours from the present to 2030s, compared to living rooms of an increase less than 150 hours.

Due to the lack of ventilation, dwellings can only rely on air infiltration to condition the indoor environment. Recently built dwellings have higher insulation and airtightness to keep them warmer in the summer. However, older dwellings have their external walls with no insulation layer and bigger gap and crack, resulting in enough air infiltration to effectively cool indoor operative temperature in summer.

4.5 Natural ventilation strategies

Natural ventilation is effective in reducing overheating hours in dwellings. Figure 4-2 shows that overheating hours in the living rooms can be dramatically reduced with natural ventilation in scenario 2 (Daytime ventilation), 3 (Night-time ventilation) and 4 (all-day ventilation). The results show that night-time ventilation has the best performance in cooling almost all types of dwellings. Compared to all-day ventilation, it can reduce more than up to 50% of overheating hours in the warm summer in the 2080s. Both of them can reduce the overheating risks in the living rooms until the 2050s except in London. The 2000s terraced houses experience the severest overheating risks, with around 200 overheating hours in London in the 2050s, increasing up to 483 in the 2080s. Although

daytime ventilation has the weakest performance on reducing the overheating risks, it can still significantly reduce overheating hours by around 50% for recently built dwellings for living rooms in 2080.

Figure 4-3 presents the cooling effect of natural ventilation on bedrooms in scenario 2 (Daytime ventilation), 3 (Night-time ventilation) and 4 (All-day ventilation). For bedrooms with daytime ventilation, significant overheating risks have been found in the dwellings in the present day. Daytime ventilation has low performance in effectively reducing overheating hours for bedrooms when the ambient temperature rises. Among the three ventilation strategies, night ventilation is the best one in mitigating overheating risks, and it can maintain thermal comfort of dwellings in the 2030s, except in London and Birmingham. Dwellings in London suffer more overheating risks, with 12 out of 18 total types of dwellings suffering overheating under the current climate. Among those 12 dwellings, early built dwellings have more overheating hours than newly built dwellings, with 85 hours in the 1900s detached houses compared to 1 hour in the 2010s detached houses.

Criteria II is stricter for assessing the overheating risk of bedrooms, and the allowance overheating hours are fewer than in living rooms. Therefore, overheating hours in bedrooms are higher. The 1900s and 1950s built dwellings have higher thermal mass but lower insulation. With daytime ventilation, the thermal mass absorbs the solar heat and release the heat to the internal building during the night, resulting in more overheating hours.
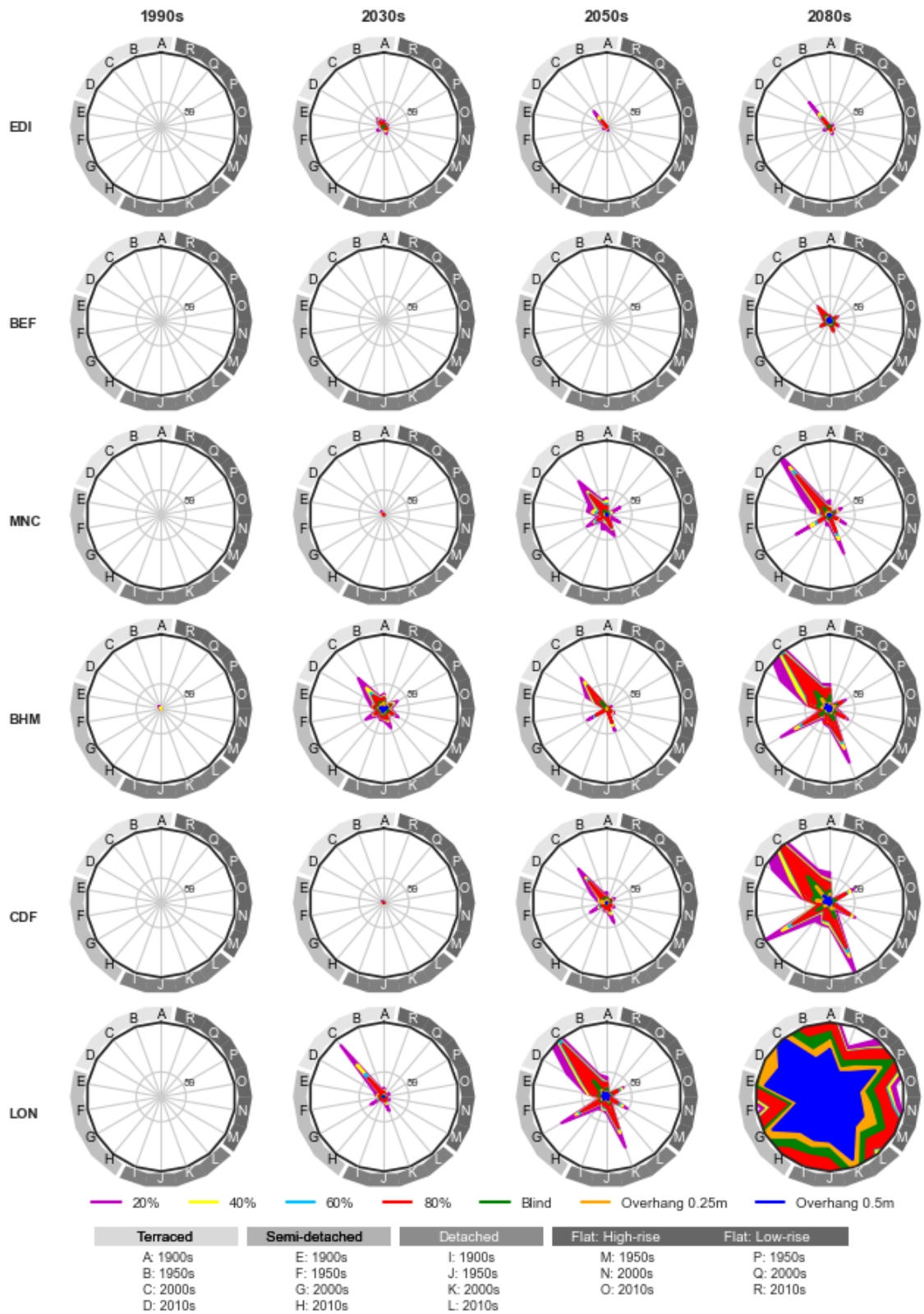
Figure 4-4 Overheating hours in living rooms with night ventilation using different opening area and shading strategies (Scenario 3 (20% - 80% opening), 5, 6, and 7) between the 1990s and 2080s. Each line in the radar plot represents the number of overheating hours for one type of dwellings. The inner circle (59 hours) accounts for the overheating threshold for living rooms in summer.
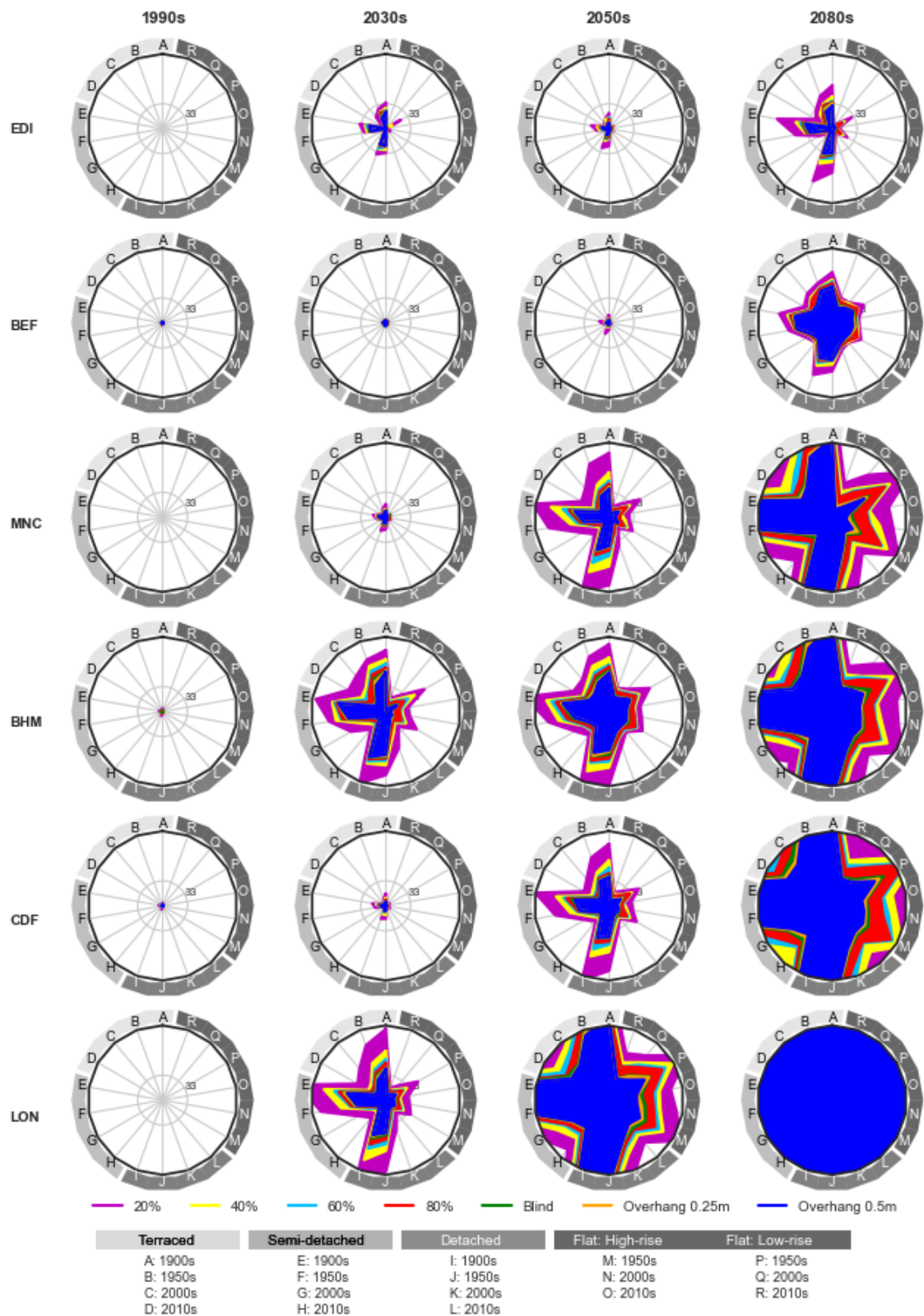
Figure 4-5 Overheating hours in bedrooms with night ventilation using different opening area and shading strategies (Scenario 3 (20% - 80% opening), 5, 6, and 7) between the 1990s and 2080s. Each line in the radar plot represents the number of overheating hours for one type of dwellings. The inner circle (33 hours) accounts for the overheating threshold for bedrooms in summer.

4.6 Opening areas of windows

Night ventilation has been found to outbid daytime and all-day ventilation in mitigating overheating risks in dwellings in the UK. However, there are still significant overheating hours found in the dwellings after the 2030s. It is worth to consider other mitigation measurements such as enlarging the window opening area and adding shading elements. Figure 4-4 illustrates the overheating hours in the living rooms with night ventilation plus overheating mitigation measures, including opening window area from 20% to 80%, and the installation of blinds and overhangs. In the figure, the inner circle indicates 3% of the total summer occupied hours (59 hours) which is the overheating threshold for the living rooms in summer. The results show that increasing the window opening area can significantly reduce the summer overheating risk. In the 2030s, the 2000s terraced houses are the only dwellings likely to suffer overhearing risks in London and Birmingham, and increasing window area to 80% of the window area can prevent overheating of the living rooms in both cities. In the 2050s, another two types of dwellings (the 2000s semi-detached and detached houses) are likely to exceed overheating threshold for living rooms in London. Maximising the window opening area to 80% can still not prevent them from overheating in the summer.

Figure 4-5 illustrates the overheating hours in the bedrooms with night ventilation plus overheating mitigation measures, including opening window area from 20% to 80%, and the installation of blinds and overhangs. The inner-circle in each radar plot indicated 1% of the total summer bedroom occupied hours (which is 33 hours as the overheating threshold for bedrooms). Any type of building with its overheating hours out of the circle is considered to have summer overheating risk in its bedrooms. The results show that opening 20% of the window area for night ventilation can eliminate overheating risk in bedrooms for all dwellings in all the cities in the present days. The results also show that the largest reduction of overheating hours is found when the opening window area increases from 20% to 40%, especially for the 1900s and 1950s dwellings. The reduction is up to 76 hours for the 1900s detached houses in London in the 2080s.

There is a total of 108 buildings considering building types, ages and locations in this study. It is worth to find out the impact of opening areas on the living rooms and bedrooms on all of these buildings as the thresholds of overheating hours are different for them. As mentioned in previous sections, they are 59 and 33 hours for living rooms and bedrooms respectively. It is also worth to looking into the 2030s' weather to see how climate will influence overheating risks for different types of dwelling in the near future. Maximum natural ventilation is assumed with opening windows for 24 hours. By ascending the overheating hours of all 108 dwellings, the impact of opening areas can be illustrated in Figure 4-6 at (a) for living rooms and (b) for bedrooms.
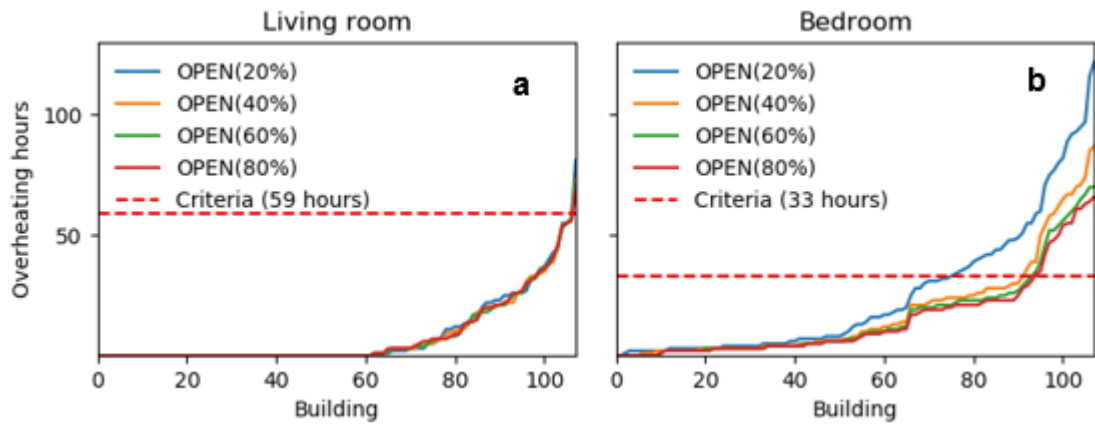
Figure 4-6 Overheating hours at (a) living rooms and (b) bedrooms in ascending orders for all 108 dwellings with opening area from 20% to 80% in the 2030s. Enlarging window areas for openings for natural ventilation is more effective in reducing overheating hours at bedrooms.

The results show that living rooms suffer low overheating risks than bedrooms in the 2030s. Even with 20% of window area opened for natural ventilation, there are only 2 of 108 dwellings overheated. The overheating hours are slightly over the threshold of 59 hours, with a maximum of 81 hours in the terraced house built in 2000 in London. The results also indicate that increasing window areas has almost no change in overheating hours in living rooms. Contrarily, bedrooms are easier to be overheated in the 2030s. The reason could be the stricter overheating threshold of 33 hours compared to 59 hours for living rooms. Additionally, room temperatures over 26 °C are also considered overheating, as high temperatures can affect sleep. The results also show that increasing the window opening area can significantly reduce the overheating hours for bedrooms. The number of overheated dwellings is reduced from 32 to 14 by increasing the opening area from 20% to 80%.

4.7 Blinds and overhangs

For living rooms, night ventilation with a maximum opening area can not prevent all dwellings from overheating risks in the 2050s, but adding blinds can effectively reduce overheating hours to under 20 for them. As shown in Figure 4-4, the 2000s terraced houses are the most vulnerable dwellings to overheating, and they can be overheated after the 2030s in certain cities even with 80% of the window area opened for night ventilation. However, the blinds can maintain overheating hours under the threshold for most of the dwellings in the 2080s. Based on night ventilation and blinds, the installation of 0.25 m overhangs can prevent dwellings from overheating risk in Cardiff in the 2080s. However, it is not enough for dwellings in London in the 2080s, even with 0.5 m overhangs.

For bedrooms, even with maximum openable area, blind and 0.5m overhang, most of the dwelling get overheated in the 2050s and 2080s. As shown in Figure 4-5, houses

75

built in the 1900s and 1950s are easier to be overheated as the ambient temperature increase in future. They can easily exceed the overheating threshold in London and Birmingham in the 2030s, in Manchester and Cardiff in the 2050s, in Edinburgh and Belfast in the 2080s. Dwellings built in the 2010s have the least overheating risk in the bedrooms in summer, keeping indoor thermal comfort in Edinburgh and Belfast in the 2080s. The results also present that the installation of blinds has better performance for flats, reducing 34 overheating hours in the 1950s flats in the 2080s. Unlike in living rooms, shading such as blind and overhang can only slightly reduce overheating hours for bedrooms because occupied hours in bedrooms are almost at night. Opening window area from 20% to 80% can maximumly reduce 123 overheating hours for the 1900s detached houses in the 2080s, but blinds and 0.5 m overhangs can only reduce 34 and 13 for the 1950s flats in the 2080s.

4.8 Overheating hours under future climates

Climate change can significantly influence overheating hours as UK dwellings are often naturally ventilated and vulnerable to adapt to future climates. The TM59 describes that indoor temperatures are generally used to assess the overheating hours for dwellings, with different overheating thresholds of 59 and 33 hours for living rooms and bedrooms, respectively. It is worth to investigate the impact of future climates on living rooms and bedrooms at all 108 dwellings regardless of building types, build years and locations. The opening window area is assumed to be 20% rather than 60% or 80% because small window opening areas are often recommended considering safety to stop children from falling. Windows are assumed to open 24 hours to investigate if the maximum natural ventilation can eliminate overheating risks. By ascending the overheating hours of all 108 dwellings, Figure 4-7 (a) and (b) illustrate the impact of future climates on living rooms and bedrooms' overheating hours.

The results show that bedrooms suffer more overheating risk than living rooms in future climates. Although no overheating hour for both rooms is found in the 1990s' weather, the numbers of overheated living rooms are 2, 18 and 48 out of a total of 108 dwellings in the 2030s, 2050s and 2080s. Compared with living rooms, the numbers of overheated bedrooms are 32, 65 and 96 out of a total of 108 dwellings under the same climates. In the near future in the 2030s, about 30% of dwellings have overheating risk in their bedrooms but less than 2% in their living rooms. The proportions of dwellings with overheated living rooms and bedrooms are abou17% and 60% in the 2050s, 45% and 89% in the 2080s. Bedrooms suffer more overheating risks than living rooms, and more concerns should be paid to adapt them to future climate change.
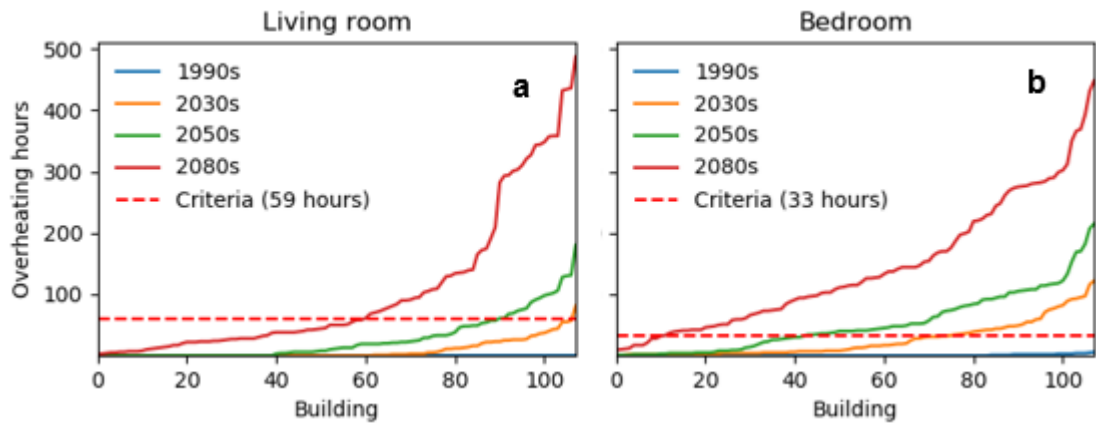
Figure 4-7 Overheating hours at (a) living rooms and (b) bedrooms in ascending orders for all 108 dwellings with an opening area of 20% from the 1900s to 2080s. The 59 and 33 are the thresholds of overheating hours for living rooms and bedrooms. The bedrooms of buildings are easier to suffer overheating risks compared to living rooms in the same year.

4.9 Discussion

The effect of natural ventilation on UK dwellings was investigated by simulating the overheating hours under different scenarios. However, these scenarios were based on the assumptions of natural ventilation and shading strategies which should be explained in this section.

Overall, night ventilation has the best performance in reducing overheating risks compared to daytime ventilation and all-day ventilation. The average overheating hours of all dwellings are 21 with night ventilation, 35 with all-day ventilation and 125 with daytime ventilation. This number is significantly increased to 784 without natural ventilation by closing the windows. The results show that natural ventilation could dramatically reduce overheating risks. Kubota et al. (Kubota, et al., 2009) pointed out night ventilation leads to lower indoor temperature than daytime and all-day ventilation but can result in high humidity in the hot-humid climate of Malaysia. If the climate of the UK develops as high as the climate projection, humidity control is needed when night ventilation is applied. Michael et al. (Michael, et al., 2017) also suggested night ventilation is the most effective compared to the other two strategies as it capitalises on the thermal inertia of the building envelope. However, no ventilation and all-day ventilation are also important for naturally ventilated dwellings as occupants are inactive in changing the ventilation status (Lai, et al., 2018).

Overheating hours can be significantly affected by floor insulation. There are more overheating hours for living rooms in the 2000s/2010s built houses, but in flats, the overheating hours decrease as the build year increases. However, the same pattern (overheating hours decrease in new dwellings) occurs for bedrooms in all types of dwellings, and the 1900s built houses has the highest overheating risk. The performance

of the living rooms and bedrooms are similar for flats but totally different for houses. To find out what makes the big difference for overheating hours in the living room and bedroom of the houses, we analysed the hourly temperature in the living room and bedroom in a semi-detached house. Figure 4-8 presents the hourly temperature of the living room and bedroom with full-day natural ventilation in a semi-detached house in the summer design day in London in the 2080s. In the living room, the 2000s built houses did have the highest indoor operative temperature. Comparing the floor conduction heat transfer rate per area between the living room and bedroom, we found out the floor cumulated conduction rate per area is much lower in living room in the 2000s and 2010s built house due to an additional insulation layer in the ground floor in modern houses. But for bedrooms, they are the same materials with wooden layers. Moreover, the living room and bedroom are on the same floor for flats resulting in a similar performance of the living room and bedroom in flats.
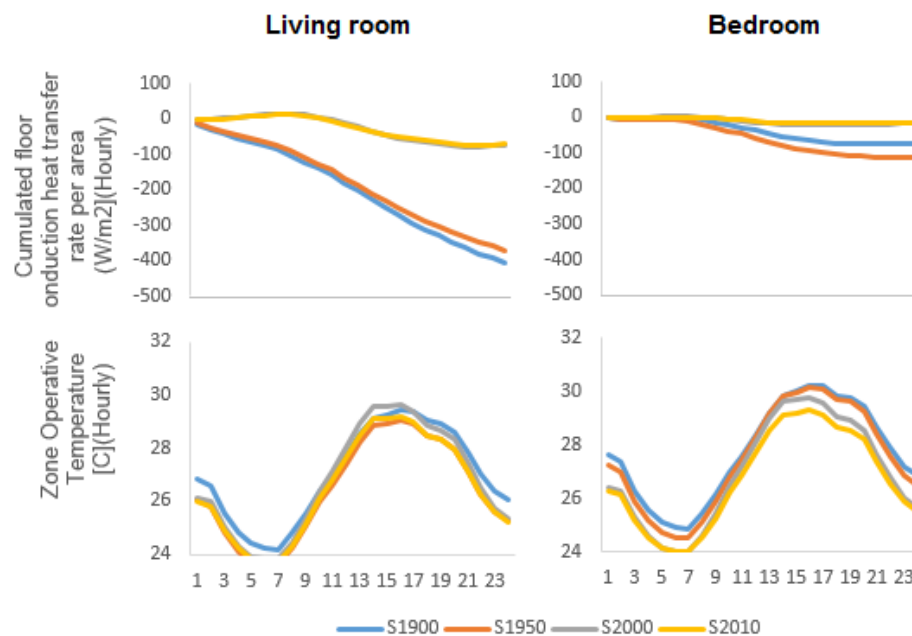
Figure 4-8 Room operative temperatures and floor conduction transfer rate per area in one day of July. Bedrooms' floors have a higher conduction heat transfer rate per area resulting in higher indoor operative temperature.

The regions of UK building stock are shown in Figure 4-9, illustrating the distribution of four main types of dwellings in the UK regions. The housing or household survey data was from 2016-2017 housing statistics, 2011 UK census data, 2016 estimates of the households and dwellings in Scotland. The results showed that most of the detached houses were in the East of England and Scotland while most of the other two types of houses (including semi-detached and terraced houses) located in London, East of England, Yorkshire and the Humber, and Scotland, around 1 million dwellings. Compared with houses, most of the flats were in London, around 1.7 million flats which

are about 1.7 times of dwellings in Edinburgh with the second-largest flats. However, Scotland land area is much larger, which is about 50 times in London. It shows the largest density of flats in London. In Cardiff, Edinburgh and Belfast, fewer houses and flats have been found in these areas compared to London, Birmingham and Manchester. The high density of flats and terraced houses indicated that a large population worked and lived in London, resulting in more terraced houses, low-rise and high-rise apartments. However, the study of overheating risk found that ground-floor living rooms in terraced houses could experience overheating risk in the 2030s in London without adequate window opening (needs 80% of total window area) for natural ventilation. Compared with the living rooms, bedrooms are easier to be overheated in the old dwellings, especially in house-type dwellings. However, all dwellings could have suffered a severe overheating risk since the 2050s in London and Birmingham. It is recommended more dwellings could be built in Scotland as the future climate in Scotland is comparably cooler than in England. The results also showed overheating in Cardiff is much lower than in London, although they are slightly different in latitudes. Urban Heat Island effect is concerned to produce microclimate in the cities resulting in higher temperature in London. Natural ventilation and wind speed could also be affected by surrounding tall buildings. Air quality is another concern as occupants might keep windows closed due to air quality and dust.
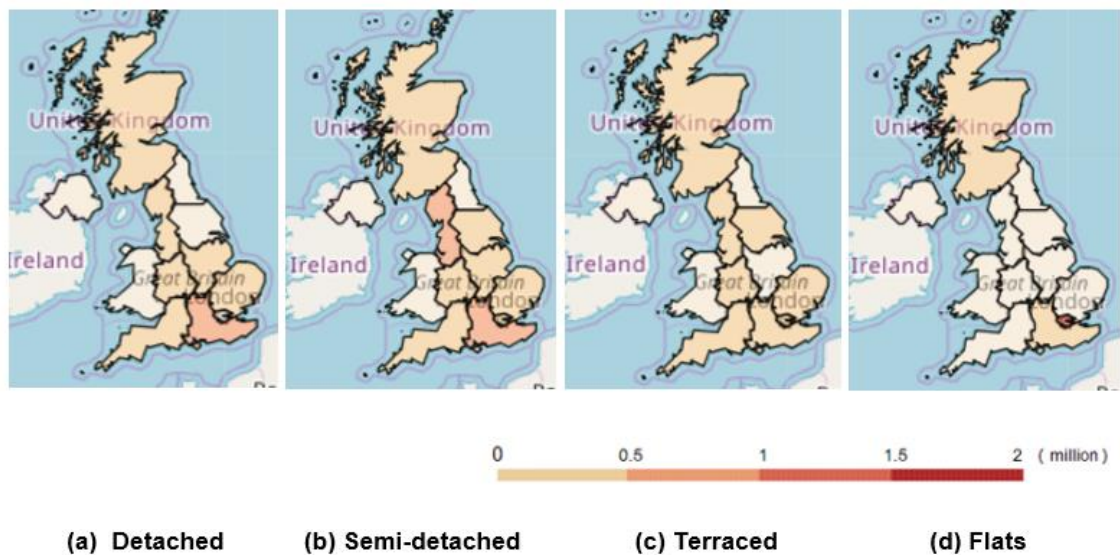
Figure 4-9 The households by dwelling types in the UK regions. The plot shows the distribution of the four types of dwellings in the UK regions. The deeper colour, the more dwellings are in this region.

Children and elderly occupants are vulnerable to overheating risk. Figure 4-10 illustrates the population estimates for different age groups (age 0-15, 16-64, 65 and over) in the UK regions in mid-2017. The largest number of children (age 0-15) is in London and South England. The largest number of elderly (65 and over) is in the South East. A

significant number of children and elderly have also been found in West England, North West Yorkshire and the East of England. Compared with children and elderly, the number of people aged between 16 and 64 is up to 5 times of them in each region, and most of them are in England (especially London and South West) and Scotland. Children and the elderly are more vulnerable to overheating risks in the summer. As shown in Figure 4-10, South East and London are lower latitude regions in the UK, but with the largest number of children and elderly people. Both regions are expected to experience high overheating risks in future. Elderly people aged 65 and over account for 19.4% of the total population in London, which is a big number as London urban area has a population of around 8.8 million (ONS, 2018). As the biggest city in the UK, London has a population of more than 8 million according to 2011's census, while 34.5% of them are the elderly and young children under 15. The number of vulnerable occupants in London is seven times more than the second largest population in Birmingham. Belfast and Cardiff have the least of the vulnerable occupants, which is about 117 thousand. Those vulnerable occupants in London are expected to experience the highest overheating risk in the 2030s. A deep concern should be paid to the large group of vulnerable people.
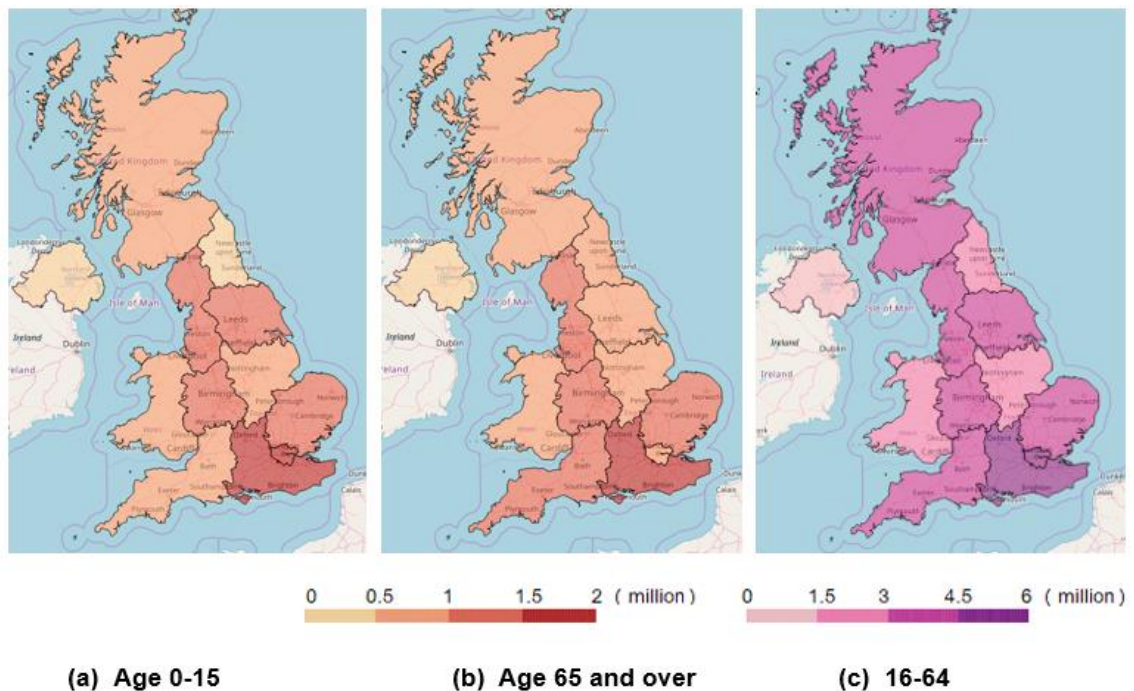
Figure 4-10 The population in the UK regions: (a) age 0-15; (b) age 65 and over; and (c) age 16-64. According to the 2017 population estimates in the UK (ONS, 2018), it showed the hexagon areas population across the UK.

4.10 Summary

UK dwellings are expected to experience a high risk of overheating in future climates. The indoor temperatures vary in different locations and time periods but are overall proportional to the outdoor operative temperatures. Dwellings in lower latitude UK cities

such as London will suffer more overheating risk. However, more than 3 million elderly and children are living in London. More attention should be paid to them as they are vulnerable to overheating while excess 'heat-related' mortality is proved to be related to overheating and heatwave.

Overheating hours can be gradually reduced by increasing the openable window area, but the reduction is not proportional to the openable area. The effect of increasing the openable area is becoming less on reducing overheating hours as a very small reduction was found when the window area increased from 60% to 80%.

Shading such as blinds and overhangs are more effective for reducing overheating hours in living rooms than in bedrooms because the shading measures could dramatically reduce solar gain during the daytime occupancy in living rooms. On the contrary, occupancy in bedrooms is from 10 pm to 7 am during the night when occupants go to sleep. Although the shading does reduce the overheating hours in bedrooms by reducing solar gain during the day to reduce indoor temperatures in the evening, night ventilation plays a more important role in reducing the overheating hours. Even without shading, night ventilation can cool down indoor temperatures effectively, reducing overheating hours significantly. In addition, the absorption of solar gain is highly related to the thermal mass of the buildings. Buildings with lower thermal mass can store less solar heat gain. Especially in light-weight high-insulated flats, they have less thermal mass than medium-weight houses which have more thermal inertia allowing houses to store heat during the day and release heat during the night. Therefore, shading is less effective for bedrooms of medium-weight houses but more effective in cooling light-weight flats.

Bedrooms are easier to be overheated due to stricter overheating assessment criteria involving 26°C as overheating threshold temperature. Bedrooms would have a high overheating risk when the outdoor temperature reaches 22°C during the night occupied hours. However, the living room would have a high risk of overheating when the outdoor temperature is over 29 °C.

This chapter pointed out that dwellings in London are expected to experience high overheating risk, but those in Edinburgh and Belfast have much less overheating risk. The largest number of vulnerable groups of people, such as children and the elderly are living in London and South East which are both low latitude regions expected to experience the highest overheating risk in the UK future climates. They are at the risk of heat-related deaths in the hot summer such as heatwaves in future. They are encouraged to move to the North of England or Scotland.

Chapter 5   Case studies and preparation

This chapter outlines the site preparation of the case study buildings for the development and implementation of the data-driven MPC for near real-time optimal control of the HVAC systems. Section 5.1 and 5.2 introduce monitoring and control systems in residential and commercial buildings. Below them, sub-sections also analyse the data collected from both sites. As metering and environmental data contain much information about buildings, systems and occupants, it provides an opportunity for a better understanding of building energy performance, indoor environment and occupants behaviours. This chapter analyses the collected data for extracting the information contained in the data before developing data-driven models in Chapter 6.

5.1 Residential building

5.1.1   Data

Data can be collected from the systems in the building. Each system can provide different types of data. As shown in Table 5-1, various data have been collected from different systems in the building. The heating system can provide temperature set-points which is a key to control the indoor temperatures. The monitoring system measures indoor dry-bulb temperature and relative humidity in each room through the temperature and humidity sensors installed in the rooms. The weather station can generate a lot of weather parameters such as outdoor temperature through the local weather station installed in the garden. These devices and sensors provide high-resolution measurements and offer the opportunity for occupants to understand the performance of their buildings. The data can also be used to develop data-driven models for predicting future energy or thermal performance, providing opportunities to adjust the control strategies for energy efficiency and thermal comfort.

Table 5-1 Monitored data from sensors and weather station.

