Cloud based development of novel data-driven algorithms for heat demand prediction to improve control of heat generation

Kevin Naik

Department of Civil Engineering

Nottingham Trent University

This thesis is submitted for the degree of

Doctor of Philosophy

School of Architecture, Design

and the Built Environment

October 2023

Declaration

Note

In Chapter 4 of my thesis, I delve into the specifics of the SEMS project, a collaboration between Nottingham Trent University, Imperial College London, University of Nottingham, and Siemens. The Energy Report features a paper on this project, which I cover in this chapter. My contribution to this paper was focused on the Heat Prediction Service, while Richard Charlesworth detailed and developed the Hypernetwork Theory and Sean Jones was responsible for the development of Electricity Demand.

Kevin Naik

October 2023

Publication

1. Naik, K. and Ianakiev, A., 2021. Heat demand prediction: A real-life data model vs simulated data model comparison. Energy Reports, 7, pp.380-388.

2. Jones, S., Charlesworth, R., Naik, K., Charlesworth, T., O'Dwyer, E., Ianakiev, A., Johnson, J., Boukhanouf, R., Gillott, M., Sellwood, V. and Aloor, J., 2021. A multi-energy system optimisation software for advanced process control using hypernetworks and a micro-service architecture. Energy Reports, 7, pp.167-175.

3. Naik, Kevin and Ianakiev, Anton, Optimising Heat Consumption at Micro-Level Using User-Centric Data-Driven Model. Available at SSRN: http://dx.doi.org/10.2139/ssrn.4108849

4. Naik, K., Ianakiev, A., Galadanci, A.S., Cucca, G., Shubham, Sun, M., 2023. Evaluating the Potential of Wind and Solar Energy in Achieving Zero Energy Ratings in Residential Homes: a Nottingham Case Study, in: Submitted to Smart Energy. Elsevier.

5. Shubham Shubham, Kevin Naik, Shivangi Sachar, Anton Ianakiev, Performance analysis of low Reynolds number vertical axis wind turbines using low-fidelity and mid-fidelity methods and wind conditions in the city of Nottingham, Energy, Volume 279, 2023.

Conference Presentation

1. Zero energy rating of residential homes leveraging wind and solar energy at Smart Energy System conference 2022
2. Heat demand prediction: A real-life data model vs simulated data model comparison at District Heating and Cooling Symposium 2021
3. Optimising heat consumption at micro-level using user centric data driven model at Smart Energy System conference 2021
4. An open source industrial 4.0 framework for district heating systems at ADBE School Conference 2020

Teaching Pedagogy

1. Siena, F.L., Russell, S., Malcolm, R., Brook, E.E., Cutts, A., Martin, L. and Naik, K., 2022. SUCCESSES & CHALLENGES OF SUPPORTING PRODUCT DESIGN EDUCATION FOR DEAF & HARD OF HEARING LEARNERS DURING A PANDEMIC: A CASE STUDY. In DS 117: Proceedings of the 24th International Conference on Engineering and Product Design Education (E&PDE 2022), London South Bank University in London, UK. 8th-9th September 2022.
2. Siena, F.L., Naik, K. and Watt, P, 2023. Utilising Hackathons To Enhance The Design Process For Electronics & Programming Learning Within Product Design Education. In DS 123: Proceedings of the 25th International Conference on Engineering and Product Design Education (E&PDE 2023).

Kevin Naik

October 2023

Acknowledgements

I would like to express my sincerest gratitude to Siemens and NTU for generously funding my PhD studies, without which this thesis would not have been possible.

I would like to extend my heartfelt appreciation to my DOS Professor Anton Ianakiev for his invaluable support, guidance, motivation, and advice throughout my PhD journey. I am also immensely grateful to my entire supervisory team, including Prof Michael White, Evtim Peytch, and Independent Assessor Dr. Robert Evans, for their constructive comments and insightful discussions that greatly contributed to the quality of my work. Additionally, I would like to acknowledge the support of Sean Jones and Richard Charlesworth from Siemens, who provided me with invaluable industrial and academic guidance through their involvement in the SEMS project.

I would also like to express my gratitude to my beloved wife, Payal Naik, and daughter, Meray Naik, for their unwavering support and understanding during this demanding period. I owe a debt of gratitude to my parents, Bhanu Naik and Kiran Naik, whose blessings and encouragement gave me the courage to pursue this challenging endeavor in a foreign land.

Furthermore, I am deeply appreciative of the help and care provided by my in-laws, Nilan Vashi and Alka Vashi, who took care of my family and household responsibilities, allowing me to focus on my research.

Special thanks go to Dhara Joshi and Mit Raval for their meticulous proofreading and editing of my work. Lastly, I am grateful to my dear friends

Sushmita Borah, Shubham, Arijit Sen, Sharmila Mujmudar, and Sherna Salim for their unwavering support, guidance, and encouragement throughout my PhD journey.

In conclusion, I am thankful to all the individuals and organizations who have supported me in this challenging and rewarding journey, and I look forward to continuing to learn and grow in my future endeavors.

Funding

1. Siemens (UK) and Nottingham Trent University provided funding for the PhD, thereby supporting this study.

2. Grant agreement No. 646511 of the European Union Horizon 2020 REMOURBAN research and innovation program supported this study.

3. The European Union's Horizon 2020 research and innovation program provided funding for a section of the study under Grant Agreement No. 691895.

4. Grant agreement No. EC grant 860101 of the European Union Horizon 2020 Marie Curie zEPHYR research and innovation program provided funding for a particular section of the study.

Abstract

This thesis explores the potential of machine learning algorithms to improve heat demand prediction in district heating systems. The study compares the performance of various machine learning methods on datasets generated for SEMS (EU H2020 SHARINGCITIES project) and real-life data collected from EU H2020 REMOURBAN project in Nottingham. The thesis demonstrates the importance of selecting the appropriate algorithm for the specific dataset and highlights the potential of machine learning in district heating. Further research in this area could lead to more efficient and sustainable energy use by optimising heat demand prediction in district heating systems.

As proof of concept, an Internet of Things (IoT) framework is implemented in this thesis. The thesis also introduces a novel human-centric approach to utilise real-life data to create an individual heat profile for district heating users and generate realistic individual heat demand of the heating system to set up an optimum heat generation mode.

One of the outcomes of the thesis is that real-time data-driven heat demand prediction is undertaken using more than one machine learning algorithms simultaneously. Additionally, a simulation tool has been developed for data analysis and can be used for further data exploration.

The thesis further investigates the theoretical approach to calculate the heat prediction of the individual homes based on customer behaviour, which can be used to make critical decisions about the energy performance of buildings. This approach is human-centric, as customer behaviour is driven by socio-economic conditions. Finally, the thesis illustrates the cost of optimisation of district heating systems by developing a test rig of a smart radiator using the IoT framework with a novel control strategy.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

1.1 Research Background

In recent times, the energy market has undergone a significant evolution with respect to technological advancements, resource utilisation, and consumption patterns. The rise in energy demand on a daily basis has led to the emergence of various opportunities to optimise heat generation. The continuous improvement of planning and optimising energy generation is critical for technological advancements. In the United Kingdom (UK), District Heating (DH) has become increasingly popular in recent times. The UK government is taking proactive measures to install and implement DH systems due to their efficiency when compared to traditional heating systems. Furthermore, they cause less carbon emissions and promote the reuse of heat from various industries [10, 11]. The adoption of renewable energy sources in the energy sector has further created new opportunities for the development and exploration of technical solutions. The simultaneous use of multiple energy sources has presented a challenge in the optimisation of energy generation, which can be categorized as a multi-vector demand problem [12, 13]. There has been a global trend among various governments to adopt District Heating (DH) systems. This trend is attributed to the numerous benefits associated with the use of these systems, including increased efficiency, decreased carbon emissions, and the ability to recycle heat from industries. By promoting the adoption of DH systems, governments hope to contribute to the achievement of global climate goals and reduce the environmental impact of energy generation. Moreover, the evolution of heat networks from the first generation to the fifth generation has brought about increased energy efficiency. The first generation of DH systems was characterized by high network temperatures and high heat losses, resulting in a low level of efficiency. However, with the development of more advanced systems, DH networks have become increasingly efficient, with the fifth generation characterized by low network temperatures, minimal heat loss, and greater flexibility in the integration of renewable energy sources [11, 14, 15]. Overall, the adoption of DH systems and the evolution of heat networks represent significant steps towards the development of sustainable and environmentally-friendly energy systems. Through continued innovation and research, DH systems are poised to become even more efficient and effective, providing a reliable source of energy for years to come. To date, research and development efforts in the District Heating (DH) sector have primarily focused on innovations in hardware and heat sources, such as coal waste, biomass, and solar technologies. In contrast, existing software solutions for DH systems are primarily focused on monitoring and control, with little emphasis on optimisation. Despite the lack of optimisation solutions, significant research has been conducted to develop tools and libraries that can be utilised for control optimisation. Indeed, optimisation is an established science, and much of the groundwork has already been established for the use of these tools in DH systems.

Machine learning is an excellent example of an optimisation technique that can be applied to DH systems. It provides a powerful means of identifying patterns and trends in energy consumption, which can be used to develop predictive models that optimise energy generation and distribution. By leveraging machine learning, it may be possible to create more efficient and effective DH systems that are better suited to meet the needs of modern energy consumers [16].

While DH systems have made significant strides in recent years, the development of software solutions that optimise energy generation and distribution is an area that requires more attention. The application of machine learning to DH systems presents an exciting opportunity for advancing the state of the art in DH software development. Continued exploration of machine learning, as well as other optimisation techniques, will help to unlock new possibilities in the optimisation of DH systems and pave the way for a more sustainable and efficient energy future.

The Internet of Things (IoT) has emerged as a critical interdisciplinary field that integrates Cloud Computing, Embedded Systems, and Data Sciences. Since its inception in 2010, IoT has gained tremendous popularity, owing to its ability to enable planning, scheduling, optimisation, and prediction, while also providing insights from the data collected. As infrastructure continues to age and demand for energy generation, storage, distribution, and demand management in cities increases, IoT has become an increasingly important tool for transforming these systems [2].

Energy systems incorporate a vast array of technologies, including power conversion systems, such as power plants, distributed renewables, transmission systems, power transmission lines, gas lines, and end-user systems [3]. Given the pressing need to save energy in every possible form, it is critical to leverage advanced technologies like IoT to optimise energy utilisation and distribution.

By integrating IoT technologies, we can develop new energy management systems that are more efficient, cost-effective, and eco-friendly. IoT can be used to collect, analyse, and respond to real-time data related to energy usage and consumption. For example, IoT sensors can be deployed to monitor energy consumption patterns in buildings, homes, and cities, which can be analysed to develop predictive models for optimising energy usage. This approach can help reduce energy waste, lower operational costs, and promote sustainability [16].

