NOTTINGHAM
TRENT UNIVERSITY

Implementation of Low Temperature District Heating in Existing Buildings

Asad Ashfaq

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

Nottingham Trent University.

School of Architecture Design and the Built Environment

May, 2020

Declaration of Authorship

I, Asad Ashfaq, declare that this thesis titled, 'Implementation of Low Temperature District Heating in Existing Buildings' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.

- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.

- Where I have consulted the published work of others, this is always clearly attributed.

- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.

- I have acknowledged all main sources of help.

- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself

Signed: _____

Date: _____

Dedicated to my lovely **Wife**

Abstract

District heating (DH) provides a gateway for the integration of low carbon technologies and renewable energy sources to achieve a sustainable carbon neural future. The low temperature district heating (LTDH), in particular, is the latest and most efficient technology which enables the possibility of combining multi-vector heat sources to the network such as, renewable energy sources, heat-pumps and waste heat from the industry. This thesis considers REMOURBAN project to investigate the implementation of low-temperature district heating in existing boiler based buildings in Nottingham. This LTDH (60/30) network intervention is first of it's kind in the UK and utilises return pipe to heat 94 flats. The study is comprised in three main parts, i.e. thermal performance modelling of buildings, hydraulic modelling of the district heating network and predictive modelling of monitored data.

The results from the first part show that retrofitting increases the energy performance of buildings by almost 50%, and the relation between the building regulations and thermal performance analysis show that with current regulations in the UK, it is unlikely to achieve the target of net-zero emission buildings (NZEB) by the year 2050. The second part of the study investigates the design and operation of an energy efficient LTDH network (from REMOURBEN). The results from the hydraulic modelling suggest that the networks should be designed with variable speed pumping, and supply water temperature should be kept constant from the plant room. This leads to the lowest energy consumption in the network. It is concluded that the energy efficiency and Δt in REMOURBAN project can be improved by reducing flow-rates both in the network and circulation pump inside the plant room. Moreover, the techno-economic analysis for the de-carbonised district heating network shows that 100% decarbonisation depends on selling excess electricity

and heat to the private consumers.

Finally, the predictive modelling suggests classical stochastic SARIMA method is good for short-horizon forecasts while modern machine learning (MLP and SVR neural networks) are best for medium and long-horizon forecasts. The GIS mapping shows that the decentralised LTDH network with multiple energy centres is the optimum strategy owing to the cost of network pipe-works and heat-losses in the network.

The overall conclusion of the study is that the implementation of low temperature district heating in existing building is possible and optimisation as well as control of flow-rates are the key factors in achieving energy efficiency in the network. The novelty of this study is that a live LTDH network intervention has been as a case study which provides a energy efficient solution for the UK. The learning from this study can be replicated to the future LTDH network project anywhere in the UK or elsewhere.

Acknowledgements

This thesis would not have been possible without the help and support from many wonderful people, to whom I wish to say a huge thank-you!

First, I would like to express my sincerest gratitude to my supervisor, Anton Ianakiev for his immense support throughout this journey. I am very much grateful for supervision meetings where he guided me in such a professional and kind way. I extend my deepest respect to Syed Mohyuddin for being part of my supervision team, and Hynda Klalib for helping me to settle in Nottingham. The role of Nottingham Trent University for the award of Vice Chancellor bursary is greatly appreciated. It would have not been possible to step up in my career without this. The meetings with Nottingham City Council, Nottingham City Homes, Melius Homes, Enviroenergy Ltd, SASIE Ltd and Acumen Projects Ltd have been really helpful. I am obliged for their help in sharing data required for this PhD research.

Many thanks to my parents and siblings, (Umair, Rabia, Ali) for always believing in me. It is due to my parent's blessing that I am able to achieve this milestone and making them feel proud. I extend my appreciation to my parents in law as well as to my family members in the UK, who were always there to talk and share things in difficult time. Especially to Anwar bhai, Tahira, Haseeb, Khadija and Ayesha. Thank you everyone for tolerating me in rainy days and being my inspiration to be a better person.

I would like to extend my sincere thanks to Mei Ren and Benjamin Jones from BuroHappold Engineering for offering an internship during my PhD and letting me work on their amazing engineering projects. This has been a wonderful learning experience which I can never forget. Many thanks goes to my colleagues there who played a great role in professional development. The mem-

Contents

List of Figures

List of Tables

List of Abbreviations

$4GDH$	4th Generation District Heating
ADF	Augmented Dickey Fuller
AMM	Abandoned Mine Methane
$ARCH$	Autoregressive conditional heteroskedasticity
ACF	Autocorrelation Function
AR	Autoregressive
$ARIMA$	Autoregressive Integrated Moving Average
$BBOE$	Bottom Bottom Opposite End
BER	Building Carbon Emission Rate
BGS	British Geological Survey
BRE	Building Research Establishment
$BREDEM$	Building Research Establishment Domestic Energy Model
BS	British Standard
$BRUKL$	Building Regulations UK, Part L
$CapEx$	Capital Investment Costs
CBM	Coal Bed Methane
$CCGT$	Combined Cycle Gas Turbines
CCS	Carbon Capture and Storage
CHP	Combined Heat and Power
$CIBSE$	Chartered Institution of Building Services Engineers
COP	Coefficient of Performance
$COP21$	Conference of the Parties

C_p	Specific Heat Capacity
DC	Direct Current Electrical Grid
DH	District Heating
$DECC$	Department of Energy and Climate Change
DNN	Deep Neural Network
DOE	Department for the Environment
DS	Danish standard
DSY	Design Summer Year
$DYMOLA$	DYnamic MOdelling LAboratory Software
DHW	Domestic Hot Water Demand
EN	Elastic Net
$EPBD$	European Performance of Building directive
EPC	Energy Performance Certificate
$ESCO$	Energy Services Company
FME	Feature Manipulation Engine Software
FNN	Feed-Forward Neural Network
$GARCH$	Generalized Autoregressive Conditional Heteroskedasticity
$GBRT$	Gradient Boosted Regression Tree
GHG	Green House Gas
GIS	Geographical Information System
GWh	Gigawatt Hours
$HIUs$	Heat Interface Units
HP	Heat Pump
$HVAC$	Heating Ventilation and Air Conditioning
ICT	Information Communications Technology
$IDAICE$	IDA Indoor Climate and Energy Software
$IESVE$	Integrated Environmental Solutions Virtual Environment

ISET	Institute for Solar Energy Supply Technologies
KDE	Kernel Density Estimation
KNN	K-Nearest Neighbors
LASSO	Least Absolute Shrinkage and Selection Operator
LCOE	Levelised Cost of Energy
LEED	Leadership in Energy and Environmental Design
LR	Linear Regression
LSTM	Long Short-Term Memory
LTDH	Low Temperature District Heating
MA	Moving Average
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MSE	Mean Squared Error
MSPE	Mean Square Percentage Error
MLP	Multi-layer Perceptrons
MSE	Mean Squared Error
MVHR	Mechanical Ventilation Heat Recovery
NABERS	National Australian Built Environment Rating System
NCM	National Calculation Method
NZEB	Net Zero Emission Buildings
OLS	Ordinary Linear Regression
OpEx	Operation and Maintenance Costs
PEX	Cross-linked Polyethylene Foam
PCAF	Partial Autocorrelation Function
PID	Proportional Integral Derivative
PMV	Predicted Mean Vote
PV	Photo-Voltaic

Q_{loss}	Heat-loss
RBF	Radial Bais Regression
RdSAP	Reduced Standard Assessment Procedure
ReLU	Rectified Linear Unit
REMOURBAN	REgeneration MOdel for accelerating the smart URBAN transformation
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
R2	R-Squared
RSS	Residual Sum of Squares
SAP	Standard Assessment Procedure
SBEM	Simplified Building Energy Model
SARIMA	Seasonal Autoregressive Integrated Moving Average
SC	Shading coefficient
SCC	Southampton City Council
SGD	Stochastic Gradient Descent
SVC	Support Vector Classification
SVM	Support Vector Machines
SVR	Support Vector Regression
SH	Space Heat Demand
TBOE	Top Bottom Opposite End
TER	Target Emission Rate
TES	Thermal Electric Storage
TRV	Thermostatic Radiator Valve
TRY	Test Reference Year
WSHP	Water Source Heat Pumps

List of publications

Following publications have been made during the course of this PhD research.

Journal:

Heat coupling of the pan-European vs. regional electrical grid with excess renewable energy, **Asad Ashfaq**, Zulqarnain Haider Kamali, Mujtaba Hassan Agha, Hirra Arshid, Pages 363-377, Volume 122 (2017), **Energy**, ISSN 0360-5442, http://doi.org/10.1016/j.energy.2017.01.084.

Cost-minimised design of a highly renewable heating network for fossil-free future, **Asad Ashfaq**, Anton Ianakiev, Pages 613-626, Volume 152 (2018), **Energy**, ISSN 0360-5442, http://doi.org/10.1016/j.energy.2018.03.155.

Features of fully integrated renewable energy atlas for Pakistan; wind, solar and cooling, **Asad Ashfaq**, Anton Ianakiev, Pages 14-27, Volume 97 (2018), **Renewable and Sustainable Energy Reviews**, ISSN 1364-0321, http://doi.org/10.1016/j.rser.2018.08.011.

Investigation of hydraulic imbalance for converting existing boiler based buildings to low temperature district heating **Asad Ashfaq**, Anton Ianakiev, Pages 200-212, Volume 160 (2018), **Energy**, ISSN 0360-5442, http://doi.org/10.1016/j.energy.2018.07.001.

Modeling of wind energy conversion system using PSCAD/EMTDC, Zulqarnain Haider Kamali, Ziyad Salameh, **Asad Ashfaq**, Pages 178-187, Volume 7 (2017), **International Journal of Renewable Energy Research**, ISSN ISSN 1309-0127, https://ijrer.org/ijrer/index.php/ijrer/article/view/5245.

Conference:

Optimisation of Low Temperature District Heating Networks using Machine Learning Methods, **Asad Ashfaq**, Anton Ianakiev, November 2018.

Hydraulic control model for the implementation of LTDH in existing boiler based buildings, **Asad Ashfaq**, Anton Ianakiev, September 2017.

Chapter 1

Introduction

With introduction of the Climate Change Act 2008, the UK is liable to reduce greenhouse gas (GHG) emissions by at least 80% until 2050 to the levels recorded in the year 1990 ([Committee on Climate Change 2010]). This has made the UK a pioneer to recognise the importance of climate change and become a part of the 2016 climate change Paris agreement. To achieve these targets, the UK has introduced several clean growth measures and initiatives including; installation of wind, solar and nuclear power plants across the country, decommissioning of conventional fossil fuel power plants, and reduction of building sector emissions by improving thermal performance and efficiency of the heat network.

The UK has the largest offshore wind power capacity installed in the world and it has commissioned world's largest off-shore wind power plant which generates the record amounts of renewable electricity. This green energy investment has made UK world leader in renewable energy technologies. In 2016, the electricity came from low carbon sources has been almost twice the level in 2010. Since 1990, the average domestic energy consumption has decreased by 17% which is due to the energy efficiency measures. Currently, the government

plans to upgrade the thermal performance of existing domestic buildings and increase the energy efficiency of non-domestic building by atleast 20% until 2030. There are further plans to end the sale of new conventional petrol and diesel cars and vans by 2040 (UKERC 2009, BEIS UK 2018*a*).

The UK low carbon economy is predicted to grow by an estimated 11 per cent per year between 2015 and 2030, which is four times faster than the rest of the economy. This means that the low carbon economy would increase from around 2% of the UK's total output to around 8% by 2030 (BEIS UK 2018*b*).

In a recent study, it is discussed that the carbon emissions to heat our buildings alone contribute almost one third of total emissions, and the share of UK's heat demand is second highest in the European Union. The heat demand of the buildings is mostly provided by the natural gas-boilers and the proportion of district heating network is merely 2%. The UK has the potential to reduce these carbon emissions by readily increasing the share of district heating network to 14% especially in the central and south-eastern regions (Ashfaq & Ianakiev 2018*b*). This will provide UK with the increased flexibility, long-term security of supply and environmental benefits.

In principle, the district heating is mainly feasible in areas of high heat demand density. This is due to major capital cost and reduction in heat losses through the pipes of district heating network. The heat demand density is higher in areas with larger population usually around the city centres. Therefore, heat mapping provides invaluable information for the identification of high density areas in order to assess the feasibility of district heating implementation. The spatial heat demand distribution in the UK is shown in Figure 1.1 and is a vital resource for the future planning of district heating network.
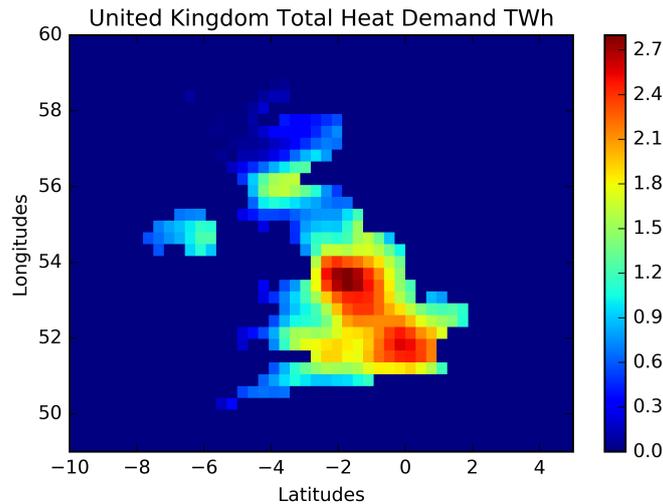
Figure 1.1: Figure shows the heat demand distribution in the UK spatial resolution of 40 x 40 km^2 for the year (2011) adapted from (Ashfaq & Ianakiev 2018*b*). The middle and eastern regions in the UK have higher heat demand compared to other regions. This highlights the importance of district heating in Nottingham and the case study used in this thesis.

Though, there are several small-scale district heating networks in the UK, the examples of large scale district heating networks are Westminster (London), Birmingham, Sheffield, Southampton, Woking, Aberdeen and Nottingham.

1.1 District heating

District heating is a network for the transmission and distribution of heat from source to the residential and commercial users using pre-insulated pipes. In modern district heating networks, the heat for space and hot water demand is supplied from a combination of centralised or decentralised sources.

Historically, the district heating was developed by the former soviet union where it was used as a heating source in east-European countries and Russia. This initial technology of district heating used water as steam (with temperature $\geq 200^oC$) for the space heating in buildings located in densely populated areas. This is also known as first generation district heating. But it did not

last long due to inefficient systems and lack of safety because of steam (Woods & Overgaard 2016). Later on, the same technology was adopted and revolutionised by Scandinavian countries. They introduced new techniques and used pressurised water as a heating medium instead of steam. This lowered the water temperature ($\geq 100^o$C) and reduced the losses and gave birth to the second generation of district heating. This incredibly reduced the heat losses and led to further lowering of water temperatures to utilise heat from other processes such as co-generation, waste heat and renewable energy sources. The third generation of district heating uses water temperature $< 100^o$C, and the fourth generation is ≤ 60 ^{o}C. In summary, the four generations of district heating are differentiated on the basis of supply water temperature from $\geq 200^o$C to $< 60^o$C (Lund et al. 2014).

1.2 Low temperature district heating system

The 4th generation (4GDH) district heating is also referred as low temperature district heating (LTDH). It's main parameter is to decrease the supply water temperature and maintain high delta t (Δ t), i.e. the difference between the supply and return water temperature in the network. In contrary to the conventional district heating network, the supply and return water temperatures are reduced from 80/40 ^{o}C to 50/20 ^{o}C. This reduces the heat-losses and provides with an opportunity for the integration of low-grade heat from renewable energy sources, waste heat from the industry into the district heating network.

As described above, the low temperature district heating system is a heat supply, transmission and distribution network with specific supply and return water temperatures in the network. According to the recommend guidelines, the operational design supply and return temperatures should be in the range
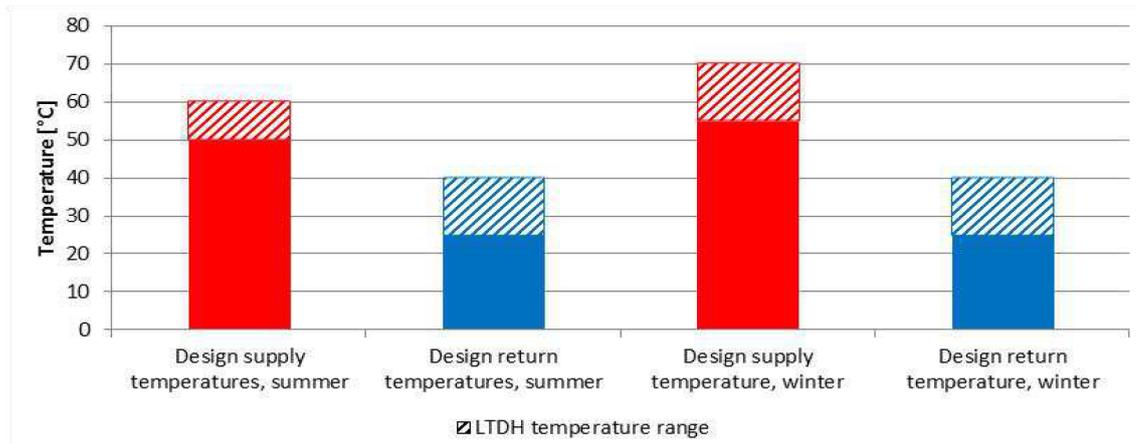
Figure 1.2: The range of design supply and return temperature of low temperature district heating according to the recommended guidelines (Olsen et al. 2014).

of 55-70°C and 25-40°C during winters, and 50-60°C and 25-40°C during summers as shown in Figure 1.2. Such a network should be able to meet consumer space heating and domestic hot water demand while maintaining optimum indoor thermal comfort as well.

The LTDH can be implemented in several ways to both new and existing buildings. The new developments can either be connected directly to the existing district heating or new standalone network. Likewise, the existing developments can also be connected to the existing network with or without the refurbishment of district heating network. These different network configurations have been shown in Figure 1.3.

1.2.1   Benefits of low temperature district heating

With recent carbon emissions reduction targets, the focus has shifted to increase the energy efficiency of the systems and reduce heat demand of buildings. These measures alone will increase the heat losses in the existing district heating network. The LTDH overcomes the heat losses issue by decreasing net-
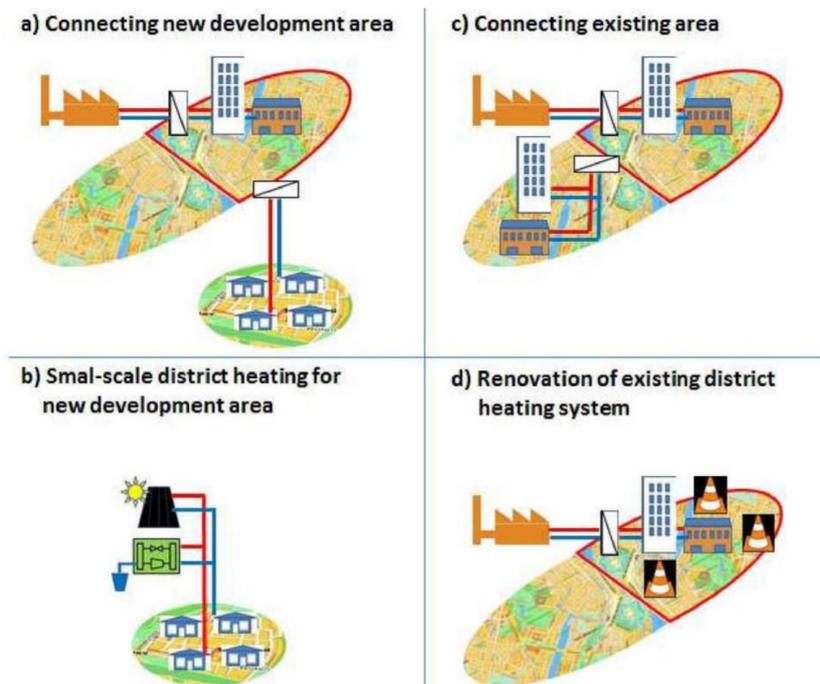
Figure 1.3: Implementaion of LTDH network in different situations (Olsen et al. 2014).

work temperatures and hence, this increases its competitiveness for low energy buildings in low energy density areas.

The district heating is commonly implemented using cogeneration. The cogeneration, also termed as combined heat and power (CHP), generates heat and electricity (power) simultaneously from a single fuel source with minimum losses. The conventional heat boilers and electricity power plants have combined losses of more than 44%. In comparison, the CHP has 20% losses for the same amount of electricity and heat generation. Figure 1.4 presents the energy flow using Sankey diagram.

The LTDH makes the possibility of using diverse low-grade heat sources to satisfy the base load heat demand such as, renewable energy sources, heat-pumps and waste heat from the industry. This makes it vital for long term resilience and security of heat supply, hence environmentally friendly. The
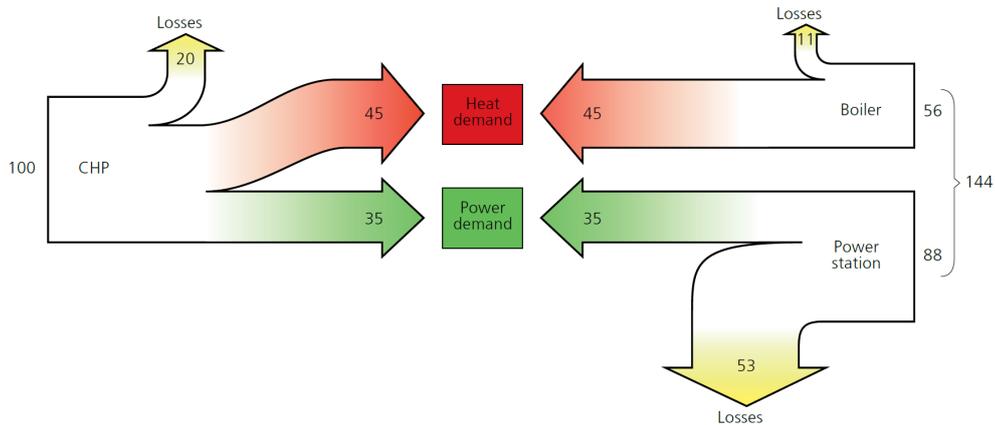
Figure 1.4: Sankey diagram illustrates CHP as more energy efficient compared to separate heat and power production ([Wiltshire et al. 2014](#)).

low water temperature in LTDH reduces the thermal stress on the pipes and ensures safety from pipe leakage, and its maintenance as well as enhances lifetime of the district heating network. The LTDH also helps space saving at the building level along with other benefits.

1.2.2 Challenges in implementation of LTDH

Even though, the low temperature district heating offers several benefits, there are still challenges associated with its application especially its implementation in existing buildings. The most important is the access to the real monitored data collection of the low temperature district heating. This provides insights to the operation of the district heating network and different parameters to maintain high delta t (Δ t), i.e. the difference between the supply and return water temperature in the network. The delta t (Δ t) also defines efficiency and heat losses from the district heating network.

Another challenge is controlling the temperature degradation across the low temperature district heating network. Its optimisation is important as the legionella growth and thermal comfort inside the buildings strongly depends

on it. Furthermore, the challenges associated with the correct thermal heat demand estimation, pipe sizing and network design define the issues of hydraulic performance and operation of the demand driven district heating network.

Finally, the economical and financial aspects of the low temperature district heating may impact the feasibility of the entire network. There are few other challenges, such as type of buildings, consumers, occupant behaviour and the stake holders involved, but in our view the above discussed four challenges are the most important and act as a barrier towards the transition to low temperature district heating.

1.3   Research aim and objectives

The aim of this research is;

'to investigate the implementation of low-temperature district heating in existing boiler based buildings'

The research question of this research are:

- To explore the impact of building regulations and investigate the thermal performance of existing and new buildings in order to achieve net-zero emission buildings (NZEB) in the UK.

- To investigate delta T for improving the energy efficiency and reduced heat losses in low temperature district heating network.

- To study the economic feasibility of renewable energy powered district heating network.

- To explore different machine learning and forecasting methods and application of most suitable technique for district heating network analysis.

- To determine the utility of GIS modelling for early stage planning and design of an optimum district heating network.

1.4   Research plan and research methods

The following diagram summarises the structure of this thesis. This thesis is split into three parts i.e. thermal performance modelling of buildings, hydraulic modelling of the district heating network and data analysis of monitored district heating network data. The underlying theory of different methods used in this thesis along with REMOURBAN LTDH project will be discussed in Chapter 2. Then, the building regulations and thermal performance calculations will be performed using the IDA-ICE and IES-VE software in Chapter 3. Subsequently, the hydraulic model of the REMOURBAN low temperature district heating network is developed and analysed in Dymola software in Chapter 4. The economic feasibility for a future fully renewable energy powered decarbonised district heating network will be examined in Chapter 5. Later, the machine learning and classical forecasting methods will be used to find the suitable method for forecasting district heating time series in Chapter 6. The role of GIS modelling in early stage planning and design of DH network will be elaborated in Chapter 7. Finally, the discussion on results and contributions to the knowledge will be discussed in Chapter 8.
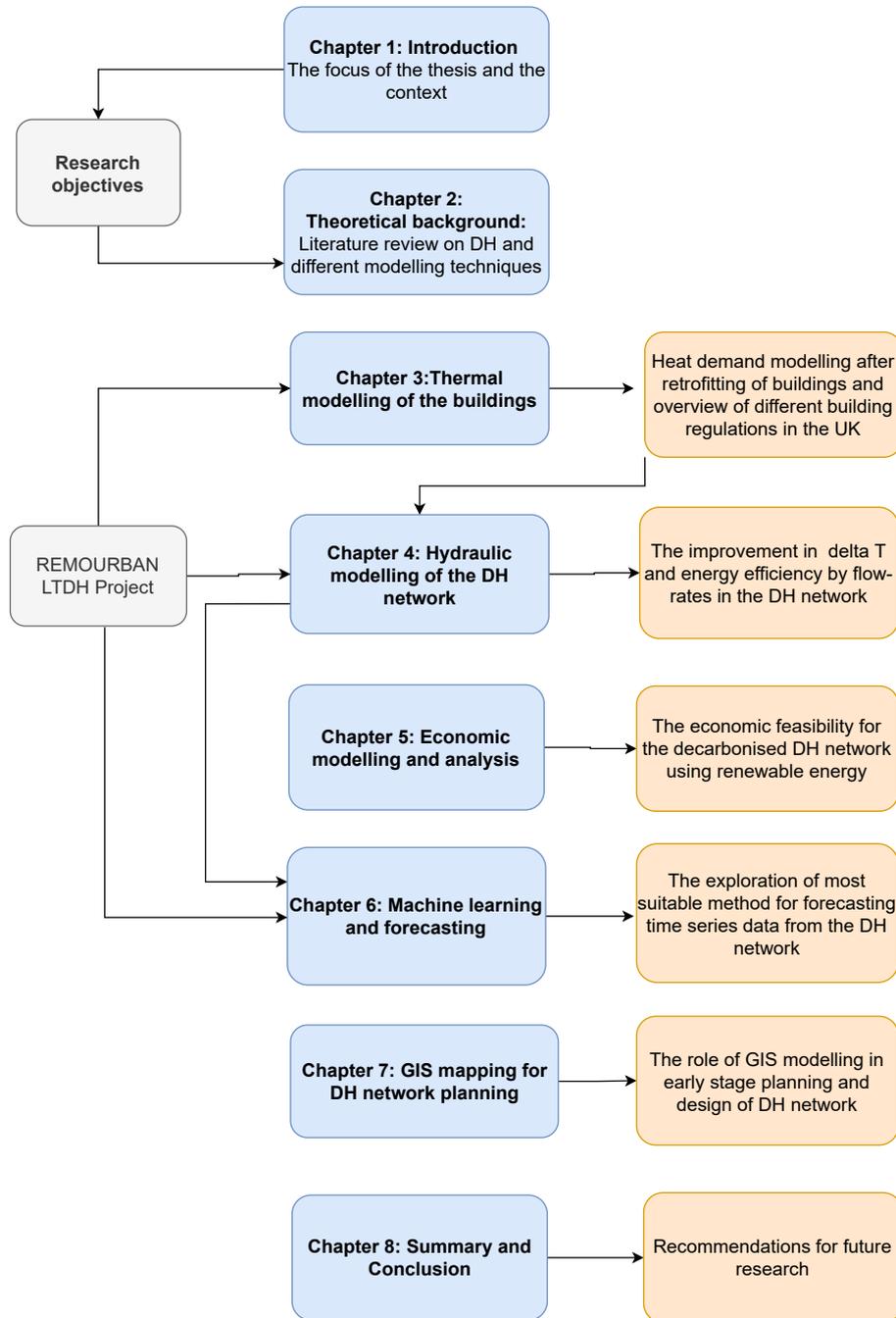
Figure 1.5: Graphical representation of PhD thesis structure.

Chapter 2

Theoretical Background

Overview

In this chapter, theoretical background of energy modelling techniques studied in this thesis is discussed. The overview of district heating progress in the UK and state of the art low temperature district heating intervention in Nottingham has been described. It is anticipated that this will be a gateway to achieve the UK's 2050 carbon emission targets. Later, the underlying thermal-hydraulic and economic modelling implemented in this thesis has been elaborated. This is followed by the theory of several classical stochastic and modern machine learning approaches for forecasting time-series data.

2.1 District heating in the UK

In EU, the heat demand of buildings represents 79% (192.5 Mtoe) of final energy consumption and United Kingdom has the second highest heat demand compared to other European countries (Ashfaq et al. 2017). The share of buildings' heat demand in the UK is almost 50% of the total energy consumption and the district heating is considered as central towards carbon neutral future

with zero emissions. In contrast to other European countries where heat is supplied by district heating, the building heat demand in the UK is usually supplied by individual gas-boilers, and the share of district heating is limited to only 2%. As per the Department of Energy and Climate Change (DECC), there are currently approximately 2000 district heating networks distributed across the UK, which supply approximately 5.5 TWh heat annually to almost 210,000 dwellings and 1,700 commercial buildings (BuroHappold Engineering 2016). Figure 2.1 shows the spatial distribution of district heating networks in the UK.
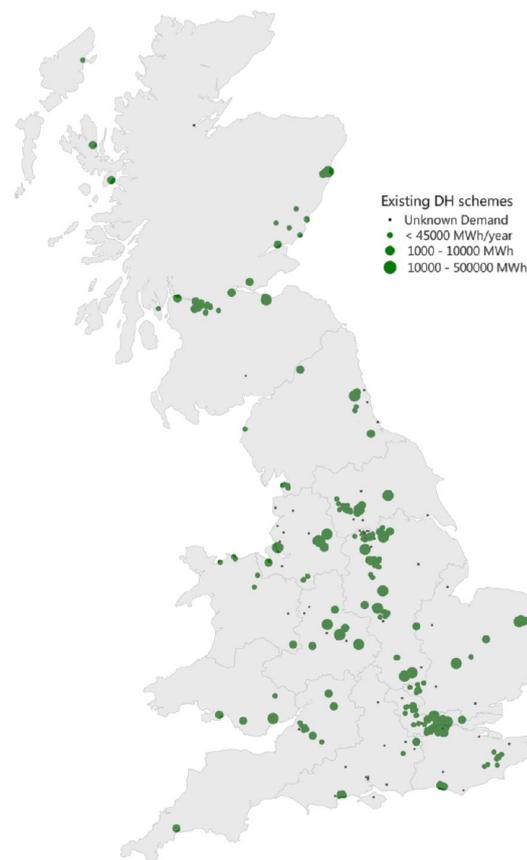
Figure 2.1: Spatial distribution of existing heat networks in Great Britain (Buro-Happold Engineering 2016).

Historically the implementation of district heating in the UK can be traced back to 1901 when the Bloom Street power plant in Manchester (UK) was built to supply steam in order to heat nearby warehouses and factories in the radius of 1.5 km (Woods & Overgaard 2016). This led to the gradual origination of several district heating networks such as, Westminster (Pimlico), Birmingham, Southampton, Sheffield, Nottingham, Woking and Aberdeen (Wiltshire et al. 2014).

The local authorities in the UK have always initiated the district heating network schemes with the aim to alleviate fuel poverty but not to generate heat for the entire district. Therefore, they have always transferred the operation to the private energy services companies (ESCO) for successful delivery. In 1986, the first large scale district heating network, Southampton District Energy Scheme, was established by Southampton City Council (SCC). It utilises geothermal energy from a deep aquifer with a heat pump together with a CHP plant and conventional boilers. Currently, this district heating network is 14 km of pipe length and generates over 40 GWh of heat, 26 GWh of electricity and 7 GWh of chilled water annually. Similarly, the Sheffield and Nottingham district heating network emerged in 1987 and 1989 as citywide district heating schemes with the main heat source as the municipal waste incineration. These two networks are still considered as the largest district energy network in the UK. Later, the city of Llanwddyn in Wales, established a biomass-fuelled district heating network to serves a variety of buildings. The brief time line for district heating network establishment in the UK has been shown in Figure 2.2.
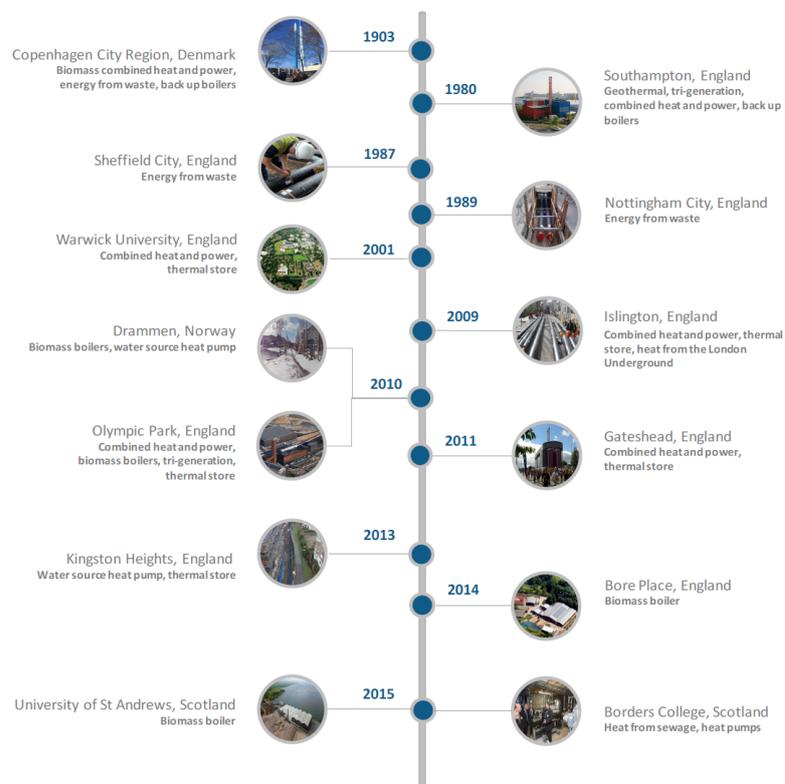
Figure 2.2: Timeline representing evolution of district heating in the UK (BEIS UK 2018*a*).

As discussed above that the UK has always considered combination of heat sources and technologies over the period of time. More specifically, these different heat sources include CHP (combined heat and power), deep geothermal, water for heat pumps, industrial waste heat, energy from waste incineration, renewable heat and urban recovered heat (London underground). Where the first district heating networks (such as Bloom Street Manchester) used steam as a heat source with an operating temperatures of over 200°C. On the other hand, the operating temperatures in recent district heating networks are lower than 40°C using low grade renewable energy as a heat source. This on-going evolution in the actual district heating technologies has led to the classification into different generations on the basis of network operating temperatures explained below and the difference among each generation is shown in Figure 2.3.

- 1st generation district heating > 200°C.

- 2nd generation district heating > 100°C.

- 3rd generation district heating < 100°C.

- 4th generation district heating < 60°C.

- 5th generation district heating < 40°C.

Among the above five generations, the 4th generation of district heating is currently popular in the UK. It is also know as low-temperature district heating and 4GDH.
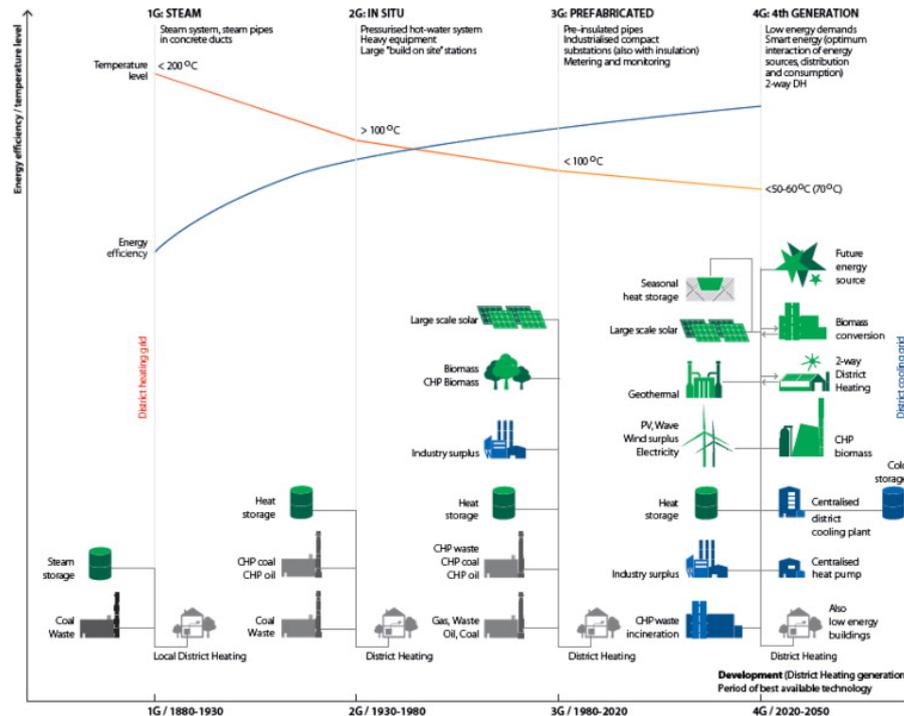
Figure 2.3: Illustration of the development in district heating technology and concept of 4th generation district heating compared to previous generations (Lund et al. 2014).

2.2   Decarbonisation of energy system pathways

The decarbonisation of heating sector and other energy system is crucial to address increasing environmental and sustainability concerns. The broad acceptance of carbon dioxide (CO_2) and other greenhouse gas (GHG) emissions for climate change has made decarbonisation an international policy priority. As part of wider international efforts, the UK Government recently set out a legal binding framework for the decarbonisation of economy from now to 2050 with substantial progress expected over the next decade. To this end, the government has published following three stage pathways in order to enhance the integration of low carbon technologies such as, renewable energy (marine,

bio-energy, wind and solar PV), carbon capture and storage (CCS), nuclear power and fuel cells (UKERC 2009).

1. Firstly, in the short term 2010-2030, the coal-fired plants will have carbon capture storage (CCS) implemented and electrical grid is decarbonarised by 80%.

2. Secondly, over the medium term to 2035, the focus will be on the de-carbonisation of power, heating, transport and residential sector. The bio-energy technologies will have significant impact on energy mix.

3. Thirdly, in the longer run to 2050, the proportion of low carbon technolo-gies will have a significant impact on our society. The share of energy from marine, solar PV and especially offshore wind power generation will increase, and fuel cells will have impact on the decarbonisation of transport sector.

The absolute decarbonisation is an immense challenge as our societies are built around carbon-based fossil fuels since many decades. In summary, the overall impact of accelerated development of low carbon technologies is not simple and straight forward, because of continuous changes in low carbon supply options and high ambitions to achieve overall decarbonisation. The en-ergy system decarbonisation involves a complex interaction between changing patterns of production, distribution and consumption. It may offer signifi-cant long-term benefits such as, alternative and affordable solution, diversity, security and sustainability.

2.3 District heating in Nottingham

Nottingham is the seventh largest metropolitan economy and ninth largest city of the UK. It has a population of around 321,550 and located in the eastern region of England known as East-Midlands. In 1989, the district heating network was initiated by British Coal Authority and currently managed by Enviroenergy Limited, owned by Nottingham City Council. The Nottingham's district heating network is considered as the largest in entire UK. The waste incinerator is used as heat source for the 68 Km of well-insulated pipe network which connects approximately 4,900 domestic and commercial users (Ianakiev et al. 2017).

The Eastcroft waste incinerator burns municipal waste to generate 52 tonnes per hour of steam at 371°C. It generates around 442 - 476 GWh of heat annually and provides steam to the Enviroenergy London Road combined heat and power (CHP) plant, where it generates electricity and transforms to medium temperature hot water. This pressurised hot water is used as a medium in the district heating network and distributed to consumers at a rated pressure of 11 bars and supply water temperature of between 85 - 120°C with seasonal variations. The district heating network has the installed capacity of 80 MW of heat and 14.4 MW of electricity. The current network operates with following priority:

1. Burning of waste in waste-incinerator.

2. Electricity generation.

3. Heat for the district heating network.

The heat generation and distribution in Nottingham has immense amount

of heat losses. Though, the waste incinerator generates around 442 - 476 GWh of heat annually, only 144 GWh of this heat is distributed through the district heating network and 60 GWh of heat is used for electricity generation. Recently, the (Ianakiev et al. 2017) has calculated that almost 21% of heat is wasted during heat transmission from waste-incinerator to the CHP plant and 36% is wasted as flue-gases during electricity generation at the CHP plant. The use of combined and heat power plant makes electricity and district heating network inter-related. It is anticipated that reduction in these heat losses will help in improving the efficiency of the entire network.

The recent climate targets has inspired Nottingham to be a pioneer in carbon-neutral city by the year 2028 and the district heating network is central to achieve city's ambitious targets. The short-term targets for the year 2020 includes 20% of energy from renewable energy and 26% reduction in carbon emissions. The waste incinerator burns around 170,000 tonnes of waste and the district heating scheme offsets approximately 27,000 tonnes of CO_2 emissions annually (Ianakiev et al. 2017).

2.3.1 Low temperature district heating intervention in Nottingham

The climate targets and heat losses in Nottingham district heating network motivates to search for alternatives. The high return water temperature of the existing district heating network provides the possibility of LTDH intervention to the surrounding regions, rather than extending high temperature network. Therefore, as part of the REMOURBAN project, the return water pipe (60^oC) of the existing district heating network passes through the heat-exchanger and used as a source for new secondary LTDH network with heat interface units

(HIUs) installed in each property. This LTDH network serves 94 residential properties located in four block of flats with the supply and return water temperature of 60^oC and 30^oC, respectively. Figure 2.4 shows block of flat for the LTDH implementation in existing boiler-based buildings.
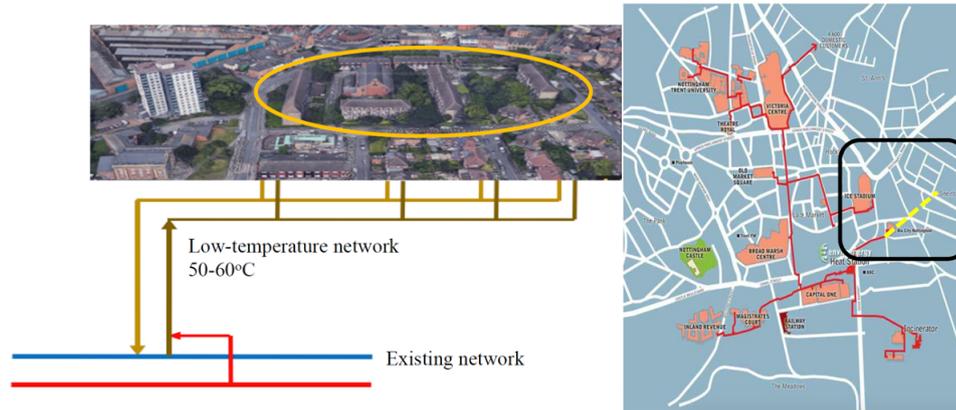
Figure 2.4: Left figure shows the basic concept of low temperature district heating intervention in REMOURBAN project - Nottingham. Right figure shows the 68 Km of Nottingham district heating network map.

The existing properties were installed with traditional natural gas boiler-based heating system and electrical heaters. As a pre-condition for the implementation of LTDH the building needs to have a good thermal performance. Since the buildings from the REMOURBAN project belongs to the Council estate, the building envelope needed to retrofit in order to achieve energy efficiency. The retrofitting included the improvement in building fabric, windows, doors, infiltration and ventilation. In addition to this, a private-wire electric supply powered by roof–mounted PV panels with battery storage has also been installed on each block of flats.

The REMOURBAN (REgeneration MOdel for accelerating the smart URBAN transformation) is a major Future Cities demonstrator project. This *EU Horizon 2020* project is EU funded for five years (2014-2019) to tackle issues related to energy, transport and ICT. This project has provided a gateway

to Nottingham for efficiency improvement and extension of existing district heating network. It is anticipated that heat losses of this secondary LTDH are 75% lower compared to the existing district heating systems. In this thesis, this first of it's kind LTDH network in the UK is investigated with following objectives.

- Reducing network's supply water temperature

- Increasing the difference between supply and return temperatures i.e. high delta t (Δ t).

2.4   LTDH network modelling methodology

The LTDH differentiates itself with utilisation of lower supply temperature and flow-rates which leads to lower pumping energy and energy consumption across the network.

Thermal demand modelling

The first step in designing and modelling a district heating begins with heat demand estimation. This is performed either by calculating energy consumption of the building (from SAP and SBEM method) along with degree day method for seasonal heat load variations, or simply implementing thermal modelling software for dynamic energy performance calculations such as, EnergyPlus, IDA-ICE, IES-VE, DesignBuilder etc.

The Standard Assessment Procedure (SAP) and SBEM (Simplified Building Energy Model) are standard methods for energy efficiency of the buildings calculations. The degree day method Equation (2.1) is commonly used for heat-load calculations (Küçüka 2007) and considers outdoor air (dry-bulb)

temperature for seasonal variations. On the other hand, the energy demand calculations from dynamic thermal modelling software are more accurate and the entire building has to be modelled using building parameters such as (orientation, exposed perimeters, U values of brick cavity walls and glass windows).

$$Q = Q_0 \left(\frac{T_{in} - T_{out}}{T_{in} - T_{out,0}} \right) \tag{2.1}$$

here T_{in} = indoor temperature(^{o}C), T_{out} = outdoor temperature(^{o}C) and $T_{out,0}$ = outdoor temperature(^{o}C) for design conditions.

The next step is the estimation of domestic hot water demand in designing district heating network.

Domestic hot water demand modelling

The diversity factor provides the instantaneous hot water demand of a dwelling and particularly useful to limit heat losses from the over-sizing of pipes. The domestic hot water demand is calculated using diversity factor from the Danish standard DS 439:2009 (Danish standard 469 2013) which is equivalent of the British standard BS 8558:2011 (Wiltshire et al. 2014) but with comparatively lower diversity factors. The expression for instantaneous domestic hot water demand calculations is given in Equation (2.2) and the diversity factor calculated from the DS 439:2009 standard with respect to number of dwellings is also shown in Figure 2.5.

$$P_{max} = 1.19 * N + 18.8 * N^{0.5} + 17.6 \tag{2.2}$$

where, Pmax(kW) is the maximum domestic hot water demand for the set of dwellings and N is the number of dwellings. According to the Equation (2.2)
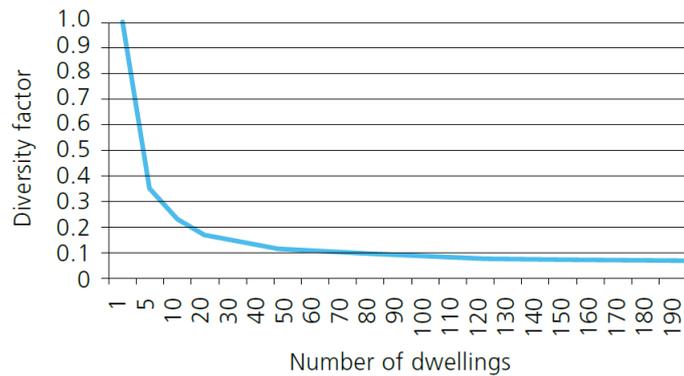
Figure 2.5: Diversity factors for the calculation of instantaneous domestic hot water demand in dwellings (Wiltshire et al. 2014).

the demand reduces with increase in number of dwellings and its 37.5 KW for a single dwelling. Once the number of dwellings increases beyond 200 the diversity factor is taken as 0.07. The hydraulic design and modelling is the next step after instantaneous domestic hot water demand calculations.

Hydraulic modelling

The hydraulic modelling is important for the operation and optimisation of the district heating network. The heat-losses in the network are reduced once the water-flow of entire network is balanced, also known as hydraulic balance, where the flow-rate and delta t (Δt) is according to the consumers heat demand. The hydraulic balance is pre-condition for achieving high delta t (Δt) in a well-functioning district heating heating network and problems with incorrect hydraulic balance leads to, low delta t (Δt) across the network, high hot-water supply temperature, higher flow-rate and heat-losses across the network as well as higher pumping cost and energy losses.

The network's design flow-rate and pipe sizes to fulfil the consumer heat demand are calculated from Equation (2.3) and the detailed methodology for pipe sizing is given later in Chapter 4 Section 4.2.2.

$$Q = \dot{m}\rho C(T_s - T_r) \tag{2.3}$$

where, Q = heat load (KW), T_s = network's supply temperature, T_r = network's return temperature, $\dot{m}$ = flow-rate (m^3/s), ρ = density (kg/m^3) and C = specific heat (kJ/kg^{o}C) of circulated hot-water. The Equation (2.3) is used to calculate the flow-rate in network with variations in heat demand and the delta t (Δt) i.e. supply and return temperature of heating network.

$$\dot{m} = K_v \sqrt{\Delta P_v} \tag{2.4}$$

K_v is the regulation capacity of a control or balancing value and ΔP_v (bar) is the controlled differential pressure across the value.

Economic modelling

The economic feasibility is determined using concept of levelised cost of energy (LCOE). The LCOE is the net present value of the unit cost from the energy generation technology over its lifetime. In comparison to other economic feasibility methods, it is more robust and provides a holistic comparison between different technologies. The (Rodriguez et al. 2015) used same method for the economic feasibility calculation of the pan-European electrical grid.

$$V = \text{CapEx} + \sum_{t=1}^{T} \frac{\text{OpEx}_t}{(1+r)^t} \tag{2.5}$$

$$\text{LCOE} = \frac{V_{sys}}{\sum_{t=1}^{T} \frac{L_t + H_t}{(1+r)^t}} \tag{2.6}$$

where, V is the net-present value of future expenses by the capacities required to fulfil the energy demand, r is the discount rate. L is the total generation in a year.

2.5 Time series prediction and forecasting

Over the recent years, several new machine learning methods have been developed to forecast time series data. These algorithms can be classified into pure machine learning and classical stochastic methods. In this thesis, both of these approaches will be applied on time series data to evaluate the best forecasting method for district heating network.

2.5.1 Machine learning methods

Traditionally, machine learning methods are categoried into three types; supervised learning, unsupervised learning and reinforced learning. Among these, supervised learning is the most common type of machine learning approach where predictions are made on known data. There are two kinds of supervised learning problems, which include classification and regression. Classification aims to predict a class label from a predefined set of possibilities. In contrast, regression tries to make predictions as a real or a continuous number. In this thesis, only regression based supervised learning methods have been explored. This is due to the nature of time series data where prediction is required as a real number.

2.5.1.1 K-nearest neighbors

K-nearest neighbors is the simplest machine learning method used for both classification and regression problem. In early 1970's, this method had numerous applications related to statistical analysis and pattern recognition.

The K-nearest neighbor method remembers the training dataset and computes the mean of K-nearest neighbors for prediction. In theory, the K-NN method requires optimisation of two parameters, the number of neighbors and the method for measuring distance between two data points in the training dataset. The Euclidean distance method is frequently used in regression problems, which is defined as,

$$d(X,Y) = \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2} \qquad (2.7)$$

The strength of K-NN method lies in its convenience, straightforward tuning and adjustment of parameters. But, it lacks the ability to handle large datasets with sparse dataset.

2.5.1.2   Linear models

The linear models are the oldest and largely studied machine learning models. In past decade these models have been further developed into multiple linear models. Due to their simple mathematical structure and ability to handle large and high-dimensional data-sets, they are considered as the first steps towards machine learning applications.

Ordinary linear regression

The ordinary linear regression (OLS) or simply linear regression is the oldest linear regression model with its origination dates back to over a century ago, even before the computers. The linear regression model expects the linear relation between the variables for parameters calculation and represented as,

$$f(X) = \beta_0 + \sum_{i=1}^{p} X_j \beta_j \qquad (2.8)$$

where, the variable X_j is an independent input vector, β_j are the unknown co-efficients or weights are calculated by the model itself to reducing mean squared error between the actual values and predictions. The OLS method derives its name from the estimation method, where the coefficients $\beta = (\beta_0, \beta_1, ..., \beta_p)$ which minimise the residual sum of squares (RSS) are calculated (Müller et al. 2016).

$$RSS(\beta) = \sum_{i=1}^{N}(y_i - f(x_i))^2 \tag{2.9}$$

$$RSS(\beta) = \sum_{i=1}^{N}(y_i - \beta_0 - \sum_{i=1}^{p} X_{i}j\beta_j)^2 \tag{2.10}$$

The linear regression model has no parameters to control over-fitting. The mean fitting error of the model is calculated from the residual sum of squares (RSS). Therefore, the below models employ shrinkage method to improve model fitting and prediction error.

Ridge regression

The ridge regression uses ordinary linear regression as base and applies penalty to the coefficient β to minimise residual sum of squares, also known as $L2$ regularisation. The ridge regression in equivalent Lagrangian form is written as,

$$\hat{\beta}^{ridge} = \text{argmin}_\beta \left\{ \frac{1}{2}\sum_{i=1}^{N}(y_i - \beta_0 - \sum_{j=1}^{p}x_{ij}\beta_j)^2 + \lambda\sum_{j=1}^{p}\beta_j^2 \right\} \tag{2.11}$$

where, λ is a penalty parameter and defines the amount of penalty on coefficients. With increase in λ, the coefficients are shrunk towards zero (Hastie et al. 2009). The ridge method can also be mathematically explained as,

$$\text{subject to} \sum_{j=1}^{p} \beta_j^2 \leq t \tag{2.12}$$

$$\hat{\beta}^{ridge} = \text{argmin}_\beta \sum_{i=1}^{N} (y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2 \tag{2.13}$$

The above equation uses $L2$ penalty to regularize the parameters.

Lasso regression

The lasso (least absolute shrinkage and selection operator) regression works on the same principle as ridge regression with minor but crucial differences. It controls the over-fitting of the model by implementing $L1$ regularisation, which uses subset selection of the data for prediction. This reduces the prediction error of the model compared to ordinary least square regression. In lasso the $L1$ penalties provides automatic feature selection. The coefficients are estimated with the addition of the penalty term (λ) to minimise the residual sum of squares (RSS) (Hastie et al. 2009). This is implemented using L1 regularisation i.e. $\lambda \sum_{j=1}^{p} |\beta_j|$ and reduces variance in the model at the expense of bias. The Equation (2.14) is expanded in the Lagrangian form in Equation (2.16).

$$\hat{\beta}^{lasso} = \text{argmin}_\beta \left\{ \sum_{i=1}^{N} (y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right\} \tag{2.14}$$

$$\text{subject to} \sum_{j=1}^{p} |\beta_j| \leq t \tag{2.15}$$

$$\hat{\beta}^{lasso} = \text{argmin}_\beta \sum_{i=1}^{N} (y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2 \tag{2.16}$$

where y is the target, x_j is the predictor, β_j is the regression coefficient, N is the total steps in the time series and p is the total predictors. The tuning of model depends on multiplier λ for regularisation (Müller et al. 2016) which controls the effect of coefficient used in the model. When λ is near 0 the model behaves similar to OLS regression and once the λ is high, the effect of coefficient is limited. Though ridge and lasso regression are quite similar, there are key differences as ridge uses $L2$ penalty $\sum_1^p \beta_j^2$ and lasso uses $L1$ penalty $\sum_1^p |\beta_j|$ (Géron 2019). The $L1$ penalty makes the solution non-linear.

Elastic net regression

The Elastic net regression provides middle ground between the ridge and lasso regression. It introduces elastic penalty for variables selection as lasso and shrinks the correlated coefficient as ridge regression. The elastic net penalty can be represented as Equation (2.17),

$$\lambda \sum_{j=1}^{p} (\alpha\beta_j^2 + (1+\alpha)|\beta_j|) \tag{2.17}$$

where, α controls the mixing ratio between the ridge and lasso's penalties. When $\alpha = 0$, the Elastic Net behaves as ridge regression and $\alpha = 1$ makes it equal to Lasso Regression (Hastie et al. 2009). In Equation (2.17) the term $(1+\alpha)|\beta_j|$ averages the highly correlated features and the term $\alpha\beta_j^2$ promotes the sparse solution of these averaged features.

Stochastic gradient decent regression

The stochastic gradient decent regression is a type of linear method which calculates the gradient at every step by picking random instances from training data-set. Due to its stochastic (random) nature, this method tends to be irregular and instead of gradually decreasing to reach minima, the loss (cost) function come into play which decreases the calculated averages by bouncing up and down (shown in Figure 2.6a). This approach helps to avoid local minima and achieves the global minima for the solution (shown in Figure 2.6b).
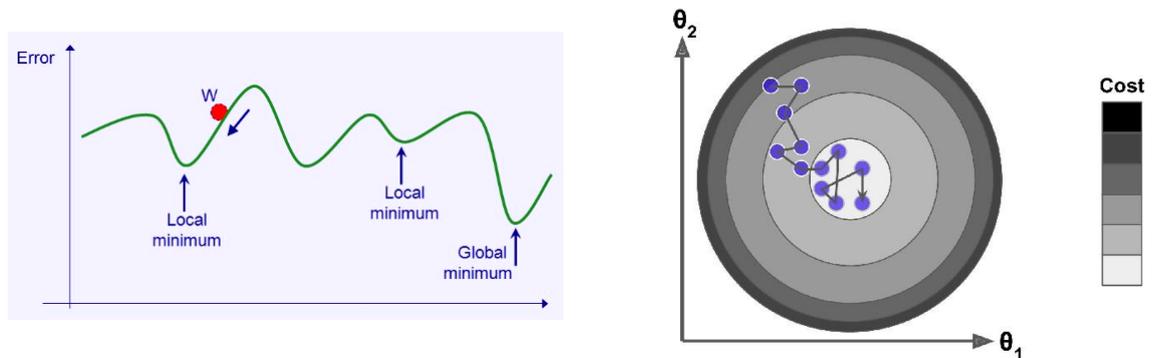
Figure 2.6: Stochastic gradient descent regression methodology to avoid local minima and achieves the global minima (Géron 2019).

During the process of reaching global minima, the algorithm struggles to settle at the minimum due to randomness. The solution to this problem lies in reducing learning rate. The learning rate defines larger steps in the beginning of calculations which decreases gradually and allows the algorithm to settle at the global minimum. This entire process is also referred as simulated annealing. The loss (cost) function is computed using the Equation (2.18), which is used in Equation (2.19) to calculate the weights in the model.

$$L = (\hat{y} - y)^2 \tag{2.18}$$

$$w_{ij}(t+1) = w_{ij}(t) + \eta \frac{\partial L}{\partial w_{ij}} \tag{2.19}$$

where, w_{ij} is the weight of the linear equation, η is the learning rate and L is the loss (cost) function for which a smaller value of L is preferred. The choice of loss (cost) function depends on the type of supervised learning method.

2.5.1.3   Decision trees

Decision trees is a well known predictive model for its wide application in classification and regression problems on time series data. The regression decision trees use real numbers as a prediction outcome, compared to the classification decision trees where the outcome represents a class to which the data belongs. The outcome from the decision trees methods depends on a hierarchy of if and else questions which can be represented in a form of a tree. The tree grows using training data by greedy recursive binary splitting (Müller et al. 2016). Each node in a tree either contains a question to be asked or a terminal node (also known as a leaf). The edges of a tree links the answers to a questions (Géron 2019). One such example tree with two features X1 and X2 is shown in Figure 2.7.

In this example, the feature Xi and split cutting point tj is selected at each node to minimise the prediction error. The prediction score in leaf nodes (R1-R5) represents the mean output from set of training data splits stepping down the tree. Hence, overall predictions of a tree depends on the the values from the leaf nodes. The stopping point of a tree is where the node does not split any further and turns into a leaf node. This is when the reduction in prediction error is less than the user-defined tolerance.
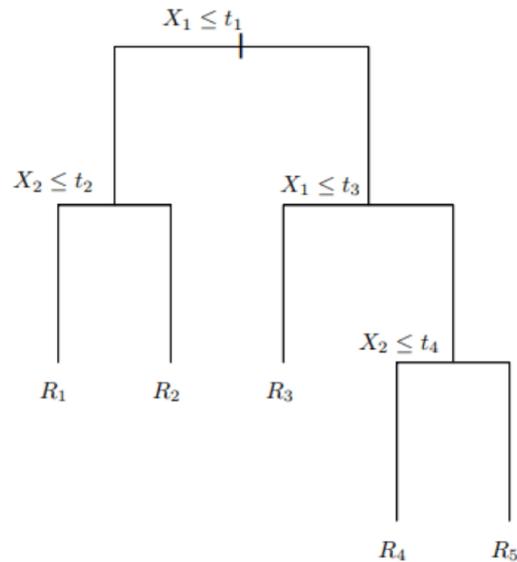
Figure 2.7: The concept of decision tree method with two input features.

Pruning

Pruning is a process to control the size of a tree, which prevent over-fitting of the model. This can be done in 2 ways; pre-pruning (or pruning) and post-pruning. In pre-pruning, tree growth is stopped at an earlier stage whereas in post-pruning a large tree is built, but less informative nodes are collapsed or merged to reduce the tree size and complexity. In principle, a large tree provides good results on the training data but the model tends to overfit on test data, and pruning prevents it by minimising variance at the cost of a slight biasness, similar to regularisation. Most common ways of performing pruning are restricting the depth of tree and limiting number of terminal nodes (leaf). The pruning is also used to fine tune the model.

2.5.1.4 Ensemble Method

Historically, the aggregate predictions using a group of prediction methods often provides a better answer than the individual predictor. The ensemble methods are examples of such grouped predictors.

Random forests

Random forest is an ensemble method which combines decision tree with bagging. In bagging method, the training data is split into multiple subsets and multiple predictions are made using same algorithm. The individual predictor methods tend to have a higher bias and variance, which can be reduced by aggregating multiple predictions (Géron 2019, Poutiainen 2019). This is the key property of bagging. Each tree is not pruned individually rather allowed to grow deeper, this results into high variance which is reduced by averaging predictions (in regression) from each tree. The predictions using bagging on bootstrapped training set on trees is calculated by:

$$\hat{f}_{bag}(x) = \frac{1}{B} \sum_{b=1}^{B} \hat{f}_b(x) \tag{2.20}$$

where $\hat{f}_b(x)$ represents the prediction of each tree b. The most important parameters of random forest is the tree depth (number of features) and the number of trees (no. of estimators), which are helpful in fine tuning the model. The predictions from random forest are always better compared to the decision trees due to bagging, but this also increases the model complexity and makes the results interrogation difficult.

Gradient boosted regression trees

Gradient boosted regression tree (GBRT) is an ensemble method which combines decision trees with boosting method. The trees are built in a series and each tree corrects the errors of the previous tree. Generally, the boosting method trains predictors sequentially, each reducing the prediction error by

correcting mistakes of its predecessor. The length of these trees tend to be shallow with a depth of one to five, and controlled by strong pre-pruning.

In gradient boosting, each tree performs good predictions on a part of the data, therefore more trees are iteratively added. The combined predictions of shallow but more number of trees help to improve the overall performance. Adding more trees in the ensemble will increase model complexity, but at the same time there are more chances of error correction on training data. In addition to pre-pruning and number of trees, a third parameter of prime importance is learning rate, which controls the ability of trees to correct the mistakes identified in the previous trees. A higher learning rate increases the complexity of model and therefore this parameter requires careful optimisation. The GBRT models are sensitive to the parameters and tuning model to optimum parameters is the key for prediction. The accuracy is better if parameters are set correctly. Therefore, the first preference in industrial applications is given to the decision trees and random forest methods, then gradient boosted regression trees.

2.5.1.5   Support vector machines

The support vector machines (SVM) are commonly used in supervised machine learning for both classification (SVC) and regression (SVR) problems. The idea in SVM is to create a hyper-plane (virtual boundary) to distinguish between different clusters of data based on their characteristics. The SVM algorithms determines the hyper-plane parameters to keep the maximum virtual boundary between the different sample of data points.

In this thesis, the support vector regression (SVR) models are used. Therefore, the theoretical background of support vector regression is discussed.