Parameters	Sensors	Location/Equipment	Unit
Dry-bulb temperature	Temperature sensor	Bedroom, living room, kitchen, hallway	°C
Relative humidity	Humidity sensor	Bedroom, living room, kitchen, hallway	%
Electricity consumption	LED pulse sensor	Electricity consumption meter	kWh
Electricity Generation	LED pulse sensor	Electricity generation meter	kWh
Outdoor dry-bulb temperature	Weather station	Outside of the building	°C
Outdoor relative humidity	Weather station	Outside of the building	%
Temperature set-point	Thermostat	Hallway	°C

5.1.2 Data analysis

Table 5-2 shows the summary for inputs between 02 January 2018 and 07 March 2019 with a 15-min temporal resolution. The measured parameters are temperature set-points of the heating system and indoor air temperatures for the three rooms (bedroom 1, bedroom 2 and living room) in the test building. For room temperatures, the peaks are slightly different because the rooms face different directions and gain different levels of sunlight at different times.

Figure 5-1 shows the probability density and normal distribution for the model inputs, which illustrate the most distributed value for each input parameter during the study period. The results point out that most of the temperature set-points fall into two ranges. The first one is from 16°C to 17°C, and the other is from 19°C to 21°C. The first one is the unoccupied temperature set-points which occur when occupants are not at home, and another one represents occupied temperature set-points when occupants stay at homes. The results also show that indoor temperatures in bedroom 1 and living room are similar, representing occupants require the same heating level in both rooms. However, indoor temperatures in bedroom 2 are slightly lower, with a decrease of 1 to 2 °C. One of the reasons room temperatures are different is that occupants require less heating in this room or may use the room less frequently and sometimes turn off heating in this room. Another one is because the rooms face different directions and gain different levels of sunlight at different times, leading to different peaks of temperatures in different rooms.

Table 5-2 Summary of input data for the forecasting model.

Parameter	Unit	Min	Max	SD[1]	Skewness	Kurtosis
setpoint	°C	9.000	23.665	1.840	0.074	-1.326
bed1-temp	°C	16.463	27.308	1.541	0.997	1.443
bed2-temp	°C	16.320	27.998	1.683	0.979	1.228
living-temp	°C	16.606	28.617	1.652	0.584	0.028

[1] SD = Standard deviation.
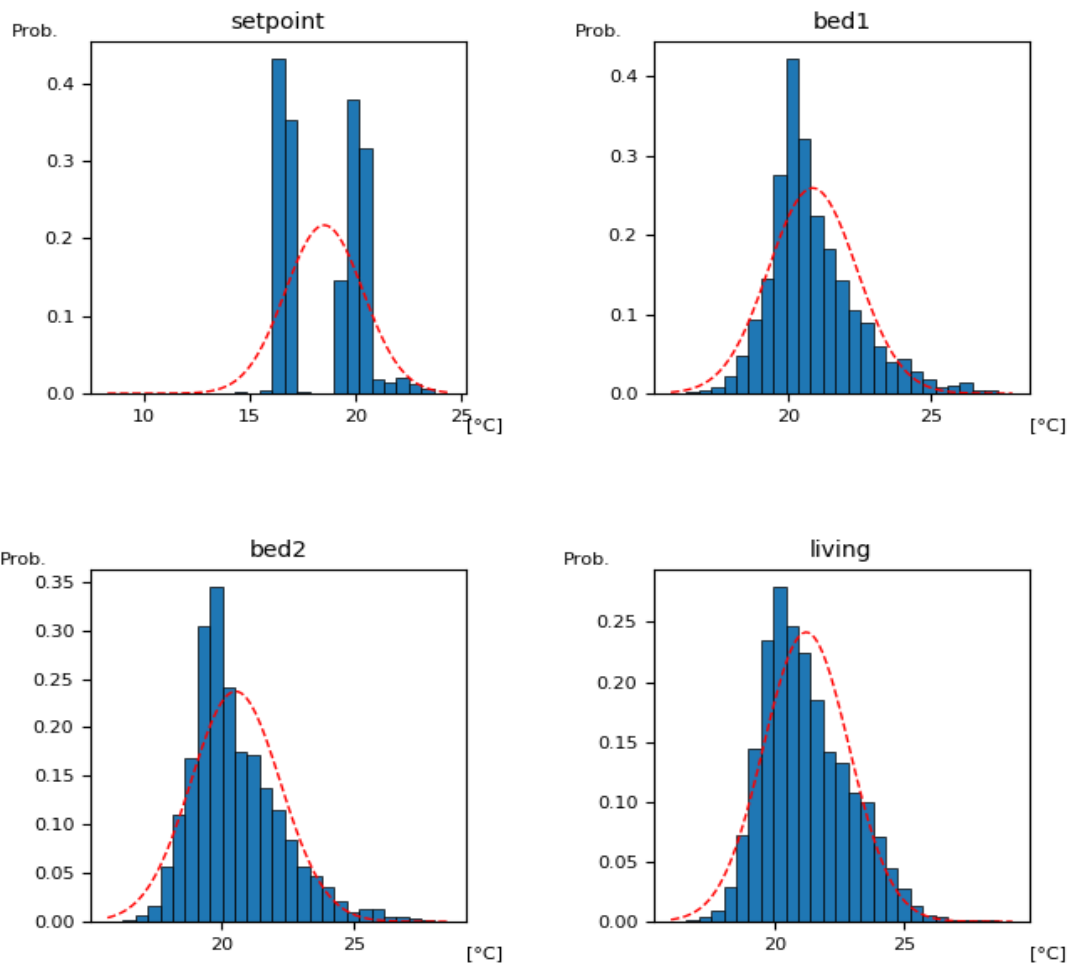
Figure 5-1 Probability density of controls and room temperatures. Red dashed lines indicate normal distribution. Data represents the period between 02 January 2018 and 07 March 2019.

5.2 Commercial building

5.2.1 data

The HVAC system on the 4th floor is used for testing MPC in this project. Figure 5-2 describes the distribution of the rooms, indoor units and ventilating pipes on the 4th floor. There are 9 indoor units from U2-10 to U2-18 on this floor. Indoor unit U2-10 and U2-18 are in the open space while other indoor units are in the office. The locations of indoor units and control panels are shown in Figure 5-2.
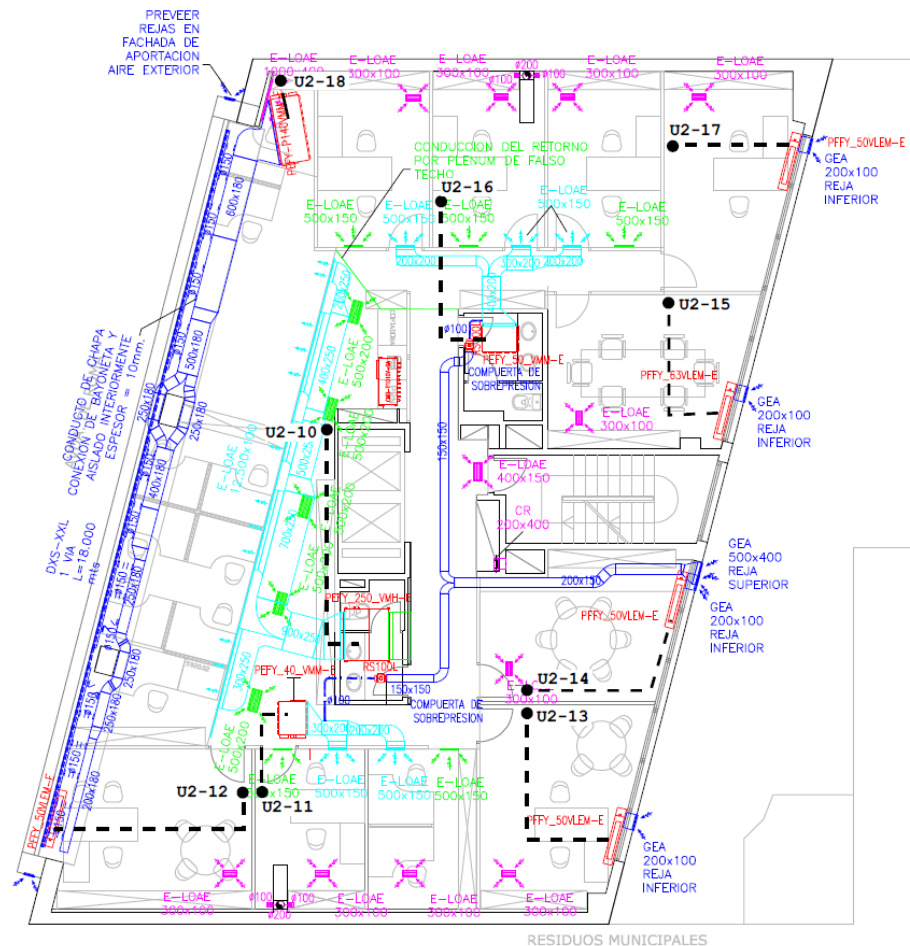
Figure 5-2 Floor plan for floor 4 of the commercial building. It describes the locations of the indoor units on the floor.

Table 2 listed all the number and type of devices on the 4th floor. The weather station is at the top of the building monitoring local weather parameters. Different types of indoor units have been installed in different places on the floor. From the table, we can find all the indoor units has the 'Read' and 'Write' function, which enables the indoor units to monitor the readings and change the control signals.

Table 5-3 describes all the parameters monitored from the given list of the devices. They are used to create a system model and optimise the control signals for energy cost saving, power control and indoor thermal comfort. The required parameters include:

1. 'energy_active' of the 'HVAC Floor 4 Meter' - total energy consumption of the HVAC on the floor level;

2. 'power_active' of the 'HVAC Floor 4 Meter' – the power of all indoor units on the floor level;

3. 'U2xx_hvac_unit_on', 'U2xx_hvac_unit_mode', 'U2xx_hvac_unit_set-point', 'U2xx_hvac_unit_temperature' - first three are used to control indoor units, last one is the indoor temperature as a result of control.

4. 'temperature_outdoor', 'humidity_outdoor' and 'solar_radiation' of the 'weather station' – they are used to describe the disturbance of outdoor weather.

Table 5-3 The list of monitoring devices for the HVAC system and weather. It listed all the devices that monitoring energy meter, power, indoor temperature and weather parameters.

Device	Device type	Read(R)/ Write(W)	Type
HVAC Floor 4 Meter	3-phase Electricity meter	R	measurement
Global Floor 4 Meter	3-phase Electricity meter	R	measurement
Indoor Unit 2-10	Ceiling concealed Indoor unit	R/W	control
Indoor Unit 2-11	Ceiling concealed Indoor unit	R/W	control
Indoor Unit 2-12	Ceiling concealed Indoor unit	R/W	control
Indoor Unit 2-13	Ceiling concealed Indoor unit	R/W	control
Indoor Unit 2-14	Ceiling concealed Indoor unit	R/W	control
Indoor Unit 2-15	Ceiling concealed Indoor unit	R/W	control
Indoor Unit 2-16	Ceiling concealed Indoor unit	R/W	control
Indoor Unit 2-17	Ceiling concealed Indoor unit	R/W	control
Indoor Unit 2-18	Ceiling concealed Indoor unit	R/W	control
T / RH - Floor 4	Temperature ang humidity sensor	R	measurement
Weather Station	Weather Station	R	measurement

Table 5-4 The parameters from the devices. It describes all the parameters that the devices measure and upload to the database.

Parameter	Description	Min	Max	Units	Device
energy_active	Active energy	---	---	kWh	HVAC Floor 4 Meter
energy_apparent	Apparent energy	---	---	kVAh	
energy_reactive_ind	Inductive energy	---	---	kVAh	
energy_reactive_cap	Capacitive energy	---	---	kVAh	
power_active	Active power	---	---	W	
power_apparent	Apparent power	---	---	VA	
power_reactive	Reactive power	---	---	VAr	
U2xx_hvac_unit_on[2]	On / Off	0 (Off)	1 (On)	---	Indoor unit (U2-10 to U2-18)
U2xx_hvac_unit_mode	Operation mode	1 (Heat)	4 (Cool)	---	
U2xx_hvac_unit_set-point	Temperature set-point	17	30	ºC[1]	
U2xx_hvac_unit_temperature	Ambient temperature	0	99.9	ºC	
temperature_indoor	Temperature	---	---	ºC	T/RH - Floor 4
humidity_indoor	Relative humidity	---	---	%	
pressure_outdoor	Barometer	880	1080	mbar	Weather station
temperature_outdoor	Outdoor temperature	-40	65	ºC	
wind_speed	Wind speed	3	241	kph	
wind_direction	Wind direction	0 (North)	360	degrees	
humidity_outdoor	Outdoor humidity	1	100	%	
uv_index	UV index	0	16	---	
solar_radiation	Solar radiation	0	1800	W/m2	
rain_day	Day rain	0	9999	mm	

5.2.2 Data analysis

Historical information for each input used in training the forecasting model of the HVAC system in the commercial building was obtained from the BMS system between the 1st of February and the 21st of October 2019 with 15-min temporal resolution.

Table 5-5 lists the measured parameters from the indoor air units, including control commands of indoor units, temperature set-points, room temperatures, as well as the total HVAC energy consumption and power on the 4th floor. Other parameters such as outdoor conditions are also measured and used in the model, including air temperature, humidity and solar radiation. The table shows the statistical description for the data through the standard deviation, Skewness and Kurtosis, giving insights into the data distribution. The skewness is used to measure the data symmetry, and the kurtosis measures the "peakedness" of the data. Besides, Figure 5-3, Figure 5-4 and Figure 5-5 illustrate the probability density and normal distribution for the model inputs, indicating the most distributed value for each input parameter.

As shown in Figure 5-3, Figure 5-4 and Figure 5-5, they describe the probability and normal distribution of the status (-1, 0 and 1 for cooling, stop working and heating), temperature set-point and temperature measured from different indoor units (from U2-10 to U2-18) deployed in the floor 4 of the commercial building. Apart from these control commands, outdoor temperature, outdoor humidity, solar radiation,floor-level energy and power are also described in the figures. The results show that most of the status of HVAC is at stop working due to no working during the night and weekend. The probability of cooling is more than heating because of the data collected between July and October. It is surprising that the temperature setpoints vary significantly in different indoor units for different rooms. One of the reason could be that the rooms face different direction gaining different level of solar heat. Indoor unit U2-10 locates in the open space. The results show that more heating and cooling are needed to condition the space. Indoor unit U2-16 is in the middle of the floor. Therefore, the room also needs more heating and cooling as it has no window for natural ventilation and solar gain. Energy and power present similar patterns as they are related to each other.

Table 5-5 Summary of monitored parameters from the commercial building from the 1st of July to the 21 of October 2019.

Parameter	Unit	Min	Max	SD [1]	Skew [2]	Kurt [3]
U210_hvac_command	int	-1	1	0.337	-2.058	3.152
U210_hvac_unit_set-point	°C	21	27	1.916	-0.610	-1.068
U210_hvac_unit_temperature	°C	21	29.5	1.314	0.498	0.825
U211_hvac_command	int	-1	1	0.330	-0.635	5.943
U211_hvac_unit_set-point	°C	19	26	2.420	-0.369	-1.825
U211_hvac_unit_temperature	°C	18.8	30.5	2.714	-0.160	-1.295
U212_hvac_command	int	-1	1	0.101	-5.829	93.182
U212_hvac_unit_set-point	°C	21	26	2.129	0.087	-1.655
U212_hvac_unit_temperature	°C	18.2	34.4	2.725	0.492	-0.003
U213_hvac_command	int	-1	1	0.180	-4.732	25.388
U213_hvac_unit_set-point	°C	21	26	2.320	-0.274	-1.790
U213_hvac_unit_temperature	°C	17.8	30.5	1.940	0.335	-0.448
U214_hvac_command	int	-1	1	0.224	-3.939	13.931
U214_hvac_unit_set-point	°C	21	26	2.332	-0.315	-1.806
U214_hvac_unit_temperature	°C	19.7	30.1	1.917	0.447	-0.391
U215_hvac_command	int	-1	1	0.241	-3.365	11.619
U215_hvac_unit_set-point	°C	21	26	2.180	-0.194	-1.682
U215_hvac_unit_temperature	°C	18.9	30.5	1.886	0.464	-0.295
U216_hvac_command	int	-1	0	0.356	-1.974	1.898
U216_hvac_unit_set-point	°C	21	27	2.025	-0.671	-1.114
U216_hvac_unit_temperature	°C	22	29	1.172	-0.094	-0.141
U217_hvac_command	int	-1	0	0.219	-4.096	14.783
U217_hvac_unit_set-point	°C	21	26	2.289	-0.284	-1.790
U217_hvac_unit_temperature	°C	20.1	30.1	1.850	0.396	-0.352
U218_hvac_command	int	-1	1	0.351	-1.827	2.690
U218_hvac_unit_set-point	°C	21	27	1.797	-0.714	-0.724
U218_hvac_unit_temperature	°C	19.5	32	1.920	0.470	0.726
temperature_outdoor	°C	7.3	37	4.848	0.356	-0.575
humidity_outdoor	%	11	92	12.949	-0.543	-0.129
solar_radiation	W/m^2	0	1204	296.992	1.155	-0.061
energy_active_hvac	kWh	0	6.1	0.635	2.182	4.668
power_active_global	W	159	19949	4341.576	1.330	0.783

[1] SD = Standard deviation, [2] Skew = Skewness, [3] Kurt = Kurtosis.
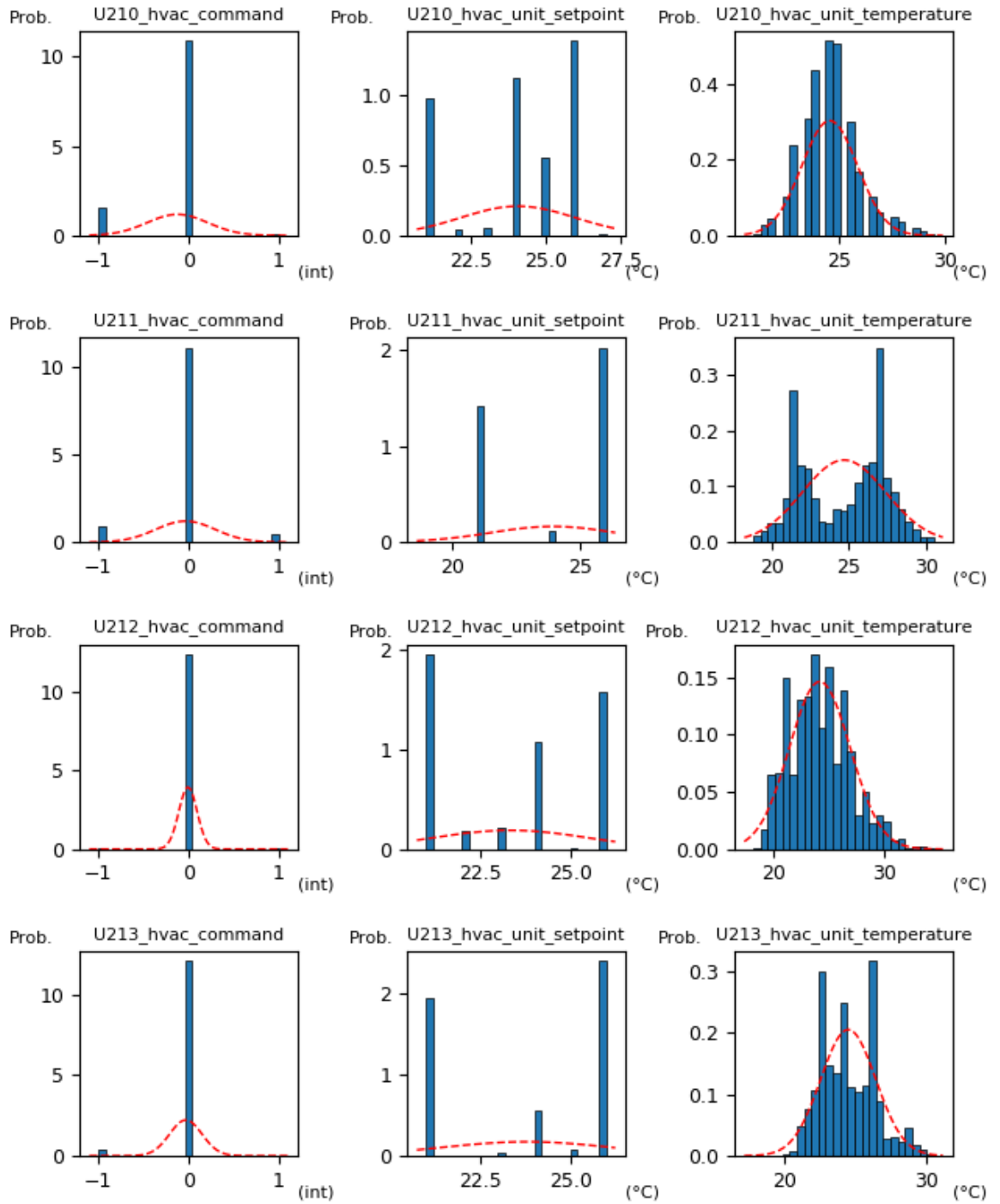
Figure 5-3 Probability and normal distribution that is a continuous probability distribution for each parameter from the 1st of July to the 21 of October (part 1).
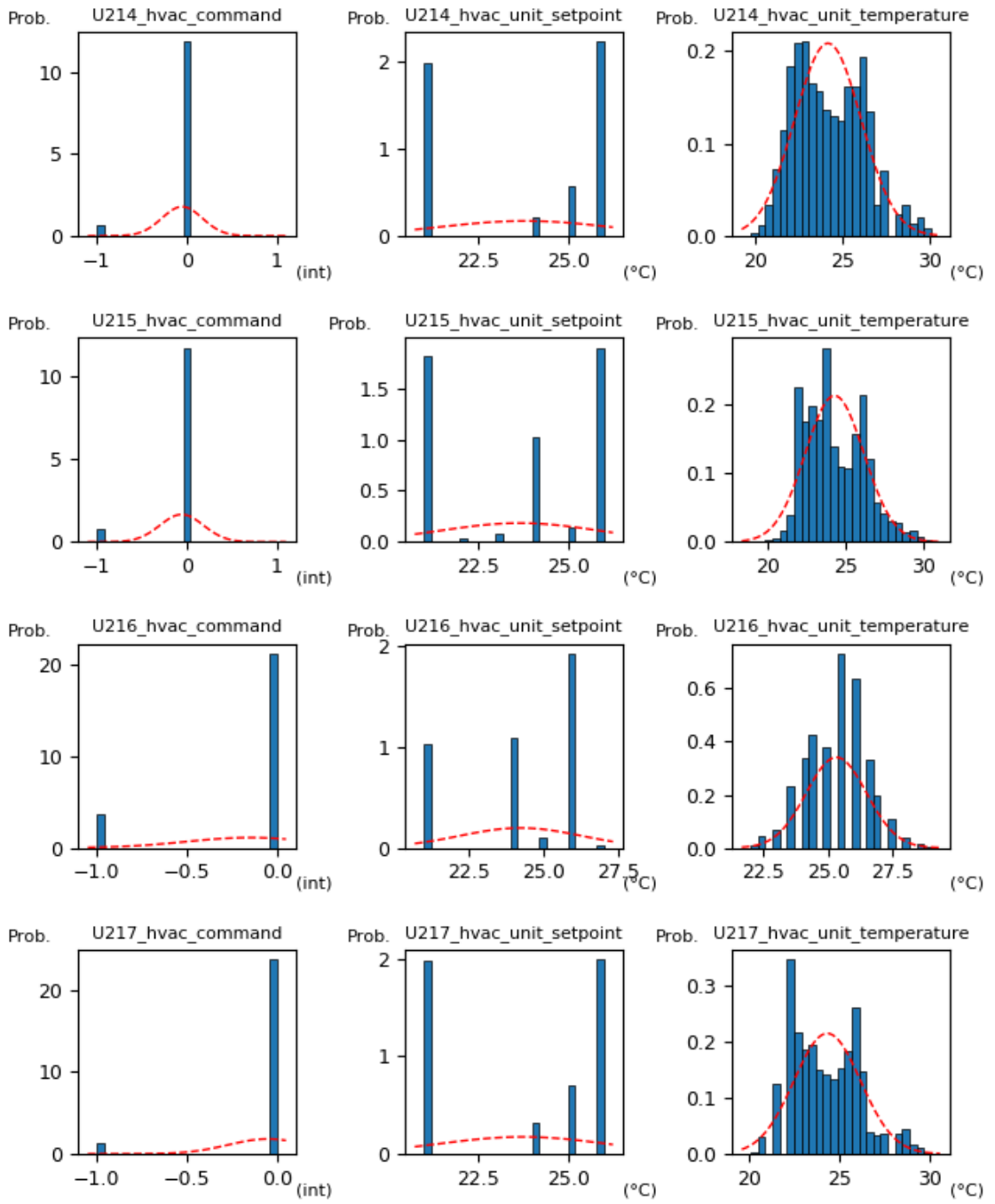
Figure 5-4 Probability and normal distribution that is a continuous probability distribution for each parameter from the 1st of July to the 21 of October (part 2).
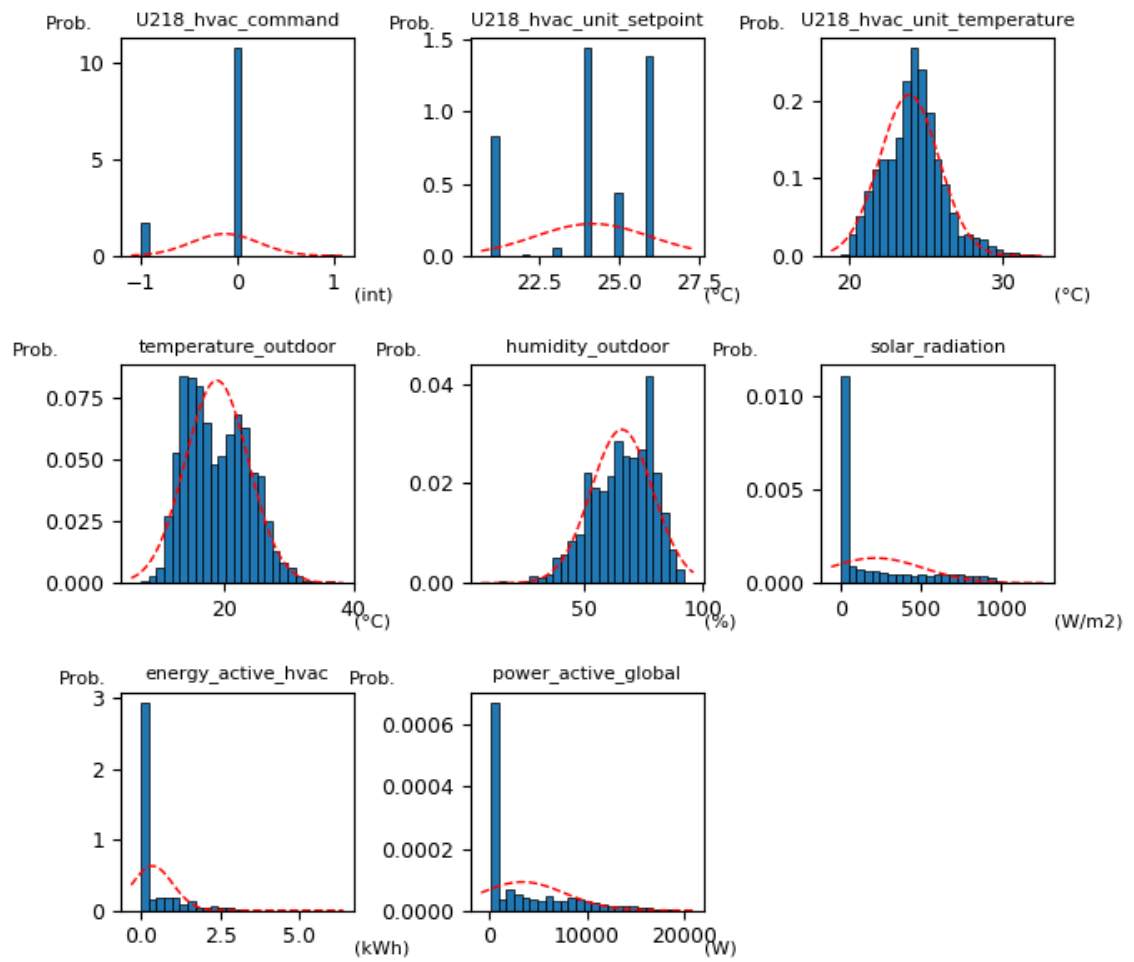
Figure 5-5 Probability and normal distribution that is a continuous probability distribution for each parameter from the 1st of July to the 21 of October (part 3).

5.3 Summary

This chapter analyses the data collected from the two case study buildings. It illustrates the temperature set-points are highly related to the occupancy in the residential building. Additionally, it demonstrates that the location of the rooms can significantly affect heating and cooling in the buildings. In the test commercial building, open space and rooms without windows require more heating and cooling to condition the indoor condition.

Chapter 6   Data-driven forecasting models

This chapter outlines the development of data-driven forecasting models for HVAC systems in residential and commercial buildings. Machine learning algorithm SVR, LSTM and RF were used to forecast indoor temperatures, energy consumption and power. More forecasting results can be found in Appendix A.

6.1 Introduction

Forecasting energy and thermal conditions in the buildings are vital for optimal control of the HVAC systems. It enables us to know how HVAC systems' operation responds to the disturbances such as weather, resulting in what levels of change in indoor comfort and energy performance. It also provides an opportunity to improve occupants' thermal comfort and the energy performance of the buildings and systems through optimisation of the controls. By integrating the forecasting models, optimisation aims to adjust the control actions with the forecast within finite time horizons. However, the optimisation needs to iteratively run the forecasting models to find an optimal solution, sometimes hundreds of thousands of times. Conventional simulation-based models such as physics-based simulation models require seconds (or minutes if the physics-based models are complex) to complete a simulation process, consuming hours or days for an optimisation. Such simulation-based models are not suitable for optimisation of HVAC controls, especially for near real-time optimal control that requires optimisation completed within a short time (such as 15 minutes) before taking the next control action.

Data-driven models are used to predict the behaviour of the HVAC systems in the case study buildings. The models can learn from past patterns in historical data to predict future events. Such patterns or relations are mapped between the input and output variables in the training stage. In this section, weather parameters, previous indoor temperatures, previous energy consumption, previous temperature set-point and time (e.g., month, day, hour) are the input variables. The output variables are future indoor temperatures, power and energy consumption for forecasting thermal and energy performance in the buildings. After the model has been trained, it will be used in the optimisation process to predict the optimal inputs (e.g., temperature set-point) at the next time step.

The data-driven model for the residential building is to forecast indoor temperatures, providing a thermal MPC to control the HVAC system to optimise the thermal comfort in the bedrooms and living rooms. The proposed MPC is to generate control signals for the next 15-minute temperature set-point for indoor thermal comfort. The data-driven forecasting model can predict indoor room temperatures with high accuracy in a short

time, normally in seconds or less than a second, making the near real-time MPC to produce optimised results within 15 minutes.

6.2 Data preparation

6.2.1 Residential building

Before training the model, it is necessary to clean the data. Data clean is essential but could be time-consuming as it may take many steps, including eliminating anomaly data, removing or filling the gaps and reproducing the data in a new resolution. Besides, input parameters/variables are determined and chosen, considering their importance to the output variables. The programming language Python is used as the coding environment for developing the data-driven models, and the package Scikit-Learn is applied for selecting the features according to the relative importance for each input feature.

Historical data are used to create data-driven models. These data are mostly time series data, and the inputs and outputs variables have been summarised in Table 6-1 and Table 6-2. Unlike the physics-based simulation model, building physics parameters such as window to wall ratio are not required. Time parameters (month, day, hour and minute), historical indoor temperatures and temperature setpoints in high resolution (per 15 minutes) are the inputs and outputs variables. Machine learning algorithms such as Neural Networks are used to create data-driven models. Input variables such as temperature set-points are also considered as the controllable variables or manipulated variables in the MPC for optimisation and control.

Table 6-1 Input variables of the forecasting model for the residential building.

Input	Parameter	Type	Value type	Min	Max	Unit
x_s	set_point	manipulated	continuous	15	35	°C
x_t^{b1}	bed1_temp	manipulated	continuous	15	35	°C
x_t^{b2}	bed2_temp	manipulated	continuous	15	35	°C
x_t^l	living_temp	manipulated	continuous	15	35	°C
x_m	month	measured	Integer continuous	1	12	Month
x_d	day	measured	Integer continuous	1	31	Day
x_h	hour	measured	Integer continuous	0	23	Hour
x_{mm}	minute	measured	Categorical	0	45	Minute

Table 6-2 The output variables of the forecasting model for the residential building.

Output	Parameter	Type	Value type	Min	Max	Unit
y_t^{b1}	next _bed1_temp	manipulated	continuous	15	35	°C
y_t^{b2}	next_bed2_temp	manipulated	continuous	15	35	°C
y_t^l	next_living_temp	manipulated	continuous	15	35	°C

6.2.2 Commercial building

Table 6-3 Input variables of the forecasting model for the commercial building.

Input	Parameter	Type	Value type	Min	Max	Unit
x_c^{u0}	U210_hvac_command	manipulated	Categorical	-1	1	N/A
x_s^{u0}	U210_hvac_unit_set-point	manipulated	Integer continuous	20	26	°C
x_c^{u1}	U211_hvac_command	manipulated	Categorical	-1	1	N/A
x_s^{u1}	U211_hvac_unit_set-point	manipulated	Integer continuous	20	26	°C
x_c^{u2}	U212_hvac_command	manipulated	Categorical	-1	1	N/A
x_s^{u2}	U212_hvac_unit_set-point	manipulated	Integer continuous	20	26	°C
x_c^{u3}	U213_hvac_command	manipulated	Categorical	-1	1	N/A
x_s^{u3}	U213_hvac_unit_set-point	manipulated	Integer continuous	20	26	°C
x_c^{u4}	U214_hvac_command	manipulated	Categorical	-1	1	N/A
x_s^{u4}	U214_hvac_unit_set-point	manipulated	Integer continuous	20	26	°C
x_c^{u5}	U215_hvac_command	manipulated	Categorical	-1	1	N/A
x_s^{u5}	U215_hvac_unit_set-point	manipulated	Integer continuous	20	26	°C
x_c^{u6}	U216_hvac_command	manipulated	Categorical	-1	1	N/A
x_s^{u6}	U216_hvac_unit_set-point	manipulated	Integer continuous	20	26	°C
x_c^{u7}	U217_hvac_command	manipulated	Categorical	-1	1	N/A
x_s^{u7}	U217_hvac_unit_set-point	manipulated	Integer continuous	20	26	°C
x_c^{u8}	U218_hvac_command	manipulated	Categorical	-1	1	N/A
x_s^{u8}	U218_hvac_unit_set-point	manipulated	Integer continuous	20	26	°C
x_m	month	measured	Integer continuous	1	12	Month
x_d	day	measured	Integer continuous	3	1	Day
x_h	hour	measured	Integer continuous	0	23	Hour
x_{mm}	minute	measured	Categorical	0	45	Minute
x_t^w	temperature_outdoor	measured	continuous	15	35	°C
x_h^w	humidity_outdoor	measured	continuous	0	100	%
x_s^w	solar_radiation	measured	continuous	0	1204	W
x_e	energy_active_hvac	measured, controlled	continuous	0	7	kWh
x_p	power_active_global	measured, controlled	continuous	0	20000	W
x_t^{u0}	U210_hvac_unit_temperature	measured, controlled	continuous	15	35	°C
x_t^{u1}	U211_hvac_unit_temperature	measured, controlled	continuous	15	35	°C
x_t^{u2}	U212_hvac_unit_temperature	measured, controlled	continuous	15	35	°C
x_t^{u3}	U213_hvac_unit_temperature	measured, controlled	continuous	15	35	°C
x_t^{u4}	U214_hvac_unit_temperature	measured, controlled	continuous	15	35	°C
x_t^{u5}	U215_hvac_unit_temperature	measured, controlled	continuous	15	35	°C
x_t^{u6}	U216_hvac_unit_temperature	measured, controlled	continuous	15	35	°C
x_t^{u7}	U217_hvac_unit_temperature	measured, controlled	continuous	15	35	°C
x_t^{u8}	U218_hvac_unit_temperature	measured, controlled	continuous	15	35	°C

Table 6-4 The forecasting model output variables for the commercial building.