Overall, IoT represents a promising technology that can be utilised to drive significant advancements in energy management systems. By leveraging the insights provided by IoT technologies, we can transform energy generation, storage, distribution, and demand management systems to build a more sustainable and efficient future.

This thesis puts forth an innovative framework that leverages the potential of the Internet of Things (IoT) for real-time heat prediction and implementation. The framework is designed to facilitate the selection of the best performance configuration by incorporating multiple prediction algorithms. It also demonstrates the efficacy of long-term prediction. Importantly, the framework illuminates new research avenues that emerge due to its use.

The motivation behind the use of IoT in this context is its proven accessibility and reliability, making it an ideal technological solution. The proposed framework is readily accessible, and as an open-source technology, it is available on demand, effortless to use, and economical.

The concept of Industry 4.0 revolves around machine-to-machine communication facilitated by embedded computing, enabling decision-making without human intervention. Industry 4.0 is a convergence of technologies such as Cyber-Physical Systems (CPS), the Internet of Things (IoT), Industrial Internet of Things (IIOT), Cloud Computing, cognitive computing, and Artificial Intelligence. At its core, industry 4.0 is reliant on cloud computing.

Considering the fundamental principles of Industry 4.0, various cloud platforms have been developed, including AWS, Azure [17], Bluemix [18], SalesForce, Google cloud platform [19], GE Predix, and Siemens MindSphere. After conducting a thorough comparison, it has been determined that Salesforce's Heroku is the preferred cloud service for this research. This is because it provides superior capabilities for deploying applications compared to other cloud services.

This thesis designed a data collection system that functions in real-time on the cloud and applies analysis to optimise or predict heat load. Additionally, explored the potential of controlling the heat source through the cloud-based system. By integrating cloud computing with data analysis and control mechanisms, to improve the efficiency and effectiveness of heat generation and management.

This thesis presents a novel approach to estimating heat requirements for individual users based on a human-centric perspective. The unique requirements and usage schedules of each customer are taken into account in order to provide more accurate heat demand predictions. The author builds upon previous work, where machine learning algorithms were used to predict heat demand from historical data, and extends it by implementing an IoT framework for district heating systems. The framework allows for more real-life testing of heat prediction algorithms and plays a crucial role in the optimisation of heat generation. By estimating heat demand more accurately, heat losses can be reduced, leading to more efficient district heating systems. The thesis presents two methodologies: a novel approach to heat demand estimation and the optimisation of district heating through the implementation of an IoT framework and home automation.

1.2 Research Gap

This section outlines the research gaps that have been identified through a thorough analysis of the existing literature. These include:

1. The need to enhance load prediction through the development and application of data-driven models that can accurately predict heat demand in real-time.

2. The requirement for a robust data collection and storage system that is capable of seamlessly integrating with cloud-based technology to ensure that accurate and timely data can be accessed and analysed as needed.

3. The potential for cloud computing to be utilised in the analysis and control of heat generation in DH systems, thereby optimising their efficiency and minimising energy waste.

1.3 Aim

The primary aim of this research is to develop a novel data-driven algorithm for demand prediction, utilising data obtained from sensors of the REMOURBAN project. The focus is on the improvement of control of heat generation on Cloud by leveraging the real-time data collection and storage capabilities provided by Cloud computing.

1.4 Objective

The primary objective of this research is to develop a comprehensive data collection and storage system for sensor data in real-time. The collected data will be utilised to train and evaluate different data-driven models for load prediction. The following objectives have been identified to achieve the primary objective:

1. To design and develop a reliable and efficient data collection and storage system for sensor data in real-time.

2. To utilise the collected data for load prediction by training and evaluating different data-driven models.

3. To validate the performance of the developed data-driven models with new data collected.

4. To deploy the validated data-driven models on Cloud for real-time analysis and control of heat generation.

5. To generate control signals for heat generation plant from a validated model to ensure the efficient and sustainable generation of heat for end-users.

1.5 Thesis Structure

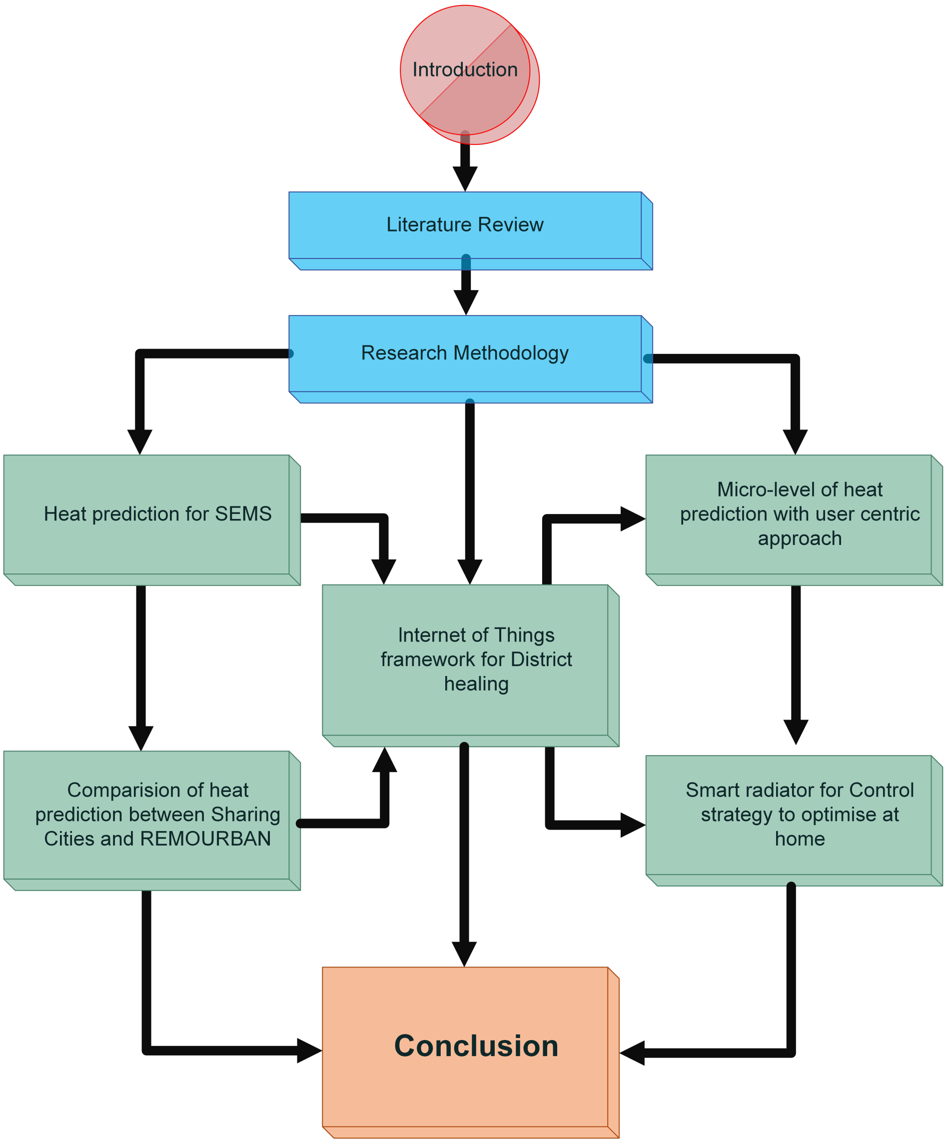


Fig. 1.1 Thesis Structure

1.6 PhD Journey

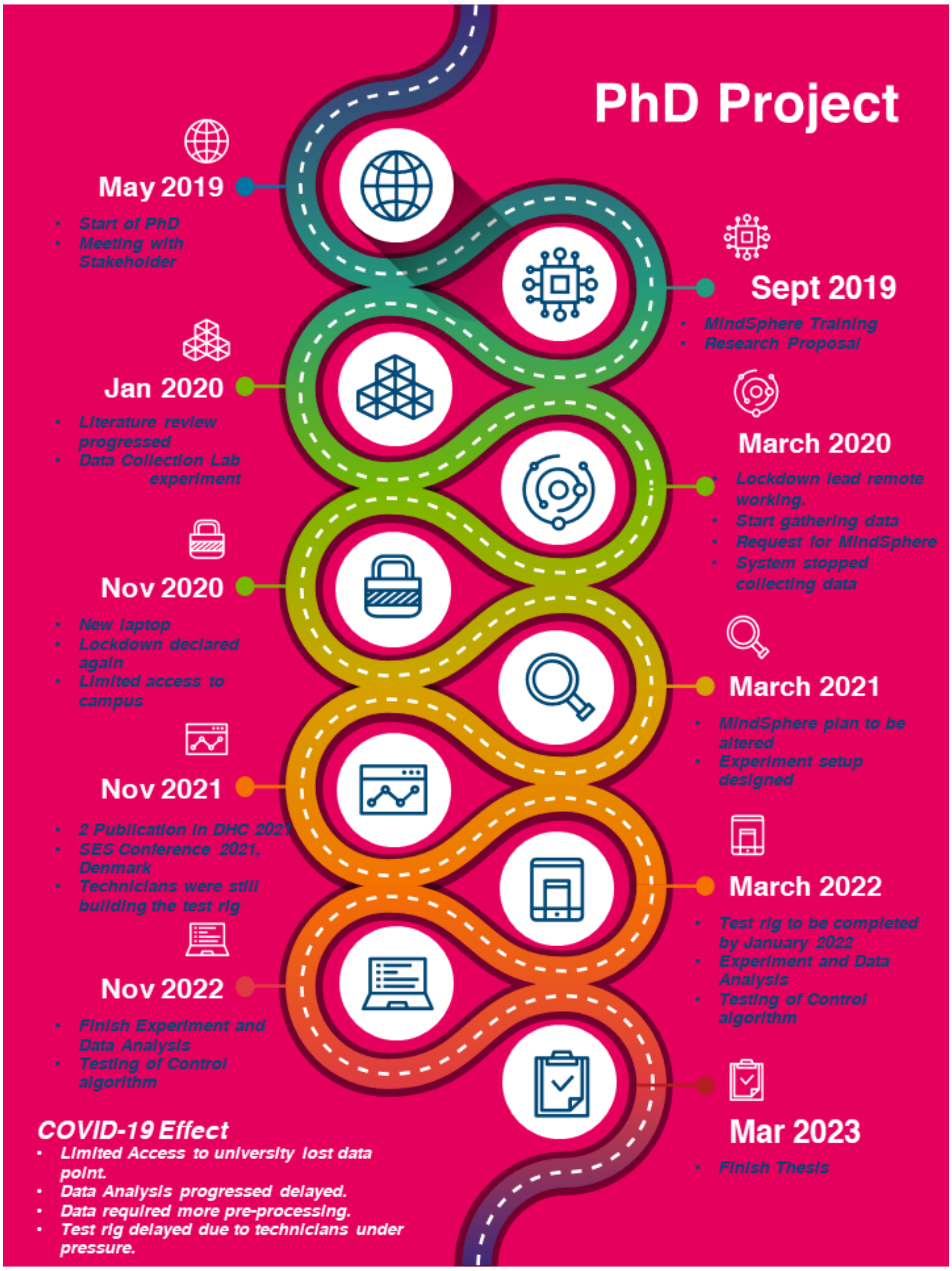


Fig. 1.2 Overview of PhD Journey

Chapter 2

Literature review

This section aims to provide a comprehensive review of the relevant literature on three key sub-topics: the evolution of district heating systems, load prediction, and cloud services. Through the analysis of research papers, articles, and websites, this section aims to identify gaps in the existing literature and consolidate the knowledge in these areas.