The SVR model determines the optimum parameters for function to predict the actual time-series observations with an error tolerance. The concept of SVR is evolved from the multiple linear regression function:

$$f(x) = x^T\beta + b \qquad (2.21)$$

where, x^T is collection of independent variables for set of samples, β is the vector weights (regression coefficients) and b defines the vector of constants. The β and b are calculated to keep the ϵ-insensitive residuals to minimum. This is computed from the empirical risk (R_{emp}) given by (Agrawal & Adhikari 2013), (Poutiainen 2019),

$$R_{emp}(\beta, b) = \frac{1}{N}\sum_{i=1}^{N}V_\epsilon(y_i - f(x_i)) + \frac{\lambda}{2}||\beta||^2 \qquad (2.22)$$

where, y_i is target observation i, $\frac{\lambda}{2}||\beta||^2$ is L2 norm penalty and V_ϵ is the insensitive loss function of ϵ with following constraints,

$$V_\epsilon(r) = \begin{cases} 0 & if\,|r| < \epsilon \\ |r| - \epsilon & otherwise \end{cases} \qquad (2.23)$$

This elaborates that the insensitive loss function V_ϵ is 0 once the absolute residual value r is smaller than ϵ and this relation is shown in Figure 2.8. In practical applications the data is mostly not linearly separable. Therefore, the SVR constructs a 'Soft Margin Hyper-plane' to allow feasible constraints. The values of β and b to minimise the empirical risk are found through the following parameters optimisation (Agrawal & Adhikari 2013), (Poutiainen 2019),
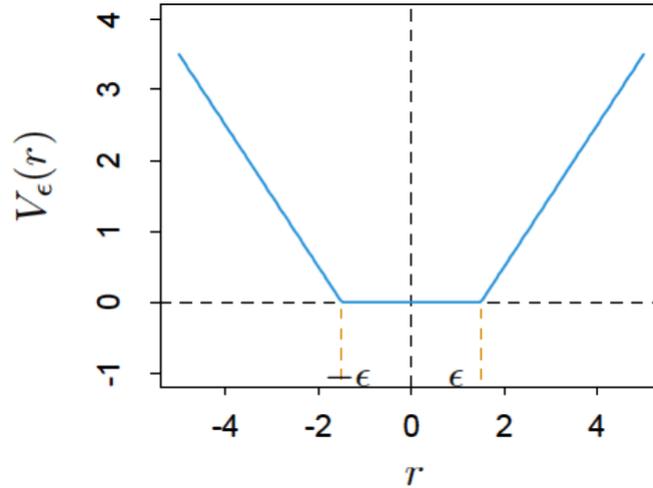
Figure 2.8: The relation between insensitive loss function (V_ϵ) and absolute residual value (r) for SVR method ([Poutiainen 2019](#)).

$$min_{\beta,b,\xi_i,\xi_i^*} = \frac{1}{2}||\beta||^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \tag{2.24}$$

subject to:

$$
\begin{aligned}
y_i - f(x_i) &\leq \epsilon + \xi_i, \quad \forall i \\
f(x_i) - y_i &\leq \epsilon + \xi_i^*, \quad \forall i \\
\xi, \xi^* &\geq 0, \quad\quad \forall i
\end{aligned}
\tag{2.25}
$$

where C is a regularisation constant used as penalty of mis-classification, ξ_i, ξ_i^* are slack variables to relax the hard-margin constraints. The above optimisation problem is know as the primal formulation and can be solved computationally efficient in terms of support vectors α_i, α_i^*. The support vectors are found from the dual formulation optimisation problem given by ([Poutiainen 2019](#)):

The two set of Lagrange multipliers are used which are $\alpha = (\alpha_1, \alpha_2, ..., \alpha_N)^T$ and $\alpha^* = (\alpha_1^*, \alpha_2^*, ..., \alpha_N^*)^T$, where $0 \leq \alpha_i, \alpha_i^* \leq C$. For

support vectors $0 < \alpha_i, \alpha_i^* \leq C$. Finally, the optimum decision hyperplane is obtained as (Agrawal & Adhikari 2013):

$$\hat{f}(x) = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)K(x, x_i) + b \tag{2.26}$$

The kernel function $K(x, x_i)$ is used for the so called kernel trick and different kernel functions are presented in Table 2.1.

Table 2.1: List of Kernel functions for support vector machines

Kernel	Function
Linear	$K(x, z) = x^T z$
Gaussian (RBF)	$K(x, z) = e^{-\|x-z\|^2}$
Polynomial	$K(x, z) = (1 + x^T z)^q, q \in N$

2.5.1.6   Neural networks

In 1943, neurophysiologist Warren McCulloch and the mathematician Walter Pitts introduced the concept of neural networks. They presented a computational model for biological neurons to perform complex computations using propositional logic. Since then, neural network have developed several methods for classification and regression problems. Recently, they are referred as "deep learning" as well. In this analysis, the Multilayer perceptrons (MLPs) method of neural networks is discussed and also known as (vanilla) feed-forward neural networks.

Multilayer perceptrons (MLPs)

The neural networks or multilayer perceptrons (MLP) are the combination of multiple layers of sigmoid perceptrons, where nodes of perceptrons are connected with each others. The limitations in traditional neural networks are removed by implementing multiple hidden layers of perceptrons in MLP method.

The neural network with more than two hidden layers is known as deep neural network (DNN). The MLP has several hidden layers which contains weight and bias neurons, except the output layer. This method is a feed-forward neural network (FNN) in which the predictions are made from the input to output direction.

Input Layer $\in \mathbb{R}^7$ Hidden Layer $\in \mathbb{R}^5$ Hidden Layer $\in \mathbb{R}^5$ Output Layer $\in \mathbb{R}^1$
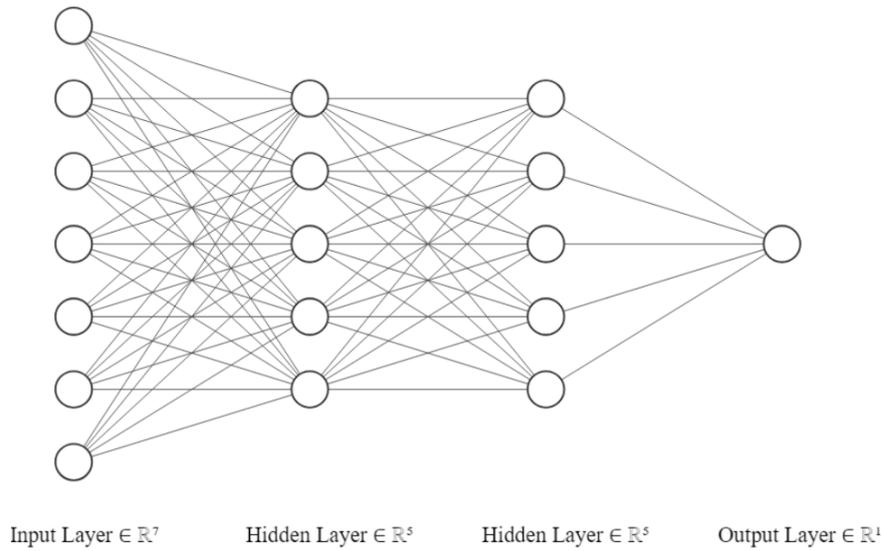
Figure 2.9: Feed forward, fully connected neural network with 7 input features, 2 hidden layers with 5 hidden units each and 1 output (Poutiainen 2019).

This is explained in Figure (2.9), where a MLP based neural network is shown, X is the input, $\hat{Y}$ is the output, and the weight w and bias b values are represented by connections between the nodes of perceptrons. The network has two hidden layers Z and T with units M and N, respectively.

$$\bar{Z} = [Z_1, Z_2, Z_3, ..., Z_M] \tag{2.27}$$

$$\bar{T} = [T_1, T_2, T_3, ..., T_N] \tag{2.28}$$

$$Z_m = \sigma(b_{0,m} + w_{0,m}^T X), m = 1, 2, 3, ..., M \tag{2.29}$$

$$T_n = \sigma(b_{1,n} + w_{1,n}^T \bar{Z}), n = 1, 2, 3, ..., N \tag{2.30}$$

where, $b_{0,m}$ is the bias and $w_{0,m}^T$ is the weight between the input and hidden layer Z with units m, the $b_{1,n}$ is the bias and $w_{1,n}^T$ is the weight between the hidden layers Z and T with units n. The percepton activation function σ controls the bias value and output from the node. The hyperbolic tangent σ_{tanh} and Rectified Linear Unit (ReLU) σ_{ReLU} are two commonly used activation functions and given below Equations (2.31) and (2.32).

$$\sigma_{tanh} = \frac{2}{1 + e^{-2x}} - 1 \tag{2.31}$$

$$\sigma_{ReLU} = \frac{1}{1 + e^{-x}} \tag{2.32}$$

The final outputs from the model is given by:

$$\hat{Y} = g(\bar{b}_2 + w_2^T \bar{T}) \tag{2.33}$$

where, $\bar{b}_2$ and w_2^T is the bias and weight between the hidden layer T and output, g is the output layer function. The neural networks are trained by using back-propagation method to estimate the optimum values for weight and bias. In back-propagation method the model first assumes random weight values for prediction and calculates prediction error from each node. Then, increases the weight and bias of each node to reduce the mean squared error (MSE) of the

final prediction. This is commonly referred as Gradient Descent method. The mathematical expression for increasing the weight of each node is given as,

$$w_{i,j}(t+1) = w_{i,j}(t) + \eta(Y_j - \hat{Y}_j)X_i \qquad (2.34)$$

where, $w_{i,j}$ is the node weight between the ith input neuron and the jth output neuron, t is the training instance in the time-series, X_i is the ith input value of the current training instance, $\hat{Y}_j$ is the output of the jth output neuron for the current training instance, Y_j is the target output of the jth output neuron for the current training instance and η is the learning rate.

The tuning of neural network hyper-parameters is crucial to the model's prediction performance. The complexity of the neural network generally depends on number of hidden layers and neurons. This makes the training of model bit slow, particularly for large datasets. Therefore, techniques such as cross-validation and early stopping is used to control model's complexity and over-fitting (Géron 2019), (Poutiainen 2019). The following hyper-parameters are required to be tuned for optimum neural network model.

- Number of hidden layers

- Number of neurons per hidden layer

- Activation functions

- Learning rate

- Learning algorithm

- Regularisation

2.5.2   Stochastic methods

In forecasting the historical data is used for future time series predictions. A time-series is a sequence of real time observations (integer or floating point) with respect to time. Generally, there are two main types of time-series, uni-variate and multi-variate. In uni-variate, there is only one variable in the time-series whereas, there can be multiple variables or features in multi-variate time-series. The classical time-series forecasting methods assumes linear relationships between model parameters, but interestingly their performance on ordinary uni-variate is better as compared to other sophisticated machine learning methods.

In this analysis, classical stochastic forecasting methods are used on uni-variate time-series data, which includes Autoregressive (AR), Moving Average (MA), and Autoregressive Integrated Moving Average (ARIMA). Before discussing forecasting models, it is necessary to describe the steps involved for dataset preparation.

Stationarity of time-series Data

A time-series is termed as stationary when its statistical properties (mean, variance and standard deviation) remains constant over time. In other words, there should be no trend or seasonality. Before implementing the autoregressive methods (AR, MA, ARMA and ARIMA models), it is necessary to have the time-series as stationary and checked using 'Dickey-Fuller Test'. The stationary of timeseries enables to extract the hidden patterns in the underlying data and helps to make correct statistical assumptions for forecasting. Therefore, the non-stationary time-series should be first transformed into stationary using methods such as aggregation, smoothing, polynomial fitting and differ-

encing. In ARIMA method, the differencing is used to effectively transform the time-series into stationary, and the SARIMA additionally incorporates seasonal differencing into ARIMA to model time-series data with periodic characteristics.

Dickey-Fuller Test

The Dicky-Fuller test is used to evaluate stationarity of the data. It assumes that the time-series is non-stationary and compares 'Test Statistic' results and 'Critical Value' at several confidence levels. The time-series is said to be stationary if the 'Test Statistic' is less than the 'Critical Value'.

2.5.2.1   Autoregressive models (AR)

The autoregression (AR) models forecasts the next step in the sequence by using number of lagged observations and dependencies between observation. The autoregression (AR) considers the linear relation between parameters and its lagged values (Jacob et al. 2020).

The autoregressive model is denoted as AR(p), where p represents the order of the 'Auto Regressive' (AR) term and refers to the number of lags of time-series considered by the model for prediction. The Equation (2.35) shows a basic form of AR model with an order p where AR(p) is represented as a linear process.

$$y_t = a + \sum_{i=1}^{p} \phi_i y_{t-i} + \varepsilon_t \qquad (2.35)$$

where, y_t is stationary variable, a is constant, p is number of lag measurements, ϕ_i is autocorrelation coefficient used in estimation and ε_t is Gaussian white noise series with zero mean and variance σ_ε^2 (Jacob et al. 2020).

2.5.2.2   Moving average models (MA)

The moving average (MA) models forecasts the next step in the sequence by using linear combination of number of previous errors (residual), the expected value and a random noise (Jacob et al. 2020). The moving average model is denoted as MA(q), where q represents the order of the 'Moving Average' (MA) term and refers to the number of lagged residual error considered by the model for prediction. The Equation (2.36) shows a basic form of MA model with an order p where MA(q) is represented as a linear process.

$$y_t = \mu + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \tag{2.36}$$

where, μ is expected value of y_t, the θ_i is weights, q is the number of historical error values and ε_t is Gaussian white noise series with with zero mean and variance σ_ε^2 (Jacob et al. 2020).

2.5.2.3   Autoregressive integrated moving average (ARIMA)

The Autoregressive Integrated Moving Average (ARIMA), also known as Box and Jenkins forecasting, is the most widely used forecasting method. The acronym ARIMA originates from Autoregression AR(p), Moving Average MA(q) models and integration I(d) for differencing. The ARIMA models forecasts the next step in the time-series as a linear function of previous observations (lags) and residual errors. It is combination of autoregressive and moving average models with differencing. The mathematical representation of the model is denoted as,

$$y_t = a + \sum_{i=1}^{p} \phi_i y_{t-i} + \mu + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \tag{2.37}$$

The behaviour of model is defined by the order of AR, MA, I terms in the following notation Equation (2.38), the model can work either as AR, MA, ARMA or ARIMA.

$$ARIMA(p, d, q) \tag{2.38}$$

where;

- p : the autoregressive order. It defines the number of previous observations.

- d : the difference order. It defines the number of differences in order to make time-series stationary.

- q : moving average order. It defines the number of previous forecast errors for moving average.

The order of AR, MA, I terms is selected by either using auto-correlation (ACF) and partial auto-correlation (PACF) plots or dedicated grid search algorithms for fine tuning of hyper-parameters such as, Pyramid Arima library in Python.

Autocorrelation and partial autocorrelation plots

The Autocorrelation (ACF) and Partial Autocorrelation (PCAF) plots helps to identify p and q values based on correlation of previous time steps. More specifically, the ACF is an estimation of linear dependence between time step observations that are represented by a lag p, whereas the PACF determines number of autoregressive terms q. The Pearson's correlation coefficient is used to access the relationship between two variables, and this is represented as a

number between 1 and -1 describing positive and negative correlation, respectively.

The ARIMA modelling has a limitation to perform well on seasonal data. So, if the time-series has a repeating pattern (seasonality), ARIMA will not be a good choice. For a time-series with seasonal element, SARIMA is used which is a extension of ARIMA.

2.5.2.4   Seasonal autoregressive integrated moving average (SARIMA)

The Seasonal Autoregressive Integrated Moving Average (SARIMA) model uses ARIMA model as base and extends the results with the inclusion of seasonal autoregression, differencing, and moving average. This extension makes the model notation as,

$$SARIMA(p,d,q)(P,D,Q)m \tag{2.39}$$

where, the order for the AR(p), I(d), and MA(q) is same as in ARIMA, and the order for AR(P), I(D), MA(Q) and m are at the seasonal level which are as follows:

- P: Seasonal autoregressive order.

- D: Seasonal difference order.

- Q: Seasonal moving average order.

- m: The number of time steps of the seasonal period.

The order of these seasonal parameters are difficult to calculate manu-

ally, therefore the dedicated grid search algorithms are used (*Pyramid SArima* library in Python).

2.6   Model optimisation and evaluation

2.6.1   Generalization, over-fitting and under-fitting

In supervised learning, a model is developed using training data in order to make precise predictions on unseen data with similar properties. The model's performance has to be generalised to make accurate predictions on new data. The main goal is to build a model which is generalised enough to make as accurate predictions as possible. The future performance of a model is assessed on a unseen testing data. Therefore, training and testing datasets are two key components of model building.

Theoretically, simple models tend to generalise better on new data. Hence, it is best practice to start with building a simplest model. This is because building a model which is too complex for the available information can lead to over-fitting of model. In this situation, the model fits (generalised) too closely to the characteristics of training data and works entirely well on the training dataset, but is unable to work on unseen testing dataset. On the other hand, if a model is too simple to learn all the different aspects in the data, it will not perform well on the training data. Selecting such a simplistic model can lead to under-fitting. However, there is an optimum position (also called sweet spot) where model can have its best performance (generalisation). This is shown in Figure 2.10. This is most important step of model building and optimisation which is achieved with the help of fine tuning of parameter. In supervised learning, a large number of data is important for model building.

This is because it can help in building a fairly complex model. Sometimes, having access to more data can work better than tweaking the parameters to fine tune the model.
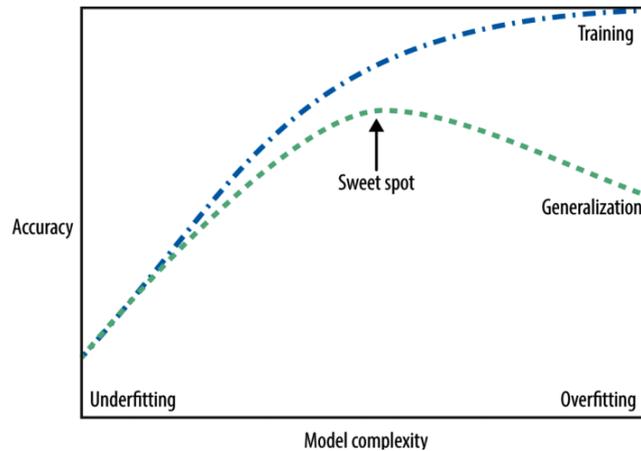
Figure 2.10: Trade-off of model complexity against training and test accuracy (Müller et al. 2016).

2.6.2 Grid search and Cross-validation

A model's generalisation performance can be increased by fine tuning key parameters. The grid search method explores all possible combinations of parameters and creates a table containing model's performance with different settings of parameters. This helps to select the most suitable parameters effectively as well as quickly and shown in Figure 2.11.

In order to assess the generalisation performance of a model, a statistical re-sampling method, cross-validation, is used which repeatedly split the entire dataset into a training and testing data. Using this method multiple models are trained. The k-fold cross-validation is the best known and commonly used type of cross-validation, where k represents number of times the data will be split and mostly selected 5 or 10. In case of 5-fold cross-validation, the data is split into five sets (folds) of nearly equal sizes. Among five folds, two to
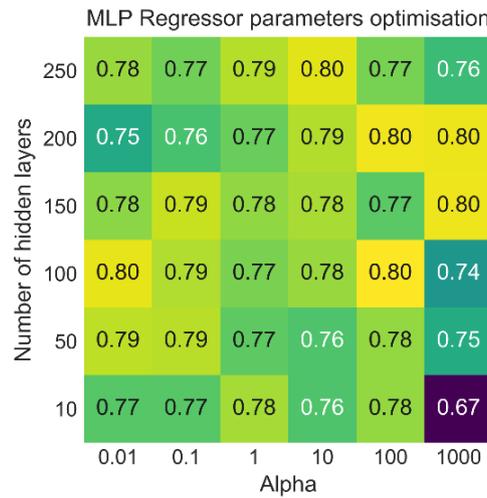
Figure 2.11: Heatmap representing example of a Grid search method to find optimum model parameters

five folds (training sets) are used to build and train the model while first fold (testing set) is used to evaluate the model's accuracy. The next model is built using fold 2 as the testing set and remaining folds (1, 3, 4, 5) as the training set. This procedure is repeated using folds 1, 3, 4 and 5 as testing sets. The accuracy is computed five times separately for each simulation of the model. This process is shown in Figure 2.12.
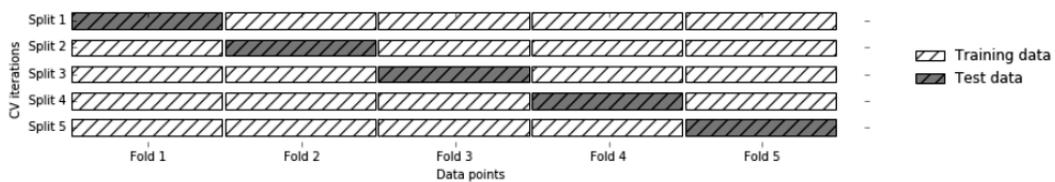
Figure 2.12: Data splitting with five-fold cross-validation (Müller et al. 2016).

The cross-validation and grid search are commonly used together as this provides a better estimation of the generalisation performance as compared to using a single split into training and a validation set. In this hybrid approach, cross-validation evaluates the performance with each parameters combination and helps in adjustment of model parameters.

One of the major benefit of cross-validation lies in its random nature of splitting the data. By having multiple splits of same data enables to assess the sensitivity of model towards selection of the training dataset. This also make the usage of data more effectively. However, no method comes without a disadvantage. The major disadvantage of cross-validation is higher computational cost and slow efficiency in comparison to using a single split of the data.

2.6.3   Model evaluation metrics

The regression based machine learning models can be evaluated using different scoring metrics. Most commonly used scoring metrics are; mean absolute error (MAE), mean absolute percentage error (MAPE), mean squared error (MSE), mean square percentage error (MSPE), root mean square error (RMSE) and R^2. In this thesis, only R^2, MAE, MAPE and RMSE have been used.

2.6.3.1   Mean absolute error (MAE)

The mean absolute error (MAE) provides the estimation of error in predicted values by calculating average of the absolute differences between the actual and predicted value. It calculates the error as an absolute value which is of linear nature because average of all absolute differences obtain an equal weight. This measure does not provide information about over or under prediction. Mathematically, it is computed as:

$$MAE = \frac{1}{N} \sum_{t=1}^{N} \left| Y_t - \hat{Y}_t \right| \qquad (2.40)$$

where, N is the total number of sample values, Y_t is the actual expected value and $\hat{Y}_t$ is the model's predicted value. This score has the ability to penalise huge errors, however, it is not sensitive to outliers value in data itself.

2.6.3.2 Mean absolute percentage error (MAPE)

The mean absolute percentage error (MAPE) is a relative error and obtained by dividing absolute error by actual value. This is considered as weighted version of MAE, therefore the optimal constant predictions for MAPE are assumed as weighted median of actual values. The MAPE is very useful in evaluation of forecasting performance. Mathematically, MAPE is computed as:

$$MAPE = \frac{1}{N} \sum_{t=1}^{N} \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| * 100 \qquad (2.41)$$

where, N is the total number of sample values, Y_t is the actual expected value and $\hat{Y}_t$ is the model's predicted value. The weight of actual sample is inversely proportional to it's predicted value. For a very small outlier value in data, MAPE would have a high biasness due to larger weight.

2.6.3.3 R-squared (R^2)

The R^2 metric is also called as the coefficient of determination. It provides an estimation of goodness of fit between fitted values by the model and actual expected values in dataset. This measure is scale free, and is not biased to the very large or small values in the dataset. This is because R^2 value between 0 and 1 determines whether the model fit is poorest or excellent fit. The Equation (2.42) is used to compute R^2.

$$R^2 = \frac{\sum_{t=1}^{N} \left(\hat{Y}_t - \bar{Y}_t\right)^2}{\sum_{t=1}^{N} \left(Y_t - \bar{Y}_t\right)^2} \tag{2.42}$$

where, N is the total number of sample values, Y_t represents actual expected values, $\bar{Y}_t$ as its mean, and $\hat{Y}_t$ is the predicated values.

2.6.3.4 Root mean square error (RMSE)

The accuracy of a model's prediction can be calculated by root mean square error (RMSE). It computes the differences between actual expected and predicated values and useful for comparing prediction errors of different models. The expression used to calculate RMSE is:

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^{N} \left(Y_t - \hat{Y}_t\right)^2} = \sqrt{MSE} \tag{2.43}$$

where, N is the total number of sample values, Y_t is the actual expected value, $\hat{Y}_t$ is the predicted value. The scale of errors is converted to the same scale as of actual expected values by taking the square root of mean squared error (MSE). The RMSE tends to penalise large errors and scales results in same units.

Chapter 3

Thermal modelling of the buildings

Overview

This chapter outlines building regulations which both domestic and non-domestic buildings must comply for the conservation of fuel and power in England/Wales. These building regulations are the minimum energy efficiency standards for the buildings. Therefore, three buildings have been considered for thermal modelling using two separate software as a case study. The thermal performance models for both domestic and non-domestic buildings are modelled using IDA-ICE and IES-VE software, respectively. Later, a Python model has been developed and implemented in IES-VE software to facilitate the users to demonstrate compliance with building regulations. Finally, the results are discussed with limitations of current building regulations.

3.1 Building regulations in United Kingdom

In 2008, the UK became first country to introduce "The Climate Change Act" to reduce 80% of greenhouse gas emissions (GHG) by 2050 to the level recorded in 1990 (Committee on Climate Change 2010). The carbon dioxide emissions from built environment sector accounts for 46% of the UK's total carbon emissions (Kelly et al. 2012). The UK's future clean growth strategy is to reduce the emissions created by heating our homes and businesses. These account for almost a third of UK building emissions. The government has further plans to upgrade the thermal performance of all homes to Energy Performance Certificate (EPC) Band C by 2030 and develop measures to support businesses and improve their energy productivity by at least 20 percent.

In the UK, the share of residential buildings is 27% with 50% of these emissions consist of space heating and hot water demand (Centre for Sustainable Energy Association for the Conservation of Energy & Moore 2008). The UK government has published certain set of building regulations for both new and existing, domestic and non-domestic buildings which ensures that the government policies are being met according to the standards. The building regulations in England and Wales are categorised into sixteen groups covering several aspects of requirements. Of these, Part L (conservation of fuel and power) and Part F (ventilation) in England and Wales focuses on the energy efficiency and carbon emission requirements.

The buildings regulation (Part L for England and Wales, Section 6 for Scotland and Part F for Northern Ireland) provide guidelines for increasing energy efficiency and performance of buildings. It set limits on energy consumption of building components, building fabric, lights and solar gain.

The Part L regulations in England and Wales is further divided for domestic buildings (L1A and L1B) and non-domestic buildings (L2A and L2B), where L1A,L2A apply to new buildings and L1B, L2B are for existing buildings. The Part L is also used for the energy demand estimation and used as a stepping down for the generation of energy performance certificates (EPC). The valid EPC is a legal requirement which the owner must posses whenever a building is sold, rented or constructed. This mandatory requirement is enforced by the European Union in legislation 'European performance of buildings directive'(EPBD) (European-Commission 2002). The EPC certificate generation methodology for existing and new buildings and its relation with Part-L buildings regulations is shown with the flow-diagram in Figure 3.1.

3.1.1 Regulations for domestic buildings

The current energy estimation methodology for the UK's domestic building is based on set of Building Research Establishment Domestic Energy Model (BREDEM) models developed and maintain by Building Research Establishment (BRE) institute. Moreover, regulations for the energy performance requirements of new domestic buildings are covered in Part L1A and existing buildings in Part L1B (Building Regulations 2010*a*,*d*). These regulations set limits on energy efficiency and carbon dioxide emissions which the building needs to achieve in order to demonstrate compliance. These performances are calculated using standard assessment procedure (SAP) tool for new domestic buildings and reduced standard assessment procedure (RdSAP) tool for old and existing domestic buildings. These assessment procedures are in-agreement with the EU energy performance of building directive (EPBD) and
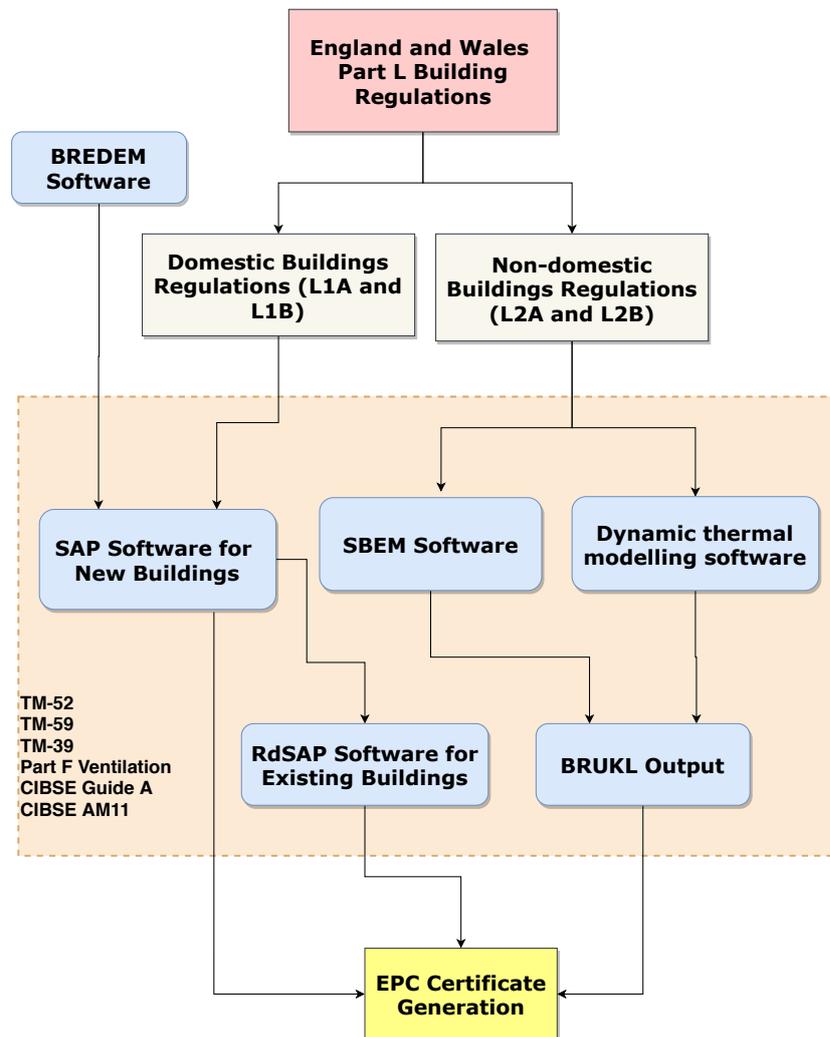
Figure 3.1: The methodology of EPC certificate generation of existing and new buildings and its relation with Part-L buildings regulations in England/Wales

hence used to generate energy performance certificates (EPCs) (Ministry of Housing & Government 2015).

Historically, the SAP was developed in 1993 as an independent calculation methodology by department for the environment (DOE) and Building research establishment (BRE). The initial version of SAP was based on BRE-DEM model 9 which integrated the Part L of building regulations and estab-

lished SAP itself as a national rating scheme for buildings performance evaluation (Anderson et al. 2015). The SAP was originally developed to address following issues:

1. Optimise the energy efficiency of all buildings.

2. Initiate improvement in energy performance of new buildings based on the SAP rating

3. Replace other duplicate energy rating schemes from the private-sector .

The application of SAP on existing buildings was limited as it was introduced for new buildings and requires enormous amount of input data, therefore a reduced version of SAP known as RdSAP was introduced. The RdSAP reduces the input data requirements and assumes typical building characteristics as well as minimises the assessment processing time. In principle, this means that any existing building with incomplete physical characteristics can be assessed and provided with an EPC certificate.

Though, these tools originate from BREDEM, there are several key differences among BREDEM, SAP and RdSAP. The BREDEM is an energy demand estimation tool and requires building's physical characteristics for calculation, but ignores several important factors such as, location specific weather data, temperature set points, space heating, hot water demand and internal gains profiles (occupancy, equipment gain, lighting and efficiency) (Kelly et al. 2012).

On the other hand, SAP provides energy performance using energy demand estimation from BREDEM and calculates economic efficiency of the building. It provides relative economical energy performance of building in the form of £/m^2 as an energy efficiency index (ranging from 1—100) which

is also known as SAP rate (Kelly 2011). Since RdSAP originates from SAP therefore it has the same methodology but assumes default standard parameters and set data for existing building's energy performance. In conclusion, the energy demand is estimated using BREDEM, whereas the energy performance is estimated using SAP and RdSAP.

3.1.2 Regulations for non-domestic buildings

The building regulations Part L2 applies to non-domestic buildings in England/Wales, where Part L2A covers regulations regarding the conservation of fuel and power for new non-domestic buildings and Part L2B is for the existing non-domestic buildings. The Part L2 sets limit on certain characteristics of building and provides energy performance guidelines that can be achieved using any method, but the carbon dioxide emissions from the building should be lower than the target emissions (Building Regulations 2010*b*,*c*).

The minimum carbon dioxide emissions to demonstrate compliance with Part L2 can either be calculated using BRE's Simplified Building Energy Model (SBEM) or the detailed dynamic simulation model (DSM) (SBEM 2010). The results are then fed into the compliance calculation software, BRUKL (Building Regulations UK, Part L), which evaluates various aspects of building according to regulations. Finally, the Energy Performance Certificate (EPC) is generated based on the results.

These are several criteria mentioned in Part L2 regulations, but the Criterion 1 directly addresses the requirements for carbon emissions. The criteria which needs to evaluate are as follows (IES Virtual Environment 2014):

- Building carbon emission rate (BER) should be lower than the target carbon emissions rate (TER).

- Limitations on the minimum building fabric and efficiency of building services.

- Limitations on solar gains.

- The constructed building should be in-line with calculated building CO_2 emission rate.

- Provision of information for the energy efficient functioning of building.

3.2   Thermal comfort assessment

Thermal comfort is defined as a state of satisfaction with thermal conditions in the surrounded environment. In the UK, it is correlated with well-being of occupants and assessed by CIBSE adaptive thermal comfort methodology set out in TM-52 and TM-59.These has been undertaken in-line with methodologies set out in the CIBSE guidance documents, including:

- The Limits of Thermal Comfort: Avoiding Overheating in European Buildings (TM-52)

- Technical Memorandum 59 Design Methodology for the Assessment of Overheating Risk in Homes (TM-59)

Compliance with these guides are widely perceived as a benchmark for assessing if overheating is likely to occur within buildings. TM-52 can be used to assess any type of building whereas TM-59 has been tailored to target overheating risks in homes. There are no mandatory building regulations that state overheating assessment is required, but directly related to design parameters of the building. It is a best practice to achieve thermal comfort with overheating assessment.

3.2.1 CIBSE TM-52

The CIBSE TM-52 thermal comfort analysis is commonly used to assess the risk of overheating in non-domestic buildings. It predicts the indoor operative temperature using methodology for adaptive thermal comfort calculation. In adaptive thermal comfort, the threshold for overheating temperature varies with respect to mean outdoor air temperature, as suggested in BS EN:15251 standard (BSI 2008). The TM-52 sets out three criteria, where 2 of the 3 criteria should be satisfied in order to pass the overheating assessment.

1. Criterion 01: The total number of occupied hours during which the difference between the operative temperature (ΔT) and set-point is greater or equal to one degree (^oC) shall not be more than 3% during summer season (May — September).

2. Criterion 02: The weighted average of exceeding over-heated indoor operative temperature hours shall not be more than 6 in a day.

3. Criterion 03: The maximum daily indoor operative temperature shall not be more than 4^oC

The CIBSE recommends using category II for TM-52 thermal comfort analysis for newly built buildings. The maximum acceptable temperature is calculated from following three types of building category given in Table 3.1. The adaptive thermal comfort at thermal bands for different building category are shown in Figure 3.2.

3.2.2 CIBSE TM-59

CIBSE TM-59 is a standardised approach to predict risk of overheating in domestic buildings designs using dynamic thermal analysis. The methodology
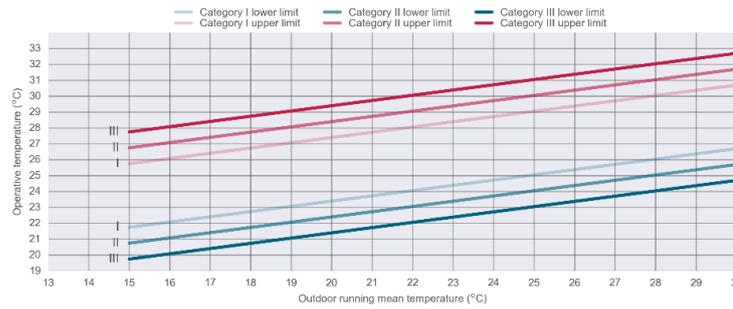
Figure 3.2: Thermal comfort bands from BS EN:15251 (BSI 2008). Research has shown that comfortable room temperatures vary with external air temperature. Occupants are generally more accepting of warmer room temperatures if the weather is also warm outside.

Table 3.1: Building categories from the standard BS:EN-15251

Category	Description	Acceptable range
Category I	High expectation level - spaces occupied by elderly and sensitive person	2
Category II	Normal expectation level - new and renovation buildings	3
Category III	Moderate expectation level – existing buildings	4

is usually applied on residential apartments but can also be extended to other residential spaces. The TM-59 methodology restricts the user by providing pre-set occupancy, internal gains, equipment and lighting profiles for the analysis. These profiles are more stringent compared to TM-52 methodology. The compliance with TM-59 is only achieved once following criteria are met:

- For living rooms, kitchens and bedrooms – The total number of occupied hours during which the difference between the operative temperature (ΔT) and set-point is greater or equal to one degree (^{o}C) shall not be more than 3% during summer season (May — September). It is same as in CIBSE TM-52 Criterion 1.

- For bedrooms only – the operative temperature during 22:00–7:00 hrs should not exceed beyond 26°C for more than 1% of annual hours, where 32 hours make 1% of annual hours.

3.3 Case study - domestic building thermal analysis

A domestic building from the REMOURBAN project is modelled in IDA-ICE 4.6.2 software and real monitored weather data is used for heat demand estimation. The monitored weather data is first discussed in Section 3.3.1 and results from thermal modelling with retrofitting are elaborated below in Section 3.3.2.

3.3.1 Monitored weather data

Three years (2014–2016) of high resolution meteorological data with temporal resolution of 15 min is obtained from the nearby weather station located at $53°3'41.62''$ N, $0°57'49.75''$ W and converted into one year of hourly data. This filters out extreme events and provides robust time-series of data. The meteorological weather data contains outdoor dry air temperature, pressure, relative humidity, incoming solar radiations, wind speed and direction. This weather data is converted into EnergyPlus format (*.epw*) climate file and then imported into IDA-ICE software for simulations.

The results show that the incoming solar radiations are reasonably high in Nottingham during months of April to September and reach up to 900 W/m^2 per hour. This shows the significance of incoming solar radiations for heat demand estimation. The outdoor air temperature varies between 2–25°C throughout the year and is never below 0°C. On the other hand, the behaviour of wind speed data is quite opposite. During summers, the hourly wind speed stays constant but fluctuates during winters with wind storms between months of December to February. These results are shown below in Figure 3.3.
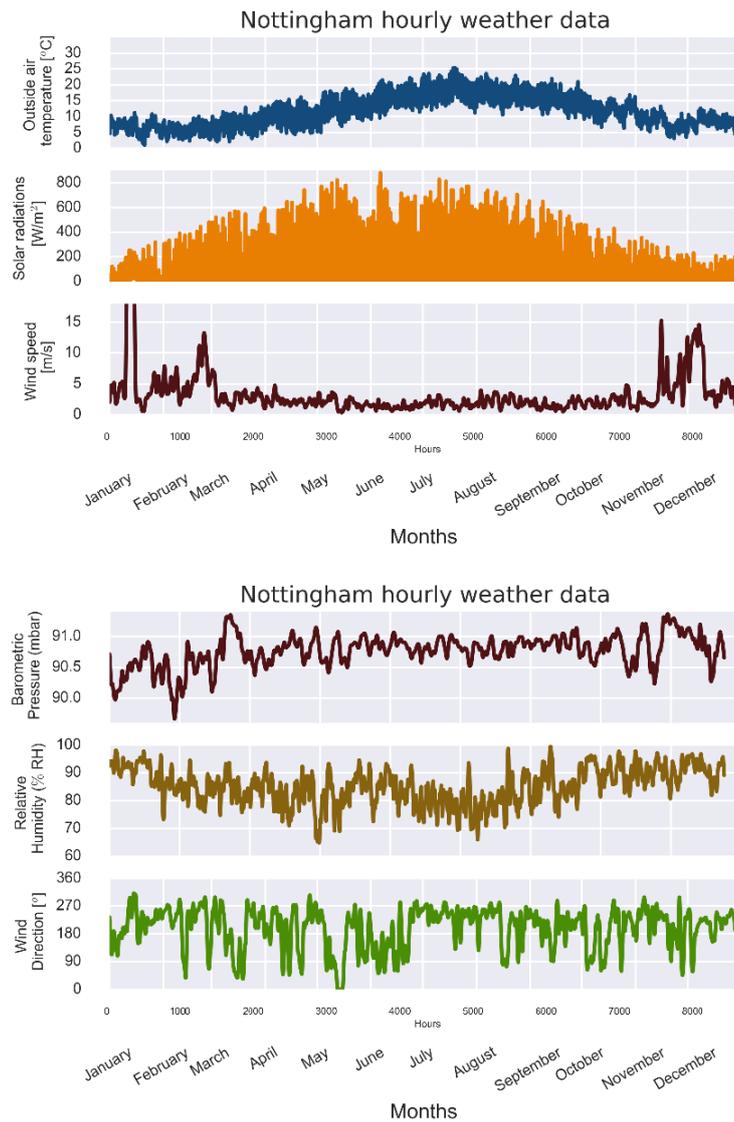
Figure 3.3: Monitored weather data from Nottingham weather station located at 53°3′41.62″ N, 0°57′49.75″ W. This data is used for thermal performance analysis and adapted from (Ashfaq & Ianakiev 2018*b*)

The data for soil temperature in Figure 3.4 is useful for heat loss estimation due to temperature difference with ground temperature. It is observed that the soil temperature fluctuates between 8–16°C and 4–8°C during summers and winters, respectively. These monitored weather time-series data are used for thermal modelling in IDA-ICE model and shown in Figures 3.3 and 3.4.
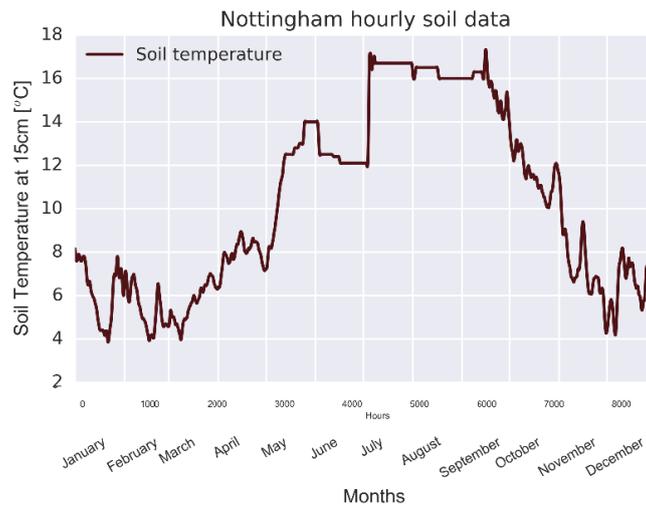
Figure 3.4: Monitored soil temperature (15 cm depth) data from Nottingham weather station located at 53°3′41.62″ N, 0°57′49.75″ W. This data is used for thermal calculations and adapted from (Ashfaq & Ianakiev 2018*b*)

3.3.2   Heat demand after retrofitting of buildings

In the UK almost 80% of existing building stock will still be operational by the year 2050 and therefore, increasing energy performance of existing buildings is vital for achieving future carbon emission targets (Ianakiev et al. 2017). Reduction of energy demand by retrofitting of existing building is fundamental for the utilisation of low carbon fuels as well as to address the issue of climate change. To this end, 94 flats (from REMOURBAN Project) have been retrofitted for the implementation of low temperature district heating. These flats were originally part of social housing, constructed from 1960's and are in four blocks. To continue with analysis, only one of the building is used for pre and post-retrofit heat demand calculations. Since, being part of social housing the construction plans were unavailable, therefore information from surveys has been used for geometry (walls, roof, windows, and floor), orientation, fabric (U-value) and glazing parameters.

The IDA-ICE software is dynamic multi-zone heat demand calculation

software and accredited to the standard DS/EN 15265 (EQUA Simulation AB 2010). It is commonly used by researchers and energy consultants in Scandinavian countries for energy performance of the buildings (EQUA Simulation AB 2013). The building is first modelled using its characteristics such as geometry, orientation, construction materials, fabrics, exposed parameters and glazing values. The natural ventilation is taken as 0.94 ac/hr which is in agreement of CIBSE Building Code (CIBSE Jul 2015). Then, the domestic internal gains for the occupancy and equipment are taken as 0.81 and 1.55 W/m^2, respectively (Tunzi et al. 2016), (Domestic Building Regulations 2013) and gain profiles are shown in Figure 3.5.
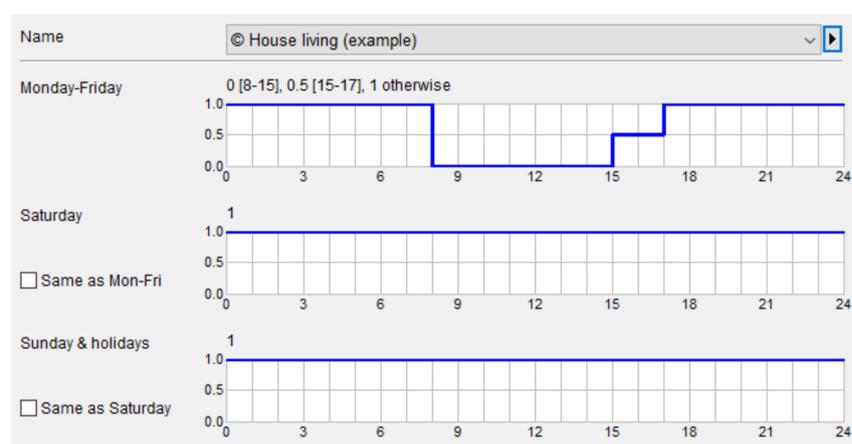
Figure 3.5: Internal gains and occupancy profiles assumed for the domestic building in case study

The design outdoor temperature is taken as -8°C for extreme conditions (Frederiksen & Werner 2013) and given in standard (Danish standard 469 2013). The weather data is an important element of heat demand estimation, therefore real monitored weather data of Nottingham is used for the dynamic thermal simulation.

The pre-retrofit buildings were built with brick cavity walls, concrete floors, roof with tiles, single glazed windows and raft-foundations. These flats

are retrofitted to attain thermal performance according to the 'UK Building Regulation - Part L1B' standard (Building Regulations 2010*a*). The post-retrofit building has shown an improvement in U-value for brick cavity walls, windows and roof to 0.3, 1.8 and 0.123 W/m²K, respectively. This deep-retrofit building provides similar energy performance to the usual practice of retrofitting. The comparison between the U-value for pre and post-retrofit buildings is provided in Table 3.2.

Table 3.2: Main fabric and glazing properties of flat before and after retrofitting building envelope. (Ashfaq & Ianakiev 2018*b*).

Component	Before retrofit U- value (W/m²K)	After retrofit U- value (W/m²K)
Wall	2.1	0.3
Glazing	2.727	1.8
Roof	0.346	0.123
Floor	2.128	2.128
Overhang	2.128	2.128
Heat demand (KWh)	23,897	11,253
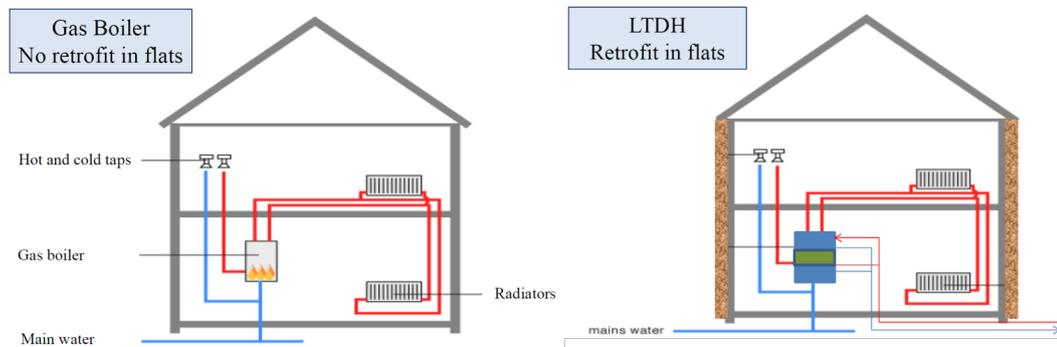
Figure 3.6: Retrofitting of flats from REMOURBAN project.

The dynamic thermal simulation results show there is significant increase in operative air temperature to 19°C for all rooms in building, except the entrance hallway which can be explained due to entrance door opening schedule and air infiltration. The overall air infiltration in post-retrofit building has decreased to the passive house standard, i.e. 0.6 ac/hr under 50 Pascal (Cui et al. 2017).

It is observed that retrofitting has reduced the energy demand by almost 52% for each flat, i.e. 23,897 to 11,253 KWh. The noticeable reduction in heat demand is at the ground floor where operative air temperature is increased by 2^oC (17°–19^oC). This explains that the improvement in thermal efficiency of wall and roof insulation, and conversion of windows to double glaze significantly increases the energy performance. These room specific pre and post-retrofit operative air temperature simulation results are further shown in Figure 3.7.
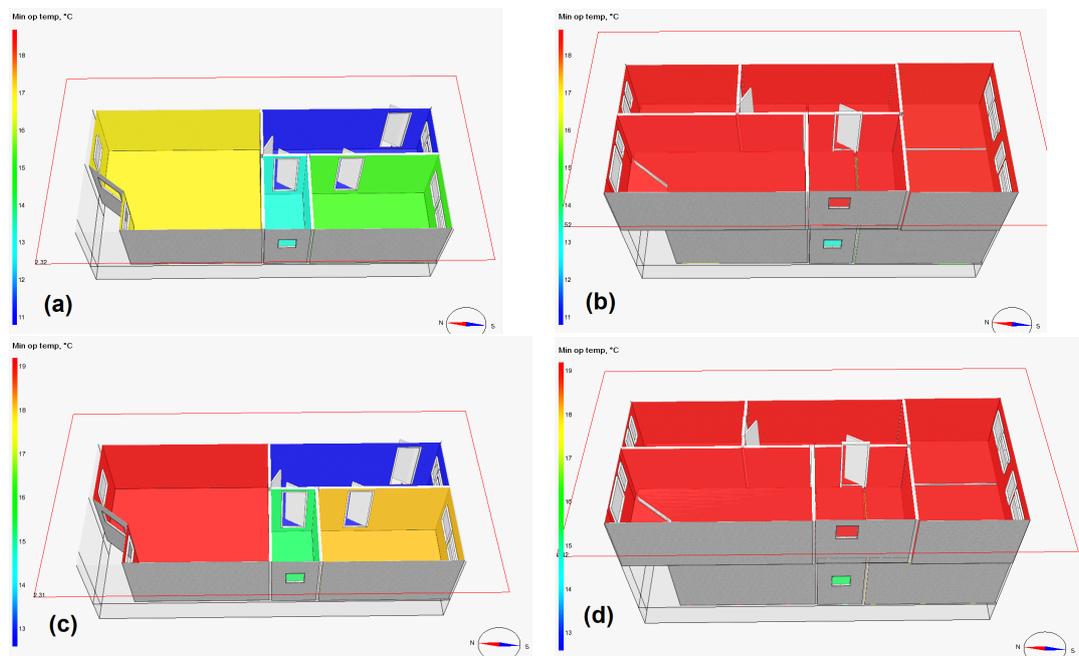
Figure 3.7: Thermal performance simulation results from the IDA-ICE software before and after the retrofitting. Figures (a,b) show the minimum operative temperature before the retrofit and Figures (c,d) show minimum operative temperature after the retrofit. These results are adapted from (Ashfaq & Ianakiev 2018*b*)
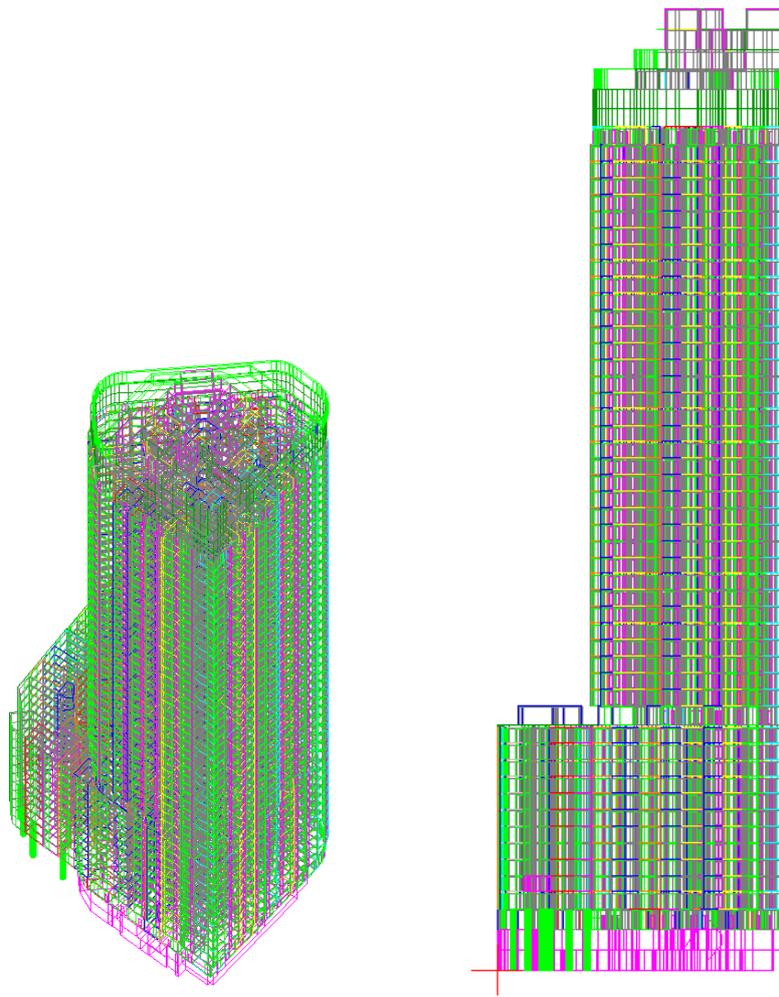
Figure 3.8: View of multi-zone thermal model in IES-VE software.

3.3.3 TM-52 thermal comfort assessments - domestic buildings

The TM-52 thermal comfort assessments analysis has been performed on the proposed 55 floor residential development site located in north-west region of England. The 55 floor residential tower will have basement of 2 floors, amenity areas and flexible-use commercial units in the ground and 1st floor, and residential apartments from 2nd to 55 floor. The dynamic thermal model simulation has been carried out using IES Virtual Environment (IES-VE) software, version 2019.1.0 and shown in Figure 3.8. Of theses 55 floors, only floor 02, 03,

14, 15 and 49 have been analysed as representative examples for over-heating assessment. This is due to computational limitations of IES-VE software and it is completely reasonable to select these floor, owing to the similarity among their orientations and structural layout.

As discussed above, the over-heating assessment depends on the Part L SAP calculations and building construction materials (fabric, glazing, etc). Therefore, the thermal model is created considering parameters from the initial Part L SAP results given below.

Weather data

It is a mandatory requirement for TM-59 thermal comfort methodology to use CIBSE future Design Summer Year (DSY) weather file for over-heating assessment. Therefore, the Manchester DSY1 2020 high emissions scenario 50th percentile weather data has been used in the assessment. Although using DSY weather data is not required by TM-52 methodology as it is more onerous than the minimum requirements, it has been employed to show good design practice and assist in mitigating potential future climate change. The Manchester DSY1 2020 high emissions scenario 50th percentile air temperature profile is shown in Figure 3.9.

Thermal modelling parameters

The thermal model considers several constant input parameters (building fabric standards and glazing properties) for over-heating assessment (TM-52 and TM-59). These parameters are summarised in Tables 3.3 and 3.4.

Moreover, the occupancy and internal gain profiles (people, equipment and lighting) for domestic building are generated and passed into the model. The internal gains for TM-52 overheating assessments are given in Table 3.5.
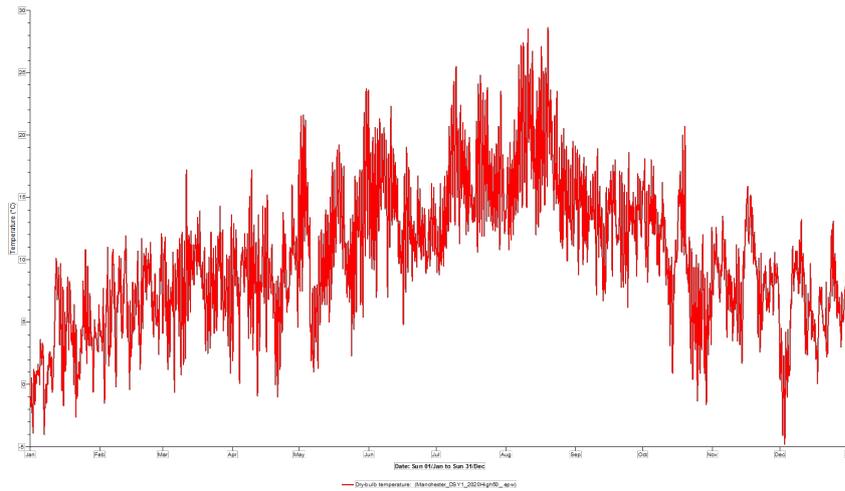
Figure 3.9: The figure shows the CIBSE future Design Summer Year (DSY) 2020 - 50th percentile high emissions scenario air temperature profile used for overheating assessment. The DSY weather data is a single year continuous data compared to Test Reference Year (TRY) data (which is a monthly averaged data).

Table 3.3: Fabric parameters

Building fabric	U-value (W/m^2 K)	Notes
Curtain wall	1.10	Curtain wall including transparent,opaque and purge panels (area weighted average)
Roof	0.12	Uninsulated – Heat loss to below zones
Ground floor	0.14	Flat roof
Air permeability		The infiltration rate has been calculated based on design air permeability of 2.75 m^3/hr/m^2 at 50 Pa

Table 3.4: Glazing properties

Criterion	Light transmission	G value	Internal shading applied
External glazing	0.71	Varies in every iteration	Yes

Table 3.5: Internal gains used for TM-52 over-heating assessment

Parameter	Internal gain profile
Occupancy	70W sensible and 45W latent per person Single bedroom – 1 person from 22:00 – 07:00 Master bedroom – 2 people from 22:00 – 07:00 2 Bed living – 2 people from 06:00 – 08:00 and 19:00 – 22:00 3 Bed living – 3 people from 06:00 – 08:00 and 19:00 – 22:00 4 Bed living – 4 people from 06:00 – 08:00 and 19:00 – 22:00
Equipment	115W from 06:00 – 08:00 and 19:00 – 23:00 450W cooking load from 07:00 – 08:00 and 19:00 – 21:00
Lighting	2W/m^2 from 06:00 – 07:00 and 20:00 – 22:00

The overall apartment ventilation strategy is proposed to be a combination of mechanical and natural ventilation. Mechanical Ventilation Heat Re-

covery (MVHR) units are included in the apartments and natural ventilation is provided through restricted opening windows for "purge" ventilation. The building has fixed external windows and natural ventilation is provided with louvers on all floors for "purge" ventilation. The floor 2 has an exception where one tower has patio doors in living room. The louvers and equivalent openable free area are modelled from the dimensions given in architecture drawings. Table 3.6 outlines the parameter used for modelling ventilation panels.

Table 3.6: Ventilation panel parameters

Parameter	Louvre	Patio door
Free area	52%	90%
Coefficient discharge	0.40	-
Max opening angle	-	90°
Opening type	louvre	window/door – side hung

The shading has considerable effect on over-heating assessment and quantified by shading coefficient (SC). It is a measure of thermal performance of glass (panel or window) in a building and described as the proportion of sunlight passing across the glass surface to the direct incoming solar radiations. It depends on the transparency of glass (colour) and degree of reflection. Table 3.7 outlines the shading coefficient for various shading devices used in this analysis.

Table 3.7: Shading coefficient of various blinds and curtains

Shading device	Shading coefficient
Net curtain (fine)	0.76
White cotton curtain	0.54
Linen blinds (cream holland)	0.40
Blackout blinds	0.05-0.10

The acoustic restrictions have been included in internal conditions applied to the thermal model assessments. These includes number of hours during

which ventilation panels will be required to remain open in living spaces. The following criteria are considered for acoustic requirements.

- The opening of windows should be avoided between 11 pm and 7 am.

- The bedroom windows can remain open during the day to provide temperature control.

- The opening of windows in living areas during the day should be limited to short periods of time.

3.3.3.1   Iteration 1

The first iteration for over-heating assessment uses parameters outlined above and assumes no shading devices in thermal model. Moreover, the glazing's g-value used in the thermal simulation is given in Table 3.8.

Table 3.8: Iteration 1 glazing and shading parameters

Parameter	Value
Glazing g value	0.40
Blinds	None

The results in Figure 3.10 show that 15% of rooms in the selected floors fail Criteria 1 & 2 for over-heating assessment. This due to high effect of incident solar radiations on windows and curtain wall (external wall) which increases the indoor temperature of rooms. The curtain wall are external walls made of large metal-framed sheets of glass and commonly used by architects to create modern outlook. The next iteration assumes the installation of internal shading devices such as, curtains and linen roller blinds.

3.3.3.2   Iteration 2

This iteration uses same parameters as in Iteration 1 and white curtains are modelled as shading devices. Their parameters are outlined in Table 3.9.
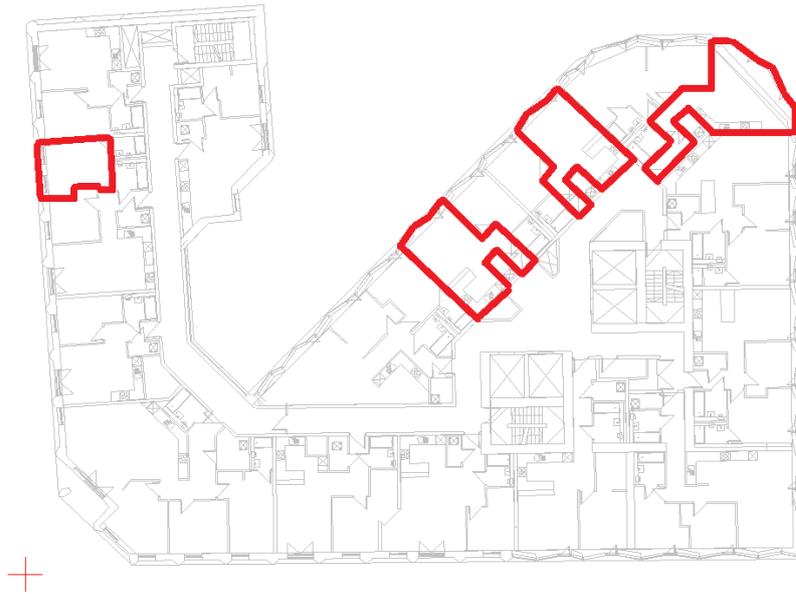
Figure 3.10: Figure show results from TM-52 Iteration 1. The zones shown in red illustrate rooms found to overheat on all levels throughout the selected floors.

Table 3.9: Iteration 2 glazing and shading parameters

Parameter	Value
Glazing g value	0.40
Blinds	Lightweight white curtains in all living spaces and bedrooms Shading coefficient = 0.54
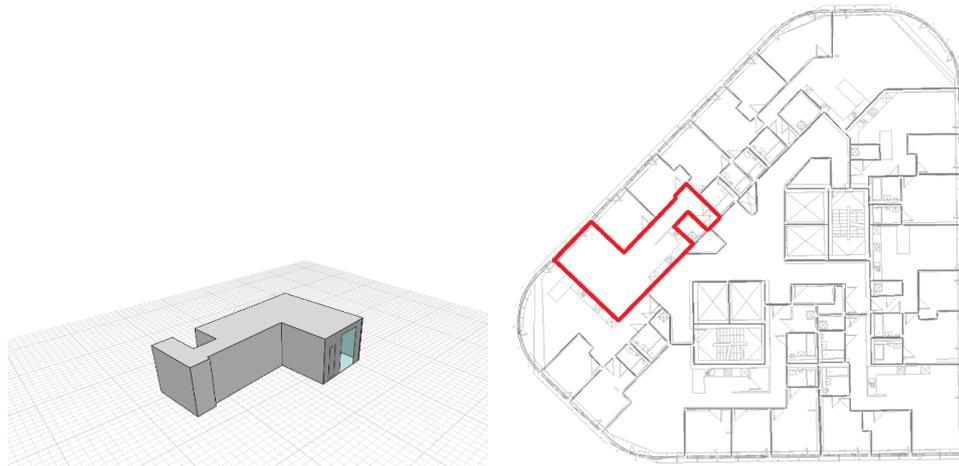
Figure 3.11: Figure show results from TM-52 Iteration 2. The zone shown in red illustrate rooms found to overheat on Floor 14, 49.