Output	Parameter	Type	Value	Min	Max	Unit
y_e	next_energy_active_hvac	controlled	continuous	0	7	kWh
y_p	next_power_active_global	controlled	continuous	0	20000	W
y_t^{u0}	next_U210_hvac_unit_temperature	controlled	continuous	15	35	°C
y_t^{u1}	next_U211_hvac_unit_temperature	controlled	continuous	15	35	°C
y_t^{u2}	next_U212_hvac_unit_temperature	controlled	continuous	15	35	°C
y_t^{u3}	next_U213_hvac_unit_temperature	controlled	continuous	15	35	°C
y_{t1}^{u4}	next_U214_hvac_unit_temperature	controlled	continuous	15	35	°C
y_t^{u5}	next_U215_hvac_unit_temperature	controlled	continuous	15	35	°C
y_t^{u6}	next_U216_hvac_unit_temperature	controlled	continuous	15	35	°C
y_t^{u7}	next_U217_hvac_unit_temperature	controlled	continuous	15	35	°C
y_t^{u8}	next_U218_hvac_unit_temperature	controlled	continuous	15	35	°C

The next time step's floor energy consumption, power and room temperatures are the output variables in the forecasting model as presented in Figure 6-4. They are also the controllable variables that are to be controlled by changing the manipulated variables in the MPC.

It is important to evaluate the importance of features for a data-driven model. The higher percentage of the feature importance shows the higher correlation between the features and the forecasted outputs. Selecting the features can be carried out by using python functions such as 'feature importance', which automatically calculates the individual importance of each input parameters to the outputs.

Table 6-5 presents the correlation between each input (last timestep) and output (next timestep) feature. The energy consumption in the next timestep is mostly correlated to the last timestep power as it contributes with 79 % of the feature importance to next timestep energy, while the last timestep energy also contributes 8 % that is the second most important feature for energy forecasting. Other input variables such as outdoor temperature have less correlation to energy with no more than 2 % of the importance. Power is mostly related to its last time step value which accounts for 88 %. The hour and energy consumption at the last timestep also contribute 2 % respectively to the power, showing that time and energy consumption can also influence the power.

Regarding the indoor temperature, it is found that the last timestep for each room temperature has a significant impact on the next timestep of that room temperature. The correlation between indoor temperatures can reach up to 99 %. It can be explained that temperature changes slowly in a short time due to thermal inertia and heat exchange.

Table 6-5 The feature importance between the inputs and outputs. It shows the correlation between the input and output features.

Input feature	Output feature										
	y_e	y_p	y_t^{u0}	y_t^{u1}	y_t^{u2}	y_t^{u3}	y_t^{u4}	y_t^{u5}	y_t^{u6}	y_t^{u7}	y_t^{u8}
x_d	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
x_h	1%	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%
x_{mm}	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
x_t^w	2%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%
x_h^w	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%
x_s^w	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%
x_e	**8%**	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%
x_p	**79%**	**88%**	1%	0%	0%	0%	0%	0%	1%	0%	1%
x_t^{u0}	0%	0%	**95%**	0%	0%	0%	0%	0%	0%	0%	0%
x_t^{u1}	1%	0%	0%	**99%**	0%	0%	0%	0%	0%	0%	0%
x_t^{u2}	1%	1%	0%	0%	**98%**	0%	0%	0%	0%	0%	0%
x_t^{u3}	1%	0%	0%	0%	0%	**98%**	0%	0%	0%	0%	0%
x_t^{u4}	0%	0%	0%	0%	0%	0%	**99%**	0%	0%	0%	0%
x_t^{u5}	0%	0%	0%	0%	0%	0%	0%	**98%**	0%	0%	0%
x_t^{u6}	0%	0%	0%	0%	0%	0%	0%	0%	**96%**	0%	0%
x_t^{u7}	1%	0%	0%	0%	0%	0%	0%	0%	0%	**99%**	0%
x_t^{u8}	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	**97%**

6.3 Data-driven forecasting models

6.3.1 Auto-Regressive Integrated Moving Average (ARIMA)

ARIMA is used as a baseline model comparing with other three machine learning-based data-driven models. The forecasting models will be used in MPC for optimal control of the building HVAC systems. The models aim to forecast the temperature, energy and power in the two case study buildings. Compared to machine learning-based forecasting models, ARIMA model uses a single variable as its input and output, meaning that the forecast of temperature or energy is based on the historical temperature and energy. Other variables, such as weather, are not contributing to the forecasting results.

The p, d and q are the most important hyper-parameters in ARIMA models. The d is the number of differencing transformations needed to get the time series data to be stationary. In Figure 6-1, the indoor temperature is overall stationary in short terms but is not stationary in the whole year due to the change of seasons. When d=0, the temperature increases significantly when the time step is around 5000 and 7000 when months are July and August. As the tested UK house has no mechanical cooling but only natural ventilation as the main cooling measurement. The indoor temperatures are obviously affected by outdoor temperatures. After differencing the time series data with d=1, the data is stationary. Increasing d from 1 to 2, no significant change has been found. Therefore, the chosen d value is 1 for the stationarity.
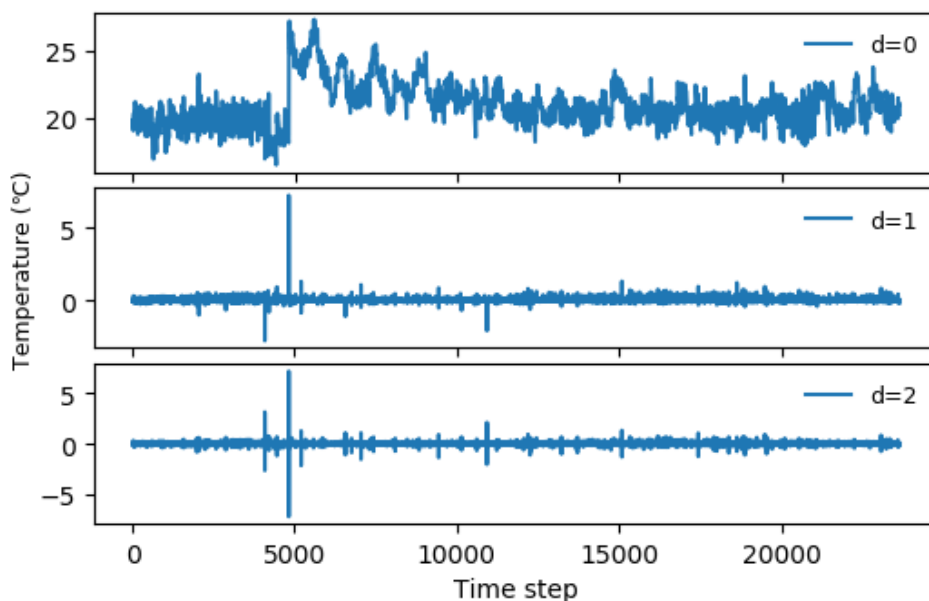
Figure 6-1 Differencing order for indoor temperature data.

Table 6-6 The AIC and BIC for different (p, d, q) for ARIMA model for indoor temperature in residential building.

ARIMA (p, d, q)	AIC	BIC
(0, 1, 0)	-36705.927	-36697.900
(0, 1, 1)	-40845.504	-40821.300
(0, 1, 2)	-42096.277	-42064.000
(1, 1, 0)	-42528.946	-42504.700
(1, 1, 1)	-43023.440	-42991.200
(1, 1, 2)	-43021.570	-42981.200
(2, 1, 0)	-42974.686	-42942.400
(2, 1, 1)	-43021.562	-42981.200
(2, 1, 2)	-43019.492	-42971.100

The AIC and BIC values of indoor temperature time series data for different p and q have been indicated in Table 6-6. The lowest AIC and BIC are found when the values are both 1 for p and q. Hence the ARIMA model is of order (1, 1, 1) for forecasting indoor temperature in the residential building and forecasting results have been compared with measured data in Figure 6-2 a.

Figure 6-2 Correlations between measured and predicted values using ARIMA for (a). the residential building's room temperature and (b). room temperature; (c). energy; (d). power in the commercial building

Figure 6-2a illustrates the correlation between the measured and predicted indoor temperature in residential building. The R squared value (R^2) is used to indicate the forecasting performance. The R^2 is between 0 and 1, while the high R^2 represents a more precise regression. The figure presents that ARIMA is accurate in forecasting the next

timestep indoor temperature in the residential building as the R^2 reaches 0.994. The trend line and equation show the linear regression between the measured and predicted value. The closer the points to the trend line means the higher forecasting accuracy. In Figure 6-2a, the scatter points are nearly overlapped with the red trend line showing the high accuracy of the forecasting results.

The same methods have been applied in the ARIMA forecasting models for indoor temperature, energy and power in the commercial buildings, showing in Figure 6-2b, c and d. Figure 6-2b illustrates the ARIMA also performs well in forecasting indoor temperature in the commercial building with an R^2 value of 0.924. It is less accurate compared to the forecasting performance on the indoor temperatures in the residential building. By investigating the actual temperatures in the commercial building, the indoor temperatures are grouped in an interval of 0.5 °C. The distribution of the actual indoor temperatures indicates that temperatures are controlled and maintained by the HVAC systems in the commercial building. The figure also shows that the largest difference between the measured and forecasted indoor temperatures is less than 2 °C. Additionally, the difference is slightly larger when the indoor temperature is between 24 and 25 °C, and smaller when the actual indoor temperature is more than 26 °C.

Figure 6-2c and d illustrate the correlation between the measured and predicted data for energy and power in the commercial building. The R^2 values are 0.823 and 0.899, respectively, less than those in the indoor temperature forecasting models. The scatter points are spread in a wider band instead of concentrating around the trend line. There are more points distributed far away from the trend line, indicating the inaccurate forecasting of those points.

6.3.2 Support Vector Regression (SVR)

SVR has been used to forecast the temperature, energy and power, and the results are demonstrated in Figure 6-3. The results show that SVR models have almost the same performance in forecasting indoor temperatures in both residential and commercial buildings. The correlations between the measured and predicted data are 0.99 and 0.927, respectively, compared to 0.994 and 0.924 using ARIMA methods. The results also reveal that SVR performs better than ARIMA in forecasting energy consumption, resulting in a higher correlation of 0.907 compared to 0.823 using ARIMA. In Figure 6-3 c, the scatter points are spreading in a narrower band, showing a smaller difference between the measured and predicted energy data. However, SVR has poor performance than ARIMA in forecasting power. The correlation is 0.83, which is much less than 0.899 with ARIMA methods.

Figure 6-3 Correlations between measured and predicted values using SVR for (a). the residential building's room temperature and (b). room temperature; (c). energy; (d). power in the commercial building.

6.3.3 Long Short-Term Memory (LSTM)

LSTM has been paid more attention to its sequence learning which enables it in capturing the temporal patterns (Kong, et al., 2017). In developing the LSTM models, learning rate and training epoch are two of the most important hyper-parameters. The optimisation of the two hyper-parameters is carried out to find out the best combination of them. Figure 6-4 and Figure 6-5 illustrates how learning rate and epoch affect the model training using the loss of training and testing. The results find out that the learning rate of 0.01 is too big for the model to learn the problem. Even when the epoch is increased from 500 to 1000, no significant improvement can be found. With the learning rate of 0.001, the volatility is much less compared to the training with the learning rate of 0.01. With the epoch increasing from 500 to 1000, the testing loss does improve and shows a smoother learning progression.
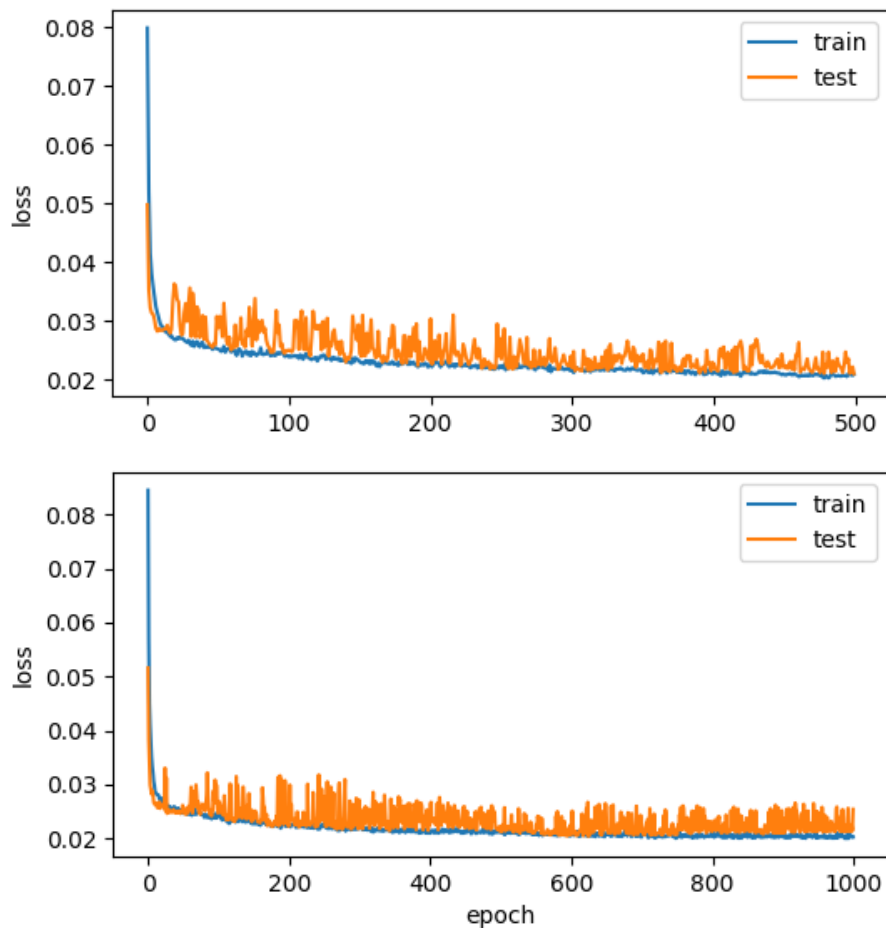
Figure 6-4 Learning rate of 0.01 for 500 and 1000 epochs.
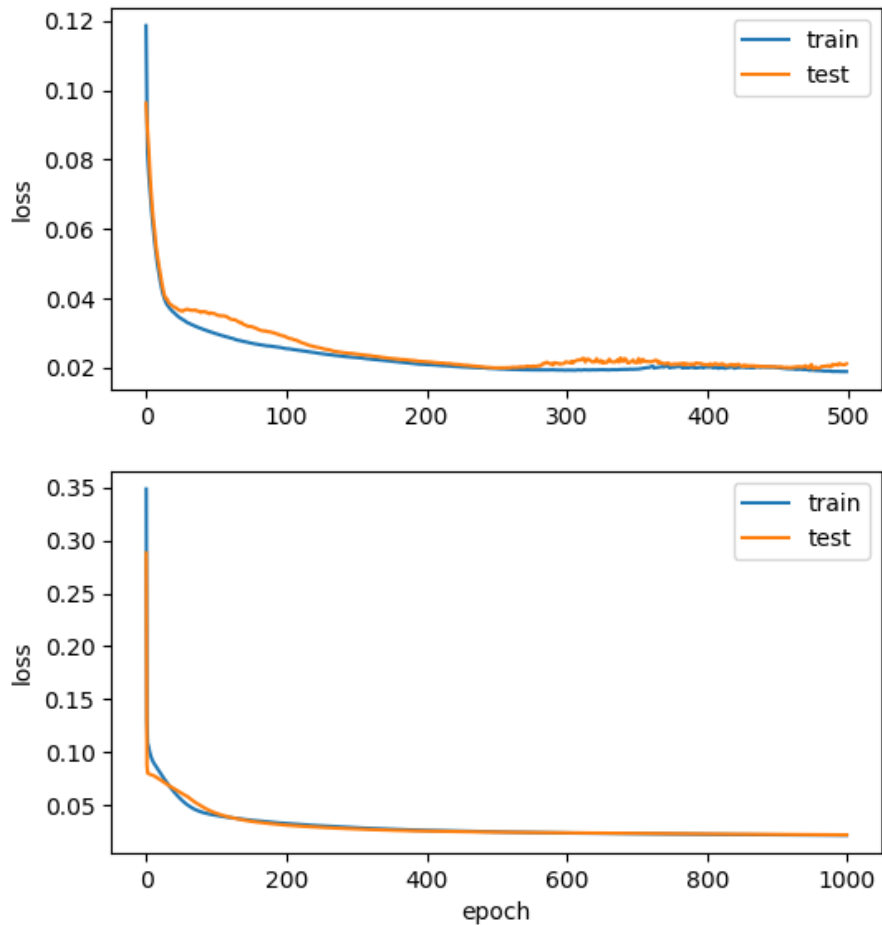
Figure 6-5 Learning rate of 0.001 for 500 and 1000 epochs.

The correlation between the measured and predicted indoor temperature, energy and power in residential and commercial buildings are illustrated in Figure 6-6. The results present that LSTM is as accurate as ARIMA in forecasting indoor temperatures in both residential and commercial buildings. The differences in R^2 are 0.002 and 0.004 compared to ARIMA models. However, LSTM performs better than ARIMA in the forecast of energy and power. The R^2 for measured and predicted energy reaches 0.903, which is much higher than 0.823 using ARIMA. The scatter figure also presents a higher density of points, showing less difference between the measured and predicted energy. For power forecasting, the correlation is slightly higher, with an increase of 0.1 in R^2. LSTM shows higher accuracy than ARIMA in forecasting the environmental and energy-related elements in this section.
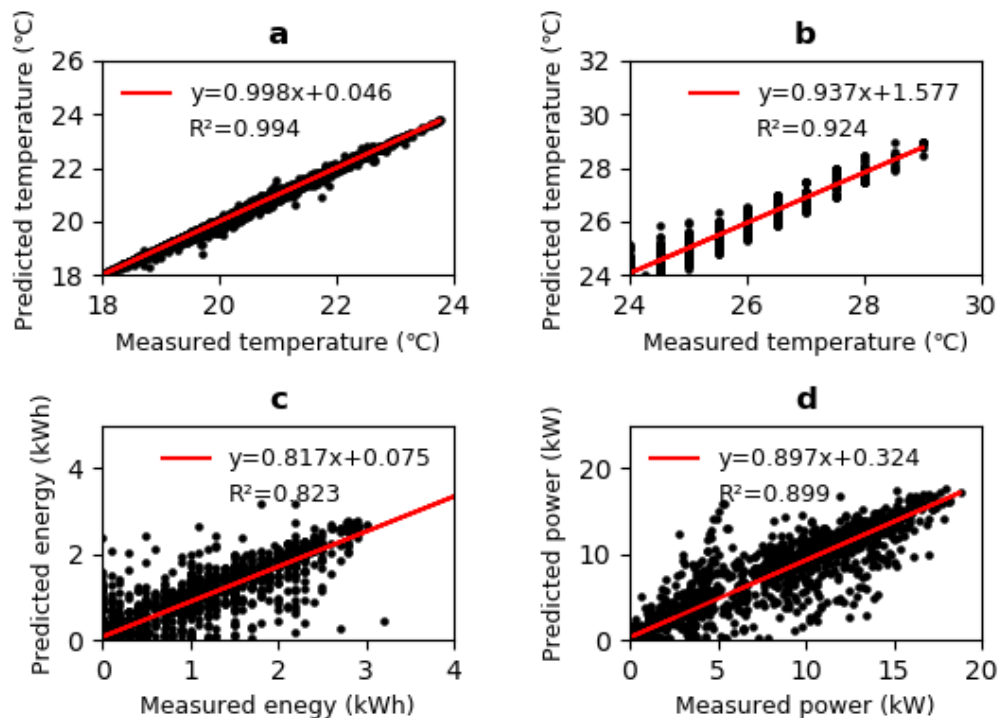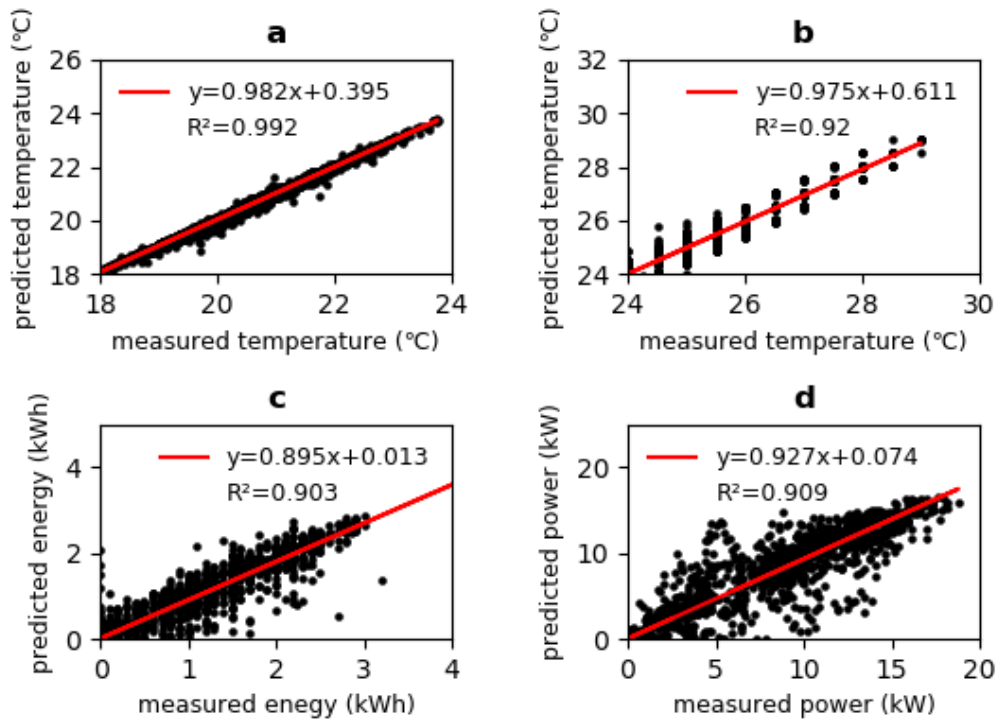
Figure 6-6 Correlations between measured and predicted values using LSTM for (a). the residential building's room temperature and (b). room temperature; (c). energy; (d). power in the commercial building.

6.3.4 Random Forest (RF)

RF is proposed for time series forecasting. Figure 6-7 illustrates the correlation between the measured and predicted data from the four different models. As shown in Figure 6-7 a and b, RF performs nearly the same as ARIMA in forecasting indoor temperatures. Compared to ARIMA, The R^2 decreases slightly for predicted indoor temperatures in the residential building but increases the same amount of R^2 in the commercial building. However, RF has the highest accuracy in forecasting energy and power, compared to ARIMA and the other two machine learning algorithms. The R^2 is 0.922 from the energy forecasting model with RF, compared to 0.823 with ARIMA, 0.907 with SVR and 0.903 with LSTM. For the power forecasting model with RF, the R^2 is 0.92 compared to 0.899 with ARIMA. 0.83 with SVR and 0.909 with LSTM. Figure 6-7 c and d also shows the distribution of scatter points is denser in a narrower band around the tend line, stating that RF is better for forecasting energy and power in the tested commercial building.
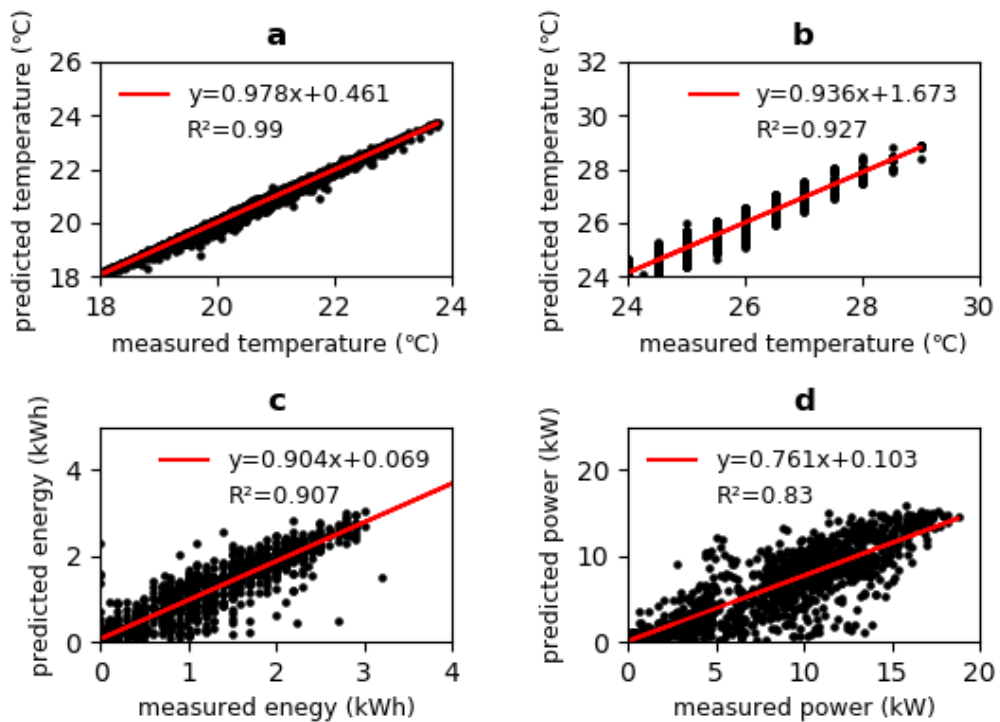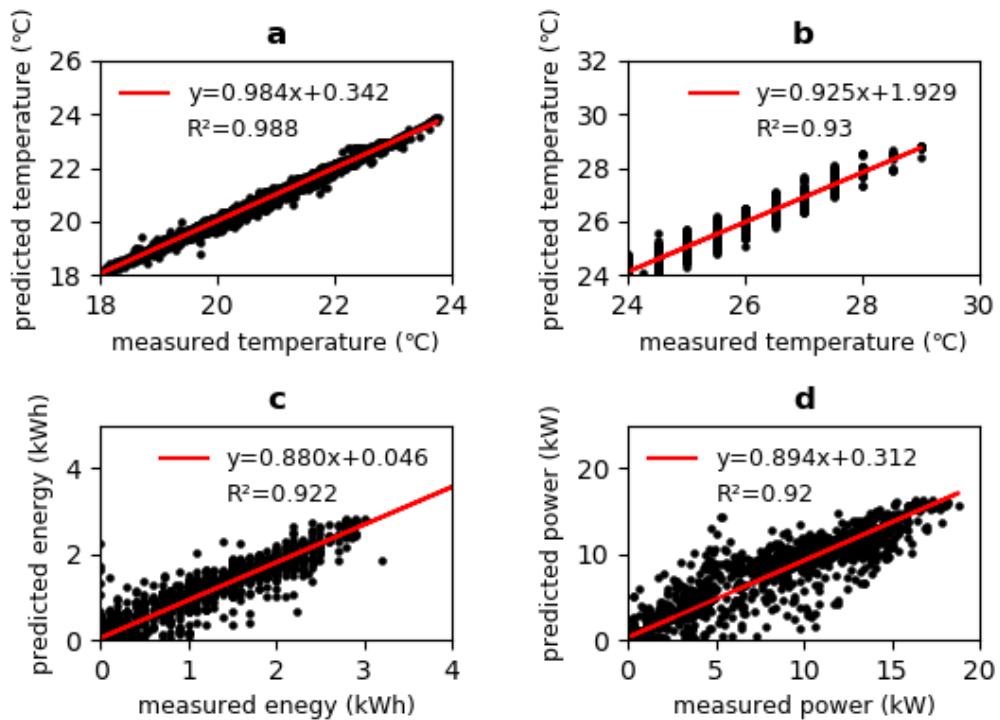

Figure 6-7 Correlations between measured and predicted values using RF for (a). the residential building's room temperature and (b). room temperature; (c). energy; (d). power in the commercial building.

6.4 Comparison of the machine learning algorithms

Forecasting models have been developed using different machine learning algorithms, and the forecasting results have been compared with the ARIMA model. The models have been analyzed by investigating the correlations between the measured and predicted data to provide an estimation of forecasting performance using R^2 value. The results present that all the machine learning-based models and ARIMA modes have high accuracy in forecasting indoor temperatures for both residential and commercial buildings. Other than the indoor temperature, machine learning-based models are more accurate than ARIMA in forecasting energy and power, except SVR, which shows less accuracy with a lower R^2 value in forecasting power compared to ARIMA.

Although R^2 has been used to analyze the performance of the forecasting models, there are more error measures such as MAE, RMSE and MAPE for error analysis. These error measures can provide more angles to analyze the forecasting performance among the different machine learning-based data-driven models.

6.4.1   Overall error analysis

Table 6-7 illustrates different types of errors to measure the forecast accuracy for machine learning-based data-driven forecasting models. From the R^2 scores, the results show that LSTM has the best performance in forecasting indoor temperatures in the residential building. RF model gains the highest accuracy in predicting indoor temperature, energy and power in the commercial building. As R^2 is used to explain how well the models fit the data or how much percentage of the variance can be explained, higher R^2 presents better forecasting performance. Compared to R^2, MAE, RMSE and MAPE are used to analyze the errors between the forecasted and actual values. Lower values represent higher forecasting accuracy.

Using MAE, RMSE and MAPE for error analysis, LSTM shows the highest accuracy in forecasting indoor temperatures in residential buildings. For forecasting indoor temperature in the commercial building, LSTM has the lowest MAE and MAPE value, but RF has the lowest RMSE and R^2. For forecasting energy, RF methods show the least error with the smallest MAE, RMSE and MAPE scores. RF has the best performance in forecasting energy in commercial building. For forecasting power, RF has the least MAE and RMSE scores, but LSTM has the least MAPE score. Considering the highest R^2 score with the RF model, RF is better than LSTM in forecasting power in commercial building. Overall, LSTM has the best performance in forecasting indoor temperatures in the residential building, but RF performs better in predicting indoor temperature, energy and power in commercial buildings.

Table 6-7 Forecasting error for whole testing data.

Element	Algorithm	R^2	MAE	RMSE	MAPE
Indoor temperature*	SVR	0.990	0.069	0.096	0.337
	LSTM	0.992	0.063	0.084	0.308
	RF	0.988	0.075	0.105	0.364
Indoor temperature**	SVR	0.927	0.196	0.258	0.777
	LSTM	0.919	0.189	0.271	0.746
	RF	0.930	0.195	0.251	0.772
Energy**	SVR	0.907	0.132	0.224	27.187
	LSTM	0.903	0.122	0.229	32.627
	RF	0.922	0.115	0.206	26.836
Power**	SVR	0.830	1099.899	1977.659	65.104
	LSTM	0.909	707.847	1449.478	44.076
	RF	0.920	701.790	1357.284	47.777

Note: * for the residential building and ** for the commercial building

Time can play an important role in forecasting accuracy. For example, energy consumption can be different on weekdays and weekend. Table 6-8 and Table 6-9 illustrate the errors of machine learning-based forecasting models on weekdays and weekend. The R^2 scores decrease for all the models in forecasting indoor temperatures in the residential, revealing all the machine learning algorithms present more accurate results in forecasting weekdays' indoor temperature in the residential building. However, the decreases in MAE, RMSE and MAPE show that the forecasting accuracy improves at the weekend.

For the commercial building, the higher R^2 scores are found in forecasting indoor temperature at the weekend, with an average of 0.94 compared to 0.89 on weekdays. The decrease of forecasting errors in the LSTM algorithm shows that the LSTM model has better performance in forecasting weekend's indoor temperatures in the commercial building. However, the increase of errors in SVR and RF models at the weekend shows they are less capable of forecasting the weekend's indoor temperature compared to the LSTM model.

LSTM models continue to show their better performance in forecasting weekend's energy and power, with higher R^2 scores and lower MAE and RMSE values. Although R^2 scores in SVR and RF models decrease in forecasting energy and power at the weekend, the MAE and RMSE values also decrease, showing an improving performance. Therefore, all machine learning algorithms have better performance in forecasting energy and power at the weekend with reduced errors of MAE and RMSE.

Although R^2 scores and various error measures (MAE, RMSE, MAPE) have been used to analyze the performance of the forecasting models with different machine learning

algorithms, there is no analysis for the details of the predicted results. For example, it is not possible to know how much difference is between the prediction and forecast. The difference between the predicted results and actual measurements can be used to enhance the understanding of the forecasting performance. The frequency distribution of the difference can be virtually investigated using a histogram, from which it is clear to see the difference and its distribution. The histogram aiming to compare the prediction and measurements for different forecasting models (regarding the indoor temperature, energy and power) has been demonstrated in Figure 6-8.

Table 6-8 Forecasting error for weekdays.

Element	Algorithm	R^2	MAE	RMSE	MAPE
	SVR	0.991	0.070	0.098	0.342
	LSTM	0.993	0.063	0.086	0.312
Indoor temperature*	RF	0.989	0.076	0.108	0.369
	SVR	0.887	0.194	0.257	0.776
	LSTM	0.889	0.193	0.272	0.767
Indoor temperature**	RF	0.893	0.194	0.250	0.774
	SVR	0.897	0.161	0.265	26.011
	LSTM	0.892	0.146	0.257	31.473
Energy**	RF	0.912	0.147	0.245	27.374
	SVR	0.790	1496.662	2377.279	55.629
	LSTM	0.890	890.218	1667.566	41.248
Power**	RF	0.901	956.533	1631.553	44.948

Note: * for the residential building and ** for the commercial building

Table 6-9 Forecasting error for weekends.

Element	Algorithm	R^2	MAE	RMSE	MAPE
	SVR	0.987	0.067	0.093	0.327
	LSTM	0.990	0.062	0.081	0.300
Indoor temperature*	RF	0.985	0.073	0.100	0.352
	SVR	0.940	0.200	0.259	0.778
	LSTM	0.942	0.177	0.266	0.693
Indoor temperature**	RF	0.942	0.199	0.254	0.768
	SVR	-0.540	0.069	0.076	153.085
	LSTM	0.935	0.074	0.131	37.094
Energy**	RF	-0.025	0.046	0.062	84.417
	SVR	-0.923	229.499	305.470	85.889
	LSTM	0.929	358.901	895.190	50.259
Power**	RF	0.095	142.948	209.525	53.982

Note: * for the residential building and ** for the commercial building
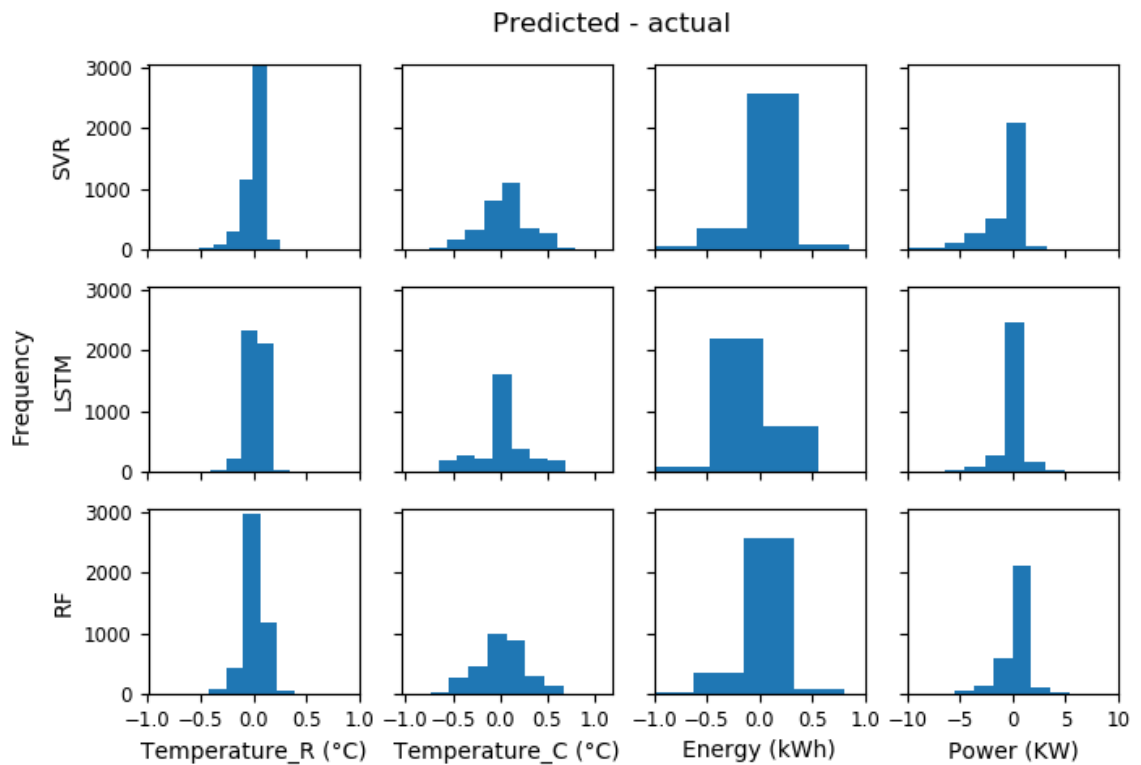
Figure 6-8 Frequency distribution of the difference between the measurements and prediction using different machine learning algorithms. (Temperature_R and Temperarture_C represent for indoor temperatures in the residential building and commercial building, and energy and power are from the commercial building).