The sub-topic on the evolution of district heating systems provides an overview of the historical development of these systems. This includes an examination of the key milestones in the evolution of district heating systems, as well as an analysis of the various technologies and approaches that have been used in the design and implementation of these systems. By reviewing the evolution of district heating systems, this section seeks to provide a comprehensive understanding of the state-of-the-art in this field and identify any areas where further research is needed.

The second sub-topic, load prediction, explores the different methods that have been developed for predicting the demand for district heating services. This includes an examination of the various factors that impact the demand for district heating, as well as an analysis of the various models and techniques that have been developed for predicting future demand. Through a thorough review of the available literature, this section aims to identify the strengths and weaknesses of existing load prediction models and highlight areas where further research is needed to improve the accuracy and effectiveness of these models.

The final sub-topic, cloud services, examines the use of cloud computing in district heating systems. This includes an analysis of the various cloud-based services that can be used to improve the efficiency and reliability of district heating systems, as well as an evaluation of the benefits and challenges associated with the use of cloud-based technologies in this context. By providing a comprehensive review of the literature on cloud services in district heating, this section aims to provide insights into the current state-of-the-art in this field and identify areas where further research is needed to advance the use of cloud computing in district heating systems. Overall, this section seeks to consolidate the existing literature on the evolution of district heating systems, load prediction, and cloud services, and identify any gaps in the current knowledge. By providing a comprehensive review of the literature, this section aims to contribute to the development of more efficient and effective district heating systems, with the potential to benefit society and the environment.

2.1 Evolution of District Heating

District Heating is a heating method that has its roots in ancient Rome, where heated baths were a common feature. The concept has evolved over time, and today it refers to a system that transports heat through hot water in a network of pipes that are connected to a central station. This section provides an overview of the evolution of district heating from its initial stages to the present day, outlining the development of 1^{st} to 5^{th} generation district heating systems.

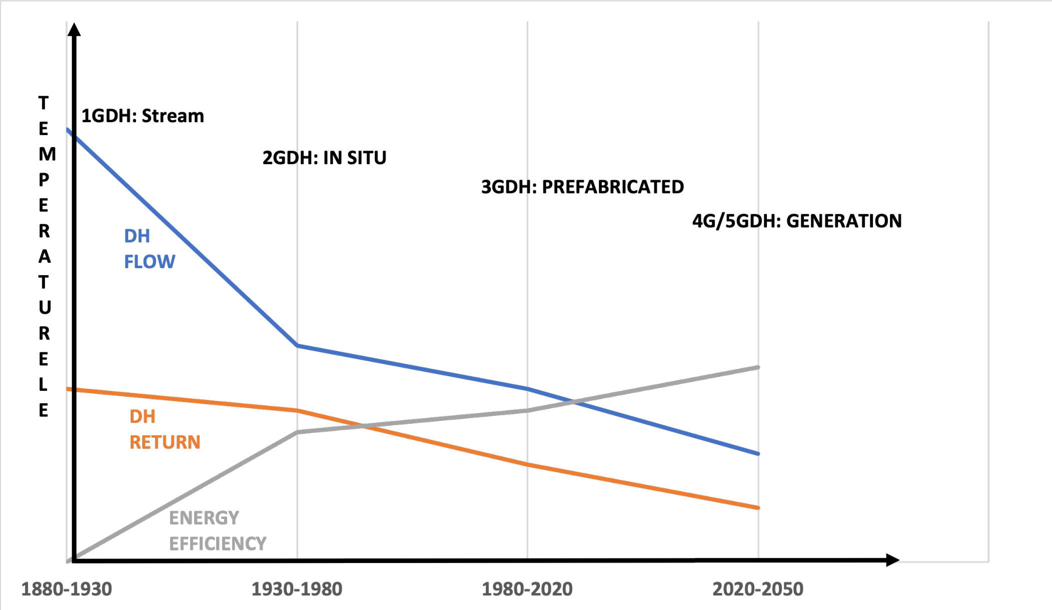


Fig. 2.1 History of District Heating [1]

2.1.1 1^{st} Generation

The first generation of district heating systems emerged in the late 19^{th} and early 20^{th} centuries, and involved the use of steam generated from coal-fired power plants to supply heat to nearby buildings. This approach was effective, but had limited efficiency and environmental sustainability.

The heat transfer medium used in the district heating system operates at high temperatures of up to 300°C and pressures of 20 bar. This high operating temperature and pressure result in a higher risk of corrosion and also contribute to significant heat loss. These factors render such systems outdated and inefficient. Despite these limitations, there are still some district heating systems that continue to operate with these conditions. For example, the Parisian Urban Heating Company operates a district heating system in Paris that uses these high-temperature and high-pressure conditions to distribute heat to buildings [15, 10, 1]).

However, it is important to note that there are potential risks associated with the use of such outdated systems. The high operating temperature and pressure levels can lead to safety hazards, and the system's inefficiency can result in higher energy costs and environmental impacts. Therefore, it is important to consider modern and more sustainable alternatives for district heating systems, such as lower operating temperatures and the use of renewable energy sources. These solutions can help to reduce costs, improve energy efficiency, and reduce environmental impacts.

2.1.2 2 nd Generation

The second generation of district heating systems emerged in the 1960s and 1970s, and involved the use of hot water as a heat carrier. This was a more efficient and sustainable approach, and it allowed for the integration of renewable energy sources.

The period following the Second World War saw significant development in the field of district heating as part of reconstruction efforts. One major improvement was the introduction of two-pipe closed-loop systems, which consisted of a high-temperature water pipe under high pressure and a second pipe for the return of condensed water to the central station. These systems, collectively known as second-generation district heating (2GDH), were instrumental in improving the hydraulic pumps and facilitating easier management of the fluid. As a result, the overall efficiency of district heating systems increased by approximately 50%.

Despite these improvements, 2GDH systems still have some limitations. For instance, the high capital expenditure (CAPEX) required to install the system, including thermal storage tanks and heat exchangers, presents a significant challenge. Additionally, the system suffers from high heat loss, which further reduces overall efficiency. Moreover, the lack of control over the temperature and flow of the heating medium is a significant limitation.

To overcome these limitations, alternative approaches, such as the use of smart grid technologies and thermal energy storage, have been proposed. These modern solutions offer improved control over the heating medium and the ability to integrate renewable energy sources, which can further enhance the efficiency and sustainability of district heating systems. Overall, these developments demonstrate the ongoing efforts to improve district heating systems and overcome the limitations of earlier approaches [15, 10, 1]).

2.1.3 3^{rd} Generation

The emergence of third-generation district heating systems in the 1990s marked a significant milestone in the evolution of district heating. This approach involved the integration of combined heat and power (CHP) plants to generate both electricity and heat, thereby enhancing energy efficiency, reducing carbon emissions, and improving overall system reliability.

One of the major advantages of third-generation district heating systems is their operating temperature, which is lower than that of earlier generations. As district heating gained popularity in the 1970s, energy-efficient and cost-effective heat exchangers were developed to facilitate lower operating temperatures. Additionally, the pipes used in these systems are made of more compact materials and have better thermal insulation, which significantly reduces heat loss. These advancements have resulted in district heating systems that are easy to maintain and have a high uptime, which has led to their widespread adoption in many developed countries [15, 10, 1].

Third-generation district heating systems also offer individual control, providing users with greater flexibility to manage their energy consumption. With continued development, 3G district heating has become the most commonly used template in the current scenario. This underscores the growing importance of sustainability, energy efficiency, and cost-effectiveness in district heating systems, which will likely continue to drive future advancements in the field.

2.1.4 4^{th} Generation

The emergence of the fourth generation of district heating systems in the 2000s marked a significant advancement in the field. This generation involved the integration of advanced technologies such as heat pumps and thermal storage systems, leading to greater flexibility and control, and improved system efficiency.

One of the core features of fourth-generation district heating is its focus on the reuse of wasted heat generated from various sources, such as industry, data centers, and others. This approach has led to the development of low-temperature district heating, which supplies heat at a temperature less than 70°C, thereby reducing heat loss. In this era of fourth-generation district heating, governments have played a critical role in promoting these systems, leading companies to invest in new technologies, and improve heat exchangers, pipes, and other components [15]. Moreover, the development of the fourth generation of district heating systems has opened up new research opportunities in the field of supply/demand management. There is a significant scope for research in load prediction, which is discussed in the next section. This underscores the growing importance of research in the field to address emerging challenges, improve system efficiency, and promote sustainability [15, 10, 1].

2.1.5 5^{th} Generation

The development of the fifth generation of district heating systems is a promising avenue for even greater energy efficiency and sustainability. The focus of this generation is on the integration of renewable energy sources, smart grid technologies, and energy storage systems. One of the main aspects of this approach is the further exploration of low-temperature district heating, which is aimed at reducing heat losses and reusing heat within buildings by installing a reversible heat pump in the building plant room. However, the implementation of such systems requires a significant capital investment, and the larger pipe diameter required and the additional electricity needed to operate heat pumps may also increase costs. Nonetheless, the theoretical potential for increased energy efficiency and sustainability is a strong incentive for further research and development. Figure 1 shows that the operating temperature decreases as the generation number increases, and the number of energy sources also increases, leading to greater efficiency [15, 10].

Overall, the evolution of district heating systems reflects a gradual shift towards greater efficiency, sustainability, and integration of advanced technologies. As the concept continues to evolve, it is likely that we will see even greater improvements in the performance and sustainability of these systems. The comprehensive review of the evolution of district heating from the 1st Generation to the 5th Generation provides valuable insights into the development of more efficient and sustainable heating systems. The review highlights some significant research gaps that can be addressed to improve district heating systems, such as the design and development of innovative materials for components, the development of new policies for implementation, and the optimisation of control strategies.

Furthermore, the review identifies load prediction and intelligent district heating as critical areas of research to enhance the efficiency of district heating systems. However, addressing all these research gaps in one initiative may be challenging, as the scope of research is vast. Therefore, this research focuses on load prediction and control strategy, which can significantly improve the efficiency of district heating systems.