The results in Figure 3.11 show that the installation of shading devices improves the results and only 2 rooms (1.5%) of rooms in selected floors are unable to demonstrate thermal comfort with over-heating assessment. The

worst performing space during summer is '3 Bed living room' on floor 49 which fails to achieve compliance with both Criteria 1 & 2. This is due to high internal gains from the equipment and incoming incident solar radiations. The in-depth analysis in Figures 3.12 and 3.13 shows that the operative temperature in 3 Bed living room on Floor 49 rises during unoccupied hours. This is due to equipment gains which results in mildly uncomfortable conditions between 18:00 -21:00 hours during month of August.
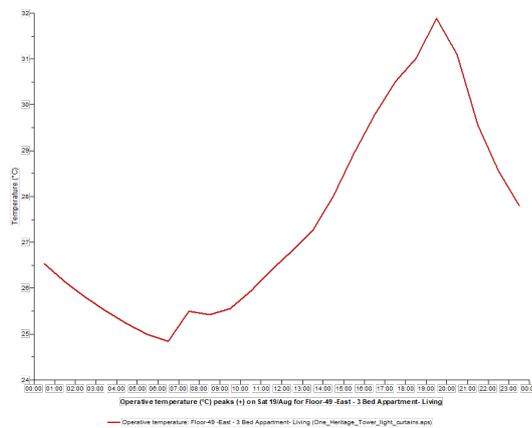
Figure 3.12: Figure shows the peak operative comfort temperature (°C) result in the failing zone (3 Bed living room on Floor 49) from TM-52 Iteration 2.
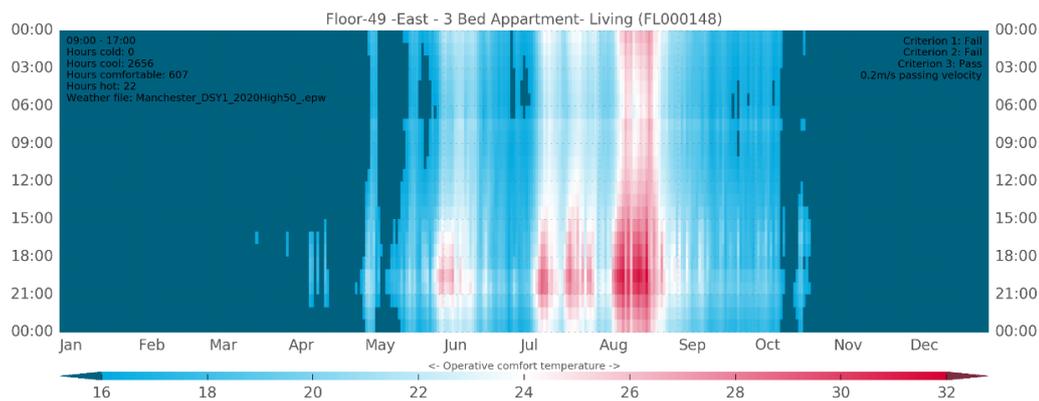
Figure 3.13: Figure shows the operative comfort temperature (°C) result in the failing zone (3 Bed living room on Floor 49) throughout the year from TM-52 Iteration 2.

The thermal comfort can be achieved by providing mechanical cooling to these two failed rooms, however the next iteration considers another type of

shading devices and provides an option for mitigating the need for mechanical cooling for all apartments.

3.3.3.3 Iteration 3

This iteration assumes same parameters as in Iteration 2 and linen blinds are used as a shading devices with parameter outlined in Table 3.10.

Table 3.10: Iteration 3 glazing and shading parameters

Parameter	Value
Glazing g value	0.40
Blinds	linen blinds (cream holland) in all living spaces and bedrooms Shading coefficient = 0.40

The results show the thermal comfort is achieved for TM-52 overheating assessment and it is recommended to use linen blinds (cream holland) or shading devices with similar shading coefficient to reduce the effect of incoming solar radiations and mitigating risk of over-heating.

3.3.4 TM-59 thermal comfort assessment - domestic buildings

As discussed earlier, the TM-59 over-heating assessment methodology restricts the user to implement pre-set internal gain profiles (occupancy, equipment and lighting) and CIBSE future Design Summer Year (DSY) weather data files. The air temperature profile used in this assessment is provided in Figure 3.9 and the mandatory internal gain profiles from TM-59 are given in Figure 3.14.

Moreover, in TM-59, blinds and shading devices can be used only, if specifically included in the design and are not obstructive to the natural ventilation.

Number of people	Description	Peak Load Sensible	Latent	00–01	01–02	02–03	03–04	04–05	05–06	06–07	07–08	08–09	09–10	10–11	11-12	12-13	13-14	14-15	15-16	16-17	17-18	18-19	19-20	20-21	21-22	22-23	23-24
				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	Single bedroom occupancy	75	55	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.7
2	Double bedroom occupancy	150	110	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	1	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.7
2	Studio occupancy	150	110	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1-bed: living/kitchen occupancy	75	55	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
1	1-bed: living occupancy	75	55	0	0	0	0	0	0	0	0	0	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0	0
1	1-bed: kitchen occupancy	75	55	0	0	0	0	0	0	0	0	0	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0	0
2	2-bed: living/kitchen occupancy	150	110	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
2	2-bed: living occupancy	150	110	0	0	0	0	0	0	0	0	0	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0	0
2	2-bed: kitchen occupancy	150	110	0	0	0	0	0	0	0	0	0	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0	0
3	3-bed: living/kitchen occupancy	225	165	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
3	3-bed: living occupancy	225	165	0	0	0	0	0	0	0	0	0	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0	0
3	3-bed: kitchen occupancy	225	165	0	0	0	0	0	0	0	0	0	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0	0
	Single bedroom equipment	80		0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.13
	Double bedroom equipment	80		0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.13
	Studio equipment	450		0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.24	0.24	0.24	0.24	0.24	0.24	0.24	0.24	0.24	1	1	0.44	0.44	0.24	0.24
	Living/kitchen equipment	450		0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.24	0.24	0.24	0.24	0.24	0.24	0.24	0.24	0.24	1	1	0.44	0.44	0.24	0.24
	Living equipment	150		0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	1	1	1	1	0.4	0.4
	Kitchen equipment	300		0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	1	1	0.17	0.17	0.17	0.17
	Lighting profile	2	(W/m²)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0

Figure 3.14: Heat gains and thermal environmental conditions for human occupancy taken from TM-59 methodology and based on CIBSE Guide A. A full break down of these gains with descriptions can be found in the TM-59 over-heating assessment guide.

The curtain wall and external glazing are part of fixed façade with no opening capabilities. Hence, blinds and other shading devices are not obtrusive to natural ventilation required to achieve compliance with TM-59 overheating assessment. The natural ventilation is provided with louvers in each room which are controlled separately and only open, once the internal dry bulb temperature is above 22°C and the room is occupied.

3.3.4.1   Iteration 4

This iteration uses same parameters as in Iteration 3 but with additional mandatory parameters of TM-59 methodology given in Table 3.11. The glazing g-value and other parameters used in this iteration is given in Table 3.8.

Table 3.11: Iteration 4 glazing and shading parameters

Parameter	Value
Glazing g value	0.40
Blinds	Linen blinds in all living spaces and bedrooms Shading coefficient = 0.40

The results in Figure 3.15 show that 93% of rooms in selected floors are unable to demonstrate thermal comfort with over-heating assessment and failure of all Criteria is evident in most rooms on each floor. These unexpected results

Figure 3.15: Figure show results from TM-59 Iteration 4. The zones shown in red illustrate rooms which are found to overheat in selected floors.

compared to TM-52 iteration 3 can be explained by the mandatory internal gain profiles (occupancy, equipment and lighting) in TM-59 methodology.

This issue can be resolved by either using shading devices with lower shading coefficient or installing mechanical cooling during summers. Alternatively, the glazing with lower g-value can be used which will reflect back the incoming incident solar radiations. However, this might not be the most optimum solution as the building may require more heating during winters and might not comply with Part L SAP regulations. Therefore, a sweet-spot needs to be determined for g-value. The next iteration assumes using blinds with lower shading coefficient (black out blinds) and reduced glazing g-value.

3.3.4.2   Iteration 5

This iteration uses all parameters as in Iteration 4 and glazing g-value is reduced along with shading devices with lower shading coefficient. These parameters are outlined in Table 3.12.

Table 3.12: Iteration 5 glazing and shading parameters

Parameter	Value
Glazing g value	0.35
Blinds	Blackout blinds in all living spaces and bedrooms Shading coefficient = 0.05
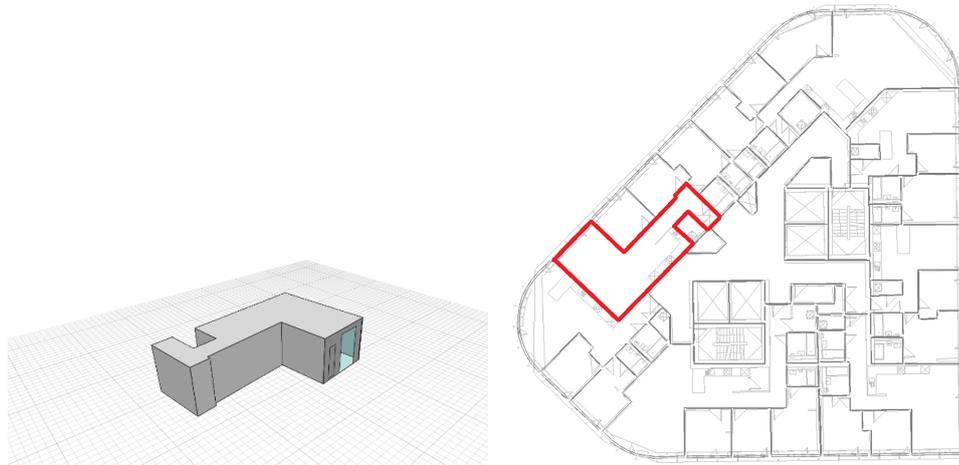
Figure 3.16: Figure show results from TM-59 Iteration 5. The zones shown in red illustrate rooms which are found to overheat on floors 49.

The results in Figure 3.16 show that only 0.77% of rooms in the selected floors are unable to achieve thermal comfort with over-heating assessment. The only failed room is 3 bed apartment's living room on Floor 49, which is the same room failed over-heating assessment in several iteration. As mentioned earlier, this room experiences high internal gains from the equipment and fails to achieve compliance with both Criteria 1 & 2. It is recommended that this room should be installed with mechanical cooling during summer season, rather than applying alternate solutions on the entire building (such as improving g-values). The Table 3.13 enlists number of hours during which ventilation panels are expected to remain open for natural ventilation.

Table 3.13: Acoustic analysis for the building

Louvre	Average hours / day during summer	Percentage time - windows open during summer
Occupied spaces	4.33	18%
Bedrooms Only	7.20	30%

3.4   Case study - non domestic building thermal analysis

The second case study considers a stadium as a non-domestic building for thermal performance and compliance analysis in IES-VE software. This stadium is currently under construction in North-West of the UK. The IES-VE (Integrated Environmental Solutions Virtual Environment) software contains several applications that are all linked together with a single interface for the simulation. The software contains module for building geometry, solar radiations analysis, thermal modelling analysis, daylight analysis, heating, ventilation and air conditioning (HVAC) systems analysis.

The entire building is modelled in IES-VE software and dynamic thermal simulation is performed for the Part-L building regulations compliance analysis. The software provides flexibility to run compliance analysis using SBEM or dynamic thermal simulations and uses BRUKL software for carbon emissions calculations. The thermal modelling calculations for building regulation compliance are like a black box, but an innovative model has been developed in Python to calculate the amount of PV generation (power and area) required to demonstrate compliance for any project. This model along with results are discussed below.
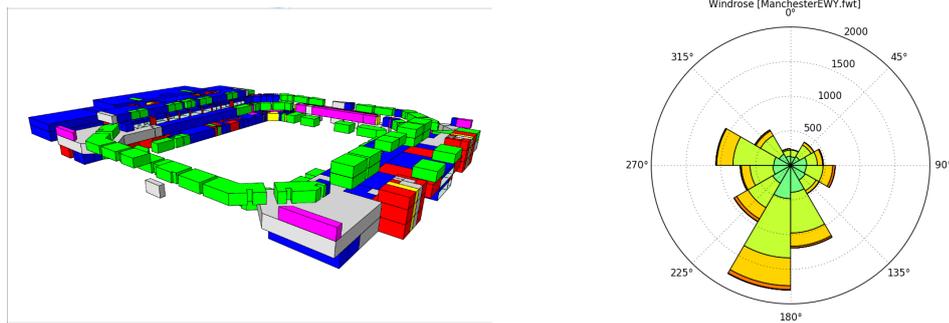
Figure 3.17: The left figure presents the aerial view of the non-domestic building (stadium) considered as a case study, and the right figure presents wind rose graph from the Manchester climate file recommended by CIBSE.

3.4.1   Part L compliance assessment

The IES-VE software creates three type of models for each building i.e. real, actual and notional model. The real building model is created with geometry, construction materials, operational parameters, set-points, occupancy profiles and all other parameters given by the user for thermal simulations. The actual building model is based on real building model, but considers UK NCM methodology parameters (operational parameters, set-points, internal gains and ventilation rate) for thermal simulations. The user has the flexibility to alter only few parameters in actual building model i.e. lighting, construction materials, and auxiliary systems fan power. The software automatically generates the notional building model from the actual building model and it doesn't allow user to modify its properties. The notional building has the standard default values for all parameters (construction materials, fabric, glazing area, operational parameters, set-points, internal gains and ventilation rate). The notional buildings model is used as a benchmark for carbon emissions against actual building model and calculations from the BRUKL calculator. The building passes the Part-L compliance regulations only if the carbon emissions from

the actual building are lower or equal than the carbon emissions of the notional building (target carbon emissions).

First the building is modelled using its construction properties and building service equipment parameters such as, geometry, orientation, construction materials, insulation, HVAC and domestic hot water systems (Non-Domestic Building Regulations 2013). Then, the internal gain profiles (lighting, people and equipment), ventilation rates, heating and cooling set-points are set as per building/room usage, and parameters given in NCM guidelines, CIBSE Part-F and CIBSE Part A (National Calculation Methodology 2013). Finally, the dynamic thermal simulation is performed using recommended weather files provided by CIBSE, which leads to carbon emission calculation. The building fabric thermal resistance (U-values) used in the thermal model are given in Table.3.14.

Table 3.14: Building parameters for thermal performance simulation in IES-VE software

Component	Units	Value
Building Fabric U-values		
External wall	W/m^{2}K	0.15
Ground floor and exposed floor	W/m^{2}K	0.22
Roof	W/m^{2}K	0.18
General glazing	W/m^{2}K	1.5
Glazing Properties		
Light transmission		0.6
G-value		0.3
Air Permeability and Thermal Bridge Properties		
Air permeability		10
Thermal bridge		10% of U-value

A model has been developed in Python programming language, which first performs thermal simulation, runs Part-L compliance and then analyses the carbon emission from both actual and notional building. If the building emissions rating (BER) is greater than target emissions rating (TER), then it
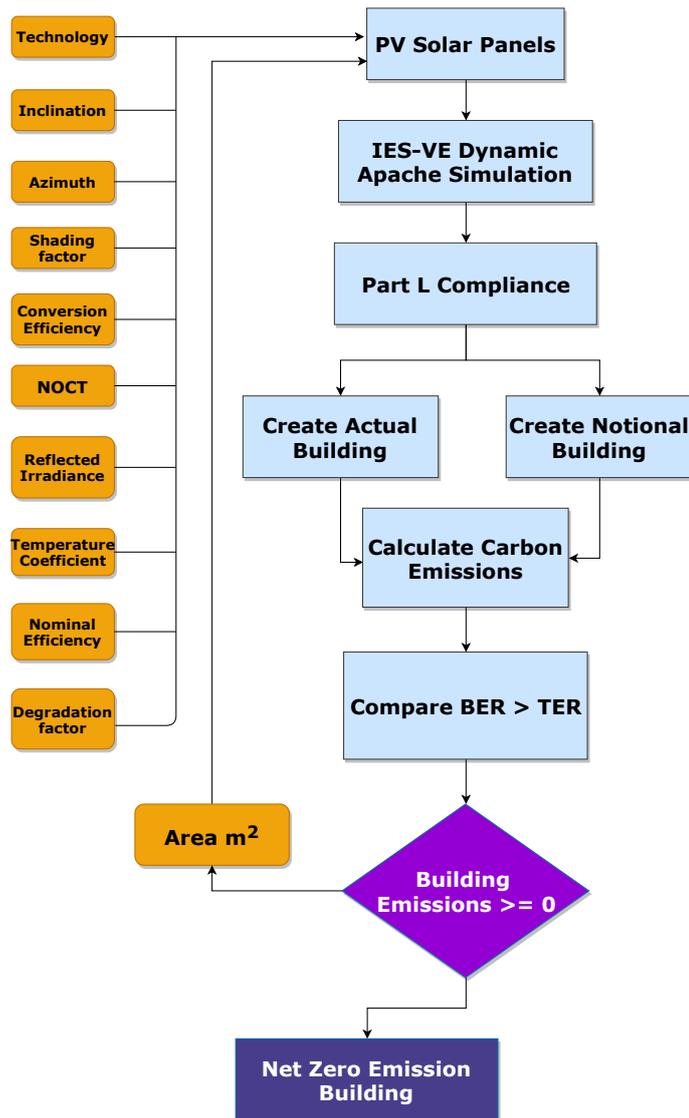
Figure 3.18: Figure represents the methodology of developed Python based PV generation model for calculating required amount of PV power generation in IES-VE software.

calculates the PV generation required for getting under compliance regulations and net-zero carbon emissions. The Part-L allows the use of renewable energy sources to reduce the building carbon emissions. The calculation methodology of the Python based model is shown in Figure 3.18.

The developed model provides the possibility to investigate the energy, power and carbon emission time-series as well as the breakdown of carbon

Table 3.15: The solar panel parameters used in simulation

Properties	Value	PV types	Value
Area	1–15000	Degradation factor	0.990
Azimuth	180.0	Electrical conversion efficiency	0.949
Shading factor	1.0	Module nominal efficiency	0.174
Inclination	15.0	Normal Operating Cell Temperature	45.0
Type of model	Parametric panel	Reflected irradiance	1000
Technology	Monocrystalline Silicon	Temperature coefficient	0.0040

emissions by end use for both actual and notional buildings. It also calculates the minimum PV area required to demonstrate compliance and ultimately, net zero emissions. In addition, this provides flexibility to change or adjust PV solar cell parameters given in Table 3.15.
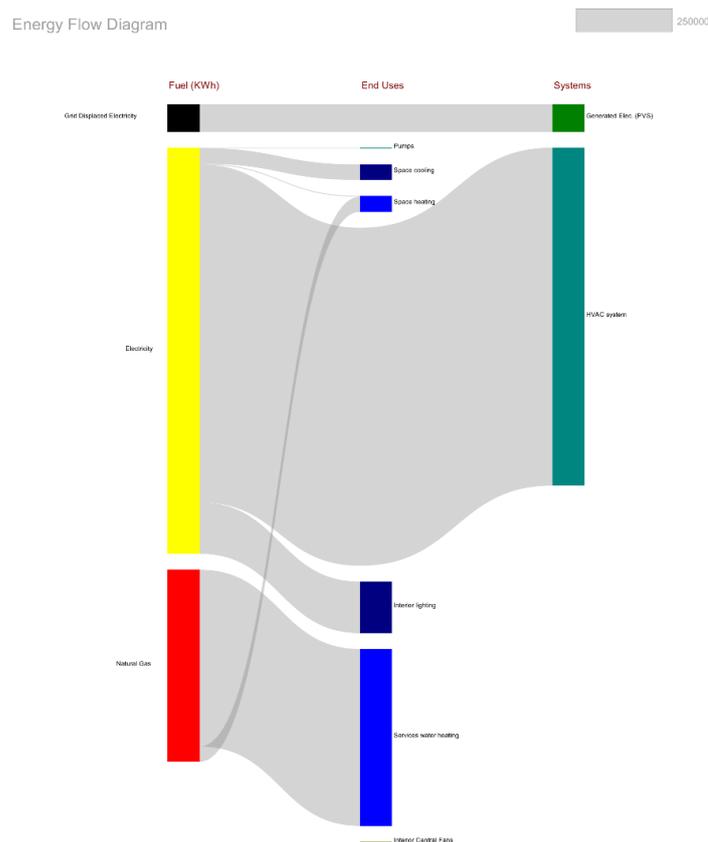
Figure 3.19: Sankey diagram illustrates energy from source to consumption in thermal performance of simulated non-domestic building.

The above generated model is applied to the non-domestic building and the data shows that the current building emissions (BER) are higher than the
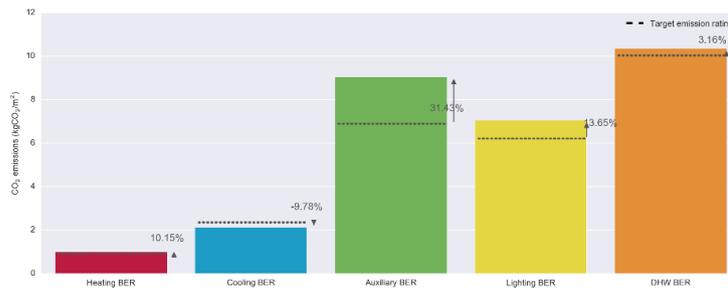
Figure 3.20: The barplot shows comparison of carbon emission between the actual building carbon emissions (BER) and notional building carbon emissions (TER).

target building emissions (TER), and the PV solar generation of 260 MWh is required in order to demonstrate Building Regulation Part L compliance. This reduces the building emissions (BER) and the building passes compliance requirements with BER being 2.57% lower than the TER. Moreover, the share of domestic hot water and auxiliary energy in the actual building (BER) is higher than the notional building (TER) and the share of carbon emissions from the HVAC system's auxiliary energy is highest among other emissions. This is due to the hot water circulation heat-losses and high energy consumption of the HVAC's fan system. It is very common issue in non-domestic buildings and mostly compensated by PV solar generation or other renewable energy sources. These results are further shown in Figure 3.20.
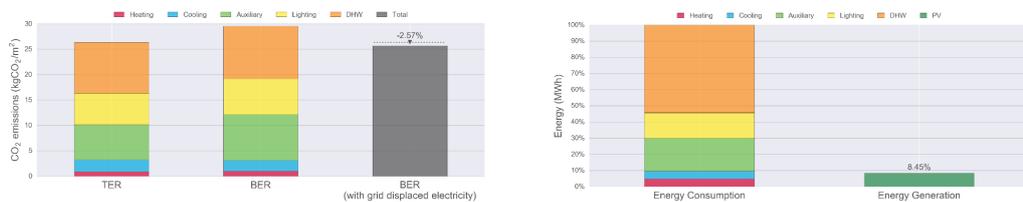
Figure 3.21: The figures shows breakdown of carbon emissions by end use among the actual building carbon emissions (BER) and the notional building carbon emissions (TER), and the reduction in carbon emission required with renewable energy generation to demonstrate building regulation compliance

The Figure 3.21 demonstrates the breakdown of carbon emissions by end use and shows the renewable energy generation and reduction of carbon emis-

sions required to achieve compliance. The results further explain that the 260 MWh of PV solar generation can be achieved with the solar panel array of 2000 m^2 to demonstrate Part L compliance and this will remove around 110,000 kgCO$_2$/m^2 of carbon emission from the atmosphere. These results are further shown below in Figure 3.22 and the breakdown of seasonal monthly solar (PV) electricity generation is shown in Figure 3.23.
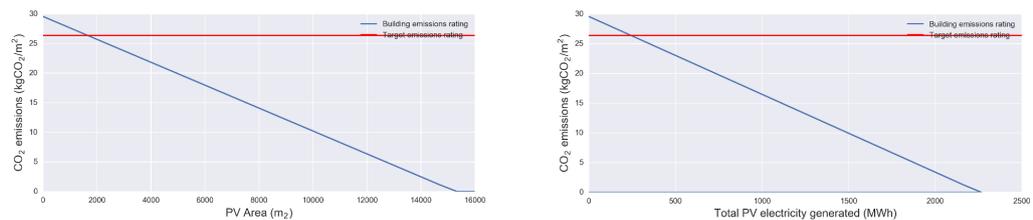
Figure 3.22: Results from the Python based developed model in IES-VE software. The left figure represents the relation between carbon emissions and PV area. The right figure represents the corresponding amount of PV generation. These figures show that PV solar generation of 260 MWh is required in order to demonstrate Part L compliance, and generation of 2326 MWh is required in order to achieve net-zero emissions for the non-domestic building being used as a case-study.
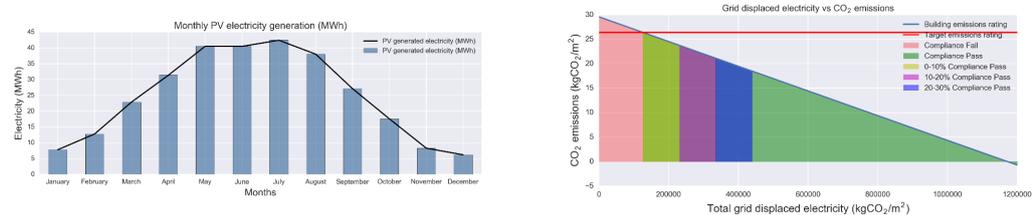
Figure 3.23: The left figure represents the seasonal monthly electricity generation by the 2000 m^2 of PV solar panel required to demonstrate Part L compliance. The right figure quantities the relation between carbon emissions to the increase in renewable energy generation from PV panels. It represents the amount of carbon emissions which needs to be mitigated for the performance beyond target carbon emissions.

3.4.2   Net-zero carbon emissions assessment

As the government is committed to climate change targets, the developed model is extended to quantify the improvement where non-domestic building will achieve net-zero emissions. The aim in this section is to calculate the

minimum amount of solar (PV) generation required to achieve net-zero carbon emissions. This will highlight the significance of an energy efficient and low carbon building design.

As shown in Figures 3.22 and 3.23, the renewable energy generation of 2326 MWh is required in order to achieve net-zero emissions for the non-domestic building used as a case-study. This will offset almost 11 million $kgCO_2/m^2$ of carbon emissions to make it net-zero emission building (NZEB). This can be demonstrated with the PV solar panel array of almost 15000 m^2. These calculations assumes that mono-crystalline silicon solar cells with an inclination of 15° facing south direction and parameters given in Table 3.15. The Python model provides flexibility to change any of these parameters. The Figure 3.24 shows the breakdown of end use carbon emissions for a net-zero emission building (NZEB).
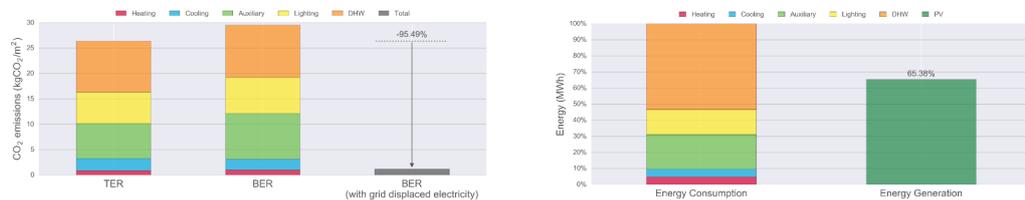
Figure 3.24: The figures shows breakdown of carbon emissions by end use among the actual building carbon emissions (BER) and the notional building carbon emissions (TER), and the reduction in carbon emission required with renewable energy generation to achieve net-zero emissions

3.4.3 TM-52 thermal comfort assessment - non domestic buildings

In this section, the TM-52 thermal comfort assessment has been performed for non-domestic buildings using IES-VE dynamic thermal modelling software which is in-agreement with CIBSE AM-11. The CIBSE TM-52 suggests using Predicted Mean Vote (PMV) to assess the risk of over-heating in conditioned

space connected with HVAC (Heating, ventilation, and air conditioning) system which is usually in the case of non-domestic buildings.

The Predicted Mean Vote (PMV) is a thermal scale originally developed by Fanger in 1970's which predicts the thermal comfort (experienced by occupants) within the scale of Cold (-3) to Hot (+3). It considers air temperature, mean radiant temperature, air velocity, relative humidity, occupant's activity and clothing insulation. The CIBSE Guide A recommends using Predicted Mean Vote (PMV) to assess the annual thermal comfort and should be maintained above -0.5 in winters and below +0.5 in summers, which is similar to the PMV range given in ASHRAE 55. The PMV values with corresponding comfort bands are given in Table 3.16.

Table 3.16: Predicted Mean Vote (PMV) standard comfort bands.

PMV Value	Corresponding sensation
-3	Cold
-2	Cool
-1	Slightly cool
0	Neutral
1	Slightly warm
2	Warm
3	Hot

The IES-VE software provides PMV for over-heating analysis and provides capability to use Python for detailed analysis. Therefore, a Python based model has been used for the detailed analysis which provides result for the recommended air velocity to achieve thermal comfort with over-heating assessment. This Python model looks at all occupied spaces and runs CIBSE TM-52 thermal comfort analysis method to determine whether those spaces show a risk of overheating. In addition, an air movement velocity is calculated for spaces which fail this assessment to indicate where movement of air could

help meet passing requirements. The method used for comfortable temperature estimation is from the standard BS:EN-15251.
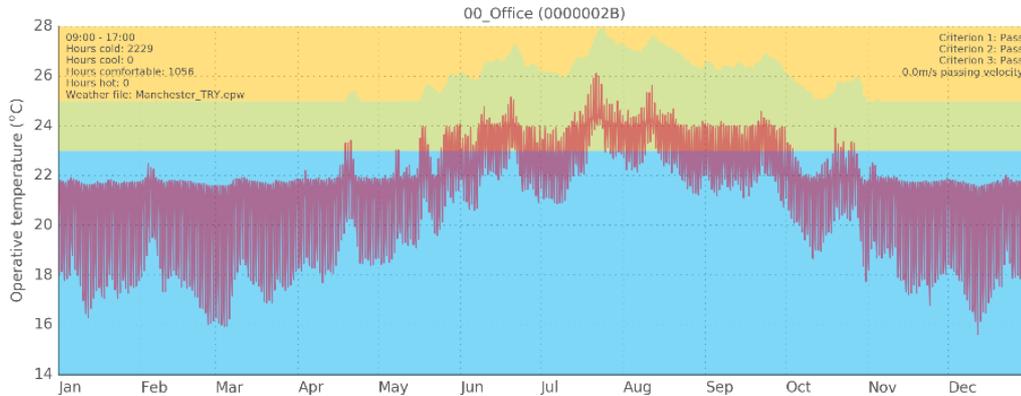
(a) Operative air temperature time-series throughout the year
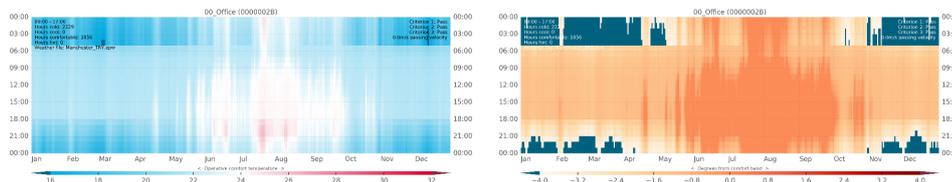
(b) Heatmap of comfort temperature

(c) Heatmap of PMV

Figure 3.25: Results from TM-52 risk of over-heating assessment for the office at ground floor in non-domestic building. These results are from the Python model developed in IES-VE software. As per TM-52 guidelines, the over-heating criteria are assessed between 9:00–17:00 hours during summers (May–September).

The occupied spaces in the non-domestic building are assessed against the temperature set-points given in Table 1.5 of the CIBSE Guide A and a high level of internal occupancy and equipment gains are assumed. Therefore, providing accurate thermal comfort during the worst possible condition is necessary. The steady state simulation is used to calculate the maximum heating and cooling loads in order to size heating, cooling and ventilation system. The dynamic thermal simulation provides hourly results for the in-depth analysis through out the year. Hence, the HVAC system (air delivery) is sized using steady state simulation and considering the supply of pre-conditioned cooled air at design flow rates and temperatures determined by the set-point
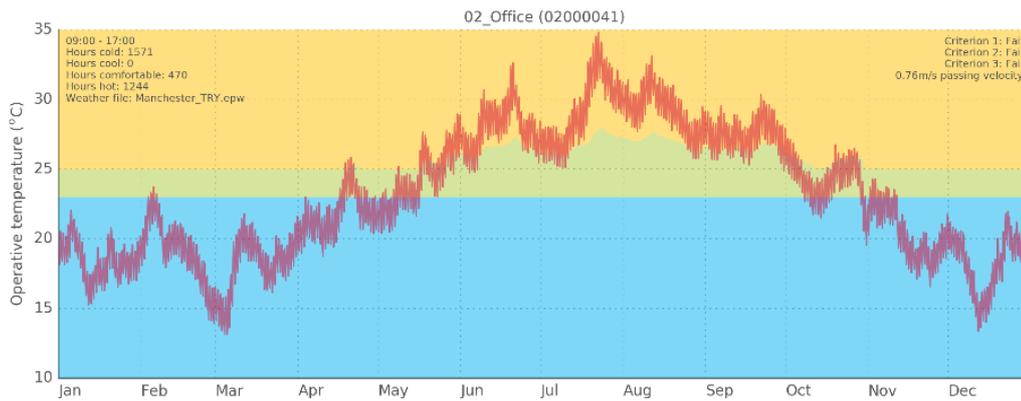
(a) Operative air temperature time-series throughout the year
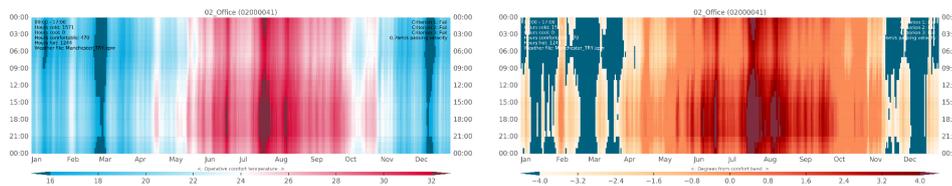
(b) Heatmap of comfort temperature

(c) Heatmap of PMV

Figure 3.26: Results from TM-52 risk of over-heating assessment for the office second floor in non-domestic building. These results are from the Python model developed in IES-VE software. As per TM-52 guidelines, the over-heating criteria are assessed between 9:00–17:00 hours during summers (May–September).

of between 16°C–26°C. The supply air volume is varied according to the internal room temperature and CO_2 concentration. The clothing level is assumed for the summer and outdoor activities (clo=1) and the activity level is taken as 'seated at rest' for the offices and restaurants, and 'sedentary work and standing' for the changing rooms and kitchen.

The results in Figures 3.25, 3.26 show that overall 85% (81/95) of the spaces in non-domestic building have passed TM-52 risk of over-heating assessment, where 83% (79/95) meet criterion I, 92% (87/95) and 85% (81/95) pass criterion 2 and 3, respectively. It is found that all important spaces attain the thermal comfort, except the toilets and few offices on the ground and second floor. This suggests that it is due to the low air ventilation (fresh incoming air velocity). The restaurants on each floor are observed to pass the TM-52

criteria and the PMV value observed by the occupants remains in the range between (-0.8–+0.8). On the other hand, offices on the second floor fail almost all criteria and minimum of 0.72m/s incoming fresh air velocity is required to overcome the risk of over-heating during the summer season (May–September).

3.5   Discussion and recommendations

The building regulations Part L1A and L1B apply on domestic buildings for the conservation of fuel and energy. The energy consumption for newly built domestic buildings is estimated using SAP software and for existing domestic buildings it is estimated using RdSAP software. In-spite of common perception, the SAP and RdSAP does not calculates the energy efficiency of a building but provides the cost-effective performance of a building by estimating the relative cost of energy. It is due to fact that both SAP and RdSAP are based on the BREDEM model, which does not consider several important parameters for calculations such as, occupancy behaviour, variations in internal gains, room air temperature, weather data and energy prices and forecasts. Therefore, it leads to the incorrect economic results for the energy efficiency and carbon emission calculations. This inherent flaw of SAP and RdSAP effects the underlying EPC estimates and forecasting of end-use energy savings. Hence, this presents with large variations between the estimated and actual energy performance of the domestic buildings and limits the application of energy efficiency techniques. This difference between the estimated energy and actual energy consumption of the buildings is also known as the performance gap.

The SAP and BREDEM methodology is outdated as it has never been validated for low energy buildings. Moreover, there is also profound need for

experimentation and validation on newly built houses to accurately measure the impact of energy efficiency measures. Since, a handful of newly built houses have failed to demonstrate minimum criteria for the compliance. The government has plans to update and introduce the latest version of SAP10 by 2020, which might overcome some of the limitations of current SAP2012.

While, looking at the building regulations Part L2A and L2B for the non-domestic building, the dynamic simulation modelling is well developed and robust compared to SBEM and BREDEM methodology. The NCM methodology is more detailed and apply strict limitations on the building services system. This leads to better energy performance outcome compared to the SAP and RdSAP. But, NCM methodology has few limitations for the designers, for example it always assumes the supply of hot water to toilets and in few scenarios leads to the pretty higher and unrealistic share of hot water demand. Moreover, there are no templates for certain building usage such as car-parks and needs to be modelled as a buffer space to demonstrate compliance.

Furthermore, the TM-52 over-heating assessment for thermal comfort is very well written standard with significant applications. However, there are certain things which needs to be added into it. It should be necessary for the designers to provide the hourly operative temperature comfort analysis to the owner. Moreover, the TM-52 standard must be extended to calculate the risk of over-heating for the entire year, instead of during the summer season i.e. May to September only.

In addition, the performance gap should be considered seriously and TM-54 (Evaluating Operational Energy Performance of Buildings at the Design Stage) should be integrated into the building regulations Part L in the UK.

Similarly, the Leadership in Energy and Environmental Design (LEED) for the buildings green certification in the US receives same recommendation regarding the estimated building energy performance and carbon emissions, and the need for liable testing and validation was proposed in the study conducted by Wedding and Brown (Wedding & Crawford-Brown 2007, 2008). The buildings should be made with the aim of design to perform instead for the sake to demonstrate building regulations compliance. The Australian buildings regulation NABERS (National Australian Built Environment Rating System) is a plausible example of how future buildings should be designed and regulated.

In conclusion, the building regulation methodologies are central to the UK government policy for estimation and improvement of the building stock. Though, SBEM, BREDEM, SAP and RdSAP are currently used methodologies, there is considerable performance gap which needs to be addressed to meet energy efficiency and future climate change targets.

3.6   Summary

In this chapter, the relationship between the building regulations and thermal performance of domestic and non-domestic buildings has been discussed along with significant contributions regarding the improvement of building regulations in the UK. Firstly, the energy performance of an existing domestic building has been studied before and after retrofit. The results show that the energy consumption reduces by almost 50% after the retrofit and the risk of over-heating is determined by the type of glazing being selected. Moreover, a non-domestic building, yet newly built is considered and a novel Python model has been implemented to analyse energy efficiency using dynamic simulations. The Python model quantifies the building carbon emissions (BER)

and calculates the renewable energy generation required to demonstrate build-
ing regulations compliance and ultimately net-zero emissions. Consequently,
it is concluded that with current building regulations in the UK, which forms
the basis for a building design, it is unlikely to achieve the target of net-zero
emission buildings (NZEB) in existing domestic and newly built non-domestic
buildings by the year 2050.

Chapter 4

Hydraulic modelling of the district heating network

Overview

This chapter discusses the design and operation of an energy efficient LTDH network and uses a real LTDH network (REMOURBAN project) as a case study. Firstly, the monitored data from the space heating systems of existing boiler based buildings is used to investigate hydraulic imbalance. Then, a novel hydraulic model developed in Python programming language is implemented to explore four different control scenarios in order to identify the optimum pumping strategy with least amount of pumping power, energy consumption and heat-losses in the low temperature district heating (LTDH) network. Secondly, the analysis is extended to the physical component based modelling using Dymola programming language, where the energy efficient operation of the network is discussed with optimum design flow-rates parameters, pump control strategies and integration of multiple heat sources in the network. Finally, the results are validated using monitored data from the REMOURBAN

project and recommendations for the energy efficient design and operation of the LTDH network have been made.

4.1 Project description

The LTDH has evolved to be the most energy efficient district heating technology and its design and operation requires careful consideration. The high return water temperature in Nottingham district heating network shows sufficient capacity for a secondary LTDH network intervention to the nearby areas using return pipe of the network, rather than extending the existing high temperature district heating network and discussed in Section 2.3.1.

This secondary LTDH intervention for 94 flat consumers is first of it's kind in the UK to utilise return water pipe of the district heating network. This will provide a gateway to Nottingham in efficiency improvement and extension of existing district heating network. It is anticipated that lowering supply water temperatures will further reduce energy losses in the network by almost 75% compared to the existing district heating networks. The overall aim is to reduce energy losses and evaluate hydraulic performance of this LTDH network (60/30) with integration of multiple heat sources.

In this study, a real time LTDH network (REMOURBAN project) has been used as a case-study and the design and operation of this LTDH network has been investigated for 94 flats in four blocks i.e. Byron, Keswick, Haywood and Morley court. The solution for improving its energy efficiency have been discussed in two parts. The first part explores the effect of supply water temperatures with variations in flow-rates and optimum pumping strategy from the plant room. This analysis for the optimum pumping strategy is performed on the Byron court (31 out of 94 flats). Then, the second part uses Dymola

software and investigates energy efficiency of the network with various design flow-rates parameters, pump control strategies and integration of multiple heat sources in the network.

4.2 Python programming based model

The methodology for Python programming based hydraulic model is divided into two section namely, heat demand modelling and hydraulic modelling. The heat demand for each flat is calculated in IDA-ICE software using real monitored weather data for Nottingham. Then, the hydraulic modelling is performed using a mathematical hydraulic model developed in Python programming language which provides flexibility to integrate several thermal models in the analysis. This methodology is shown graphically in Figure 4.1.

4.2.1 Heat demand modelling of the buildings

The heat demand is calculated using dynamic thermal modelling in IDA-ICE 4.6.2 software. This software is a powerful design tool and commonly used to optimise building's energy usage and thermal comfort of occupants. This is performed by creating a model of the proposed building considering its geometry, geographical location, orientation and weather data, building fabric properties, building services such as heating, cooling and equipment gains, and building usage (heating and cooling set-points) and occupancy profiles. Then, the model is simulated with results as mentioned in Section 3.3.

4.2.2 Hydraulic modelling of the LTDH network

The developed hydraulic model takes into account a number of assumptions in order to minimise the computation processing time without compromising accuracy. These include; a) the medium (water) is in-compressible, b) no
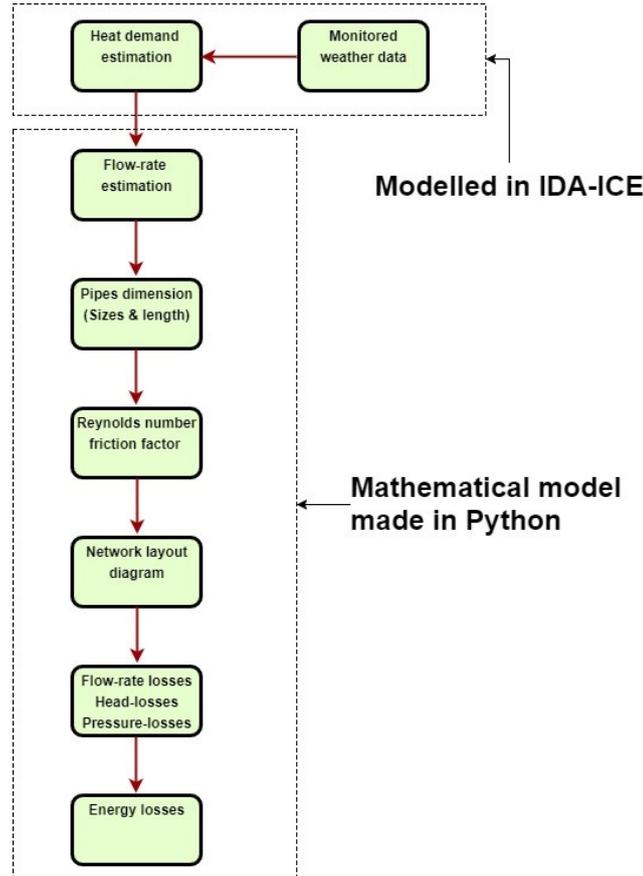
Figure 4.1: The methodology of the hydraulic model developed in Python programming language

leakage from the pipes and c) the district heating pipes (supply and return) are symmetrical.

The first step in Python based model is to calculate the flow velocity (v) and flow-rate (q) in each pipe at the design supply (T_s) and return (T_r) water temperature across the district heating network. This is done by first taking the heat demand (Q) calculated earlier in Section 4.2.1. The specific heat of water (C_p) and density (ρ) decreases with the increase in water temperature and are calculated using Equations (4.2, 4.3). This leads us to accurately calculate the mass-flow rate (m) using Equation (4.4).

$$Q = mC_p(T_s - T_r) \tag{4.1}$$

$$C_p = 4209.1 - (132.8 * 10^{-2} * T_{water}) + (143.2 * 10^{-4} * T_{water}^2) \qquad (4.2)$$

$$\rho = 1000.6 - (0.0128 * T_{water}^{1.76}) \qquad (4.3)$$

$$m = \frac{Q}{C_p(T_s - T_r)} \qquad (4.4)$$

where, the relation between flow-velocity (v) and mass-flow rate (m) with area (A) is given as,

$$m = \rho.A.v \qquad (4.5)$$

$$v = \frac{m}{\rho.\frac{\pi}{4}.d^2} \qquad (4.6)$$

The diameter (d) of pipes is taken from the project documents and the Equation (4.6) is used to calculate flow-velocity with the constraint to limit the maximum allowable flow-velocity of less than or equal to 2 m/s, as per recommendations from the design standards (CIBSE Jul 2015). Similarly, the relation between volumetric flow-rate or simply flow-rate (q) and flow-velocity (v) is

$$q = A.v \qquad (4.7)$$

$$q = \frac{\pi}{4}d^2.v \tag{4.8}$$

The Equation (4.8) is used to calculate the flow-rate (q) in pipes. In district heating network the flow of water is limited to the design values with control valves. The flow-rate across the control valve for the given flow regulation capacity (K_v) and differential pressure across the valve (Δ P) is given by

$$q = K_v\sqrt{\Delta P} \tag{4.9}$$

Pressure-loss calculation

The pressure-loss in pipes depends on the friction factor, flow-rate, diameter and length of each pipe. The friction factor (f) is further influenced by the dimensionless Reynold's number of flow (R_e), flow-rate (q), kinematic viscosity (ν), dynamic viscosity (μ), and roughness of inner pipe surface (e) (Çengel 2007). The district heating pipes are commonly made of steel with the roughness of $0.05mm$ and the friction factor (f) is calculated using following equations from the principles of fluid dynamics.

$$\nu = \frac{\mu}{\rho} \tag{4.10}$$

$$R_e = \frac{4.q}{\pi.d.\nu} \tag{4.11}$$

if Re is ≤ 2000 then,

$$f = \frac{64}{R_e} \tag{4.12}$$

otherwise, the f is calculated using Swamee-Jain Equation (Çengel et al. 2012),

$$f = \frac{0.25}{\left[log \frac{e/D}{3.7} + \frac{5.74}{Re^{0.9}} \right]^2} \tag{4.13}$$

the head-loss (Δh_f) and pressure-loss (ΔP) in pipes is estimated by Darcy equation as,

$$\Delta h_f = \frac{f.l.v^2}{2.g.d} = \frac{8.f.l.q^2}{\pi^2.d^5.g} \tag{4.14}$$

$$\Delta P = \rho.g.\Delta h_f \tag{4.15}$$

Pumping power and energy consumption calculation

The shaft pumping power (P_a) and electricity consumption (E) in the district heating network are estimated using following equations,

$$P_a = \Delta P.q = \frac{\rho.g.\Delta h_f.q}{\eta} \tag{4.16}$$

$$E = \frac{S * P_a}{\eta_m} \tag{4.17}$$

where, S is the security factor assumed as 1.1, g is the gravitational constant, η and η_m are the efficiency of pump (0.85) and electric motor (0.70), respectively.

Thermal resistance and heat-loss calculation

The district heating pipes experience thermal resistance which impacts the outlet water temperature and heat-losses in the network. Its calculation is

important in low temperature heating systems as the supply temperatures are already at the threshold to fulfill consumers demand. The thermal resistance relies upon composition material of pipes and temperature difference between the ground surface and water inside the pipes.

The thermal resistance model implemented in this analysis, calculates heat-losses of each pipe in network by taking into account the effect of individual components in pipes. It ignores the resistance of steel pipe mantle due to its insignificant effect, but considers the resistance from insulation (R_i) and ground surface (R_g) with respect to the thermal conductivity of their composition materials (Palsson 1997). The model also takes into account the effect of thermal resistance (R_h) due to heat transfer between the supply and return temperature pipes of the district heating network. These thermal resistances have been modelled in the Python based developed model and the mathematical expressions to define them are,

$$R_i = \frac{1}{2\pi\lambda_i} * ln\left(\frac{D_m}{D_o}\right) \tag{4.18}$$

$$\lambda_i = 0.0196734 + 0.000080747303 * T_{water} \tag{4.19}$$

where, λ_i is the thermal conductivity for insulation PEX foam and determined by temperature, wetness levels and aging of the insulation (Lund & Mohammadi 2016). Its value varies between 0.024 - 0.026 W/m^{o}C and is calculated using expression given in Equation (4.19) to consider variations in water temperature, and validated from experimental results (Gabrielaitiene et al. 2010).

$$R_g = \frac{1}{2\pi\lambda_g} * ln\left(\frac{2H}{D_m} + \sqrt{\left(\frac{2H}{D_m}\right)^2 - 1}\right) \tag{4.20}$$

$$H = S_d + 0.0685 * \lambda_g \tag{4.21}$$

The S_d is the depth of pipe from the surface soil level, λ_g is the thermal conductivity of ground and assumed constant. H is the effective burial depth and the resistance due to convective and radiation heat transfer between the air and ground surface (Palsson 1997, Bøhm 2010).

$$R_h = \frac{1}{2\pi\lambda_g} * ln\left(\sqrt{\left(\frac{2H}{S_c}\right)^2 + 1}\right) \tag{4.22}$$

$$S_c = L_c + Dm \tag{4.23}$$

The L_c is distance between the supply and return water pipe, D_m is the pipe diameter. Several other thermal resistances between different elements inside the pipe are modelled using following equations from the principles of fluid dynamics (Palsson 1997). These include water-insulation (R_{wi}), ground-surroundings (R_{gu}) and insulation-ground (R_{ig}) as shown in Figure 4.2).

$$R_{wi} = \frac{1}{2\pi\lambda_i} * ln\left(\frac{1 + \frac{D_m}{D_o}}{2}\right) \tag{4.24}$$

$$R_{gu} = \frac{1}{2\pi\lambda_g} * ln\left(\frac{4H}{D_m + D_g} + \sqrt{\left(\frac{4H}{D_m + D_g}\right)^2 - 1}\right) \tag{4.25}$$

$$R_{ig} = R_i + R_g - R_{wi} - R_{gu} \tag{4.26}$$
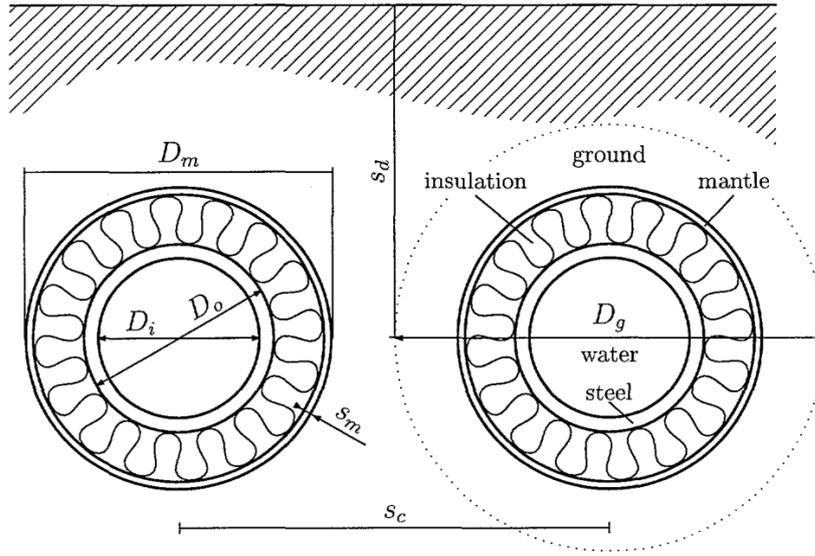
Figure 4.2: Cross section of a pair of supply are return district heating pipes buried in ground (Palsson 1997).

The scaling factor (θ) is then calculated using above thermal resistances to find the true steady state heat-loss (Palsson 1997).

$$\gamma_{supply} = \frac{\Delta T_r}{\Delta T_s} = \frac{T_{rp} - T_u}{T_{sp} - T_u} \tag{4.27}$$

$$\gamma_{return} = \frac{\Delta T_s}{\Delta T_r} = \frac{T_{sp} - T_u}{T_{rp} - T_u} \tag{4.28}$$

$$\theta = (R_i + R_g)\frac{(R_i + R_g) - \gamma R_h}{(R_i + R_g)^2 - R_h^2} \tag{4.29}$$

The T_u is the soil temperature and the factor γ_{supply} for supply pipe, is the proportion of temperature difference between the return and the supply water. Similarly, γ_{return} for the return pipes, is the opposite of γ_{supply}. Moreover, the total heat transmission resistance (R_{total}) and overall heat transfer coefficient (U) are computed by adding all thermal resistances between different elements of the pipe

$$R_{total} = R_{wi} + R_{gu} + R_{ig} \qquad (4.30)$$

$$U = \frac{1}{R_{total}}\theta \qquad (4.31)$$

Finally, the distribution heat-loss (Q_{loss}) and outlet water temperature (T_{outlet}) per pipe calculated from thermal resistance model is given by,

$$Q_{loss} = \sum_{t}^{8760} U * l * (T_{supply} - T_u) \qquad (4.32)$$

$$T_{outlet} = (T_{inlet} - T_u) exp^{-\frac{U.L}{C_{p}.m}} + T_u \qquad (4.33)$$

where, the outlet water temperature (T_{outlet}) is inlet water temperature (T_{inlet}) for the next pipe in the network and depends on the soil temperature (T_u), mass flow-rate (m) and heat-loss from the pipe. Equation (4.33) is adapted from the (Liu et al. 2016) and explains that, the outlet temperature is directly proportional to the mass flow-rate and indirectly proportional to the length as well as the overall heat transfer coefficient of pipe.

4.3 Monitored space-heating - data analysis

The existing boiler based heating systems in the UK operates at constant flow-rate and the hydraulic balance is achieved by controlling supply water temperature with heat consumption variations. This strategy maintains the hydraulic balance, but reduces the delta t (Δt) and system's energy efficiency.

In the REMOURBAN project, there has been no changes in the previously installed space-heating system inside the flats, but only gas-boilers are replaced

Table 4.1: List of radiators installed in a flat from REOMRBAN project (Ashfaq & Ianakiev 2018*b*).

No.	Room	Radiator size	Radiator type	Power output (W)
1	Hallway	1100 x 600	Single panel, single convector	1,100
2	Lounge	1600 x 600	Double panel, double convector	2,845
3	Bathroom	500 x 600	Double panel, double convector	889
4	Kitchen	400 x 600	Double panel, double convector	711
5	Bedroom - 1	1400 x 600	Single panel, single convector	1,400
6	Bedroom - 2	1000 x 600	Single panel, single convector	1,000
7	Bedroom - 3	400 x 600	Double panel, double convector	711

with the heat interface units (HIU) which is connected to the district heating network. Therefore, the space-heating system data (heat demand, supply and return water temperature data) is monitored from one of the flat to understand thermal comfort and current operation of system before the implementation of LTDH. The schematic illustration of space-heating system loop inside the flat is shown in Figure 4.3 and the detailed description of the installed hydronic radiators is given in Table 4.1.
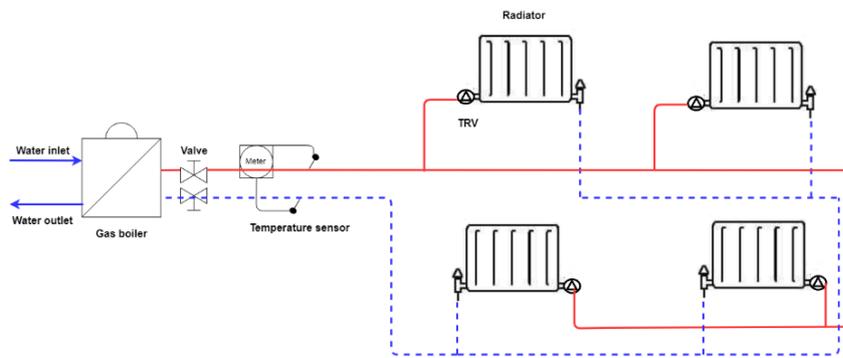
Figure 4.3: Schematic of the space-heating system loop inside the flat from RE-MORBAN project. The space-heating system is double string system with thermostatic radiator valves (TRVs) installed on plate radiators (Ashfaq & Ianakiev 2018*b*).

The existing flat contains individual gas-boiler where the room temperature is maintained by TRVs installed on radiators. The monitored supply and return water temperature data has been plotted against the outdoor air temperature as shown in Figure 4.4. It is observed that the supply water tem-
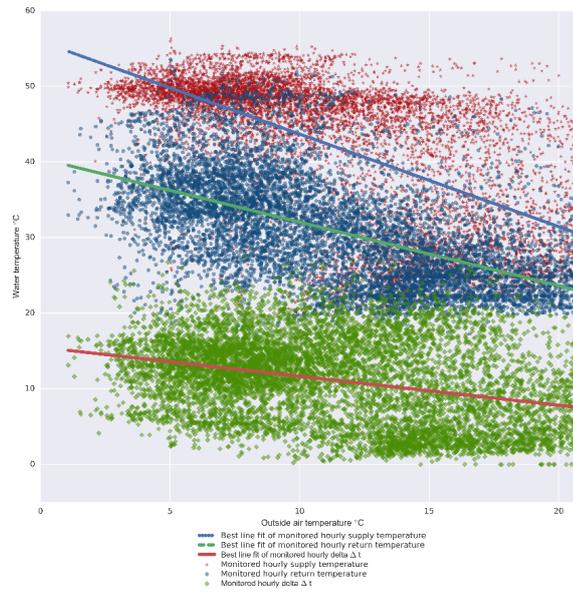
Figure 4.4: Hourly monitored supply and return water temperature data of space-heating system from an existing boiler based building. The scatter plot depicts the relationship between network temperatures and outdoor air temperature. The best line of fit for the hourly monitored supply and return water temperature depicts negative correlation between the water temperature and outside air temperature (Ashfaq & Ianakiev 2018*b*).

perature decreases rapidly with increase in outside temperature compared to the return water temperature. The Δt of space-heating system is highest when the outside temperature is 0°C and lowest around 23°C, respectively. However, the average Δt throughout the year remains around 11 ^{o}C. The best line of fit for the hourly monitored supply and return water temperature depicts negative correlation between the water temperature and outside air temperature. This methodology for the hydraulic imbalance evaluation is adapted from Zhang et al. (Zhang et al. 2016) and the results are in agreement with these findings.

It is concluded that in order to maintain high Δt, the regulation of return water temperature with respect to outside temperature is more important compared to the supply water temperature. Moreover, the variations in return

water temperature compared to supply water temperature, at the same outdoor temperature, is due to the hydraulic imbalance in space-heating system. This hydraulic imbalance issue can be explained due to the over-sizing of room radiators and other control equipment inside the building. It is commonly assumed that the installation of TRVs on radiators automatically regulates the flow-rate with respect to indoor temperature and improves the Δt, but the hourly monitored space-heating system data suggests that it is not the case and the Δt still remains relatively low.
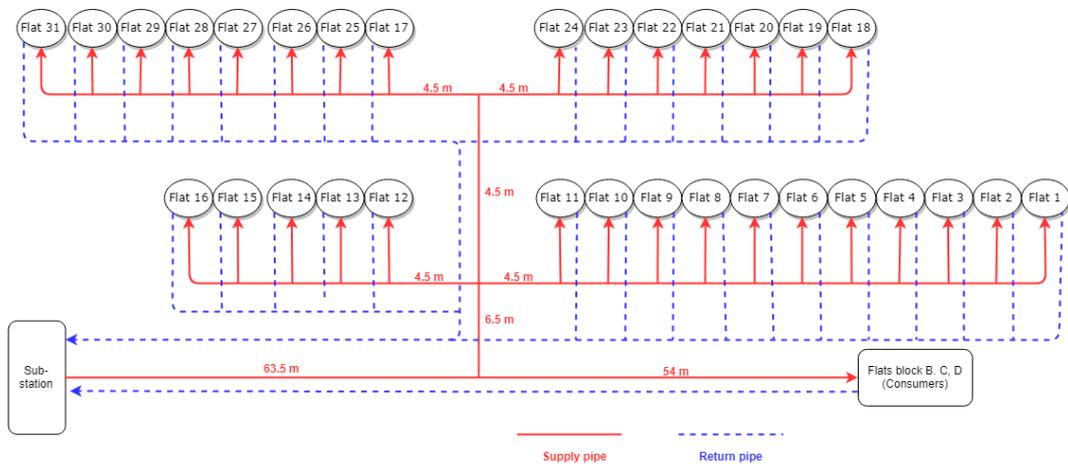
Figure 4.5: The network layout of low temperature district heating network from REMOURBAN Project in Nottingham, UK (Ashfaq & Ianakiev 2018*b*).

4.4   Results and analysis

4.4.1   Hydraulic modelling using Python

Initially, one of the buiding (Byron court) consisting of 31 flats is selected for modelling and all parameters are taken from the REMOURBAN project. The schematic of building network is shown in Figure 4.5. The hydraulic model from Section 4.2 is used for four different scenarios to evaluate the most appropriate pumping strategy. The scenarios 1 and 2 assumes constant flow-rate, whereas scenarios 3 and 4 are with variable flow-rate. The flow-rates,

pumping power, energy consumption and heat-losses are compared among all scenarios and the optimum scenario is concluded.

4.4.1.1 Scenario 1 and 3 - constant supply water temperature

The scenarios 1 and 3 assume constant supply water temperature of 60°C from the plant room. The flow-rates are kept constant in scenario 1 and variable in scenario 3 using variable speed pumps.

The results show that variable flow-rates in scenario 3 reduces energy consumption of the LTDH network compared to the scenario 1 by 63% (from 964 KWh to 360 KWh). This reduction suggests that variable speed pump increases energy efficiency of the network during partial heat consumption. Similarly, the head-loss at each flat in scenario 3 is comparatively lower than in scenario 1. The comparison between energy consumption and head-loss among different operational scenarios is shown in Figures 4.6, 4.7, 4.8 and elaborated in Table 4.2.

While comparing heat-losses it is found that in scenario 3, the heat-losses in LTDH network decreases from 62% to 47%. This reduction by almost 14% shows the heat-losses have strong dependence on the supply water temperature than flow-rate variations. On the other hand, the flow-rate and pumping power are comparatively higher in summer. This is due to low Δt and shown in Figure 4.9.

4.4.1.2 Scenario 2 and 4 - variable supply water temperature

The scenarios 2 and 4 consider situation where the supply water temperature varies with respect to outdoor air temperature from the plant room. The flow-rates are constant in scenario 2 and variable in scenario 4 using variable speed pumps and weather compensation control valves in the LTDH network.
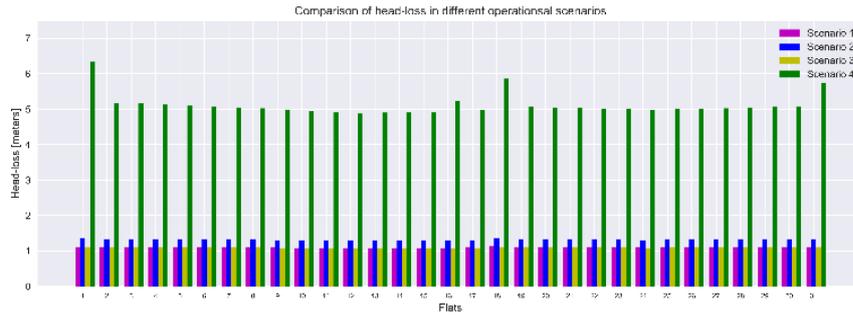
Figure 4.6: Head-loss comparison to each flat in different operational scenarios (Ashfaq & Ianakiev 2018*b*).
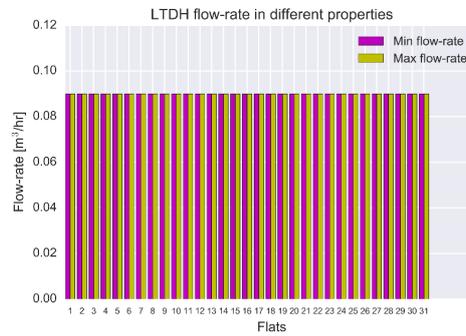
Figure 4.7: Figure represents the maximum and minimum flow-rate restricted to each flat in the LTDH network (Ashfaq & Ianakiev 2018*b*).

Table 4.2: Comparison between different operational scenarios for the LTDH network (Ashfaq & Ianakiev 2018*b*).