Figure 6-8 illustrates the frequency distribution of the difference between the actual measurements and prediction for different elements (indoor temperature, energy and power) using SVR, LSTM and RF. The results found that the forecast for indoor temperatures is more accurate for residential building than for commercial building. Most of the differences between the measurements and the prediction are within 0.2 °C for residential building models. Compared to the residential building, there are some differences up to 0.6°C. For indoor temperatures in the residential building, although the LSTM has the highest R2 value, the SVR and RF show more results with errors within 0.1 °C. For indoor temperatures in the commercial building, LSTM has a higher frequency of differences within 0.1 °C. However, the differences can reach up to 0.7 °C using LSTM while the other two have the most differences, less than 0.6 °C. For energy forecasting, SVR and RF show better performance as most of the differences are from -0.2 to 0.2 °C while they are between -0.5 and 0 °C with LSTM. It has also been proved by the lowest R^2 and the highest RMSE and MAPE scores that the LSTM is less accurate than the other two algorithms in forecasting energy. For power forecasting, although RF has the highest R^2, the LSTM model shows that most of the differences are within 1.2 kW.

6.4.2 The forecast for peaks

The peaks in building environmental and energy measurements are important and more attention should be paid to forecasting these peaks. Peak indoor temperatures can be thermal risks to occupants if those temperatures are out of the thermal comfort range. Additionally, these peak temperatures may indicate unnecessary heating demand in the heating periods. The peak energy consumption is more important as those peaks indicate the highest energy consumption at certain points. With the time of use energy tariffs, occupants can shift some energy consumption to other points to save energy cost. For peak power, commercial buildings such as offices may use kWmax energy tariffs, which charge different rates of energy according to the highest power. Therefore, the energy cost is not only related to total energy consumption but also to the peak energy or power.

Indoor temperatures in the residential building

Figure 6-9 illustrates the indoor temperatures and the peaks in the residential building. Those peaks have been highlighted with 'x' labels in the figure. As the season of the testing sample is around Spring, the indoor temperatures change between 18 °C and 24°C, while all the selected peaks are assumed to be over 20 °C. The number of the peaks are 304 out of the total number of 4740 measurements. By comparing the real measurements with the forecasted peaks using different algorithms, the errors have been illustrated in Table 6-10. The results show that LSTM has the best performance in forecasting the peak indoor temperatures in the residential building. It has the highest R^2 score of 0.988 compared to 0.982 by SVR and 0.980 by RF. Additionally, LSTM also has the least MAE, RMSE and MAPE values among the three algorithms.

Table 6-10 The errors of forecasted peaks of indoor temperatures in the residential building.

Algorithm	R^2	MAE	RMSE	MAPE
SVR	0.982	0.067	0.104	0.315
LSTM	0.988	0.053	0.084	0.253
RF	0.980	0.079	0.110	0.370
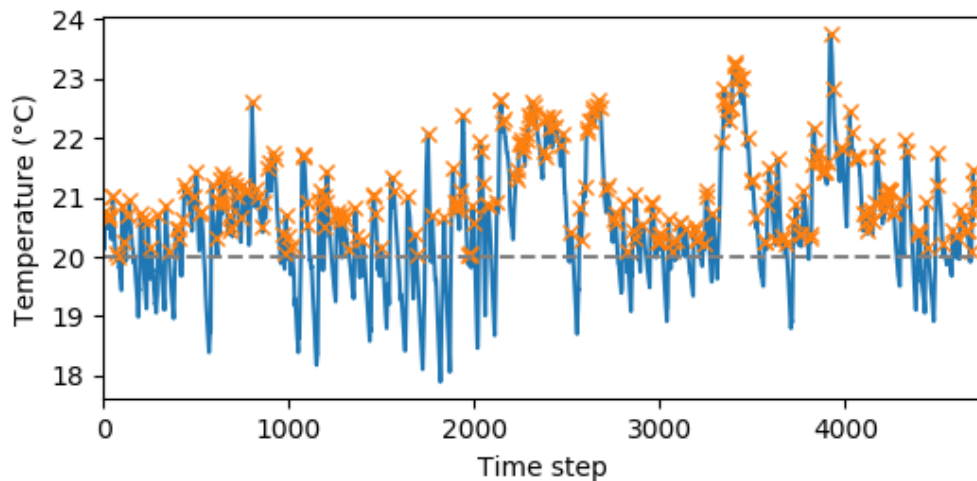
Figure 6-9 The peaks of indoor temperatures in the residential building.

Indoor temperatures in the commercial building

Figure 6-10 illustrates the indoor temperatures and the peaks in the commercial building. The number of the peaks are 333 out of the total number of 3066 measurements between September and November. These peaks have been highlighted with 'x' labels in the figure, and the highest ones are over 28 °C. Compared to the residential building, most of the peaks of indoor temperatures in the commercial building are much higher due to the season and location of the building. The high temperatures over 26 °C almost occurred during the weekend or several hours before the opening time on Monday. With the 26°C as the maximum temperature setpoint during these months, the indoor temperatures during the working hours are between 24 °C and 26 °C, controlled by the HVAC system. Due to the HVAC control, the indoor temperatures often stay in certain temperatures, resulting in most of the peaks with 0.5 °C intervals, including 25 °C, 25.5 °C and 26 °C.

By comparing the real measurements with the forecasted peak using different algorithms, the errors have been illustrated in Table 6-11. The results show that SVR has the best performance in forecasting the peak indoor temperatures in the commercial building. It has the highest R^2 score of 0.864 compared to 0.803 by LSTM and 0.856 by RF. Additionally, SVR also has the least MAE, RMSE and MAPE values among the three algorithms.
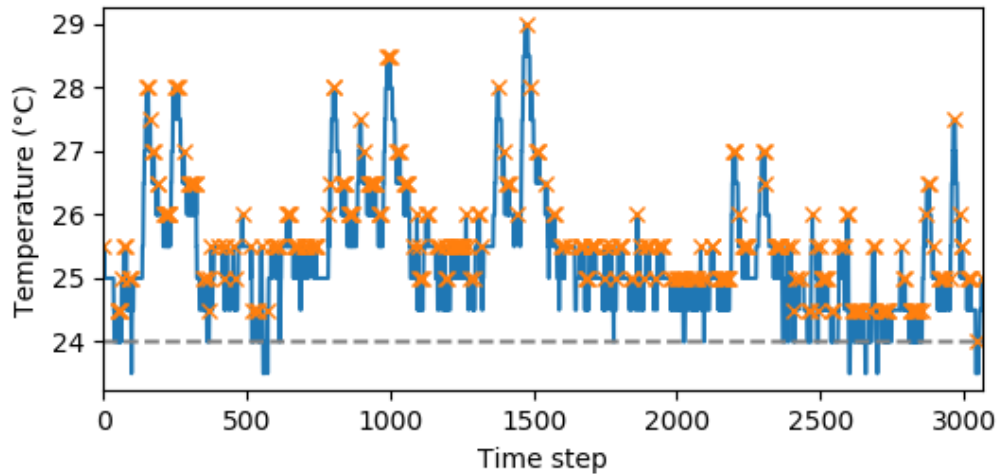
Figure 6-10 The peaks of indoor temperatures in the commercial building.

Table 6-11 The errors of forecasted peaks of indoor temperatures in the commercial building

Algorithm	R^2	MAE	RMSE	MAPE
SVR	0.864	0.276	0.318	1.079
LSTM	0.803	0.336	0.383	1.313
RF	0.856	0.288	0.328	1.123

Energy

The peaks of energy consumption represent the increasing energy demand such as heating and cooling during the day in the office. The high demand for energy consumption normally requests a high supply of power, resulting in an increasing rate of energy. It is possible to shift these demand through methods such as pre-heating or cooling. Shifting demand to the time with a lower energy rate can also save energy cost. Therefore, the forecast for peak energy consumption is important and helpful to the optimisation of the operation of the building system.

Figure 6-11 illustrates the floor level energy consumption and the peaks in the commercial building. The number of the peaks are 224 out of the total number of 3066 measurements between September and November. The peaks have been highlighted with 'x' labels in the figure, while the highest peaks are above 4 kWh. The results show that almost half of the peaks are over 2 kWh, and some of them can even exceed 3 kWh.
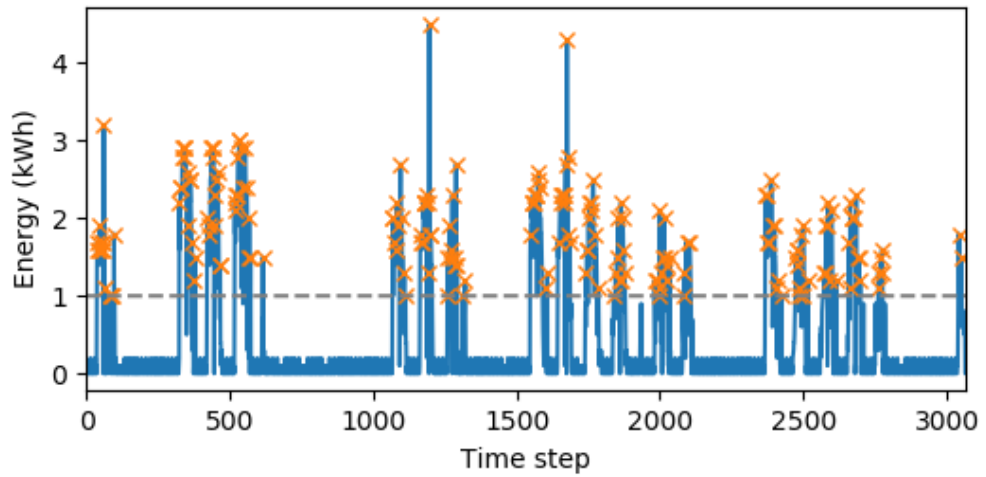
Figure 6-11 The peaks of energy consumption in the commercial building.

Compared with the forecast of peaks indoor temperatures, the forecast of the peak energy consumption is less accurate. By comparing the real measurements with the forecasted peaks with different algorithms, the errors have been illustrated inTable 6-12. The results show that RF has the best performance in forecasting the peak indoor temperatures in the residential building. It has the highest R2 score of 0.574 compared to 0.446 by SVR and 0.292 by LSTM. Additionally, RF also has the least MAE, RMSE and MAPE values among the three algorithms.

Table 6-12 The errors of forecasted peaks of energy consumption in the commercial building.

Algorithm	R^2	MAE	RMSE	MAPE
SVR	0.446	0.331	0.500	21.437
LSTM	0.292	0.390	0.566	24.869
RF	0.574	0.325	0.439	20.409
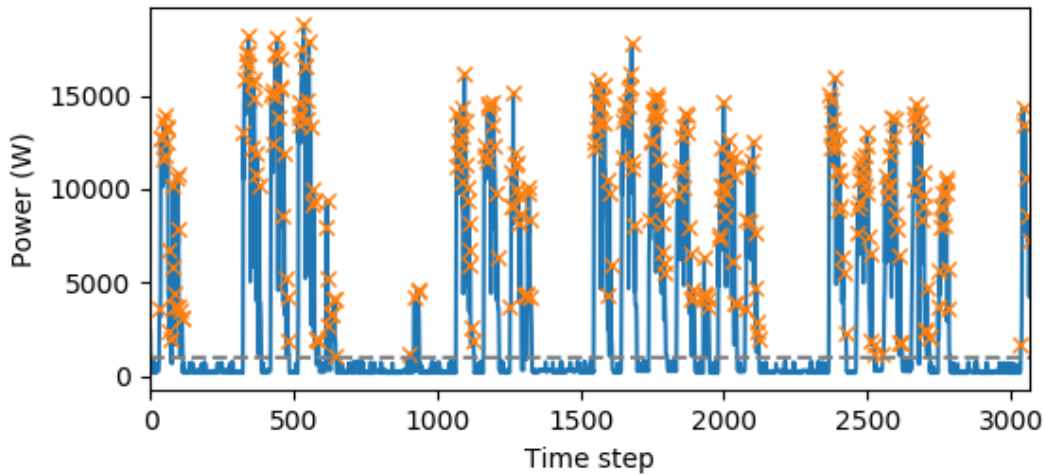
Figure 6-12 The peaks of power in the commercial building.

Table 6-13 The errors of forecasted peaks of power in the commercial building.

Algorithm	R²	MAE	RMSE	MAPE
SVR	0.037	3634.906	4309.945	41.722
LSTM	0.560	2126.266	2913.455	24.361
RF	0.609	2079.282	2748.071	23.540

Power

Same as the peak energy consumption, the peaks of power generally represent for the increasing energy demand during the day in the office. Figure 6-12 presents the same patterns as energy peaks in Figure 6-11, indicating high energy consumption draw high power at the same time. High power results in not only higher energy consumption but also higher cost because the energy rate is increased if the maximum power exceeds a certain value. As shown in Table 7-2 and Table 7-3, the energy rate in peak hours increases from 0.105507 to 0.111586 €/kWh if the maximum power of the whole building exceeds 157.5 kW. By shifting energy demand to other time with lower power, the peak power can be reduced. With foreseen power peaks, it is possible to shift demand and reduce energy cost with a lower energy rate. Therefore, the forecast for peak power is vital and meaningful to the optimisation of the operation of the building system.

Figure 6-12 illustrates the power and the peaks in the commercial building. The number of the peaks are 319 out of the total number of 3066 measurements between September and November. These peaks have been highlighted with 'x' labels in the figure, and the highest ones are exceeding 15000 W. The results show similar accuracy as the forecast of energy peaks. By comparing the real measurements with the forecasted peaks with different algorithms, the errors have been illustrated in Table 6-13. The results show that

RF has the best performance in forecasting the peak indoor temperatures in the commercial building. It has the highest R^2 score of 0.609 compared to 0.037 by SVR and 0. 0.560 by LSTM. Additionally, RF also has the least MAE, RMSE and MAPE values among the three algorithms.

6.4.3 Day-ahead forecasting

Day-ahead sub-hourly forecasting models for residential building's indoor temperatures, commercial building's indoor temperatures, energy and power have been developed. The needs of the day-ahead forecast are for developing MPC for the next 24 hours. The sub-hourly (per 15 minutes) models are repeated for 96 times (24 hours) to gain the day-ahead forecasting. In this process, the outputs of the sub-hourly forecasting models for the last time step are used as the inputs in the next time step. In the residential building, the day-ahead forecast of indoor temperatures is integrated with the MPC for controlling heating to maintain the indoor temperatures. In the commercial building, the day-ahead forecast of indoor temperatures, energy and power are applied in the MPC for controlling indoor units to maintain the indoor temperatures and save energy cost. As the MPC for the next 24 hours is highly dependent on the day-ahead forecasting results, it is necessary to investigate the day-ahead forecasting performance of the models.

Indoor temperatures in the residential building

The day-ahead forecasts of indoor temperatures in the residential building using SVR, RF, and RNN have been illustrated in Figure 6-13. The results show that the SVR model has the best performance in forecasting day-ahead indoor temperatures in the residential building. The prediction by the RF model underestimates the actual values slightly compared to the SVR method. The worst performance is found in the forecast by the LSTM model, which is much lower than the actual values.

Indoor temperature in the commercial building

The day-ahead forecasts of indoor temperatures in the commercial building using SVR, RF and RNN have been illustrated in Figure 6-14. The results show that the SVR model has the best performance in forecasting day-ahead indoor temperatures in commercial building. The prediction by the LSTM model overestimates the actual values slightly compared to the SVR method. The worst performance is found in the forecast by the RF model, which is much higher than the actual values. The prediction by the RF is flat compared to the other two methods and can not capture the patterns. However, both SVR and LSTM are capable of following the patterns and capturing the valley in the prediction.
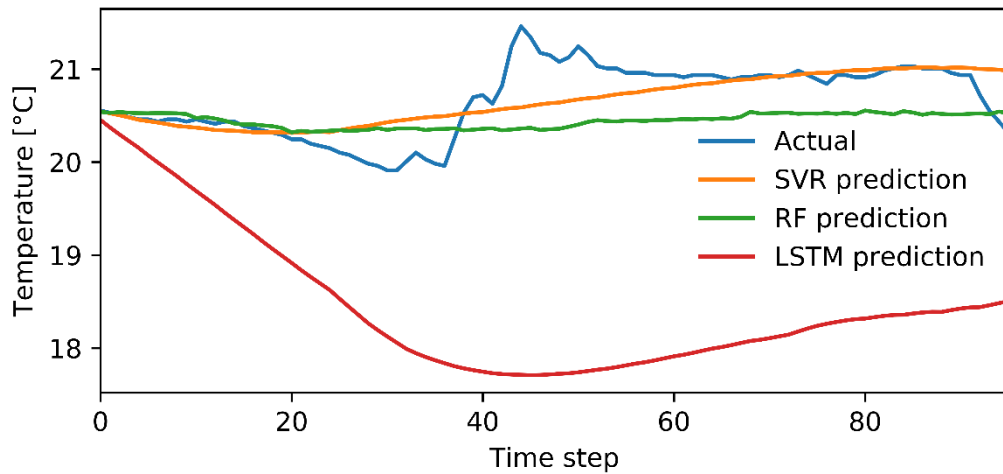
Figure 6-13 Day-ahead forecast of indoor temperature in the residential building.
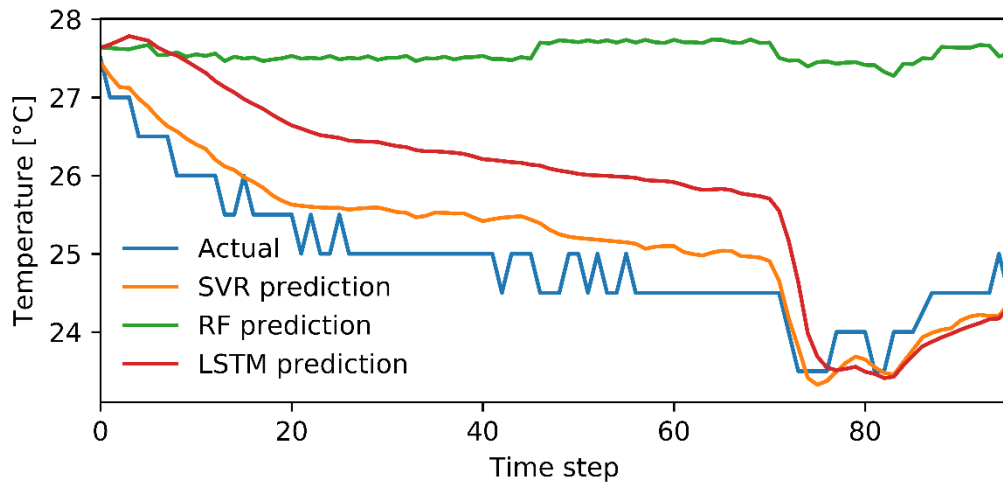
Figure 6-14 Day ahead forecast of indoor temperature in the commercial building.

Energy

Figure 6-15 presents the day-ahead forecast of energy in the commercial building using SVR, RF and RNN. The results indicate that both the SVR model has the best performance in predicting the day-ahead energy consumption. Although LSTM shows better performance in the time step from 80 to 96, its prediction gives zero energy consumption in the low consumption period (time step from 0 to 70), which underestimate the actual values and is not realistic in a commercial building. Compared to LSTM, SVR provides a prediction of reasonable non-zero energy consumption in this period. RF performs similarly to SVR in the low consumption time steps, but it can not capture both peaks. Therefore, RF has the worst performance in forecasting the day-ahead energy consumption in commercial building.
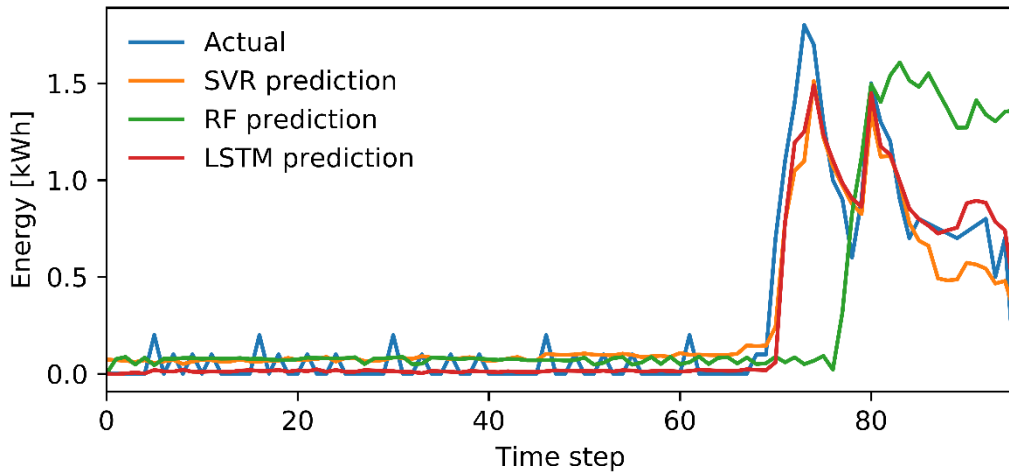
Figure 6-15 Day ahead forecast of energy in the commercial building.

Power

Figure 6-16 depicts the day-ahead forecast of power in the commercial building using SVR, RF and RNN. The results illustrate that the LSTM model has the best performance in predicting day-ahead power. Although both LSTM and SVR have a few time steps (2 to 3 steps) delayed in predicting the peaks, LSTM performs better than SVR as its predicted peaks are closer to the actual values. RF performs similarly as SVR and LSTM in the low consumption time steps from 0 to 70, but the predicted peaks are much more delayed than using the other two methods. Therefore, RF has the worst performance in forecasting the day-ahead power in the commercial building.
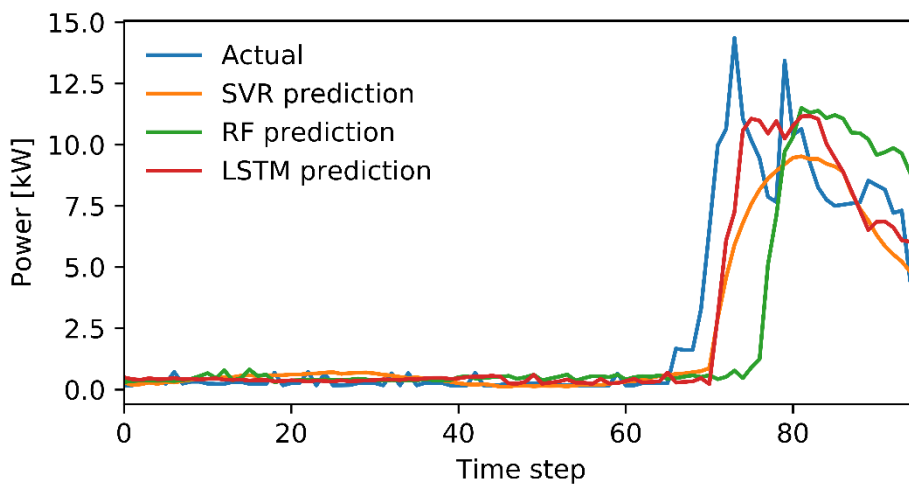
Figure 6-16 Day ahead forecast of power in the commercial building.

6.5 Summary

This chapter developed data-driven models using different machine learning algorithms, including SVR, LSTM and RF. A statistical method ARIMA was used as a baseline model to compare with them. From the above comparison and analysis (Section 6.3 to 6.4), it is obvious that each machine learning algorithm has its advantages and disadvantages in certain cases. All the machine learning algorithms show high accuracy of up to 0.99 in R^2 values in forecasting indoor temperatures in the residential building. The accuracy decreases slightly to up to 0.92 for forecasting indoor temperatures (1 of the 9 room temperatures as an example) in the commercial building. RF outperforms the other two algorithms in forecasting global energy and power for floor 4 in the commercial building, with both R^2 values of 0.92. The accuracy of the forecast of power is decreased at weekdays and increased at weekends. In forecasting the peaks, the forecasting accuracy of indoor temperatures decreases slightly, but they drop significantly from up to 0.92 to up to 0.61 for forecasting energy and power. Increasing the forecasting steps from 1 to 96 for day-ahead forecasting, SVR performs better than the other two algorithms in matching the one-day patterns.

Chapter 7   Near real-time model predictive control

This chapter introduces the development of data-driven MPC for HVAC systems in both case study buildings. Section 7.2 proposes an MPC for optimizing the heating system in a residential building for improving indoor thermal comfort. Section 7.3 develops an MPC for optimizing the heating and cooling in a commercial building for minimizing the energy cost by integrating the energy tariffs with the constraints in maximum power and indoor temperatures in a comfortable range.

7.1 Introduction

This chapter reports on the development of MPC controllers for the HVAC system in the commercial building. First, it introduces the background information for the test site, including the HVAC system and the methods for data collection are described. Secondly, the optimisation problem is defined after a brief review of building control approaches and the key MPC techniques, including the development of models, the considerations of disturbances and their applications. Then, the procedures and methodologies for the development of MPC are discussed, including the training and testing for the data-driven model. After that, the developed evolutionary algorithm based MPC is investigated on the control of modes and set-point temperatures of the HVAC units in the building. The results highlight the potential benefits of the application of MPC in achieving the optimisation objectives.

7.2 Residential building

7.2.1   Objective function

The optimisation problem is identified through the design of the objective function and constraints. The thermal system in the Cardiff test site is a central heating system that is composed of a boiler and radiators in rooms in a residential house.

The objective function, expressed in Equation 1, shows the optimisation target is to minimize the deviation of the actual indoor temperature from the reference comfort temperature at each time step. The goal is to improve the thermal comfort of occupants in the building. It demonstrated the difference between the indoor temperature (T_k) and the reference temperature (R_k) from time step k to N should be minimized. As there are three room temperature monitored, each monitored indoor temperature should be considered ($T \in [T_1, T_2, T_3]$, T_1 for bedroom 1, T_2 for bedroom 2, and T_3 for the living room).

$$\min \sum_{k=1}^{N} (\,|\,T_k - R_k\,|\,) \tag{1}$$

s.t. $\qquad\qquad T_{lb} \leq T_k \leq T_{ub}$

7.2.2 Constraints

The constraints are described in Equation 2 and 3, illustrating the temperature constraints for each room (T_1, T_2, T_3 for bedroom 1, bedroom2 and the living room). The aim of the constraints is to limit indoor air temperatures within the given comfort ranges.

$$18 \leq T_k \leq 23 \text{ (winter)} \tag{2}$$
$$22 \leq T_k \leq 26 \text{ (summer)} \tag{3}$$
$$T_k = [T_1, T_2, T_3]$$

7.2.3 Optimisation algorithm

An evolutionary algorithm is used to solve the objective function by satisfying the constraints based on the forecasting model. The forecasting model is used for forecasting the change of the heating system or building environment in this case. A suitable optimisation algorithm should be adopted to solve the formed optimisation problem. Due to the use of a data-driven forecasting model, the evolutionary algorithm is selected to solve the nonlinear and black box model. The evolutionary algorithm-based method can be used to address the cost-effective problem by generating better solutions in new generation/ offspring.

7.2.4 Results and discussion

The MPC has been developed based on a data-driven model and evolutionary optimisation method. To test the performance of MPC in different seasons, a summer day in August and a winter day in March have been selected to represent a summer and winter scenario. Table 7-1 illustrates the settings of heating temperature set-points and design temperature ranges in MPC for summer and winter scenarios. These two scenarios are used to investigate the performance of MPC on a residential heating system in the two seasons. In winter scenario, the building requires intensive heating for thermal comfort. While in the summer scenario, few amounts of heating is needed. The results illustrate the MPC improves the indoor thermal comfort compared with original real controls in both scenarios.

Table 7-1 Summer and winter scenarios for testing MPC

Scenario	Heating temperature set-points during the day (°C)	Heating temperature set-points during the night (°C)	Design temperature range for thermal comfort (°C)
Summer	19 – 23	17 – 19	22 – 26
Winter	19 – 23	17 – 19	18 – 23

In the winter scenario, MPC is tested on a winter day in March 2018. The aim of MPC is to allow more hours with the indoor temperatures falling into the design temperature range from 18 °C and 23 °C. In the MPC, the upper and lower temperature set-points are used as constraints for optimizing indoor temperatures. Through the temperature set-points, MPC has extra opportunities in turning on and off the heating in the residential building. The temperature set-points are between 17 °C and 19 °C during the night, 19 °C and 23 °C during the day. Figure 7-1 shows the comparison of indoor temperatures and temperature set-points between the MPC and real control in the winter test day with control per 15 minutes (total 96 time steps in a day). The results show that MPC provides a more comfortable indoor environment by increasing the number of hours within the design temperature range. First of all, room temperatures below 18 °C start to increase earlier with MPC, reducing hours of room temperatures under the lower limit. Secondly, MPC also reduces the chance of room temperatures over the upper limit. In addition, living room temperatures can exceed 23 °C in some time steps with real control, but such peaks are not found with MPC. Finally, the control with MPC has fewer fluctuating changes in the indoor temperature over the time horizon.

In the summer scenario, MPC is tested on a winter day in August 2018. The aim of MPC is to achieve more indoor comfort by reducing hours of indoor temperatures out of the design temperature range from 22 to 26 °C. The temperature set-points are constrained between 17 °C and 19 °C during the night, 19 °C and 23 °C during the day. Figure 7-2 illustrates a summer scenario with the room temperature controlled with and without MPC. The results show that MPC is capable of constraining the indoor temperatures within the design temperature range in more time steps. Without MPC, room temperatures are between 25 and 27 °C. There are many hours with room temperatures over 27 °C. By applying MPC, room temperatures fall below 26 °C in all three rooms. In the living room, it also reduces the chance of low temperatures under 22 °C. The results present that MPC maintains indoor temperatures well by constraining indoor temperatures not only below the upper limit (26 °C) but also above the lower limit (22°C).

From the scenarios, MPC proves to maintain the indoor temperatures within the design temperature range. It can also reduce the fluctuation of indoor temperatures by changing room temperatures slowly. To optimise indoor thermal comfort, MPC takes advantages

of forecasting models to foresee the change of temperatures and optimises the controls per 15 minutes during the day.

Winter scenario
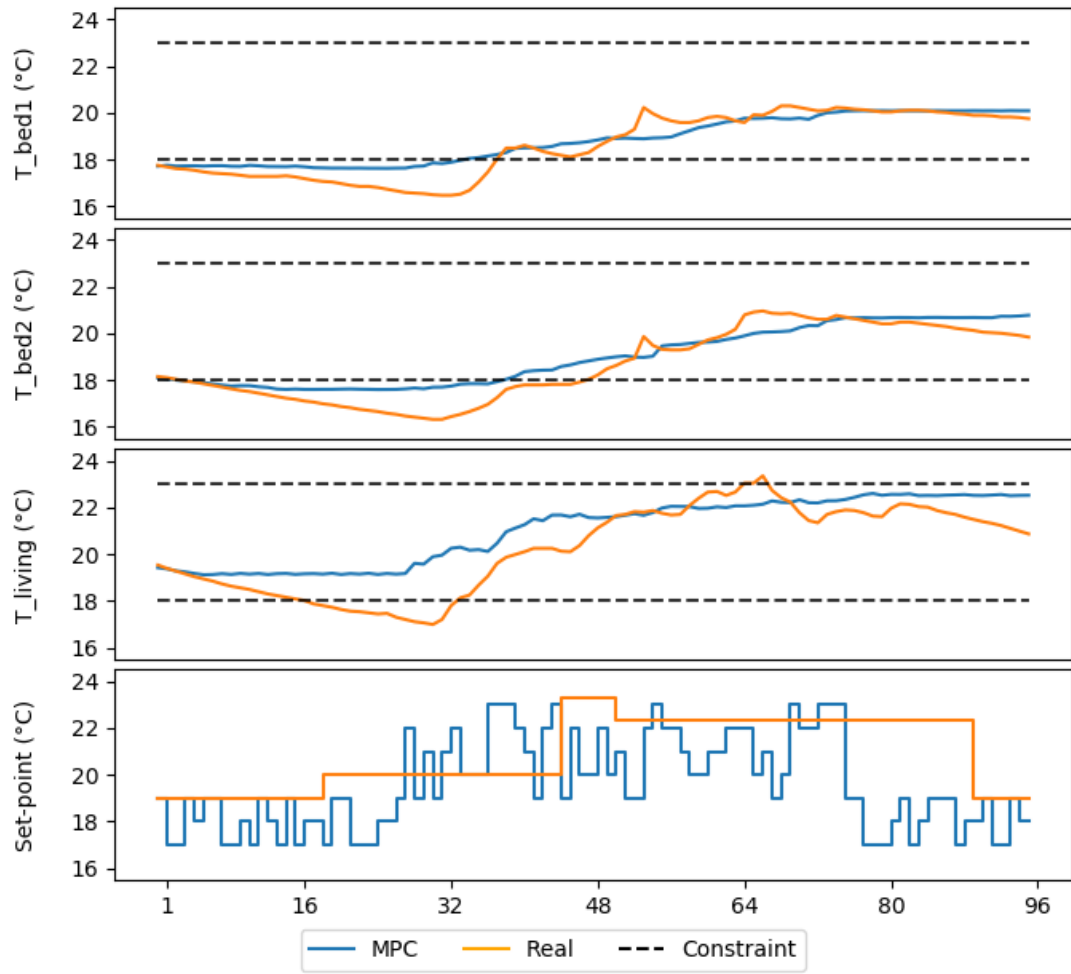
Figure 7-1 Comparison between MPC and real control for the day ahead controls of a residential heating system and the resulting room temperatures on a winter day.