This review emphasises the need for further research to improve the efficiency and sustainability of district heating systems, which can contribute to reducing greenhouse gas emissions and improving the quality of life for people in urban areas. The findings of this research can provide valuable insights into the development of more efficient and sustainable district heating systems, which can benefit the research community, policymakers, and the public.

Load prediction is a crucial aspect in district heating systems, as it plays a pivotal role in achieving optimal efficiency in heat generation. The accurate prediction of load demand is instrumental in guiding the decision-making process for the purchase and generation of heat. Furthermore, load prediction is vital in ensuring that heating systems operate smoothly and in a timely manner, facilitating effective scheduling, infrastructure development, and raw material planning.

2.2 Heat Prediction

Load prediction is essentially the forecasting of demand for heat, and as such, it is critical for the smooth functioning of district heating systems. To gain a comprehensive understanding of the existing research gap, an extensive literature review of various papers has been conducted in this section. The concept of load prediction for district-heating systems has been a critical area of research since the early 2000s. In 2002, [20] emphasized the importance of load prediction and proposed a basic model based on social behaviour and outdoor temperature [20]. In order to improve the operation of district-heating systems, it is necessary for energy companies to implement reliable optimisation routines, both computerized and manual, to enhance system operation. However, to develop a production plan for heat-producing units, an accurate prediction of heat demand is required. This can be achieved by taking into account the outdoor temperature and the social behaviour of consumers, as these factors have the most significant influence on demand. Based on this insight, a load prediction model is proposed in this thesis.

Several methodologies have been developed for heat-load forecasting, but many of them fail in practice due to a lack of measured data and uncertainties in weather forecasts. In such situations, a simpler model may provide predictions that are as good as those generated by more sophisticated approaches. This is evidenced by the analysis of applications discussed in this thesis.

The proposed method by Dotzauer, for predicting heat demand in a district heating system is based on the relationship between load and outdoor temperature and social behaviour of consumers. Despite the simple construction of the model, its predictions are comparable to those generated by other, more complex techniques. The paper presents both theoretical and practical results, which demonstrate the feasibility and effectiveness of the proposed approach [20].

Following Dotzauer's emphasis on load prediction in 2002, several researchers have explored load prediction for district heating. One approach used time-series analysis with the incorporation of outdoor temperature to predict heat demand for district heating [21]. Subsequently, with the increase in computational power, the application of machine learning algorithms became feasible, and many researchers began to utilise them for heat demand prediction in district heating. Thus, machine learning algorithms have become a popular tool for district heating companies to improve their load prediction and optimise their production plans.

In 2014-2015, researchers proposed a novel approach for heat load forecasting using an online machine learning algorithm that could automatically use new data for heat prediction. The main motivation behind proposing an online approach was to enable real-time processing, which is critical for district heating systems. Notable works in this area include those by [22] and [23]. The "An Online Machine Learning Algorithm for Heat Load Forecasting in District Heating Systems" by Provatas, discusses the importance of heat load forecasting in district heating optimisation, where energy companies need to estimate the amount of energy required to meet market demand while minimising peak boiler usage and optimising energy generation. An online machine learning algorithm is suggested for heat load forecasting using online bagging and the Fast Incremental Model Trees with Drift Detection (FIMT-DD) as the base model, along with a mechanism for handling missing values, measurement errors, and outliers. The algorithm is evaluated using operational data from the Karlshamn District Heating network, and two approaches for aggregating data from network nodes are investigated. The algorithm is shown to accurately forecast heat load with a mean absolute percentage error of 4.77%, making it a viable alternative to state-of-the-art algorithms for heat load forecasting. The algorithm is memory-efficient and can process data in real time, providing a concrete foundation for operational usage of online machine learning algorithms in the domain of district heating.

The paper by [22], discusses the importance of demand forecasting in district heating and cooling systems, as it plays a crucial role in overall energy efficiency efforts. The authors present the current status and results of their work in developing and implementing online machine learning algorithms for demand forecasting, which includes decision tree-based regression algorithms, neural network-based approaches, and ensemble solutions. The practical implementation and commissioning of the system in two different operational settings are also described. The authors found that the demand predictions have a robust behaviour within acceptable error margins and can be used in applications such as intelligent network controllers for district heating. However, the forecasting ability deteriorated when the models were confronted with scenarios not covered by their training data, and continuous re-training is necessary for the models to adapt to new datasets. Several potential improvements to increase the predictive ability of the system are identified, such as adding short-term historical data and the day of the year for seasonable purposes. The study also shows that artificial neural networks provide the best forecasting ability among the studied algorithms and can handle data outside the training dataset.

The paper entitled "A review of District Heating Systems: Modelling and optimisation" provides an overview of load prediction techniques used in the past. The article highlights the complexity level of district heating systems, which is measured by four parameters: the number of technologies utilised, the number of users, the temporal profile, and spatial concerns. Additionally, the paper categorizes district heating systems based on geographical conditions, scale, heat density, and end-user demand [24].

Over the past few decades, researchers have utilised various predictive models, including regression, Artificial Intelligence algorithms (ANN), and fuzzy neutral networks. Among these models, artificial neural networks are more widely used for load prediction. Additionally, support vector machines (SVM) are used in cases where small data sets are available. However, these models have certain limitations, such as over-fitting and the amount of data available. At this point, one model for an entire district level is preferred, and all predictive models strive for accuracy, total energy, and computational time [24].

The use of machine learning and deep learning algorithms has opened up new opportunities for load prediction research since 2016. Research groups have explored different approaches to load prediction using both forward and data-driven models. As noted earlier, heating depends on various factors, including building and physical conditions, for which complex energy simulations are commonly used. Forward models are well-developed and can be simulated using software programs such as EnergyPlus and TRNSYS [25].

The paper by [25], discusses the development of a Modelica® library for the modelling of thermal-energy transport in district heating systems within the framework of the AMBASSADOR project, which is funded by the European Commission under FP7. The library comprises detailed models of the distribution and consumption components commonly found in district heating systems. The paper presents the detailed models of the components, their validation against other software such as Apros® and IDA-ICE®, and the development of reduced mathematical models implemented in Simulink®.

The detailed models of the distribution pipe, distribution network, and hot water storage models were validated with models implemented in other software. The validation showed that the detailed models captured the related phenomena with a reasonable degree of accuracy. However, the comparison of the detailed models developed in Modelica®, Apros®, and IDA-ICE® showed that, while most of the models performed similarly, they did not reproduce the dynamics in the same way. The paper analyses the relevant causes of the differences and procedures to control them.

The reduced models were developed to meet the computational cost of the detailed models. The reduced models take the most relevant system dynamics into account but do not include all the possible terms. The most relevant dynamics in the pipes were identified assuming some simplifications, such as the decoupling of the mass flow rate calculation from the temperature calculation. The reduced models were validated against the corresponding detailed physical models. The results showed that the reduced models performed well in the simulations, but they gave a less accurate delay when implemented in the DSP due to the fixed sample time imposed by the platform.

In conclusion, the paper demonstrated that the detailed and reduced developed models are performing sufficiently well for the application at hand. However, further improvements are required when larger systems are simulated.

Consequently, data-driven models have gained momentum and have become a popular approach for predicting load. These models employ statistical methods to predict the energy demand based on historical data. In a recent study conducted in 2018, the operational thermal load forecasting was performed using machine learning algorithms such as linear regression, artificial neural networks, support vector machine, and extremely randomized trees regressor. The data used in this study was obtained from a district heating system in Sweden. Among the methods considered, artificial neural networks demonstrated superior performance under test conditions [26–28]. This thesis presents an expert system for forecasting thermal load, which is an important component for optimising district heating and cooling systems. The system combines four different data-driven methods, including linear regression, extremely randomized trees regression, feed-forward neural network, and support vector machine, to create a robust and accurate forecast. The system was tested using a dataset of 27 months from 10 residential buildings in Rottne, Sweden, and was found to outperform the individual methods. The expert system was able to track the best performing expert, which was the ANN with full feature set, and adding different experts added robustness to the forecaster and reduced susceptibility to changes in the DHS. The expert system is easily expandable to include new experts that meet the fit and predict interface provided by scikit-learn. Future research will focus on integrating this expert system into a DHS control solution for peak shaving to limit the thermal peak load and avoid the use of biodiesel burners.

The study titled "Forecasting District Heating Demand using Machine Learning Algorithms" by [29] investigates the potential of machine learning and data-driven models for short-term load forecasting in the solar community. The paper compares the performance of four different models, namely artificial neural networks (ANN), support vector machine (SVM), decision trees, and linear regression, to determine the most accurate model for load prediction. Notably, the focus of the study is on the comparative analysis of different models, rather than proposing any novel improvements to these models.

The results of the study indicate that both ANN and SVM are effective in capturing the general trend of demand, while the performance of the decision tree model is highly dependent on the size of the leaf. The authors' findings highlight the potential of machine learning algorithms for load prediction in district heating systems and provide insights into the suitability of different models for specific applications [29, 30].

Recent literature has highlighted the use of deep learning models for load forecasting of demand. One study by [31] investigates the application of deep learning techniques in combination with various machine learning techniques previously discussed [32, 33]. The authors applied polynomial linear regression, Ridge regressor, Lasso regressor, and Deep neutral net on two case studies, and the results showed that the deep learning technique performed well on both case studies.

In another study, [34] compared the performance of the feature fusion long short term memory (FFLSTM) model with other models such as back propagation, support vector regression, regression tree, random forest regression, gradient boosting regression, and extra trees regression. The result showed that FFLSTM outperformed all other algorithms. This study demonstrates the potential of deep learning models, such as FFLSTM, for improving load prediction accuracy.

The future prospects of the research on district heating demand forecasting based on machine learning and data-driven models show a possibility of developing a new model to improve accuracy. Additionally, there is an interest in exploring the research area of real-time information exchange infrastructure. However, the challenges and limitations of current models are also acknowledged, such as the risk of over-fitting and the limited availability of data. These challenges highlight the need for further research to develop more accurate and reliable load forecasting for district heating systems.

In recent years, two branches of Artificial Intelligence have dominated research in district heating demand forecasting, namely machine learning and deep learning. Both of these approaches rely on the size of available data for better performance. Various research groups have explored both forward and data-driven models for heat prediction. Heating systems depend on many factors such as the type of building and its physical condition, for which complex energy simulation programs have been developed.

The literature review highlights the significance of heat prediction and data-driven models. It can be concluded that the most popular methods used for heat prediction are Linear Regression, Decision Tree, and Artificial Neural Network. In this research, these three methods are evaluated on different datasets for heat prediction. By exploring these methods, we aim to contribute to the ongoing research and development in the field of district heating demand forecasting.