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Flow-rate	constant	constant	variable	variable
Supply water temperature	constant	variable	constant	variable
Maximum flow-rate from plant room (m³/hr)	5.34	5.87	5.31	11.71
Minimum flow-rate from plant room (m³/hr)	5.34	5.87	0.82	0.05
Heat-loss from district heating pipes (MWh)	14.86	9.34	14.85	9.34
Maximum pumping power (KW)	70	92.68	68.82	734.07
Energy consumption (KWh)	964	1261	360	1189
Heat-losses in LTDH network (MWh)	62.25%	9.6%	47.66%	11.13%

The results show that variation in supply water temperature increases the energy consumption of the network. The energy consumption increases to 1261 KWh in scenario 2 and 1189 KWh in scenario 4. This increase in energy
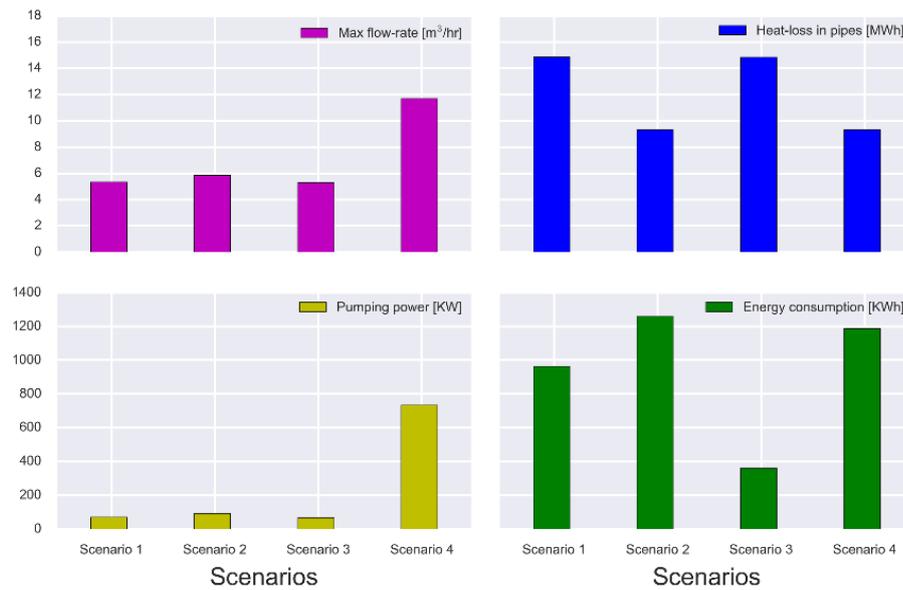
Figure 4.8: Comparison of results between different operational scenarios. First row represents maximum flow-rate from the plant room and heat-loss from district heating pipes in the ground during heat transmission. The second row represents pumping power and energy consumption of the LTDH network(Ashfaq & Ianakiev 2018*b*).

consumption can be explained by the increase in required flow-rates due to the variations in supply water temperature from the plant room. Similarly, the flow-rates and head-loss in scenario 2 and 4 are higher compared to the scenarios with constant supply temperature. These energy consumption and head-loss results are shown in Figures 4.6, 4.8 and elaborated in Table 4.2.

It is observed that the heat-loss from district heating pipes in both scenarios reduces by 37%. Though, the flow-rate and pumping power are comparatively higher in scenario 4, the heat-loss in scenario 4 is even lower than scenario 3. These results are shown in Figures 4.9 and 4.10. It is concluded that reducing supply water temperature is more effective compared to flow-rate variations for reduction in heat-losses in the LTDH network. The overall heat-losses in LTDH network in scenarios 2 and 4 are reduced to just 9% and

11%, respectively. These heat-losses are minimal compared to other scenarios.

These results along with comparison are further elaborated in Table 4.2.
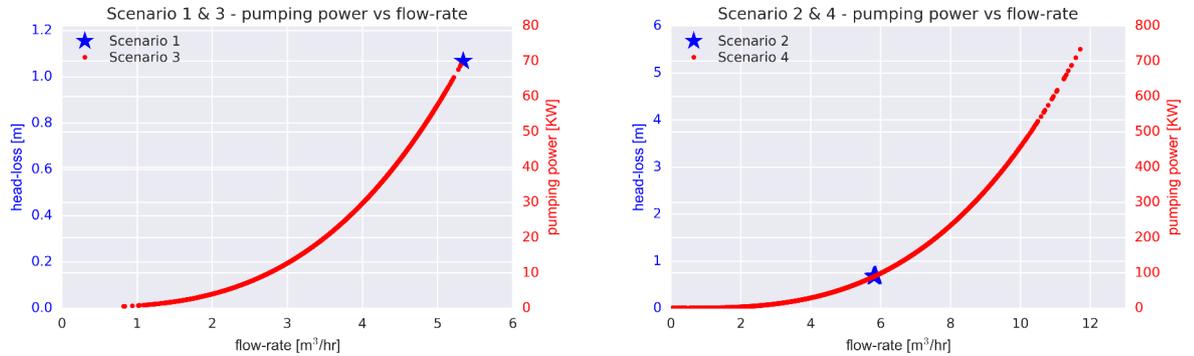
Figure 4.9: Comparison of pumping power and flow-rates between different operational scenarios for the LTDH network. Left figure compares result for scenarios 1 and 3, whereas right figure compares result for scenarios 2 and 4 (Ashfaq & Ianakiev 2018*b*).
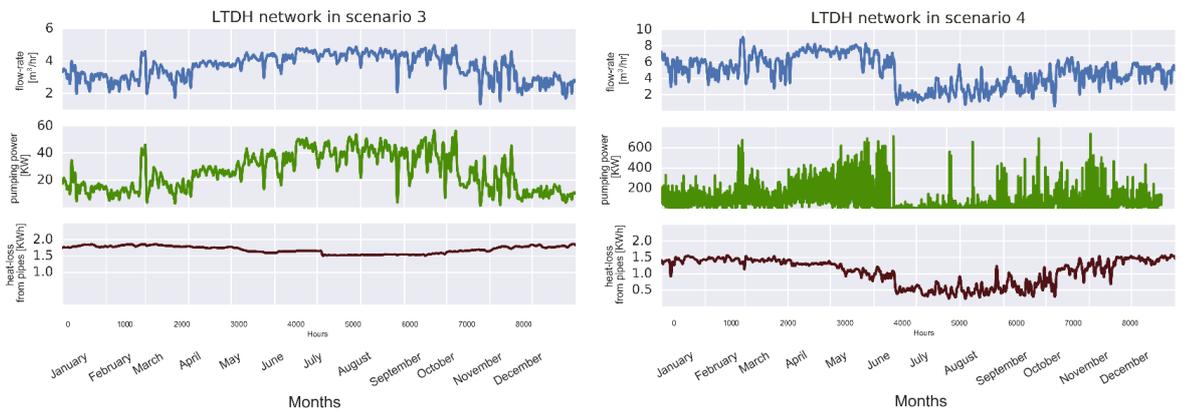
Figure 4.10: Comparison between the operational scenario 3 and 4. First and second row represents hourly flow-rate and pumping power from the plant room, whereas the third row represents hourly heat-loss from district heating pipes in the ground during heat transmission (Ashfaq & Ianakiev 2018*b*).

4.4.2 Hydraulic modelling using Dymola

The results from previous Section 4.4.1 provide basis for the pumping strategy to be implemented in the low temperature district heating network. The analysis is further extended on the entire LTDH network of 94 flats (four blocks

shown in Figure 2.4, 4.11) and the detailed physical component based thermo-hydraulic model is developed in Dymola software version 2019. The Dymola (DYnamic MOdelling LAboratory) is a user interface and uses Modelica language compiler owned and maintained by Dassault Systemes. The Modelica programming language itself is an open source language. In this analysis, the licensed HVAC library (for Heating Ventilation and Air Conditioning) is used to investigate the effect of different parameters on the energy efficiency of the network. To this end, all flats in the LTDH network are simulated for different set of parameters and for the sake of convenience the term iteration will be used in rest of the chapter. The baseline iteration refers to the existing network and used as a benchmark for comparing against other iterations.

Figure 4.11: Aerial view of 94 flats from the REMOURBAN project site

The entire LTDH network is divided into several modules where each module represents a separate building in the network i.e. main plant room, Byron court, Keswick court, Haywood court and Morley court. The LTDH network modules modelled in Dymola are shown in Figures 4.12, 4.13, 4.14. The actual LTDH network uses heat from the return pipe of existing district heating network (primary network) as a heat source to the secondary LTDH network from the heat-exchanger and shown in Figure 2.4. For the sake of

simplicity, the heat source of this secondary LTDH network is assumed as gas-boiler in the simulations. In the plant room, a control loop is also modelled on the circulation pump which switches it off once the outside air temperature becomes equal or greater than 14°C. The plant room also contains a thermal heat storage and assumes the direction of inside flow as counterflow. This increases the heat transfer inside thermal heat storage.
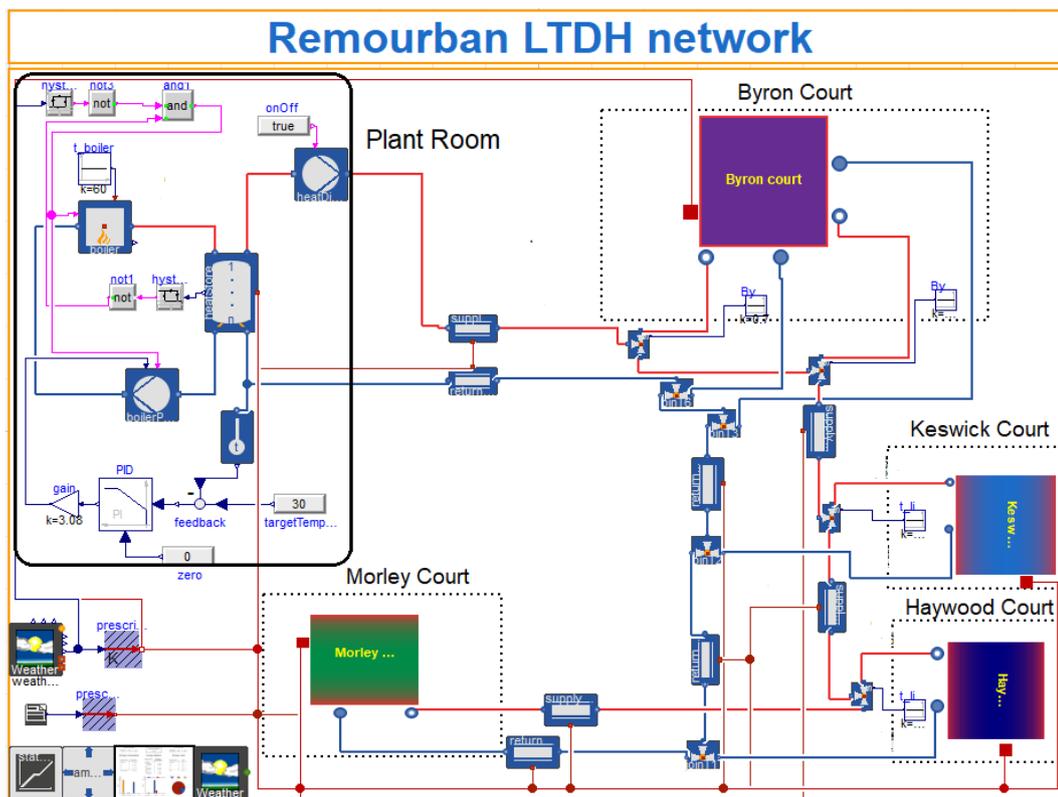
Figure 4.12: LTDH network model in Dymola split into several separate modules

The LTDH network contains three types of flats i.e one bedroom, two bedroom and three bedroom. Each building has different topology among Byron, Keswick, Haywood and Morley courts. Therefore, each building is modelled separately depending on the type and location of each flat, pipe sizes and parameters (such as, construction, heat demand profiles, heat-losses and internal gains). The model of different flats along with the internal gain profile considered in simulations is shown in Figure 4.14.
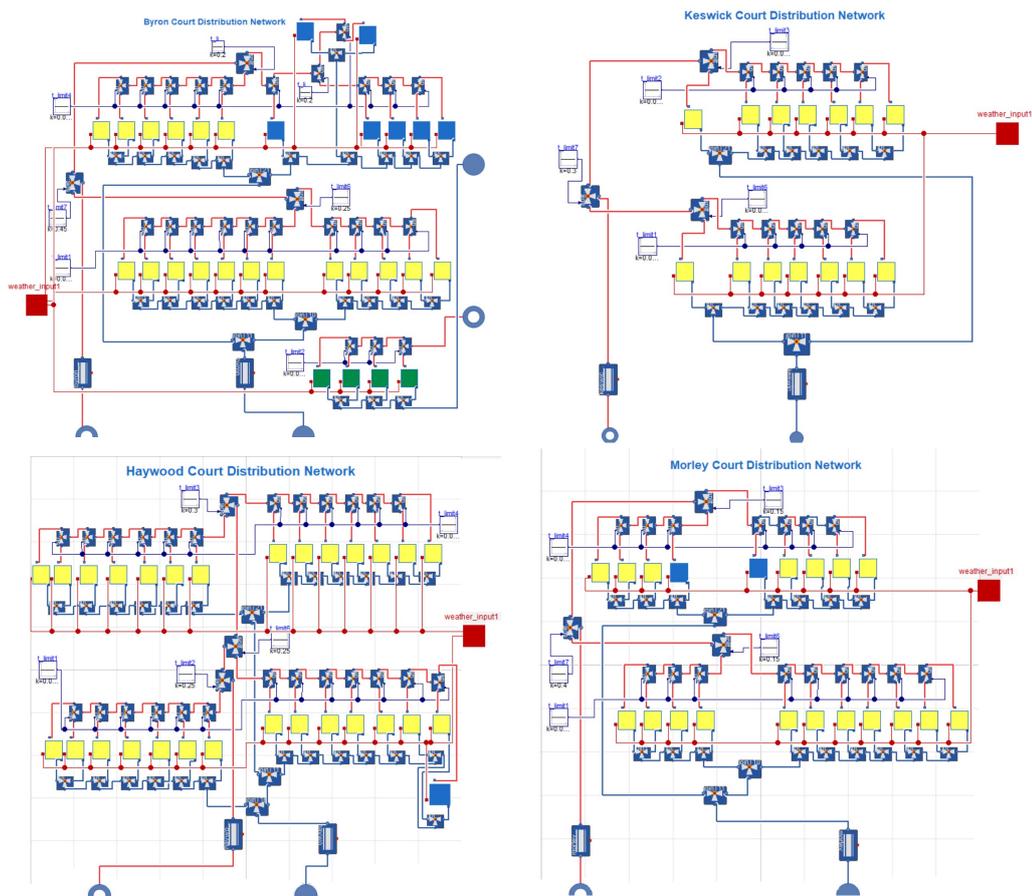
Figure 4.13: Modules for Byron, Keswick, Haywood and Morley court of the RE-MOURBAN LTDH modelled in Dymola software. Where, green boxes represent three bedroom flat, yellow boxes represents two bedroom flat and blue boxes represent one bedroom flat. The red box shows the integration of outside weather data for heat demand estimation of the each flat.

The simulation uses exact parameters from the REMOURBAN project (also given in Appendix) and weather data (outside air temperature, solar radiations, relative humidity and soil temperature) as discussed in Section 3.3.1. These simulations replicate the operation of existing district heating network and are performed at hourly intervals for the month of highest heat demand i.e. January. This provides a meaningful comparison among different iterations.
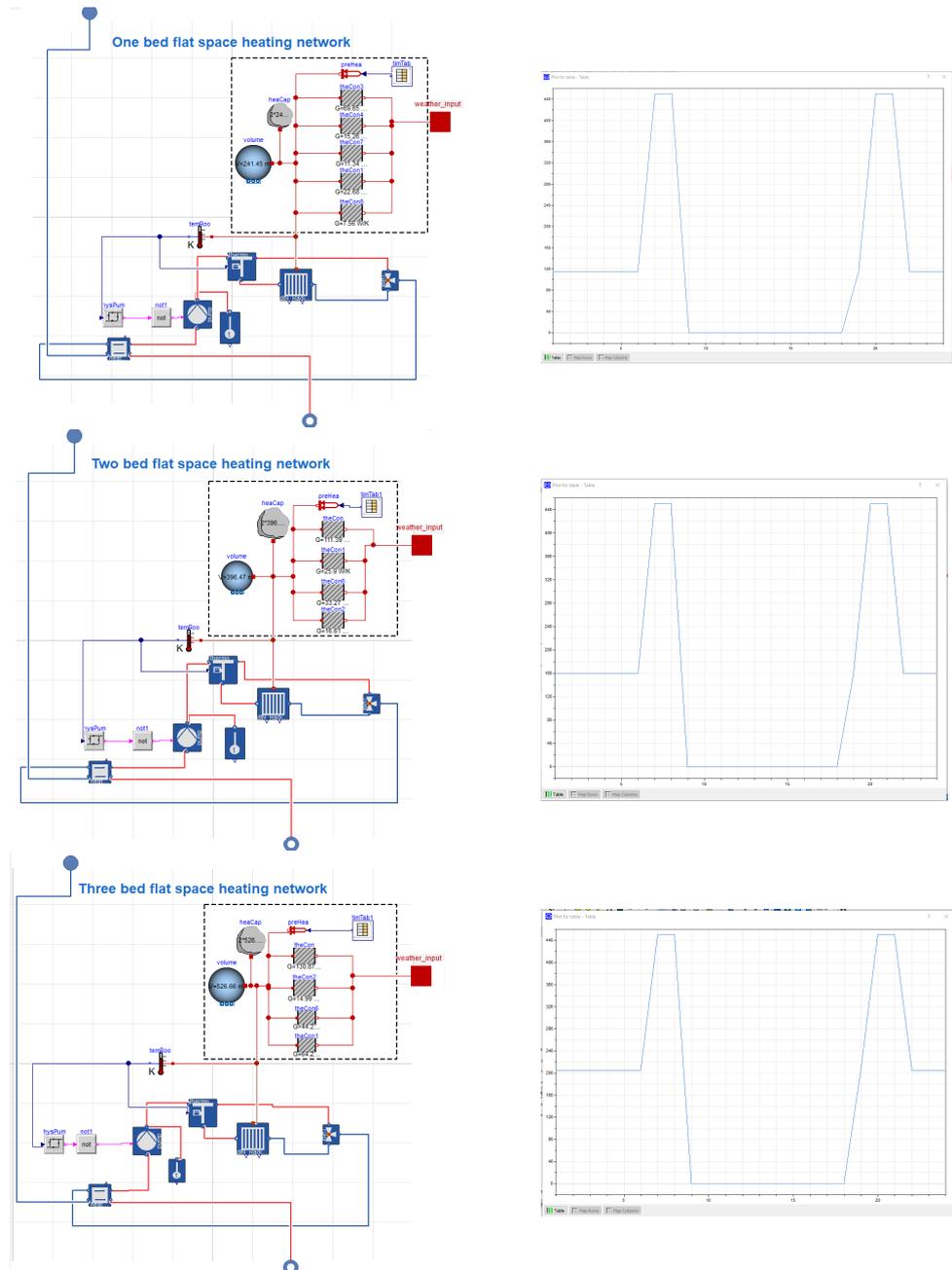
Figure 4.14: Sub-modules showing one, two and three bedroom flats along with their respective internal gains modelled in Dymola software.

4.4.2.1 Baseline iteration

The baseline iteration is the representation of existing low temperature district heating network. The network is simulated with its default parameters to understand the operation, energy consumption and heat losses of the network.
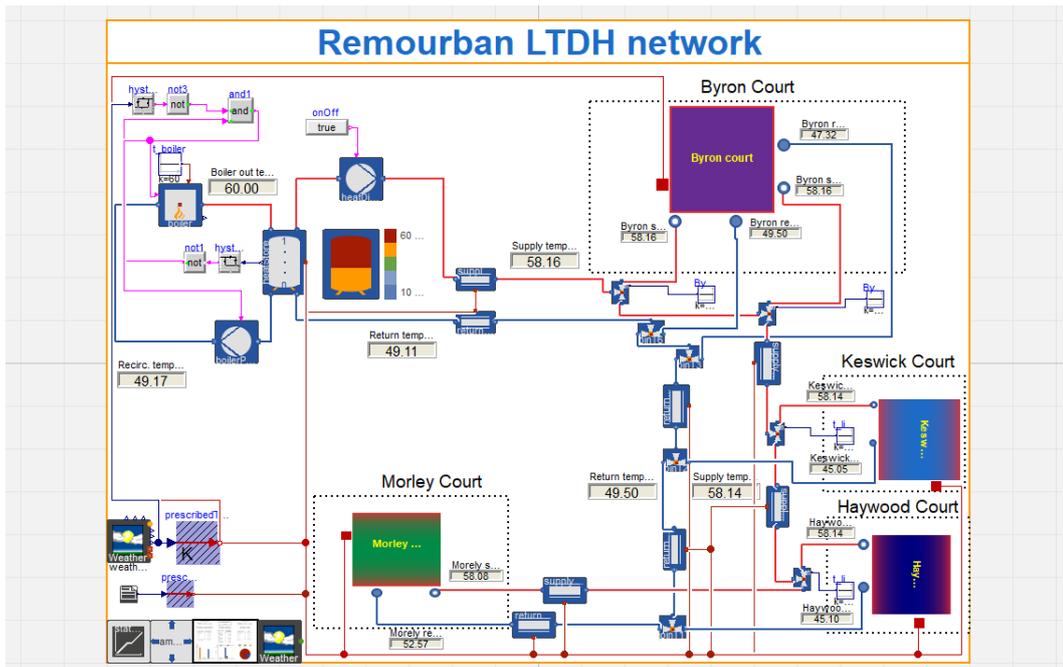
Figure 4.15: Results from the baseline iteration simulation depicting the performance of existing REMOURBAN LTDH network

The simulation in Figure 4.15 shows that the network has huge amount of energy demand and the network's supply water temperature degradation is not an issue. Due to the transmission heat-losses, the temperature of supply water decreases to 58^oC once it reaches to the Byron court and remains consistent for all the buildings, even for the Morley court which is the furthest building in the network. Moreover, there are significant heat losses in the plant room, especially, from the circulation pump and thermal heat storage. The circulation pump is designed to operate at reasonably high flow-rate and responsible for the high water temperature inside the thermal heat storage. It works as buffer in the network and the water temperature varies between $50 - 60^oC$ between upper/middle and lower sections of the thermal heat storage.

The comparison between supply and return water temperature for different buildings in the network and indoor air temperatures of flats is shown in Figure 4.16. This variation in return water temperature in the network and

Figure 4.16: Results from baseline model simulation. Comparison between supply and return water temperature for different buildings in the network and indoor air temperature in flats

indoor air temperature can be explained due to the heat demand profiles of different flats.

Table 4.3: Overview of flow-rates and other parameters in baseline model of LTDH network

Location	Pipe size (mm)	Pipe length (m)	Flow-rate (kg/s)	Serving
Main plant room	90	71	5.14	Entire network
Byron court under-croft	90	5	1.45	31 flats
Keswick court under-croft	75	56.5	0.65	14 flats
Haywood court under-croft	75	22	1.36	29 flats
Morley court under-croft	76	63	1.68	20 flats
Circulation pump	-	-	4.08	Plant room

The simulation results show that the Δt is very low (2^oC) in the network and the return water temperature is 49.11^oC in the plant room. This high return water temperature can be explained due to the high flow-rates in the network. The flow-rates considered in LTDH network simulations for different buildings are given in Table 4.3.

The simulation predicts that the overall CO_2 emissions for the month of January are around 37,513 kg. The energy statistic results in Figure 4.17 shows that the share of electricity utilised for pumping the pressurised hot water in the network is 31% (21,607 KWh) and the gas used as a fuel for heat generation

Figure 4.17: Energy statistics results from baseline low temperature district heating network modelled in Dymola software depicting the performance of existing network

through the boiler is 69% (174,243 KWh) in the total energy consumption. The network generates approximately 151,347 KWh and consumes 145,561 KWh of heat. This difference of approximately 5,900 KWh between the generated and consumed heat in the network reflects the high amount of heat being wasted in the district heating network.

4.4.2.2 Iteration 1

The Iteration 1 is the optimisation of existing low temperature district heating network. The flow-rates in the network are reduced and the operation and energy consumption of the network is investigated. A PID (proportional integral derivative) controllers based feedback loop is also implemented to regulate mass flow-rate of the circulation pump with respect to the return water temperature of the network.

The simulation results in Figure 4.18 shows that the Δt of network increases to the designed parameters and the operational energy demand de-
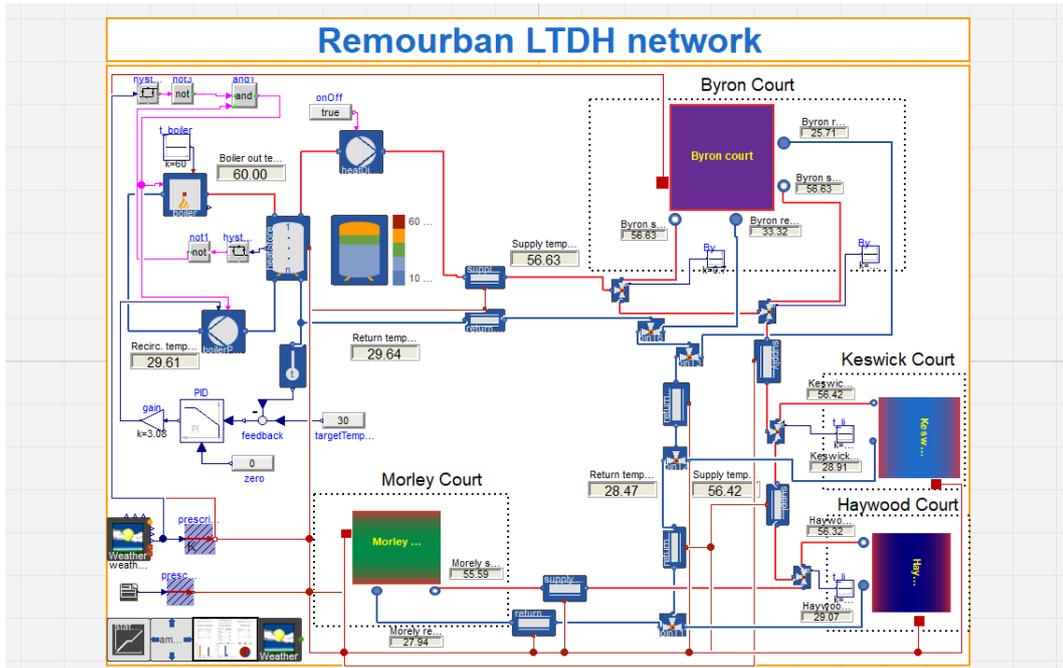
Figure 4.18: Results from the Iteration 1 simulation depicting the performance of optimised LTDH network

Table 4.4: Overview of flow-rates and other parameters in Iteration 1 of LTDH network

Location	Pipe size (mm)	Pipe length (m)	Flow-rate (kg/s)	Serving
Main plant room	90	71	2.4	Entire network
Byron court under-croft	90	5	0.8	31 flats
Keswick court under-croft	75	56.5	0.36	14 flats
Haywood court under-croft	75	22	0.75	29 flats
Morley court under-croft	76	63	0.49	20 flats
Circulation pump	-	-	3.08	Plant room

creases significantly. In comparison to the baseline model, this decrease in energy demand and return water temperature is attributed to the reduction in flow-rates in the network. These reduced flow-rates considered in the LTDH network for different buildings are given in Table 4.4.

The temperature of supply water degrades slightly and it enters the network at around 56°C which remains consistent for all the buildings. Moreover, the return water temperature is reduced to 29°C in the plant room. This is due to the installation of PID controller which regulates flow-rate of the cir-

culation pump and maintains water temperature in the thermal heat storage. The upper section of thermal heat storage has the water temperature of 50^oC and drops in the lowest sections having temperature of 10^oC. The comparison between supply and return water temperature for different buildings in the network and indoor air temperature in flats is shown in Figure 4.19.

Figure 4.19: Results from iteration 1 simulation. Comparison between supply and return water temperature for different buildings in the network and indoor air temperature in flats

Figure 4.20: Energy statistics results from iteration 1 low temperature district heating network modelled in Dymola software with optimum parameters.

The simulation predicts that the overall CO_2 emissions are reduced to around 24,124 kg for the month of January. The energy statistics results in

Figure 4.20 show that the percentage share of electricity and gas remains at 31% and 69%, but in comparison to the baseline model, the total energy consumption is reduced to just 13,820 KWh and 112,157 KWh, respectively. The district heating network generates approximately 98,789 KWh and consumes 95,908 KWh of heat. The difference between the generation and consumption is reduced to around approximately 2,881 KWh.

4.4.2.3 Iteration 2

The iteration 2 investigates operation of the network with integration of solar thermal power plant as a heat-source. The flow-rate of inlet water to the solar thermal plant is controlled using a sophisticated control loop, which regulates the pump's speed with respect to the outlet temperature and the water temperature inside thermal heat storage. The solar thermal plant uses water with 47% propylene glycol as a working medium and transfers the heat into the thermal heat storage with one heating coil and shown in Figure 4.21.

The simulation results show that the Δt of the network remains very low even with the optimisation of flow-rates. The maximum achievable outlet temperature from the solar thermal plant is 9°C and owing to the distribution heat-losses in thermal heat storage the supply water temperature in the network is limited to 8.86^{o}C. Moreover, the average return water temperature in the network is calculated as 8.60^{o}C. The comparison between supply and return water temperature for different buildings in the network and indoor air temperature in flats is shown in Figure 4.22.

In comparison to the baseline model, the heat from solar thermal plants is not enough to fulfill the space heat demand and not feasible to use as a main heat source especially during winter months (January). The solar thermal
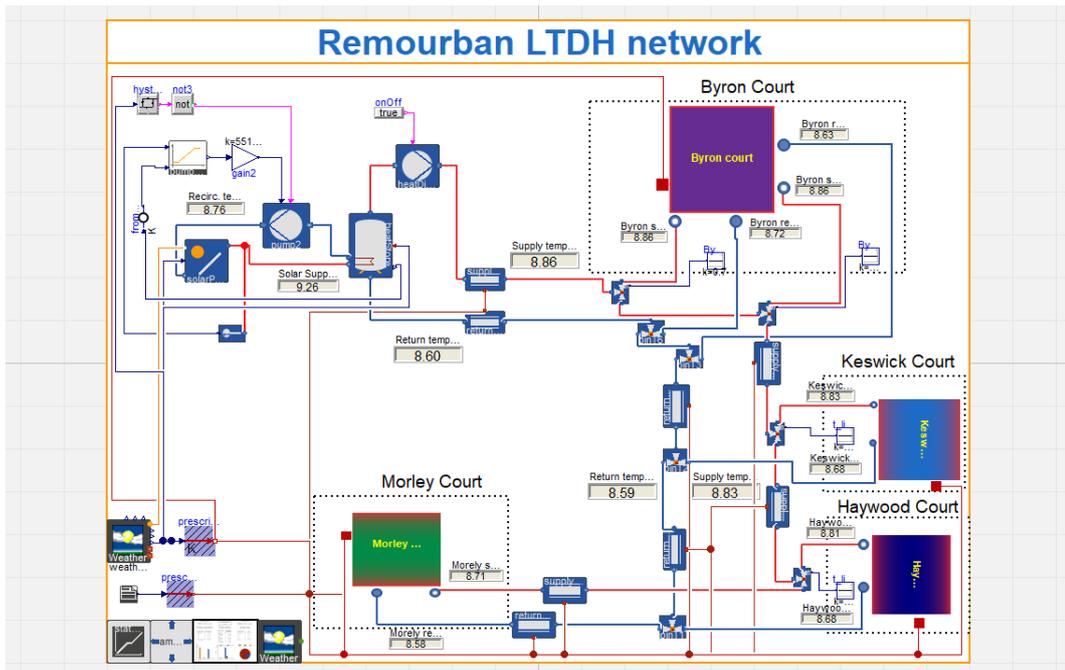
Figure 4.21: Results from the Iteration 2 simulation depicting the network's performance with integration of solar thermal power plant as a heat-source to the LTDH network
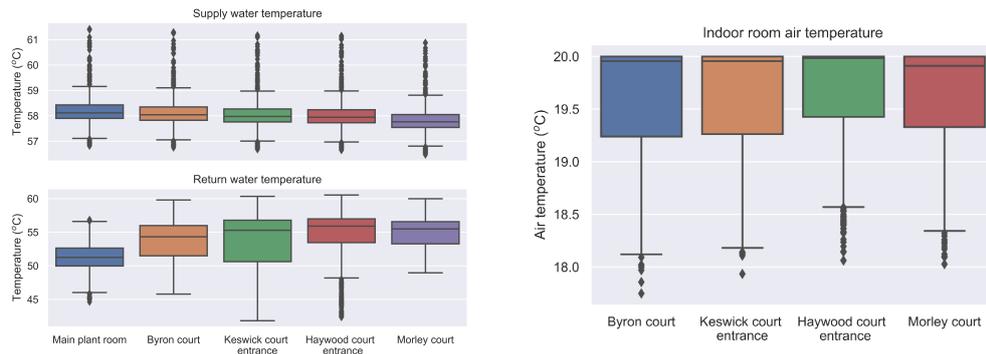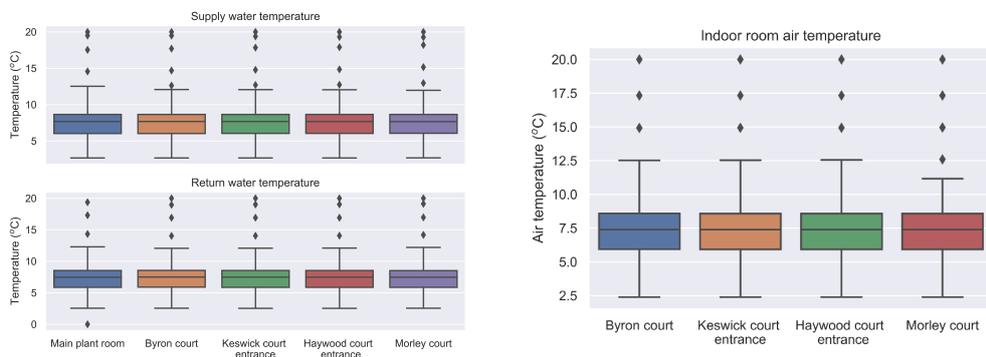

Figure 4.22: Results from iteration 2 simulation. Comparison between supply and return water temperature for different buildings in the network and indoor air temperature in flats

plants operate at low flow-rates and therefore can be used as a secondary heat source for underfloor heating purposes.

The simulation predicts the overall CO_2 emissions are reduced to just 5,460 kg, but at the expense of limited amount of heat generation. The energy statistics results in Figure 4.23 shows that the percentage share of electricity
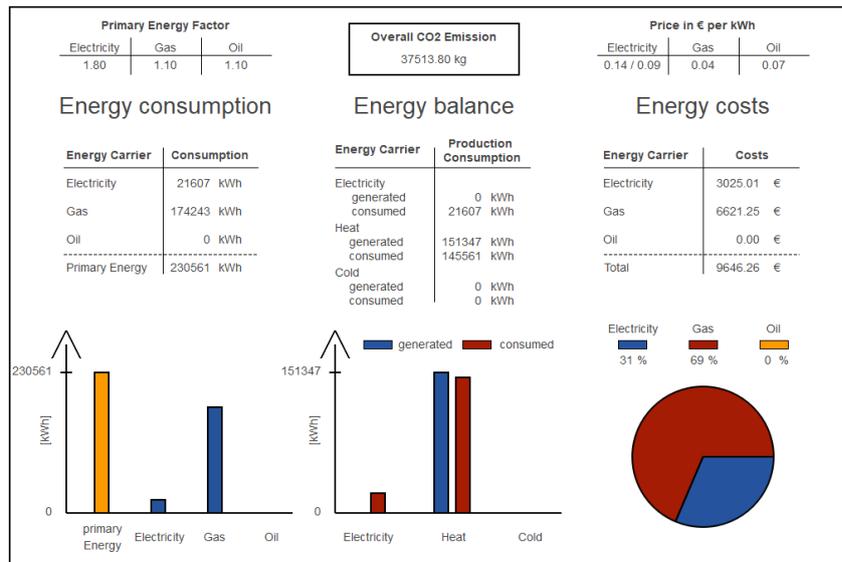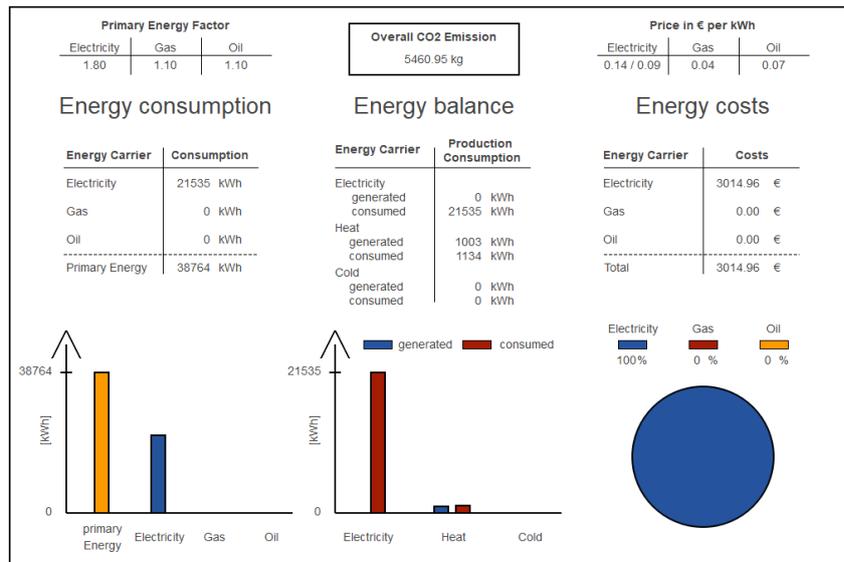
Figure 4.23: Energy statistics results from iteration 2 low temperature district heating network modelled in Dymola software with integration of solar thermal power plant.

is 100% and the network consumes approximately 21,535 KWh of electricity and 1,134 KWh of heat for the month of January.

4.4.2.4 Iteration 3

The iteration 3 investigates the optimised LTDH network (from iteration 1) with the integration of solar thermal power plant. In other words, it evaluates the operation and energy consumption once both iteration 1 and iteration 2 networks are combined together. This provides an opportunity to analyse a multi-vector network supplied with heat from multiple energy sources.

The simulation results in Figure 4.24 shows that the network uses thermal heat storage with two heating coils as the network uses two mediums i.e (incompressible water and water with 47% propylene glycol in the solar thermal plant thermal plant). The solar thermal plant preheats the return water from the network before feeding it into the boiler. The network also contains two separate control loops to regulate the return water temperature in the network,
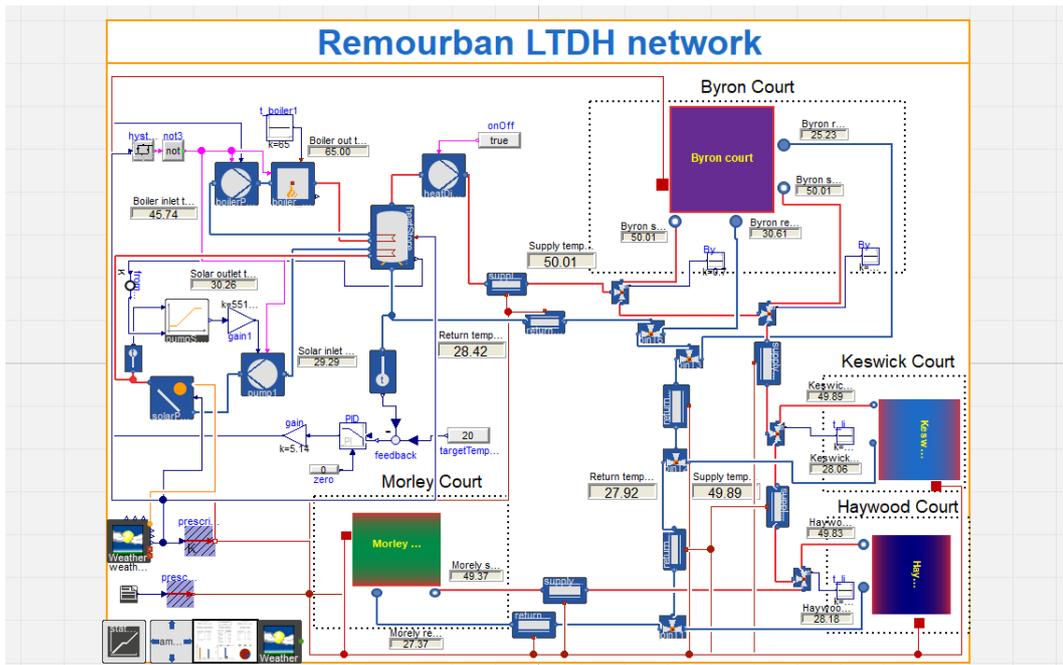
Figure 4.24: Results from the Iteration 3 simulation depicting performance of optimised LTDH network (from iteration 1) with the integration of solar thermal power plant

i.e. a PID (proportional integral derivative) based feedback controller on the circulation pump before the boiler and the pumping speed control loop on the solar thermal plant.

It is observed in Figure 4.24 that even the output water temperature from the boiler and solar thermal plant is around 65^oC and 30.26^oC, the supply water temperature decreases to 50^oC. This is due to the transmission heat losses in thermal heat storage and LTDH network. The Δt of network is maintained to designed parameters using a PID controller (similar to iteration 1). This regulates the heat output from the boiler with respect to the return water temperature. The comparison between supply and return water temperature for different buildings in the network and indoor air temperature in flats as shown in Figure 4.25.

The simulation predicts the overall CO_2 emissions are increased to 35,946
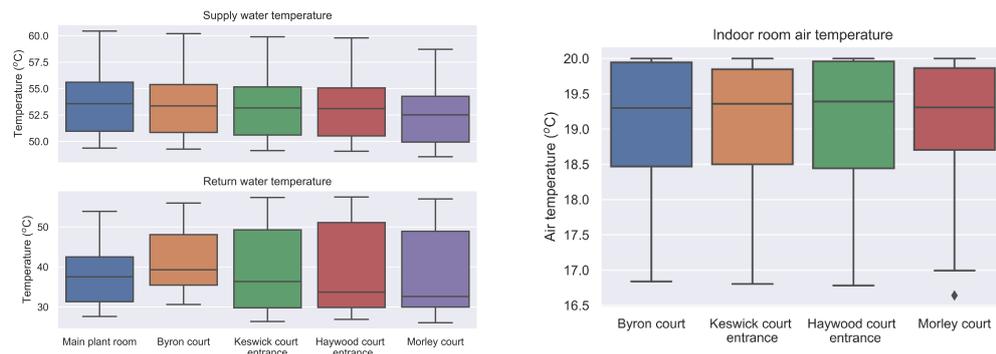
Figure 4.25: Results from iteration 3 simulation. Comparison between supply and return water temperature for different buildings in the network and indoor air temperature in flats
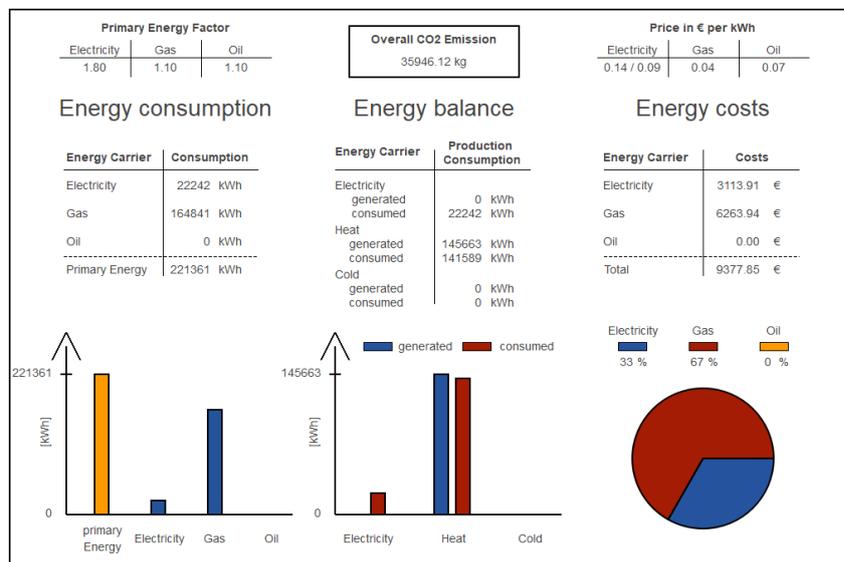
Figure 4.26: Energy statistics results from iteration 3 low temperature district heating network modelled in Dymola software.

kg for the month of January. The energy statistics results in Figure 4.26 show that the percentage share of electricity and gas is increased to 33% and 67%, and the electricity and heat consumed by the network is approximately 13,820 KWh and 112,157 KWh, respectively.

4.4.2.5 Iteration 4

The iteration 4 is another configuration of iteration 3 where the solar thermal plant pre-heats the water inside the thermal heat storage, before feeding it

into the gas-boiler which is directly connected to the LTDH network. The parameters in iteration 4 are same as in iteration 3, and evaluates the energy consumption with reliable heat source connected directly to the network.
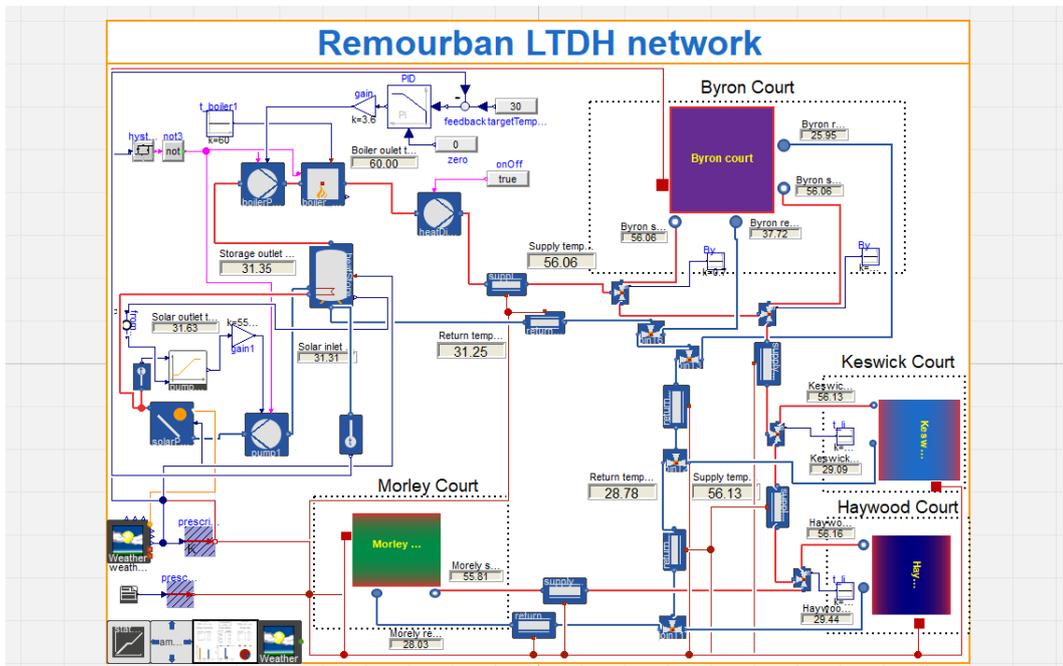
Figure 4.27: Results from the Iteration 4 simulation, where the solar thermal plant pre-heats the water inside the thermal heat storage, before feeding it into the gas-boiler directly connected to the LTDH network.

The simulation results in Figure 4.27 show that the solar thermal plant pre-heats the supply water and increases its temperature to 31.35^{o}C. This is then heated in boiler for upto 60°C and then feed to the LTDH network, where its temperature degrades to 56°C outside the Byron court due to transmission heat-losses. The Δt of network remains inline with the designed parameters as the return water temperature of network is maintained from the PID controller. The comparison between supply and return water temperature for different buildings in the network and indoor air temperature in flats is shown in Figure 4.28.
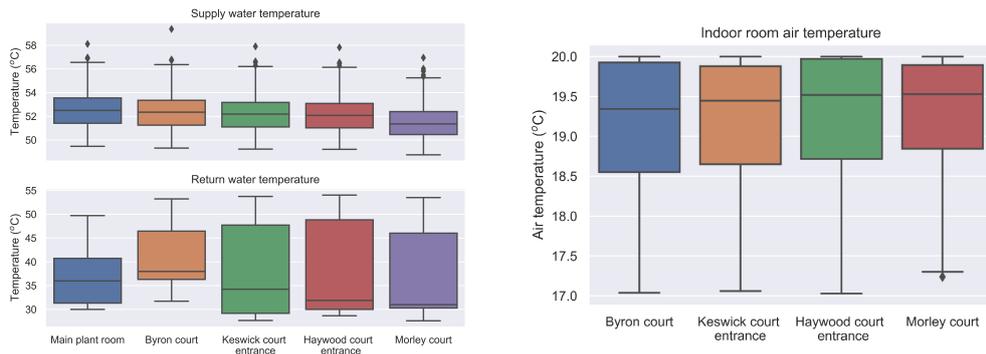
The simulation predicts the overall CO_2 emissions are increased to 36,004
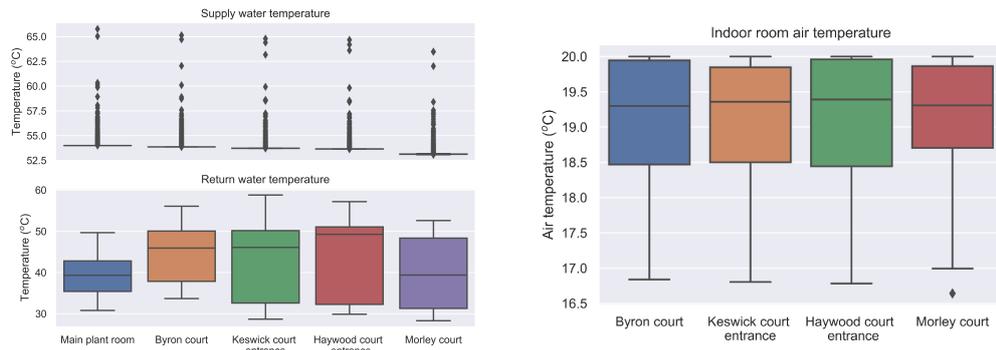
Figure 4.28: Results from iteration 4 simulation. Comparison between supply and return water temperature for different buildings in the network and indoor air temperature in flats
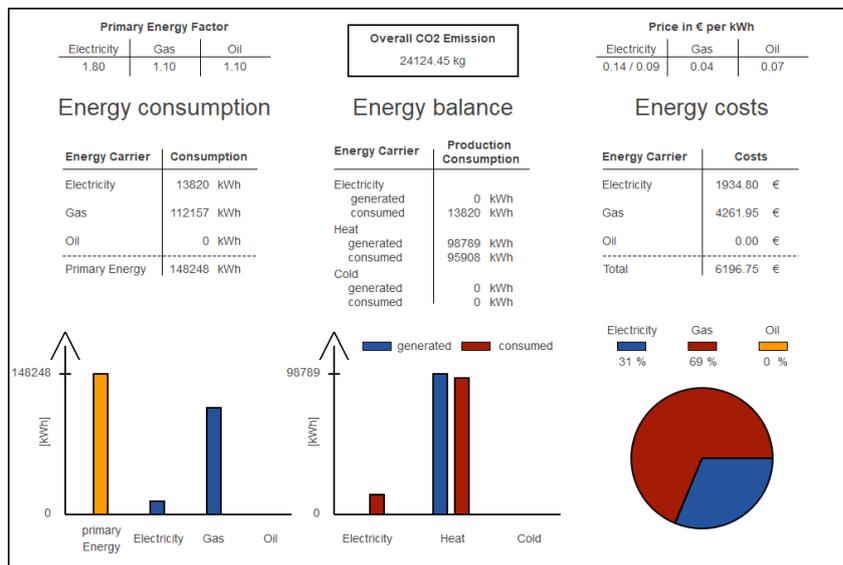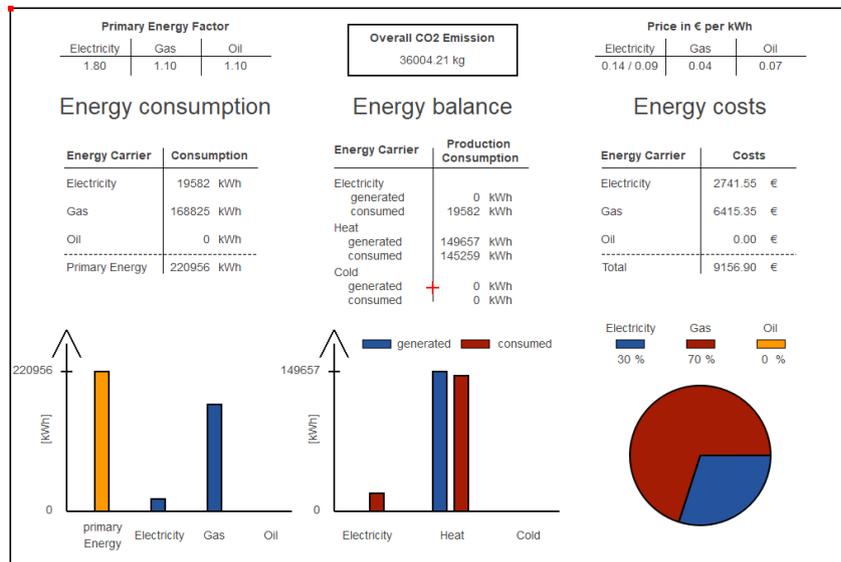

Figure 4.29: Energy statistics results from iteration 4 low temperature district heating network modelled in Dymola software.

kg for the month of January. The energy statistics results in Figure 4.29 show that the percentage share of electricity and gas remains at 30% and 70%, and the electricity and heat consumed by the network is approximately 19,582 KWh and 145,259 KWh, respectively.

4.4.3   Validation of results using real monitored data

The validation of developed LTDH model is necessary for the credibility of simulation results. Therefore, the hourly monitored data for the month of January is obtained from the REMOURBAN LTDH network project. The data is from 4$^{\text{th}}$ to 28$^{\text{th}}$ January and used for the model validation. It should be noted that this project is still operational in the testing phase and several parameters are changing on daily basis. Thus, acquiring meaningful and consistent hourly monitored data has been a real challenge. Moreover, a few flats from the Byron court had not been connected during this time.

A boxplot is a statistical method for data representation and informs whether the data is symmetrical, grouped, skewed or contains outliers. The data distribution is based on five parameters; minimum, first quartile (Q1), median, third quartile (Q3), and maximum and shown in Figure 4.30.
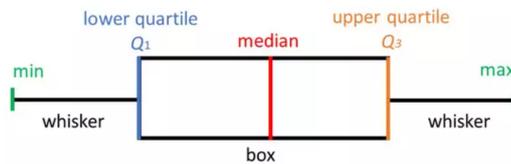
Figure 4.30: Graphical representation of the boxplot.

In Remourban project several meters are installed both in the network and inside the flats. Since the aim is to perform validation and operation of LTDH network, therefore the data from meters installed only in the network is used (i.e. main plant room, Byron, Keswick, Haywood and Morley courts). The Figure 4.31 shows the box plot for comparison of different parameters in the monitored from the REMOURBAN LTDH network.

On analysing the monitored data in Figure 4.31, it is evident that the supply water temperature remains relatively similar for different meters in the
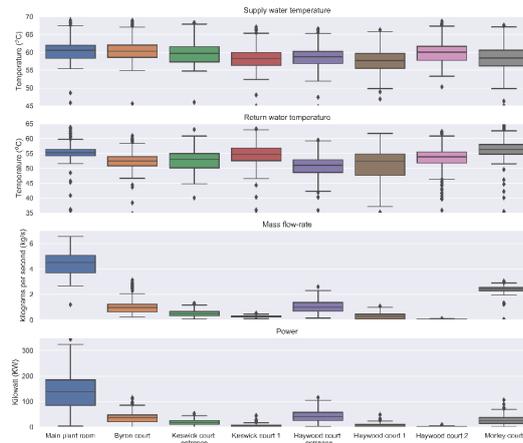
Figure 4.31: Boxplot representing comparison of different parameters in the monitored from the REMOURBAN LTDH network
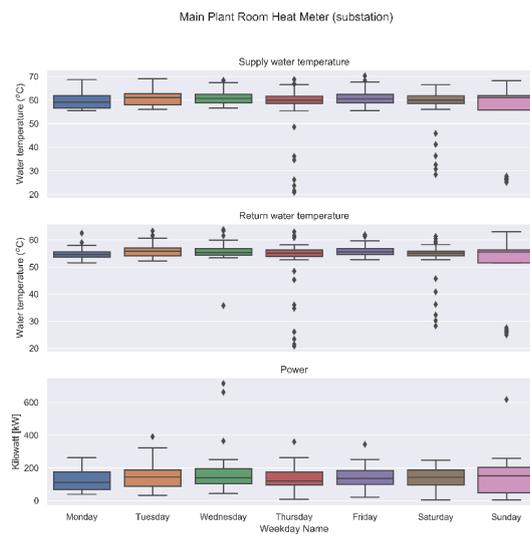
Figure 4.32: Boxplot representing days of the week comparison in monitored data from the main plant room

network compared to the return water temperature which varies throughout the network. The median of supply and return water temperature is found to be between 58-60°C and 52–56°C, respectively. Moreover, the return water temperature data contains several outliers. While analysing the flow-rate and power, it is observed that surprisingly the Haywood and Morley court receives reasonably high flow-rate but less power is consumed. Specifically for the

Morley court which receives water with high flow-rate and returns this to the plant room through the bypass valve in the network. This explains that the network needs to be balanced either by reducing flow-rate from the main plant room or by optimising control valves installed in the network.
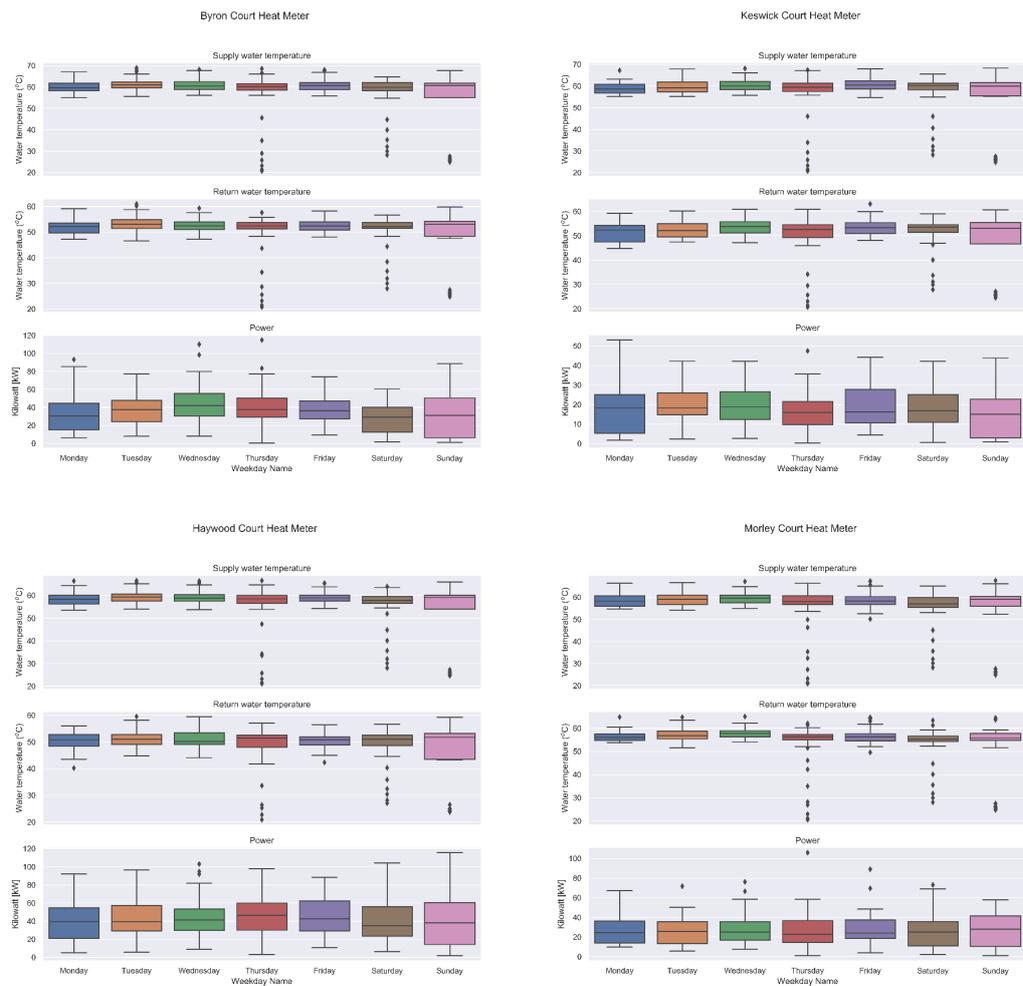
Figure 4.33: Boxplot representing days of the week comparison in monitored data from Byron, Keswick, Haywood and Morley courts

The weekday analysis in Figure 4.32 shows that both supply and return water temperature from the main plant room is found to remain constant throughout the week except on Sunday, where the trend shows deviation.

Likewise, the court data in Figure 4.33 shows similar trend and deviation is prominent on Sunday.
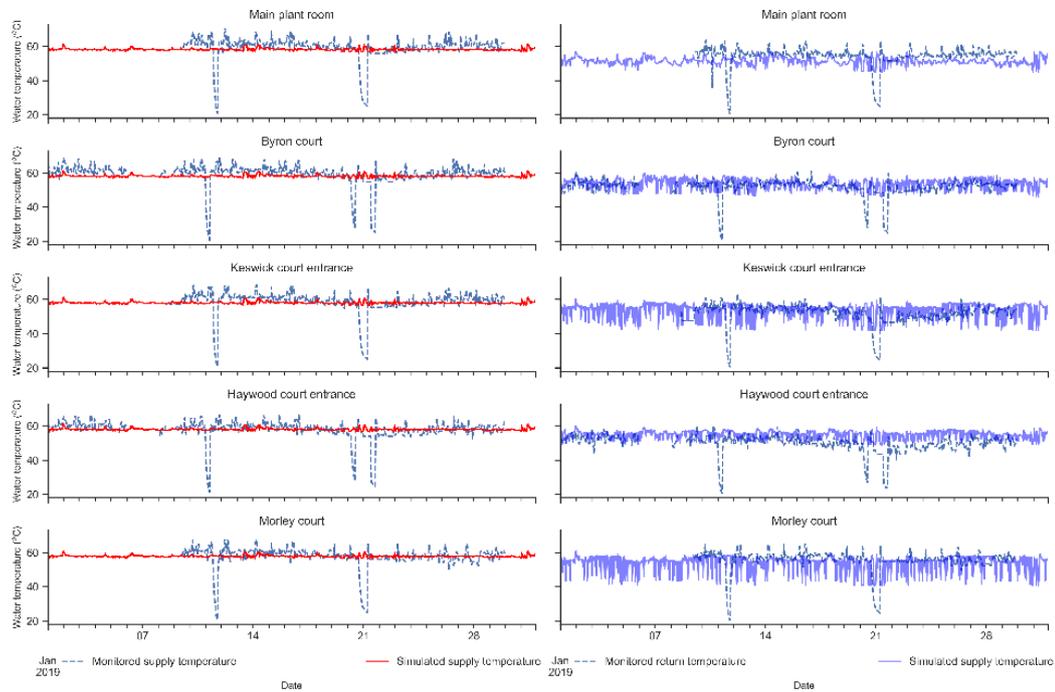
Figure 4.34: Validation of the baseline model by comparison with monitored from 4th to 28th January. The left panel compares supply water temperature whereas the right panel compares return water temperature.

The above monitored data is used for validation of the LTDH model simulated in Dymola model. The Figure 4.34 plots the simulated supply and return water temperature calculated from the baseline LTDH model developed in Dymola with the hourly monitored data from the REMOURBAN project. It is observed that both simulated supply and return water temperature matches perfectly with the monitored data. There few mismatches in the simulated return water temperature of Keswick and Morley court which can be explained due to the thermal performance of existing radiators installed inside the flats. This validates the developed Dymola model and simulation results (discussed in previous Section) as well as highlights the significance on flow-rates optimisation for the energy efficienct operation of LTDH network.

4.5 Discussion and recommendations

In the UK, the traditional district heating networks are designed to operate on high supply water temperature with constant flow-rates. This common practice impacts the energy efficiency due to high transmission heat-losses in the network. However, increasing the delta t (Δt) and reducing return water temperature improves the energy efficiency of the network.

The results from the first part of the hydraulic modelling analysis suggest that the LTDH network should be designed with variable speed pumping, and supply water temperature should be kept constant from the plant room. This leads to the lowest energy consumption in the network as elaborated in Section 4.4.1.1 (scenario 3). On the other hand, the constant speed pumping is a predominant strategy for the district heating networks in the UK. The heat-losses in the existing district heating network operating on constant flow-rate can be reduced significantly once the weather compensation valves are installed for the regulation of supply water temperatures according to the outside temperature. Additionally, reduction in supply water temperature enables the use of heat from renewable energy resources and other low-grade waste heat sources.

The monitored supply and return water temperature data of an existing boiler based space-heating system inside the flat suggests that their conversion to LTDH is technically feasible as a supply temperature are already below 60°C. Owing to very high return temperature, these systems need to be hydraulically balanced first. This is because hydraulic imbalance in existing boiler based space-heating systems causes enormous heat losses. The regulation of return temperature is vital for achieving high Δt and energy efficiency.

This can be easily achieved by installation of TRVs on low temperature hydronic radiators. The function of TRV's is to control the flow through each radiator with respect to the room temperature and make sure that it never exceeds the maximum limit.

These results suggest that alongside the importance of flow-rates and pipe sizing in the LTDH network, the focus should be given to the installation of thermostatic radiator valves on radiators. In the case of high-rise buildings, the TRVs alone are unable to maintain the hydraulic balance in existing space-heating system in buildings. Therefore, either pressure independent thermostatic radiator valves should be used or the pre-setting function of TRV's on radiator with balancing valves and differential pressure controller should be employed (Zhang et al. 2016). The pre-setting function controls the amount of water passing through the radiator.