Summer scenario
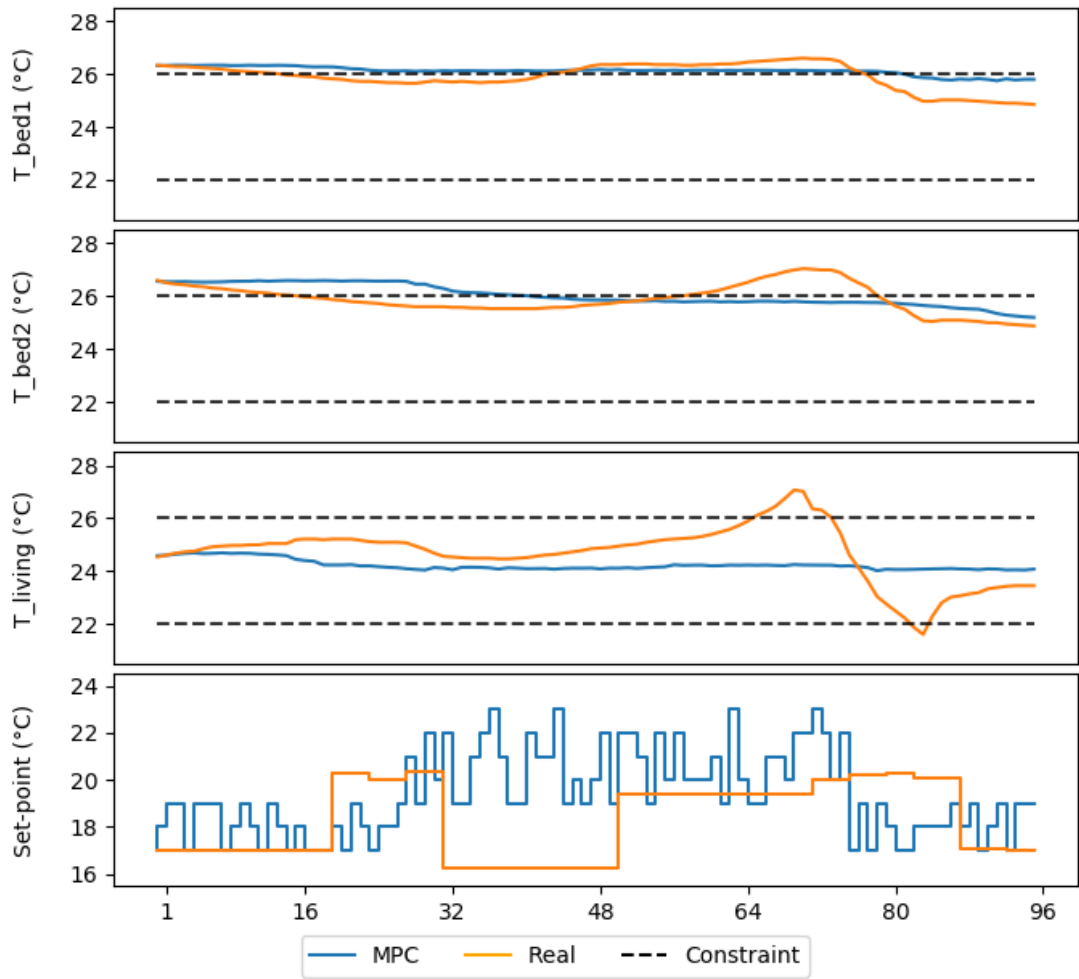
Figure 7-2 Comparison between MPC and real control for the day ahead controls of a residential heating system and the resulting room temperatures on a summer day.

7.3 Commercial building

7.3.1 Time of use

ToU optimisation is one of the most traditional DR by shifting loads from high-price intervals to low-price intervals. It has two objectives: (1) minimize the energy cost; (2) maintain indoor thermal comfort. As the HVAC system consumes a large part of the total energy in the building, ToU use case will consider the optimisation of the HVAC system in the office building. With the ToU, it provides different energy rates in a day. Load shifting can be implemented by rescheduling the operation of the indoor units to achieve the minimum energy cost. Thermal comfort can be maintained by restricting the indoor temperature within the comfort temperature zone.

The daily energy prices for each ToU period are listed in Table 7-2. While the energy rate for office building has been shown in Figure 7-3, indicating the energy price rate in (1) winter from October to April; (2) and summer from April to October.

Table 7-2 Prices of energy in each of the ToU periods.

ToU periods	Prices of energy (€/kWh)
Peak hours	0.105507
Partial-peak hours	0.091094
Off-peak hours	0.062543
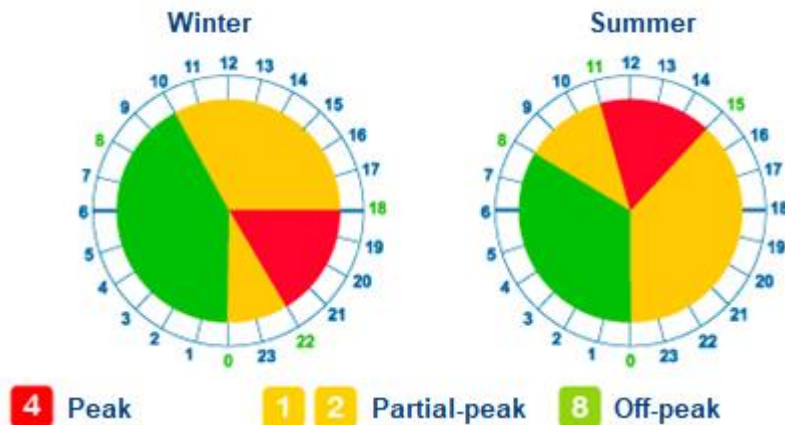
Figure 7-3 Hourly energy tariff in winter and summer in different hours.

7.3.2 Maximum power control (kWmax)

kWmax has cross-effects with ToU as the energy rate can be influenced when the maximum power is exceeded. Besides the ToU for energy consumption, the kWmax is another use case about penalties for power demand. The penalties are that a higher rate of the energy price will be charged when the power exceeds the maximum power in the contract. The current contract that the office signed with the energy grid indicates the

maximum power of 150 kW in each time period. The demand charges based on the contract are demonstrated in the following table.

Table 7-3 The demand charges for kWmax use case.

Maximum power (kW)	The rate of demand charges
kWmax ≤127.5	127.5kW* (# of days in month) *(Time period rate)
127.5 < kWmax < 157.5	(Max power demanded in month) *(# of days in month) *(Time period rate)
kWmax > 157.5	[(Max power demanded in month) +2*(Max power demanded in month-157.5kW)] *(# of days in month) *(Time period rate) New Time period rate: Peak hours: 0.111586 Partial-peak hours: 0.066952 Off-peak hours: 0.044634

7.3.3 Objective function

The objective is to find out the optimised future control signals (HVAC mode, on/off and temperature set-point) to minimise the energy cost of the HVAC. As the energy cost is the result of the energy price rate and energy consumption, the future total cost can be calculated through aggregating the energy cost at each time step in future. The variation of the energy price comes from the use case Time of Use (ToU), which makes the MPC consider not only energy consumption but also the energy tariff.

The optimisation problem can be illustrated as:

$$min \sum_{k=1}^{N} (E_k \times R_k) \qquad (1)$$

s.t.
$$P_k \leq P_{max}$$

$$T_{lb} \leq T_k \leq T_{ub}$$

Equation 1 shows the objective function for the MPC. It demonstrated that the total energy cost from time step k to N should be minimised. At time step k, the energy cost is equal to energy consumption(E_k) multiplying the rate of the energy price (R_k). There are two constraints: (1) the power (P_k) is constrained to be less than maximum power (P_{max}) and indoor temperature (T_k) is within the comfort temperature band between lower bound temperature (T_{lb}) and upper bound temperature (T_{ub}). As there are 9 indoor unit thermostats, each monitored indoor temperature should be considered (T = [T_1, T_2, T_3, T_4, T_5, T_6, T_7, T_8, T_9]; T_1 to T_9 are indoor temperature from each thermostat).

7.3.4 Constraints

Use case kWmax provides a combination of the energy tariffs for building consuming different maximum power in a month. The contract signed with the grid allows the building to have a flexible energy rate. For example, the current contract mentioned maximum power is 150 kW, and the hourly energy price is € 0.105507 for peak hours. However, if the maximum power is over 157.5 kW, the hourly rate is increased to € 0.111586.

From kWmax use case, we found that maximum power has a significant influence on the energy cost. It provides the flexibility for demand response for shifting the power from peak periods to low periods. As the MPC is based on a floor-level model, the global power for the 4th floor is used as the constraint. The historical data shows that the maximum power can be up to 20000 W in the summer days. Equation 2 shows the maximum power of 10000 W has been set as the constraint for MPC to investigate if the MPC can reduce or shift the maximum power in future time steps.

$$P_k \leq 10000 \qquad (2)$$

Indoor temperature is the most important factor regarding the thermal comfort of the occupants. Therefore, the prioritised target for the user or building manager is to optimise the control of the indoor unit to maintain the indoor temperature within a comfortable range.

As the indoor temperature is the most concerned thermal parameter, out of the comfort zone can result in health and safety issue. Therefore, the MPC is aimed at maintaining the indoor temperature within the lower and upper indoor temperature. In this case, Equation 3 illustrated the lower and upper temperatures are set up to 20 °C and 26 °C for temperature constraint for each room/space.

$$20 \leq T_k \leq 26 \qquad (3)$$

$$T_k = [T_1, T_2, T_3, T_4, T_5, T_6, T_7, T_8, T_9]$$

Where T_k represent the indoor temperature for each unit as listed in the following table. The variable T_k (from T_1 to T_9) represents the indoor temperature from room thermostat sensor U2-10 to U2-18 on the floor plan.

7.3.5 Optimisation algorithm

Optimisation of the control on indoor units for future time steps is built based on the system forecasting model, objective and constraints. The evolutionary algorithm has

been used as the optimisation method for optimising the future control of the indoor units based on the data-driven model. It can generate a group and evolve solutions generation by generation. The equally optimal solution is called Pareto non-dominated front. The selection of the optimal solution of the Pareto front is dependent on the objectives and constraints.

The evolutionary algorithm initialises the population based on the identified type and range of the manipulated inputs. With the system model, the output variables such as next time step energy are forecasted. The forecasted results at time step k can be used as the inputs for time step k+1. Repeatedly, the model can forecast the next N (N = 96 for a day ahead forecasting) time step system states. The sorting process is based on the non-domination criteria of the population. The objective of the optimisation algorithm is to improve the adaptive fit of a group of optimal solutions to a Pareto front.

7.3.6   Results and discussion

The MPC has been developed based on a data-driven model and evolutionary optimisation method. The aim is to minimize the energy cost of the commercial HVAC system. In order to test the performance of MPC on the heating and cooling periods, a summer day in July and a winter day in February have been selected. Since different temperature set-points and modes may cause a great difference in energy use, different scenarios are used to show different settings of temperature setpoints and modes in the MPC in the two test seasons, as shown in Table 7-4.

Table 7-4 Scenarios for testing MPC with different settings in winter and summer days

Scenario	Season	Temperature set-points at working hours (°C)	Temperature set-points at non-working hours (°C)	Mode at non-working hours (1 for heating, 0 for OFF, -1 for cooling)
1	Summer	20 – 26	20 – 26	-1
2	Summer	20 – 22	20 – 22	-1
3	Summer	20 – 26	20 – 28	0
4	Winter	20 – 26	20 – 22	1
5	Winter	20 – 26	20 – 22	0

It is important to find how MPC minimizes cost by shifting energy and power while maintaining indoor temperatures. In the listed 5 scenarios, the performance of MPC is investigated by looking into the resulting energy, power and indoor temperatures.

In scenario 1, MPC is applied with a wide range of temperature set-points from 20 °C to 26 °C and mode of -1 for turning on indoor units for cooling on a summer day. Figure 7-4 compares MPC with real control by looking into the resulting floor-level energy, power and room temperatures in the next 96 time steps. The 96 time steps account for 24 hours as each time step represents 15 minutes. Compared with real control, the peak energy

128

consumption and power can be significantly reduced during the day. However, during the off-peak hours, such as time step from 0 to 20, the energy consumption with MPC is slightly more than real control. This is due to the energy consumption for pre-cooling in these time steps to maintain indoor temperatures. The results show indoor temperature can be well maintained as most of the rooms have temperatures below 26 °C with the MPC. In addition, the results also show MPC changes the indoor temperature in a more stable way than real control. With real control that relies on temperature setpoints to condition the space, the mode of indoor units is constantly -1, meaning indoor units are always turned on for cooling in the working hours. While during the night, the energy consumption is close to zero as indoor units are turned off at non-working hours. Results show MPC can shift load from peak hours to lower-load hours by turning on and off the indoor units at different times.

In scenario 2, temperature set-points with 20 and 22 °C for lower and upper limits have been given in MPC for cooling on a summer day, as shown in Figure 7-5. Compared with scenario 1, it is a narrower range of temperature set-points. Because there are some indoor units, for example, indoor unit U211, showing that room temperatures are still out of the design comfort band. The guess here is that the upper limits of temperature set-points and comfort band are both 26 °C which may lead to inadequate cooling as no cooling will be offered when temperatures are 26 °C. The results show the temperature in indoor unit U211 is reduced quickly from the temperatures above the 26 °C after time step 32, allowing more hours getting into the design comfort band between 20 and 26 °C. The results also find a significant change of average temperatures, showing more hours are within the design comfort band. It means that indoor temperature across the floor is more likely to be with the comfort band, showing an overall improvement of thermal comfort.

In scenario 3, temperature set-points are set to be 20-26 °C between 6 am and 7 pm, but indoor units are turned off with the mode of 0 between 7 pm and 6 am on a summer day. As the building is normally closed between those non-working hours, cooling is not necessary during non-working hours. Although the upper limit of temperature set-point is 28 °C during non-working hours, no cooling is provided during these hours. This scenario is used to compare with scenario 1 to see if energy consumption can be reduced by turning off indoor units during non-working hours and how indoor temperatures are influenced. Figure 7-6 shows that the energy, power and indoor temperatures under MPC and real control. The results find that the peak energy and power have almost no change, but indoor temperatures from indoor units are slight higher during non-working hours. The average temperature also illustrates that indoor temperatures are nearly the

same as scenario 1 in working hours but slightly higher during non-working hours. The results also show that energy consumption is reduced during non-working hours.

In scenario 4, a winter day was tested with MPC by setting temperature set-points to 20-26 °C in working hours and 20-22 °C in non-working hours. Figure 7-7 shows the energy, power, indoor temperatures, temperature set-points and modes with and without MPC in scenario 4. Reducing the upper limit of temperature set-points from 26 to 22 °C is to reduce the heating times in non-working hours. The aim is to save energy but keep room temperatures at an acceptable level during the night in winter. The results show MPC significantly reduces peak energy consumption and power while improves indoor thermal comfort by constraining temperatures within the comfort band between 20 and 26 °C. First of all, MPC reduces peak energy and power by shifting loads, resulting in the maximum power under the constraint of 10000W. Secondly, MPC maintains indoor temperatures well by constraining them within the comfort band. Thirdly, compared with real control, MPC changes the indoor temperature in a smoother way. Finally, it finds real control has both heating and cooling modes on the winter day, which may lead to energy waste. MPC provides only one mode that is heating in winter days to reduce energy waste caused by frequently changing the mode.

In scenario 5, indoor units are turned off during non-working hours. Figure 7-8 shows that energy consumption and power are significantly reduced in the working hours, which are almost the same as scenario 4. There is a slight reduction in energy consumption in non-working hours. Also, the change in indoor temperatures is very small. Looking in to the indoor unit U211, it finds that the indoor temperature is slightly reduced. The reason for little change in indoor temperatures may be that the temperature set-points in non-working hours in scenario 4 are between 20 and 22 °C. It only turns on heating when indoor temperatures are below 20 or 22 °C. However, most of the indoor temperatures during these hours are already higher than 20 °C, resulting in a similar performance in both scenarios.

Scenario 1

Figure 7-4 The comparison between the MPC and real control in energy, power and indoor temperatures in scenario 1.

Scenario 2

Figure 7-5 The comparison between the MPC and real control in energy, power and indoor temperatures in scenario 2.

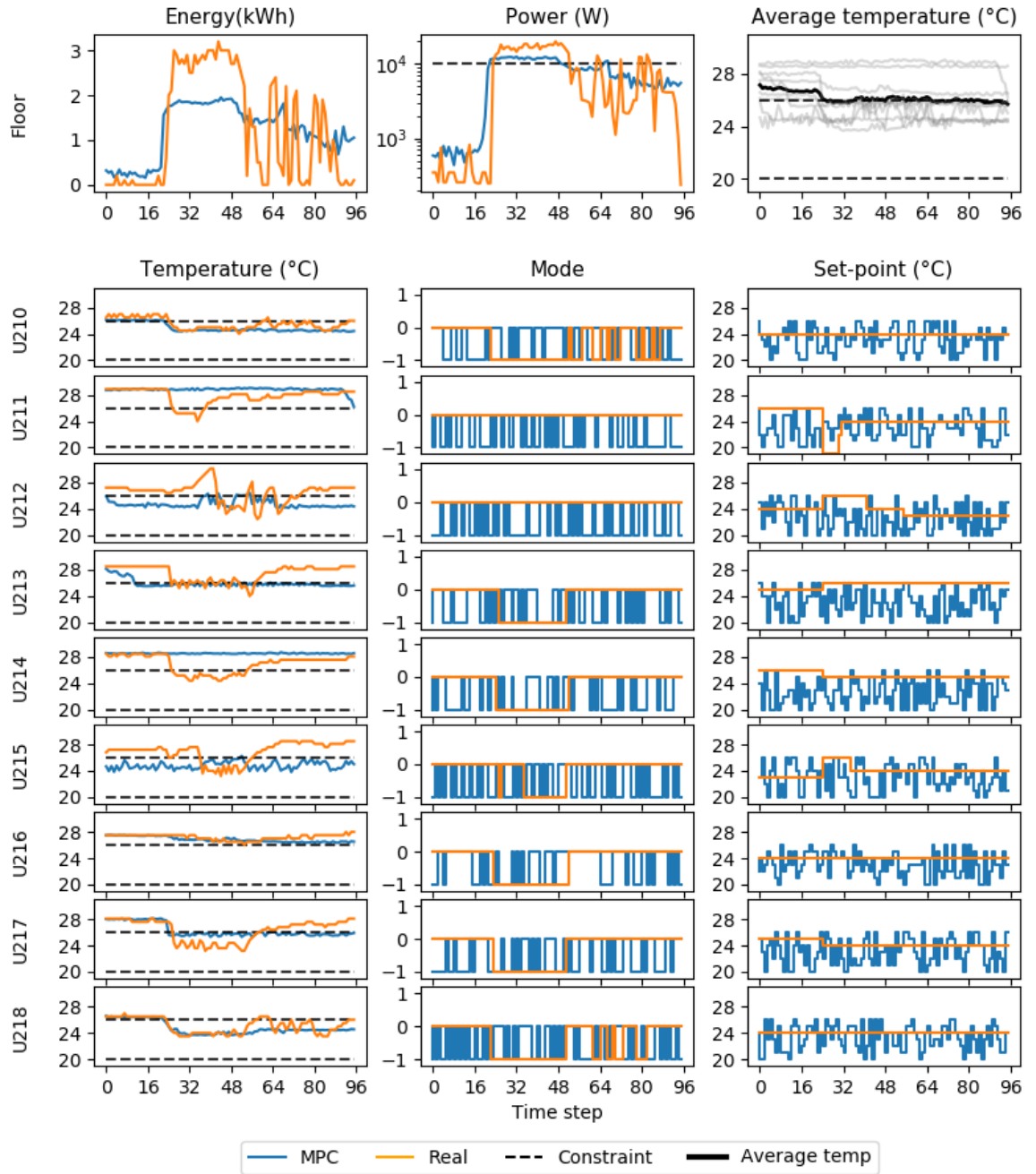## Scenario 3

Figure 7-6 The comparison between the MPC and real control in energy, power and indoor temperatures in scenario 3.

Scenario 4

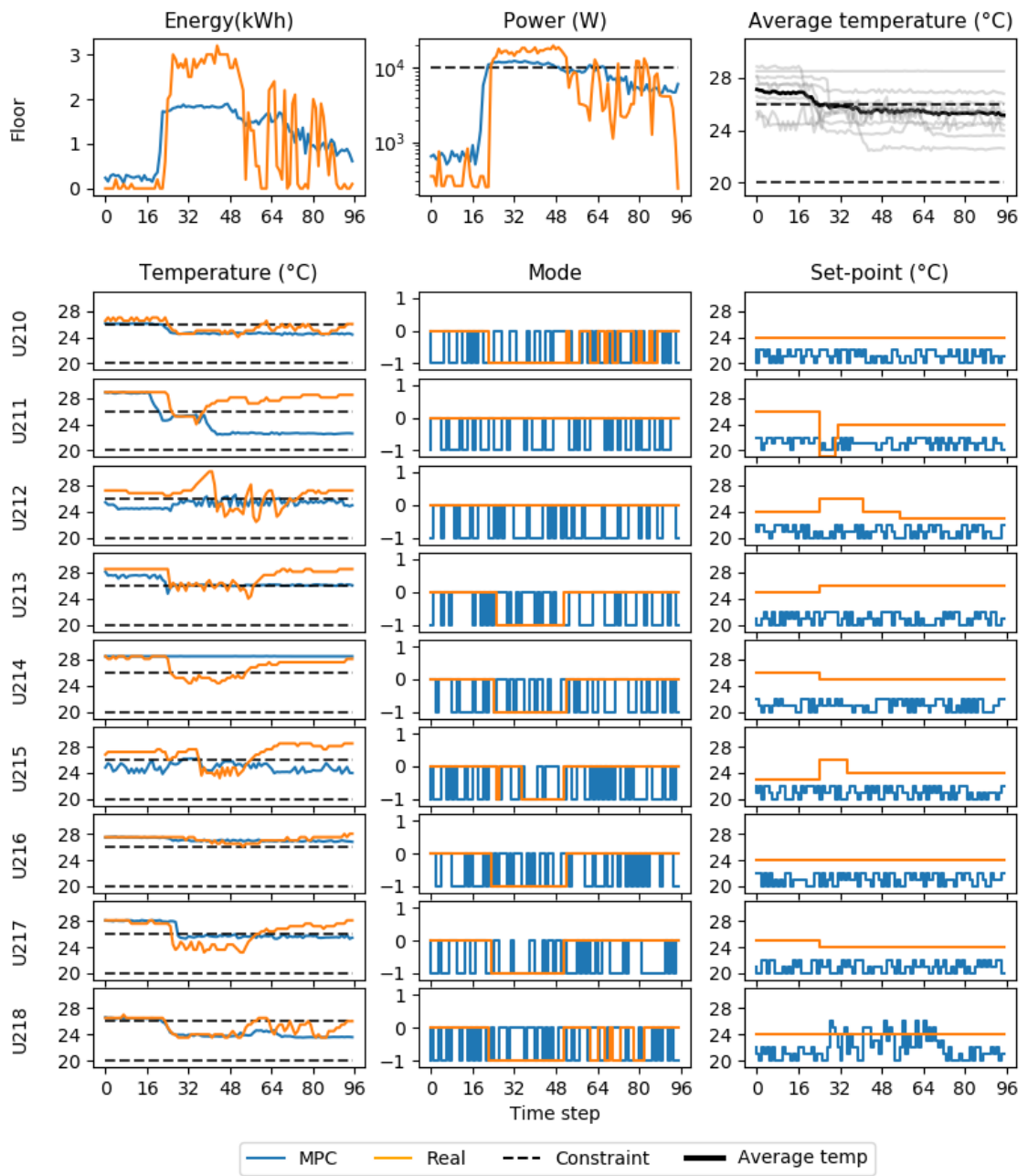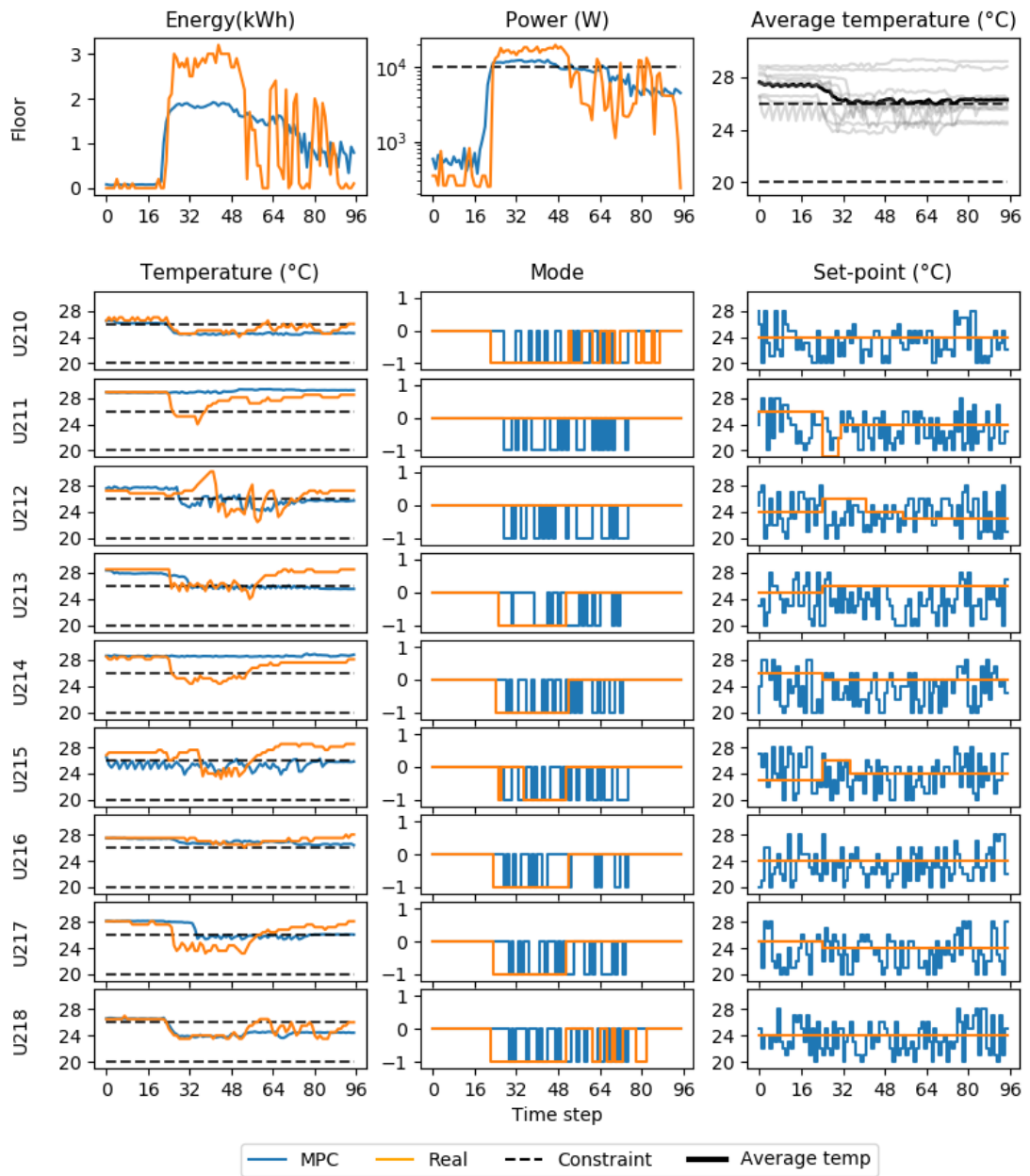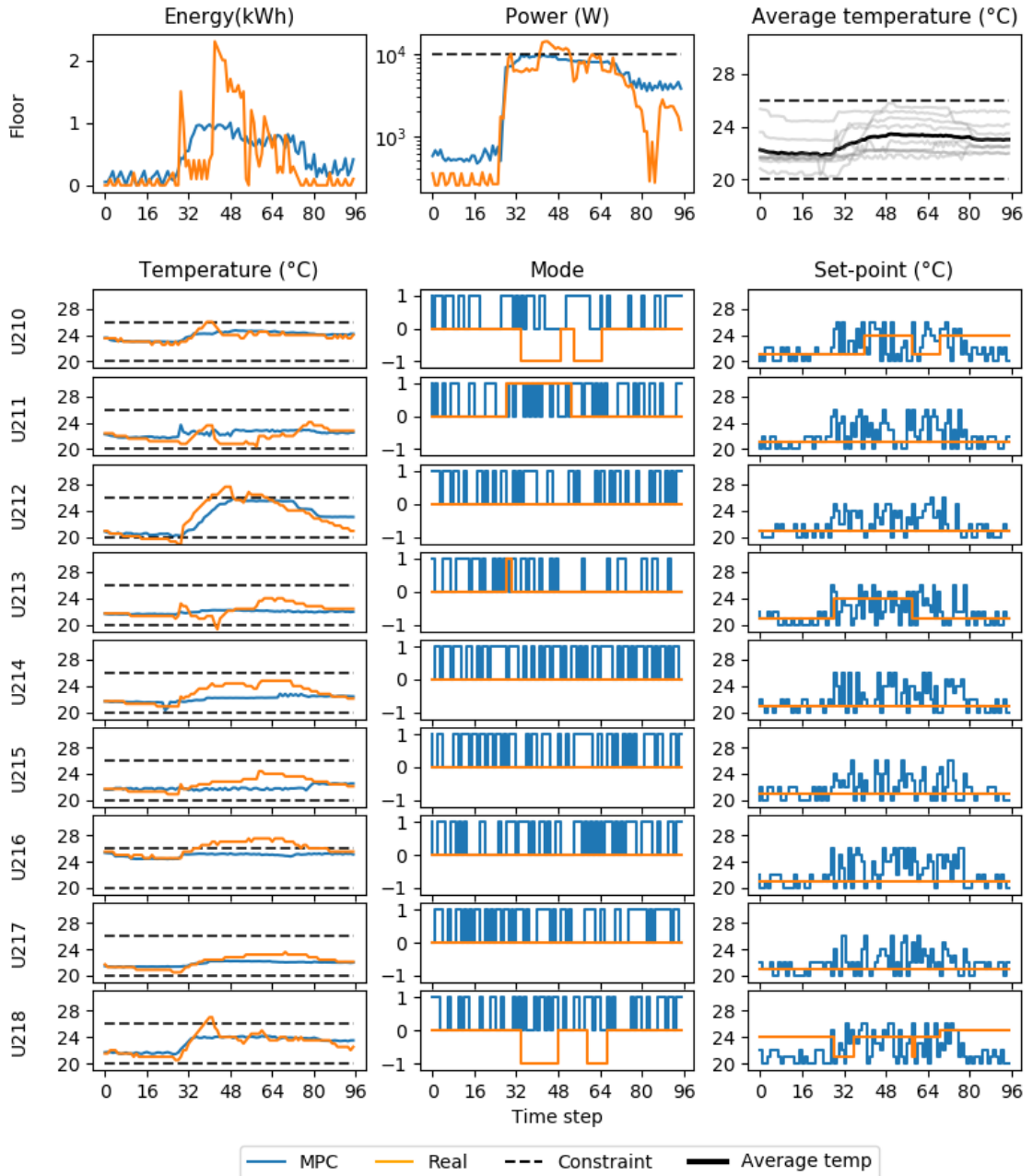Figure 7-7 The comparison between the MPC and real control in energy, power and indoor temperatures in scenario 4.
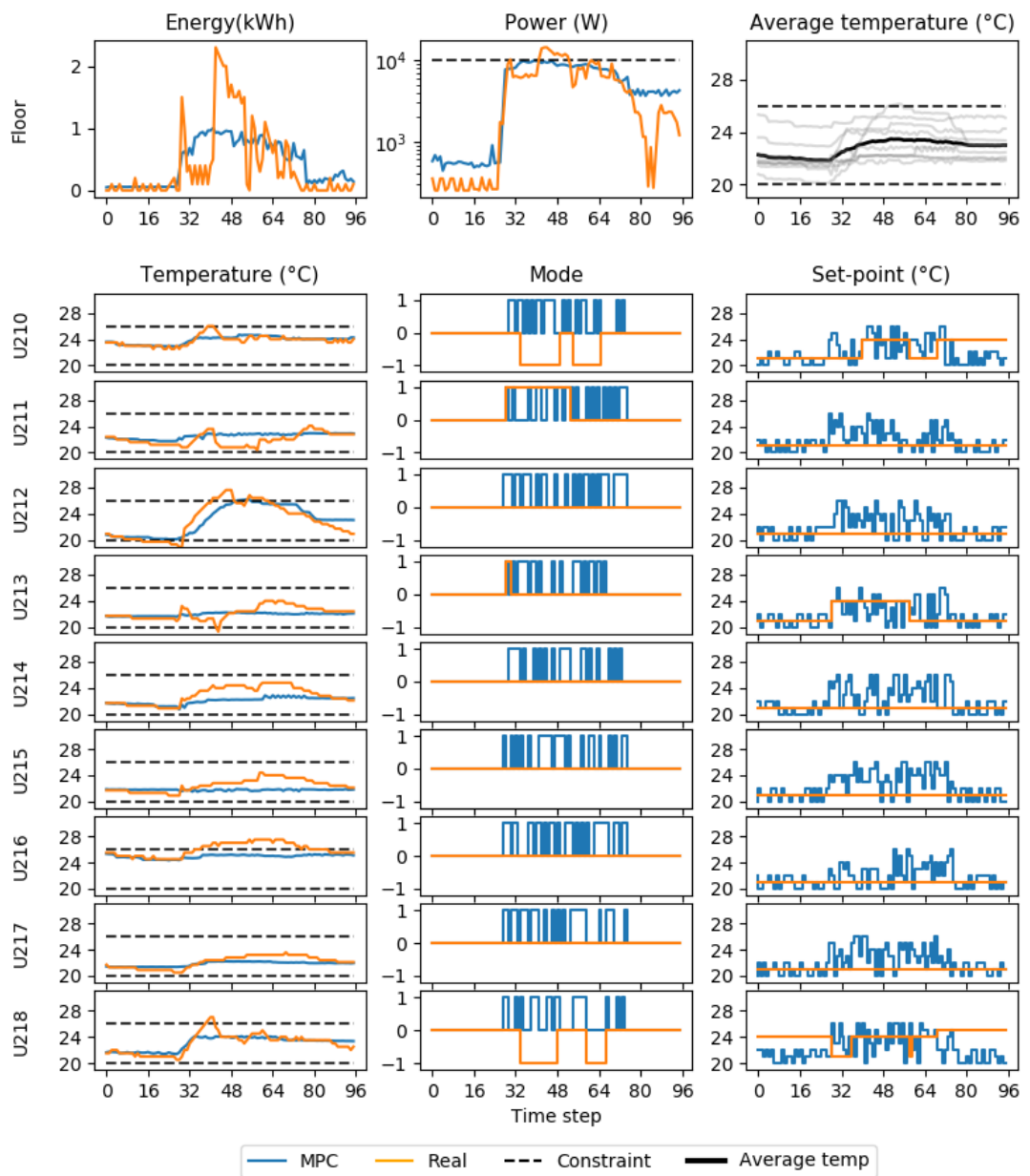
Scenario 5

Figure 7-8 The comparison between the MPC and real control in energy, power and indoor temperatures in scenario 5.

Time steps for optimisation

The objective of the MPC is to minimise the energy consumption in N time steps while N = 1 is for 15 minutes and N = 96 is for a day. In order to investigate the influence of the time steps on the MPC, different time steps for optimisation are tested on energy consumption, power and average indoor temperature in scenario 5.

Figure 7-9 presents the results of the MPC for different time steps from 16 to 96 on energy consumption. The mpc_16 represents the MPC for 16 timesteps, and mpc_96 represents the MPC for 96 timesteps. From the real energy consumption, the peak energy consumption occurs between the time step 32 and 64. It is important to find if the MPC with more than 64 timesteps (mpc_64, mpc_80 and mpc_96) would have a significant influence on the peak energy consumption. The average energy consumption per time step for mpc_64 and mpc_80 are nearly the same with the value of 0.82 kWh, compared with mpc_96 with the value of 0.84 kWh. Compared to the mpc_96, the mpc_64 and mpc_80 have less energy used during peak time. Especially for time step between 64 and 80, the mpc_96 estimates an energy consumption of 9.2 kWh but the mpc_80 estimates 5.9 kWh, which is much less. The reason could be that the mpc_96 tries to shift more load in the previous few time steps resulting in more estimated load after timestep 64 although the optimisation with more steps cannot overtake others with fewer optimisation steps. However, more optimisation steps would involve the optimisation or load shifting for more hours, covering the peak and low load period.

Figure 7-10 shows the results of the MPC for different time steps from 16 to 96 on power. The results demonstrate that different time steps have little influence on power consumption. After time step 32, all the MPC (mpc_32, mpc_48, mpc_64, mpc_80 and mpc_96) can maintain the power under 10000 W.