Table 2.1 Summary of essential Features

Classification	Attribute	Citations	Number of Research Paper
	Outdoor Temperature	[35–44] [45–54] [55–64] [65–72] [73–82] [83–91] [20, 92–100] [101–106, 14, 1, 16, 20]	74
	Ultraviolet	[104, 107, 108, 106, 109, 110]	6
	Wind Direction	[104, 106, 111, 112]	4
	Heat Index	[101–104, 108, 106, 109, 105, 113, 114] [14, 115, 1, 111]	14
	Dew-point	[105, 106, 110]	3
	Wind Chill	[104, 107, 114, 112]	4
	Wind Gust	[104, 112, 111]	3
	Pressure	[102, 16, 116]	3
Weather	Elevation	[106]	1
Patterns	Solar Radiation	[35, 39, 43–49, 53] [55, 57–59, 61–63] [64, 65, 54, 71, 35, 75–77, 79–81, 117, 83–85] [86, 88, 90, 93, 95, 102, 104, 107, 14, 106, 108, 113, 109, 105, 118]	45

Table 2.1 Summary of essential Features

Classification	Attribute	Citations	Number of Research Paper
	Humidity	[35, 39–45, 47, 49, 51, 53, 55, 57] [58–61, 63, 64, 68, 35, 73–76] [77, 79, 80, 84, 85, 91, 93, 95, 105, 106, 16]	36
Human Factors	Wind	[35, 39, 43–45, 47, 57–59, 61] [74, 76, 77, 85, 86, 88, 90, 92, 94, 95] [104, 107, 114, 16, 111, 112, 62, 63, 73]	23
	Precipitation	[73, 74, 76, 77, 105]	5
	Time stamp	[35–42, 44, 45][46–48, 52, 56, 61, 63, 67, 68, 60] [69, 70, 72, 35, 73–75, 78, 79, 85] [87, 91, 20, 20, 94–96, 102, 103, 113][16, 118]	40
	Building Activity	[44, 47, 50, 72, 96]	5
	Utilization Level	[35, 44, 46, 47, 57–59, 62, 76, 77, 81]	11

Table 2.1 Summary of essential Features

Classification	**Attribute**	**Citations**	**Number of Research Paper**
	Indoor Environment	[48, 53, 57–59, 62, 66, 54, 73, 75–77]	12
Building Specifications	Internal Architecture	[35, 117, 119, 82, 120, 121, 83, 122–124][97–99, 125–133][134, 84, 135–138]	25
	Building Skin	[79, 81, 117, 139, 119, 82, 120, 121, 83, 140] [122–131][132–134, 84, 135–138, 97–99, 141]	29
	Building Exposure	[81, 117, 121, 140, 122–124, 97–99][125–134] [84, 135–138, 141]	23
	Climate Control	[35, 53, 58, 54, 75, 139, 120]	7

Table 2.1 presents a summary of the essential features that hold significant value in the process of model training, thereby leading to the generation of highly accurate models for heat prediction. These features serve as crucial components of the machine learning algorithm, which enable the model to learn patterns and relationships between the input and output data. The incorporation of these features is paramount to ensuring the accuracy of the predictive model, as it facilitates the model's ability to generalize to new and

unseen data. Therefore, a comprehensive understanding of these features is vital in the development of successful predictive models for heat prediction.

2.3 Machine Learning Algorithms

Machine Learning (ML) is a subfield of Artificial Intelligence (AI) that enables computers to learn from data, without being explicitly programmed. In Machine Learning, algorithms automatically identify patterns and relationships in large datasets, and use this information to make predictions or decisions about new, unseen data.

In traditional programming, humans explicitly code the rules and logic necessary for a computer program to execute a specific task. In contrast, Machine Learning relies on algorithms that can learn from data and identify patterns on their own. These algorithms can be trained on vast amounts of data and can use statistical models to predict outcomes or identify patterns that may not be apparent to humans.

There are three main categories of machine learning algorithms: supervised learning, unsupervised learning, and reinforcement learning. Supervised learning involves training a model on labelled data, where the outcome is already known. Unsupervised learning involves training a model on unlabelled data, where the outcome is not known. Reinforcement learning involves training a model to take actions in an environment in order to maximize a reward signal.

Machine learning is increasingly used in a wide range of applications, including image and speech recognition, natural language processing, recommender systems, fraud detection, and autonomous vehicles. As more data is generated and collected, the potential for Machine Learning to make breakthroughs in various fields is constantly growing.

2.3.1 Supervised Learning Algorithms

2.3.1.1 Nearest Neighbour Classification

K-Nearest Neighbors (KNN) is a type of nearest neighbor classifier that is defined by its ability to classify unlabelled examples by assigning them the class of similar labeled examples. Despite the simplicity of this idea, KNN and other nearest neighbor methods are indeed extremely powerful. Here are a few reasons why:

1. Non-parametric: KNN is a non-parametric method, which means that it does not make any assumptions about the underlying distribution of the data. This makes KNN more versatile than many other machine learning algorithms, which may require strong assumptions about the data in order to be effective.

2. Effective with non-linear data: KNN is particularly effective with non-linear data, since it does not require the data to be linearly separable. This means that KNN can handle complex decision boundaries, and can work well with a wide variety of datasets.

3. Robust to noisy data: Since KNN uses the "nearest neighbors" of each data point to make predictions, it can be more robust to noisy data than other algorithms. Outliers and other anomalies are less likely to have a significant impact on the final predictions, since they are only a small part of the overall data.

4. Easy to implement: KNN is a simple and intuitive algorithm that can be implemented without much difficulty. This makes it a popular choice for researchers and practitioners who are just starting out with machine learning.

KNN is not perfect, and there are situations in which other algorithms may be more appropriate. For example, KNN can be computationally expensive with very large datasets, and may not be as interpretable as other methods. Nevertheless, KNN and other nearest neighbor methods remain a popular and effective choice for many machine learning tasks.

Strengths:

- Simple and easy to implement: K-Nearest Neighbors (KNN) is a straightforward and effective machine learning algorithm that can be implemented without much difficulty.
- Non-parametric: KNN does not make any assumptions about the underlying data distribution, which makes it more versatile than many other machine learning algorithms.
- Fast training phase: KNN has a fast training phase since it only involves storing the training data.

Weaknesses:

- Lack of interpretability: KNN does not produce a model that can help us understand how the features are related to the class. This lack of interpretability can be a disadvantage if we need to explain how the algorithm makes its predictions.
- Sensitivity to the choice of k: KNN requires the selection of an appropriate k value, which can impact the quality of the results.
- Slow classification phase: The classification phase of KNN can be slow, especially when dealing with large datasets.

- Preprocessing requirements: Nominal features and missing data require additional processing before they can be used with KNN, which can be a time-consuming task.

When compared to other machine learning algorithms, KNN stands out due to its non-parametric nature and ease of implementation. Decision trees are more interpretable, but may not perform well with high-dimensional data, while support vector machines can handle high-dimensional data more efficiently but require more preprocessing and feature engineering. Random forests are generally more accurate, but can be computationally expensive, especially with large datasets.

KNN is most appropriate when the dataset is small and low-dimensional, the relationships between features and class labels are non-linear, and the dataset contains noisy data or missing values. When the algorithm does not need to be highly interpretable and the dataset is balanced, KNN is a good choice. In contrast, the SVM or decision tree algorithm may be more appropriate for high-dimensional datasets, while random forests or other ensemble methods may be more efficient for very large datasets.

2.3.1.2 Naive Bayes Classification

Bayesian methods have their roots in the foundational work of Thomas Bayes, an 18th-century mathematician who formulated principles to describe event probabilities and how they should be updated with new information. Today, Bayesian methods are widely used and offer a number of strengths, including simplicity, speed, and high effectiveness, particularly with noisy and missing data. Bayesian methods are also capable of handling relatively few or very large numbers of training examples. However, these methods rely on the assumption of equally important and independent features, which may not be accurate in many situations. Bayesian methods are also not ideal for datasets with numerous numeric features and may produce estimated probabilities that are less reliable than predicted classes. Nevertheless, Bayesian methods provide an easy way to obtain estimated probabilities for predictions, which can be useful in various contexts.

2.3.1.3 Decision Trees Classification

Decision tree learners use a tree structure to represent the relationships between the features of a dataset and the target variable, which can be a categorical or continuous variable depending on the type of problem. The tree is constructed by recursively partitioning the data based on the values of the features, so that each split creates more homogeneous subsets of the data with respect to the target variable. The resulting tree can be used to make predictions for new instances by following the branches of the tree from the root to a leaf node, where a prediction is made based on the majority class or mean value of the training instances that reach that node. Decision tree learners are popular in machine learning due to their simplicity, interpretability, and ability to handle both categorical and numerical features.

The Random Forest algorithm is a powerful ensemble method that uses multiple decision trees to improve model accuracy and reduce the risk of overfitting. It is a versatile algorithm that can handle a mixture of data types, including numeric and categorical variables, and is commonly used for both classification and regression problems. Additionally, it is less sensitive to outliers and requires less data preprocessing and normalization than some other models. However, it can be relatively slow to create models and make predictions, especially on large datasets, and can be prone to overfitting if the number of trees is too high. It is also somewhat less interpretable than a single decision tree.

The C5.0 algorithm is a powerful and versatile classifier that can handle different types of features and missing data. It also has an automatic learning process and can exclude unimportant features. However, it has some weaknesses, such as being prone to overfitting or underfitting, and having trouble modelling some relationships due to its reliance on axis-parallel splits. Additionally, large decision trees can be difficult to interpret and the decisions they make may seem counter intuitive.

2.3.1.4 Linear Regression Numeric prediction

Regression analysis is a statistical technique used to model the relationship between a dependent variable and one or more independent variables. The goal is to find a mathematical equation that can predict the value of the dependent variable based on the values of the independent variables. Linear regression is a common type of regression analysis, but there are also other types, such as logistic regression, polynomial regression, and ridge regression, that can be used depending on the nature of the data and the research question.

Multi Linear Regression is a commonly used technique for modelling numeric data and can be adapted to a wide range of modelling tasks. However, it does make strong assumptions about the data, and the form of the model must be specified by the user in advance. Additionally, it only works with numeric features, so categorical data requires extra processing.

Regression trees are a machine learning method that combines the advantages of decision trees with the ability to model numeric data. This approach does not require the user to specify the model in advance, and it uses automatic feature selection, which enables it to work with a vast number of features. Moreover, regression trees may fit some types of data much better than linear regression, but they require a large amount of training data. One of the main challenges of regression trees is that it is difficult to determine the overall net effect of individual features on the outcome. Additionally, large trees can become more challenging to interpret than a regression model. However, regression trees remain a useful approach for many prediction problems.

2.3.1.5 Neural Networks

An Artificial Neural Network (ANN) is a computational model that seeks to mimic the response of biological brains to sensory inputs. It accomplishes this by mapping a set of input signals to an output signal using a model derived from our understanding of neural networks. ANNs have gained popularity due to their flexibility and applicability in a wide range of problems, including classification and numeric prediction.