This is a common perception that the entire heating network is prone to a low Δt. However, the monitored data shows that it is hydraulic imbalance issue inside the buildings which leads to high return water temperatures in the district heating network. The hydraulic imbalance impacts the Δt and is necessary to maintain high Δt across the network. Its imbalance makes the district heating networks to operate at high supply temperatures with high flow-rates and large pumping capacities are needed in the network. This also leads to lower efficiency, low Δt, high return water temperature and uneven distribution of heat in the district heating network. The district heating network is hydraulically balanced when the flow-rate and the Δt are in accordance to the consumers demand (Boysen & Thorsen 2007).

Radiator inlet water connection

The hydraulic imbalance and low delta t (Δt) issue inside the buildings can also be attributed to the conventional radiator connection practice still being used. In the UK, the hydronic radiators are frequently connected in the BBOE (bottom, bottom, opposite end) configuration as per the BS-3521 standard (BS EN 442-1 2014) which leads to high flow-rate and low delta t (Δt) across the heating network. This standard provides guideline for the hydronic radiator connections in the conventional heating network with high supply water temperature. Though, the recent focus has been on the reduction supply water temperatures and improving Δt of the heating networks, the hydronic radiators are still being installed in the same configuration.
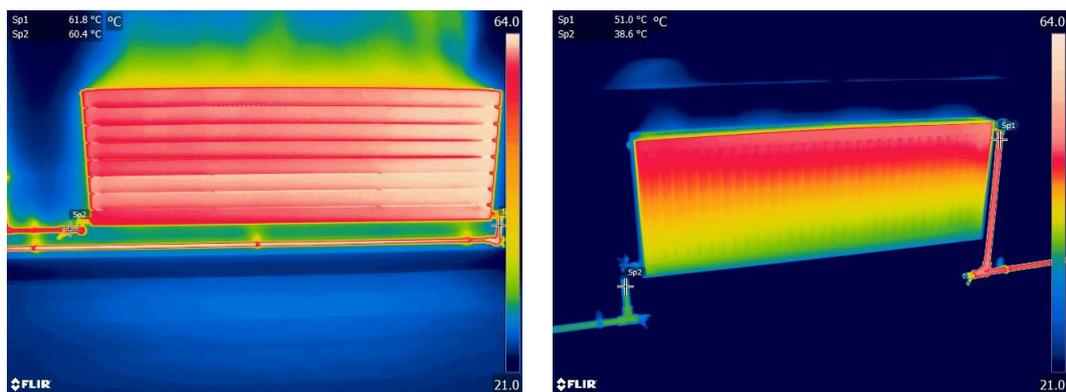
Figure 4.35: Thermo-graphic images show comparison between different radiator connections from the same building. Left figure show results for the radiator connected with bottom, bottom, opposite end (BBOE) connection. Right figure show results for the radiator connected with top, bottom, opposite end (TBOE) configuration.

Therefore, two radiators with different connection configurations (BBOE - bottom, bottom, opposite end and TBOE - top, bottom opposite end), from the same building, have been taken as a case study. As shown in thermographic images in Figure 4.35 a considerable difference is observed in the actual operation of BBOE configuration radiator in the buildings compared to

experimental results in (Calisir et al. 2017). In BBOE configuration, the bottom end of radiator has higher temperature than the top and delta t (Δt) across the radiator is found to be even less than 1%. The installation of TRVs is found to be ineffective in this type of configuration, as it can not improve delta t (Δt) across the radiator and operates like a shut on/off valve. This can be understood due to high amount of scaling and fouling effect due to radiator connections in BBOE configuration and causes most of the heat to be lost in pipes or outside the building. On the other hand, when radiators are connected in the TBOE configuration the delta t (Δ t) is found to be around 12. These findings are similar to discussed in (McIntyre 1986, Ward 1991) and the heat-losses are increased by almost 30% with radiators connected with BBOE than the TBOE configuration.

Recommendations

The efficiency of LTDH heating networks depends on the correct design of network parameters and optimisation of space heating systems inside the buildings. Once the heating network is imbalanced, it impacts the Δt and energy losses. The network operates at high supply temperatures with high flow-rates and large pumping capacities. The network can also be hydraulic imbalance because of the erroneous heat demand estimations, wrong flow-rates calculations, over-sized and unnecessary valves, pumps or even flow-limiters in the main district heating network. It can also be due to lack of balancing, flow-limiters and other flow control equipment (Boysen & Thorsen 2007). In summary, the heating network needs to be balanced for energy efficient operation at design parameters.

It is recommended that following measures will help in increasing efficiency of the heating networks in the UK.

1. District heaing pipes must not be over-sized deliberately for reducing flow-rate in the network. This is because it leads to higher heat-losses, head-loss, pumping power and energy consumption.

2. Installation of TRV on hydronic radiators and pressure independent thermostatic radiator valves (TRVs) in the case of high rise buildings.

3. Hydronic radiators installation in TBOE (top bottom opposite end) configuration against the existing practice of BBOE (bottom bottom opposite end) configuration.

Furthermore, the correct space heating demand estimation and guidelines for the domestic hot water demand calculation needs to be updated in the UK. If above discussed recommendations are implemented in the LTDH networks, then considerable energy consumption savings along with positive environmental impact are expected to achieve.

4.6   Summary

The results from this chapter show that in order to maintain high delta t (Δt), the regulation of return water temperature with respect to the outside temperature is more important compared to the supply water temperature. Moreover, the hydraulic modelling is performed in two parts, and the results from the first part of LTDH modelling analysis suggest that the LTDH network should be designed with variable speed pumping, and supply water temperature should be kept constant from the plant room. This leads to the lowest energy consumption in the network. On the other hand, the heat-losses in the the existing

district heating networks, which predominantly operate at constant flow-rate in the UK, can be reduced by installation of weather compensation valves (regulation of supply water temperature according to the outside temperature). Furthermore, in the second part, the energy efficiency and Δt from all iterations are compared against the baseline model and the results show that both energy efficiency and Δt in iteration 1 is higher than the baseline model. For all other iterations Δt appears to be less than the baseline model. These data suggests that the LTDH network in REMOURBAN project can be improved by reducing flow-rates both in the network and circulation pump inside the plant room (i.e. iteration 1).

Chapter 5

Cost minimised decarbonised district heating network

Overview

In this chapter the cost minimised design of the de-carbonised district heat network is discussed. The city of Aarhus, Denmark is taken as a case study, where its hourly heat demand and heat generation cost data is used for the analysis. The heat network is coupled to the fully renewable based electrical grid and excess generation is used by the heat-pump and thermal heat storage. The optimum solution is suggested for the entire network with least amount of backup generation capacity, thermal heat storage capacity, natural gas boiler capacity and levelised cost of energy. Finally, the economic feasibility for the optimum network design is suggested and compared with the actual heat generation costs for the existing fossil-fuel based district heat network. The rationale behind using Aarhus district heating network instead of Nottingham was due to the availability of high quality heat demand and cost time-series data which was not available for Nottingham at that time. Additionally, the

Aarhus has similar population as of Nottingham with 300,000 inhabitants and it's district heating network is considered as among five most energy efficient and modern network in the world.

5.1   Background

With increased concerns of climate change and global trend to sustainability, the focus has shifted towards mitigation of carbon emissions with low cost energy systems and this has been the main subject of Paris (COP21) agreement (IEA & IRENA 2017). Not long ago, the contribution of CO_2 from US, European Union (EU-28), China and India was reported around 61% of global emissions (Olivier et al. 2015). In response to this, the EU is willing to take initiative to control greenhouse gas emissions mainly with the transformation of heating sector by the year 2050 (McKinsey et al. 2010). Moreover, several other studies have also emphasised the de-carbonisation of heating sector instead of electrical grid for achieving carbon neutral future (Lund et al. 2015, 2016), and a roadmap has been provided for obtaining energy efficiency in the heating sector in the EU (Hansen et al. 2016).

To support this, few studies have identified the spatial distribution of heat demand scattered around Europe. This has resulted into detailed analysis of the district heating network and heat consumption by the buildings as well as the availability of surplus heat resources in the EU countries (Persson et al. 2014, Möller & Nielsen 2014). Moreover, the future of district heating in Europe is found promising (Connolly et al. 2014), and the potential savings from using heat pumps for the conversion of excess energy into heat to feed into district heating network has also been explored (Ashfaq et al. 2017).This leads to a consideration that the future expansion of district heating networks

is only economic feasible with the addition of waste heat sources (Grundahl et al. 2016, Dominković et al. 2017). The geographical distribution of heat demand in Denmark and renewable energy power generation time-series is shown in Figure 5.1.
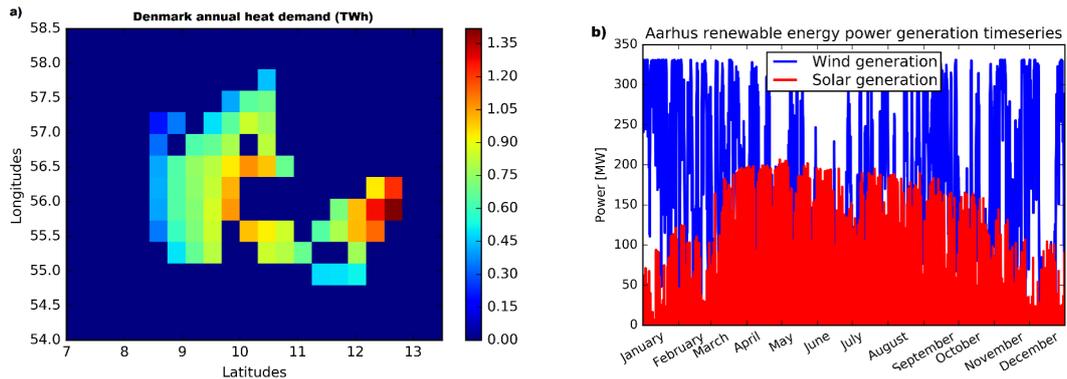
Figure 5.1: Figure (a) shows the heat demand distribution in the Denmark at spatial resolution of 40 x 40 km^2 for the year (2011). Figure (b) shows the renewable energy time series for the city of Aarhus, Denmark. These results are adapted from (Ashfaq & Ianakiev 2018*a*).

The existing district heating networks in Europe are mostly supplied by the combined heat and power plants (CHP) and fuel based boilers. If these have to be replaced then their holistic impact need to be addressed by taking into account technical and economic aspects. There are studies on the decarbonisation of EU's electrical grid using 100% renewable energy (Connolly et al. 2016, Rodríguez et al. 2014) and recommendations for the interconnection of multiple energy sources to achieve least cost solution. (Mathiesen et al. 2015, Thellufsen & Lund 2017, 2015). Though, the techno-economic feasibility for the electrification of heating sector in US is studied by coupling both networks together (electrical grid and heat networks) (Pensini et al. 2014), it needs to be explored in the EU.

Therefore, this chapter outlines the energy saving benefits with the synergy of both electrical and heating networks and addresses following questions:

- Is the integration of electric grid and heating sector technically and economically feasible?

- How intra-day demand profiles effects the electrified district heating network?

- Is a comparison between electrified and the existing district heating network, with variation in renewable energy penetration (γ) and wind/solar mix (α^W), cost-effective?

- What is the viability of heating network decarbonisation using electrification?

5.2   Energy system modelling

This analysis considers a futuristic highly renewable energy-based network, where the wind and solar (PV) generation are taken as renewable energy sources and other sources are assumed as instantaneous backup power generation sources (conventional energy generation, hydro-electric storage lakes, biomass). The modelling has been divided into two parts: technical modelling and the economic modelling for the analysis. The modelling is graphically represented from the flowchart in Figure 5.2.

5.2.1   Electrical grid modelling

The electrical grid is modelled as Direct Current (DC) grid with unconstrained power sources from wind and solar (PV) power generation at hourly intervals i.e.*t*. The transmission and distribution losses are not considered, which is
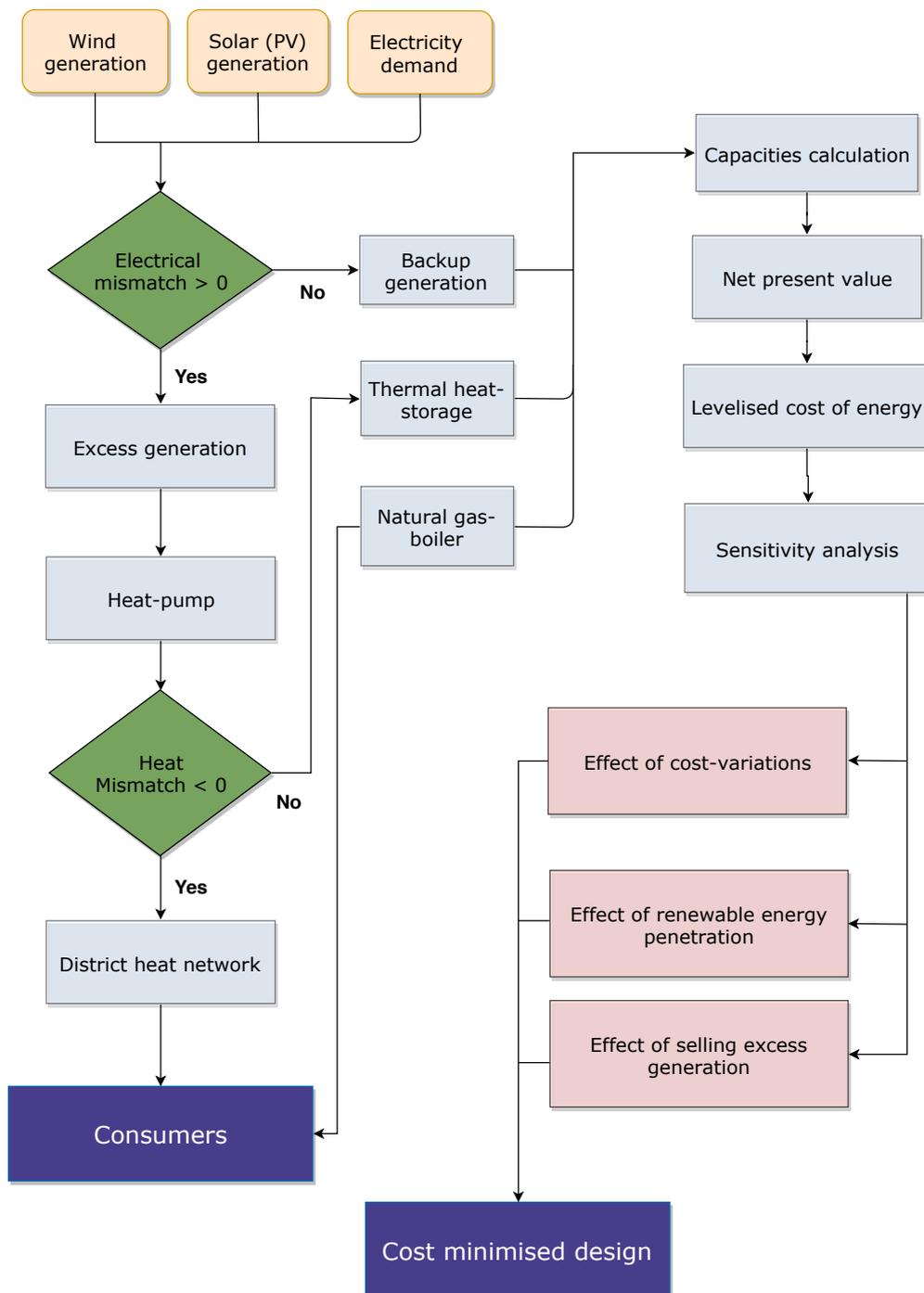
Figure 5.2: The methodology of cost minimised design of a decarbonised district heating network analysis

an acceptable assumption and helpful to identify barriers in the DC electrical grid. Therefore, 8 years (2000-2007) of Denmark's renewable power generation (wind and solar) and electrical demand data is obtained from the ISET dataset (Bofinger et al. November, 2008). First, Aarhus electrical demand hourly time-series is generated by scaling Denmark's electrical demand time series to match the annual electricity demand (1,603 GWh) given in (Rasmussen 2012). Then, the hourly wind, solar (PV) generation and electrical demand time-series are normalised to their average value, represented by symbol $\langle . \rangle$. This methodology of electrical grid modelling is adapted from (Becker, Rodriguez, Andresen, Schramm & Greiner 2014, Rodríguez et al. 2014, Schlachtberger et al. 2016).

The wind/solar mix (α^W) determines the amount of wind and solar PV generation (G^W and G^S) in the total power generation (G). The renewable energy penetration factor (γ) defines the total renewable power generation (wind and solar) to the electrical demand, where the (γ) value of 1 and 1.5 represents the renewable power generation is equal and 50% more than the electrical demand, respectively. This modelling is shown by the equations below:

$$G(t) = G^W(t) + G^S(t) \tag{5.1}$$

$$\alpha^W = \frac{\langle G^W \rangle}{\langle G \rangle} \tag{5.2}$$

$$\gamma = \frac{\langle G \rangle}{\langle L_E \rangle} \tag{5.3}$$

$$G(t) = \gamma . \langle L_E \rangle [\alpha^W . G^W(t) + (1 - \alpha^W) G^S(t)] \tag{5.4}$$

The excess generation (P^{ex}) and backup generation (B) is calculated from the difference between power generation and electrical demand, known as the electrical mismatch (Δ_E). The hourly excess generation and backup generation time-series are computed from the amount of power generation greater than the electricity demand and vice versa i.e. ($|\Delta_E|_+$ and $|\Delta_E|_-$).

$$\Delta_E(t) = G(t) - L_E(t) \tag{5.5}$$

$$P^{ex}(t) = |\Delta_E(t)|_+ \tag{5.6}$$

$$B(t) = |\Delta_E(t)|_- \tag{5.7}$$

In this modelling, the excess generation time-series is directly used by the heat-pump for the heating and required backup storage capacity ($\mathcal{K}^B$) is calculated from the average of backup storage time-series. This capacity estimation method is adapted from the studies in (Becker, Frew, Andresen, Zeyer, Schramm, Greiner & Jacobson 2014, Becker et al. 2015, Dahl 2015).

$$\mathcal{K}^B = \left\langle \sum_{t=1}^{8760} B(t) \right\rangle \tag{5.8}$$

5.2.2 Heat network modelling

The heat demand (L_H) is comprised of space heating demand (Q_{sh}) and domestic hot water demand (Q_{dhw}). The heat network is coupled with the electrical grid and the excess electricity is used to fulfil the heat demand. Therefore, the calculation of heat mismatch is vital to this research. The heat mismatch
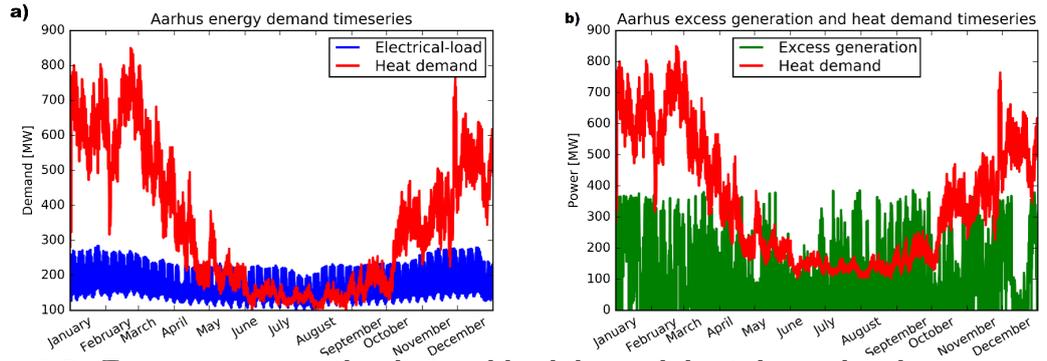
Figure 5.3: Figures compares the electrical load demand, heat demand and excess renewable energy generation time-series for the city of Aarhus, Denmark. These results are adapted from (Ashfaq & Ianakiev 2018*a*).

(Δ_H) is the difference between the heat demand (L_H), and the product of excess generation (P^{ex}) with heat pump coefficient of performance (COP). The hourly heat demand and generation cost time-series data (2011-2015) is collected from the Aarhus Municipality and the annual figures of heat demand are compared with those given in (Rasmussen 2012) for validation.

$$L_H(t) = Q_{sh}(t) + Q_{dhw}(t) \tag{5.9}$$

$$\Delta_H(t) = P^{ex}(t) * \text{COP} - L_H(t) \tag{5.10}$$

Similar to Section 5.2.1, the excess heat energy and heat deficit energy in the coupled network is calculated from the hours with positive heat mismatch and negative heat mismatch ($|\Delta_H|_+$, $|\Delta_H|_-$), respectively. The thermal heat storage (E) can be used to store this excess heat energy, and natural gas-boilers (N) to cover the heat deficit energy

$$E(t) = |\Delta_H(t)|_+ \tag{5.11}$$

$$N(t) = |\Delta_H(t)|_-\tag{5.12}$$

The 99% quantiles of the thermal heat storage energy time-series distribution $p(E)$ is used to calculate the required thermal heat storage capacity ($\mathcal{K}^E$). This is done to mitigate the effects of severe conditions. On the other hand, the average of heat deficit energy time-series is used to calculate the required natural gas-boiler capacity ($\mathcal{K}^N$) . These required backup capacities depend upon the share of wind and solar (PV) generation and the amount of renewable energy generation in the network. This method for capacity estimation is used in studies (Rodriguez et al. 2015, Rodríguez et al. 2014).

$$q = \int_0^E p(E)dE\tag{5.13}$$

$$\mathcal{K}^E = E^{99\%}\tag{5.14}$$

$$\mathcal{K}^N = \left\langle \sum_{t=1}^{8760} N(t) \right\rangle\tag{5.15}$$

Heat-pump coupling

The heat pump is used for the conversion of excess renewable energy generation into heat for the district heating network. The working principle of the heat-pump based coupling is divided into three following steps as shown in the Figure 5.4.

1. The excess renewable energy generation is converted into heat by the heat pump and supplied to the district heating network with consumers.

2. In case of positive heat-mismatch, the excess amount of heat is stored into
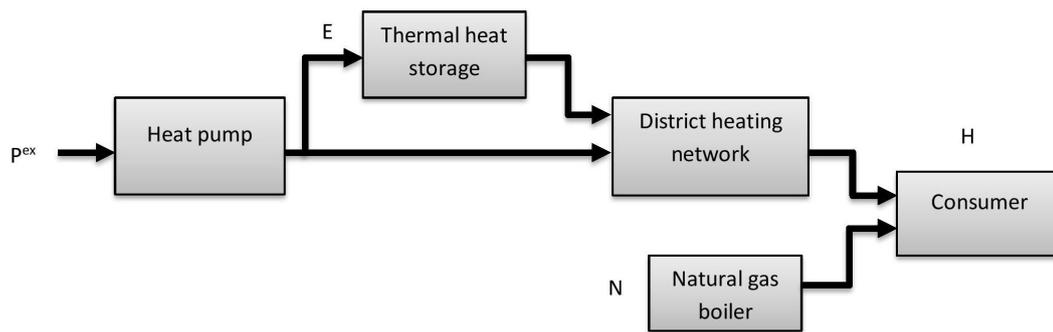
Figure 5.4: Process flow diagram of heat-pump based coupling

the centralised thermal heat storage and utilised by the district heating network upon its need by the consumers.

3. Finally, the on-site natural gas-boilers are available for the consumers as backup, when there is no thermal heat energy in the network.

The heat coupling assumes several efficiencies such as; heat pump with COP of 3, thermal heat storage, district heating network and natural gas boiler with the efficiency of 90%, 90% and 100%, respectively. These efficiencies have been considered in different studies (Pensini et al. 2014), (Ashfaq et al. 2017)

The following figures compare the electrical load demand, heat demand and excess renewable energy generation time-series for the city of Aarhus, Denmark.

5.2.3   Coupled network analysis

The hourly analysis is vital for the optimum demand forecasting and operation of the coupled energy network. Therefore, the heat demand, excess generation, thermal heat storage energy and natural gas-boiler energy profiles are analysed during the course of a day as well as throughout the year.

The daily heat demand profile is found to match people commute pattern, which is highest at 9 hrs and gradually decrease afterwards with an average
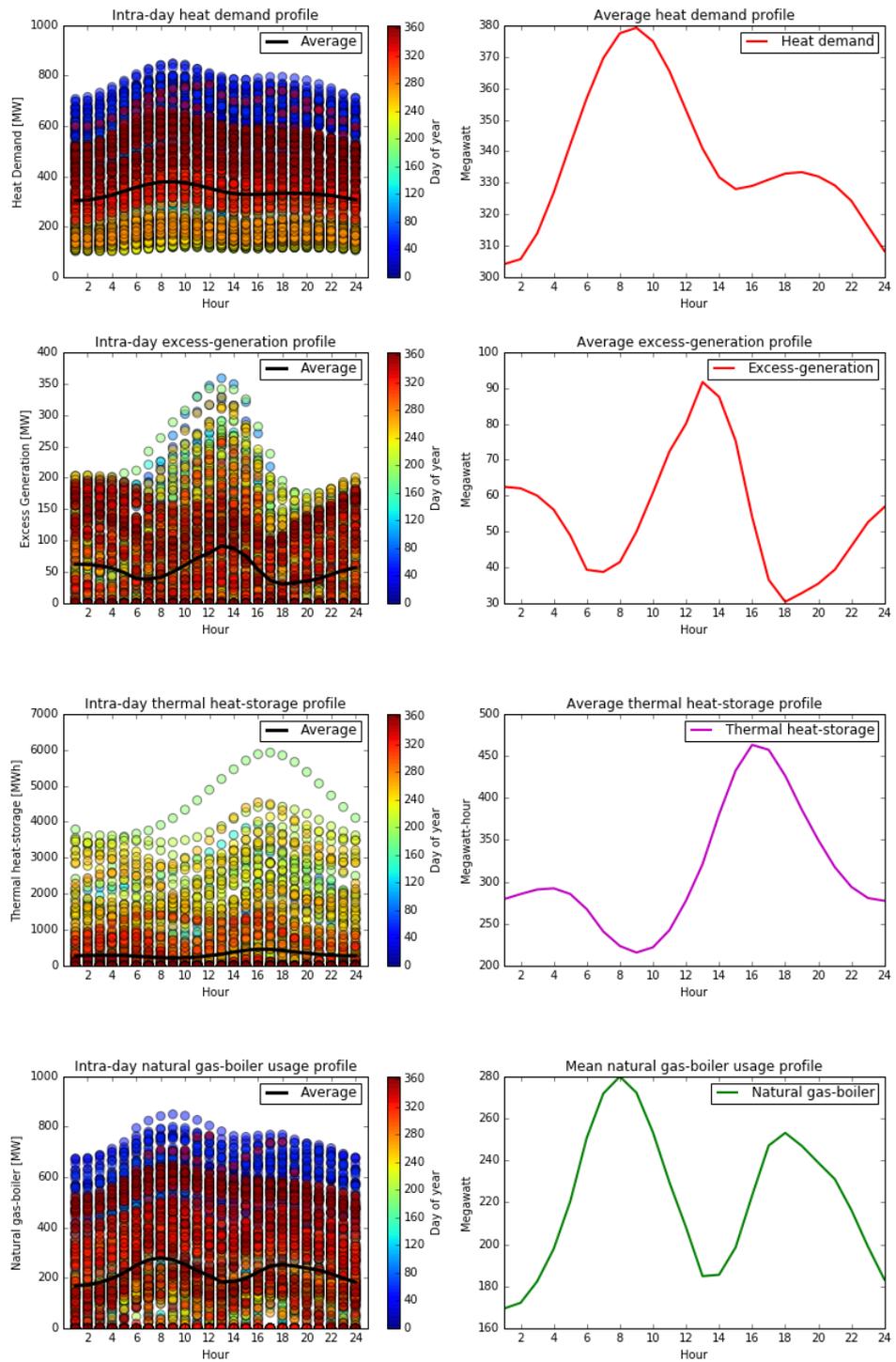
Figure 5.5: The left panels represents the hourly comparison between different components of heat-pump based coupling assumed for the city of Aarhus. The right panels represent the averaged profiles at each hour. The digit 0 in colour bar depicts the 1st January and 365 depicts 31st December. These results are adapted from (Ashfaq & Ianakiev 2018*a*).

daily variation between 310–380 MWh. The excess renewable energy generation and thermal heat storage energy profiles show a completely different trend. Both of these are minimum in the morning and rises quickly to the highest level around the afternoon, and abruptly decrease to their lowest levels during off-peak hours. This opposite behaviour provides the possibility of using thermal heat storage energy for the heating purposes. The average daily variation of excess renewable energy and thermal heat storage has been observed between 30–90 MWh and 220–470 MWh, respectively.

On the other hand, the natural gas-boiler energy profile shows similar trend to the daily heat demand. It is highest in the morning, due to absence of heat from the thermal heat storage and lowers in the afternoon with an increase in off-peak hours. The natural gas-boiler shows the average daily variation of 170–280 MWh. These profiles are shown are shown in Figure 5.5.

Furthermore, the time-series is analysed for the in-depth hourly investigation of the heat coupling in the network. It is observed that the excess heat is stored into thermal heat storage during summer season, and backup heat from the natural gas-boiler is used throughout the year with maximum during winter season.

Moreover, the natural gas boiler capacity reduces considerably by designing thermal heat storage capacity at 99% quantile compared to the 90%. In addition, the natural gas-boiler capacity reduces dramatically beyond the renewable energy generation (γ) of 140% and wind/solar mix (α^W) of 0.6, respectively as compared in Figure 5.6(c,d). These finding are in agreement with some other studies ((Ashfaq et al. 2017), (Pensini et al. 2014).

The middle column in Figure 5.7 depicts scenario, when the renewable
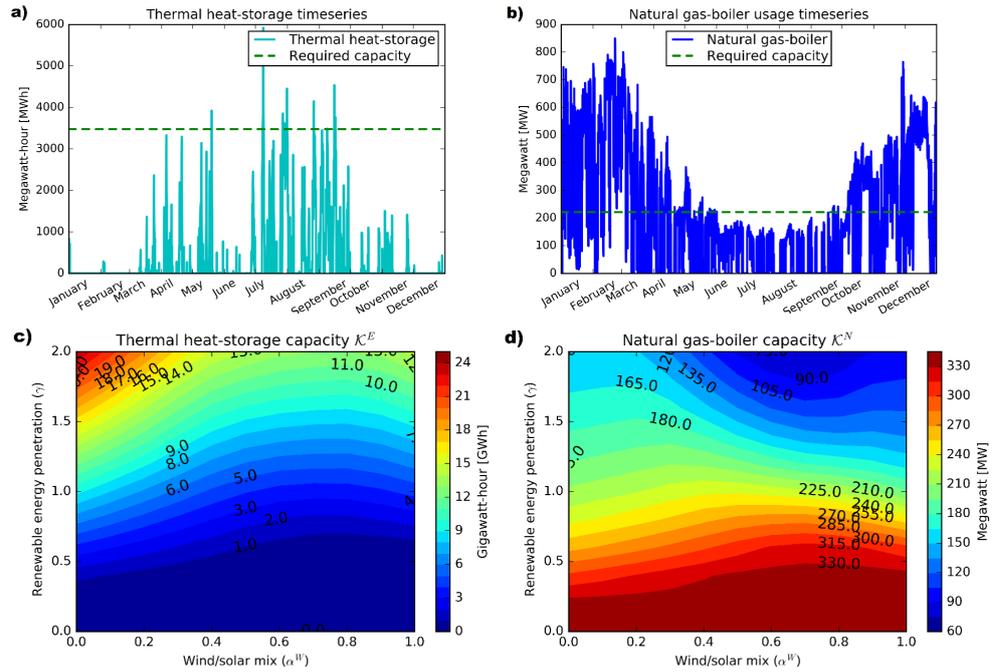
Figure 5.6: The left panels illustrates the capacity of thermal heat storage required in the network, whereas the right panels illustrates the capacity of gas-boiler. These results are adapted from (Ashfaq & Ianakiev 2018a).

energy penetration is equal to electrical load demand ($\gamma=1$). It is observed that both backup generation capacity and thermal heat storage capacity reduces with the increase in wind/solar mix and found to be minimum at $\alpha^W = 0.8$. Whereas, the natural gas boiler capacity initially increases with wind/solar mix and then starts decreasing after the wind/solar mix $\alpha^W \geq 0.6$. The natural gas boiler capacity is minimum at wind only mix ($\alpha^W = 1$). This behaviour of capacities is similar to that observed by (Ashfaq et al. 2017).

Minimum backup and storage capacity

Three different scenarios are considered in order to find the minimum backup and thermal heat storage capacity. It is found that the backup generation and natural gas-boiler capacity depends on the renewable energy generation (γ),
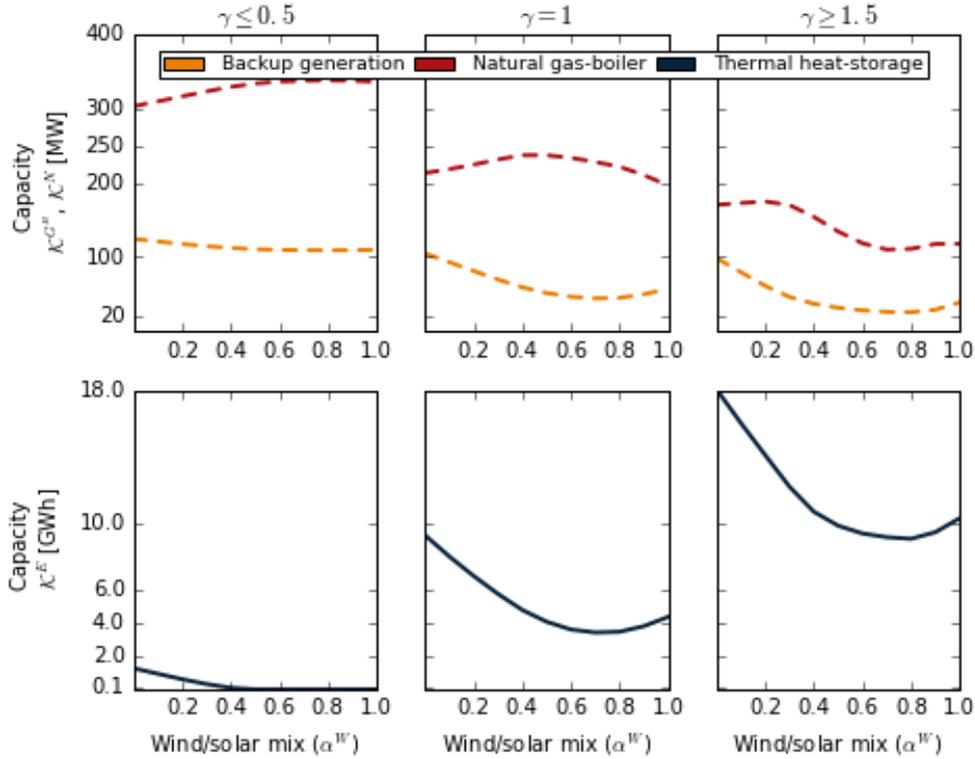
Figure 5.7: The required energy capacities for the decarbonised district heating network for different levels of renewable energy generation. The first row illustrates the capacity of gas-boiler and backup generation required in the network, whereas the second row illustrates the capacity of thermal heat storage. The columns represents the renewable energy penetration factor (γ) i.e. scaling of the total renewable power generation (wind and solar) to the electrical demand. These results are adapted from (Ashfaq & Ianakiev 2018*a*)

while thermal heat storage capacity depends on the wind/solar mix (α^W) in the network.

It is observed that at (γ=1) in the middle column of Figure 5.7, both backup energy and thermal heat storage capacity are minimum at wind/solar mix of 0.8 whereas, the natural gas-boiler capacity is minimum at wind/solar ($\alpha^W = 1.0$). The scenario with (γ=0.5) shows that both backup energy and natural gas-boiler capacity are at maximum levels and the thermal heat storage capacity is merely absent. On the other hand, at (γ=1.5) in the right column, the backup generation and natural gas-boiler capacity reduces considerably

due to increase in supply of heat from the thermal heat storage, and minimum at wind/solar mix (α^W=0.8).

Hence, in conclusion, the renewable energy penetration ($\gamma \geq$1.5) with wind/solar mix (α^W=0.8) is the most suitable option for the heat coupling of the electrical grid with excess renewable energy. The detailed economic modelling of the above three scenarios is described in the next section.

5.3   Economic modelling and analysis

The economic feasibility for the optimum configuration of renewable energy generation with the wind/solar mix is determined using concept of levelised cost of energy (LCOE). The LCOE is the net present value of the unit cost from the energy generation technology over its lifetime. In comparison to other economic feasibility methods, it is more robust and provides a holistic comparison between different technologies. (Rodriguez et al. 2015) used the same method for the economic feasibility calculation of the pan-European electrical grid.

There are two main categories of costs for any power generation technology. The initial capital investment (CapEx) costs and the operation and maintenance (OpEx) costs. The CapEx cost is an initial investment of the project whereas, OpEX costs are recurring monthly or annual costs during the technology's lifetime. The OpEx costs are further sub-divided into fixed monthly or annual operation and maintenance expenses, and variable fuel expenses.

The CapEx and OpEx costs considered for several renewable energy technologies have been adapted from the study on pan-European electrical grid by (Rodriguez et al. 2015). Additionally, the costs for the heating network

Table 5.1: Cost and other parameters assumed for different technologies for the levelised cost of energy calculations. This table is adapted from (Ashfaq & Ianakiev 2018*a*).

Technologies	CapEx Fixed [€/W]	OpEx Fixed [€/KW/year]	OpEx Variable [€/MWh]	Lifetime [years]	Transmission & distribution efficiency [%]
Wind turbines - onshore	1.0	15.0	-	30	100
Solar photo-voltaic (PV)	1.5	8.5	-	30	100
CCGT turbines	0.90	4.5	56	30	100
Heat pump	0.61	4.3	-	20	COP=3
Thermal heat storage	0.0027	0.03	-	35	90
DH network	-	2	-	30	90
Natural gas boiler	0.18	9.5	17.0	16	100

technologies have been taken from study on the United States by (Pensini et al. 2014). The backup generation is provided using Combined Cycle Gas Turbines (CCGT), and excess generation costs in the electrical grid is computed from the excess amount of wind and solar PV power generation costs. In calculations, the discount rate (r) is used as 4% and CapEx cost for the district heating network is not included due the existing network is already in-place at Aarhus, Denmark. The variable OpEx fossil fuel costs are assumed to remain unchanged in next 30 years for the CCGT (56 €/MWh) and natural gas-boilers (17 €/MWh). Table 5.1 enlists the cost used in this study for different technologies.

$$\text{Cost}_{\text{€}/t} = \text{Power}_{\text{MWh}/t}.\text{Cost}_{\text{€}/\text{MWh}} \qquad (5.16)$$

The costs for all technologies consider the net-present value (V) of future expenses by the capacities required to fulfill the energy demand. The following equation computes net-present value (V) with the discount rate (r) of 4%.

$$V = \text{CapEx} + \sum_{t=1}^{T} \frac{\text{OpEx}_t}{(1+r)^t} \tag{5.17}$$

$$V_{elec} = V_{G^W} + V_{G^S} + V_B \tag{5.18}$$

$$V_{heat} = V_{HP} + V_E + V_{DH} + V_N \tag{5.19}$$

where, T denotes the lifetime of technology. V_{elec} represents the net-present value for the wind turbines-onshore (G^W), solar PV (G^S) and CCGT turbines (B). The V_{heat} represents the net-present value for the heat pump (HP), thermal heat storage (E), district heating network (DH) and natural gas boiler (N). Then, the LCOE is calculated as,

$$V_{sys} = V_{elec} + V_{heat} \tag{5.20}$$

$$\text{LCOE} = \frac{V_{sys}}{\sum_{t=1}^{T} \frac{L_t + H_t}{(1+r)^t}} \tag{5.21}$$

The LCOE divides total energy generation expenses to the sum of total energy demand during the lifespan of network (Rodriguez et al. 2015). In Equation (5.21), the sum of discounted energy generation costs is divided by the discounted future energy demand. It is important to note that the wind and solar (PV) generation are two main power generation sources. Whereas, the combined cycle gas turbines (CCGT), thermal heat storage and natural gas-boilers are only for backup power and heat sources.

5.3.1 Economic optimal network

The economic feasibility for both renewable based standalone (electrical and heating) and coupled networks is calculated using LCOE from the Equation(5.21). The capacities and costs used for the analysis are taken from the Section 5.2.3 and Table 5.1. First, the current heat generation cost from the existing district heating network is discussed and then, the economically optimum configuration for each network with the breakdown of costs is discussed by considering three scenarios. Where, each scenario for different amount of renewable energy generation (γ) is represented in separate columns, and networks are represented by rows. The first and second rows represents the fully renewable energy based electrical grid and heating network with heat-coupling, and the third row elaborates coupled networks (electrical and heating).

The analysis of heat generation cost for the existing district heating network shows that the heat demand for the city of Aarhus is 2.9 TWh which is 5% of the Denmark's heat demand. This heat is generated with an average of 40 €/MWh per hour and approximately 2.2 million € worth of heat can be supplied from the thermal heat storage with the installation of heat-pump based coupling. These results are shown in Figure 5.8.

The LCOE results show that when renewable energy generation (γ=1) is equal to the electrical load demand, then the cost for the fully renewable based electrical grid, district heating network and coupled networks varies between 120 - 170, 28 – 35 and 65 - 80 €/MWh, respectively. The results further confirm that the energy costs are decreased to 45 €/MWh for the coupled network, as compared to the existing district heating network costs of 40 €/MWh alone. The LCOE for each network decreases with increase in
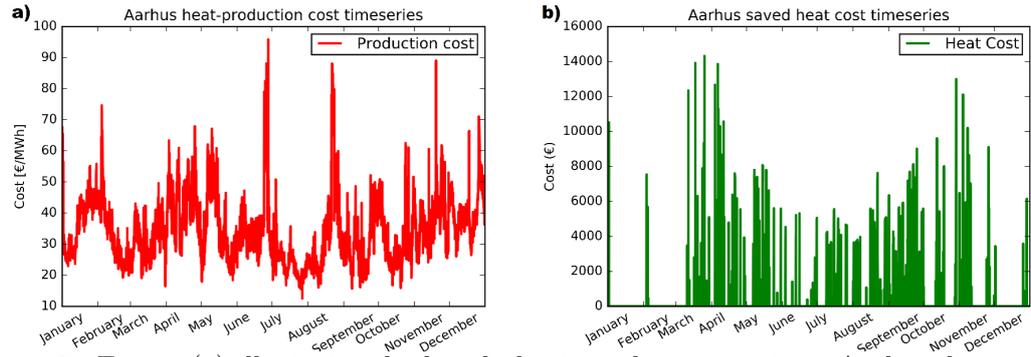
Figure 5.8: Figure (a) illustrates the hourly heat-production costs in Aarhus district heating network and Figure (b) illustrates the cost of heat savings with the implemenation of heat pump based coupling. These results are adapted from (Ashfaq & Ianakiev 2018*a*)

wind power generation, and minimum at wind power generation of 85% and solar power generation of 15%.

Furthermore, when renewable energy penetration (gamma >= 1.5), then the cost for the fully renewable based electrical grid, district heating network and coupled networks varies between 180 - 260, 18 - 28 and $45 - 65$ €/MWh, respectively. This increase is due to increase in renewable energy generation costs at heat-pump costs in the network. The LCOE is minimum at wind power generation of 80% and solar power generation of 20%. Finally, the last scenario, when renewable energy penetration ($\gamma \leq 0.5$) then, the increase in the cost of backup generation and natural gas-boiler is observed, and the LCOE is minimum at wind power generation of 100%. The breakdown of LCOE cost is shown in Figure 5.9.

In conclusion, the coupling of networks results into a minimum LCOE cost and the optimum solution for the de-carbonised future heating network can be achieved with renewable energy penetration ($\gamma \geq 1.5$), and wind power generation of 80% and solar power generation of 20%.

The above economic analysis would be more valuable with the inclusion of
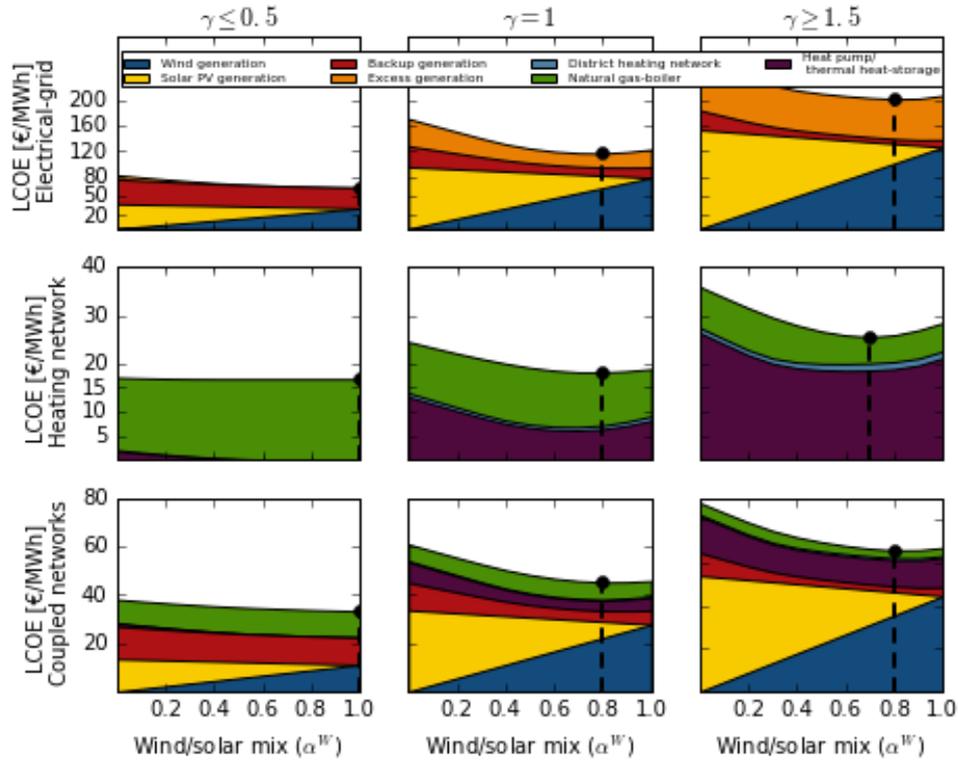
Figure 5.9: Levelised cost of energy for the decarbonised district heating network for different levels of renewable energy generation. The first and second rows illustrate the costs of different technologies in the decarbonised electrical grid and district heating network. The third row illustrates the costs once both networks are coupled together. The columns represents the renewable energy penetration factor (γ) i.e. scaling of the total renewable power generation (wind and solar) to the electrical demand. These results are adapted from (Ashfaq & Ianakiev 2018*a*).

sensitivity analysis. Therefore, the sensitivity analysis is explored to analyse the effect of different parameters with variations in cost assumption. Then, the LCOE of different technologies is studied with the increase in renewable energy penetration (γ) and selling of excess generation.

5.3.2 Effect of cost-variations

This section considers six scenarios to analysis the LCOE with variations in cost assumption for different technologies. As the focus is on the decarbonisation of the heating network therefore the cost assumptions for the
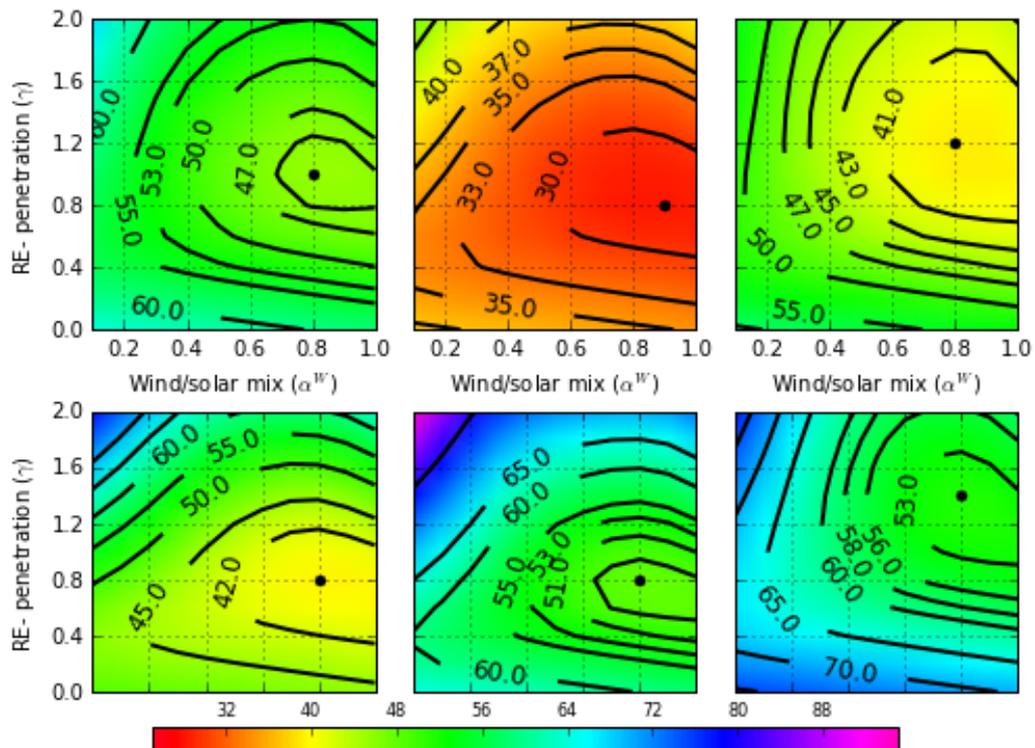
Figure 5.10: Effect of cost-variations on levelised cost of energy for the decarbonised district heating network. Panel (a) represents LCOE from the initial cost assumptions, Panels (b) and (c) represents LCOE with 50% reduction in costs assumed for the electrical grid and district heating network, Panel (d) assumes once both electrical grid and district heating network costs are reduced by 50%. Finally panels (e) and (e) assumes once the cost are increased by 50% for the heat pumps and natural gas boiler, respectively. The dark dot represents the optimum solution. These results are adapted from (Ashfaq & Ianakiev 2018*a*).

heating network are varied in four scenarios and the optimum cost-minimum solution is calculated from high-low scenarios in order to determine the optimum renewable energy penetration (γ) and wind/solar mix (α^W). The capacities, CapEX and OpEX expenses, and other assumptions are taken from the Section.5.2.3 and Table 5.1.

The first scenario in Figure 5.10 is for the comparison and considers the coupled network with cost assumptions given in Table 5.1 and the optimum LCOE is 45 €/MWh. The second and third scenario in Figure 5.10(b),(c) investigated the LCOE when cost assumptions are decreased by 50% for the

electrical grid and heat network, and the optimum LCOE is reduced to 28 and 39 euro/MWh, respectively. It is observed that the decrease in heating network cost assumptions increases the renewable energy generation (γ) for the optimum LCOE.

The fourth scenario in Figure 5.10(d), is opposite to the third scenario and investigated LCOE when cost assumptions are increased by 50% for the electrical grid and heat network. The optimum LCOE is 41 €/MWh, but the renewable energy generation (γ) is decreased to 0.8. The fifth and sixth scenarios in Figure 5.10(e),(f) evaluates LCOE when cost assumptions for the heat-pump and natural gas-boilers are increased by 50% in the heat network, and the optimum LCOE is increase to 50 and 52 €/MWh, respectively. It is observed that the increase in heat-pump cost assumptions has significant impact on the renewable energy generation ($\gamma = 0.8$) for the optimum LCOE, compared to the natural gas-boiler.

In summary, the renewable energy generation (γ) of the coupled network depends on the cost assumed for the heat-pump and natural gas-boiler in the heat network, and lowest LCOE is determined from the costs assumed for the wind and solar(PV) power generation.

5.3.3 Effect of renewable energy penetration

This section investigates the impact of renewable energy penetration (γ) on the LCOE for each component. It is observed that the LCOE for excess generation is greater than backup energy generation after the renewable energy penetration ($\gamma = 0.9$). The excess generation costs increase sharply after renewable energy penetration ($\gamma \geq 0.6$), and the wind and solar(PV) power generation increases to the maximum of 45 €/MWh and 130 €/MWh, respec-

tively. Moreover, the major share of costs in the coupled network is initially from the backup energy and back-up heat generation, but once the renewable energy penetration (γ) increases beyond 0.6, the cost of backup generation and natural gas-boiler begins to decrease from 18 to 4 €/MWh and 60 to 5 €/MWh, respectively. The LCOE costs for the heat-pump increases from 2 to 26 €/MWh, and thermal heat storage and district heating network increases to up to 2 €/MWh. These results are further shown in Figure 5.11.

In summary, the LCOE for the fully renewable energy based electrical grid, heat network and coupled network with variations in renewable energy generation is between 60 to 260, 18 to 34 and 45 to 80 €/MWh, respectively. The LCOE in coupled network is reduced by 50%, and LCOE for heat-pump is greater than the natural gas-boiler beyond the renewable energy penetration ($\gamma \geq 1.3$).

5.3.4   Effect of selling excess generation

Recently, there has been a great push towards the increase penetration of renewable energy and using excess generation in the electric grid or decarbonisation of the heat network. Therefore, the effect on LCOE with selling excess generation to the electrical grid (private wire) is discussed in this section.

It is found that when the cost of supplying excess generation is zero, then the LCOE of supplying excess generation in the coupled network is 45 €/MWh and the LCOE is minimum at renewable energy penetration ($\gamma = 1$) and wind and solar PV power generation of 85% and 15%, respectively. However, when the cost of excess generation is increased to 54 €/MWh, then the LCOE is 48.3 €/MWh whereas, the renewable energy penetration ($\gamma =$
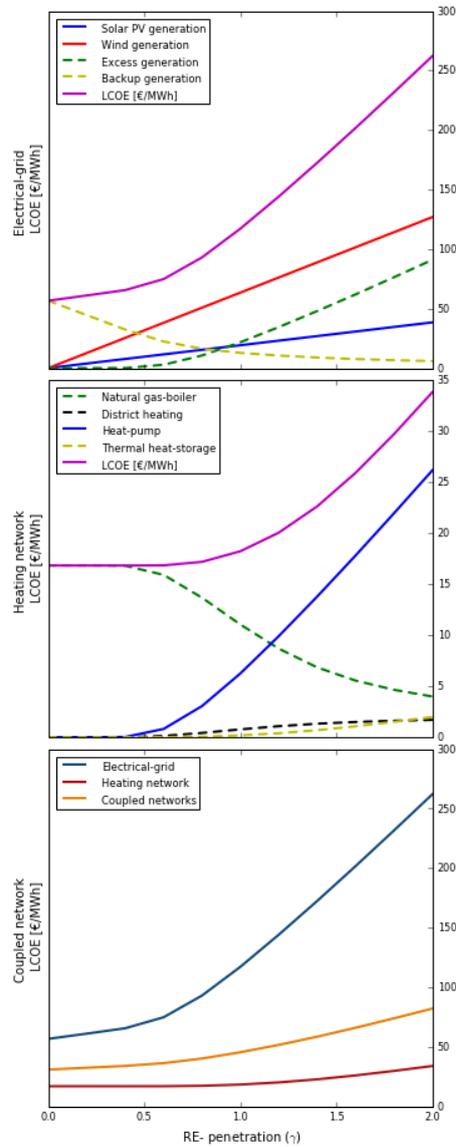
Figure 5.11: Effect of renewable energy generation to the LCOE of decarbonised energy networks. First and second panels illustrates effects on the decarbonised electrical grid and district heating network costs, and the third panel illustrates the effect once both networks are coupled. These results are adapted from (Ashfaq & Ianakiev 2018*a*).
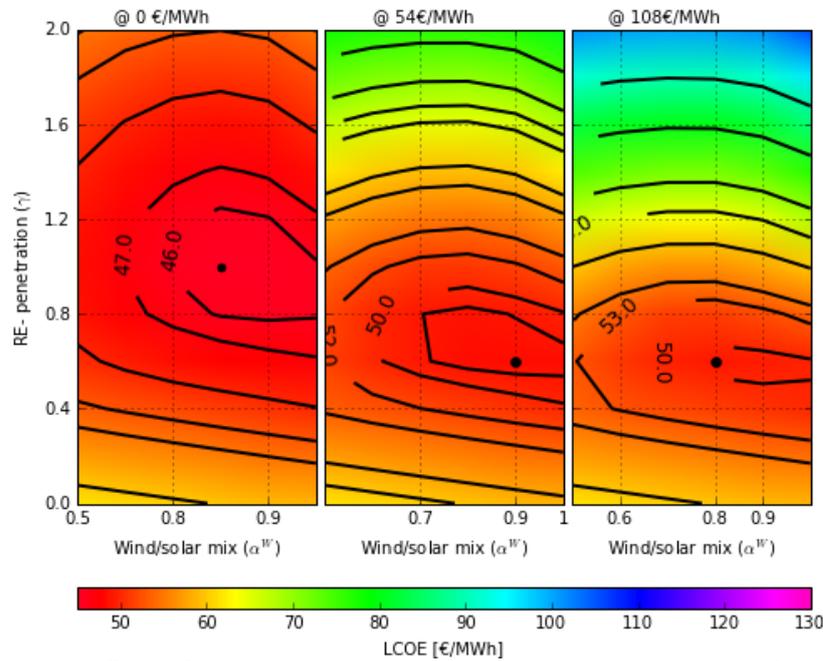
Figure 5.12: Effect of selling excess renewable energy generation of decarbonisation of district heating network. The left panel represents LCOE when the excess generation is sold free. Middle and right panels represent LCOE once the excess generation is sold at 54 and 108 €/MWh, respectively. The dark dot represents the optimum solution. These results are adapted from (Ashfaq & Ianakiev 2018*a*).

0.7) is decreased. Moreover, when the excess generation is increased twice to 108 €/MWh, then the LCOE is 49 €/MWh whereas, the wind and solar(PV) power generation are reduced to 80% and 20%, respectively. These results are shown in Figure 5.12.

It is concluded that selling excess generation from the coupled network does not have a strong effect on the LCOE, unless the renewable energy penetration increases to twice the amount of current electrical load. This can be vital for the de-carbonised fossil free future.

5.4 Discussion and conclusion

In this chapter, the real heat demand and heat generation cost data for the city of Aarhus, Denmark is used to analyse the cost-minimised design for the

decarbonisation of heat network. It adapts the idea of fully renewable pan-European electricity grid from earlier study by (Rodriguez et al. 2015) and incorporates heat network into it. Moreover, this methodology is quite robust and has already been used in several studies including; calculation of optimum wind and solar power generation (Heide et al. 2010, Andresen et al. 2014), the battery storage in electrical grid (Rasmussen et al. 2012, Heide et al. 2011), transition to pan-European renewable electrical grid (Becker, Rodriguez, Andresen, Schramm & Greiner 2014, Rodríguez et al. 2014, Dahl 2015), interconnected fully renewable based US electrical grid (Becker, Frew, Andresen, Zeyer, Schramm, Greiner & Jacobson 2014, Becker et al. 2015) and backup energy storage flexibility in large-scale renewable systems (Schlachtberger et al. 2016).

The analysis for the optimum network explains that the heat demand profile throughout the day is in agreement with people commuting to work and designing thermal heat storage capacity to the 99% quantiles of the energy usage time-series is beneficial in reducing the amount of capacities required in the network. Moreover, the data described in Section 5.2.3 reveals that the coupled networks (electrical grid and heating network) can be useful in avoiding demand variations.

The economic feasibility analysis suggests that the renewable energy based electrical grid, heat network and coupled networks are found to be minimum at 120, 35 and 45 €/MWh respectively. This further explains that the coupled networks with the wind power generation (85%) and solar (PV) generation (15%) have the LCOE of 45 €/MWh, in comparison to the heat generation cost from the existing district heating network alone is 40 €/MWh. In-future,

the costs for fully renewable energy based stand-alone electrical grid will be increased beyond 260 €/MWh, which will limit the future renewable energy penetration into the electrical grid. Therefore, the sector coupling of the electrical and heat network is necessary, as the energy costs will be decreased by around 50%.

There has been a considerable variation in cost assumptions in different studies and makes the LCOE comparison slightly challenging. The costs assumed in (Jacobson & Delucchi 2011) is approximately 50% lower than assumptions in (Arup & Ltd 2011). Furthermore, the CapEx, OpEx expenses, capacity factors and the technology lifetimes are different in studies, (EIA, U.S. Energy Information Administration 2015, Steward et al. 2009, Schaber et al. 2012, Fürsch et al. 2010), respectively. Therefore, the CapEx and OpEX costs, capacity factors and technology lifetime assumptions are taken from previous studies and sensitivity analysis has been performed.

The sensitivity analysis results reveal the dependence of LCOE on the cost assumed for wind and solar (PV) power generation. It is observed that the reduction in wind and solar (PV) power generation costs directly impact the optimum renewable energy penetration (γ). However, the renewable energy penetration (γ) depends on the cost assumed for the heat-pump and natural gas-boiler in the heat network. Furthermore, the LCOE for heat-pump is found to be greater than natural gas-boiler beyond the renewable energy penetration of 130%.

In summary, the carbon-neutral future with fully renewable based electrical grid and heat network is achievable, but this will lead to fluctuations in heat demand compared to the existing district heat network, unless demand

side management, energy from waste and other sources is utilised. Finally, the excess generation can be sold up to 108 €/MWh without significant impact on LCOE of the coupled network, and the de-carbonised fossil free future is possible. This can be instrumental towards economic viability of the fossil-free future heat network.

5.5   Summary

This chapter provides an insight into the techno-economic analysis for the de-carbonised district heat network and the results show that the futuristic fully renewable powered electrical grid and heating network is not economically feasible, unless both networks are coupled together which reduces the energy costs by almost 50%. Additionally, the per unit decarbonisation cost depends on wind and solar (PV) power generation cost. The amount of excess renewable energy generation from the electrical grid is determined by per unit cost of heat-pumps and backup energy technologies (thermal heat storage and natural gas-boiler). Finally, the 100% decarbonisation of the district heat network is possible but depends on usage of private wire (i.e. selling excess electricity and heat to the private consumers).

Chapter 6

Machine learning and forecasting

Overview

In this chapter, the aim is to investigate various forecasting and prediction methods on the time-series data. The forecast error and accuracy of each method is compared to determine the robustness and limitations of each method. The performance of regression based supervised machine learning methods have been compared against classical stochastic methods for time-series forecasting.

6.1 Forecasting in district heating networks

The district heating is a network of heat generation source, hot water supply, cold return transmission pipes and heat distribution equipment. The heat generation source can be combined heat and power (CHP), heat pumps, heat boilers, geothermal and solar thermal plants. In a traditional district heating network, the heat is transferred to buildings with the circulation of supply hot

water of 70-120°C and the return cold water is around 40-65°C (Frederiksen & Werner 2013), which is then heated and re-circulated.

The district heating network contains complex dynamics, and the energy efficient operation of district heating network depends on the optimum balance between the pressure, flow-rate and temperature difference between the supply and return water temperature in pipes. The variations in pressure and flow-rate travels through the network with speed of light, whereas the variations in temperature difference are slower and can take hours to show their effect (Gabrielaitiene et al. 2010). The heat required by the consumers can be fulfilled by regulating either the flow-rate or temperature difference between the supply and return water temperature in the network, as shown in Equation (2.3). Therefore, the correct short-term forecasting is crucial for the energy efficient operation of the district heating network.

Forecasting in district heating network is vital for its operation. There are three types of forecast horizons depending on their purpose, i.e. short (24-168 hrs), medium (<1 month) and long-term (>1 year) forecasts (Poutiainen 2019). The short-term forecast is of peculiar importance when it comes to energy efficiency and optimisation. The accurate short-term forecasts are useful to minimise the peaks, demand side management and imbalance in the network.

6.2   Time-series data analysis

The time-series data analysis aims to analyse and extract the hidden patterns inside the data before implementation of machine learning and classical forecasting methods. The continuous variations in high-resolution data makes time-series predictions challenging. Therefore, the time-series data analysis is helpful in identifying the correct forecasting method to be used.