Figure 7-11 illustrates the results of the MPC for different time steps from 16 to 96 on the average indoor temperature. The mpc_96 shows that it tries to shift load earlier than other MPC with fewer optimisation steps, resulting in quick response in increasing the temperature. Besides, the change of the indoor temperature is smaller than mpc_64 and mpc_80. However, all of them can keep the indoor temperature within the comfort zone with a temperature range from 20 °C and 26 °C.
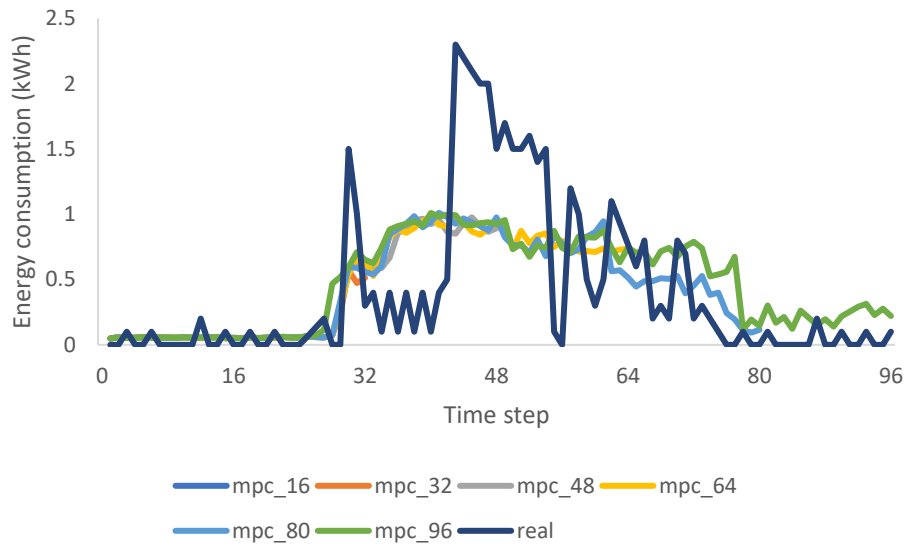
Figure 7-9 Energy consumption with the MPC and real in different time steps.
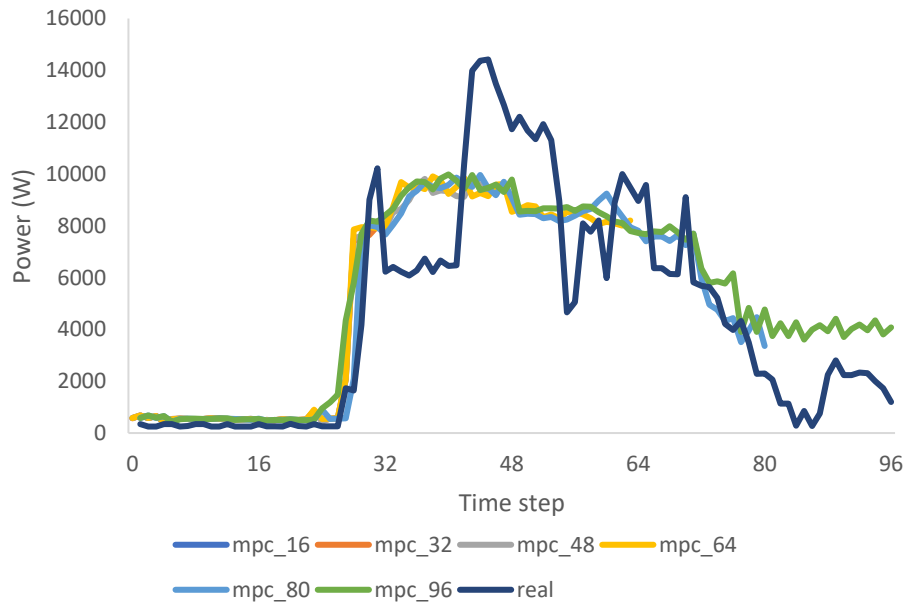
Figure 7-10 Power with the MPC and real control in different time steps.
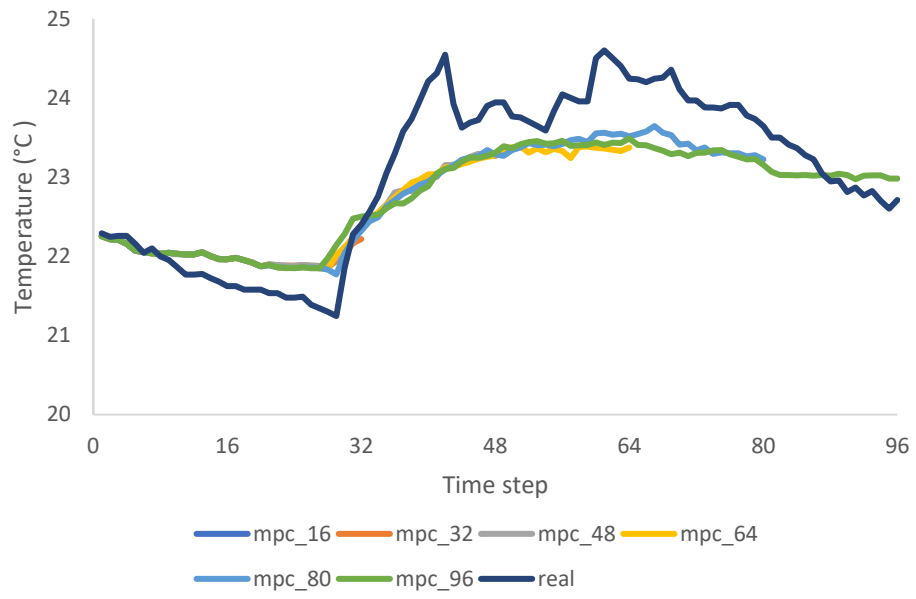
Figure 7-11 Average indoor temperature with the MPC and real control in different time steps.

Execution time

The development and testing for MPC were carried on one computer with Intel® Core (TM) i7-6700 Processor 3.4 GHz and 8 GB DDR4, 2133 MHz memory. The computing time for MPC has been illustrated in Figure 7-12. Previous MPC scenarios are based on 100 optimisation loops. With the same optimisation loops, the computing time for 96 timesteps is around 7 minutes. Four-hour (16 timesteps) MPC requires about one and a half minutes. The MPC with 100 optimisation loops using EVOLUTIONARY can significantly reduce the energy consumption, at the same time, constrain the indoor temperature and power within the given constraints.

Figure 7-13 shows the computing time for 96 timesteps MPC increases dramatically when the optimisation loops run more than 200 times. The MPC can consume more than 70 minutes with 1000 loops.
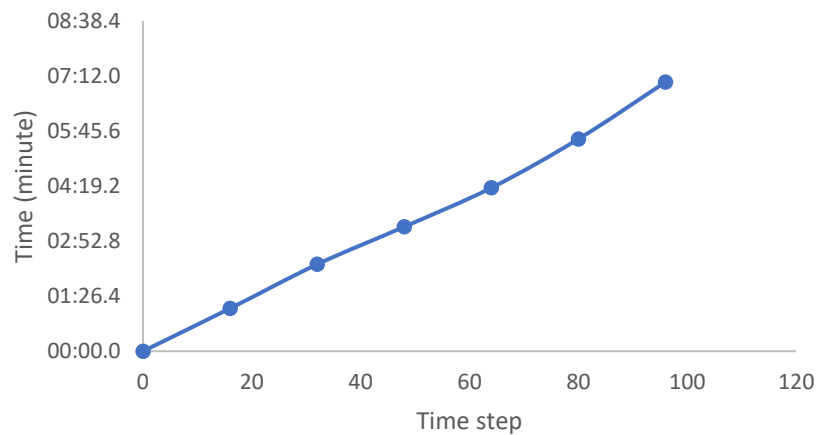
Figure 7-12 Computing time for MPC running for different time step with 100 optimisation loops.
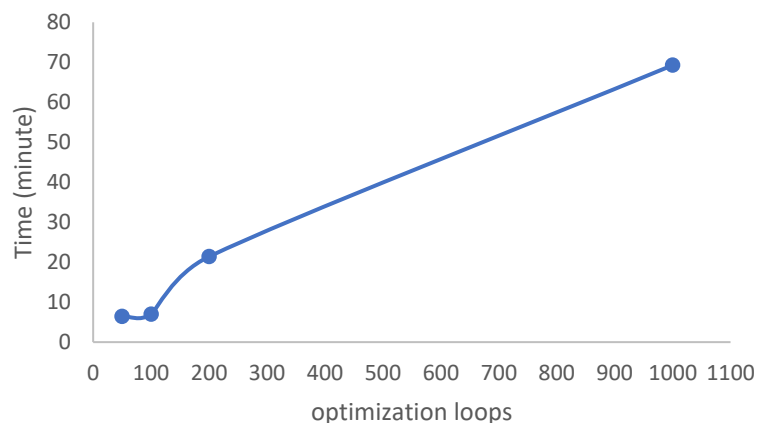
Figure 7-13 Computing time for 96 timesteps MPC with different optimisation loops.

7.4 Summary

This Chapter detailed the development of MPC optimiser, which provides day-ahead schedules for the HVAC systems. Several key aspects related to the MPC are discussed in this chapter. The MPC has been developed based on an evolutionary optimisation method and is suitable to solve energy optimisation based on data-driven models. The MPC was developed to control the HVAC system to ensure indoor thermal performance. The results show that it can efficiently limit indoor air temperatures within the given thermal comfort range.

Data-driven models have no algebraic function, most linear and non-linear optimisation method cannot be used to solve the optimisation problem. Due to its derivative-free, an evolutionary optimisation method is suitable to solve optimisation based on the data-driven model. In the residential building, the MPC is developed to control the heating system. The results show that it can improve thermal comfort by allowing more temperatures falling in the comfort temperature range. In the commercial building, the MPC is developed to control 9 indoor units on the 4th floor in a commercial building. It is found that it can efficiently shift the peak load and power to lower periods. At the same time, the indoor temperature and power are well constrained with our identified constraints. With the data-driven MPC controller, the peak energy consumption could be reduced by up to 36% and 15% for the peak power.

Chapter 8   Conclusion and contribution

This chapter revisits the proposed research questions and hypothesis in Chapter 1 and summarises the work and findings to answer them. Besides, a summary of key contributions to the body of knowledge will be presented. Following this, the limitations of the study will be discussed.

8.1 Main research findings

This section starts by answering the research questions raised in Chapter 1. Each research question will be addressed based on the findings from the previous chapters. The research hypothesis will be discussed in the following sub-section.

8.1.1   Increasing overheating risk

The first question was:

What is the impact of climate change on existing buildings?

This research question was identified in the literature review in Chapter 2 and further discussed in Chapter 4. Section 2.1 reviewed the impact of climate change and increasing overheating risk on buildings and occupants, shedding light on increasing overheating risk in non-air-conditioned buildings and the benefits of bringing controls to building elements such as windows. Chapter 4 carried out the study of overheating risk of UK dwellings in current and future climates, enhancing that control of window opening for natural ventilation is more effective than fixed scheduled natural ventilation. Additionally, it proved that natural ventilation is not enough to meet cooling demand, requiring additional cooling, such as installing HVAC systems for air conditioning in future climates due to climate change. The findings emphasised that HVAC installation could expand to meet higher cooling demand in future, and effective control strategies should be applied to HVAC operation for energy conservation and thermal comfort in buildings.

8.1.2   Potential of the sensor network and IoT

The second question was:

What are the benefits of using the sensor network and IoT for HVAC controls?

Chapter 2 presented the importance of building energy metering and environmental monitoring for improving the building environment and reducing energy cost. Additionally, it also suggested the data-driven models for developing MPC for the optimisation of HVAC operation. Collecting energy and environmental data was a basic need for developing data-driven models for the systems and buildings. A low-cost sensor network was built in the residential case study building for collecting high-resolution data,

comparing with the BMS system in the case study commercial building. Chapter 5 presented the data collected from the sensor network and analysed the data for understanding the systems and occupants behaviours before applying the data-driven approaches. Information such as electricity profile and temperature patterns can help to validate the forecast of energy and indoor environment. It also demonstrated IoT devices such as Nest thermostats for controlling indoor temperatures. Such a smart device can be integrated within the sensor network monitoring system, providing a framework for developing data-driven models and experimenting with MPC in real buildings in Chapter 6 and 7. Therefore, the sensor network and IoT provide an opportunity in developing a low-cost monitoring and control system for applying advanced control strategies such as MPC for HVAC control.

8.1.3   Data-driven modelling for MPC

The third question was:

> *How are HVAC models and tools used for simulating the behaviour of HVAC systems while considering their scope and limitations?*

Having demonstrated the potential energy saving that could be achieved by optimising the control for operational systems such as HVAC systems in buildings, the focus then turns to discuss modelling techniques for HVAC systems. This research question was discussed in Chapter 2, which has reviewed widely-used modelling techniques for HVAC operation and highlighted the data-driven approach was better than physics-based and hybrid methods for optimation-based operational control of the HVAC systems. The limitations for physics-based and hybrid approaches are the complexity of modelling and cost in time and computing, which prevent MPC optimisation from completing the modelling for hundreds and thousands of runs to have an optimal solution.

Data-driven models for HVAC systems were developed in Chapter 6, illustrating the forecast of indoor temperatures, energy and power in two case-study buildings. The forecasting accuracy could reach more than 0.99 in R^2 values for indoor temperatures, 0.92 for energy and power. Section 6.4.2 presents that machine learning algorithms are good at forecasting indoor temperatures, with R^2 values of 0.99 and 0.86 for residential and commercial indoor temperatures. Compared with indoor temperatures, the forecast of peaks of power and energy are less accurate, with R^2 values of around 0.6. Section 6.4.2 found that data-driven models are capable of forecasting day-ahead (96 timesteps) temperatures, energy and power, providing supports for the optimation by using the MPC in Chapter 7.

8.1.4 Optimisation for near real-time control for HVAC

The fourth question was:

> *Can the optimisation of HVAC operational control lead to indoor comfort, energy efficiency and cost reduction?*

Chapter 6 tested the high performance of data-driven models for forecasting energy and indoor environment. Then Chapter 7 expanded the study to the optimisation of HVAC operational control in buildings for comfort and energy efficiency. Section 7.2 proposed an MPC to optimise the heating system for reducing the fluctuation of temperatures in a residential building to achieve additional thermal comfort. The results showed that the optimisation through MPC could start heating in advance to allow the indoor temperature to fall in the comfort temperature range in advance. Section 7.3 developed an MPC to optimise the HVAC system in a commercial building. It proved that MPC for HVAC operational control is effective in reducing energy consumption, managing maximum power, saving energy cost and sustaining thermal comfort.

8.1.5 Scalability of the data-driven MPC

The final question was:

> *Can the integration of data-driven methods and MPC ease the deployment of energy management and control strategies for control systems in buildings on a wider scale and aid building automation?*

Chapter 3 (methodology) demonstrated a low-cost building monitoring and control system, providing an opportunity of using a sensor network and IoT devices for optimising the control systems in buildings. The work completed in Chapter 6 and 7 show how the data-driven models and MPC can be developed for near real-time optimal control of the HVAC systems in a residential building and a commercial building. The case studies in both types of buildings illustrated that the approach of data-driven MPC could utilize the flexibility and optimise the operation of control systems for the targets of energy, cost and comfort in different types of buildings. Therefore, the question was answered that the approach is scalable and adaptable for control systems such as HVAC systems in buildings for the conservation of energy and cost without compromising thermal comfort.

8.1.6 Revisiting the Hypothesis

Research questions are linked to the hypothesis. After discussing the research questions in previous sections, The conclusion comes to validate the research hypothesis, which is restated here.

Data-driven based model predictive control can allow near real-time optimal control of the HVAC systems through integrating machine learning, the Internet of Things and automated control to enhance the building performance by tackling the uncertainties of changing weather and dynamic system behaviours, and integrating demand response.

By answering the research questions during the research, the hypothesis was verified. It can be concluded that the research hypothesis is true that the data-driven model predictive control can allow near real-time optimal control for the operation of the HVAC systems. Each stage of this research is a piece of study contributing to verifying the research hypothesis. Initially, it investigated the impact of future climates on the existing residential buildings and introduced the needs for optimisation of control systems in the building for energy efficiency and thermal comfort. Additionally, the research developed a building metering and environmental monitoring system for collecting high-resolution and high-quality data, preparing data for developing data-driven forecasting models for energy and indoor environment. Finally, MPC controllers were developed for near real-time operational control for the HVAC systems in both types of buildings for thermal comfort and energy cost reduction. The MPC integrated IoT for building monitoring and automated control, utilizing machine learning methods for learning and forecasting energy and environmental parameters, and integrating demand response to make use of the flexibility of load and elasticity of the demand.

8.2 Contribution to Knowledge

This research has proposed a data-driven MPC for near real-time control of HVAC systems for thermal comfort, power management and energy cost reduction. The contributions can be stated as below.

1. This study assesses the impact of climate change on UK dwellings that the overheating risks are increasing in future climates, identifying the need for optimization of control systems in the operation of buildings. Flat type dwellings can be overheated for the whole summer without opening windows for natural ventilation. Even with natural ventilation, overheating risk can be found in old dwellings built between the 1900s and 1950s in London and Birmingham.

2. This study integrates the sensor network and IoT techniques in a residential building to implement a building monitoring system and control devices wirelessly, illustrating the potential of deploying an automated control system and opportunities for optimized control using a data-driven approach.

3. This study develops data-driven forecasting models using different machine learning algorithms, including SVR, LSTM and RF, comparing their forecasting performance in predicting peaks and day-ahead temperature, energy consumption and power. The forecasts of the peaks of energy and power are less accurate than the forecasts of the peaks of indoor temperatures, with R^2 values of 0.92 compared to 0.6. SVR performs better than the other two machine learning algorithms in day-ahead forecasting.

4. This study develops data-driven MPC for taking into account weather forecasts, system behaviours, energy demand and energy tariffs in a residential building and a commercial building, providing near real-time optimal control in every 15 minutes.

5. This study demonstrates that the MPC can take advantages of energy tariffs and demand flexibility (by shifting load from the high-demand periods to low-demand periods) to save the total energy cost with the constraints of ToU, kWmax and comfort temperature range. With the data-driven MPC controller, the peak energy consumption could be reduced by up to 36% and 15% for the peak power.

8.3 Research limitations

The limitations of this research are outlined as follows:

1. Occupant behaviour is one of the biggest challenges that influence the accuracy of the forecasts. In residential buildings, users might change their behaviour which is often an unpredictable factor. For instance, occupants may change their preferences, such as opening windows for extra air. Such change can influence indoor temperatures as extra cooling or heating is needed to balance the heat exchange. Occupant behaviour is often unpredictable and not recorded in the data. Therefore, forecasting performance can be influenced by the change of occupant behaviour. More information from occupants may be needed to improve forecasting performance.

2. Another limitation of this study lies in the fact that the sensors for energy metering and environment monitoring can lose data due to the drop of the internet or have abnormal data from sensors. The sensor network is made by connecting different kinds of sensors from different suppliers. They are mostly wireless sensors, leading to interruption of signals sharing the same radio frequency, resulting in disconnection in monitoring and gaps in data. Frequent checks can reduce the chance of losing too much data due to the drop on the internet. Interpolation or imputation methods could be used to deal with missing data.

3. An additional limitation is due to the tuning of models. During the training stage, many factors will affect the accuracy, including the choice of inputs and length of the data. Especially in LSTM models, many hyper-parameters need to be tuned to find an optimal combination of hyper-parameters. Optimisation algorithms can be used to optimise the selection of parameters. For example, GA finds the optimised hyper-parameters to minimise the error by trying different combinations. Although optimisation has been applied to the training process, the hyper-parameters may not necessarily be the optimal one due to the time and the selection of optimisation algorithms.

References

Afram, A. & Janabi-Sharifi, F., 2014. Review of modeling methods for HVAC systems. *Applied Thermal Engineering,* 67(1-2), pp. 507-519.

Afram, A. & Janabi-Sharifi, F., 2014. Theory and applications of HVAC control systems--A review of model predictive control (MPC). *Building and Environment,* Volume 72, pp. 343-355.

Afram, A., Janabi-Sharifi, F., Fung, A. S. & Raahemifar, K., 2017. Artificial neural network (ANN) based model predictive control (MPC) and optimization of HVAC systems: A state of the art review and case study of a residential HVAC system. *Energy and Buildings,* Volume 141, pp. 96-113.

Ahmad, M., Mourshed, M., Yuce, B. & Rezgui, Y., 2016. Computational intelligence techniques for HVAC systems: A review. *Building Simulation,* 9(4), pp. 359-398.

Ahmad, M. W. a. M. M. a. R. Y., 2017. Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption. *Energy and Buildings,* Volume 147, pp. 77-89.

Ahmad, M. W. et al., 2016. Building energy metering and environmental monitoring--A state-of-the-art review and directions for future research. *Energy and Buildings,* Volume 120, pp. 85-102.

Ahmed, M. et al., 2016. *Artificial neural network based controller for home energy management considering demand response events.* Putrajaya, IEEE.

Alvarez, F., Troncoso, A., Riquelme, J. & Ruiz, J., 2011. Energy time series forecasting based on pattern sequence similarity. *IEEE Transactions on Knowledge and Data Engineering,* 23(8), pp. 1230--1243.

Arshak, K. & Twomey, K., 2002. Investigation into a novel humidity sensor operating at room temperature. *Microelectron,* 33(3), pp. 213-220.

Arteconi, A., Hewitt, N. & Polonara, F., 2013. Domestic demand-side management (DSM): Role of heat pumps and thermal energy storage (TES) systems. *Applied Thermal Engineering,* 51(1), pp. 155-165.

Artmann, N., Manz, H. & Heiselberg, P., 2007. Climatic potential for passive cooling of buildings by night-time ventilation in Europe. *Applied energy,* 84(2), pp. 187-201.

Ascione, F. et al., 2016. Simulation-based model predictive control by the multi-objective optimization of building energy performance and thermal comfort. *Energy and Buildings,* Volume 111, pp. 131-144.

ASHRAE, 2015. *ASHRAE Handbook—HVAC Applications.* USA: American Society of Heating, Refrigerating and Air-Conditioning Engineers.

Aydinalp, M., Ugursal, V. I. & Fung, A. S., 2002. Modeling of the appliance, lighting, and space-cooling energy consumptions in the residential sector using neural networks. *Applied energy,* 71(2), pp. 87-110.

Azadeh, A., Ghaderi, S. & Sohrabkhani, S., 2008. Annual electricity consumption forecasting by neural network in high energy consuming industrial sectors. *Energy Conversion and Management,* 49(8), pp. 2272-2278.

Azadeh, A., Ghaderi, S., Tarverdian, S. & Saberi, M., 2007. Integration of artificial neural networks and genetic algorithm to predict electrical energy consumption. *Applied Mathematics and Computation,* 186(2), pp. 1731-1741.

Azadeh, A. & Tarverdian, S., 2007. Integration of genetic algorithm, computer simulation and design of experiments for forecasting electrical energy consumption. *Energy Policy,* 35(10), pp. 5229-5241.

Balan, R. et al., 2011. Parameter identification and model based predictive control of temperature inside a house. *Energy and Buildings,* 43(2), pp. 748-758.

Belic, F., Hocenski, Z. & Sliskovic, D., 2015. *HVAC control methods-a review.* Cheile Gradistei, 2015 19th International Conference on System Theory, Control and Computing (ICSTCC).

Ben-Nakhi, A. & Mahmoud, M., 2004. Cooling load prediction for buildings using general regression neural networks. *Energy Conversion and Management,* 45(13), pp. 2127-2141.

Blank, T., Eksperiandova, L. & Belikov, K., 2016. Recent trends of ceramic humidity sensors development: A review. *Sensors and Actuators B: Chemical,* Volume 228, pp. 416-442.

Braun, J. & Chaturvedi, N., 2002. An inverse gray-box model for transient building load prediction. *HVAC&R Research,* 8(1), pp. 73-99.

Buonanno, G., 2000. On field characterisation of static domestic gas flowmeters. *Measurement,* 27(4), pp. 277-285.

Cascetta, F. & Rotondo, G., 2015. Effects of intermittent flows on turbine gas meters accuracy. *Measurement,* Volume 69, pp. 280-286.

Ceylan, H. & Ozturk, H., 2004. Estimating energy demand of Turkey based on economic indicators using genetic algorithm approach. *{Energy Conversion and Management,* 45(15), pp. 2525-2537.

Chen, B. J., Chang, M. W. & lin, C. J., 2004. Load forecasting using support vector machines: A study on EUNITE competition 2001. *IEEE transactions on power systems,* 19(4), pp. 1821-1830.

Cheng, J. et al., 2018. Heatwave and elderly mortality: An evaluation of death burden and health costs considering short-term mortality displacement. *Environment international,* Volume 115, pp. 334-342.

Childs, P., Greenwood, J. & Long, C., 2000. Review of temperature measurement. *Review of scientific instruments,* 71(8), pp. 2959-2978.

Chua, K., Chou, S., Yang, W. & Yan, J., 2013. Achieving better energy-efficient air conditioning--a review of technologies and strategies. *Applied Energy,* Volume 104, pp. 87-104.

CIBSE, 2006. *Environmental Design, CIBSE Guide A..* 7th ed. London: CIBSE.

CIBSE, 2017. *Design methodology for the assessment of overheating risk in homes, TM59: 2017,* London: CIBSE.

DBEIS&Ofgem, 2013. *Smart meters: information for industry and other stakeholders,* London: Department for Business, Energy & Industrial Strategy and Ofgem.

DBEIS, 2018. *Smart meters: a guide,* London: Department for Business, Energy & Industrial Strategy.

Deb, C. et al., 2017. A review on time series forecasting techniques for building energy consumption. *Renewable and Sustainable Energy Reviews,* Volume 74, pp. 902-924.

Díaz, G., Sen, M., Yang, K. & McClain, R. L., 2001. Dynamic prediction and control of heat exchangers using artificial neural networks. *International Journal of Heat and Mass Transfer,* 44(9), pp. 1671-1679.

Dimitroulopoulou, C., 2012. Ventilation in European dwellings: A review. *Building and Environment,* Volume 47, pp. 109-125.

Dimitroulopoulou, C. et al., 2005. *Ventilation, air tightness and indoor air quality in new homes.* 1 ed. London: BRE Bookshop.

Dong, B., Cao, C. & Lee, S., 2005. Applying support vector machines to predict building energy consumption in tropical region. *Energy and Buildings,* 37(5), pp. 545-553.

Dounis, A. & Caraiscos, C., 2009. Advanced control systems engineering for energy and comfort management in a building environment—A review. *Renewable and Sustainable Energy Reviews,* 13(6), pp. 1246--1261.

Duff, M. & Towey, J., 2010. Two ways to measure temperature using thermocouples feature simplicity, accuracy, and flexibility. *Analog dialogue,* 44(10), pp. 1-6.

Edison Electric Institute, 2014. *Handbook for Electricity Metering.* s.l.:Edison.

Ekonomou, L., 2010. Greek long-term energy consumption prediction using artificial neural networks. *Energy,* 35(2), pp. 512-517.

Elliot, A. et al., 2014. Using real-time syndromic surveillance to assess the health impact of the 2013 heatwave in England. *Environmental research,* Volume 135, pp. 31-36.

EnergyPlus, 2020. *EnergyPlus.* [Online]
Available at: https://energyplus.net/
[Accessed 24 June 2020].

eQUEST, 2018. *eQUEST.* [Online]
Available at: http://www.doe2.com/equest/
[Accessed 24 June 2020].

Erol, A., Okur, S., Yağmurcukardeş, N. & Arıkan, M. Ç., 2011. Humidity-sensing properties of a ZnO nanowire film as measured with a QCM. *Sensors and Actuators B: Chemical,* 152(1), pp. 115-120.

ESP-r, 2020. *ESP-r.* [Online]
Available at: http://www.esru.strath.ac.uk/Courseware/ESP-r/tour/
[Accessed 24 June 2020].

European Commission, 2008. *EU energy and transport in figures : statistical pocketbook 2007/2008,* Luxembourg: Directorate-General for Energy and Transport, European Commission.

European Commission, 2012. *Analysis of options beyond 20% GHG emission reductions: Member State results ,* Brussels: European Commission.

European Commission, 2012. *Energy efficiency directive,* s.l.: European Commission.

European Commission, 2014. *A policy framework for climate and energy in the period from 2020 to 2030,* Brussels: European Commission.

European Committee for Standardization, 2002. *EN 12464-1:2002. Light and lighting. Lighting of work places. Part 1: Indoor work places,* Brussels, Belgium: BSI Standard Publication.

European Union, 2020. *Long-term low greenhouse gas emission development strategy of the EU and its Member States.* [Online]
Available at: https://unfccc.int/documents/210328
[Accessed 01 06 2020].

Fattaheian-Dehkordi, S., Fereidunian, A., Gholami-Dehkordi, H. & Lesani, H., 2014. Hour-ahead demand forecasting in smart grid using support vector regression (SVR). *International transactions on electrical energy systems,* 24(12), pp. 1650-1663.

Ferreira, P., Ruano, A., Silva, S. & Conceicao, E., 2012. Neural networks based predictive control for thermal comfort and energy savings in public buildings. *Energy and buildings,* Volume 55, pp. 238-251.

Ferreira, P., Ruano, A., Silva, S. & Conceicao, E., 2012. Neural networks based predictive control for thermal comfort and energy savings in public buildings. *Energy and Buildings,* Volume 55, pp. 238-251.

Foucquier, A. l. et al., 2013. State of the art in building modelling and energy performances prediction: A review. *Renewable and Sustainable Energy Reviews,* p. 272–288.

Foucquier, A. et al., 2013. State of the art in building modelling and energy performances prediction: A review. *Renewable and Sustainable Energy Reviews,* Volume 23, pp. 272-288.

Frontczak, M. & Wargocki, P., 2011. Literature survey on how different factors influence human comfort in indoor environments. *Building and Environment,* 46(4), pp. 922-937.

Fu, Y., Li, Z., Zhang, H. & Xu, P., 2015. Using support vector machine to predict next day electricity load of public buildings with sub-metering devices. *Procedia Engineering,* Volume 121, pp. 1016-1022.

Genet, J. & Schubert, C., 2011. *Designing a metering system for small and medium-sized buildings,* SE: Technical Report SEMED310007EN.

Givoni, B., 1994. *Passive low energy cooling of buildings.* New York: John Wiley & SonS.

Gonzalez, P. A. & Zamarreno, J. M., 2005. Prediction of hourly energy consumption in buildings based on a feedback artificial neural network. *Energy and Buildings,* 37(6), pp. 595-601.

Goyal, S. & Barooah, P., 2012. A method for model-reduction of non-linear thermal dynamics of multi-zone buildings. *Energy and Buildings,* Volume 47, pp. 332-340.

Goyal, S., Ingley, H. A. & Barooah, P., 2013. Occupancy-based zone-climate control for energy-efficient buildings: Complexity vs. performance. *Applied Energy,* Volume 106, pp. 209-221.

Greater London Authority, 2008. *The London Climate Change Adaptation Strategy,* London: Greater London Authority.

Green, H. et al., 2016. Mortality during the 2013 heatwave in England – How did it compare to previous heatwaves? A retrospective observational study. *Environmental research,* Volume 147, pp. 343-349.

Grozinger, J. et al., 2014. *Overview of Member States information on NZEBs: Background paper— Final Report,* s.l.: s.n.

Gupta, R. & Gregg, M., 2012. Using UK climate change projections to adapt existing English homes for a warming climate. *Building and Environment,* Volume 55, pp. 20-42.

Havas, M., 2006. Electromagnetic hypersensitivity: biological effects of dirty electricity with emphasis on diabetes and multiple sclerosis. *Electromagnetic Biology and Medicine,* 25(4), pp. 259-268.

Haykin, S., 1994. *Neural networks: A comprehensive foundation: Macmillan college publishing company.* New York: Macmillan college publishing company.

Holland, J., 1992. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence.* s.l.:MIT press.

Hong, W.-C., 2013. *Intelligent Energy Demand Forecasting.* New York: Springer.

Hu, J. & Karava, P., 2014. A state-space modeling approach and multi-level optimization algorithm for predictive control of multi-zone buildings with mixed-mode cooling. *Building and Environment,* Volume 80, pp. 259-273.

Hu, M., Xiao, F., Jorgensen, J. B. & Li, R., 2019. Price-responsive model predictive control of floor heating systems for demand response using building thermal mass. *Applied Thermal Engineering,* Volume 153, pp. 316-329.

International Code Council, 2016. *International Energy Conservation.* [Online]
Available at: http://www.iccsafe.org
[Accessed 04 04 2019].

Jiang, Y. et al., 2016. Outage management of distribution systems incorporating information from smart meters. *IEEE Transactions on Power Systems,* 31(5), pp. 4144-4154.

Kalogirou, S., 2006. Artificial neural networks in energy applications in buildings. *International Journal of Low-Carbon Technologies,* 1(3), pp. 201-216.

Kalogirou, S. & Bojic, M., 2000. Artificial neural networks for the prediction of the energy consumption of a passive solar building. *Energy,* 25(5), pp. 479-491.

Khan, Y., Khare, V. R., Mathur, J. & Bhandari, M., 2015. Performance evaluation of radiant cooling system integrated with air system under different operational strategies. *Energy and Buildings,* Volume 97, pp. 118-128.

Kim, M., Man, K. L. & Helil, N., 2019. Advanced Internet of Things and Big Data Technology for Smart Human-Care Services. *Journal of Sensors .*

Kleiber, C. E., 2017. Radiation from wireless technology elevates blood glucose and body temperature in 40-year-old type 1 diabetic male. *Electromagnetic biology and medicine,* 36(3), pp. 259-264.

Kolokotroni, M., Perera, M., Azzi, D. & Virk, G., 2001. An investigation of passive ventilation cooling and control strategies for an educational building. *Applied Thermal Engineering,* 21(2), pp. 183-199.

Kong, W. et al., 2017. Short-term residential load forecasting based on LSTM recurrent neural network. *IEEE Transactions on Smart Grid,* 10(1), pp. 841-851.

Kubota, T., Chyee, D. T. H. & Ahmad, S., 2009. The effects of night ventilation technique on indoor thermal environment for residential buildings in hot-humid climate of Malaysia. *Energy and buildings,* 41(8), pp. 829-839.

Kusiak, A., Li, M. & Tang, F., 2010. Modeling and optimization of HVAC energy consumption. *Applied Energy,* 87(10), pp. 3092-3102.

Kwadzogah, R., Zhou, M. & Li, S., 2013. *Model predictive control for HVAC systems—A review.* s.l., s.n.

Lai, D. et al., 2018. Ventilation behavior in residential buildings with mechanical ventilation systems across different climate zones in China. *Building and Environment,* Volume 143, pp. 679-690.

Lee, K.-H. & Braun, J., 2008. Evaluation of methods for determining demand-limiting setpoint trajectories in buildings using short-term measurements. *Building and Environment,* 43(10), pp. 1769-1783.

Lee, Y.-S. & Tong, L.-I., 2012. Forecasting nonlinear time series of energy consumption using a hybrid dynamic model. *Applied Energy,* Volume 94, pp. 251-256.

Li, D. H., Wan, K. K., Yang, L. & Lam, J. C., 2011. Heat and cold stresses in different climate zones across China: a comparison between the 20th and 21st centuries. *Building and Environment,* 46(8), pp. 1649-1656.

Li, P. et al., 2015. Simulation and experimental demonstration of model predictive control in a building HVAC system. *Science and Technology for the Built Environment,* 21(6), pp. 721-732.

Lomas, K. J. & Porritt, S. M., 2017. Overheating in buildings: lessons from research. *Building Research and Information,* 45(1-2).

Lomas, K. & Kane, T., 2013. Summertime temperatures and thermal comfort in UK homes. *Building Research & Information,* 41(3), pp. 259-280.

Luan, W., Irving, M. R. & Daniel, J. S., 2002. Genetic algorithm for supply restoration and optimal load shedding in power system distribution networks. *IEE Proceedings-Generation, Transmission and Distribution,* 149(2), pp. 145-151.

Mai, W. & Chung, C., 2016. *Model predictive control based on thermal dynamic building model in the demand-side management.* Boston, IEEE, pp. 1-5.

Ma, J., Qin, J., Salsbury, T. & Xu, P., 2012. Demand reduction in building energy systems based on economic model predictive control. *Chemical Engineering Science,* 67(1), pp. 92-100.

Manjarres, D. et al., 2017. An energy-efficient predictive control for HVAC systems applied to tertiary buildings based on regression techniques. *Energy and Buildings,* Volume 152, pp. 409-417.