Despite its many strengths, ANNs have several weaknesses. One significant challenge is the computational intensity required for training, particularly if the network topology is complex. Overfitting is also a common problem, whereby the network learns to memorize the training data, rather than generalizing to new examples. Additionally, the resulting model is often a complex black box, making it difficult, if not impossible, to interpret the underlying decision-making process.

2.3.1.6 Support Vector Machines

The Support Vector Machine (SVM) is a type of model that can be conceptualized as a surface that creates a boundary between data points plotted in multidimensional space, representing examples and their feature values. SVM can be applied to classification or numeric prediction problems, and it is known for being less susceptible to noisy data and overfitting. However, finding the optimal model requires testing various combinations of kernels and model parameters, and training can be slow, especially when the input dataset has a high number of features or examples. Despite being easier to use than neural networks, SVM can produce a complex black box model that is challenging or even impossible to interpret.

2.3.2 Unsupervised Learning Algorithms

2.3.2.1 Association Rules Pattern detection

Association rule learning is a popular approach for efficiently searching large databases for rules, which has been extensively studied. To reduce the number of itemsets to search, many heuristic algorithms have been proposed. Among them, the Apriori algorithm, introduced by Rakesh Agrawal and Ramakrishnan Srikant in 1994, has become widely used in the field.

The Apriori algorithm has the capability to handle large transactional datasets, resulting in rules that are easy to comprehend. It is especially useful for data mining and discovering unexpected knowledge in databases. However, the algorithm may not be as helpful when dealing with small datasets, and care must be taken to separate genuine insights from spurious conclusions that may arise from random patterns.

2.3.2.2 K-means clustering

Clustering is a machine learning task that groups similar items into clusters, without the use of predefined categories. The k-means algorithm is a widely used method for clustering.

Advantages:

- Uses simple principles that are easy to understand.
- Can be adapted to overcome its limitations.
- Performs well in real-world scenarios.

Disadvantages:

- Less sophisticated than other modern clustering methods
- Results may not always be optimal due to the algorithm's reliance on chance.
- Requires prior knowledge of the number of clusters in the data.
- Not suitable for non-spherical clusters or clusters with widely varying densities.

2.4 Control Strategy

In this section, the discussion revolves around the optimisation control strategy in district heating systems [34, 33, 58, 40, 59, 87, 101, 118, 3]. Several optimisation control strategies have been proposed in the literature, such as PID-based control strategy [142], hybrid fuzzy logic [143], and Mixed Integer Linear Programming [144, 145]. Among these methods, Mixed Integer Linear Programming has been the most commonly used method for optimising district heating systems.

In recent years, new control strategies have been developed to improve the efficiency of district heating systems. For instance, in 2018, [28] proposed a novel model-predictive control strategy for district heating systems, which showed promising results. The concept of model-predictive control involves predicting the future behaviour of the system and using this information to make real-time control decisions.

As the research progressed, more in-depth analysis and evaluation of these control strategies will be explored to determine their effectiveness in optimising the performance of district heating systems.

2.5 Cloud Service

In this section, a comprehensive overview of various cloud service providers is presented, with a specific focus on comparing these providers with Siemens MindSphere. Cloud services are increasingly becoming a popular choice for businesses due to the pay-as-you-use model, which allows for greater flexibility and cost-effectiveness. While the primary role of the cloud is to provide infrastructure-as-a-service, in this section, the focus is on the platform-as-a-service provided by these cloud service providers. With industry 4.0 in mind, further literature has been studied and compared to evaluate the most suitable cloud service provider [146, 147].

One of the cloud service providers reviewed is Azure IoT Hub, a plug-and-play cloud service provided by Microsoft Azure that allows for bidirectional communication with billions of IoT devices, stream analytics, and machine learning capabilities. Recently, Microsoft has added digital twin and IoT features to the Azure IoT Hub, making it an even more attractive option for businesses [148, 17]. Another popular cloud service provider is Google, whose cloud platform is increasingly gaining popularity. The relevant product for this proposal is the Google Cloud IoT, which offers machine learning, big query, and data visualization capabilities [19]. IBM Bluemix is another notable cloud service provider that offers IBM Watson IoT with AI capabilities [18]. Finally, Amazon is another major player in the cloud services market and holds the largest share globally. AWS IoT Core is their cloud platform that provides SDK for all devices to establish an easy connection with AWS services. Amazon has also partnered with major hardware development companies such as Intel, Qualcomm, and Broadcom to leverage their cloud and hardware capabilities [148, 147].

Apart from the above-mentioned cloud service providers, various other platform-as-a-service (PAAS) options are also available, such as Dell IoT, Hitachi Lumada, Oracle IoT cloud platform, Bosch IoT Suite, SAP Leonardo, Cisco IoT, and Salesforce IoT. These PAAS options share several common features, such as scalability, real-time data processing, storage, analysis, and applications in different real-life scenarios. Considering the application of this proposal, a PAAS platform based on industrial 4.0 is the most suitable option. Two major companies in the industrial IoT platform market are GE Predix and Siemens MindSphere [147, 149].

GE Predix is the first platform designed for industry data capturing and analysis, with several advantages, such as strong partnerships, talent, experience, knowledge of the industry, and a flexible and robust platform that allows for real-time visualization. However, the platform has several limitations, such as an unreliable and buggy admin panel, limited analytics capabilities, and a lack of developers with domain expertise. On the other hand, Siemens MindSphere is an open IoT operating system designed for industries to create their applications and analytics. Like GE, Siemens has strong partnerships with various industries. MindSphere uses an OPC system that ensures interoperability, secure connections, and third-party connections. Moreover, MindSphere offers different types of storage options such as on-site, public, or private cloud. Additionally, MindSphere is an open standard, plug and play, open interface, and cloud infrastructure [146]. Siemens has dedicated support for MindSphere, an API open to third parties to develop applications, multiple data sources can be analysed, and tested in manufacturing facilities. MindSphere is hosted on AWS and SAP HANA, leveraging their infrastructure capabilities. After reviewing the advantages of MindSphere, it is concluded that it is an ideal solution for this project.

The Covid-19 pandemic has had a significant impact on various aspects of life, including the business world. With many companies implementing work from home policies and restrictions on physical interactions, there has been a shift towards online platforms and virtual solutions. This has led to an increase in demand for cloud-based services and digital solutions, which can be accessed remotely.

Given the unprecedented circumstances caused by the pandemic, it became necessary to consider alternative options for the implementation of the proposed solution. While Siemens MindSphere was identified as an ideal solution, the uncertainty and disruption caused by the pandemic necessitated a reassessment of available options. After a thorough review of various alternative platforms, Salesforce Heroku was identified as a suitable alternative with comparable capabilities to Siemens MindSphere [146].

Salesforce Heroku is a cloud-based platform that provides developers with a powerful set of tools to build, deploy, and manage applications. It is a Platform-as-a-Service (PaaS) offering that supports multiple programming languages and provides easy integration with various data sources, making it an ideal solution for building scalable web applications. Additionally, Salesforce Heroku provides robust security features, data management capabilities, and reliable performance, making it a suitable choice for enterprise-level solutions. Figure 2.2 provides an overview of the cloud services that were investigated for this research.

Despite the shift towards Salesforce Heroku, it is worth noting that Siemens MindSphere remains a highly effective solution for Industry 4.0 applications. Its open IoT operating system for industry design and analytics, secure connection, and interoperability features make it a leading platform for industrial IoT. However, given the pandemic's impact on the project, Salesforce Heroku was deemed the most appropriate alternative to ensure the project's success.

Fig. 2.2 Different Cloud Services

2.6 Overview of Software

In this section, we present a review of various software solutions based on the advantages highlighted on their respective websites. While it is challenging to review all available software solutions, we focus on significant ones for a better understanding of the proposed system. This review is essential for making informed decisions on the best software solution to use for the project.

With the rise of Industry 4.0, software solutions have become integral to the efficient operation of various industries. These solutions provide functionalities such as data analysis, real-time visualization, and communication with IoT devices, among others. The adoption of software solutions has increased dramatically over the years, and several vendors now provide software solutions for different industries.

One of the most significant advantages of software solutions is the ability to automate various processes, making operations faster and more efficient.

In addition, software solutions are scalable, making it possible to add more features as the needs of the industry change. Another advantage is that software solutions can be customized to suit specific needs.

2.6.1 Bentley

In the context of planning, design, construction, and operations, Bentley Systems is a comprehensive software solution provider. The District Energy Network Planning and Design solution, which Bentley offers, is particularly noteworthy for its ability to optimise network performance, reduce operational costs, and improve project and service delivery. Additionally, the GIS-based tool offered by this software helps to design paths better, which can be particularly useful for infrastructure design and management. By using Bentley's District Energy Network Planning and Design solution, organizations can benefit from advanced optimisation algorithms and simulations that can improve the efficiency of their energy networks, leading to cost savings and more sustainable operations [150].

2.6.2 AVEVA

AVEVA is a software solution for district energy management that provides a range of features aimed at optimising the performance of district energy networks. Specifically, AVEVA's district energy management software is designed to minimise heat losses, maximize storage capacity, and predict the behaviour of the network. These capabilities enable the software to improve the efficiency of energy distribution and reduce energy waste, thereby reducing operational costs and improving the overall sustainability of the network [151].

2.6.3 Globema

The Globema software solution for district heating management is a Geographic Information System with powerful features for data access, reporting, statistics, investment planning, and resource management. The system is designed to enable fast and convenient access to data, allowing for efficient management and planning of district heating networks. Globema's functional module is specifically designed for district heating and provides additional features to facilitate the management of district heating networks [152].

2.6.4 Passivsystems

The Passivsystems platform is a cloud-based IoT home energy services platform designed to collect data from homes, provide data storage, and manage in-home actuators. The District Heating service suite offers various services such as Automated Meter Reading, Network Load Management, Pre-payment, Smart Valve Control, and PassivLiving HEAT. With these features, the platform can optimise energy usage, reduce energy waste, and ultimately reduce energy costs for customers. Passivsystems also provides an API that enables easy integration with third-party systems, as well as a range of analytics and reporting tools to enable better monitoring and management of energy usage [153].

2.6.5 Essential Control

The software solution called Essential Control provides remote monitoring and control capabilities for both large and small sites, enabling reduction of on-site engineer time and improvement of energy efficiency, ultimately resulting in reduced running costs of heating schemes. Additionally, this software helps to cut carbon emissions and improve local air quality [154].

It's important to narrow down the scope of the proposed framework in order to focus on the key objectives and outcomes. This can help ensure that the research is feasible and realistic within the given resources and time frame. By focusing on the development of a data-driven model for load prediction and control strategy, the proposed framework can address the current gap in real-time information exchange infrastructure and heat prediction in the district heating system. As the literature review suggests, other features like GIS-based tools, smart valve control, and network load management can be included in the framework as well, but their implementation can be considered as future work.