The time-series data being used in this research is from the REMOUR-BAN Project, where 1 year (8760 timsteps) of data is monitored from one block of flats. The aim here is to effectively forecast the supply and return water temperature using several machine learning and forecasting methods.
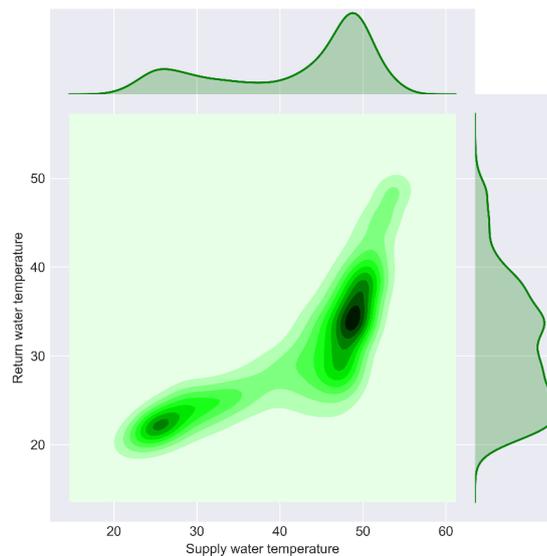
Figure 6.1: Univariate kernel density estimate plot of the supply and return water temperature time-series data

To this end, the 1 year raw data has been plotted to gain insights into the existing trends. Figure 6.1 shows that once the supply water temperature is between 45-55°C then the return water temperature varies between 25-40°C. Figure 6.2 shows distribution of supply and return water temperature using a histogram. The data informs that the supply water temperature tends to remain in the range of 45-50°C whereas the return water temperature is mostly in the range of 25-40°C. Figure 6.3 shows the hourly comparison between supply and return water temperature throughout the year using a heatmap. The data shows that both supply and return water temperatures are lower in summer which is an expected behaviour. However, the return water temperature

is observed higher than expected during months of November and December

especially during 10:00-24:00 hrs, which is an abnormal fluctuation.
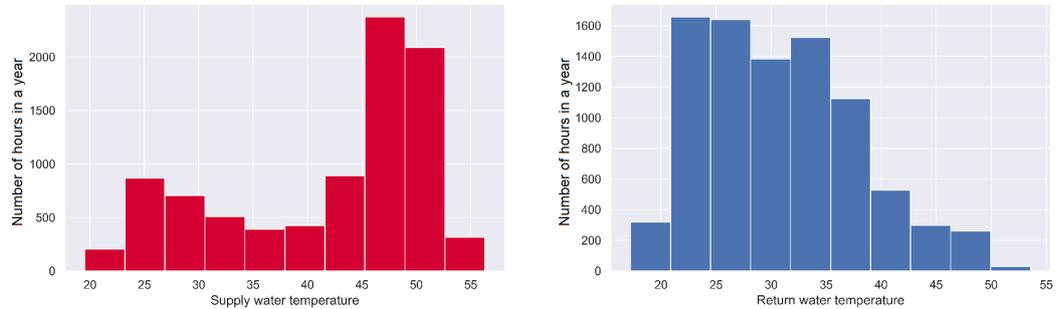
Figure 6.2: Histogram of the supply and return temperature time-series data
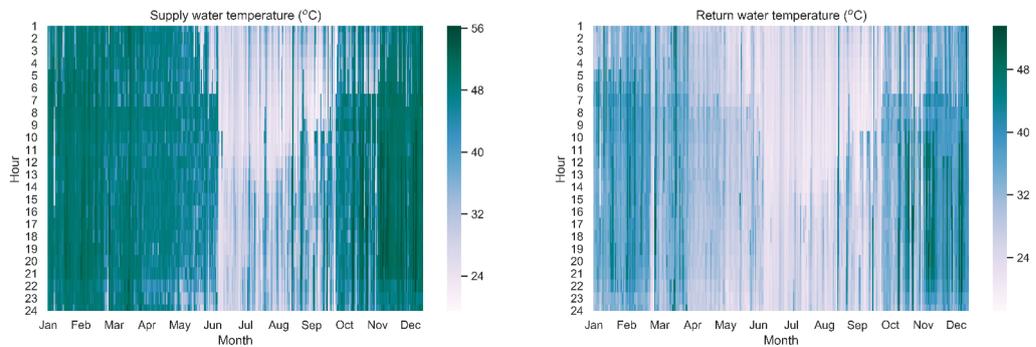
Figure 6.3: Heat-map representing hourly comparison between supply and return water temperature throughout the year

6.2.1   Dataset processing

The dataset is initially normalised using *MinMax* scaling method and the data

is split into two (i.e. train and test data set) with a ratio of 75/25. The training

data set is used for creating the model, and the test data set is for evaluating

model's performance. This splitting of data helps in creating a robust model

which generalises well and works good on unseen new data.

The dataset contains several dependent and independent feature and their

correlation has been shown in the Figure 6.4. The data show that the corre-

lation between supply and return water temperature is found to be greater
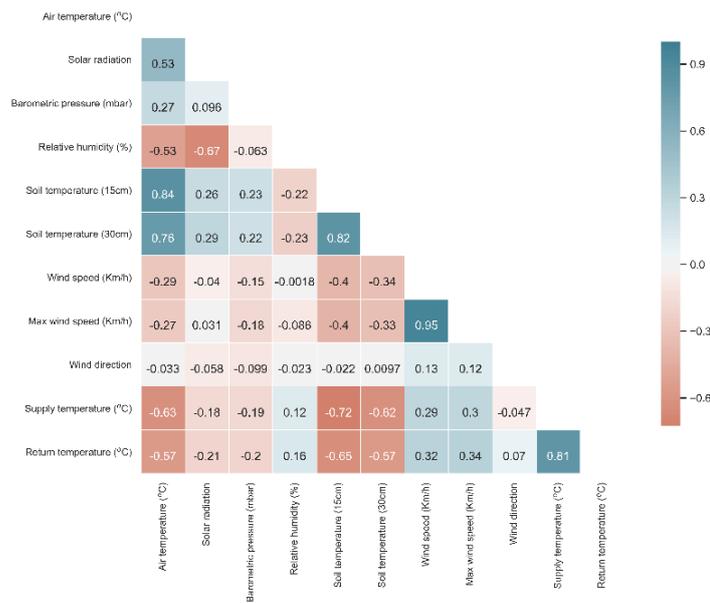
Figure 6.4: Heatmap representing the correlation between different features in the dataset

than the outside air temperature data. This explains that the weather data (outside air temperature, solar insolation, humidity and wind data) alone is not enough for forecasting using machine learning methods, and the accuracy of machine learning methods improve once the return water temperature is used to forecast the supply water temperature, and vice versa. The intention here is to use the above described monitored data to build a machine learning model, which can forecast the supply and return water temperature of the district heating network in future.

It should be noted that one should be very careful with the input dataset preparation for machine learning models, and the final model must not have a random walk. The random walk behaviour makes the user believe that the model is working and predicting but in-reality, the model is just following the input data. In this thesis, this fact has been careful considered by first creating two separate training datasets (supply and return), and then removing the

corresponding known temperature dataset i.e. 8760 timesteps of the known supply and return water temperature. This leads to two final training datasets each with nine features (9x8760 timesteps).

6.3   Supervised machine learning methods

In this section, the supervised machine learning methods are implemented for the prediction analysis over the time-series data-set explained above. The results are evaluated and optimum parameters are selected.

In supervised learning the prediction model is built using a known input and output datasets. The supervised machine learning methods are used to train the model to generate predictions on new unknown datasets. The supervised machine learning methods are available for both classification and regression prediction problems. The supervised learning methods include linear regression, logistic regression, neural networks, decision tree, Support Vector Machine (SVM), random forest, naive Bayes, and k-nearest neighbor.

The regression based supervised machine learning methods are particularly powerful in extracting relation and patterns in time-series data for forecasting and prediction. To this end, several machine learning models are developed and the time-series dataset of nine features is split into two parts i.e. training and testing data. Where the machine learning models are developed on the training data and future predictions are evaluated using unseen testing data.

The following supervised machine learning regression methods from the Python's *scikit-learn library* are used for the analysis:

- k-Nearest Neighbors

- linear regression (LR)

  - Ordinary least square regression

  - Lasso regression

  - Ridge regression

  - Elastic net (EN)

  - Stochastic gradient descent (SGD)

- Decision trees

- Ensemble methods

  - Random forest

  - Gradient boosted decision trees

- Support vector machines (SVM)

  - Radial bais regression (RBF)

  - Linear regression

  - Polynominal regression

- Neural networks

In this thesis, the regression methods are used as the aim is to predict continuous values from the data-set and classification methods (such as Naive Bayes) are not useful. The cross-validation is applied on the data before fitting

the models. This is to remove biasedness in the machine learning model. The models are carefully tuned for the generalised conditions using Grid Search method and the accuracy metrics are applied on both training and testing datasets to investigate the possibility of over-fitting and under-fitting.

6.3.1 K-nearest neighbors (K-NN)

The K-Nearest neighbors is one of the simplest method among other predictive methods. This is because it considers only two parameters, which includes number of neighbors and distance between two data points (Müller et al. 2016). The K-NN method does not require a lot of parameter adjustment and it is highly recommended to use as a baseline method before embarking into advanced methods. On the current dataset, the predictions are initially performed by selecting small number of neighbours and Euclidean distance is selected as the measure of distance between two data points.

The quality of prediction is evaluated using R2 score metric (coefficient of determination). The R2 score varies between 0 to 1, where 1 represents the predicted dataset values from the model are prefect, and 0 represents the predicted values are not good enough. This R2 score is calculated for both training and predicted datasets.

The results show that the prediction score depends on the number of neighbours considered in the K-NN model. The R2 score for the predicted supply water temperature dataset with 1 neighbour increases from 0.42 to 0.56 with 9 neighbours. Moreover, the training dataset score elaborates that the model's performance increases with the number of neighbours but upto a certain limit. The training dataset score is initially 1 which represents over-fitting of the model, but the over-fitting decreases to 0.9 once the neighbours
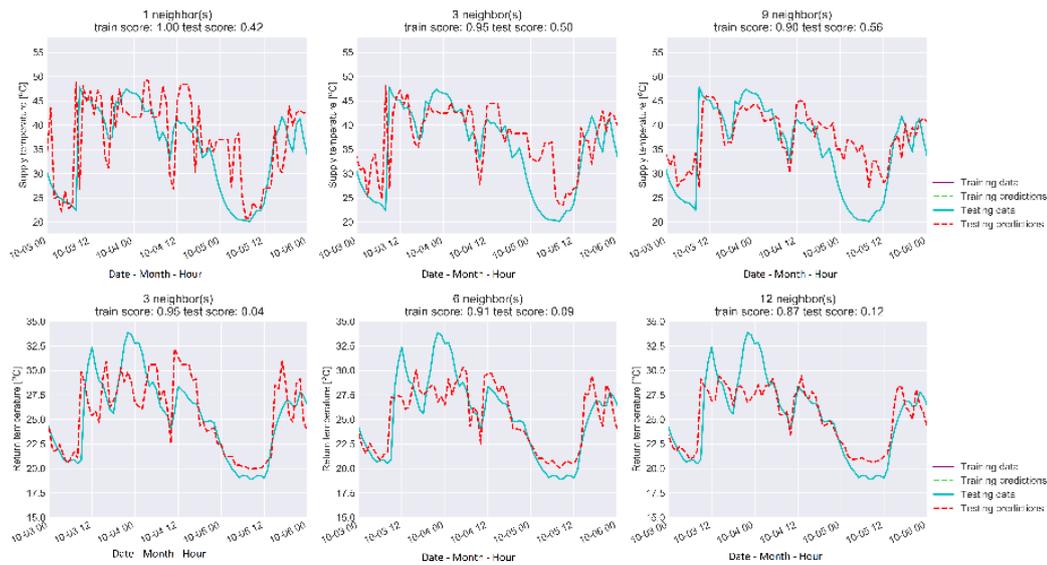
Figure 6.5: Predictions using K-NN machine learning method. Three panels in the first row compares the model's prediction performance results with different parameters on supply water temperature data, whereas the three panels in second row compares results on return water temperature data. The hourly data in Cyan colour represents testing data and Red colour represents prediction on testing data from 3rdto 5th October.

are increased. The training and predicted dataset scores for the supply and return water temperature are shown in Figure 6.5. The optimum prediction score of 0.56 (i.e 56%) has been achieved with 9 neighbours in the model, which is reasonable performance for a simple prediction method.

In conclusion, the K-NN model is good for small datasets, easy to implement and provides reasonable performance without fine tuning or adjustments. On the other hand, predictions become slow and inefficient on large dataset with several features. The model also doesn't perform on sparse data (where most of the values are 0), which makes it not preferred for practical problems. Nevertheless, it is good algorithm for baseline analysis before embarking to advanced methods.

6.3.2 Linear models

Linear models are primarily based on basic linear equation principle and their predictions are based on linear function of the input features. Over the period of time, these have been developed and evolved into several sub methods which have been used here. These are considered good for very large datasets with high dimensions.

6.3.2.1 Ordinary least squares regression

Ordinary least squares (OLS) also known as linear regression is the most classic and easy method for regression. The model automatically calculates the slope (w) and offset (b) parameters from the training dataset to minimise the mean squared error between the training data and predictions. It also considers all features from the data for predictions.

Figure 6.6: Predictions using ordinary least square regression method. The top panel shows the predictions on supply water temperature data and lower panel shows the predictions on return water temperature data. The data in figure presents the hourly prediction results from January to December. Where the Purple and Cyan colour represent actual training and testing data, whereas Green and Red colour represent predictions on the training and testing data.

Figure 6.6 show results using OLS method and the R2 for supply and return water temperature appears to be 0.63 (i.e 63 %) and 0.40 (i.e 40 %) respectively. These predictions are clearly better than the previously used k-NN method and implies that the OLS model provides better predictions on larger datasets. In OLS method, the user can not control the model parameters which leads to the limited flexibility in controlling model's complexity for reducing over or under-fitting of the model.

6.3.2.2 Ridge regression

Ridge regression is an enhanced version of OLS method and provides the ability to control model complexity as well as over or under-fitting by introducing additional constraints known as L2 regularisation. In Ridge regression model, the L2 regularisation is controlled using parameter alpha (α) which explicitly restricts over-fitting of the model. In principle, the coefficients of slope (w) should be closer to zero.

Figure 6.7: Grid Search for parameter optimisation - ridge regression method. The left panel shows parameter grid for supply water temperature data and right panel shows for return water temperature data.

Theoretically, increasing alpha to a certain level reduces the value of slope

coefficient closer to 0. The grid search method is used to find the best value of alpha. The entire process is repeated for different values of alpha (i.e, 0, 10, 100, 1000, 1000) and maximum iterations (i.e, 0, 1, 10, 100, 1000, 10000). As shown in Figure 6.7, the optimum value of alpha for supply and return water temperature is found to be 100 and 10,000 respectively. The number of iterations does not effect the predictions and increasing alpha beyond 1000 effects the model's prediction ability. This model takes into account all features from the data for predictions.
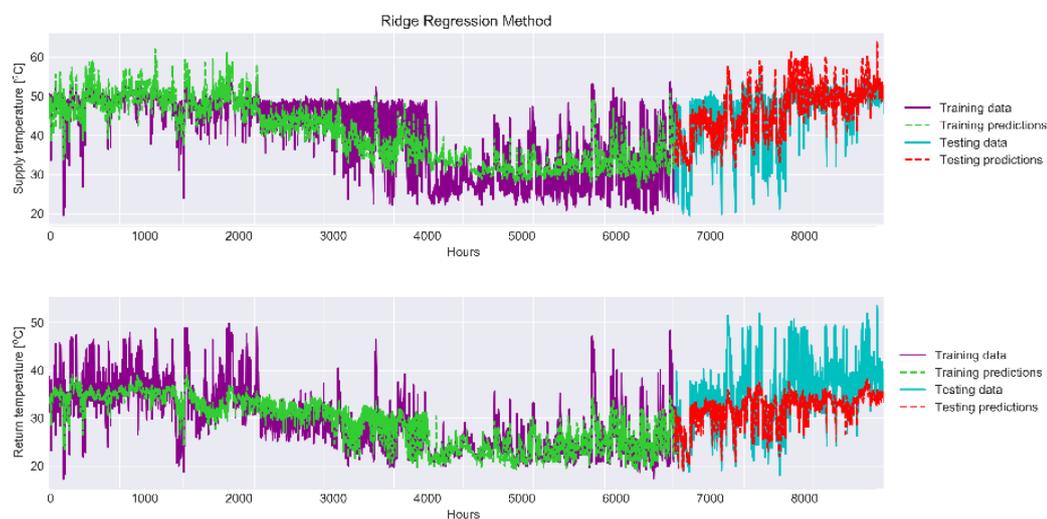
Figure 6.8: Predictions using ridge regression method. The top panel shows the predictions on supply water temperature data and lower panel shows the predictions on return water temperature data. The data in figure presents the hourly prediction results from January to December. Where the Purple and Cyan colour represent actual training and testing data, whereas Green and Red colour represent predictions on the training and testing data.

The Figures 6.8 and 6.9 show that Ridge regression method (alpha=10) provides slightly better predictions compared to OLS method. The R2 score for supply and return water temperature is found to be 0.67 (i.e 67 %) and 0.43 (i.e 43 %) respectively. This suggest that the regularisation of slope coefficients improves the model accuracy.

Figure 6.9: 72 hours predictions using ridge regression method. The hourly data in Cyan colour represents testing data and Red colour represents prediction on testing data from $3^{rd}$to 5th October. The left panel shows the predictions on supply water temperature data and right panel shows the predictions on return water temperature data.

6.3.2.3 Lasso regression

Lasso (Least absolute shrinkage and selection operator) regression is similar to Ridge method, but uses L1 regularisation (instead of L2) and considers only limited features from the training dataset. These features are selected automatically with the coefficients of slope (w) being set closer to zero. The remaining unselected features are entirely ignored. Similar to Ridge regression, Lasso regression takes into account alpha and maximum iterations where maximum iterations is a key contributing factor to find the optimum predictions.
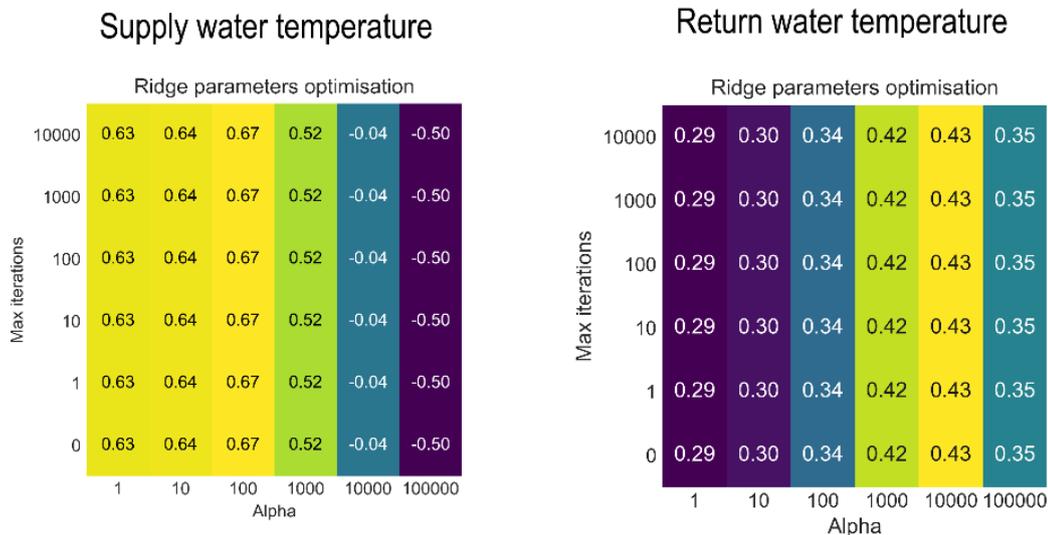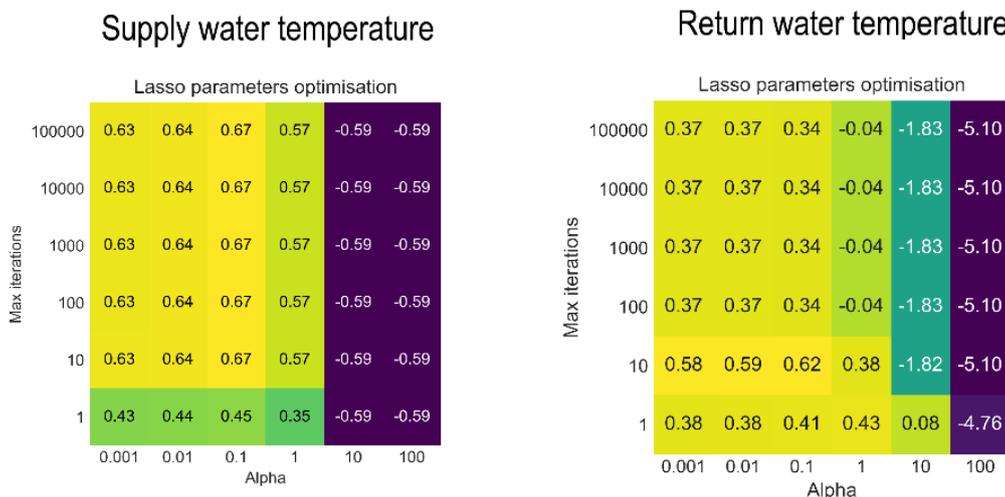

Figure 6.10: Grid Search for parameter optimisation - lasso regression method. The left panel shows parameter grid for supply water temperature data and right panel shows for return water temperature data.

As show in Figure 6.10, the grid search method is implemented to find the

optimum combination of alpha and maximum iterations. The results suggest that alpha ($\alpha > 1$ and < 0.001) leads to over-fitting of the model and the number of iterations greater than 10 are necessary to achieve better predictions.

Figure 6.11: Predictions using lasso regression method. The top panel shows the predictions on supply water temperature data and lower panel shows the predictions on return water temperature data. The data in figure presents the hourly prediction results from January to December. Where the Purple and Cyan colour represent actual training and testing data, whereas Green and Red colour represent predictions on the training and testing data.

Figure 6.12: 72 hours predictions using lasso regression method. The hourly data in Cyan colour represents testing data and Red colour represents prediction on testing data from 3rd to 5th October. The left panel shows the predictions on supply water temperature data and right panel shows the predictions on return water temperature data.

The Figures 6.11 and 6.12 show that Lasso regression method (alpha=0.1, max iter=10) provides similar predictions compared to Ridge method. However, it is only selecting 5 features out of 10 for our dataset. The R2 score for supply and return water temperature is found to be 0.67 (i.e 67 %) and

0.62 (i.e 62 %) respectively. This suggests that if the dataset is larger and contains several features then Lasso provides simpler model compared to the Ridge regression.

6.3.2.4 ElasticNet regression

The ElasticNet regression method is a combination of both Lasso and Ridge regression methods and therefore, has both L1 and L2 regularisations. The regularisation parameter L1 ratio regulates the mix between Lasso and Ridge regression, and it varies between 0 and 1. When L1 = 0, the elastic net regression model behaves similar to Ridge regression and considers all features of the dataset. On the other hand, when L1 = 1 the model considers limited features and acts similar to Lasso regression.
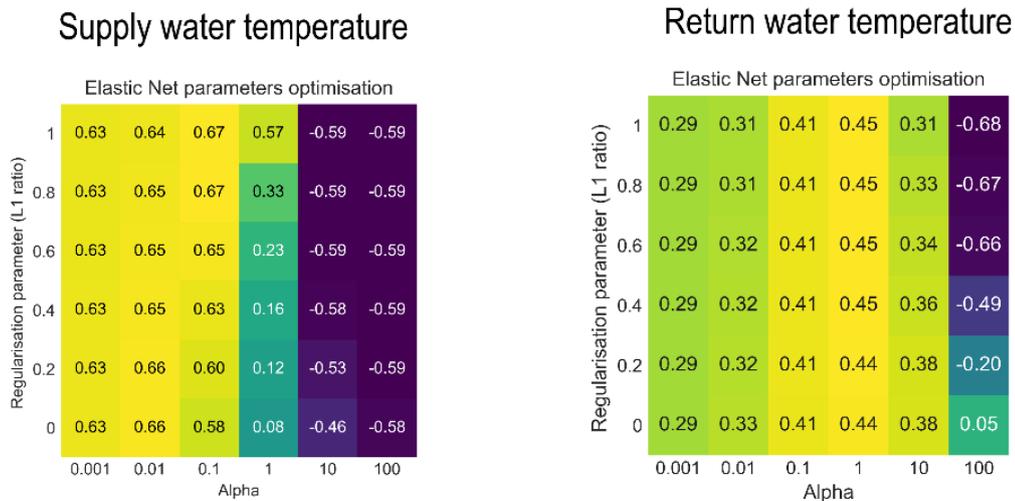
Figure 6.13: Grid Search for parameter optimisation - ElasticNet regression method. The left panel shows parameter grid for supply water temperature data and right panel shows for return water temperature data.

In ElasticNet regression, three parameters needs to be optimised. These three parameters include; alpha, maximum iterations and regularisation parameter L1 ratio. Firstly, the grid search is performed to find the optimum maximum number of iterations. Then, the second grid search was performed

to find the optimum combination of alpha and L1 ratio, while keeping the maximum number iterations constant chosen at first grid search. In our dataset, the maximum number of iterations of 10 leads to the optimum R2 score, this was used to find L1 ratio and alpha. It is found that the elastic net model selects 6 features with optimum parameters as compared to the Lasso in previous Section, which selected only 5 features to perform predictions.
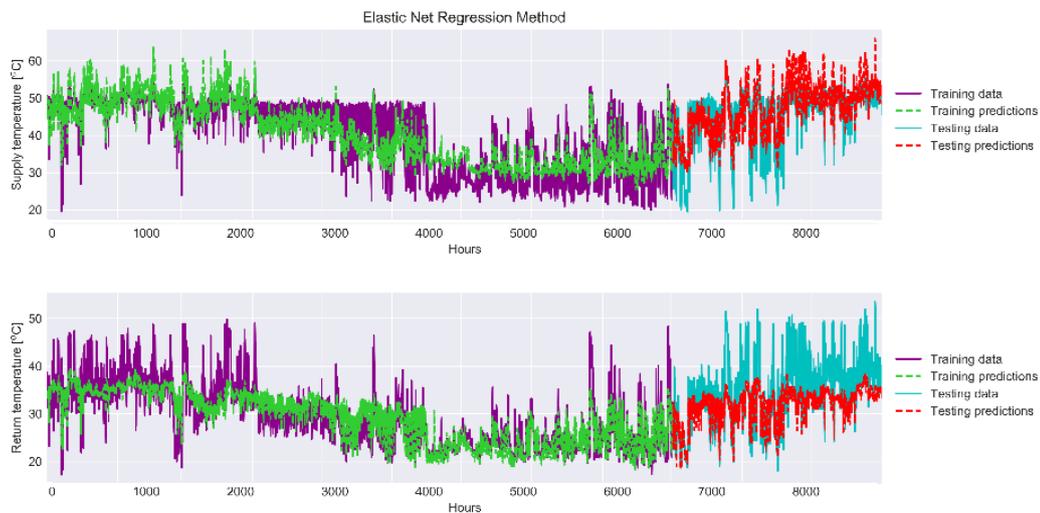
Figure 6.14: Predictions using ElasticNet regression method. The top panel shows the predictions on supply water temperature data and lower panel shows the predictions on return water temperature data. The data in figure presents the hourly prediction results from January to December. Where the Purple and Cyan colour represent actual training and testing data, whereas Green and Red colour represent predictions on the training and testing data.
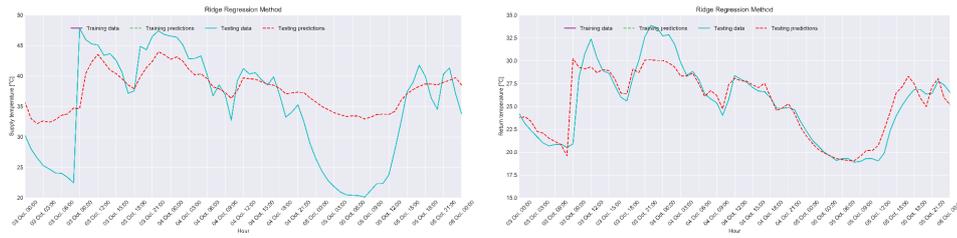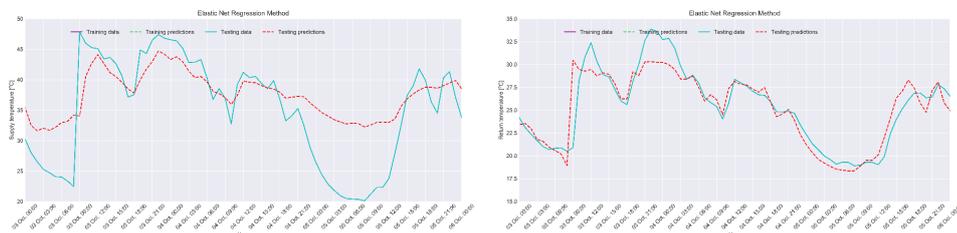

Figure 6.15: 72 hours predictions using ElasticNet regression method. The hourly data in Cyan colour represents testing data and Red colour represents prediction on testing data from 3rdto 5th October. The left panel shows the predictions on supply water temperature data and right panel shows the predictions on return water temperature data.

As shown in Figure 6.13, the results for supply water temperature suggest

that the alpha ($\alpha > 1$) leads to over-fitting of the model and the L1 factor of 0.8 with 6 features in dataset leads to the highest R2 score for the predictions. The Figures 6.14 and 6.15 show that the R2 score for supply and return water temperature is found to be 0.67 (i.e 67 %) and 0.45 (i.e 45 %), respectively.

6.3.2.5   Stochastic gradient descent regression

Stochastic Gradient Descent (SGD) regression is a stochastic method which randomly picks the chunks of training data and computes coefficient slopes for the optimum prediction. This versatile model is unique as it uses different loss functions and penalties to implement different linear models. Using L1 and L2 penalties, it can work either as Lasso or Ridge method respectively. SGD is different from other linear models because it uses stochastic approach to stop iterations on the training data as soon as training error reaches minimum. This is known as early stopping and makes its simulations on large datasets faster compared to other linear methods.

The SGD regression model depends on the learning rate defined in the parameters. SGD regression model considers all features from the data and does not solely depend on the number of iterations or L1 ratio as in Lasso and Ridge regression methods. The results show that increasing in learning rate makes the model to learn more from the training data which results into an improvement in R2 score on predictions. However, after a while the prediction score (R2) starts decreasing which is due to the over-fitting of the model. This is where early stopping comes into play and the optimum parameters for the model are obtained.

As show in Figure 6.16, the results for supply water temperature suggests that the learning rate (eta = 0.01) and alpha ($\alpha = 0.01$) provides the best R2
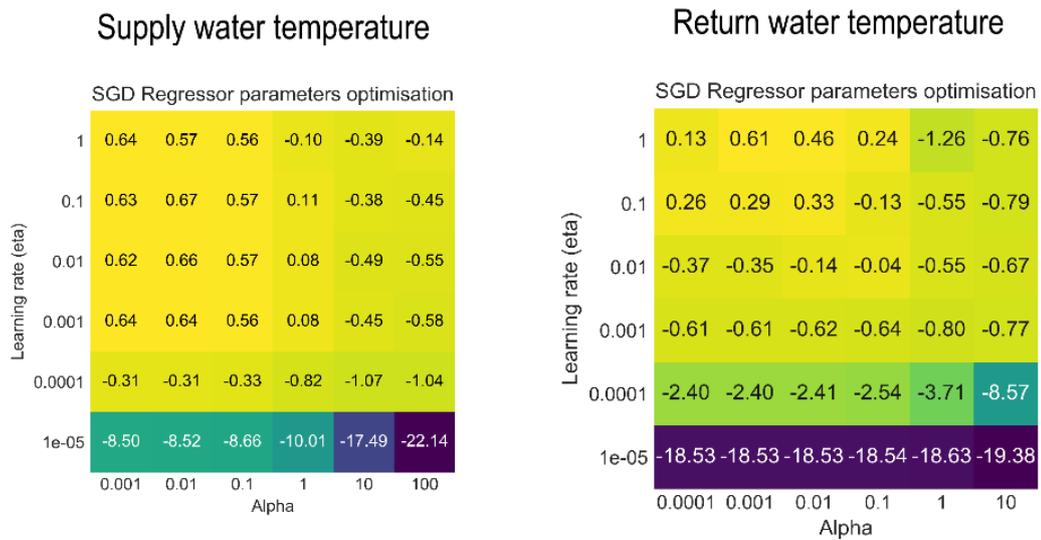
Figure 6.16: Grid Search for parameter optimisation - stochastic gradient descent regression method. The left panel shows parameter grid for supply water temperature data and right panel shows for return water temperature data.
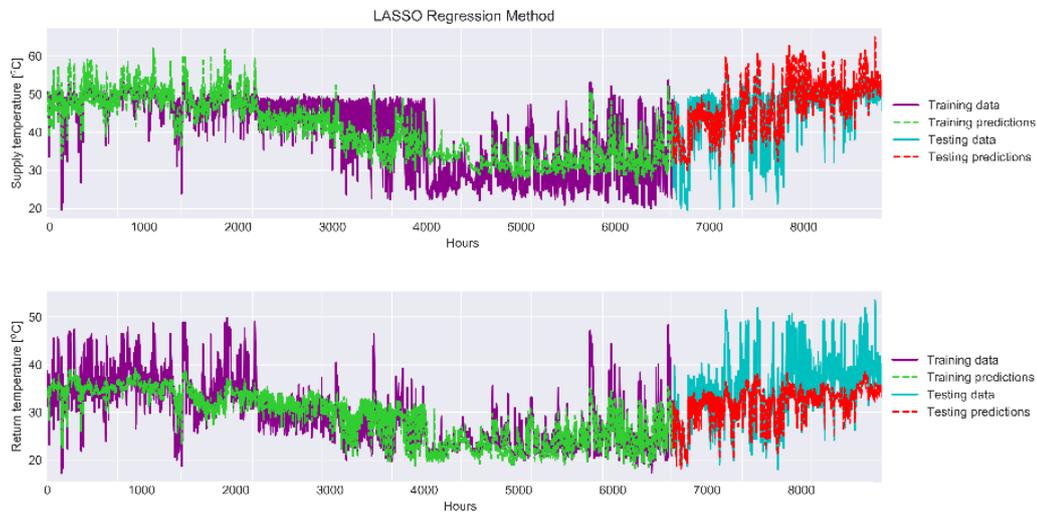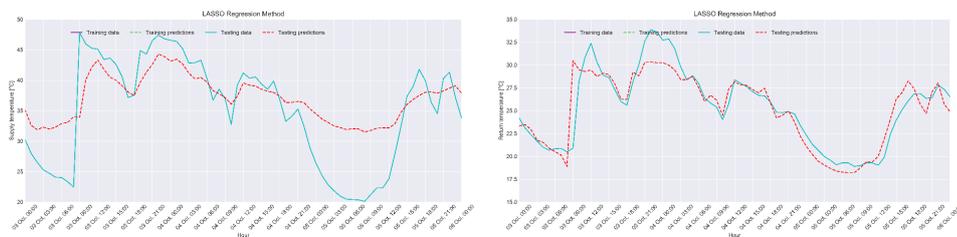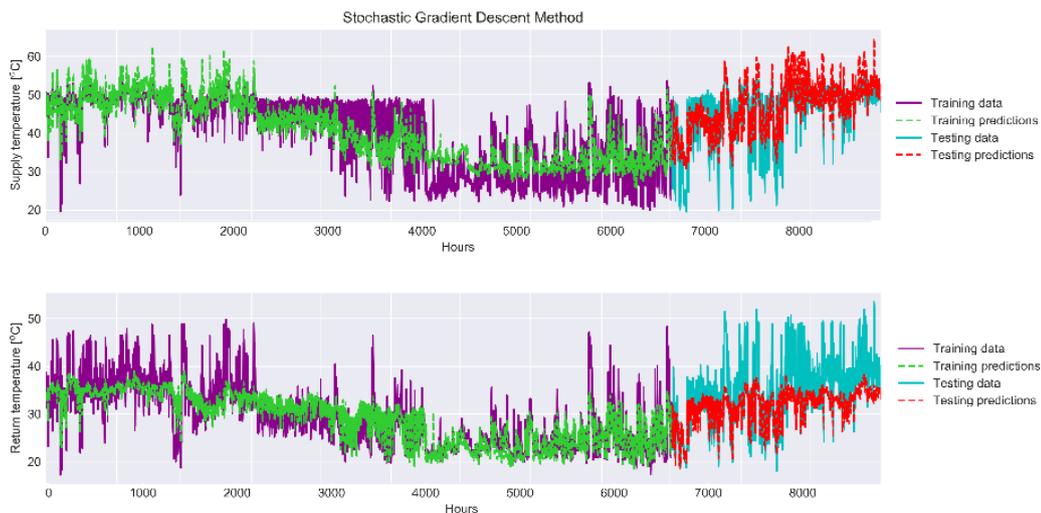

Figure 6.17: Predictions using stochastic gradient descent regression method. The top panel shows the predictions on supply water temperature data and lower panel shows the predictions on return water temperature data. The data in figure presents the hourly prediction results from January to December. Where the Purple and Cyan colour represent actual training and testing data, whereas Green and Red colour represent predictions on the training and testing data.

score of 0.67 (i.e 67 %). Similarly, for return water temperature the learning rate (eta = 1) and alpha ($\alpha = 0.001$) provides the best R2 score of 0.61 (i.e 61 %). The model loses its accuracy as soon as the learning rate goes below
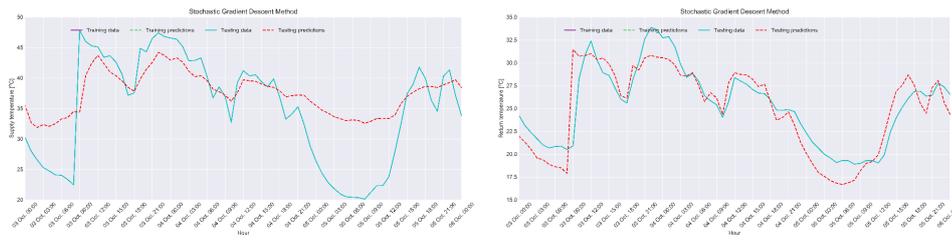
Figure 6.18: 72 hours predictions using stochastic gradient descent regression method. The hourly data in Cyan colour represents testing data and Red colour represents prediction on testing data from 3rd to 5th October. The left panel shows the predictions on supply water temperature data and right panel shows the predictions on return water temperature data.

1e-6. The time-series of these results are further compared in Figures 6.17 and 6.18.

6.3.3 Decision trees

Decision trees are very powerful with the capability of extracting information from a complex non linear data. This is the reason they are widely used in machine learning and forms the basis of several ensemble methods. They work on the principle of a tree where nodes represent questions and leaves contains the answers (predictions). Just like other machine learning methods, decisions trees may also over-fit the data. This can be controlled by either pre-pruning or post-pruning of the tree (more details in Section 2.5.1.3).

In this thesis decision trees are implemented using *scikit-learn* library and pre-pruning is performed by limiting the maximum depth of a tree. If the tree depth is not fixed, it can grow arbitrarily deep and complex, therefore, in this analysis the tree depth is explored from 1 to 10, and random state from 1 to 500. The number of random state is used by the random generator as a seed value for predictions.

Figure 6.19 shows that the R2 score of the model depends on the maximum tree depth and not the random state considered in the model. For supply

Supply water temperature

Decision Tree Regressor parameters optimisation

Return water temperature

Decision Tree Regressor parameters optimisation
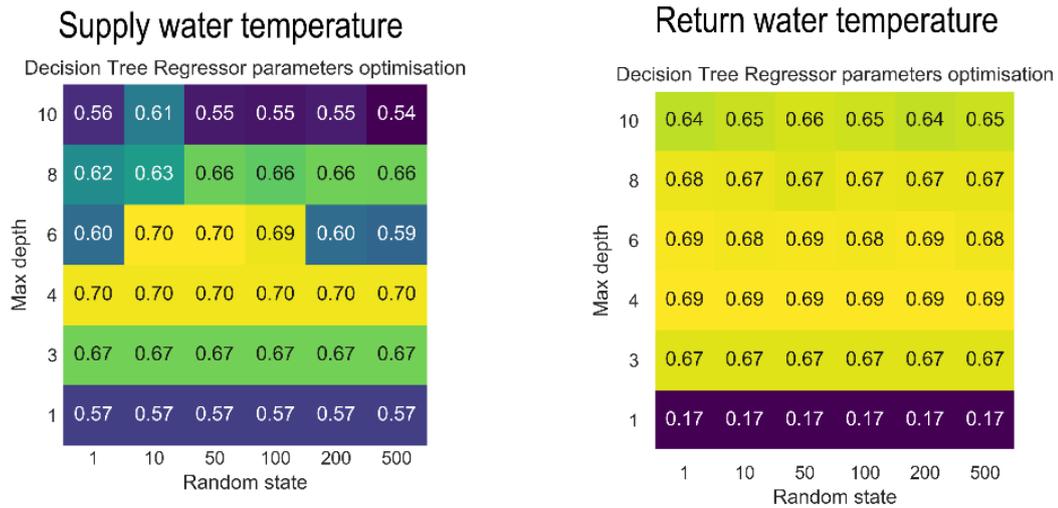
**Figure 6.19:** Grid Search for parameter optimisation - decision trees regression method. The left panel shows parameter grid for supply water temperature data and right panel shows for return water temperature data.
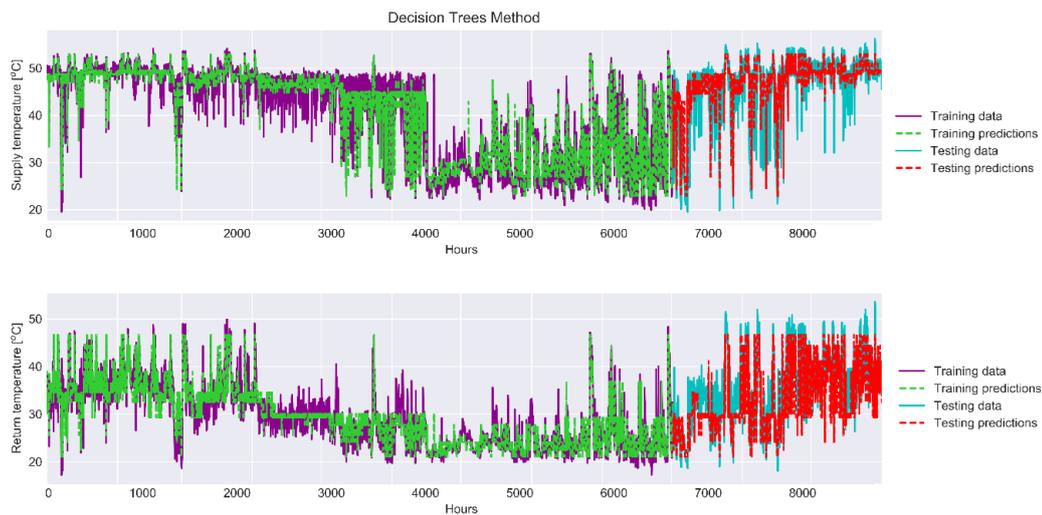
Figure 6.20: Predictions using decision trees regression method. The top panel shows the predictions on supply water temperature data and lower panel shows the predictions on return water temperature data. The data in figure presents the hourly prediction results from January to December. Where the Purple and Cyan colour represent actual training and testing data, whereas Green and Red colour represent predictions on the training and testing data.

water temperature, the optimum R2 score of 0.70 (i.e 70 %) is obtained with tree depth of 6, and stays constant with varying random states. It appears the R2 score increases with tree depth up to a certain limit which is due to under-fitting at lower values of tree depth and over-fitting at higher values.
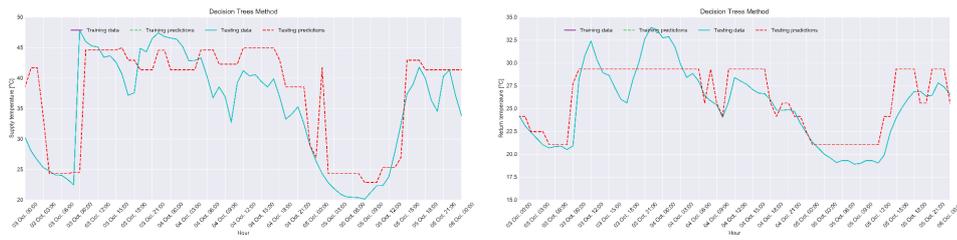
Figure 6.21: 72 hours predictions using decision trees regression method. The hourly data in Cyan colour represents testing data and Red colour represents prediction on testing data from 3rdto 5th October. The left panel shows the predictions on supply water temperature data and right panel shows the predictions on return water temperature data.

For return water temperature, optimum R2 score of 0.69 (i.e 69 %) is obtained with same tree depth. The R2 score is unaffected by different values of random state. It appears that the tree depth of 4 is suitable for return water temperature data with 10 features and helps in tuning the model. In conclusion, the prediction from decision tree models are computationally less intensive and easy to understand because of their simplicity. The time-series of these results are further compared in Figures 6.20 and 6.21.

6.3.4 Ensemble methods

The predictions of different regressors can be aggregated to achieve better results as compared to the individual regressor method. This technique is referred as Ensemble (group) learning, and the algorithm is called Ensemble method. All ensemble methods uses decision trees as building blocks. Among other ensemble methods, the random forest and gradient boosted decision trees are considered to have effective performance on wide range of datasets.

6.3.4.1 Random forests

The random forest is an ensemble method which originates by combining decision trees with bagging method (details in Section 2.5.1.4). Therefore, it

contains parameters for both regressors and always perform better than decision trees on its own.

A random forest model is built by defining the number of tree to create, which is also known as n-estimator parameter and one of the key parameter required for tuning the model. The significance of each feature in the dataset is evaluated by plotting mean feature importance graph in the *Scikit learn*,this represents the most commonly used features by these trees and varies between 0 and 1, where 0 means that feature information is not utilised by the trees and 1 means the relative feature is used.
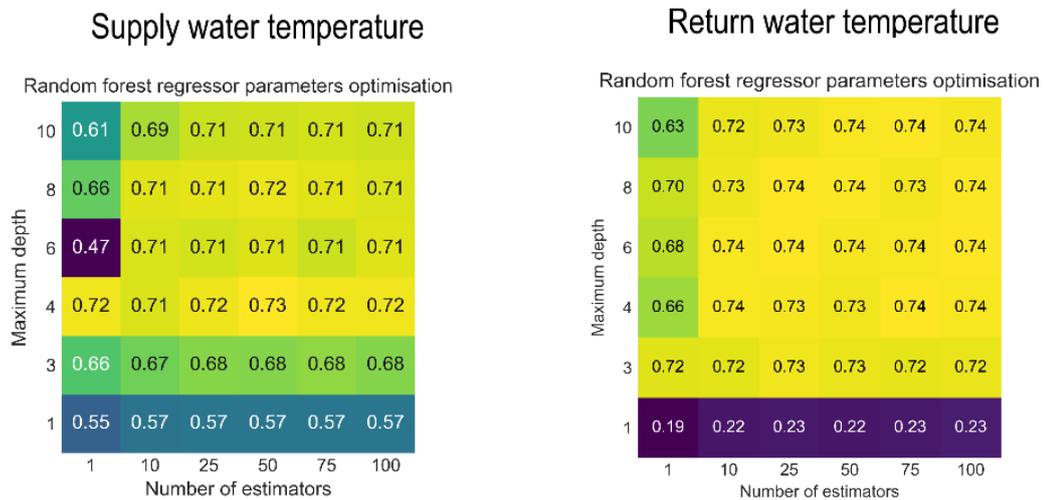
Figure 6.22: Grid Search for parameter optimisation - random forests regression method. The left panel shows parameter grid for supply water temperature data and right panel shows for return water temperature data.

Figure 6.22 shows that the R2 score of the model depends on the maximum depth of tree and number of trees (i.e n-estimator). The model is initially under-fitted at lower tree depth but once it reaches at the depth of 4 and number of trees to 10, then the R2 score considerably improves. For supply water temperature, the optimum R2 score of 0.72 (i.e 72 %) is obtained with number of trees 25 and tree depth of 6. For return temperature, the R2 score of 0.74 (i.e 74 %) is achieved with same number of trees but with the tree
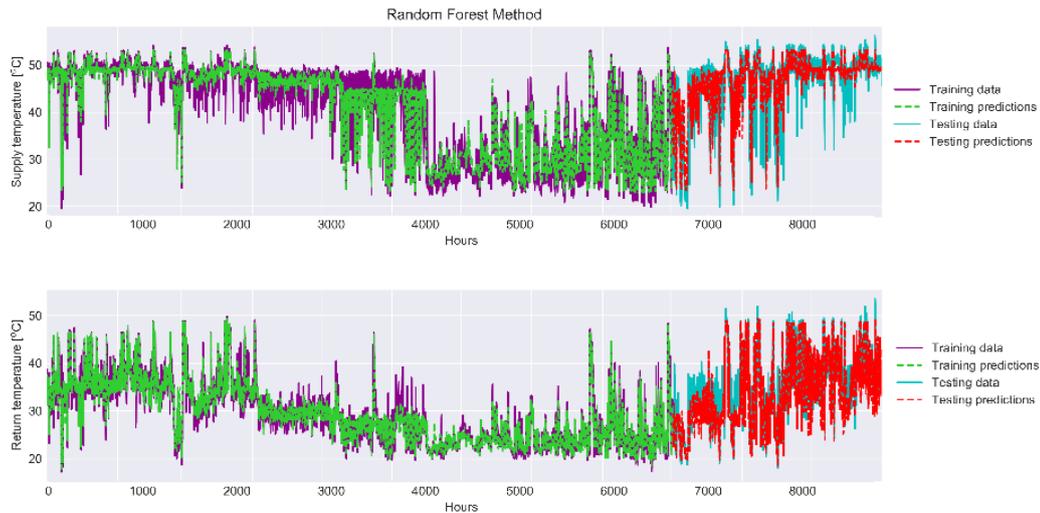
Figure 6.23: Predictions using random forests regression method. The top panel shows the predictions on supply water temperature data and lower panel shows the predictions on return water temperature data. The data in figure presents the hourly prediction results from January to December. Where the Purple and Cyan colour represent actual training and testing data, whereas Green and Red colour represent predictions on the training and testing data.
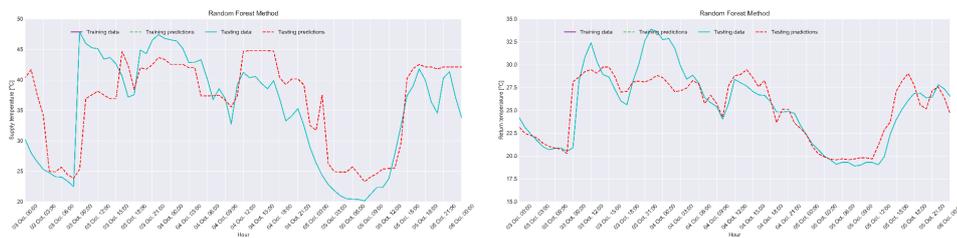
Figure 6.24: 72 hours predictions using random forests regression method. The hourly data in Cyan colour represents testing data and Red colour represents prediction on testing data from 3rd to 5th October. The left panel shows the predictions on supply water temperature data and right panel shows the predictions on return water temperature data.

depth of 8. It appears that the estimator of 25 is suitable for our dataset with 10 features, and a key parameter in tuning the model. The time-series of these results are further compared in Figures 6.23 and 6.24.

In summary, random forests are very robust and powerful methods. In particular, they are useful for understanding the important features and helps in narrowing down the feature selection from high-dimensional sparse data.

6.3.4.2 Gradient boosting regression trees

Gradient Boosting Regression Trees (GBRT) is a type of ensemble method which works by constructing trees in a sequential manner, where residual error from the predecessor tree is used to minimise the prediction error. In addition to the parameter for number of trees (n-estimator) and tree depth in random forest, the GBRT method uses the learning rate for accurate predictions. The learning rate is a key parameter which controls the correction of errors from the previous trees.

Figure 6.25 shows that the R2 score of the model depends on the learning rate and number of trees (i.e n-estimator). The model is initially under-fitted at lower learning rate compared to the number of trees, but once it reaches at 0.1, then the optimum R2 score is achieved.
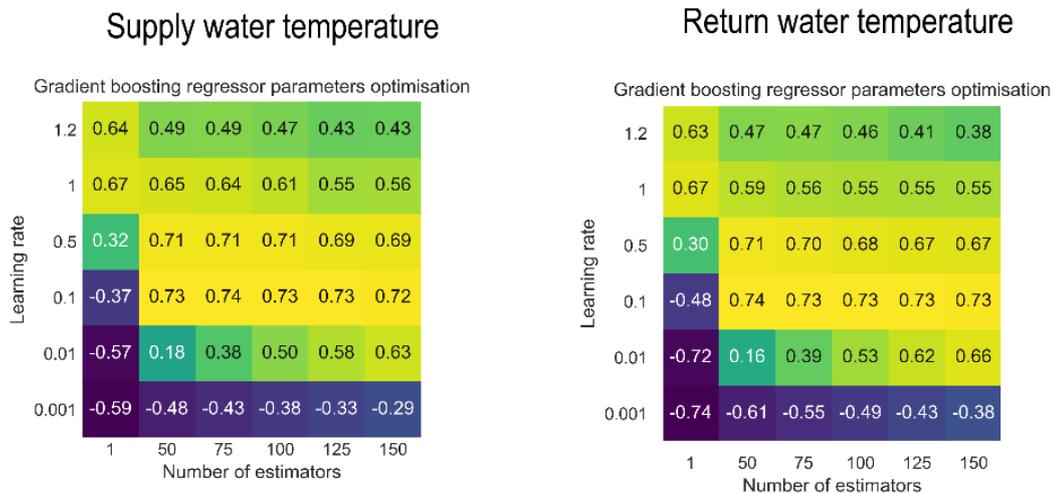
Figure 6.25: Grid Search for parameter optimisation - gradient boosting regression trees method. The left panel shows parameter grid for supply water temperature data and right panel shows for return water temperature data.

For supply water temperature, the optimum R2 score of 0.74 (i.e 74 %) is obtained with leaning rate of 0.1 and number of trees of 75. For return temperature, similar R2 score is achieved with same learning rate but 50 number of trees. It appears that the learning rate of 0.1 is suitable for our dataset and
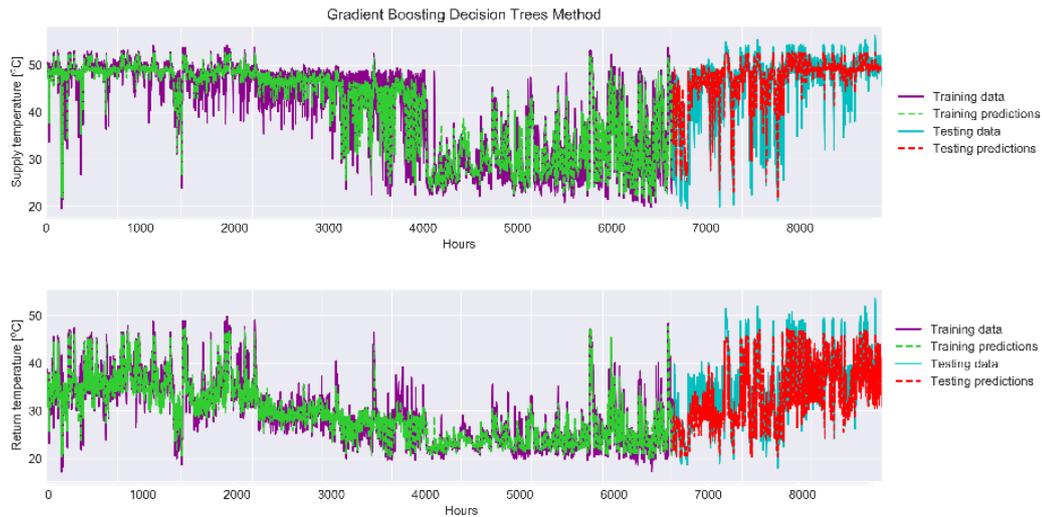
Figure 6.26: Predictions using gradient boosting regression trees method. The top panel shows the predictions on supply water temperature data and lower panel shows the predictions on return water temperature data. The data in figure presents the hourly prediction results from January to December. Where the Purple and Cyan colour represent actual training and testing data, whereas Green and Red colour represent predictions on the training and testing data.

Figure 6.27: 72 hours predictions using gradient boosting regression trees method. The hourly data in Cyan colour represents testing data and Red colour represents prediction on testing data from 3rd to 5th October. The left panel shows the predictions on supply water temperature data and right panel shows the predictions on return water temperature data.

a key parameter in tuning the model. It is observed that the learning rate and number of trees are interconnected as higher learning rate with higher number of trees leads to over-fitting and vice versa. The time-series of these results are further compared in Figures 6.26 and 6.27.

In summary, gradient boosted methods tend to be more accurate than random forest and other ensemble methods, due to the sensitive parameter

tuning and optimisation. Their predictions are mostly better with low learning rate, number of trees and tree depth.

6.3.5 Support vector machines

The Support Vector Machines (SVM) also known as kernelized support vector machines are extended methods which construct a set of hyperplanes in a high dimensional space for more complex models (not defined in the input space). A simple SVM model represents data examples in space as 2D points where different classes of data are separated by a clear gap. New data are then mapped into that gap and predictions are made on the basis of class they fall on.

These are very powerful and versatile machine learning methods which are well suited for regression and classification of complex but medium sized datasets. There are several variation of SVM depending on different types of kernels. These include linear, polynomial and RBF (Gaussian) kernels.

6.3.5.1 Linear kernel

The SVM linear kernel creates a linear margin to separate the data with the goal to create a widest margin between the decision boundaries. The width of this margin is controlled by tuning hyperparameter C of the linear kernel, where higher values of C calculates smaller width.

The Figure 6.28 shows that the results depends on the regularisation parameter C and the R2 score increases with an increase in the value of C. Once the optimum R2 score is reached and model starts to over-fit and under-fit at higher values of C. The optimum R2 scores of 0.74 (74 %) and 0.68 (68 %) for supply and return water temperature are obtained with a C of 10 and 1,

Supply water temperature

Return water temperature

Figure 6.28: Grid Search for parameter optimisation - support vector regression - linear kernel method. The left panel shows parameter grid for supply water temperature data and right panel shows for return water temperature data.

respectively. The time-series of these results are further compared in Figures 6.29 and 6.30.

Figure 6.29: Predictions using support vector regression - linear kernel method. The top panel shows the predictions on supply water temperature data and lower panel shows the predictions on return water temperature data. The data in figure presents the hourly prediction results from January to December. Where the Purple and Cyan colour represent actual training and testing data, whereas Green and Red colour represent predictions on the training and testing data.

Due to their linear nature, the application of these models in low-dimensional spaces is limited. Compared to linear models the SVM linear
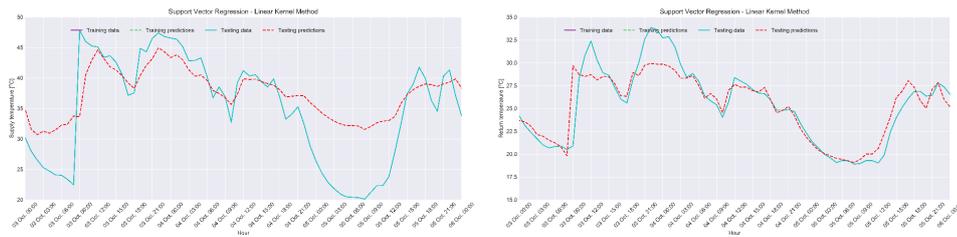
Figure 6.30: 72 hours predictions using support vector regression - linear kernel method. The hourly data in Cyan colour represents testing data and Red colour represents prediction on testing data from 3rd to 5th October. The left panel shows the predictions on supply water temperature data and right panel shows the predictions on return water temperature data.

models are slower, computationally more expensive, sensitive to data scaling and parameters.

6.3.5.2   Polynomial kernel

The polynominal kernel is a type of non-linear model for solving the non-separable datasets. It generates huge number of features using parameter *degree* of the kernel. This is also referred as the *kernel trick* in SVM. This enables to achieve the same performance of the model as if many polynomial features were added without actually having to add them. Since these features were not added at the first place, therefore they do not cause any combinatorial explosion. All possible polynomials up to a certain degree of original features are calculated by this kernel.

The Figure 6.31 shows that results depend on the regularisation parameter C and *degree* of the polynomial kernel. The model is initially under-fitted but, once the regularisation parameter C and *degree* exceeds 1, then R2 score increases tremendously. The optimum R2 scores of 0.87 (87 %) and 0.84 (84 %) for supply and return water temperature are obtained with a C of 10,000 and *degree* of 3, respectively.

It appears that the high regularisation parameter C and the *degree* are
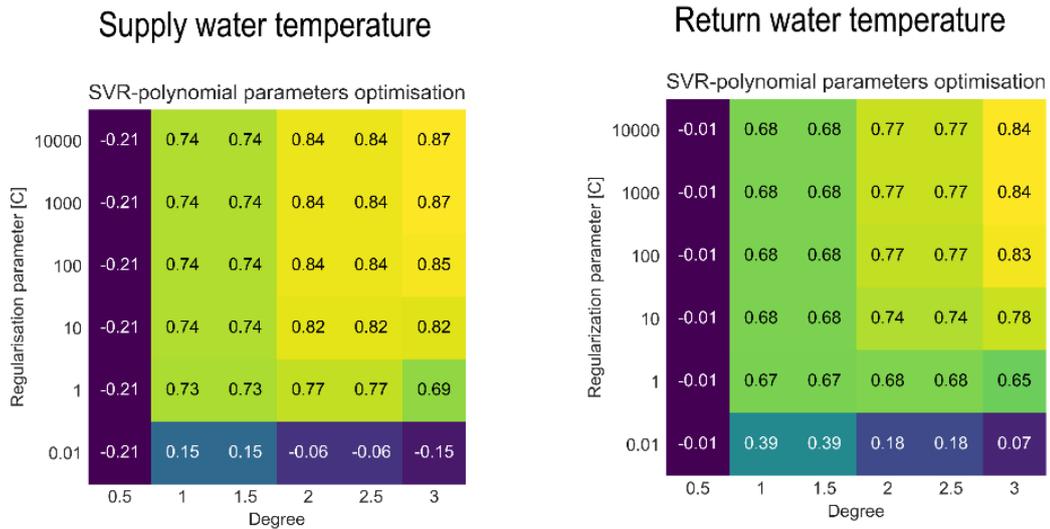
Figure 6.31: Grid Search for parameter optimisation - support vector regression - polynomial kernel method. The left panel shows parameter grid for supply water temperature data and right panel shows for return water temperature data.

Figure 6.32: Predictions using support vector regression - polynomial kernel method. The top panel shows the predictions on supply water temperature data and lower panel shows the predictions on return water temperature data. The data in figure presents the hourly prediction results from January to December. Where the Purple and Cyan colour represent actual training and testing data, whereas Green and Red colour represent predictions on the training and testing data.
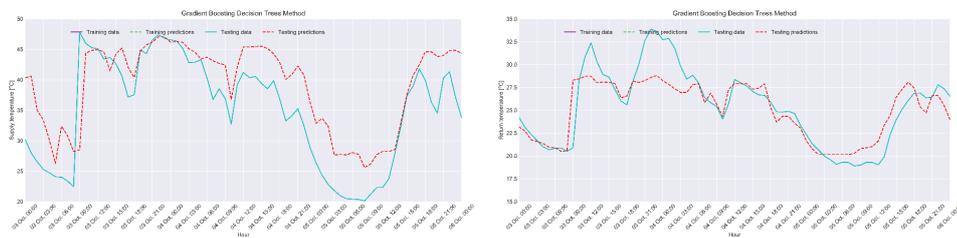
the key parameter in tuning of the model. The time-series of these results are

further compared in Figures 6.32 and 6.33.
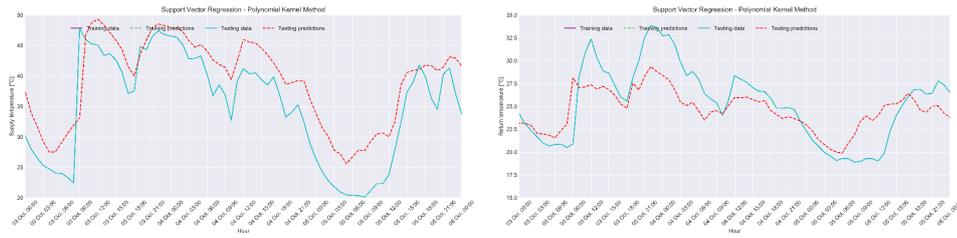
Figure 6.33: 72 hours predictions using support vector regression - polynomial kernel method. The hourly data in Cyan colour represents testing data and Red colour represents prediction on testing data from 3rd to 5th October. The left panel shows the predictions on supply water temperature data and right panel shows the predictions on return water temperature data.

6.3.5.3 Radial basis function kernel

The Gaussian or radial basis function (RBF) kernel finds the relation by calculating the Euclidean distance between two data points in an infinite-dimensional feature space. The different values of hyperparameters C and *gamma* controls the shape and width of the decision boundary of the Gaussian kernel. The smaller values of the *gamma* causes the decision boundary curve wider and vice versa. On the other hand, the regularisation hyperparameters C regulates the over and under-fitting of the model.
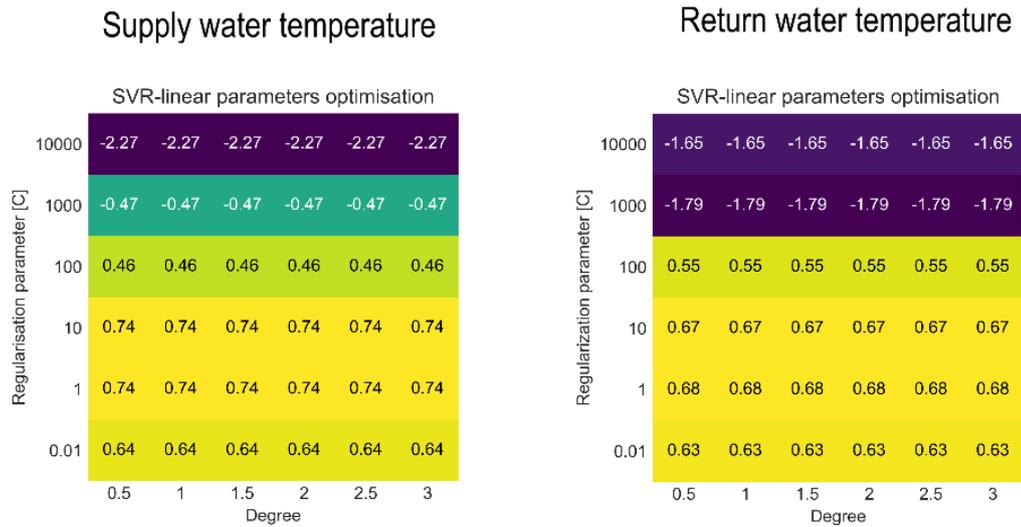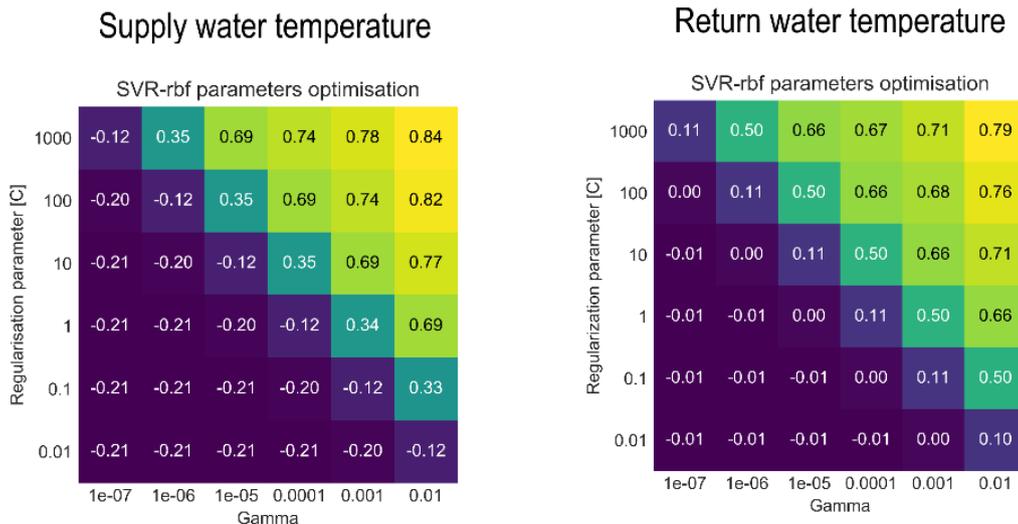

Figure 6.34: Grid Search for parameter optimisation - support vector regression - radial basis function kernel method. The left panel shows parameter grid for supply water temperature data and right panel shows for return water temperature data.