Mavrogianni, A. et al., 2012. Building characteristics as determinants of propensity to high indoor summer temperatures in London dwellings. *Building and Environment,* Volume 55, pp. 117-130.

Mazzeo, D. et al., 2020. EnergyPlus, IDA ICE and TRNSYS predictive simulation accuracy for building thermal behaviour evaluation by using an experimental campaign in solar test boxes with and without a PCM module. *Energy and Buildings,* Volume 212, p. 109812.

McLeod, R., Hopfe, C. & Kwan, A., 2013. An investigation into future performance and overheating risks in Passivhaus dwellings. *Building and Environment,* Volume 70, pp. 189-209.

Met office, 2013. *Hot dry spell July 2013.* [Online]
Available at: https://www.metoffice.gov.uk/weather/learn-about/weather/case-studies/heat-wave-july2013
[Accessed 12 04 2020].

MHCLG, 2018. *Englsih Houseing Survey Headine Report, 2016-17.* London: Ministry of Housing, Communities & Local Government (MHCLG).

Michael, A., Demosthenous, D. & Philokyprou, M., 2017. Natural ventilation for cooling in mediterranean climate: A case study in vernacular architecture of Cyprus. *Energy and Buildings,* Volume 144, pp. 333-345.

Micro-Chip Technologies, 2010. *NTC Thermistors.* [Online]
Available at: http://www.microchiptechno.com/ntc_thermistors.php
[Accessed 7 4 2019].

Mitchell, G., Bahadoorsingh, S., Ramsamooj, N. & Sharma, C., 2017. *A comparison of artificial neural networks and support vector machines for short-term load forecasting using various load types.* Manchester, IEEE.

Mourshed, M., Kelliher, D. & Keane, M., 2003. Integrating building energy simulation in the design process. *IBPSA News,* 13(1), pp. 21-26.

Namdar, A. & Berenji, H., 2014. *A comparison of computational intelligence techniques for energy time series forecasting.* Beijing, IEEE.

Nasrollahi, N. & Shokri, E., 2016. Daylight illuminance in urban environments for visual comfort and energy performance. *Renewable and sustainable energy reviews,* Volume 66, pp. 861-874.

Neto, A. & Fiorelli, F., 2008. Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption. *Energy and buildings,* 40(12), pp. 2169-2176.

Nicol, F., 2013. *The Limits of Thermal Comfort: Avoiding Overheating in European Buildings: CIBSE TM52, 2013,* s.l.: CIBSE.

Omega, 2018. *Temperature measurement equipment.* [Online]
Available at: https://www.omega.co.uk/prodinfo/temperature-measurement.html
[Accessed 07 04 2019].

ONS, 2018. *Estimates of the population for the UK, England and Wales, Scotland and Northern Ireland.* [Online]
Available at:
https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationest imates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland
[Accessed 15 07 2018].

ONS, 2018. *Population estimates for the UK, England and Wales, Scotland and Northern Ireland: mid-2017.* [Online]
Available at:
https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationest imates/bulletins/annualmidyearpopulationestimates/mid2017
[Accessed 15 07 2018].

Pai, P. F. & Hong, W.-C., 2005. Support vector machines with simulated annealing algorithms in electricity load forecasting. *Energy Conversion and Management,* 46(17), pp. 2669-2688.

Pandharipande, A. & Caiced, D., 2015. Smart indoor lighting systems with luminaire-based sensing: A review of lighting control approaches. *Energy and Buildings,* Volume 104, pp. 369-377.

Park, D. et al., 1991. Electric load forecasting using an artificial neural network. *IEEE transactions on Power Systems,* 6(2), pp. 442-449.

Parker, S. A. et al., 2015. *Metering Best Practices. A Guide to Achieving Utility Resource Efficiency, Release 3.0,* Richland: Pacific Northwest National Lab.

Parsons, K., 2009. Maintaining health, comfort and productivity in heat waves. *Global Health Action,* 2(1), pp. 20-57.

Parsons, K., 2014. *Human thermal environments: the effects of hot, moderate, and cold environments on human health, comfort, and performance.* 3rd ed. Florida: Taylor&Francis.

Pope, S., 2019. *Are AMR Devices Any Safer Than Smart Meters?.* [Online]
Available at: https://www.thehealthyhomeeconomist.com/amr-compared-with-smart-meters/
[Accessed 28 4 2019].

Prasad, A. M., Iverson, L. R. & Liaw, A., 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems,* 9(2), pp. 181-199.

Public Health England (PHE), 2014. *Heatwave plan for England - Making the case: the impact of heat on health - now and in the future,* London: Department of Health London.

Rallapalli, H. S., 2010. *A comparison of EnergyPlus and eQUEST whole building energy simulation results for a medium sized office building,* Arizona: Arizona State University.

Rawat, N., Rana, S., Yadav, B. & Yadav, N., 2016. *A review paper on automatic energy meter reading system.* New Delhi, IEEE.

RoSPA, 2014. *Royal Society for the Prevention of Accidents (RoSPA) Home Safety Position Statements.* [Online]
Available at: https://www.rospa.com/rospaweb/docs/advice-services/home-safety/rospa-home-safety-position-statements.pdf
[Accessed 18 09 2018].

Salakij, S., Yu, N., Paolucci, S. & Antsaklis, P., 2016. Model-Based Predictive Control for building energy management. I: Energy modeling and optimal control. *Energy and Buildings,* Volume 133, pp. 345-358.

Santamouris, M. & Kolokotsa, D., 2015. On the impact of urban overheating and extreme climatic conditions on housing, energy, comfort and environmental quality of vulnerable population in Europe. *Energy and Buildings,* Volume 98, pp. 125-133.

Serale, G. et al., 2018. Model predictive control (MPC) for enhancing building and HVAC system energy efficiency: Problem formulation, applications and opportunities. *Energies,* 11(3), p. 631.

Setiawan, A., Koprinska, I. & Agelidis, V., 2009. *Very short-term electricity load demand forecasting using support vector regression.* Atlanta, IEEE.

Shaikh, P. H. et al., 2014. A review on optimized control systems for building energy and comfort management of smart sustainable buildings. *Renewable and Sustainable Energy Reviews,* Volume 34, pp. 409-429.

Shan, K., Wang, S., Yan, C. & Xiao, F., 2016. Building demand response and control methods for smart grids: A review. *Science and Technology for the Built Environment,* 22(6), pp. 692-704.

Shariatzadeh, F., Kumar, N. & Srivastava, A. K., 2016. Optimal control algorithms for reconfiguration of shipboard microgrid distribution system using intelligent techniques. *IEEE Transactions on Industry Applications,* 53(1), pp. 474-482.

Siroky, D. S., 2009. Navigating random forests and related advances in algorithmic modeling. *Statistics Surveys,* Volume 3, pp. 147-163.

Smarra, F. et al., 2018. Data-driven model predictive control using random forests for building energy optimization and climate control. *Applied energy,* Volume 226, pp. 1252-1272.

Smith, S. K. & Sincich, T., 1992. valuating the forecast accuracy and bias of alternative population projections for states. *International Journal of Forecasting,* 8(3), pp. 495-508.

Sun, Q. et al., 2015. A comprehensive review of smart energy meters in intelligent energy networks. *IEEE Internet of Things Journal,* 3(4), pp. 464-479.

Tan, K. M. et al., 2005. High relative humidity measurements using gelatin coated long-period grating sensors. *Sensors and Actuators B: Chemical,* 110(2), pp. 335-341.

Teli, D., Langer, S., Ekberg, L. & Dalenbäck, J.-O., 2018. *Indoor Temperature Variations in Swedish Households: Implications for Thermal Comfort.* Kiruna, Springer, pp. 835-845.

Tham, S. et al., 2020. Indoor temperature and health: a global systematic review. *Public health,* Volume 179, pp. 9-17.

Togun, N. & Baysec, S., 2010. Genetic programming approach to predict torque and brake specific fuel consumption of a gasoline engine. *Applied Energy,* 87(11), pp. 3401-3408.

Tong, A., 2001. Improving the accuracy of temperature measurements. *Sensor Review,* 21(3), pp. 193-198.

Tong, Z. et al., 2016. Energy saving potential of natural ventilation in China: The impact of ambient air pollution. *Applied Energy,* Volume 179, pp. 660-668.

TRNSYS, 2019. *TRNSYS.* [Online]
Available at: http://www.trnsys.com/
[Accessed 24 June 2020].

UKCP, 2009. [Online]
Available at: http://ukclimateprojections-ui.metoffice.gov.uk
[Accessed 12 7 2018].

UN Environment & International Energy Agency (IEA), 2017. *Towards a zero-emission, efficient, and resilient buildings and construction sector. Global Status Report 2017.,* Paris: United Nations Environment Programme.

United Nations, 2015. *Adoption of the Paris Agreement,* Paris: United Nations.

US EIA, 2018. *How much energy is consumed in US residential and commercial buildings?,* US: The U.S. Energy Information Administration.

Vellei, M., Herrera, M., Fosas, D. & Natarajan, S., 2017. The influence of relative humidity on adaptive thermal comfort. *Building and Environment,* Volume 124, pp. 171-185.

Venkateshan, S. P., 2015. *Mechanical measurements.* Chichester: John Wiley & Sons Ltd.

Voyant, C. et al., 2017. Machine learning methods for solar radiation forecasting: A review. *Renewable Energy,* Volume 105, pp. 569-582.

Wu, S. & Sun, J.-Q., 2012. A physics-based linear parametric model of room temperature in office buildings. *Building and Environment,* Volume 50, pp. 1-9.

Wu, X. et al., 2008. Top 10 algorithms in data mining. *Knowledge and information systems,* 14(1), pp. 1-37.

Yang, L., Yan, H. & Lam, J. C., 2014. Thermal comfort and building energy consumption implications – A review. *Applied Energy,* Volume 115, pp. 164-173.

Yeo, T., Sun, T. & Grattan, K., 2008. Fibre-optic sensor technologies for humidity and moisture measurement. *Sensors and Actuators A: Physical,* 144(2), pp. 280-295.

Yugui, C., 2013. Electric energy demand forecast of Nanchang based on cellular genetic algorithm and BP neural network. *Indonesian Journal of Electrical Engineering and Computer Science,* 11(7), pp. 3821-3825.

Zhang, R., Lam, K. P., Yao, S.-c. & Zhang, Y., 2013. Coupled EnergyPlus and computational fluid dynamics simulation for natural ventilation. *Building and Environment,* Volume 68, pp. 100-113.

Zhao, H., Jivraj, S. & Moody, A., 2019. 'My blood pressure is low today, do you have the heating on?'The association between indoor temperature and blood pressure. *Journal of hypertension,* 37(3), pp. 504-512.

Zhao, H. & Magoulès, F., 2010. Parallel support vector machines applied to the prediction of multiple buildings energy consumption. *Journal of Algorithms & Computational Technology,* 4(2), pp. 231-249.

Zhao, H. x. & Magoulès, F., 2012. A review on the prediction of building energy consumption. *Renewable and Sustainable Energy Reviews,* 16(6), pp. 3586-3592.

Zhao, J., Lam, P., Ydstie, E. & Karaguzel, T., 2015. EnergyPlus model-based predictive control within design--build--operate energy information modelling infrastructure. *Journal of Building Performance Simulation,* 8(3), pp. 121-134.

Zhou, K., Fu, C. & Yang, S., 2016. Big data driven smart energy management: From big data to big insights. *Renewable and Sustainable Energy Reviews,* Volume 56, pp. 215-225.

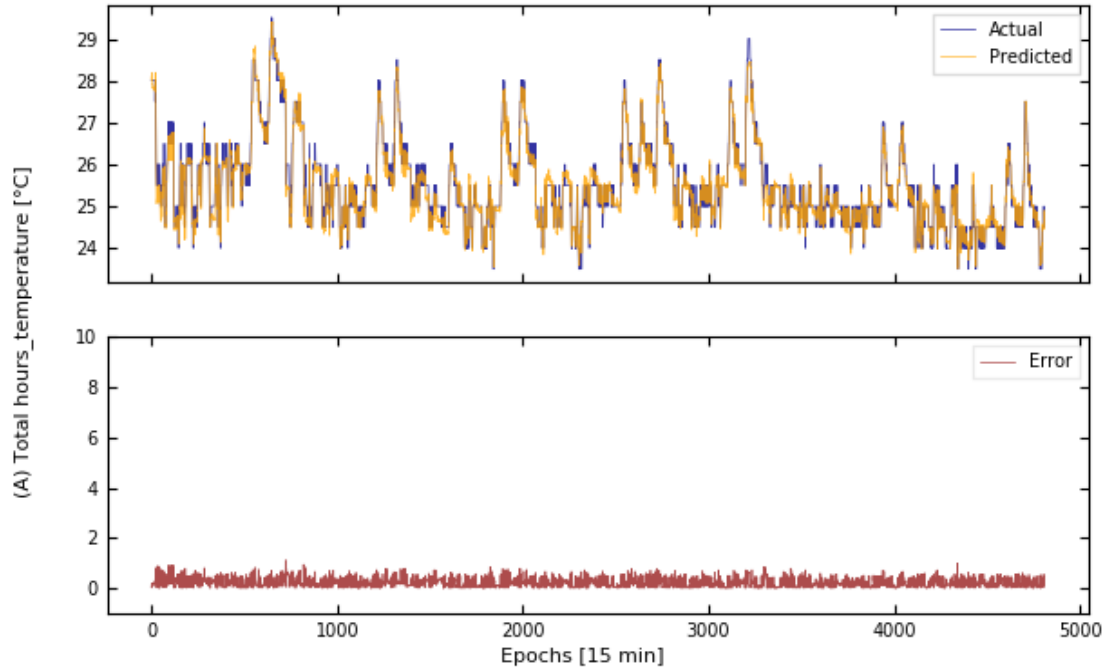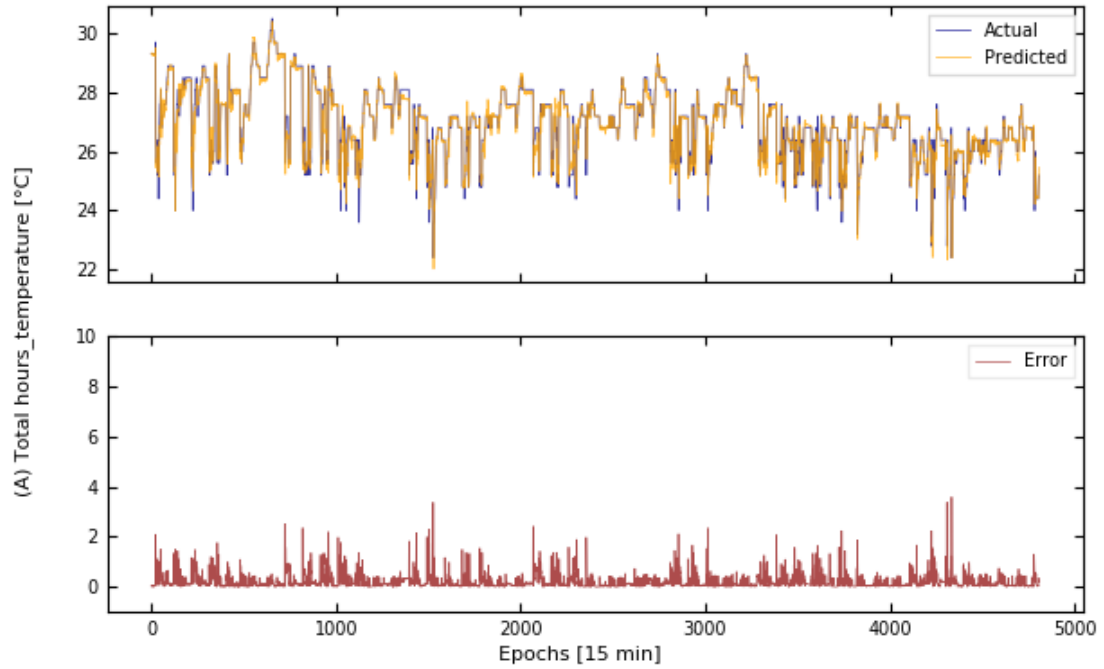# Appendix A   Forecasting results

Indoor unit U2-10

Figure A- 1 Forecasting results and errors of air temperature for indoor unit U2-10.

Table A- 1 Statistics of the forecasting results of air temperature for indoor unit U2-10.

Scenario	Description	MAE	MSE	RMSE	R^2
A	Total hours	0.215	0.077	0.277	0.928
B	One week	0.196	0.064	0.253	0.871
C	Weekday	0.198	0.069	0.264	0.857
D	Weekend	0.202	0.077	0.277	0.920
E	Working hours	0.223	0.080	0.283	0.870
F	Off hours	0.205	0.073	0.270	0.897
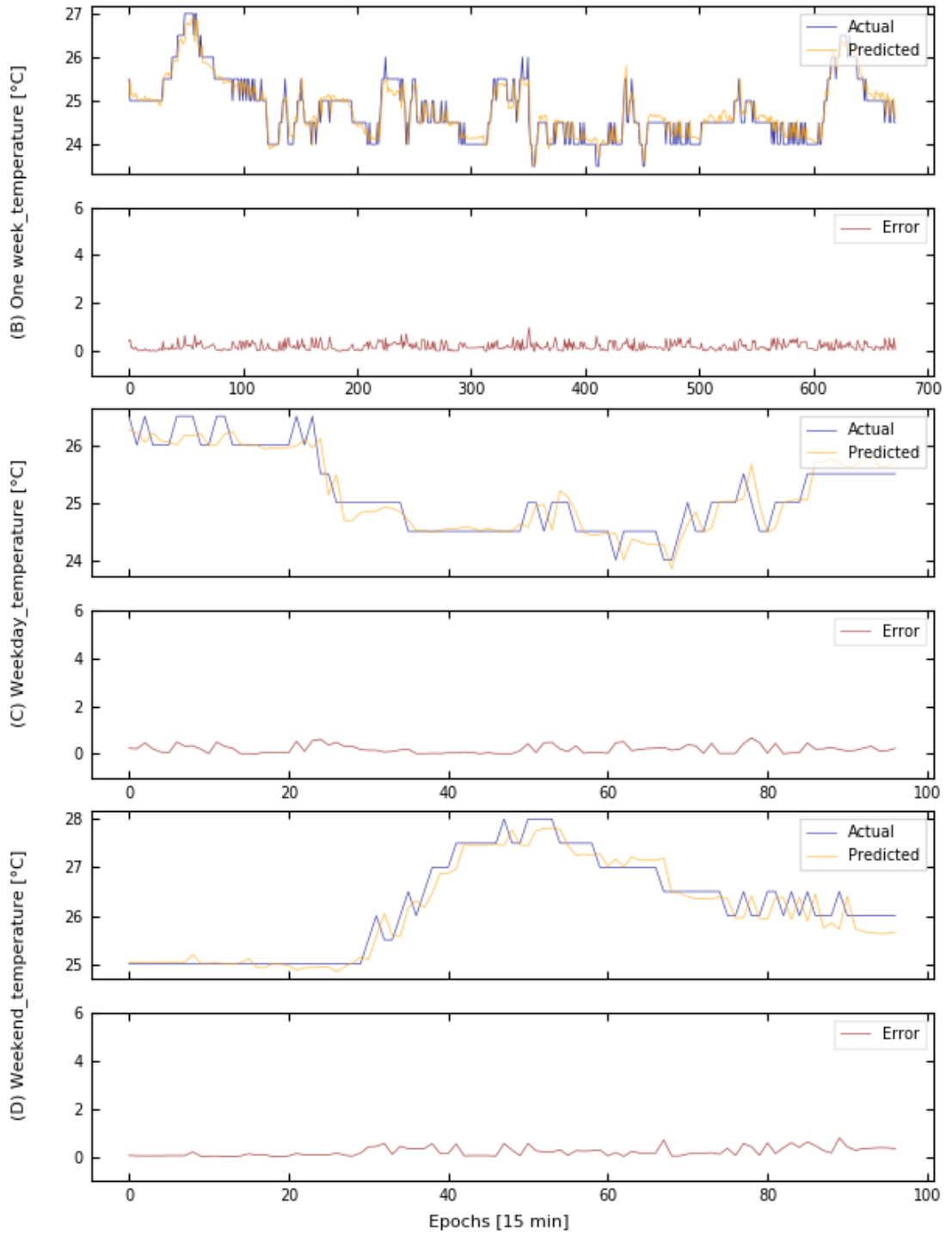
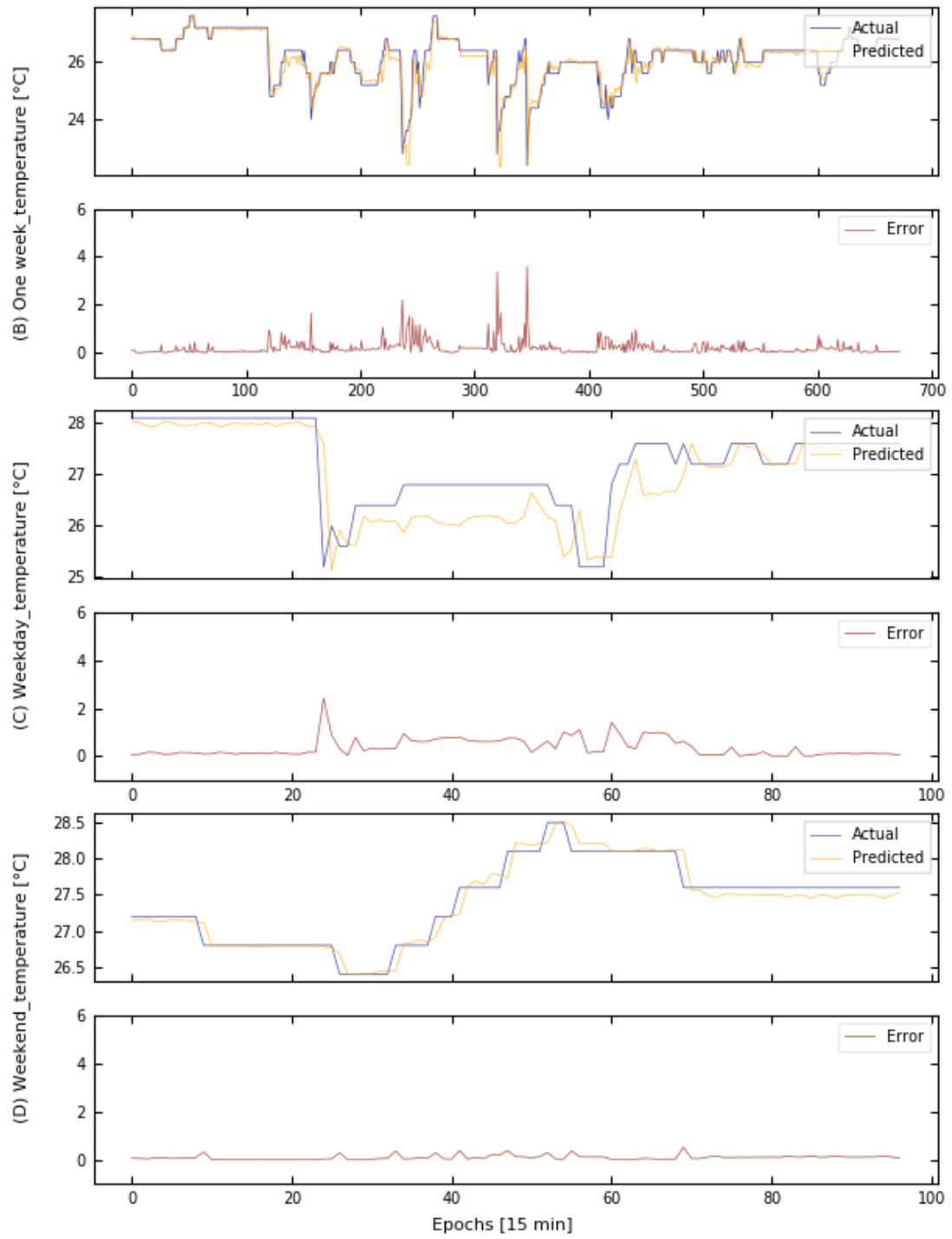Figure A- 2 Forecasting results and errors of air temperature for indoor unit U2-10 during the different scenarios.
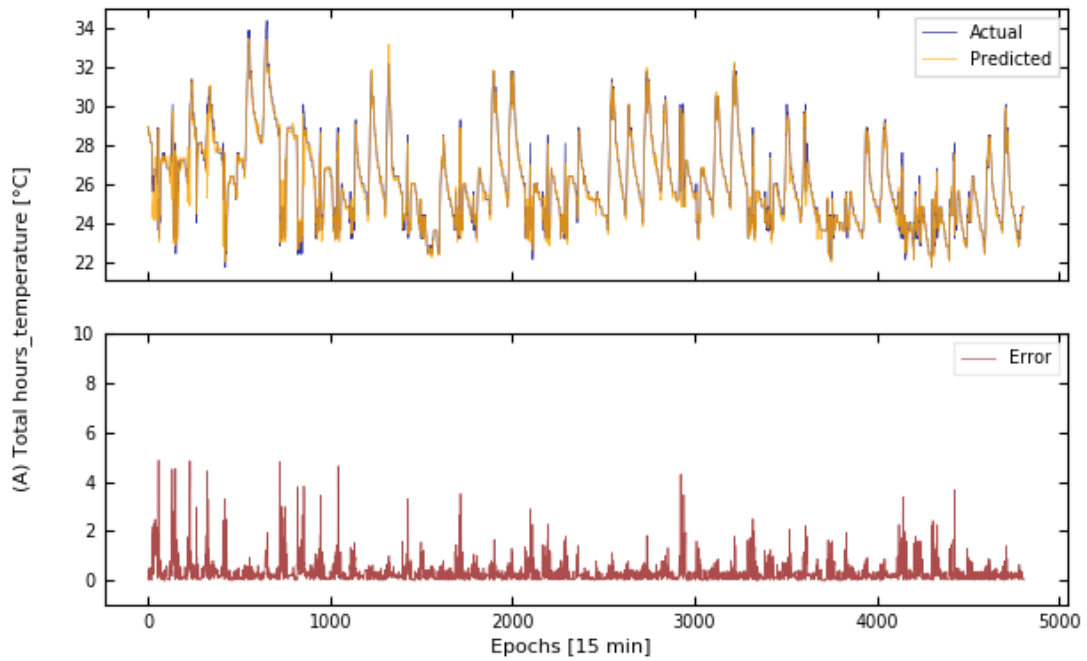
Indoor unit U2-11

Figure A- 3 Forecasting results and errors of air temperature for indoor unit U2-11.

Table A- 2 Statistics of the forecasting results of air temperature for indoor unit U2-11.

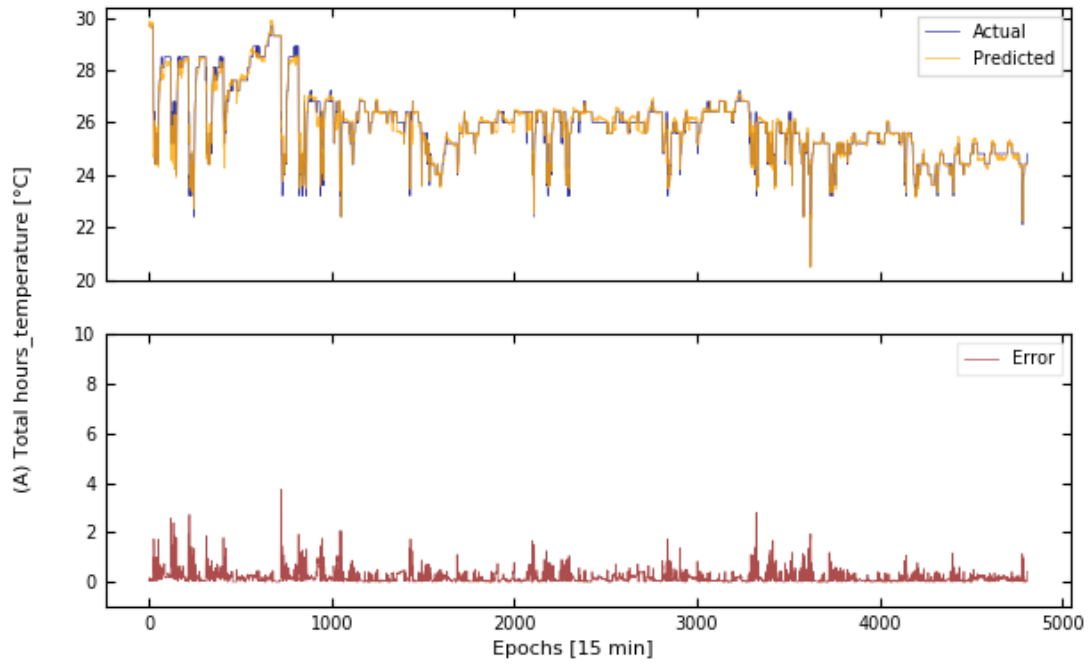| Scenario	Description	MAE	MSE	RMSE	R²
A	Total hours	0.203	0.125	0.353	0.908
B	One week	0.196	0.133	0.365	0.797
C	Weekday	0.371	0.290	0.538	0.526
D	Weekend	0.087	0.017	0.132	0.944
E	Working hours	0.332	0.245	0.495	0.831
F	Off hours	0.162	0.095	0.309	0.920

Figure A- 4 Forecasting results and errors of air temperature for indoor unit U2-11 during the different scenarios.

Indoor unit U2-12

Figure A- 5 Forecasting results and errors of air temperature for indoor unit U2-12.

Table A- 3 Statistics of the forecasting results for indoor air temperature for indoor unit U2-12.

Scenario	Description	MAE	MSE	RMSE	R²
A	Total hours	0.305	0.301	0.549	0.935
B	One week	0.307	0.253	0.503	0.887
C	Weekday	0.382	0.362	0.602	0.721
D	Weekend	0.200	0.073	0.271	0.987
E	Working hours	0.504	0.662	0.814	0.852
F	Off hours	0.194	0.120	0.346	0.949
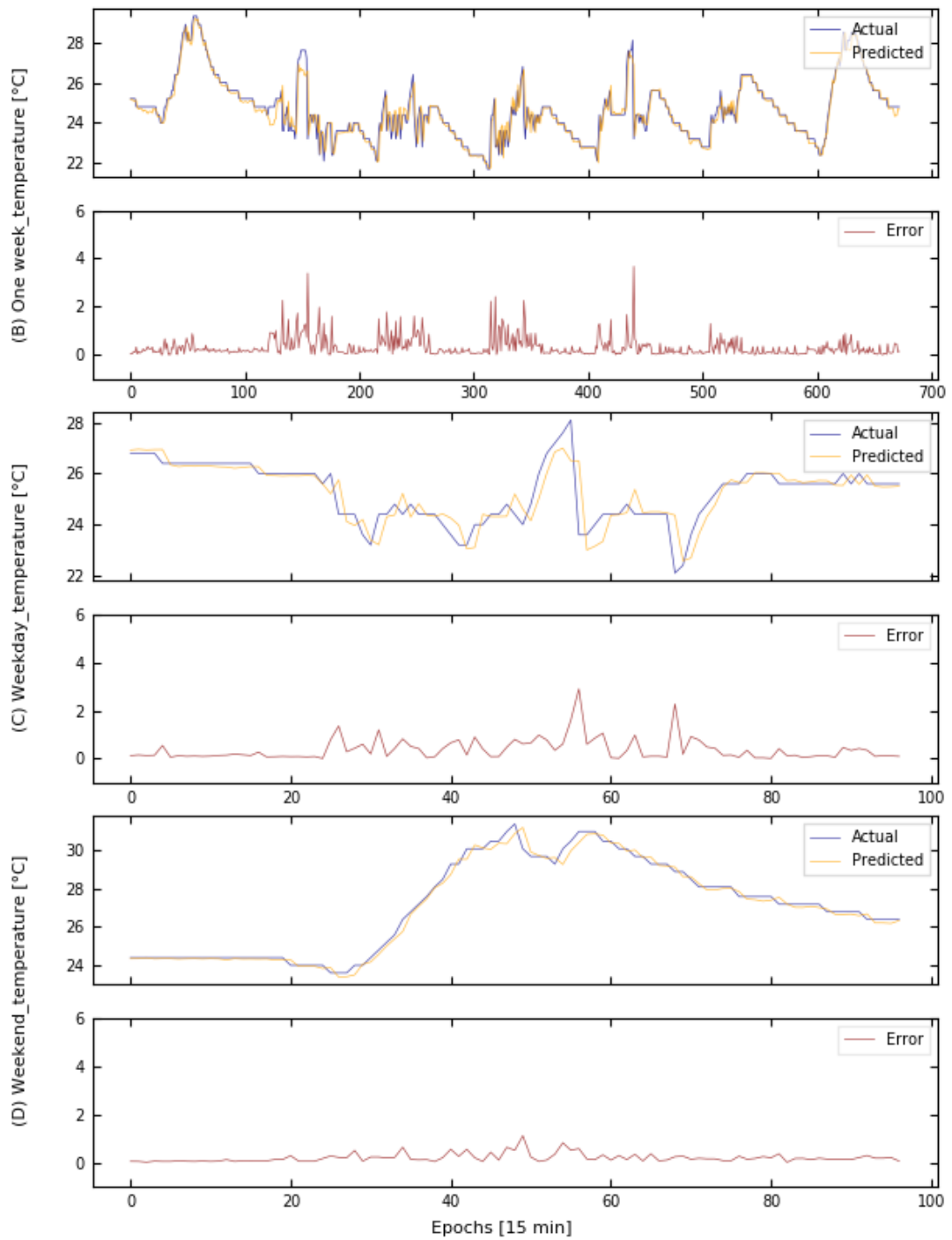
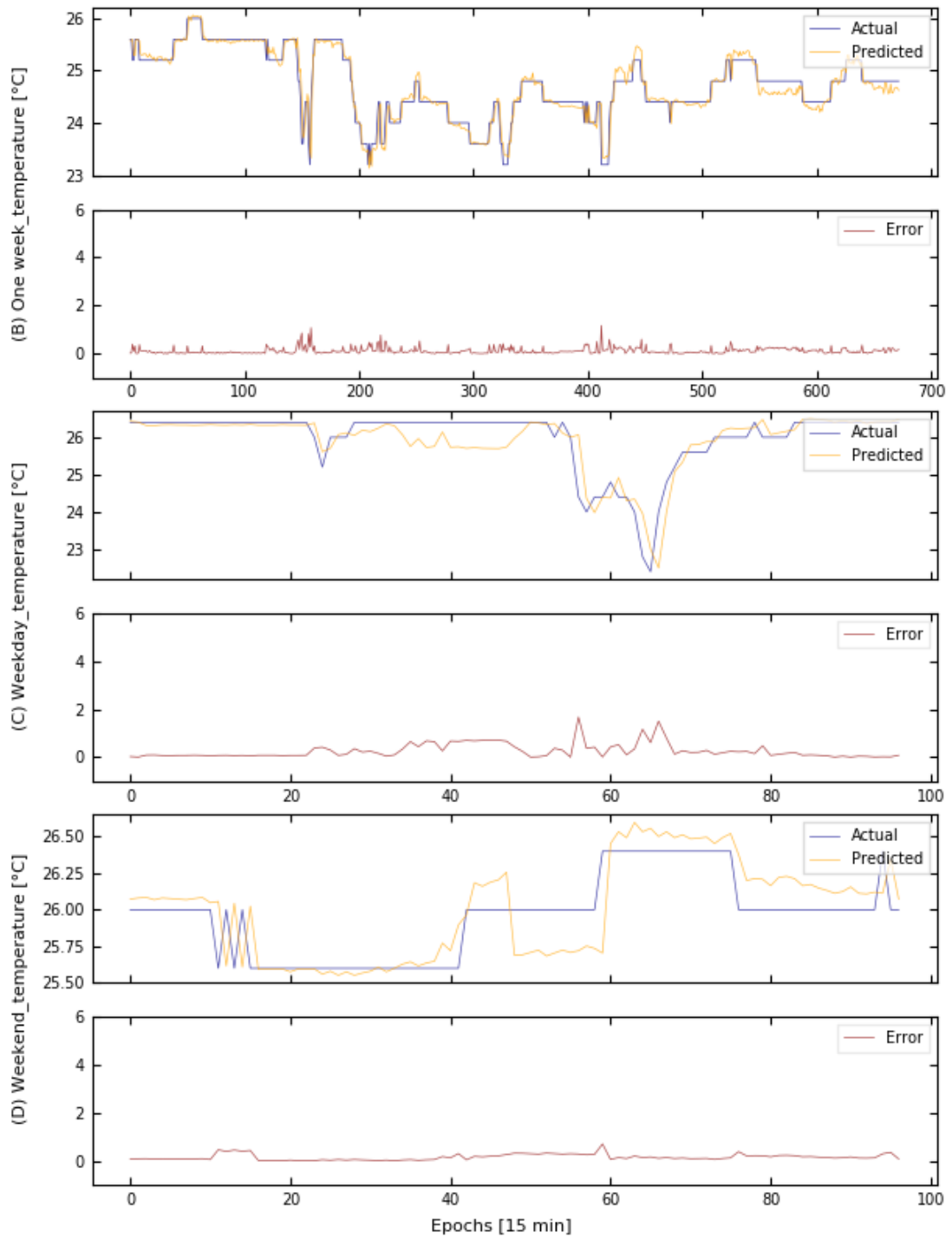Figure A- 6 Forecasting results and errors of air temperature for indoor unit U2-12 during the different scenarios.
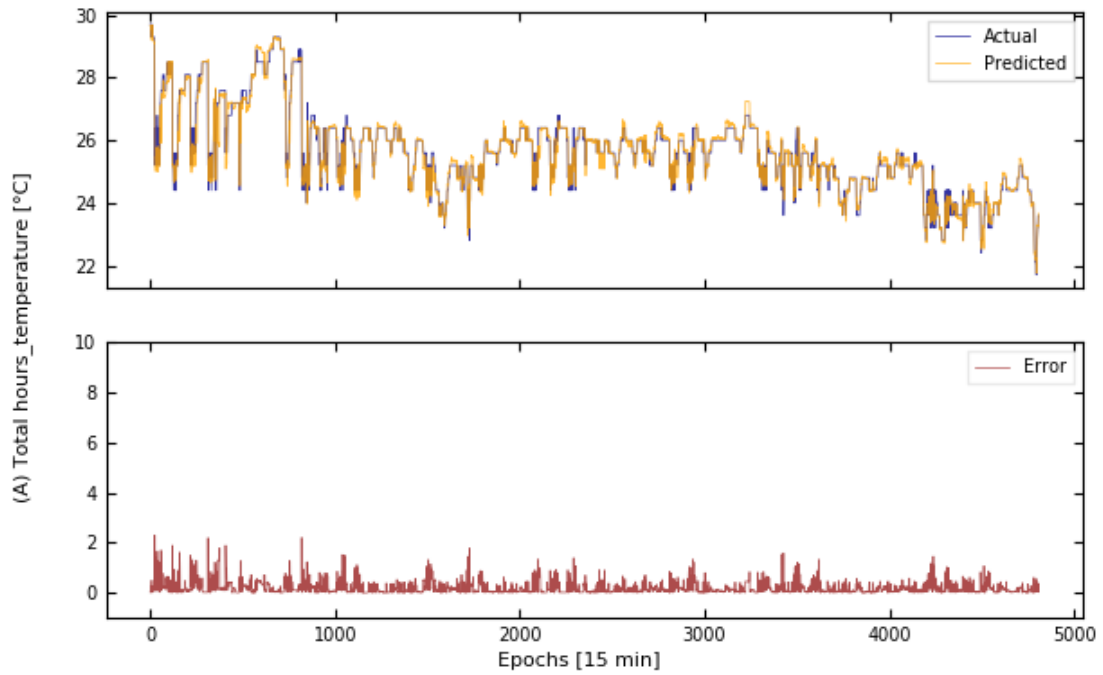
Indoor unit U2-13

Figure A- 7 Forecasting results and errors of air temperature for indoor unit U2-13.