2.7 Hypernetwork theory

Hypernetworks Theory (HT) is a conceptual framework used to describe complex systems, which is used as the basis for the software solution in question. The HT is a multilevel and multidimensional structure that represents the physical, logical, or mathematical structure of a system or combination of systems [155–162]. The HT consists of interrelated Hypersimplices, which are grouped into levels by nodes or vertices [163, 164]. Each vertex represents a part of the system, and by grouping these parts in logical ways, they form a whole. Each whole can then be a part of some other whole, forming additional levels. The collection of parts is known as a simplex.

What distinguishes the HT is its ability to represent an overlap of systems, or views of systems. This overlap enables the creation of models that provide individual and combined results. The logical way in which the parts are grouped is through R, the n-ary relation (multidimensionality). The use of HT-based orchestration coupled with a microservice architecture allows for the creation of a flexible and scalable software solution that can handle complex systems and produce reliable results.

Fig. 2.3 The Fundamental Construction of Multilevel Systems

Fig. 2.4 Hierarchical aggregation in multilevel systems

The Hypernetworks Theory (HT) provides a powerful framework for modelling complex systems, both physical and mathematical, allowing for the creation of multilevel, multidimensional structures that capture the interrelationships between different components of a system. Johnson introduced the concept of traffic, which refers to the values that flow across the system, and introduced an equivalent to the n-ary relation R, called , which enables the creation of mathematical models that can be integrated into the same framework [160].

The Hn model is used by SEMS to choreograph the data collection, forecasting, and optimisation of district heating networks. The bottom level vertices of the Hn represent the physical components of the network, such as sensors and meters, as well as external data sources, such as weather and CO2 forecasts. These bottom level vertices are then used to feed the next level up, where more complex models are created based on the traffic flowing through the system. Figures 2.3 and 2.4 present a comprehensive depiction of the fundamental principles underlying multilevel systems and hierarchical aggregation. These visuals aim to enhance the understanding of the Hypernetwork theory.

One of the key advantages of the Hn model is its ability to capture the overlap of different systems, enabling models to be created that provide individual and combined results. Each vertex in the Hn represents a part, and by grouping these parts in logical ways, they form a whole. Each whole can then be a part of some other whole, forming levels of interrelated Hypersimplices. This hierarchical structure allows for a comprehensive representation of the district heating network, from the physical components to the mathematical models that optimise its performance[155, 160, 164].

It is worth noting that the vertices in the Hn model do not have knowledge of the whole(s) of which they are a part, and are solely responsible for generating traffic that flows through the system. However, the whole needs to know the traffic created by its parts, which is specified by the Hn model. This way, the Hn model enables the optimisation of the entire district heating network by capturing the complex interrelationships between its components, both physical and mathematical.

2.8 Overview of Internet of Things

The idea of perfect IoT framework, should be accessible from anywhere around the world for a client. As the name suggest the framework uses the internet to connect devices, available anytime and anywhere by GSM (Global Systems for Mobile) [165]. Many standards have already been developed in the field of IoT and are available to use directly from the cloud providers like Microsoft Azure, Amazon, GE, Siemens, IBM, and several others. One of the advantages of cloud service is that it works on the pay-as-you-use model. The primary role of the cloud is to provide infrastructure-as-a-service. However, all cloud systems are based on platform-as-a-service provided by them. Azure IoT Hub is a service provided by Microsoft Azure. It is a plug and play cloud without writing code, bidirectional communication with billions of IoT devices, stream analytics and Machine Learning [17]. Digital Twin and IoT are two features added by Microsoft recently [148]. Another cloud service provider is Google.

The popularity of the Google cloud platform is increasing. The product which is relevant to the proposal is Google cloud IoT. It has Machine learning,

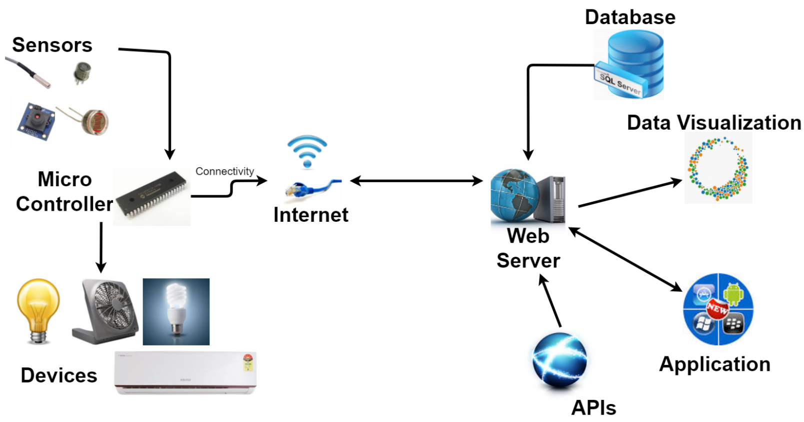


Fig. 2.5 Overview of IoT System [2]

big query, and data visualisation capabilities [19]. IBM Bluemix has IBM Watson IoT with AI (Artificial Intelligence) capabilities [18]. Amazon is another company which holds the largest share globally in cloud services. AWS (Amazon Web Services) IoT Core provides SDK (Software Development Kit) for all devices to establish easy connection with AWS services. Amazon has partnered with major hardware developing companies like Intel, Qualcomm, and Broadcom to leverage the cloud and hardware [148, 149]. Apart from this, using AWS infrastructure, several AWS as well as non-AWS IoT platforms are hosted. Other Platform-As-A-Service (PAAS) providers available are Dell IoT, Hitachi Lumada, Oracle IoT cloud platform, Bosch IoT Suite, SAP (Systems Applications and Products), Leonardo, Cisco IoT, Siemens MindSphere, and Salesforce IoT. All PAAS has few common features that are scalable, real-time, storage, analysis, and different real-life applications. As all the cloud providers provide similar service any cloud can be used. The proposed framework becomes particularly important and special services such as heat demand prediction for District Heating can be developed to leverage the advantages of cloud services [5, 166–168].

It's interesting to see how this open-source framework for IoT developed by Kevin and Supriya has inspired the proposed framework for IoT-based District

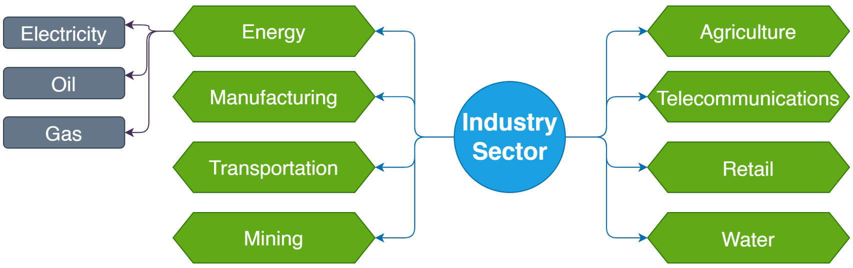


Fig. 2.6 IIoT in Different Industry [3]

Heating. The use of embedded electronics like sensors, micro-controllers, and micro-processors, along with application software like web services and app development, are essential components of the proposed framework. The incorporation of these elements allows for the collection and exchange of real-time information, which is then utilised to predict heat demand and optimise the distribution of heat in the district heating system. The proposed framework not only aims to increase the energy efficiency of the system but also strives to make it more scalable, reliable, easily accessible, and cost-effective. Overall, the proposed framework appears to be well-aligned with the broader goals of leveraging IoT to create more sustainable and efficient energy systems (Figure 2.5). Figure 2.6 provides a comprehensive overview of the identified research gap in the first review of District heating, with a specific focus on the integration of IIoT (Industrial Internet of Things).

2.9 Thermal Characteristics of Building Materials: An Exploration of Physical Properties

The performance of a building envelope in relation to heat transmission is heavily influenced by the thermal-physical properties of the materials used. While this paragraph aims to provide insight into the most significant properties, it is not intended as a comprehensive handbook on the topic. Instead, only those parameters relevant to the current research will be addressed. Prior to that, an overview of the heat transmission mechanism will also be presented.

2.9.1 Heat transmission

According to [169], heat refers to the energy that is exchanged between a system and its surrounding environment, driven by a difference in temperature between the two. Heat always flows from the system at a higher temperature to one at a lower temperature, with the total energy remaining constant. Heat transfer can be categorized into three primary mechanisms: conduction, convection, and radiation [170].

Within a building envelope, heat is transferred via various structural components such as walls, roofs, floors, and windows, utilising the aforementioned mechanisms. In the following section, an overview of these processes will be presented.

2.9.1.1 Conduction

As defined by [171], conduction refers to the transfer of heat energy through a material without causing any displacement in the basic positions of its constituent molecules. According to [172], conductive heat transfer is primarily driven by the combined transport of energy via free electrons and molecular vibration[170].

Equation 2.3 expresses Fourier's Law, which measures the heat flow through a surface per unit area.

$$
\begin{aligned}
q &= -\lambda \nabla T && (2.1)\\
&= -\lambda \left(i\frac{\partial T}{\partial x} + j\frac{\partial T}{\partial y} + k\frac{\partial T}{\partial z} \right) && (2.2)\\
&= -\lambda \frac{\partial T}{\partial x} && (2.3)
\end{aligned}
$$

$$
\begin{aligned}
Q &= Aq && (2.4)\\
&= -A\lambda \frac{\partial T}{\partial x} && (2.5)
\end{aligned}
$$

Where:

- Q = heat flow (W)
- A = area (m^2)
- q = heat flux (W/m^2)
- λ = thermal conductivity (W/mK)
- $\frac{\partial T}{\partial x}$ = temperature gradient.

Thermal conductivity is a crucial parameter that characterizes the rate at which heat transfers through a material under specific conditions. It is defined as the heat flow in Watts that crosses a thickness of 1 meter of material with a temperature difference (δT) equal to 1 Kelvin and a surface area of 1 square meter. Different materials have different conductivity values, with materials possessing high λ values typically referred to as conductors, such as metals, while those with low to moderate λ values are known as insulators. The reciprocal of conductivity is known as resistivity, which is measured in meters x Kelvin per Watt (mK/W) [170].

2.9.1.2 Convection

Convection is a type of heat transfer that occurs between a surface and a fluid flowing over it, as described by [173]. Convection is typically the primary mode of heat transfer in liquids and gases, whereas it is absent in solids. The transfer of heat resulting from convection can be quantified using Newton's law of cooling:

$$Q = hA\,(T_s - T_\infty) \tag{2.6}$$

Where:

- Q = heat flow (W)
- h = convection coefficient (W/m^2 K)
- A = area (m^2)
- T_s = surface temperature
- T_∞ = fluid temperature at distance from the surface

The convection coefficient, which characterizes the rate of heat transfer by convection, is influenced by various factors such as fluid properties (e.g., density, viscosity, and specific heat), fluid velocity, and surface geometry. These properties vary with temperature and location, resulting in varying convection coefficients throughout the system. Convection can be categorized into two types: natural convection and forced convection. Natural convection occurs when the heated fluid expands and becomes less dense, causing it to rise and displace colder fluid, leading to the formation of convection currents. Forced convection, on the other hand, involves the use of external means such as fans or pumps to move the fluid [171, 170].