The Figure 6.34 shows that results depend on the regularisation parameter *C* and *gamma*. The model is initially under-fitted but once the regularisation parameter *C* and *gamma* increases, the R2 score also starts increasing. The optimum R2 scores of 0.84 (84%) and 0.79 (79%) for supply and return water temperature are obtained with a *C* of 1000 and *gamma* of 0.01, respectively.
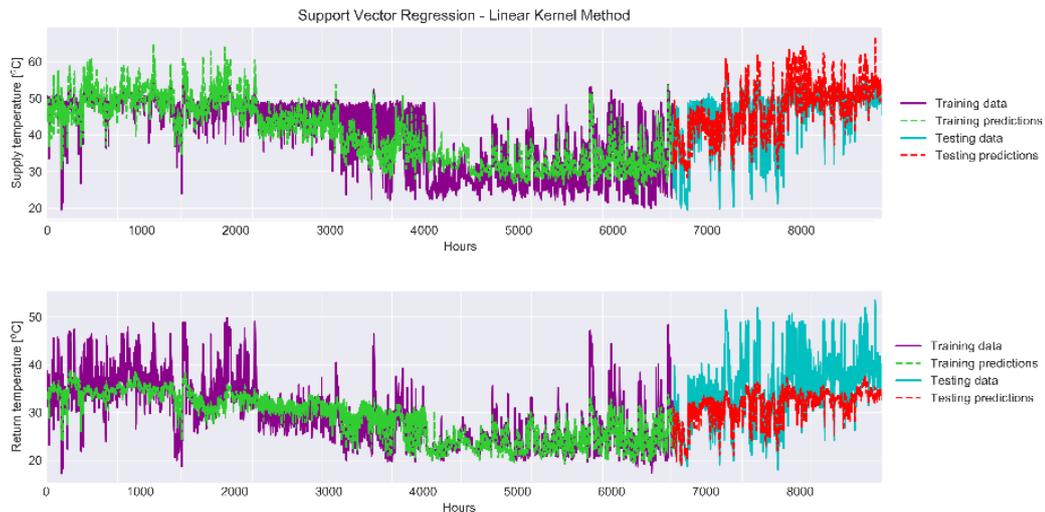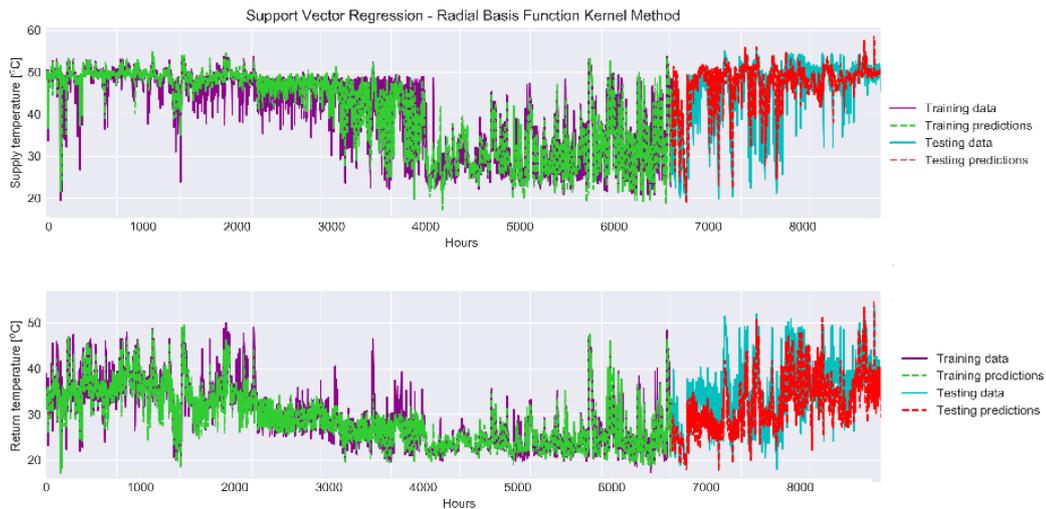

Figure 6.35: Predictions using support vector regression - radial basis function kernel method. The top panel shows the predictions on supply water temperature data and lower panel shows the predictions on return water temperature data. The data in figure presents the hourly prediction results from January to December. Where the Purple and Cyan colour represent actual training and testing data, whereas Green and Red colour represent predictions on the training and testing data.
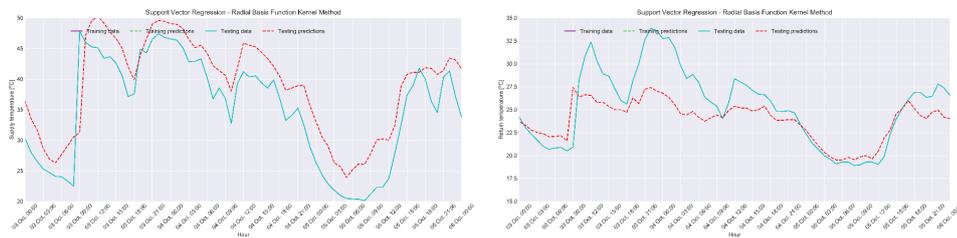
Figure 6.36: 72 hours predictions using support vector regression - radial basis function kernel method. The hourly data in Cyan colour represents testing data and Red colour represents prediction on testing data from 3rd to 5th October. The left panel shows the predictions on supply water temperature data and right panel shows the predictions on return water temperature data.

It appears that the higher regularisation parameter *C* and the *gamma*

values have considerable impact in tuning of the model. The time-series of these results are further compared in Figures 6.35 and 6.36.

6.3.6   Neural networks - multi-layer perceptron

The Multi-layer Perceptron (MLP) class of neural networks can be viewed as generalisations of linear models that collects information from multiple stages to reach at a decision. The model has to learn several coefficients or weights across three layers i.e. input, hidden and output layer.

The neural networks can be optimised by tuning several different parameters. This offers flexibility as well as difficulty to choose the correct parameters. However, most commonly tuned parameters are; the number of hidden layers, number of nodes in the hidden layer and the regularisation parameter (alpha).

In order to compute weighted sum for each of hidden layer, an activation function (*relu* or *tanh*) is used which contributes in the final weight or score. These functions enable neural networks to perform better than a simple learning method by learning more. It is observed that *relu* activation function works better on current time-series data.

The Figure 6.37 shows that the model's complexity increases with increase in number of hidden layers and results depends on the regularisation parameter (alpha) and maximum number of iterations. It is observed that the MLP neural network model under-fits at higher value of alpha parameters for both supply and return water temperature analysis. The analysis is performed using 'adam' solver with maximum number of iterations for the supply and return water temperature are fixed at 1000 and 100, respectively. The optimum R2 scores of 0.80 (80%) and 0.78 (78%) for supply and return water temperature

are obtained at hidden layers of 100 and alpha of 0.01 and 10, respectively.

The time-series of these results are further compared in Figures 6.38 and 6.39.
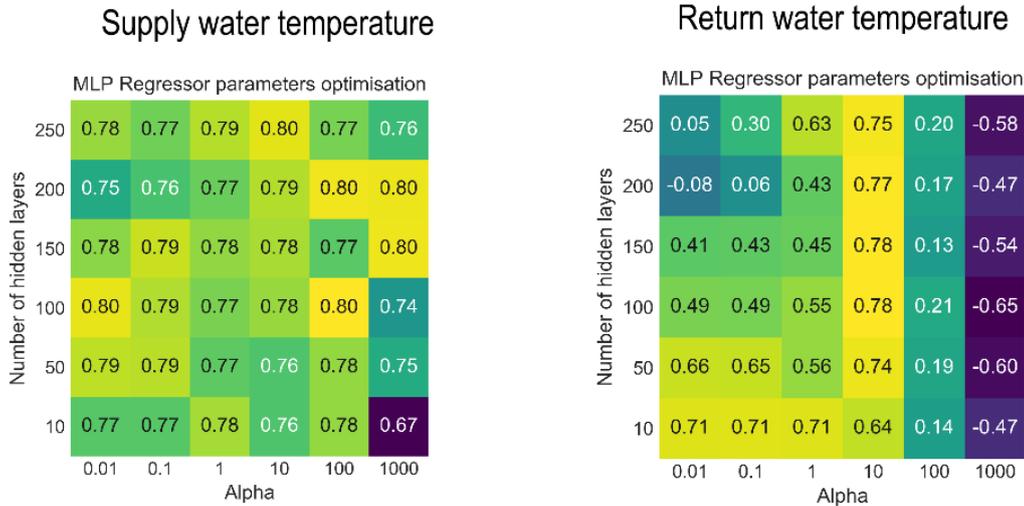
Figure 6.37: Grid Search for parameter optimisation - neural networks - multilayer perceptron method. The left panel shows parameter grid for supply water temperature data and right panel shows for return water temperature data.

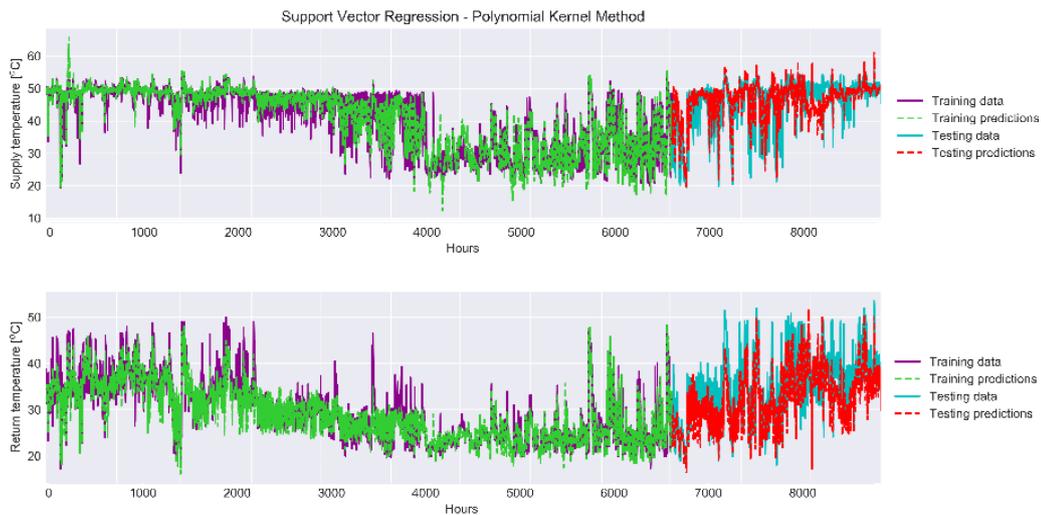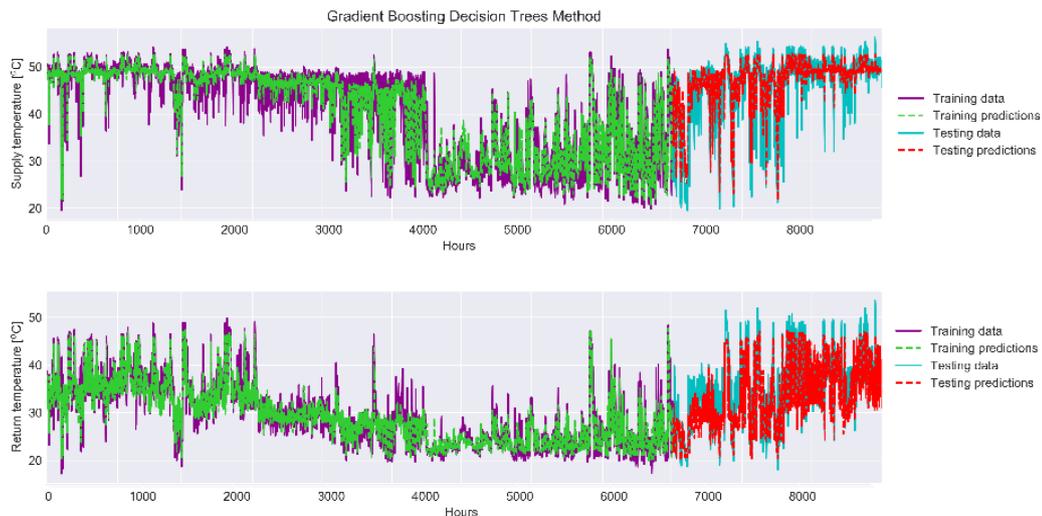**Figure 6.38:** Predictions using neural networks - multi-layer perceptron method. The top panel shows the predictions on supply water temperature data and lower panel shows the predictions on return water temperature data. The data in figure presents the hourly prediction results from January to December. Where the Purple and Cyan colour represent actual training and testing data, whereas Green and Red colour represent predictions on the training and testing data.

In summary, MLP neural network models perform better compared to other supervised learning models, but they can be very complex specially for
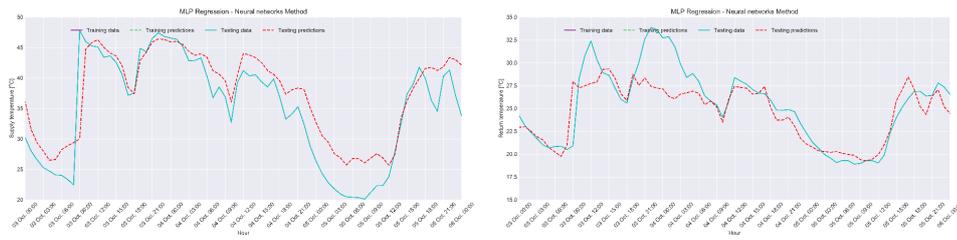
Figure 6.39: 72 hours predictions using neural networks - multi-layer perceptron method. The hourly data in Cyan colour represents testing data and Red colour represents prediction on testing data from 3rd to 5th October. The left panel shows the predictions on supply water temperature data and right panel shows the predictions on return water temperature data.

large datasets, with more number of hidden layers. Though, the computational

time of neural networks is smaller than the support vector machine models,

the optimisation of parameters for different solvers is challenging. This takes

large amount of time to train new models with large datasets.

6.4   Time-series forecasting

In this section, the classical stochastic method, Auto-Regressive Integrated Moving Average (ARIMA) is discussed for time-series forecasting analysis. It is a widely used forecasting method which is combination of Auto-Regressive (AR) and Moving Average (MA) models.
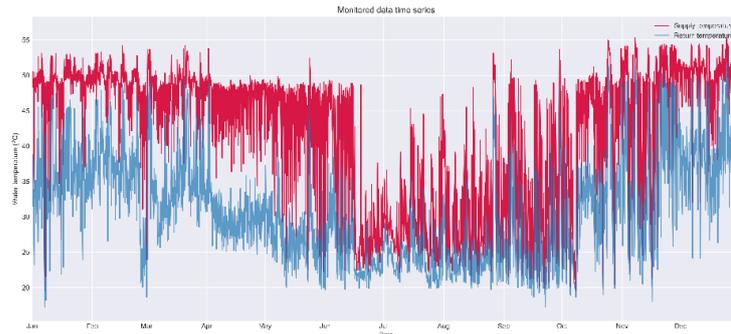
Figure 6.40: 1 year of hourly monitored supply and return water temperature time-series data used for forecasting

6.4.0.1   Stationarity of time-series data

As discussed in Section 2.5.2, it is a pre-condition that the time-series should be stationary before the implementation of forecasting methods. A time-series is stationary once its trend and seasonality has been removed. To check this, Augmented Dickey Fuller (ADF) test is used. There are two main criteria used to evaluate the results predicted from ADF test. Firstly, the P-value should be less than 0.05, and secondly the test statistic should be less than the 10% critical value. If both of these conditions are met then this confirms that time-series is stationary.

In order to determine the stationarity of the time-series data (shown in Figure 6.40), the Augmented Dickey Fuller (ADF) test from Python's statsmodel library has been used. Its results for the supply and return water temperature data shows that the P-value is less than 0.05 but test statistic is

not less than the 10% critical value, which confirmed that our time-series data

is not stationary and have trend and seasonality which needs to be removed.

There are two main methods used to investigate seasonality from complex

datasets. These include decomposition and differencing methods.
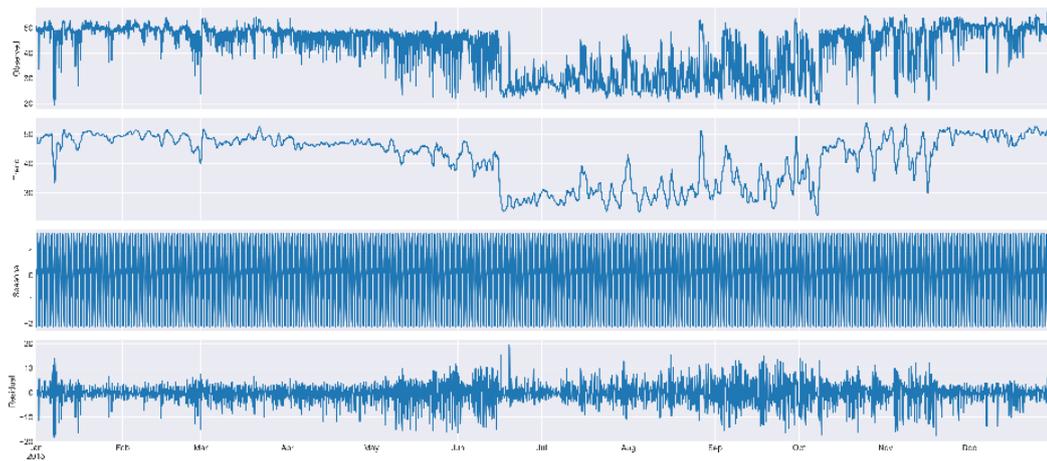
Figure 6.41: Results from seasonal decompose function on supply water temperature data. First row represents the observed time series, second and third rows represent the hidden trend and seasonality in the data. The last row represents the residual data before and after removal of trend and seasonality. This residual data is used by forecasting methods for model development.
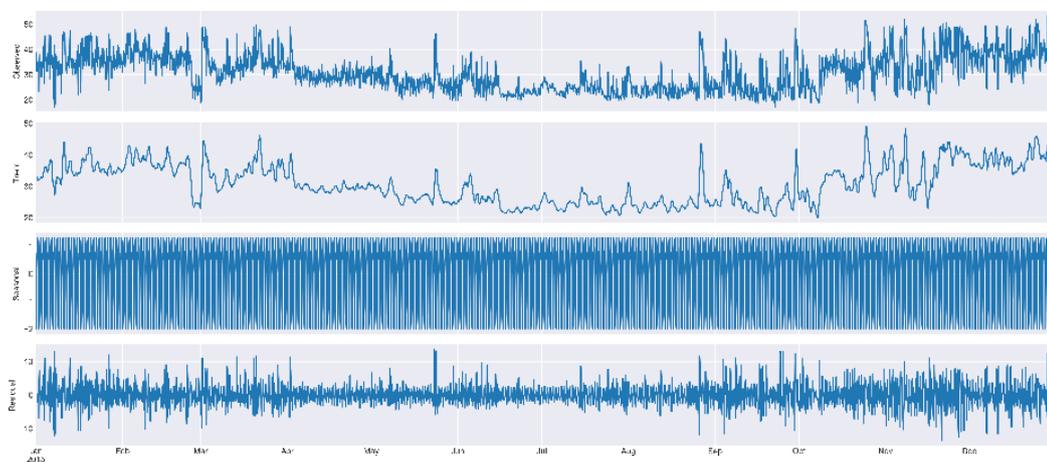
Figure 6.42: Results from seasonal decompose function on return water temperature data. First row represents the observed time series, second and third rows represent the hidden trend and seasonality in the data. The last row represents the residual data before and after removal of trend and seasonality. This residual data is used by forecasting methods for model development.

Firstly, in this analysis, seasonal decompose function, from *Python's Statsmodels library*, is used to extract trend, seasonality and residuals from time-series of the supply and return water temperature data (as shown in Figures 6.41 and 6.42). It is observed that the mean and standard deviation of the supply and return water temperature data varies throughout the year. Their trend is high in winter but low in summer, and contain high amount of constant seasonality. The residual data is used by the forecasting models, as it is the original data once the trend and seasonality has been removed. These results have helped to decide that which method (ARIMA or SARIMA) should be fitted to the data. Since, the data contains seasonality, the SARIMA method should be used.
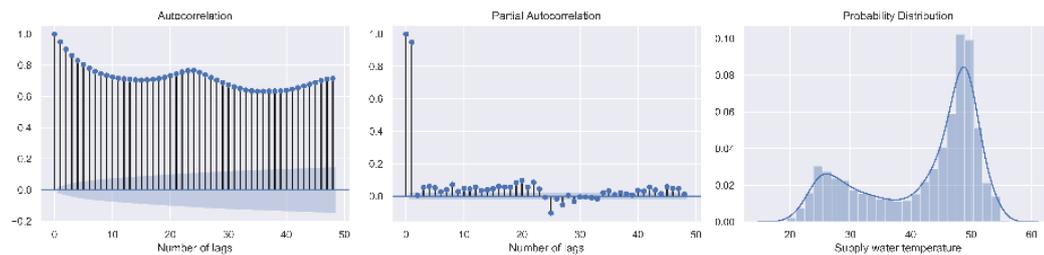
Figure 6.43: Results from Autocorrelation (ACF) and partial autocorrelation function (PACF) plots on supply water temperature data. The ACF and PACF plots are used to investigate the transformed data after differencing. The blue shaded area represents the significance interval and the lagged correlation coefficients located within this region are not critical for the analysis.
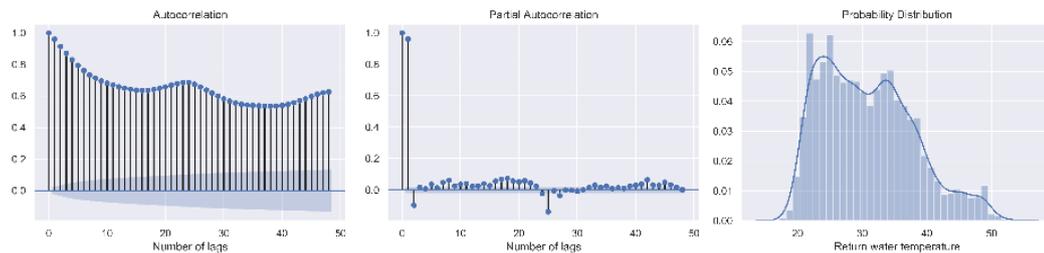
Figure 6.44: Results from Autocorrelation (ACF) and partial autocorrelation function (PACF) plots on return water temperature data. The ACF and PACF plots are used to investigate the transformed data after differencing. The blue shaded area represents the significance interval and the lagged correlation coefficients located within this region are not critical for the analysis.

Secondly, the differencing method is used to remove the seasonality form the data. The autocorrelation (ACF) and partial autocorrelation function (PACF) plots are used to investigate the transformed data after differencing. The blue shaded area (in ACF and PACF plots) represents the significance interval and the lagged correlation coefficients located within this region are not critical for the analysis. The ACF plot shows that both supply and return water temperature time-series are highly correlated at the lag = 24 i.e. correlation(y_t, $y_t - 24$), which explains that the hourly data is correlated at the same hour every day (Figure 6.43 and 6.44).
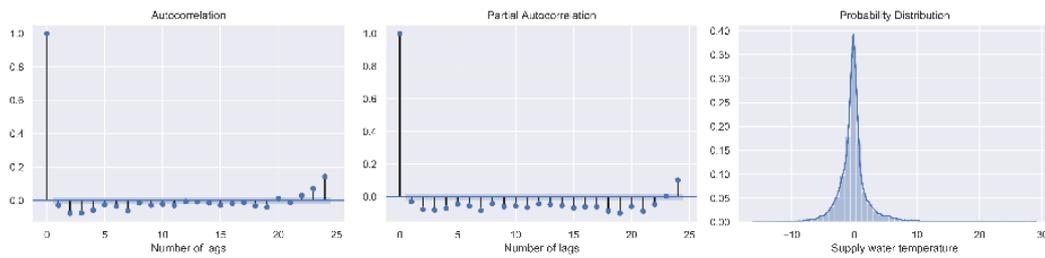
Figure 6.45: Results from Autocorrelation (ACF) and partial autocorrelation function (PACF) plots on supply water temperature data after 1st differencing.
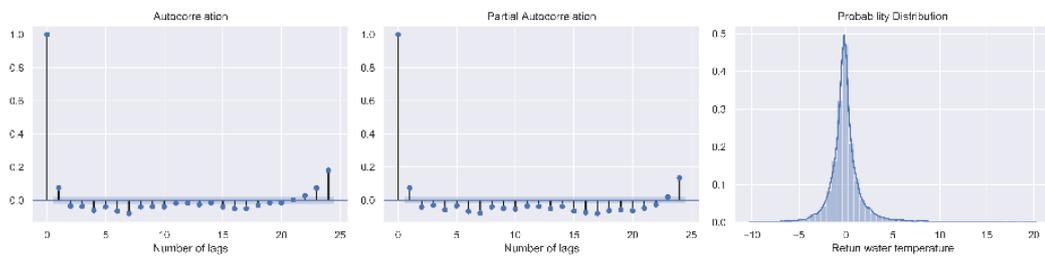
Figure 6.46: Results from Autocorrelation (ACF) and partial autocorrelation function (PACF) plots on return water temperature data after 1st differencing.

The PACF calculates the partial correlation between time-series and its own lag values, and is unaffected by any other lags. The PACF plot shows that lag = 1 is highly correlated, and 1st-order differencing will be sufficient to remove the seasonality from these time-series. Considering these results, 1st-order differencing is performed, and the autocorrelation (ACF) at lag =1 is found to be negative or lies around the significant interval, as shown in Figures

6.45 and 6.46. This confirms that PACF results discussed above and the data

is not over-differenced.

Table 6.1: Results from Augmented Dickey Fuller (ADF) test

Component	No differencing	1st differencing	2nd differencing
Supply water temperature			
ADF test statistic	-2.16	-12.64	-16.86
P-value	0.219949	0	0
Number of lag used	120	120	120
Number of observations used	8639	8639	8639
Critical value (1%)	-3.43	-3.43	-3.43
Critical value (5%)	-2.86	-2.86	-2.86
Critical value (10%)	-2.56	-2.56	-2.56
Outcome - Time Series is	Non-Stationary	Stationary	Stationary
Return water temperature			
ADF test statistic	-2.81	-12.68	-17.01
P-value	0.05	0	0
Number of lag used	120	120	120
Number of observations used	8639	8639	8639
Critical value (1%)	-3.43	-3.43	-3.43
Critical value (5%)	-2.86	-2.86	-2.86
Critical value (10%)	-2.56	-2.56	-2.56
Outcome	Non-Stationary	Stationary	Stationary

After the application of differencing method, the ADF test is performed

again to confirm if trend and seasonality has been removed from the time-

series. This time, both the conditions are fulfilled and the dataset is found to

be stationary, as shown in table 6.1. In this analysis, the differencing is used

in SARIMA method to forecast the time-series.

6.4.1 Autoregressive integrated moving average (ARIMA)

ARIMA is a linear regression based forecasting method which integrates both

autoregressive (AR) and moving average (MA). The term I represents the

order of differencing, which provides the capability to capture temporal effects

of the time-series data. The ARIMA model with hyperparameters can be

represented as $ARIMA(p,d,q)$, where p is the number of lag observations (past time values); d is number of differencing order; and q is size of moving average window.

In summary, the ARIMA is a versatile method, where a model can behave either as AR, ARI, IMA, MA, ARMA, or ARIMA based on the selected hyperparameters p,d and q. Despite this flexibility, ARIMA has limitations to handle data with seasonality, and the method discussed below is used to overcome this limitation.

6.4.2 Seasonal autoregressive integrated moving average (SARIMA)

The Seasonal Autoregressive Integrated Moving Average (SARIMA) is an extension to ARIMA, which can handle both trend and seasonal component of a time-series. This method has hyperparameters $(p,d,q)(P,D,Q)m$, where p,d and q are same as for the ARIMA, while P is the number of seasonal lag observations; D is seasonal differencing order; Q is size of seasonal moving average window, and m is time interval of repeating seasonal pattern. The seasonal terms are similar to the other terms and repeating of past seasonal values.

6.4.2.1 Forecasting using SARIMA method

The SARIMA model is implemented using the *Python's Statsmodels library.* The seasonal parameter m has been set to constant and rest of the parameters are tuned by using *auto.arima* function from the *Python's Pypyramid library*, instead of using ACF and PCAF plots for tuning of individual hyperparameters. The function iterates using several combinations of hyperparameters and provides the optimum parameters with lowest Akaike Information

Criterion (AIC) value for the SARIMA model, as shown in Figure 6.47. The AIC provides an estimate for the goodness of the statistical model for given data and used for the comparison for model selection (Chen et al. 2018).
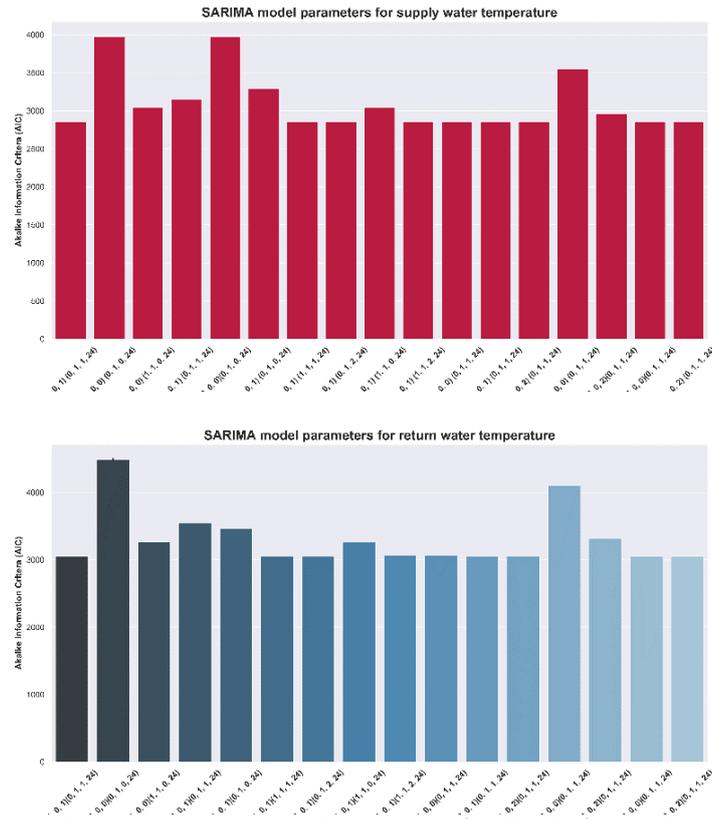
Figure 6.47: The optimum parameters with lowest Akaike information criterion (AIC) value for the SARIMA model. The top panel represent results for supply water temperature, whereas the lower panel represent results for return water temperature.

It is found that the for supply water temperature the SARIMA model $(1,0,1)(1,1,2)24$ is optimum with AIC of 2853, and for return water temperature the SARIMA model $(2,0,1)(0,1,1)24$ is with AIC of 3054. The comparison between other parameters for forecasting is shown in Figure 6.47.

The above optimum parameters for SARIMA are used to perform one week forecast. The data for the last week of December is selected to test the forecasting on supply and return water temperature. For both, Figure 6.48
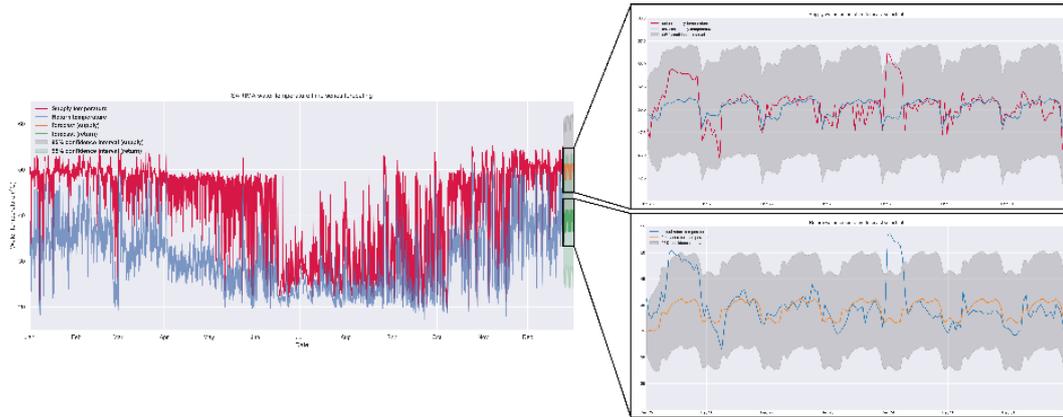
Figure 6.48: Results from SARIMA forecasting model with one week forecast (25th and 31th December). The results compares the actual and the forecasted data with grey region showing the confidence interval. The top right panel represents supply water temperature and the bottom right panel represents the return water temperature data.

shows that the forecast follows the trend of actual data except on few hours of 25th and 29th December. This change in the trend of actual data could be due to holiday season around Christmas time when the occupant's behaviour is expected to vary. The SARIMA also model provides the confidence interval along with predicted time-series, which is shown as the grey region around the forecast. This represents the 95% likelihood of the upper and lower range of possible prediction values from the model. The 95% of prediction interval is useful in evaluating model's capability, and varies between 36.6 − 54.8 for the supply water temperature data and 23.9 − 38.8 for the return water temperature data.

In order to evaluate the fitness of model, there are certain tests which needs to be performed. If these tests meet the passing criteria, only then we can have a complete confidence on our model. To this end, four statistical tests are performed on residual data for model evaluation. Figures 6.49 and 6.50 show the results for supply and return water temperature. The top left plot shows the standardised residual, where the normalised residual error is

plotted against the time. For an ideal model, there should be no obvious pattern in the residuals with a uniform variance. Similar observation has been made in supply and return water temperature models as their plots show that the residual errors are scattered around a mean of zero without any noticeable trends over time which seems adequate for our model performance. The top right plots show KDE (kernel density estimation) curve for supply and return water temperature. For optimum model, the KDE curve should resemble to the normal distribution. The top right plots in Figure 6.49 and 6.50 suggest that residual (orange curve) follow a normal distribution (green curve) with mean equal to 0.
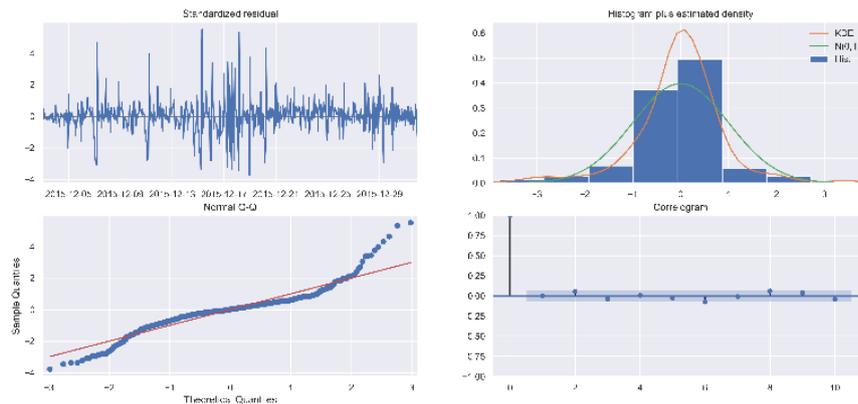
Figure 6.49: SARIMA model diagnostics for supply water temperature. The top left and top right panels show the standardised residual plot and KDE (kernel density estimation) curve. The bottom left and bottom right panels show the quantile-quantile (Q-Q) plot and correlogram (autocorrelation of the residual error).

The bottom left plot shows quantile-quantile (Q-Q) plot. For a perfect model, the residual errors should fall on a straight line to show that they are normally distributed. The bottom right plots in Figures 6.49 and 6.50 are the autocorrelation (ACF) of the residual error and also called Correlogram. The correlogram is of particular significance.In an ideal model, there should not be any autocorrelation because that could indicate a remaining trend in residual
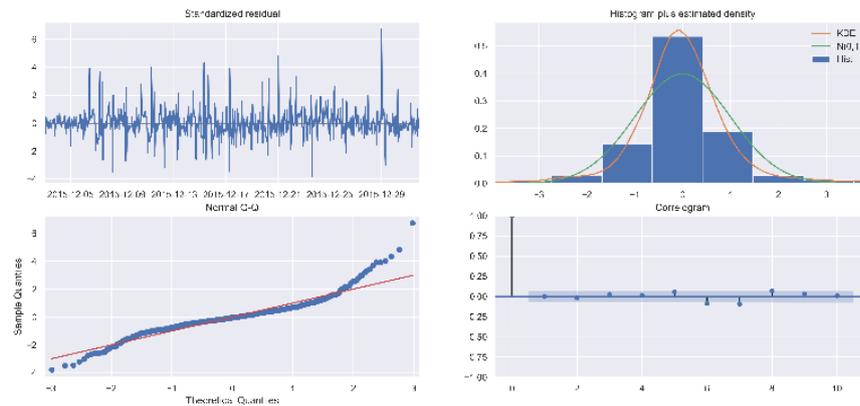
Figure 6.50: SARIMA model diagnostics for return water temperature. The top left and top right panels show the standardised residual plot and KDE (kernel density estimation) curve. The bottom left and bottom right panels show the quantile-quantile (Q-Q) plot and correlogram (autocorrelation of the residual error).

error, which needs to be explained by the model. In correlogram, the 95% of correlations for lag should be in blue area. The results show that almost 80% of data points lie on a straight line (red), and enough to accurately forecast. The residuals are found to be fairly flat and located in blue shaded region, which shows 95% of confidence interval. All these statistical tests conclude that these model are good fit to the data and can confidently be used for forecasting of supply and return water temperature.

To test the robustness of the SARIMA forecasting model, the supply and return water temperatures forecasting analysis is extended up to 3 weeks. More specifically, forecasting results are compared for one-day, one week, two weeks and three weeks. Table 6.2 and 6.3 shows different evaluation metrics for the supply and return water temperatures. It is observed that error dramatically decreases when the forecasting window size is reduced. For the supply water temperature, the mean absolute percentage error (MAPE) is found to be 4.05% for the 3 weeks forecast, which decreases to 2.71% and 2.14% for the 2-week and 1-week forecast, respectively. The MAPE for the 1-day (24-hr) forecast is
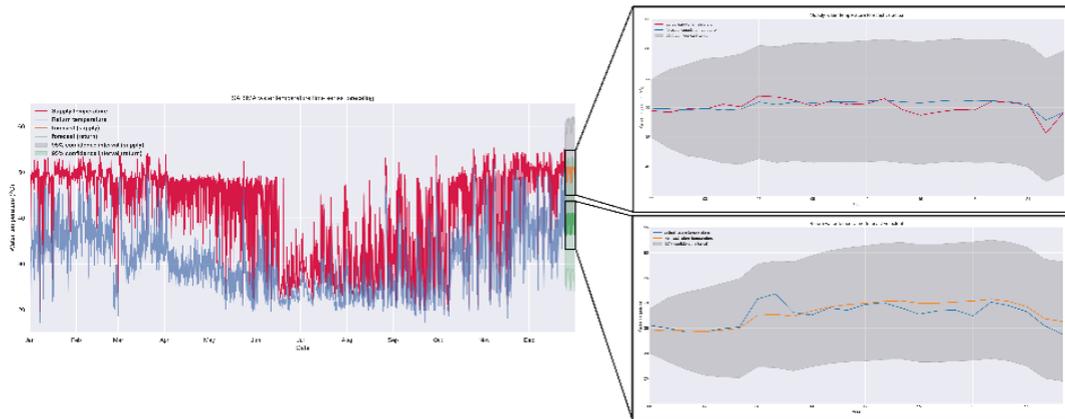
Figure 6.51: Results from SARIMA forecasting model with 24 hour forecast (31th December). The results compares the actual and the forecasted data with grey region showing the confidence interval. The top right panel represents supply water temperature and the bottom right panel represents the return water temperature data.

found to be 1.55% and Mean absolute error (MAE) is reduced to 0.768, which means the model is able to predict the supply water temperature with an error of just 0.7 ^{o}C. Moreover, the RMSE is almost reduced by almost 60% once the forecast window is reduced from 3-weeks to 1-day (24-hr).

Table 6.2: Prediction accuracy comparison at different forecasting horizons for supply water temperature

Metric	24 hours	1 week	2 weeks	3 weeks
Mean absolute error (MAE)	0.768	1.096	1.331	1.860
Mean absolute percentage error (MAPE)	1.555%	2.148%	2.714%	4.056%
Root mean square error (RMSE)	1.022	1.693	2.175	3.134
95% Prediction Interval	37.4 − 54.8	36.6 − 54.8	35.8 − 54.8	35.1 − 54.8

Similarly, the return water temperature forecast represents the same trend. The mean absolute percentage error (MAPE) is found to be 8.5% for the 3 weeks forecast, which decreases to 7.3% and 6.49% for the 2-week and 1-week forecast, respectively. The MAPE for the 1-day (24-hr) forecast is found to be 3.16% and Mean absolute error (MAE) is reduced to 0.768, which means the model is able to predict the return water temperature with an error

Table 6.3: Prediction accuracy comparison at different forecasting horizons for return water temperature

Metric	24 hours	1 week	2 weeks	3 weeks
Mean absolute error (MAE)	1.209	2.795	3.021	3.303
Mean absolute percentage error (MAPE)	3.169%	6.498%	7.317%	8.527%
Root mean square error (RMSE)	1.6098	4.413	4.383	4.458
95% Prediction Interval	$24.4 - 38.8$	$23.9 - 38.8$	$23.4 - 38.8$	$23 - 38.8$

of just 1.2^oC. Moreover, the RMSE is almost reduced by almost 40% once the forecast window is reduced from 3-weeks to 1-day. The 1-day (24-hr) forecast results for the supply and return water temperature is shown in Figure 6.51.

6.5   Discussion on results

In recent years, there has been a significant development in machine learning methods for time-series data forecasting. There has been a continuous debate regarding the application of machine learning methods in district heating network over the classical statistical forecasting methods. Therefore, in above Sections, various supervised machine learning methods and classical forecasting methods have been applied to attain the best forecasting method, which offer best predictions with respect to lowest forecast error and higher accuracy on the time-series data of district heating network. Moreover, the results from the best supervised machine learning method are compared against the classical stochastic forecasting method to determine the robustness and limitations of each method.

Generally, the machine learning methods can be categorised into three: supervised learning, unsupervised learning and reinforced learning. The supervised machine learning is the most common type of machine learning approach
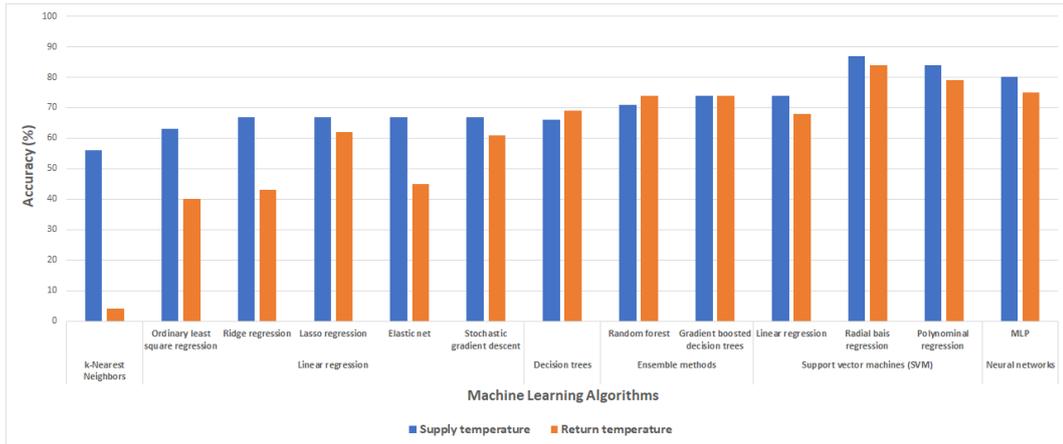
Figure 6.52: The barplot shows comparison between different machine learning methods. The results show that the support vector regression (SVR) and multi-layer perceptron (MLP) method have the best forecasting performance compared to others. Though, the accuracy of SVR-rbf kernel is higher, the computation duration of MLP based neural network algorithm is way less and the forecasting accuracy is comparable.

where predictions are made on the known data. There are two kind of supervised learning problems i.e. classification and regression. Classification aims to predict a class label from a predefined set of possibilities. In contrast, regression tries to make predictions as a real or continuous number. Here, only regression based supervised learning methods have been explored, as real number is required to forecast the supply and return water temperature in the district heating network.

The results from machine learning methods suggest that the support vector regression (SVR) and multi-layer perceptron (MLP) method have the best forecasting performance compared to others. Though, the accuracy of SVR-rbf kernel is higher, the computation duration of MLP based neural network algorithm is way less and the forecasting accuracy is comparable. Moreover, the ensemble methods (random forest and gradient based random trees) are computationally efficient (less expensive) and the forecasting accuracy is reasonably well. As expected the forecasting accuracy of K-NN and linear models

(ordinary least square, ridge, lasso, elastic net and stochastic gradient descent) is comparatively lower in the case of return water temperature.

Another point to discuss is which linear method should be used for machine learning. The results from the above Sections suggest that ElasticNet regression method is the best. Though, Stochastic Gradient Descent (SGD) methods is good, the ElasticNet is more reliable method. The ElasticNet method is also superior to the Lasso, as it provides flexibility to control model complexity from features selection and the model can even be reduced to linear kernel-based support vector regression. The Ridge and Lasso methods can be used for early stage predictions but ordinary least square (OLS) method should always be avoided.

The results further elaborate that correlation between supply and return water temperature data is found to be greater than the outside air temperature data. This explains that the weather data (outside air temperature, solar insolation, humidity and wind data) alone is not enough for forecasting using machine learning methods, and the accuracy of machine learning methods improves significantly once the return water temperature is used to forecast the supply water temperature, and vice versa. This explains that it is vital to use the supply and return water temperature of the district heating network for forecasting. Figure 6.52 shows the barplot for the comparison between different supervised machine learning methods.

Additionally, the traditional stochastic forecasting method, Seasonal Auto-Regressive Integrated Moving Average (SARIMA) is used and the results are better compared than linear and other low performance machine learning models. SARIMA is a variant of ARIMA method. It is the combination of

Auto-Regressive (AR) and Moving Average (MA) models and considers the seasonality in input data.

The results from fine-tuned SARIMA model on hourly data explains that the accuracy is very good for short-term forecasts (24 -168 hrs), and the forecast accuracy decreases as soon as the forecast horizon increases i.e. medium (<1 month) and long-term (>1 year). Once the forecasting horizon reduces from 168 to 24 hrs, the Mean absolute percentage error (MAPE) for supply water temperature decrease from 4% to 1.5% for the supply water temperature and 8.5% to 3.1% for the return water temperature, respectively.

6.5.1   Time-series forecasting versus machine learning prediction

The time-series forecasting is usually associated in financial and economic sectors by ARIMA modelling or Box–Jenkins method. Though, ARIMA methods is widely being used in the industry along its several variations (ARIMAX, SARIMA, SARIMAX), it has inherent issues and limitations which restrict its application, for example, the model assumes linear relation among variables and constant standard deviation in errors. This issue can be overcome by modelling the change in variance over time with the integration of ARCH or GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model into ARIMA model, but it makes the parameters optimisation challenging and computationally expensive (Siami-Namini & Namin 2018). Therefore, machine learning and other techniques are preferred over ARIMA models.

The best performing MLP machine learning model and SARIMA forecasting method are compared in order to determine each one's performance at different forecasting horizons. To this end, the MAE, MAPE and RMSE are
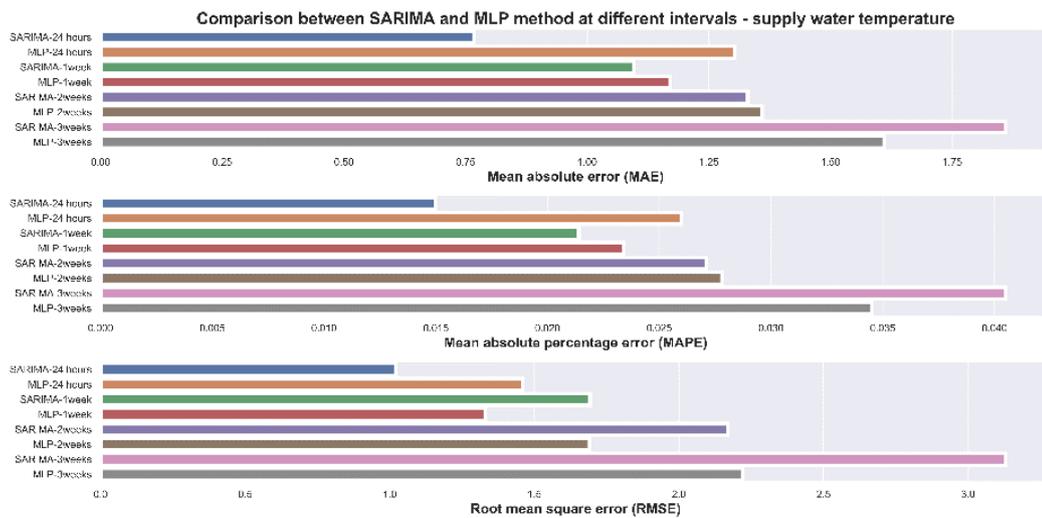
Figure 6.53: Barplot comparing the SARIMA and MLP forecast for supply water temperature at different forecasting horizons.
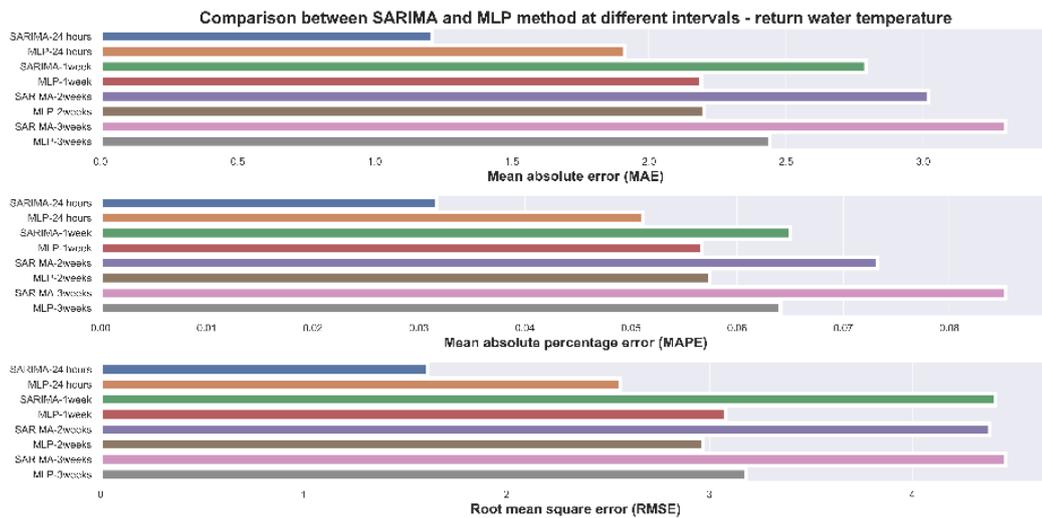
Figure 6.54: Barplot comparing the SARIMA and MLP forecast for return water temperature at different forecasting horizons.

evaluated for supply and return water temperature at 24 hours, 1 week, 2 weeks and 3 weeks forecasting horizons as shown in Figures 6.53 and 6.54. The results suggest that the traditional stochastic SARIMA method out-performs compared to MLP machine learning method for short-term forecasts (24 hours), but once the forecast horizon (2 and 3 weeks) increases the MLP neural networks out-performs SARIMA in terms of precision and accuracy. This data is also shown in Table 6.4. The forecasting time-series for both supply and return

water temperature at 24 hours and 2 weeks forecasting horizons is shown in

Figures 6.55 and 6.56.
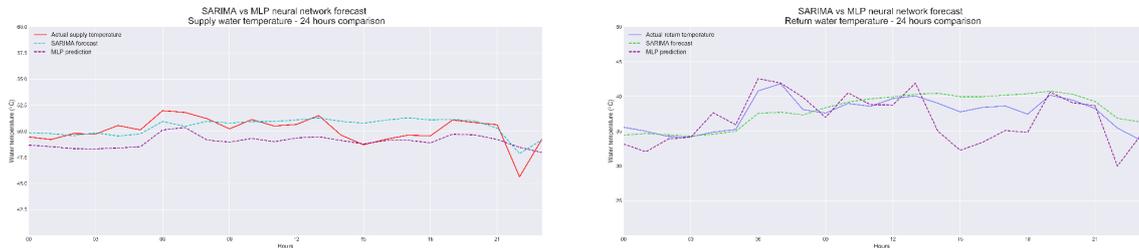
Figure 6.55: Comparison between SARIMA and MLP forecast for 24 Hours.
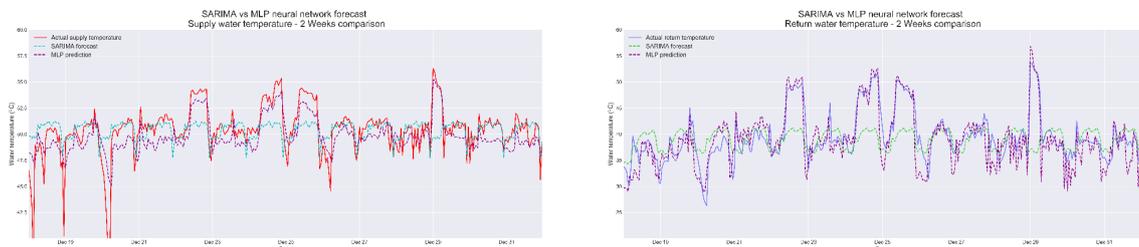
Figure 6.56: Comparison between SARIMA and MLP forecast for 2 Weeks data.

There has always been a debate which forecasting, and machine learning method should be used on the district heating networks and in conclusion, the traditional statistical methods are best for short-term time-series forecasts, and machine learning methods for medium- and long-term forecasts. It is foreseen that a hybrid model using autoregressive (SARIMA) modelling for short-term time-series forecast and unsupervised machine learning for the medium- and long forecasts as well as prediction of extreme events will be practical and another avenue to drive the research for energy modelling in district heating systems. Moreover, this behaviour of several algorithm methods for short-term forecasts is also discussed in (Makridakis et al. 2018, Siami-Namini & Namin 2018) and Figure 6.57 represents the performance of several algorithmic methods.
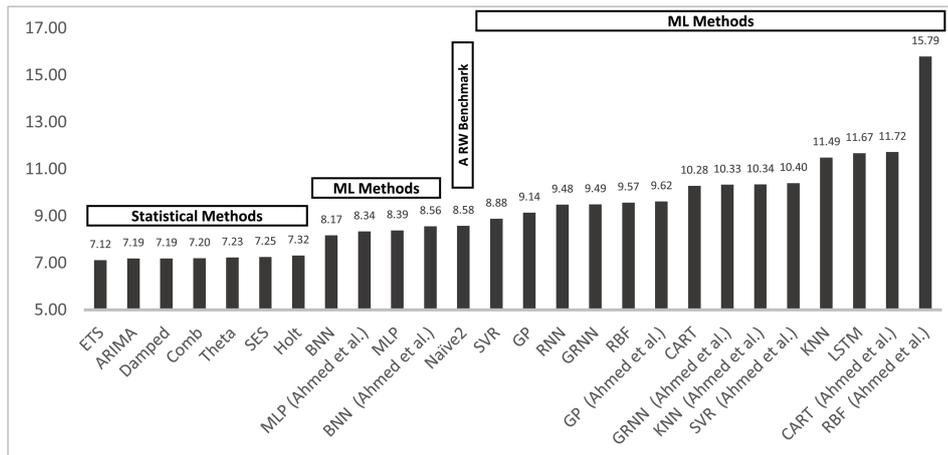
Figure 6.57: The short-term forecasting performance comparison among all algorithmic models available for time-series. The ARIMA models especially (SARIMA) performs as best overall (Makridakis et al. 2018).

It is imperative to discuss the latest advancements in the field of unsupervised machine learning for time-series forecasting. In recent years, the machine learning regression models (Support Vector Regression (SVR), Random Forests (RF)) and unsupervised machine learning models (deep learning) have been developed to overcome the limitations of traditional forecasting methods and quite popular both in the industry and researchers.

The deep learning methods (Recurrent Neural Network (RNN), and Long Short-Term Memory (LSTM)) can identify the structure and patterns hidden inside of dataset and models the statistical relation between variables in several deep hierarchy layers. The deep learning neural network model have the capability to learn nonlinear relationships and determine complexity in noisy time-series inputs data. Though, this makes these deep learning models complicated and difficult to train, the results are remarkable compared than traditional stochastic methods.

The RNN and LSTM deep learning methods are good to learn patterns in substantial dataset of multi-variate sequences of time-series without the need

to specify prediction time window, and therefore found applications in many disciplines including computer science, natural language processing, handwriting recognition, speech recognition, measuring impact of a certain news, as well as applications in economics and finance data such as, time-series prediction and volatility prediction.

Table 6.4: Comparison between MLP and SARIMA methods in order to determine best performance at different forecasting horizons

Metric	SARIMA method				MLP neural network method			
	24 hours	1-week	2-weeks	3-weeks	24 hours	1-week	2-weeks	3-weeks
Supply water temperature								
Mean absolute error (MAE)	0.768	1.096	1.33	1.86	1.304	1.17	1.36	1.61
Mean absolute percentage error (MAPE)	1.50%	2.14%	2.71%	4.05%	2.60%	2.34%	2.78%	3.45%
Root mean square error (RMSE)	1.022	1.693	2.17	3.13	1.46	1.33	1.69	2.22
Return water temperature								
Mean absolute error (MAE)	1.21	2.79	3.02	3.3	1.91	2.19	2.2	2.44
Mean absolute percentage error (MAPE)	3.17%	6.50%	7.32%	8.53%	5.11%	5.67%	5.74%	6.41%
Root mean square error (RMSE)	1.61	4.41	4.38	4.46	2.56	3.08	2.97	3.18

6.6   Summary

Time-series prediction and forecasting has number of applications in our lives and the main objective in this chapter has been to investigate which methods offers the best predictions with lower forecast errors and higher prediction accuracy for the DH network. The machine learning can be applied on time-series data for forecasting and prediction, however finding the best method remains a challenge as this depends on inherent characteristics of the data. Therefore, multiple machine learning and classical statistical methods have been explored to find the most suitable method for the supply and return water temperature of the district heating network. The results show that the correlation between supply and return water temperature data is found to be greater than the outside air temperature data. This explains that the weather data (outside air temperature, solar insolation, humidity and wind data) alone is not enough for forecasting using machine learning methods, and the accuracy of machine learning methods improves significantly once the return water temperature is used to forecast the supply water temperature, and vice versa. Therefore, it is vital to use the supply and return water temperature from the district heating network along with outside air temperature for forecasting. In conclusion, the traditional statistical models provide better results for short-term forecasts and machine learning model are more appropriate for long-term forecasts modelling. It is foreseen that a hybrid model using autoregressive (SARIMA) modelling for short-term time-series forecast and unsupervised machine learning for the medium- and long forecasts. Additionally the prediction of extreme events will be practical and another avenue to drive the research for energy modelling in district heating systems.

Chapter 7

GIS mapping for district heating network planning

Overview

This chapter discusses application of GIS mapping tools in the domain of district heating network. The technical and economical feasibility of a potential district heating network is assessed using GIS based model developed in FME software, which first calculates shortest path for the district heat network, pipe sizes, hydraulic analysis and then maps results on GIS interface. As this chapter focuses on early planning of district heating network, therefore it needed a site with unexplored potential of district heating network and alternate sources. The city of Nottingham has already got a well-established district heating network and therefore was not suitable for this study.

7.1 Introduction

In the beginning of 2019, a potential district heating network site for a university was identified in the North of England. The university campus consists of

a combination of new state-of-the-art facilities alongside older buildings. The masterplan is currently evolving with some buildings currently under construction or under design stage, and the university has long term aspirations which are in-line with the development of a heat network. To this end, the energy masterplanning is performed and its outcome suggested to explore the feasibility of setting up a district heating network. This would lead to an effective reduction of carbon emissions, operating costs and increased resilience of the entire development site.

The aim of this feasibility study is assessment of spatial distribution of energy demand in the campus and appraisal of the most appropriate heat technology for a district heating network to maximise the economic viability and carbon emission reduction. Moreover, the feasibility study is part of government's strategy for decarbonising heat supply and supporting growth of a low carbon economy in the UK.

Figure 7.1: The university campus is divided into four clusters, namely: Lower Mountjoy, East Hill Colleges, West Hill Collesges and Howlands Farm.

The precise planning and design of the district heating network is very important therefore, the campus is divided into four clusters, namely: Lower Mountjoy, East Hill Colleges, Howlands Farm and West Hill Colleges as shown in Figure 7.1.

7.2   Methodology

The GIS mapping is of foremost importance for the feasibility analysis of a potential district heating network. Therefore, a GIS based model has been developed using FME software, which first takes geographical location of the buildings along with their corresponding heat demand as an input and then: a) calculates the shortest path for district heat network, b) performs hydraulic modelling and, c) finally maps results on GIS interface. The hydraulic modelling calculates the pipe length and sizes, flow-rates and pressure gradient for each pipe segment of the district heating network.

The analysis is performed in two steps which includes: 1) identification of the buildings and their electricity and heat demand mapping, 2) district heating network design, pipe sizing and hydraulic calculations.

7.3   Results and Discussion

7.3.1   Electricity and Heat Demand Mapping

In order to get a better understanding of energy demand, the residential and non-residential buildings are identified across all four clusters described above. This identification has been shown in Figure 7.2. The residential buildings are shown in blue colour whereas non-residential buildings are coloured in red. This helps to visualise the distribution of the heat demand across the campus.
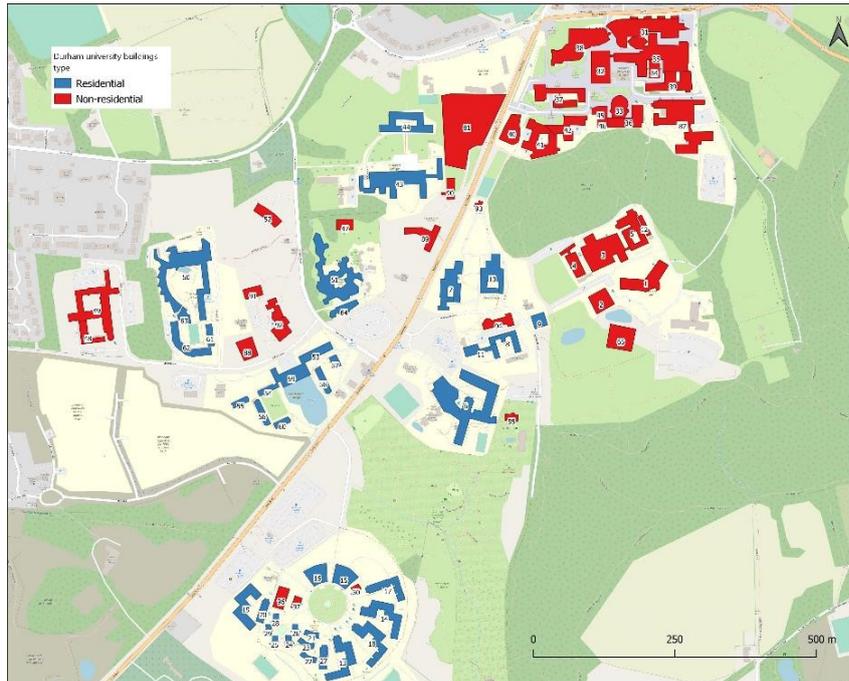
Figure 7.2: Types of buildings in university's DH network where residential and non-residential buildings are shown in blue and red respectively.
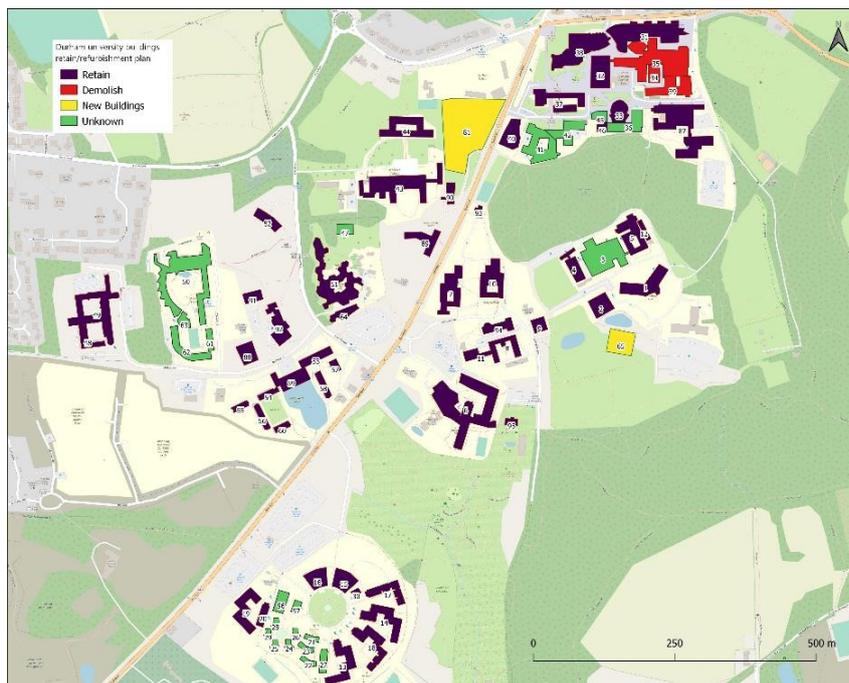
Figure 7.3: The refurbishment plan shows location of buildings to be retained, demolished, new and unknown types represented in deep purple, red, yellow and green respectively.