Table A- 4 Statistics of the forecasting results of air temperature for indoor unit U2-13.

| Scenario	Description	MAE	MSE	RMSE	R²
A	Total hours	0.179	0.095	0.309	0.943
B	One week	0.113	0.031	0.175	0.922
C	Weekday	0.264	0.163	0.404	0.750
D	Weekend	0.150	0.039	0.197	0.490
E	Working hours	0.257	0.159	0.398	0.880
F	Off hours	0.163	0.090	0.300	0.954

Figure A- 8 Forecasting results and errors of air temperature for indoor unit U2-13 during the different scenarios.
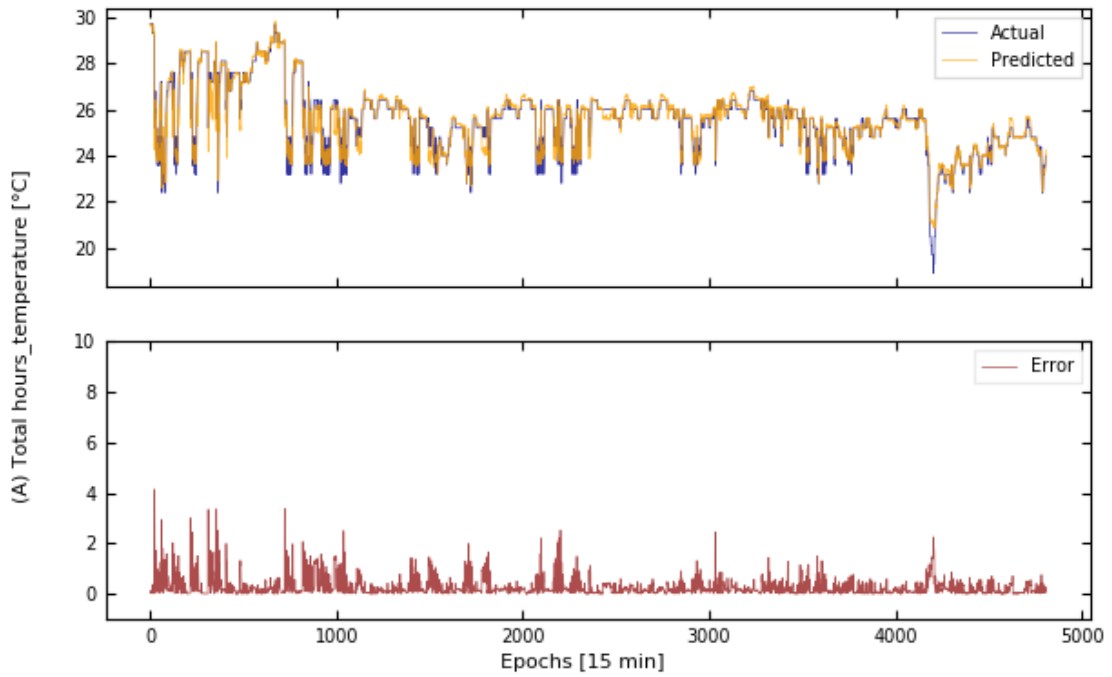
Indoor unit U2-14

Figure A- 9 Forecasting results and errors of air temperature for indoor unit U2-14.

Table A- 5 Statistics of the forecasting results of air temperature for indoor unit U2-14.

| Scenario	Description	MAE	MSE	RMSE	R^2
A	Total hours	0.183	0.094	0.307	0.943
B	One week	0.171	0.071	0.266	0.877
C	Weekday	0.246	0.163	0.404	0.618
D	Weekend	0.107	0.023	0.150	0.865
E	Working hours	0.271	0.165	0.406	0.863
F	Off hours	0.148	0.073	0.271	0.965
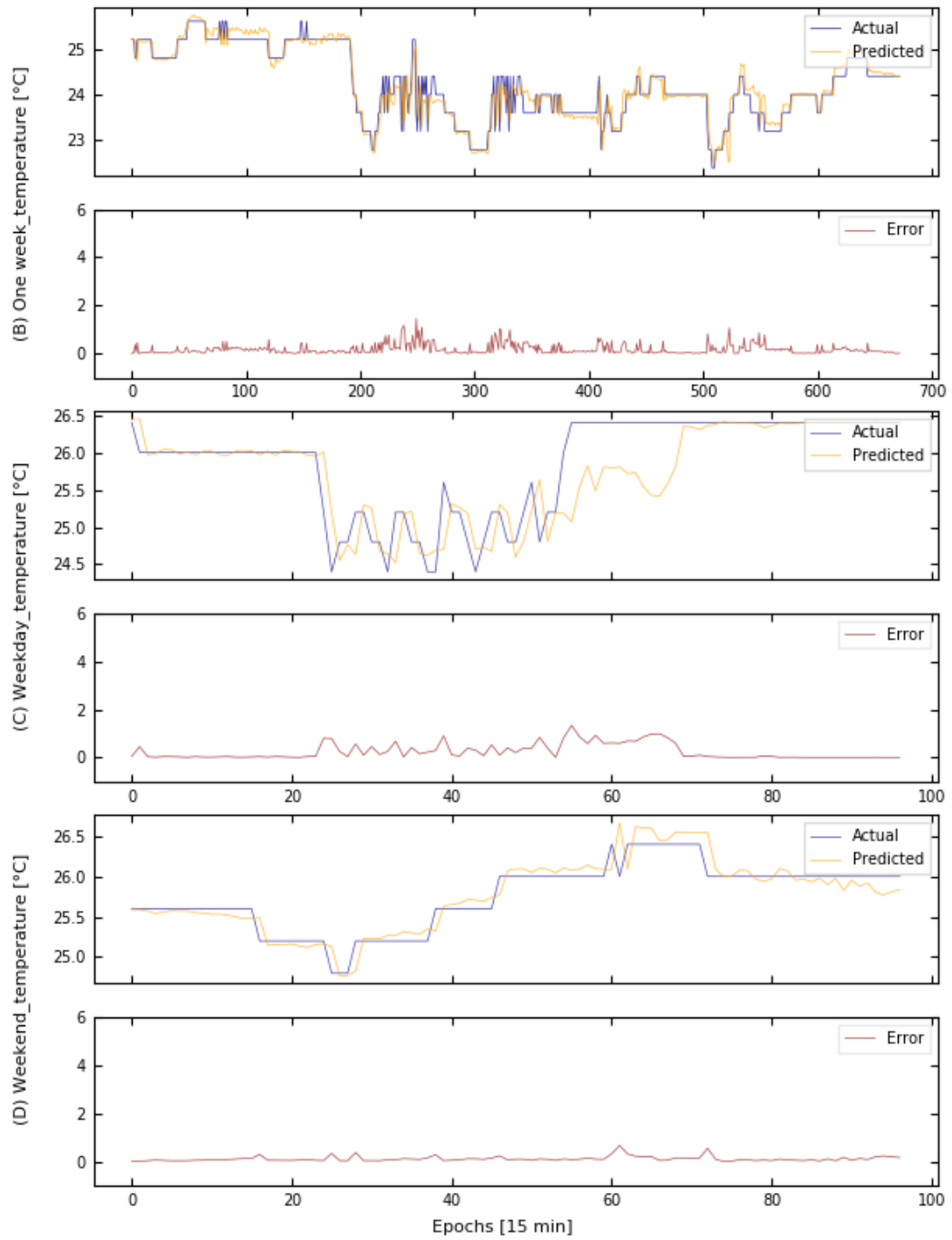
Figure A- 10 Forecasting results and errors of air temperature for indoor unit U2-14 during the different scenarios.
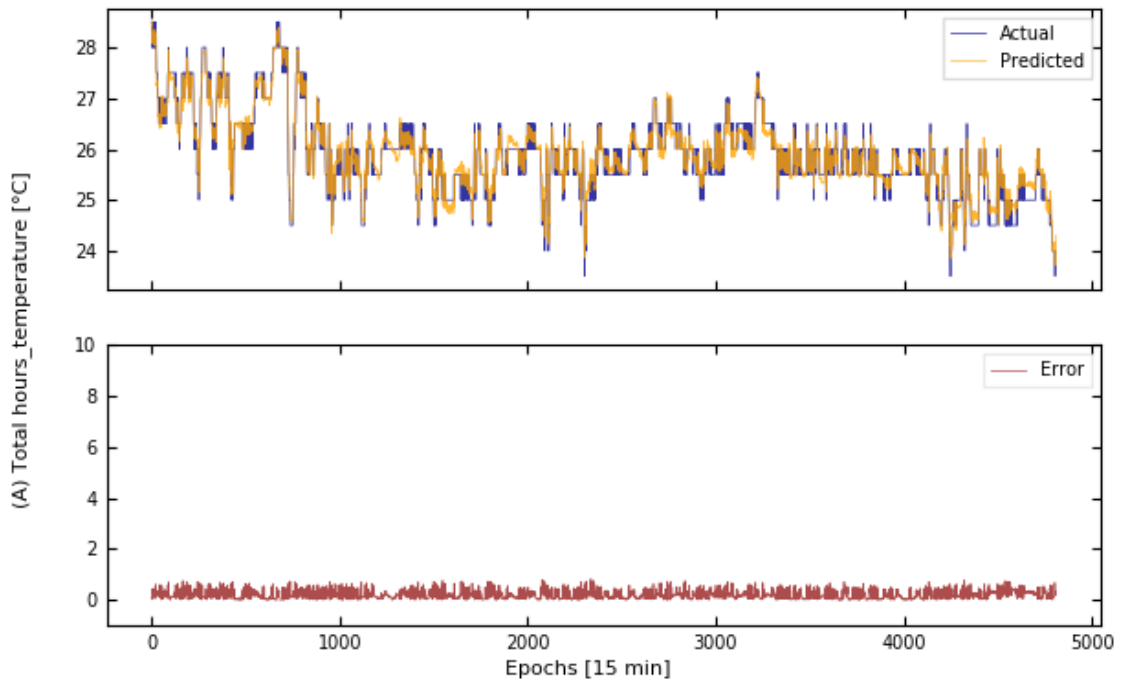
Indoor unit U2-15

Figure A- 11 Forecasting results and errors of air temperature for indoor unit U2-15.

Table A- 6 Statistics of the forecasting results of air temperature for indoor unit U2-15.

| Scenario	Description	MAE	MSE	RMSE	R^2
A	Total hours	0.250	0.173	0.416	0.913
B	One week	0.208	0.122	0.350	0.932
C	Weekday	0.349	0.231	0.481	0.774
D	Weekend	0.154	0.038	0.194	0.663
E	Working hours	0.348	0.264	0.514	0.839
F	Off hours	0.239	0.191	0.437	0.920
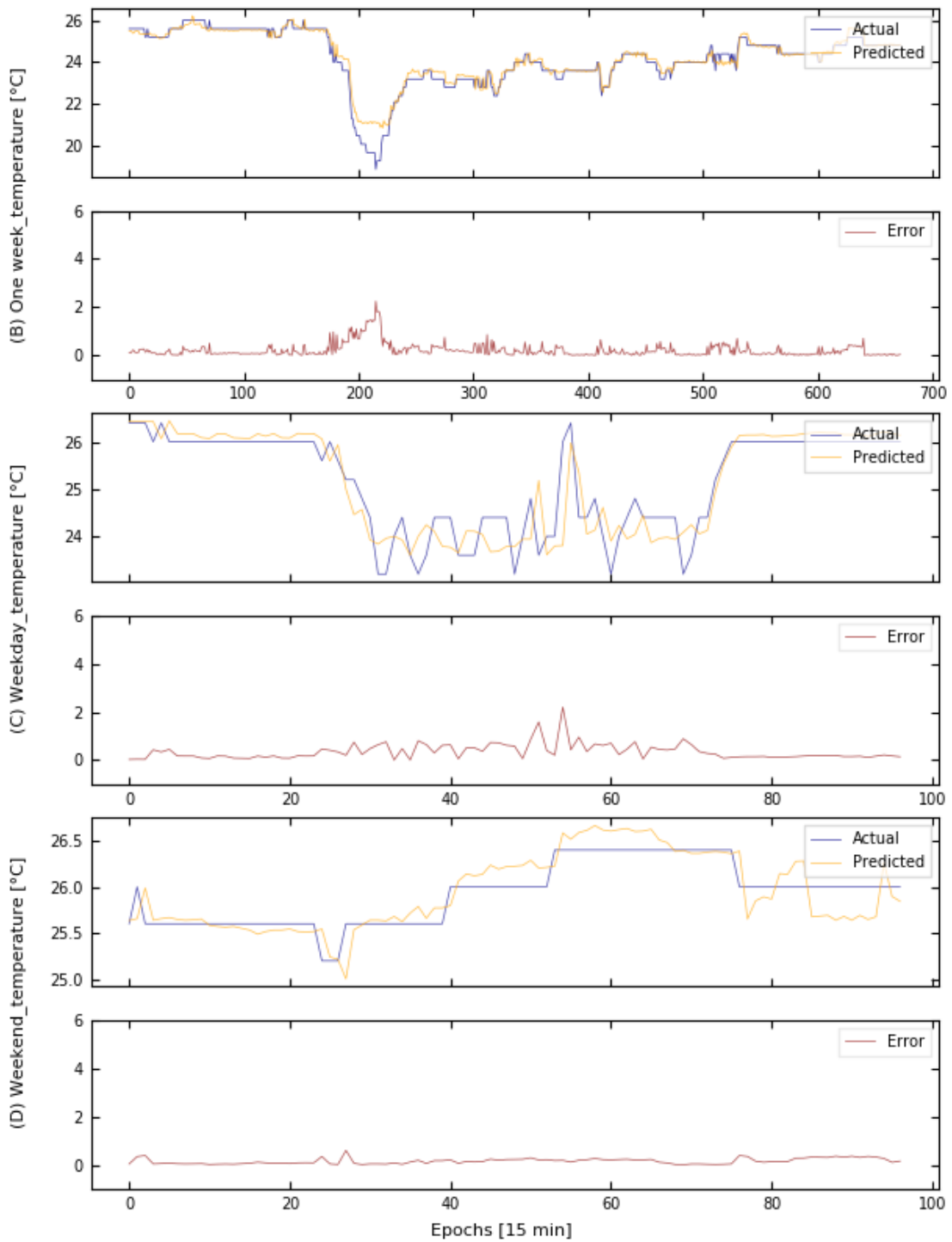
Figure A- 12 Forecasting results and errors of air temperature for indoor unit U2-15 during the different scenarios.

Indoor unit U2-16

Figure A- 13 Forecasting results and errors of air temperature for indoor unit U2-16.

Table A- 7 Statistics of the forecasting results of air temperature for indoor unit U2-16.

Scenario	Description	MAE	MSE	RMSE	R²
A	Total hours	0.188	0.061	0.247	0.907
B	One week	0.206	0.070	0.265	0.781
C	Weekday	0.186	0.069	0.263	0.809
D	Weekend	0.115	0.027	0.163	0.696
E	Working hours	0.198	0.070	0.264	0.871
F	Off hours	0.182	0.060	0.244	0.924
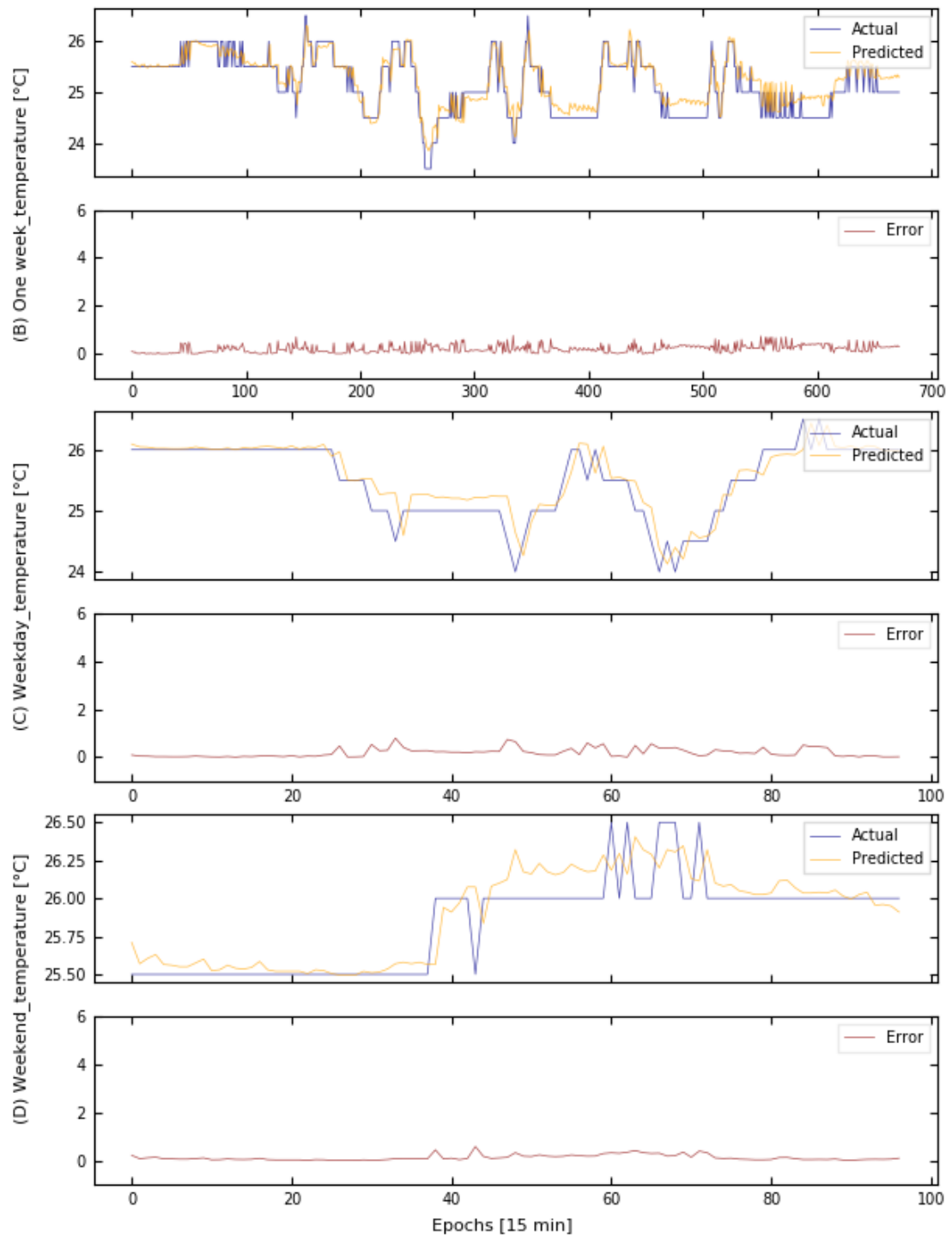
Figure A- 14 Forecasting results and errors of air temperature for indoor unit U2-16 during the different scenarios.
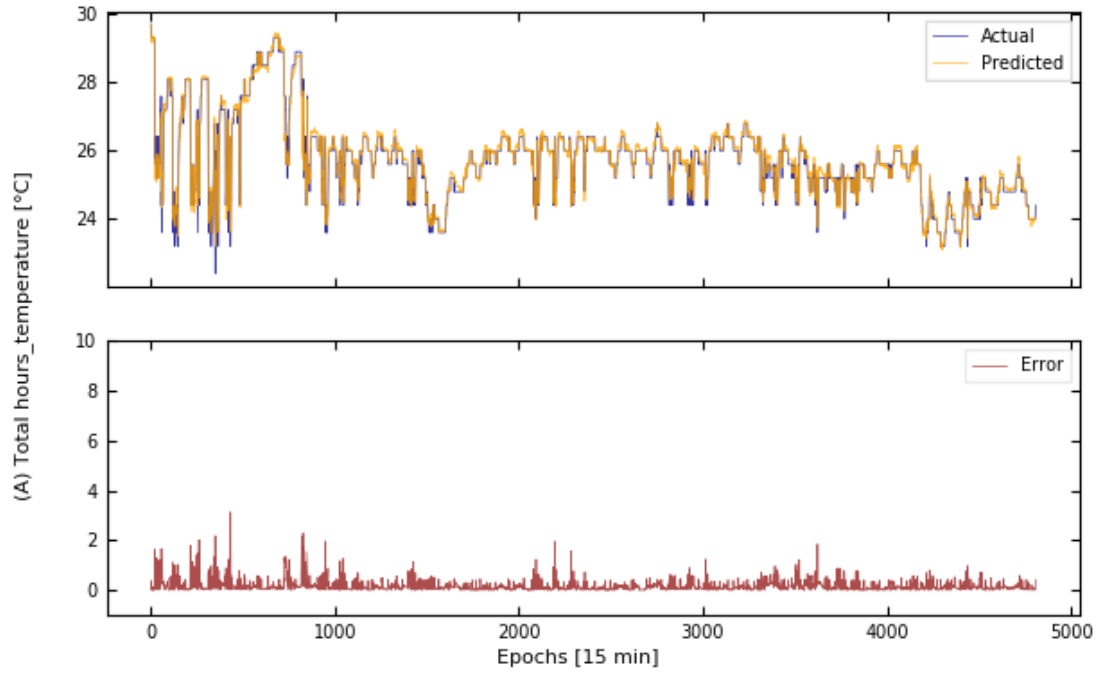
Indoor unit U2-17

Figure A- 15 Forecasting results and errors of air temperature for indoor unit U2-17.

Table A- 8 Statistics of the forecasting results of air temperature for indoor unit U2-17.

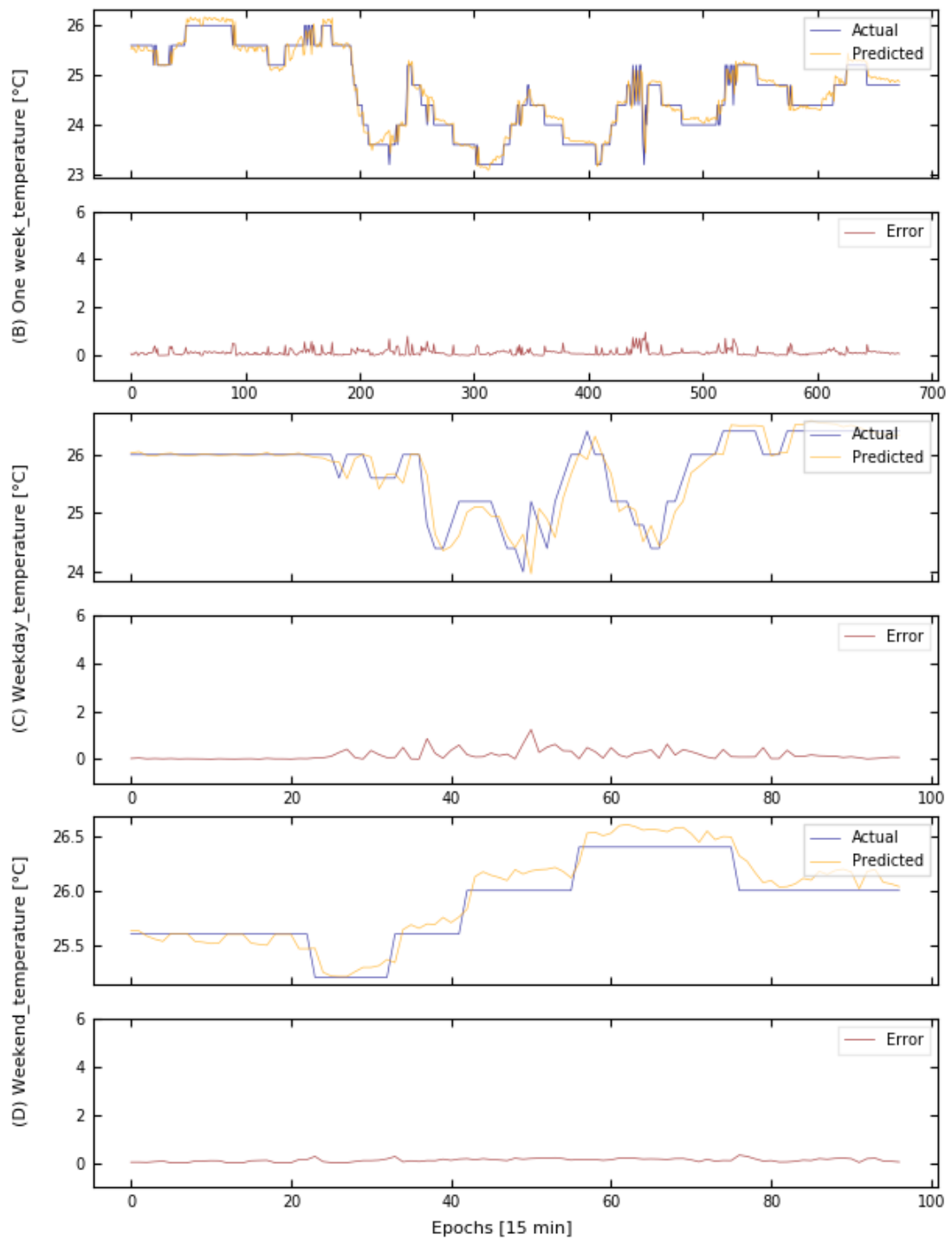| Scenario	Description	MAE	MSE	RMSE	R^2
A	Total hours	0.163	0.074	0.272	0.947
B	One week	0.127	0.035	0.187	0.946
C	Weekday	0.175	0.076	0.276	0.797
D	Weekend	0.111	0.017	0.131	0.873
E	Working hours	0.242	0.143	0.378	0.836
F	Off hours	0.135	0.051	0.227	0.970

Figure A- 16 Forecasting results and errors of air temperature for indoor unit U2-17 during the different scenarios.
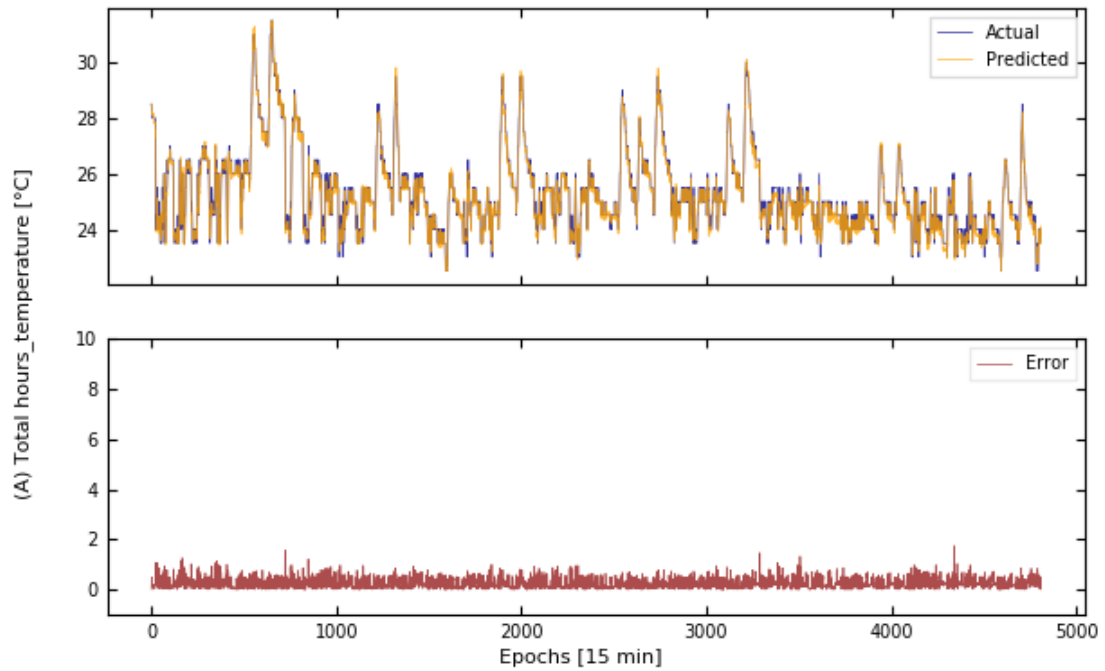
Indoor unit U2-18

Figure A- 17 Forecasting results and errors of air temperature for indoor unit U2-18.

Table A- 9 Statistics of the forecasting results of air temperature for indoor unit U2-18.

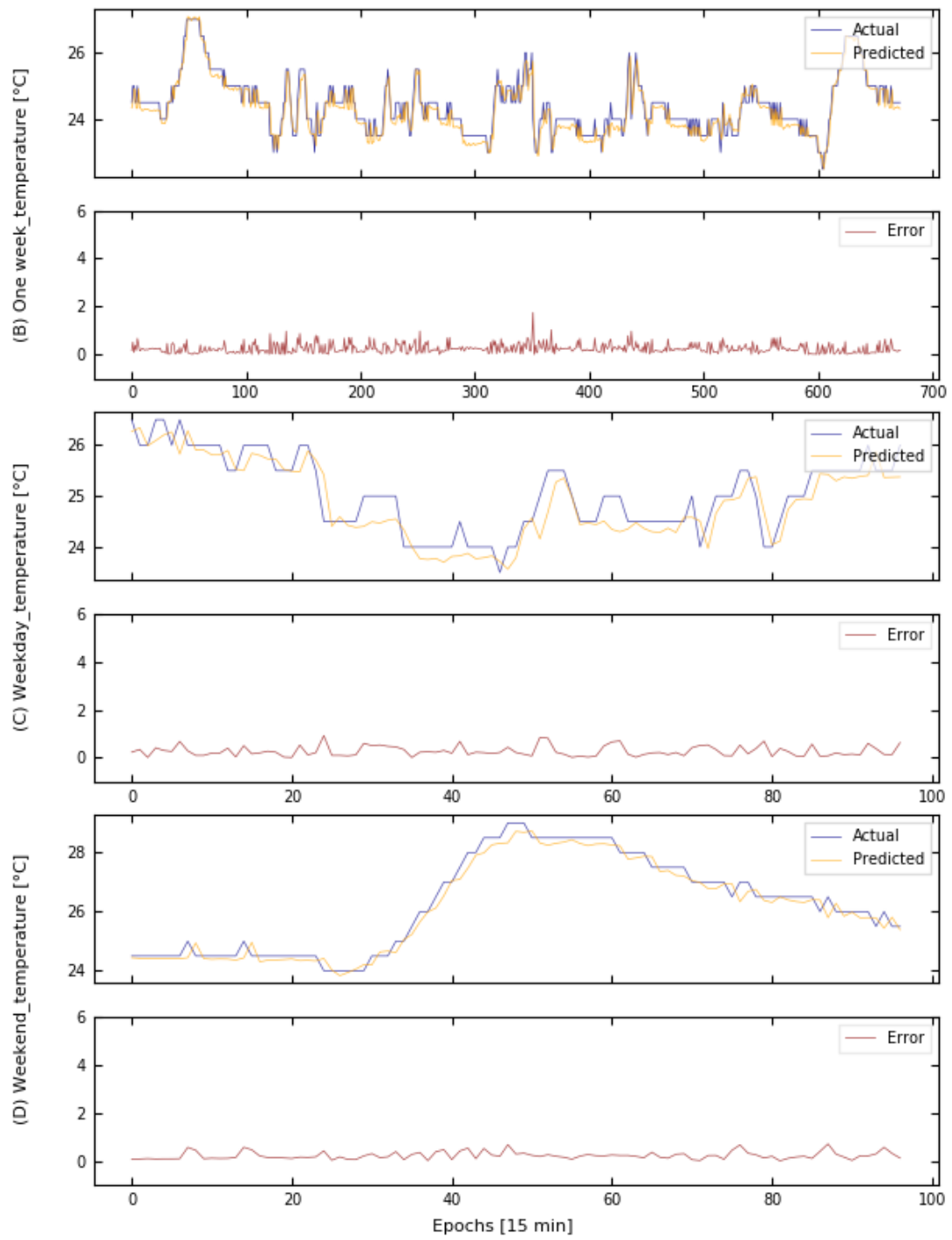| Scenario	Description	MAE	MSE	RMSE	R²
A	Total hours	0.243	0.102	0.320	0.951
B	One week	0.246	0.099	0.315	0.857
C	Weekday	0.269	0.118	0.343	0.785
D	Weekend	0.224	0.075	0.274	0.970
E	Working hours	0.277	0.129	0.359	0.893
F	Off hours	0.236	0.096	0.310	0.920

Figure A- 18 Forecasting results and errors of air temperature for indoor unit U2-18 during the different scenarios.