Additionally, Figure 7.3 shows the refurbishment and up-gradation plan of the area and clearly represents the location of existing, new buildings, to be

retained and demolished buildings. For the areas to be demolished, there will be new buildings built in their place, thus the demand is assumed to remain same.

Figure 7.4 shows the annual electricity demand geographically distributed around the area. It is observed that the electricity demand of non-domestic buildings is three to four times greater than the residential buildings, and the non-residential buildings in the Lower Mountjoy (northern cluster) have the highest electricity demand. The data elaborates that the Howlands Farm (southern cluster) has mostly residential buildings with electricity demand up to 500 MWh, which is around four times lower than the non-residential buildings.
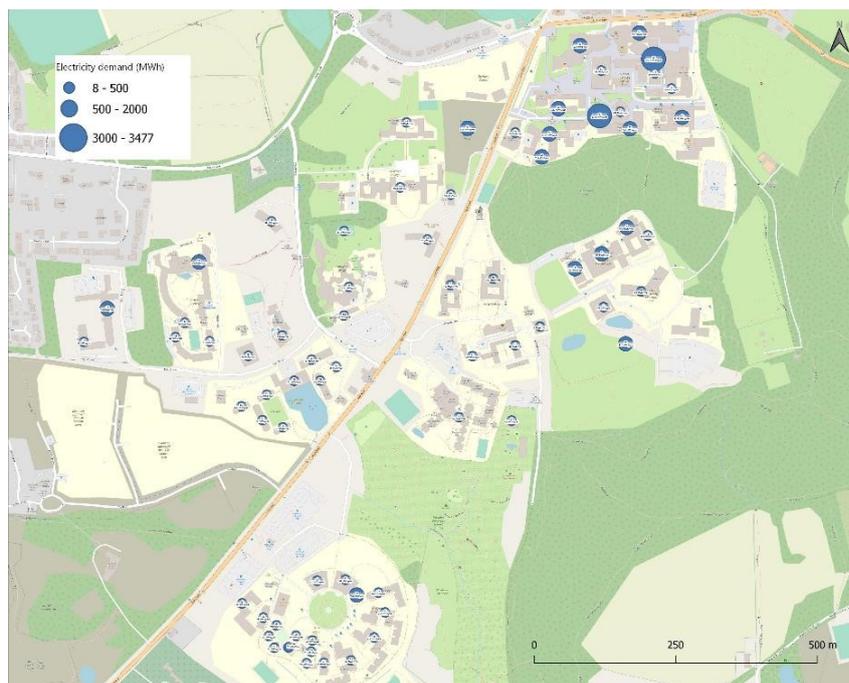
Figure 7.4: Annual electricity demand, of buildings in university's DH network, shown as blue circles of varying size depending on the amount of electricity demand.

Figure 7.5 shows the heat demand of the entire campus. The heat demand shows similar trend to electricity demand and the heat demand being twice in non-residential buildings compared to the residential buildings. Both
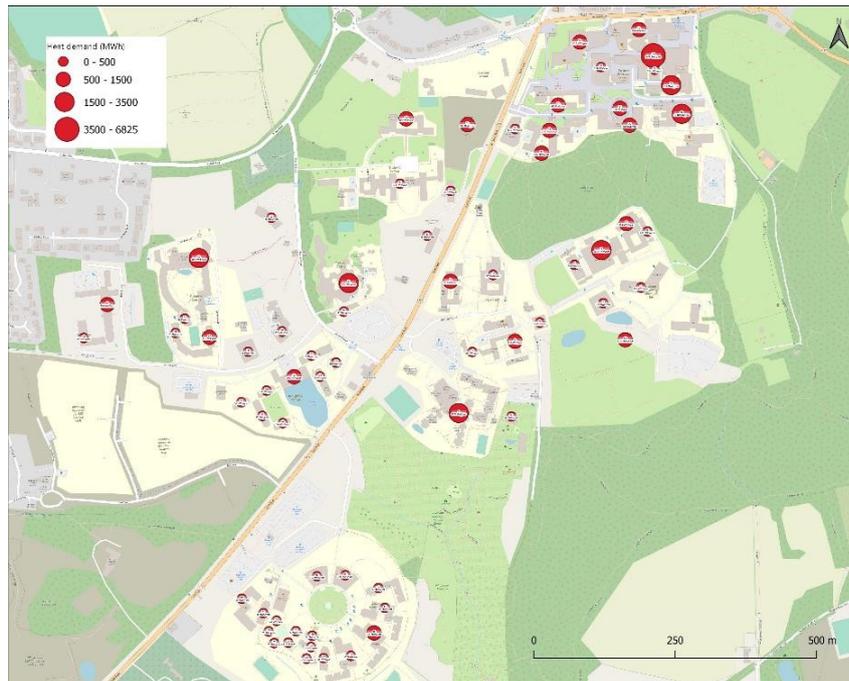
Figure 7.5: Annual heat demand shown as red circles of varying size reflecting heat demand.

figures show the maximum load of the site is located at the Lower Mountjoy (northern cluster). These electricity and heat demands are obtained from the meters located at the buildings and used for district heating network network modelling.

7.3.2   Hydraulic modelling of the district heating network

The hydraulic modelling of the district heating network is carried out using hydraulic model developed in *FME software*. It calculates the shortest path of pipes from the energy centre to the consumers, pipe sizes and the pressure gradient (Pa/m) for each segment of the network. The pipe sizes (for main trunk and distribution pipes) are calculated using flow-rate calculations for each section of the district heating network. The flow-rate calculations takes into account constraints such as; maximum permissible pressure drop and flow

velocity, peak thermal demand, delta t (Δt) and other parameters given in Table 7.1. The hydraulic modelling is performed to investigate two scenarios of the district heat network, i.e 1) centralised network 2) decentralised (stand-alone) network. Table 7.2 shows the capacity of individual energy center for each cluster.

Table 7.1: Description of key parameters for the potential district heating network

Parameter	Value	Unit
DH delta t (Δt)	20	^{o}C
Water specific heat capacity	4.181	KJ/Kg/K
Water density	1000	Kg/m3
Pipe roughness factor	0.05	mm
Max pressure drop, connection pipes	300	Pa/m
Max pressure drop, trunk mains	150	Pa/m
Max velocity, connection pipes	1.5	m/s
Max velocity, trunk mains	3	m/s

Table 7.2: Description of peak heat load for each cluster.

Name of cluster	Peak heat load (MW) Centralised network	Peak heat load (MW) De-centralised network
Northern cluster - Lower Mountjoy	5	6
Eastern cluster - East Hill Colleges	3	4
Western cluster - West Hill Colleges	3	3
Southern cluster - Howlands Farm	2	2
Total	**13**	**15**

7.3.2.1   Scenario - 1: Centralised district heating network

In the centralised scenario, a single energy centre with the capacity of 19 MW is assumed at the Lower Mountjoy (northern) cluster. Figure 7.6 shows the result for optimum district heating network along with pipe sizes, length and corresponding cost estimations for each pipe segment, are given in Table 7.3. It appears that a cost of ≈ £8 millions is needed for a total length of 6517

meters, comprising the entire district heating network. These results will be used as a benchmark to compare against Scenario 2.
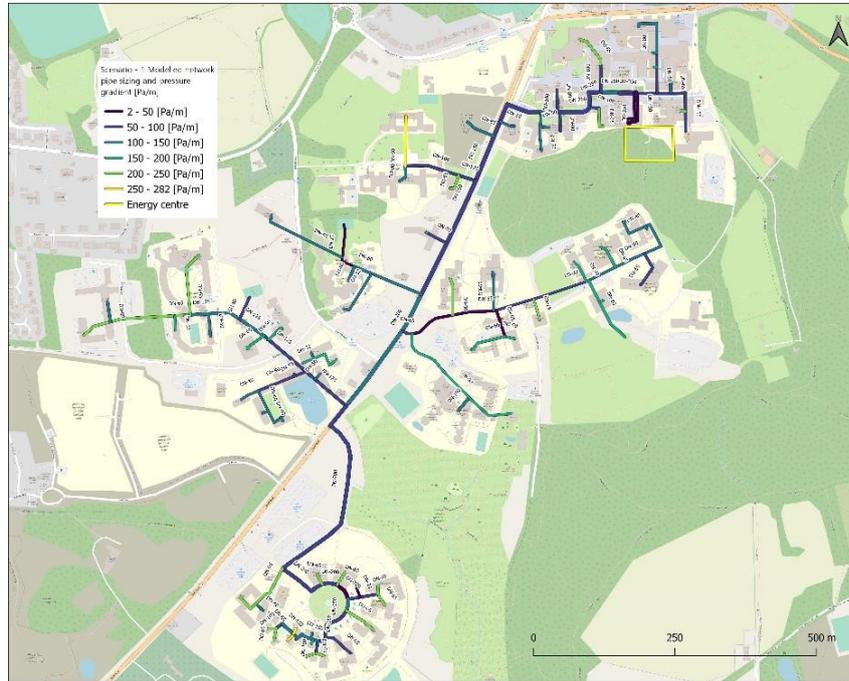
Figure 7.6: Scenario-1, centralised heat network with a single main heating plant of 19 MW. The legend and pipe colours represents the pressure gradient (Pa/m) and pipes are labeled with minimum required sizes.

Table 7.3: Description of pipe lengths and costs calculated from Scenario-1 - centralised heat network with a main heating plant of 13 MW.

Pipe Size	Length (m)	Cost (£), includes civil trenching costs
DN-20	132	137,940
DN-25	63	65,835
DN-32	606	633,270
DN-40	475	496,375
DN-50	724	756,580
DN-65	1327	1,386,715
DN-80	650	685,750
DN-100	468	505,440
DN-125	367	466,090
DN-150	189	251,370
DN-200	753	1,069,260
DN-250	605	1,161,600
DN-350	93	232,500
DN-450	65	165,750
Total	**6,517**	**8,014,475**

It can be noticed, that this centralised network uses several pipes with higher diameter. The DN-200 pipe is used as main trunk pipe for heat transmission between the clusters and DN-65 pipe is mostly used as connection pipe for heat distribution to the buildings.

7.3.2.2 Scenario - 2: Decentralised district heating network

In the decentralised (standalone) scenario, separate energy centres are assumed for each cluster. Figure 7.7 shows the result for optimum district heating network along with pipe sizes, length and corresponding cost estimations for each pipe segment. Table 7.4, 7.5, 7.6 and 7.7 show breakdown of pipe sizes, lengths, and the cost estimations for each energy centre dedicated to single cluster. The cumulative pipe cost of the network is calculated to be ≈ £ 6.3 millions for a total length of 5788 meters which is 21% less than in Scenario 1.
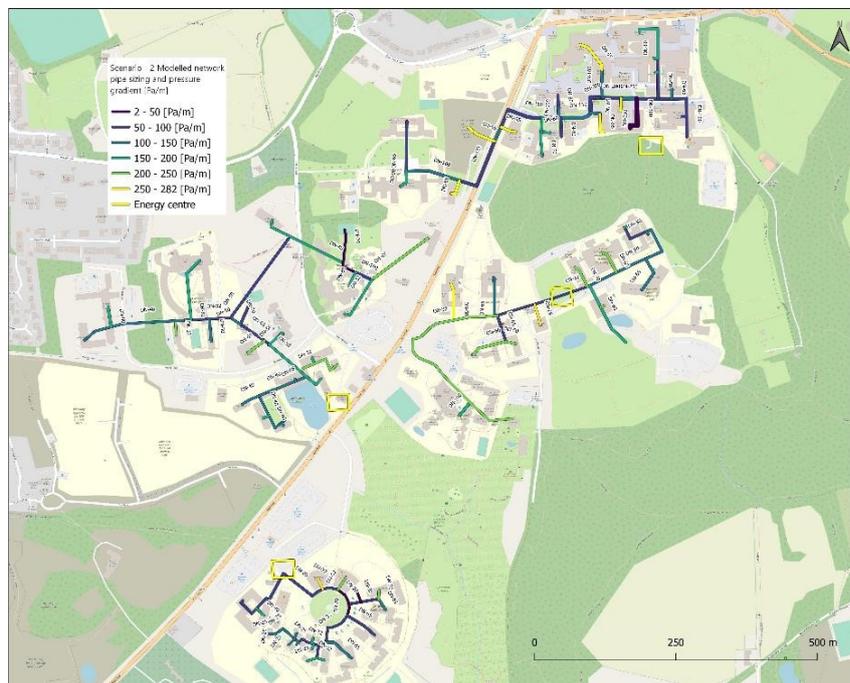
Figure 7.7: Scenario-2, decentralised (stand-alone) heat network with separate heat plant for each cluster. The legend and pipe colours represents the pressure gradient (Pa/m) and pipes are labeled with minimum required sizes.

It can be noticed in Table 7.8 that the decentralised network (Scenario 2)

Table 7.4: Description of pipe lengths and costs calculated for Lower Mountjoy cluster in Scenario-2, decentralised heat network of 6 MW.

Pipe Size	Length (m)	Cost (£), includes civil trenching costs
DN-20	1	1,045
DN-40	4	4,180
DN-50	139	145,255
DN-65	343	358,435
DN-80	196	206,780
DN-100	355	383,400
DN-125	146	185,420
DN-150	185	246,050
DN-200	82	116,440
DN-250	93	178,560
DN-300	65	135,200
Total	**1,609**	**1,960,765**

Table 7.5: Description of pipe lengths and costs calculated for West Hill Colleges cluster in Scenario-2, decentralised heat network of 3 MW.

Pipe Size	Length (m)	Cost (£), includes civil trenching costs
DN-20	91	95,095
DN-25	1	1,045
DN-32	260	271,700
DN-40	547	571,615
DN-50	418	436,810
DN-65	249	260,205
DN-80	221	233,155
DN-100	32	34,560
Total	**1,819**	**1,904,185**

Table 7.6: Description of pipe lengths and costs calculated for East Hill Colleges cluster in Scenario-2, decentralised heat network of 4 MW.

Pipe Size	Length (m)	Cost (£), includes civil trenching costs
DN-25	38	39,710
DN-32	576	601,920
DN-40	46	48,070
DN-50	415	433,675
DN-65	227	237,215
DN-80	121	127,655
DN-100	30	32,400
Total	**1,453**	**1,520,645**

Table 7.7: Description of pipe lengths and costs calculated for Howlands Farm cluster in Scenario-2, decentralised heat network of 2 MW.

Pipe Size	Length (m)	Cost (£), includes civil trenching costs
DN-20	72	75,240
DN-25	153	159,885
DN-32	214	223,630
DN-40	45	47,025
DN-50	15	15,675
DN-65	408	426,360
Total	**907**	**947,815**

Table 7.8: Cost comparison between the centralised and de-centralised (standalone) district heating network

Pipe size (mm)	Scenario 1	Scenario 2	Difference
DN-20	137940	171380	24.24%
DN-25	65835	200640	204.76%
DN-32	633270	1097250	73.27%
DN-40	496375	670890	35.16%
DN-50	756580	1031415	36.33%
DN-65	1386715	1282215	-7.54%
DN-80	685750	567590	-17.23%
DN-100	505440	450360	-10.90%
DN-125	466090	185420	-60.22%
DN-150	251370	246050	-2.12%
DN-200	1069260	116440	-89.11%
DN-250	1161600	178560	-84.63%
DN-300	-	135200	-
DN-350	232500	-	-
DN-450	165750	-	-
Total	8014475	6333410	20.98%

is economically feasible because of its lower costs and length of pipe required as compared to the centralised network (Scenario 1). The is due to pipe sizes with comparatively lower diameter. In scenario 1, the pipes greater than DN-100 are mostly used as a trunk pipes for heat transmission which contributes to excess cost of pipes. On the other hand, the scenario 2 uses pipes lower than DN-65 as trunk pipes. In addition to this, the heat-losses will be more

in Scenario 1 as the district heating network requires greater amount of pipe length.

7.3.3 Abandoned coal mines mapping

In feasibility studies, another objective is to look for alternate natural resources around the potential district heating site for improving the project's outlook. To this end, the area around the university was explored using resource maps from British geological survey (BGS) UK. This is performed to ascertain the location of different kind of coal mines, and possibility of water usage from the flooded abandoned coal mines. This is a very common situation in northern and eastern regions of England, since over the years these regions have been heavily dependent on coal for power generation. Additionally, these abandoned coal mines have sufficiently great amount of methane reserves which can be drawn out to generate power and heat in combined heat and power plants.
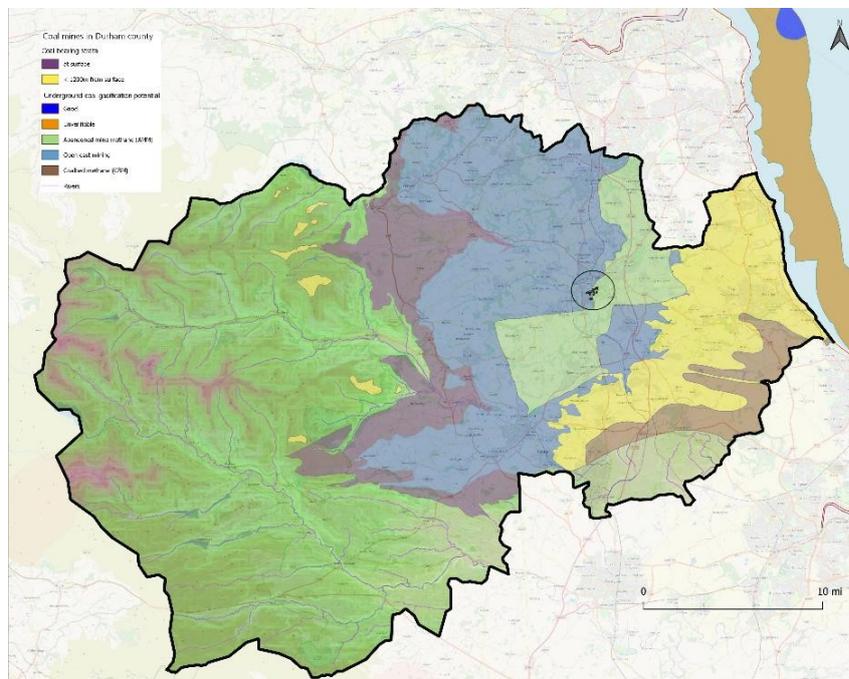
Figure 7.8: The distribution of coal resources available around the university site

The results show that the university lies in the region of open cast min-

ing and there are numerous abandoned coal mines. Moreover, the coal bed methane (CBM) resources are far away from the university and the underground coal gasification has limited potential in the sea. However, there is a technical potential to obtain flooded water along with abandoned mine methane (AMM) and coal bed methane (CBM), which can be a valuable fuel resource for energy centre in a district heating network. The location of resources is further shown in 7.8.

The abandoned mine methane (AMM) is fuel gas (mainly methane) which can be removed and utilised following the closure of the coal mine (Jones et al. 2004). On the other hand, the coalbed methane is a methane-rich gas found naturally within coal mines and can be explosive once mixes with air. It is also known as firedamp in mining industry (Holloway et al. 2005, Environment Agency UK July 2014).
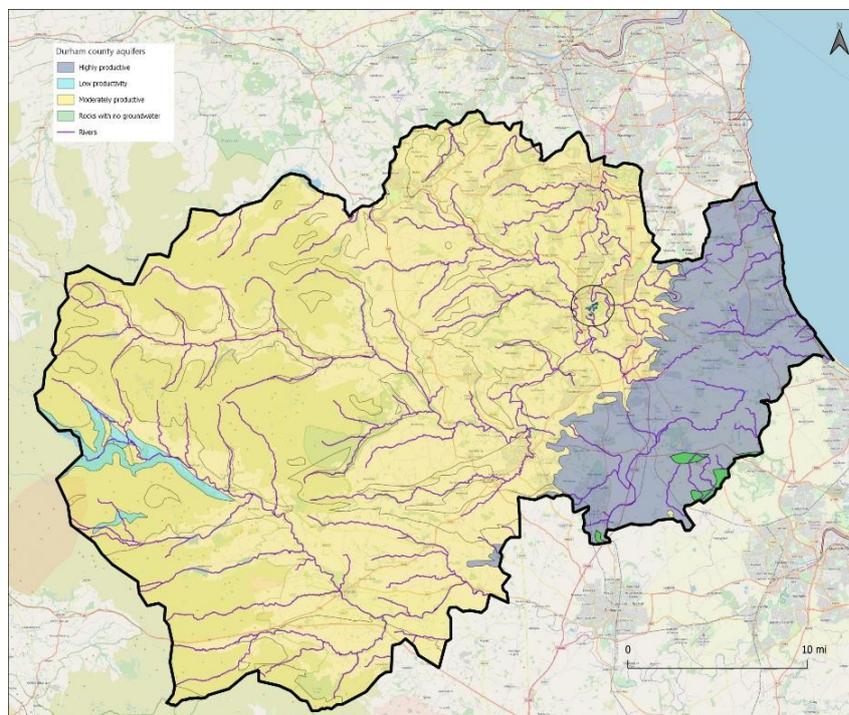
Figure 7.9: The hydro-geological maps around the university site.

The hydro-geological maps provides valuable information about aquifer

by providing hydrochemical characteristics of the underlying rocks and their usefulness for supplying groundwater. This is particularly useful for the installation of water source heat pumps (WSHPs). The results show the university is in the region of moderately productive aquifers and the highly productive aquifers are nearby. This exemplifies the possibility of using underground water for the district heating network, and shown by results in Figure 7.9.

7.4 Summary

In this chapter, the application of GIS based tools has been explored using a self-developed model, and the feasibility of a real district heating network has been discussed. The possibility of a centralised district heating network with one energy centre has been investigated and compared with a decentralised (standalone) network with multiple energy centres. The results supports the decentralise network based on the cost of network to fulfill the demand and reduction in heat losses in the network. Moreover, there is a strong evidence to utilise water along with abandoned mine methane (AMM) and coal bed methane (CBM) in the district heating network from the abandoned coal mines as an alternate energy resource. It is suggested the water source heat pumps (WSHPs) should be implemented for the utilisation of underground water in the district heating network.

Chapter 8

Discussion and Conclusions

The main aim of the research in this thesis was to investigate the implementation of low-temperature district heating in existing boiler based buildings using different energy modelling techniques. In order to achieve this, first the thermal performance of buildings was studied followed by energy efficient design of a network in different operational scenarios. This was complemented by economically optimal solution for the decarbonised heating network with integration of fully renewable electrical grid. Then, the application of machine learning and forecasting methods were explored using real monitored data from the district heating network. Finally, the role of GIS mapping was studied on a real case study project for the early stage planning and feasibility of an energy efficient district heating network.

Nottingham has set the most ambitious targets compared to other cities in the UK and aims to achieve net-zero carbon emissions by 2028. The district heating is central for achieving these targets as the heat source of current district heating is waste incineration and offsets approximately 27,000 tonnes of CO_2 emissions annually. The high return water temperature in this network has shown sufficient capacity for a secondary LTDH network intervention to

the nearby areas using return pipe. This has extended the district heating network and made it more energy efficient and greener.

8.1   Achievement of research objectives

Before the implementation of LTDH, it is necessary to increase thermal performance of the buildings. This is increased by improving its fabric, glazing and building services and needs to be in line with the building regulation in the UK. Therefore, the first research objective is addressed in Chapter 3, where the relationship between the building regulations and thermal performance of domestic and non-domestic buildings is evaluated. These building regulations are the minimum energy efficiency standards for the conservation of fuel and power in England/Wales, and forms the basis for a building design. To this end, both domestic and non-domestic buildings have been considered as a case study to investigate their thermal performance using two separate software (i.e. IDA-ICE and IES-VE).

The existing domestic buildings from the REMOURBAN project are considered in the first part of the analysis and the energy performance has been studied before and after retrofit. The results suggest that the retrofitting increases the energy performance by almost 50% and the risk of over-heating is determined by the type of glazing being selected.

In order to provide a comprehensive insight of thermal performance calculations, a non-domestic building has also been considered and the data suggests that the methodology for hot water demand calculation in building regulations is unrealistic and leads to the failure of compliance (i.e. high carbon emissions than the target building emissions). These excess carbon emissions are usually compensated by installation of PV solar generation panel to demonstrate

compliance by almost 15%. Finally, it is anticipated that with current building regulations in the UK, it is unlikely to achieve the target of net-zero emission buildings (NZEB) in existing domestic and newly built non-domestic buildings by the year 2050.

The second research objective was achieved in Chapter 4 in which the design and operation of an energy efficient LTDH network from the REMOURBAN project has been studied. This LTDH network is characterised by low supply temperature and flow-rates and high Δt i.e. the difference between supply and return water temperature. To this end, the heat demand of the buildings is used to calculate the flow-rates and hydraulic calculations.

Initially, the monitored data of an existing flat is used to investigate 25% of the network with four different control scenarios in order to find the optimum pumping strategy. This is performed using a novel hydraulic model developed in Python programming language. The results suggest that variable pumping with fixed supply water temperature from the sub-station is found to have least amount of pumping power, energy consumption and heat-losses in the network.

Then, the analysis is extended to the entire LTDH network which required robust methodology. To achieve this, Claytex licensed HVAC (for Heating Ventilation and Air Conditioning) modelling library in Dymola software is used to develop a physical component based model to analyse the optimum design flow-rates parameters, pump control strategies and integration of multiple heat sources in the network. The Dymola model was created to imitates the real LTDH network from the REMOURBAN project and referred as baseline model in the analysis. Each component (including flats) is modelled separately with

exact parameters from the network. The simulations have been performed at hourly intervals for the month of highest heat demand i.e. January. This baseline model is validated using real monitored data from the REMOURBAN LTDH network and found to in agreement.

In order to improve the network performance, several iterations were introduced in the baseline model using Dymola simulation. Each iteration involved addition of different heat source, component and network parameters. Iteration 1 optimises the efficiency of current LTDH network with reduction in flow-rates and installation of PID (proportional integral derivative) controllers based feedback loop on the circulation pump in the plant room. Iteration 2 assumes installation of standalone solar thermal plant connected to the thermal heat storage as the only heat-source to the LTDH network. Iteration 3 investigates the optimised LTDH network (from iteration 1) with the integration of solar thermal power plant. In other words, iteration 3 is the combination of both iteration 1 and iteration 2 networks. Iteration 4 considers another configuration as an extension of iteration 3, where the solar thermal plant pre-heats the water inside the thermal heat storage before feeding it into the gas-boiler connected directly to the LTDH network.

In conclusion, results from the first part of LTDH modelling analysis suggest that the LTDH network should be designed with variable speed pumping, and supply water temperature should be kept constant from the plant room. This leads to the lowest energy consumption in the network. On the other hand, the heat-losses in the existing district heating networks, which predominantly operate at constant flow-rate in the UK, can be reduced by installation of weather compensation valves (regulation of supply water temperature ac-

cording to the outside temperature). Furthermore, in the second part, the energy efficiency and Δt from all iterations are compared against the baseline model and the results show that both energy efficiency and Δt in iteration 1 is larger than the baseline model. For all other iterations Δt appears to be less than the baseline model. These results suggest that the LTDH network in REMOURBAN project can be improved by reducing flow-rates both in the network and circulation pump inside the plant room (i.e. iteration 1).

The novelty of this analysis has been to use a live LTDH network intervention as a case study which provides a energy efficient solution for the UK. The learning from this study can be replicated to the future LTDH network project anywhere in the UK or elsewhere. Furthermore, the reduction in supply water temperature will enable the possibility of using heat from renewable energy resources and other low-grade waste heat sources.

With recent environmental concerns, there is a debate regarding the electrification of the heating and other energy networks (also known as sector coupling). The research objective to investigate the economic feasibility of 100% renewable energy based district heating network is studied in Chapter 5. Where 1 year of hourly heat demand and heat generation cost data from an existing district heating network (Aarhus, Denmark) is used to investigate the economical aspects of fully excess renewable energy based heating network. The district heating network is assumed to be connected to the fully renewable based electrical grid and excess generation is used by heat-pumps and thermal heat storage. The rationale behind using Aarhus district heating network as a case study was that, the city has similar population as of Nottingham with

300,000 inhabitants and considered as among five most energy efficient district heating networks in the world.

The techno-economic analysis for the de-carbonised district heating network shows that the futuristic fully renewable powered electrical grid and heating network is not economically feasible, unless both networks are coupled together which reduces the energy costs by almost 50%. Additionally, per unit decarbonisation cost depends on wind and solar (PV) power generation cost. The cost of excess renewable energy generation from the electrical grid is determined by per unit cost of heat-pumps and backup energy technologies (thermal heat storage and natural gas-boiler). Finally, the 100% decarbonisation of the district heat network is possible but depends on usage of private wire (i.e. selling excess electricity and heat to the private consumers).

A district heating network consists of heat sources, hot water supply and cold return transmission pipes and heat distribution equipment. This creates a complex energy dynamics and the efficient operation depends on the correct estimation of heat demand which is directly related to the pressure, flow-rate and Δt in the network. The heat required by the consumers can be fulfilled by regulating either the flow-rate or temperature difference between the supply and return water temperature in the network. Therefore, forecasting of heat demand in district heating network is vital for its operation. There are three types of forecast horizons depending on their purpose, i.e. short (24-168 hrs), medium (<1 month) and long-term (>1 year) forecasts. The short-term forecast is of peculiar importance when it comes to energy efficiency and optimisation. Therefore, the accurate short-term forecasts are useful to minimise the peaks, demand side management and imbalance in the network.

As REMOURBAN is a live LTDH project and it presents the opportunity to obtain monitored data from several sensors. The fourth research objective is attained in Chapter 6, where the application of machine learning and forecasting methods is explored in the context of district heating network. The 1 year (8760 timsteps) of monitored LTDH data used in this chapter is obtained from one of the building in REMOURBAN project. This data is combined with weather data and the monitored data consists of 9 different parameters altogether. The aim is to build a model using this monitored data for the future prediction of supply and return water temperature in the district heating network. To this end, the application of several regression based supervised machine learning and classical stochastic forecasting methods is explored and results from the fine tuned models are compared to determine performance, robustness and limitation of each method.

Generally, the machine learning methods can be categorised into three: supervised learning, unsupervised learning and reinforced learning. In our analysis the supervised machine learning is used to make real or continuous number predictions on the known data. The machine learning can be applied on time-series data for forecasting and prediction, however finding the best and fine-tuned method is a challenge as this depends on inherent characteristics of the data.

The results from machine learning methods suggest that, the support vector regression (SVR) and multi-layer perceptron (MLP) method have the best forecasting performance compared to others. Though, the accuracy of SVR-rbf kernel is higher, the computation duration of MLP based neural network algorithm is way less and the forecasting accuracy is comparable. The forecast-

ing accuracy of ensemble methods (random forest and gradient based random trees) is reasonably well and are computationally efficient (less expensive). On the other hand, K-nearest neighbor (K-NN) and linear methods (ordinary least square, ridge, lasso, elastic net and stochastic gradient descent) do not perform well for forecasting. Moreover, it is observed that the forecasting accuracy of machine learning model using weather data improves greatly once the return water temperature data is used to forecast the supply water temperature, and vice versa. This explains that the weather data (outside air temperature, solar insolation, humidity and wind data) alone is not enough for forecasting using machine learning methods.

The classical stochastic Auto-Regressive Integrated Moving Average (ARIMA) forecasting method is a widely used method, and is a combination of Auto-Regressive (AR) and Moving Average (MA) models. As monitored data from district heating network contains seasonality, an extension of ARIMA known as Seasonal Autoregressive Integrated Moving Average (SARIMA) is used. It can handle both trend and seasonal component in the input time-series data. The results from fine-tuned SARIMA model on hourly data explains that the forecasting accuracy depends on the forecast horizon. The model has very good accuracy for short-term forecasts (24 -168 hrs), but it decreases as soon as the forecast horizon increases i.e. medium (<1 month) and long-term (>1 year). Still the overall accuracy of the SARIMA model is better than the low performance machine learning models (i.e. K-nearest neighbor and linear methods).

In conclusion, the performance comparison between modern machine learning and classical stochastic models for time-series forecasting suggest that

for short-horizon forecast, the traditional stochastic SARIMA method out-performs all supervised machine learning models in terms of precision and accuracy, but once the forecast horizon increases, the MLP and SVR neural networks out-performs SARIMA model.

Since, the current focus is shifting towards the installation of low carbon technologies, therefore, the GIS mapping tools can be helpful in designing LTDH networks with heat pumps. In addition, the selection of district heating network site and layout is crucial for the design and economic viability of any LTDH project. The last research objective is achieved in Chapter 7 where the application of GIS mapping tool is outlined. The early stage design and planning of a real district heating project has been used as a case study. The GIS based model has been developed in FME software which calculates shortest path for the district heat network, pipe sizes, hydraulic analysis and then maps results on GIS interface. The aim was to investigate the benefit of a centralised district heating network with one energy centre compared to a decentralised (standalone) network with multiple energy centres (such as heat pumps). The results suggested that the decentralised LTDH network with multiple energy centres is the optimum strategy owing to the cost of network pipe-works and heat-losses in the network. This also helped to identify the location of nearby heat sources and the distance between consumers.

8.2 Contribution to knowledge

This project has shown that the intervention of low temperature district heating in existing buildings can be implemented with success using the return pipe of the existing district heating networks. This has been demonstrated

by studying thermal performance of the buildings after retrofitting and energy efficient operation of the LTDH network.

The first contribution to the knowledge is that the deep-retrofitting of existing building reduces the energy consumption by almost 50% and should be performed before the implementation of LTDH. Another novel finding is that the existing boiler based buildings are already technically feasible for the conversion to LTDH, and the Δt of the REMOURBAN LTDH networks can be improved by reducing flow-rates in the network as well as the circulation pump (inside the plant room). This can be achieved by installation of PID controller inside the plant room and demonstrated with DYMOLA simulations. This knowledge should be applicable to any future LTDH networks. Another contribution to the knowledge is that futuristic de-carbonisation of district heating network is not economically feasible, unless both fully renewable powered electrical grid and heating networks are coupled together. Additionally, 100% decarbonisation of heating network depends on selling excess electricity and heat to the private consumers.

Yet another interesting finding is that predictive modelling is applicable on the LTDH network data. It has been concluded that the classical stochastic SARIMA forecasting method performs better for short term forecasting and the supervised machine learning models (SVR and MLP neural networks) are more appropriate for long term forecasting. Finally, the practical contribution of GIS mapping has shown that it can play an important role in early planning to provide a quick overview, design, layout for a potential district heating site as well as identification of alternate heat sources.

8.3 Future work

As described in this chapter, the optimisation of flow-rates is the key in achieving the energy efficiency in the LTDH network. One thing which should be explored is the use of ground source heat pumps as source of district heating network and domestic hot water is provided by electric heating inside the flats. In this way the risk of legionella will be removed and the network will be able to operate at even lower temperatures. Moreover, the flow-rate and required pipe sizes will be reduced. This will impact the heat-losses and capital cost of the network.

Since the computation power was limited and the LTDH network was simulated for only one month, but given that more computational resources in terms of high performance computing (HPC) are available, data for whole year could be simulated to extract the trends reflecting the varying demand over different seasons.

It is foreseen that moving forward, a development of hybrid model is a way to improve to forecasts the heat demand of the LTDH network. That specific hybrid model should consider SARIMA modelling for short-term forecasting and deep learning (Long short-term memory - LSTM) for medium and long term forecasting. This can also be combined with Dymola model to predict the extreme events, which is a practical avenue to drive the research for energy modelling in district heating systems. Finally, a better coordination and access to receive consistent monitored data directly from the sensors could improve better data analysis.

Appendix A

REMOURBAN LTDH network project parameters

Plant Room Schematic
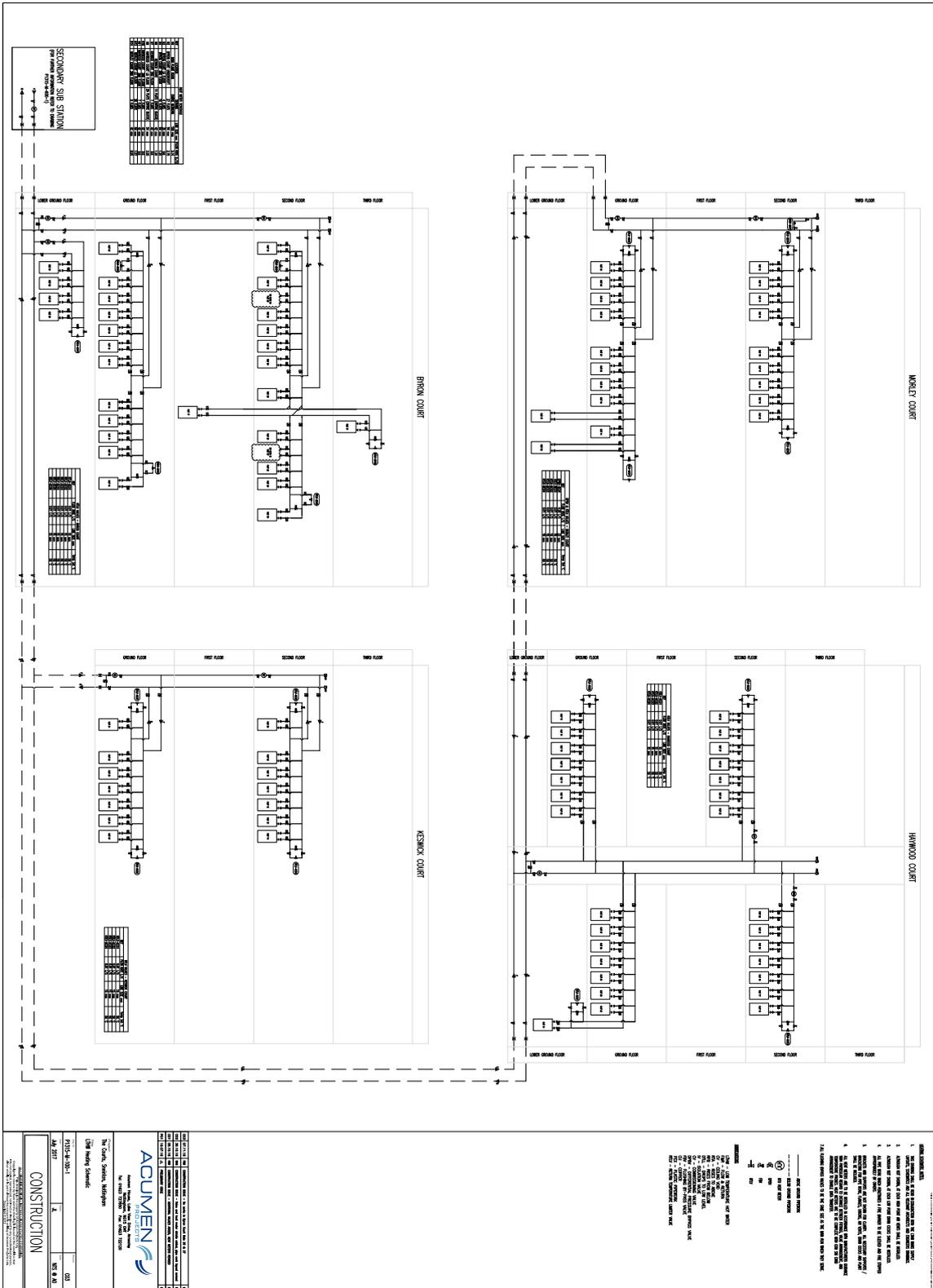
Bibliography

Agrawal, R. & Adhikari, R. (2013), 'An introductory study on time series modeling and forecasting', *Nova York: CoRR* .

Anderson, B., Chapman, P., Cutland, N., Dickson, C., Henderson, G., Henderson, J., Iles, P., Kosmina, L. & Shorrock, L. (2015), '*BREDEM 2012 - A technical description of the BRE Domestic Energy Model*', https://www.bre.co.uk/filelibrary/bredem/BREDEM-2012-specification.pdf.

Andresen, G. B., Rodriguez, R. A., Becker, S. & Greiner, M. (2014), 'The potential for arbitrage of wind and solar surplus power in denmark', *Energy* **76**(0), 49 – 58.
URL: *http://www.sciencedirect.com/science/article/pii/S0360544214002977*

Arup, O. & Ltd, P. (2011), '*Review of the generation costs and deployment potential of renewable electricity technologies in the UK*', https://www.gov.uk/government/collections/energy-generation-cost-projections.

Ashfaq, A. & Ianakiev, A. (2018*a*), 'Cost-minimised design of a highly renewable heating network for fossil-free future', *Energy* **152**, 613 – 626.
URL: *http://www.sciencedirect.com/science/article/pii/S0360544218305644*

Ashfaq, A. & Ianakiev, A. (2018*b*), 'Investigation of hydraulic imbalance

for converting existing boiler based buildings to low temperature district heating', *Energy* **160**, 200 – 212.

URL: *http://www.sciencedirect.com/science/article/pii/S0360544218312891*

Ashfaq, A., Kamali, Z. H., Agha, M. H. & Arshid, H. (2017), 'Heat coupling of the pan-European vs. regional electrical grid with excess renewable energy', *Energy* **122**, 363 – 377.

Becker, S., Frew, B. A., Andresen, G. B., Jacobson, M. Z., Schramm, S. & Greiner, M. (2015), 'Renewable build-up pathways for the us: Generation costs are not system costs', *Energy* **81**, 437 – 445.

URL: *http://www.sciencedirect.com/science/article/pii/S0360544214014285*

Becker, S., Frew, B. A., Andresen, G. B., Zeyer, T., Schramm, S., Greiner, M. & Jacobson, M. Z. (2014), 'Features of a fully renewable US electricity system: Optimized mixes of wind and solar PV and transmission grid extensions', *Energy* **72**, 443 – 458.

URL: *http://www.sciencedirect.com/science/article/pii/S0360544214006343*

Becker, S., Rodriguez, R., Andresen, G., Schramm, S. & Greiner, M. (2014), 'Transmission grid extensions during the build-up of a fully renewable pan-European electricity supply', *Energy* **64**, 404 – 418.

URL: *http://www.sciencedirect.com/science/article/pii/S0360544213008438*

BEIS UK, . (2018*a*), '*Heat Network Investment Project (HNIP), a case study brochure*', https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/691643/Heat_Network_Case_Study_Brochure.pdf. Accessed: 11 May 2020.

BEIS UK, . (2018*b*), '*The Clean Growth Strategy: Leading the way to a low carbon future*', https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/700496/clean-growth-strategy-correction-april-2018.pdf. Accessed: 11 May 2020.

Bofinger, S., von Bremen, L., knorr, K., lesch, K., Rohrig, K., Saint-Drenan, Y.-M. & Speckmann, M. (November, 2008), '*Raum-zeitliche Erzeugungsmuster von Wind- und Solarenergie in der UCTE Region und deren Einfluss auf elektrische Transportnetze*'.

Boysen, H. & Thorsen, J. E. (2007), '*Hydraulic balance in a district heating system*', https://assets.danfoss.com/documents/DOC024186450916/DOC024186450916.pdf.

BS EN 442-1, . (2014), 'BS EN 442-1:2014 Radiators and convectors. Technical specifications and requirements'.

BSI (2008), '*BS EN 15251:2007 - Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acoustics*'.

Building Regulations, . (2010*a*), '*The buildings regulations 2010, Appoved document Part-L1B Conservation of fuel and power in existing dwellings*', https://www.gov.uk/government/publications/conservation-of-fuel-and-power-approved-document-l. Accessed: 30 April 2020.

Building Regulations, . (2010*b*), '*The buildings regulations 2010, Appoved*

document Part-L2A Conservation of fuel and power in new buildings other than dwellings', https://www.gov.uk/government/publications/conservation-of-fuel-and-power-approved-document-l. Accessed: 30 April 2020.

Building Regulations, . (2010*c*), '*The buildings regulations 2010, Appoved document Part-L2B Conservation of fuel and power in existing buildings other than dwellings*', https://www.gov.uk/government/publications/conservation-of-fuel-and-power-approved-document-l. Accessed: 30 April 2020.

Building Regulations, . (2010*d*), '*The buildings regulations, Appoved document Part-L1A Conservation of fuel and power in new dwellings*', https://www.gov.uk/government/publications/conservation-of-fuel-and-power-approved-document-l. Accessed: 30 April 2020.

BuroHappold Engineering, . (2016), '*UK Spatial District Heating Analysis*', http://fes.nationalgrid.com/media/1215/160712-national-grid-dh-summary-report.pdf. Accessed: 11 May 2020.

Bøhm, B. (2010), 'Experimental determination of heat losses from buried district heating pipes in normal operation', *Heat Transfer Engineering* **Vol. 22**, pp. 41–51.

Calisir, T., Yazar, H. O. & Baskaya, S. (2017), 'Determination of the effects of different inlet-outlet locations and temperatures on pccp panel radiator heat transfer and fluid flow characteristics', *International Journal of Thermal*

Sciences **121**(Supplement C), 322 – 335.

URL: *http://www.sciencedirect.com/science/article/pii/S1290072917302545*

Centre for Sustainable Energy Association for the Conservation of Energy, . & Moore, D. R. (2008), '*Achieving optimal carbon savings from the UK's existing housing stock*', https://www.cse.org.uk/downloads/reports-and-publications/policy/how$_$low.pdf. Accessed: 30 April 2020.

Chen, P., Niu, A., Liu, D., Jiang, W. & Ma, B. (2018), Time series forecasting of temperatures using sarima: An example from nanjing, *in* 'IOP Conference Series: Materials Science and Engineering', Vol. 394, IOP Publishing, p. 052024.

CIBSE (Jul 2015), *Heat Networks: Code of Practice for the UK*, Chartered Institution of Building Services Engineers (CIBSE).

URL: *https://www.cibse.org/knowledge/knowledge-items/detail?id=a0q200000090MYHAA2*

Committee on Climate Change, . (2010), '*The Fourth Carbon Budget – reducing emissions through the 2020s*', https://www.theccc.org.uk/publication/the-fourth-carbon-budget-reducing-emissions-through-the-2020s-2/. Accessed: 30 April 2020.

Connolly, D., Lund, H. & Mathiesen, B. (2016), 'Smart energy Europe: The technical and economic impact of one potential 100% renewable energy scenario for the European Union', *Renewable and Sustainable Energy*

Reviews **60**, 1634 – 1653.

URL: *http://www.sciencedirect.com/science/article/pii/S1364032116002331*

Connolly, D., Lund, H., Mathiesen, B., Werner, S., Möller, B., Persson, U., Boermans, T., Trier, D., Østergaard, P. & Nielsen, S. (2014), 'Heat roadmap Europe: Combining district heating with heat savings to decarbonise the EU energy system', *Energy Policy* **65**, 475 – 489.

URL: *http://www.sciencedirect.com/science/article/pii/S0301421513010574*

Cui, J. M., Ianakiev, A. & Garcia-Fuentes, M. A. (2017), 'To examine appropriate deep-retrofit practice using simulation results in an EU-funded urban regeneration project', *Energy Procedia* **105**(Supplement C), 2549 – 2556. 8th International Conference on Applied Energy, ICAE2016, 8-11 October 2016, Beijing, China.

URL: *http://www.sciencedirect.com/science/article/pii/S1876610217307968*

Dahl, M. (2015), Power-Flow Modeling in Complex Renewable Electricity Networks, Master's thesis, Aarhus University, Denmark.

Danish standard 469, . (2013), '*Danish standard, DS 469-Heating and cooling systems in buildings, Denmark*'.

Domestic Building Regulations, . (2013), '*Domestic building services compliance guide*', https://www.gov.uk/government/publications/conservation-of-fuel-and-power-approved-document-l.

Dominković, D., Bačeković, I., Sveinbjörnsson, D., Pedersen, A. & Krajačić, G. (2017), 'On the way towards smart energy supply in cities: The impact of interconnecting geographically distributed district heating grids on the

energy system', *Energy* .

URL: *http://www.sciencedirect.com/science/article/pii/S0360544217303456*

EIA, U.S. Energy Information Administration (2015), 'Levelized Cost and Levelized Avoided Cost of New Generation Resources in the Annual Energy Outlook 2015'. Accessed: 2015-09-10.

URL: *http://www.eia.gov/forecasts/aeo/electricitygeneration.cfm*

Environment Agency UK, . (July 2014), '*An Environmental Risk Assessment for coal bed methane, coal mine methane and abandoned mine methane operations in England*'.

EQUA Simulation AB, . (2010), '*Validation of IDA Indoor climate and Energy 4.0 with respect to CEN Standards EN 15255-2007 and EN 15265-2007*'.

EQUA Simulation AB, . (2013), '*User Manual: IDA Indoor Climate and Energy, Version 4.5*'.

European-Commission (2002), '*Directive 2002/91/EC of the European parliament and of the council on the energy performance of buildings*', https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:2002L0091:20081211:EN:PDF. Accessed: 30 April 2020.

Frederiksen, S. & Werner, S. (2013), *District Heating and Cooling*, Studentlitteratur AB.

Fürsch, M., Golling, C., Nicolosi, M., Wissen, R. & Lindenberger, D. D. (2010), '*European RES-E Policy Analysis*', http://www.ewi.uni-koeln.de/fileadmin/user_upload/Publikationen/Studien/Politik_und_Gesellschaft/2010/EWI_2010-04-26_RES-E-Studie_Teil1.pdf.

Gabrielaitiene, I., Bøhm, B. & Sunden, B. (2010), 'Evaluation of approaches for modeling temperature wave propagation in district heating pipelines', *Heat Transfer Engineering* **29**(1), 45–56.

Géron, A. (2019), *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, O'Reilly Media.

Grundahl, L., Nielsen, S., Lund, H. & Möller, B. (2016), 'Comparison of district heating expansion potential based on consumer-economy or socio-economy', *Energy* **115, Part 3**, 1771 – 1778.
URL: *http://www.sciencedirect.com/science/article/pii/S0360544216307137*

Hansen, K., Connolly, D., Lund, H., Drysdale, D. & Thellufsen, J. Z. (2016), 'Heat roadmap Europe: Identifying the balance between saving heat and supplying heat', *Energy* **115, Part 3**, 1663 – 1671. Sustainable Development of Energy, Water and Environment Systems.
URL: *http://www.sciencedirect.com/science/article/pii/S0360544216308064*

Hastie, T., Tibshirani, R. & Friedman, J. (2009), *The elements of statistical learning: data mining, inference, and prediction*, Springer Science & Business Media.

Heide, D., Greiner, M., von Bremen, L. & Hoffmann, C. (2011), 'Reduced storage and balancing needs in a fully renewable European power system with excess wind and solar power generation', *Renewable Energy* **36**(9), 2515 – 2523.

Heide, D., von Bremen, L., Greiner, M., Hoffmann, C., Speckmann, M. &

Bofinger, S. (2010), 'Seasonal optimal mix of wind and solar power in a future, highly renewable Europe', *Renewable Energy* **35**(11), 2483 – 2489.
URL: *http://www.sciencedirect.com/science/article/pii/S0960148110001291*

Holloway, S., Jones, N., Creedy, D. & Garner, K. (2005), Can new technologies be used to exploit the coal resources in the yorkshire-nottinghamshire coalfield?, *in* J. Collinson, D. Evans, D. Holliday & N. Jones, eds, 'Carboniferous hydrocarbon geology : the southern North Sea and surrounding onshore areas', number 7, Yorkshire Geological Society, pp. 195–208.
URL: *http://nora.nerc.ac.uk/id/eprint/15896/*

Ianakiev, A., Cui, J. M., Garbett, S. & Filer, A. (2017), 'Innovative system for delivery of low temperature district heating', *International Journal of Sustainable Energy Planning and Management* **12**(0), 19–28.
URL: *https://journals.aau.dk/index.php/sepm/article/view/1633*

IEA & IRENA (2017), '*Perspectives for the Energy Transition: Investment Needs for a Low-Carbon Energy System*'.
URL: *https://www.irena.org/publications/2017/Mar/Perspectives-for-the-energy-transition-Investment-needs-for-a-low-carbon-energy-system*

IES Virtual Environment, . (2014), '*Building Regulations - Part L2 (2013) ApacheSim (DSM) User Guide*'.

Jacob, M., Neves, C. & Vukadinović Greetham, D. (2020), *Forecasting and Assessing Risk of Individual Electricity Peaks*, Springer.

Jacobson, M. Z. & Delucchi, M. A. (2011), 'Providing all global energy with wind, water, and solar power, part i: Technologies, energy resources, quan-

tities and areas of infrastructure, and materials', *Energy Policy* **39**(3), 1154 – 1169.

URL: *http://www.sciencedirect.com/science/article/pii/S0301421510008645*

Jones, N., Holloway, S., Creedy, D., Garner, K., Smith, N. J. P., Browne, M. & Durucan, S. (2004), '*UK Coal Resource for New Exploitation Technologies Final Report, Sustainable Energy & Geophysical Surveys Programme Commissioned Report CR/04/015N* ', http://nora.nerc.ac.uk/id/eprint/509526/1/CR04015N.pdf.

Kelly, S. (2011), 'Do homes that are more energy efficient consume less energy?: A structural equation model of the english residential sector', *Energy* **36**(9), 5610 – 5620.

URL: *http://www.sciencedirect.com/science/article/pii/S0360544211004579*

Kelly, S., Crawford-Brown, D. & Pollitt, M. G. (2012), 'Building performance evaluation and certification in the UK: Is SAP fit for purpose?', *Renewable and Sustainable Energy Reviews* **16**(9), 6861 – 6878.

URL: *http://www.sciencedirect.com/science/article/pii/S1364032112004595*

Küçüka, S. (2007), 'The thermal effects of some control logics used in GDHS', *Applied Thermal Engineering* **27**(8–9), 1495 – 1500.

URL: *http://www.sciencedirect.com/science/article/pii/S1359431106003401*

Liu, X., Wu, J., Jenkins, N. & Bagdanavicius, A. (2016), 'Combined analysis of electricity and heat networks', *Applied Energy* **162**(Supplement C), 1238 – 1250.

URL: *http://www.sciencedirect.com/science/article/pii/S0306261915001385*

Lund, H., Thellufsen, J., Aggerholm, S., Wittchen, K., Nielsen, S., Mathiesen, B. & Möller, B. (2015), 'Heat saving strategies in sustainable smart energy systems', *International Journal of Sustainable Energy Planning and Management* **4**(0), 3–16.

URL: *https://journals.aau.dk/index.php/sepm/article/view/684*

Lund, H., Werner, S., Wiltshire, R., Svendsen, S., Thorsen, J. E., Hvelplund, F. & Mathiesen, B. V. (2014), '4th generation district heating (4GDH): Integrating smart thermal grids into future sustainable energy systems', *Energy* **68**, 1 – 11.

URL: *http://www.sciencedirect.com/science/article/pii/S0360544214002369*

Lund, H., Østergaard, P., Connolly, D., Ridjan, I., Mathiesen, B., Hvelplund, F., Thellufsen, J. & Sorknæs, P. (2016), 'Energy storage and smart energy systems', *International Journal of Sustainable Energy Planning and Management* **11**(0), 3–14.

URL: *https://journals.aau.dk/index.php/sepm/article/view/1574*

Lund, R. & Mohammadi, S. (2016), 'Choice of insulation standard for pipe networks in 4th generation district heating systems', *Applied Thermal Engineering* **98**(Supplement C), 256 – 264.

URL: *http://www.sciencedirect.com/science/article/pii/S135943111501385X*

Makridakis, S., Spiliotis, E. & Assimakopoulos, V. (2018), 'Statistical and machine learning forecasting methods: Concerns and ways forward', *PloS one* **13**(3).

URL: *https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0194889*

Mathiesen, B., Lund, H., Connolly, D., Wenzel, H., Østergaard, P., Möller,

B., Nielsen, S., Ridjan, I., Karnøe, P., Sperling, K. & Hvelplund, F. (2015), 'Smart energy systems for coherent 100% renewable energy and transport solutions', *Applied Energy* **145**, 139 – 154.
URL: *http://www.sciencedirect.com/science/article/pii/S0306261915001117*

McIntyre, D. A. (1986), 'Output of radiators at reduced flow rate', *Building Services Engineering Research and Technology* **7**(2), 92–95.
URL: *https://doi.org/10.1177/014362448600700206*

McKinsey, KEMA, at Imperial College London, T. E. F. L., Economics, O. & ECF (2010), '*ROADMAP 2050: A practical guide to a prosperous, low-carbon Europe*'.

Ministry of Housing, C. & Government, L. (2015), '*Improving the energy efficiency of our buildings*', https://www.gov.uk/government/collections/energy-performance-certificates. Accessed: 30 April 2020.

Müller, A. C., Guido, S. et al. (2016), *Introduction to machine learning with Python: a guide for data scientists*, " O'Reilly Media, Inc.".

Möller, B. & Nielsen, S. (2014), 'High resolution heat atlases for demand and supply mapping', *International Journal of Sustainable Energy Planning and Management* **1**, 41–58.
URL: *https://journals.aau.dk/index.php/sepm/article/view/548*

National Calculation Methodology, . (2013), '*National Calculation Methodology (NCM) modelling guide (for buildings other than dwellings in England), Department for Communities and Local Government: London*', https://www.uk-ncm.org.uk.

Non-Domestic Building Regulations, . (2013), '*Non-domestic building services compliance guide*', https://www.gov.uk/government/publications/conservation-of-fuel-and-power-approved-document-l.

Olivier, J. G., (IES-JRC), G. J.-M., (IES-JRC), M. M. & (PBL), J. A. P. (2015), '*Trends in Global Co2 Emissions*', http://www.pbl.nl/en/publications/trends-in-global-co2-emissions-2015-report.

Olsen, P. K., Christiansen, C. H., Hofmeister, M., Svendsen, S., Rosa, A. D., Thorsen, J.-E., Gudmundsson, O. & Brand, M. (2014), '*Guidelines for Low-Temperature District Heating*'.

Palsson, H. (1997), Analysis of numerical methods for simulating temperature dynamics in district heating pipes, *in* '*Proceedings of the 6th International symposium on district heating and cooling simulation*', University of Iceland, Faculty of Engineering, pp. 1–20.

Pensini, A., Rasmussen, C. N. & Kempton, W. (2014), 'Economic analysis of using excess renewable electricity to displace heating fuels', *Applied Energy* **131**, 530 – 543.
URL: *http://www.sciencedirect.com/science/article/pii/S0306261914004772*

Persson, U., Möller, B. & Werner, S. (2014), 'Heat roadmap Europe: Identifying strategic heat synergy regions', *Energy Policy* **74**, 663 – 681.
URL: *http://www.sciencedirect.com/science/article/pii/S0301421514004194*

Poutiainen, Z. (2019), 'Short-term heat load forecasting in district heating systems: A comparative study of various forecasting methods'.

Rasmussen, M. G., Andresen, G. B. & Greiner, M. (2012), 'Storage and

balancing synergies in a fully or highly renewable pan-European power system', *Energy Policy* **51**, 642 – 651. Renewable Energy in China.

URL: *http://www.sciencedirect.com/science/article/pii/S0301421512007677*

Rasmussen, U. (2012), 'Water consumption in the energy sector and energy consumption in the water-sector in a danish municipality.', *Journal of Transdisciplinary Environmental Studies* **11**(1), 3 – 5.

Rodriguez, R. A., Becker, S. & Greiner, M. (2015), 'Cost-optimal design of a simplified, highly renewable pan-European electricity system', *Energy* **83**, 658 – 668.

URL: *http://www.sciencedirect.com/science/article/pii/S0360544215002212*

Rodríguez, R. A., Becker, S., Andresen, G. B., Heide, D. & Greiner, M. (2014), 'Transmission needs across a fully renewable european power system', *Renewable Energy* **63**, 467 – 476.

URL: *http://www.sciencedirect.com/science/article/pii/S0960148113005351*

SBEM (2010), '*Simplified building energy model - Building Research Establishment (BRE)*'.

Schaber, K., Steinke, F., Mühlich, P. & Hamacher, T. (2012), 'Parametric study of variable renewable energy integration in Europe: Advantages and costs of transmission grid extensions', *Energy Policy* **42**, 498 – 508.

URL: *http://www.sciencedirect.com/science/article/pii/S0301421511010081*

Schlachtberger, D., Becker, S., Schramm, S. & Greiner, M. (2016), 'Backup flexibility classes in emerging large-scale renewable electricity systems',

Energy Conversion and Management .

URL: *http://www.sciencedirect.com/science/article/pii/S0196890416302606*

Siami-Namini, S. & Namin, A. S. (2018), 'Forecasting economics and financial time series: Arima vs. lstm', *arXiv preprint arXiv:1803.06386* .

Steward, D., Saur, G., Penev, M. & Ramsden, T. (2009), 'Lifecycle cost analysis of hydrogen versus other technologies for electrical energy storage', *NREL* .

URL: *http://www.nrel.gov/docs/fy10osti/46719.pdf*

Thellufsen, J. Z. & Lund, H. (2015), 'Energy saving synergies in national energy systems', *Energy Conversion and Management* **103**, 259 – 265.
URL: *http://www.sciencedirect.com/science/article/pii/S0196890415005932*

Thellufsen, J. Z. & Lund, H. (2017), 'Cross-border versus cross-sector interconnectivity in renewable energy systems', *Energy* **124**, 492 – 501.
URL: *http://www.sciencedirect.com/science/article/pii/S0360544217302943*

Tunzi, M., Østergaard, D. S., Svendsen, S., Boukhanouf, R. & Cooper, E. (2016), 'Method to investigate and plan the application of low temperature district heating to existing hydraulic radiator systems in existing buildings', *Energy* **113**, 413 – 421.
URL: *http://www.sciencedirect.com/science/article/pii/S0360544216309574*

UKERC, . (2009), '*Decarbonising the UK Energy System: Accelerated Development of Low Carbon Energy Supply Technologies*', http://www.ukerc.ac.uk/publications/

`decarbonising-the-uk-energy-system-accelerated-development-of-low-carbon-ener`
`html`. Accessed: 11 May 2020.

Ward, I. (1991), 'Domestic radiators: Performance at lower mass flow rates and lower temperature differentials than those specified in standard performance tests', *SAGE Publications Ltd STM- Building Services Engineering Research and Technology* **12**(3), 87 – 94.
URL: *http://journals.sagepub.com/doi/pdf/10.1177/014362449101200301*

Wedding, G. C. & Crawford-Brown, D. (2007), 'An Analysis of Variation in the Energy-Related Environmental Impacts of LEED Certified Buildings', *Journal of Green Building* **2**(4), 151–170.
URL: *https://doi.org/10.3992/jgb.2.4.151*

Wedding, G. C. & Crawford-Brown, D. (2008), 'Improving the Link Between the LEED Green Building Label and a Building's Energy-Related Environmental Metrics', *Journal of Green Building* **3**(2), 85–105.
URL: *https://doi.org/10.3992/jgb.3.2.85*

Wiltshire, R., Williams, J. & Woods, P. (2014), '*A technical guide to district heating*'.

Woods, P. & Overgaard, J. (2016), Historical development of district heating and characteristics of a modern district heating system, *in* R. Wiltshire, ed., '*Advanced District Heating and Cooling (DHC) Systems*', Woodhead Publishing Series in Energy, Woodhead Publishing, Oxford, pp. 3 – 15.

Zhang, L., Gudmundsson, O., Thorsen, J. E., Li, H., Li, X. & Svendsen, S. (2016), 'Method for reducing excess heat supply experienced in typical chi-

nese district heating systems by achieving hydraulic balance and improving indoor air temperature control at the building level', *Energy* **107**, 431 – 442.

URL: *http://www.sciencedirect.com/science/article/pii/S0360544216303917*

Çengel, Y. A. (2007), *Heat & Mass Transfer: A Practical Approach*, McGraw-Hill Education Pvt Limited.

Çengel, Y. A., Cimbala, J. M. & Turner, R. H. (2012), *Fluid Mechanics [Si Units]*, McGraw-Hill Education Pvt Limited.