2.9.1.3 Radiation

Radiative heat transfer occurs between surfaces that are in direct view of each other, without requiring any medium between them. The rate of heat transfer between the surfaces is determined by the Stefan-Boltzmann law, which states that the net rate of heat transfer (in Watts per square meter) between two surfaces is proportional to the fourth power of the absolute temperature difference between them (Eq. 2.7:

$$Q = A\sigma\varepsilon\left(T_1^4 - T_2^4\right) \tag{2.7}$$

Where:

- Q = heat flow (W)
- A = area (m^2)
- σ = Stefan-Boltzmann Constant $\left(5.6703 \cdot 10^{-8}\left(\mathrm{~W/m^2~K^4}\right)\right)$
- ε = emissivity
- T_1 and T_2 = surfaces temperature (in Kelvin)

Thermal radiation is a type of electromagnetic radiation emitted by matter at all temperatures, regardless of its temperature [173].

2.9.2 Thermal transmittance (U-value) and thermal resistance (R-value)

[174] identified heat conduction through the building envelope and convective losses due to ventilation as the two primary mechanisms through which buildings lose heat. To quantify the heat transferred through a specific section of a building, such as a wall, two parameters are commonly used: the thermal

transmittance (U-value) and the thermal resistance (R-value). The U-value, also known as the overall heat transfer coefficient, represents the rate of heat flow through a surface with an area of 1 m^2 and with a temperature difference of 1 K between the two sides of the surface. In contrast, thermal resistance, which is used to calculate the U-value, is the measure of the opposition to heat flow offered by a particular component [171, 170].

The R-value is calculated as:

$$R = \frac{s}{\lambda} \tag{2.8}$$

Where:

- R = thermal resistance ($m^2 \cdot K/W$)
- s = thickness of the material (m)
- λ = thermal conductivity (W/mK)

Equation 2.8 indicates that the thermal resistance (R-value) is directly proportional to the thickness of the component and is a function of the material's thermal conductivity. The U-value (or overall heat transfer coefficient), which is used to determine the rate of heat flow through a surface, is calculated using the reciprocal of the sum of the resistance of all the layers within the component and the internal and external surface resistance. Equation 2-8 represents the general equation used to calculate the U-value [170].

$$U = \frac{1}{R_{in} + \frac{S_1}{\lambda_1} + \frac{S_2}{\lambda_2} + \cdots \frac{S_n}{\lambda_n} + R_{ex}} \tag{2.9}$$

Where:

- U = thermal transmittance (W/m^2 K)
- s = thickness of the material of each of the n layers (m)

- λ = thermal conductivity of the material of each of the n layers (W/mK)
- R_{in} = internal surface resistance ($\text{m}^2 \cdot \text{K/W}$)
- R_{ex} = external surface resistance ($\text{m}^2 \cdot \text{K/W}$)

The surface resistance R_s (R_{in} and R_{ex} in the equation) is defined as

$$R_s = \frac{1}{h_c + h_r} \tag{2.10}$$

Where:

- R_s = surface resistance ($\text{m}^2 \cdot \text{K/W}$)
- h_c = convection coefficient (W/m^2 K)
- h_r = radiation coefficient (W/m^2 K)

The coefficients hc and hr are affected by various factors such as the direction of heat flow, surface properties like emissivity, temperature, and climate conditions. Furthermore, these coefficients are distinct for the internal and external surfaces, respectively.

2.10 Overview of Space heating

The UK government has set a target of achieving net-zero carbon emissions by 2050, and district heating systems are seen as an important part of achieving this goal. District heating systems have the potential to significantly reduce carbon emissions by improving energy efficiency and increasing the use of low-carbon and renewable energy sources (Figure 2.7). The majority of district heating networks in the UK serve the residential sector, and optimising these systems can help to reduce carbon emissions from heating in homes and buildings [175] (Figure 2.8.

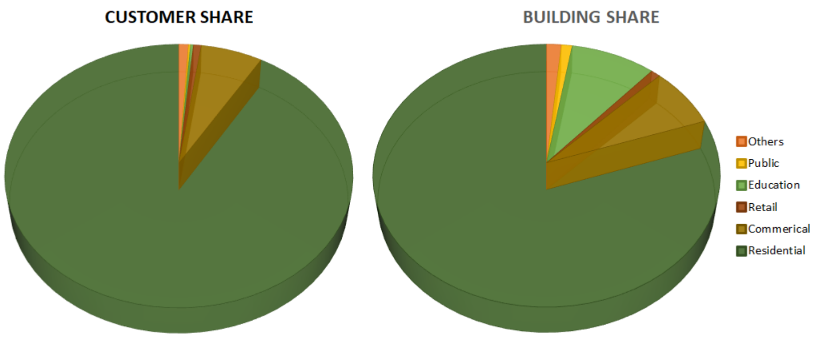


Fig. 2.7 Proportion of heat based on types of building (left) and types of customer (right) [4].

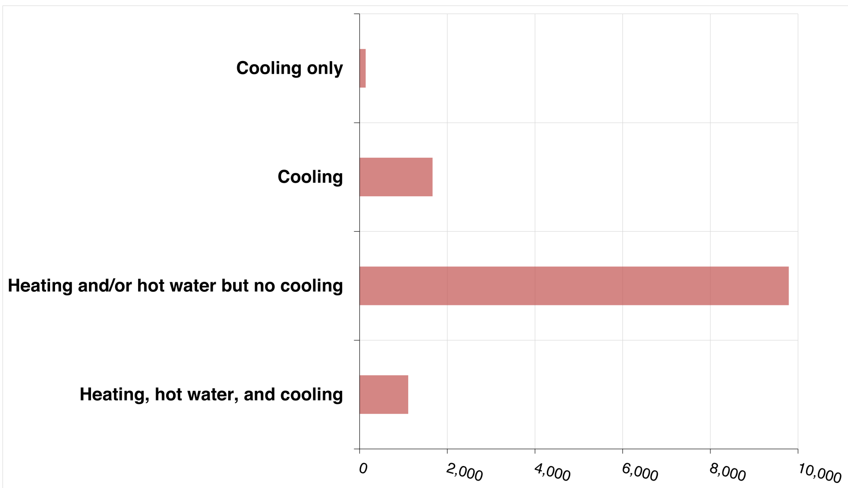


Fig. 2.8 Usage of heat network [4]

The need to reduce energy consumption in the building sector has been widely recognized due to its significant contribution to carbon emissions. In fact, according to the European Union, the building sector consumes around 40% of the total energy and is responsible for about 36% of the CO_2 emissions in the EU. Similarly, in the UK, the building sector accounts for 26% of carbon emissions. Space heating and hot water are the major contributors to household energy consumption, accounting for around 80% of the overall domestic energy demand [4]. Therefore, strategies to reduce heating energy consumption in buildings can have a significant impact on energy savings and carbon emissions.

One way to reduce heating energy consumption is to enhance building wall insulation [176]. The U-value of a building's wall is a measure of how much heat can pass through it. A higher U-value indicates that more heat can escape through the wall, leading to increased energy consumption. Building regulations in many countries set a maximum U-value for newly designed buildings, and buildings with U-values higher than this threshold may require retrofitting to improve their insulation [177, 178].

However, the energy savings resulting from enhanced insulation can be affected by several factors, such as local weather, the types of insulation used, and the materials used for insulation. Therefore, a prediction tool that can rapidly estimate the future heat loss through a building's walls can help stakeholders assess future energy savings and determine the target U-value for upgrading retrofit.

Several research works have explored the effect of insulation on building thermal performance, including the use of thermography for thermal performance evaluation and measuring the U-value of walls in buildings. Such research is essential to develop effective strategies for reducing energy consumption in buildings and achieving carbon emission reduction targets [179–181].

In addition to the impact of insulation materials and thickness on energy consumption, other factors such as local weather conditions, building design, and occupant behaviour can also affect the energy performance of buildings. For example, a study on the influence of window size and orientation on energy consumption for space heating and cooling in a residential building in Turkey found that buildings with larger south-facing windows had lower heating energy demand, while buildings with larger west-facing windows had higher cooling energy demand [181]. Another study on the impact of occupant behaviour on energy consumption in a university building in Italy showed that occupant behaviour can lead to significant variations in energy consumption, and that interventions aimed at modifying occupant behaviour can lead to substantial energy savings [182].

Overall, the literature suggests that the thermal performance of building walls can be significantly improved through the use of insulation, with external wall insulation generally outperforming internal wall insulation in terms of energy consumption for heating and cooling. However, the energy performance of buildings is affected by a wide range of factors, including insulation materials and thickness, local weather conditions, building design, and occupant behaviour, and a comprehensive approach to energy efficiency in buildings requires consideration of all of these factors [177, 178, 182, 183].

2.11 Proposed System

The implementation of district heating has gained considerable momentum in recent times, with governments in the EU/UK taking active steps and investing in the system. District heating is a well-established technology, with more than a century of history behind it. However, its recent resurgence is due to the growing emphasis on energy efficiency and the need to reduce carbon emissions. District heating systems are known for their ability to reduce energy consumption and carbon emissions by utilising waste heat and renewable energy sources. The literature review suggests that district heating is a reliable and efficient system for meeting the energy needs of a community or city. The review also highlights the need for accurate monitoring and control of district heating systems to optimise their performance and minimise energy waste. The use of advanced technologies such as the Internet of Things (IoT) and Artificial Intelligence (AI) is expected to enhance the performance of district heating systems, making them more efficient and environmentally friendly.

The proposed framework aims to integrate IoT technology with the evolving District Heating (DH) system. However, the current literature on the implementation of IoT in DH is limited. In a related tangent to the proposed framework, the concept of Smart District Heating has been discussed in the literature. For instance, [184]. highlighted the benefits of smart district heating in terms of energy savings and increased efficiency as compared to traditional DH systems [100, 184, 185]. Similarly, other studies have emphasized the significance of Information and Communication Technology (ICT) in DH. Despite the growing interest in smart DH, there is still a lack of research on the application of IoT technology in DH. The proposed framework seeks to address this gap by introducing an IoT-based solution for DH optimisation.

Figure 2.9 provides an illustrative representation of the proposed system, which begins with the transportation of heat from the heat generation sources to the buildings. These sources can range from traditional boilers and fossil fuels to renewable energy such as solar PV and heat pumps. In today's digitally connected world, buildings are equipped with a range of sensors, smart meters and internet capabilities that enable communication between various components of the system. The sensor and smart meter data is collected and stored in a data storage system, where it can be used for a range of purposes

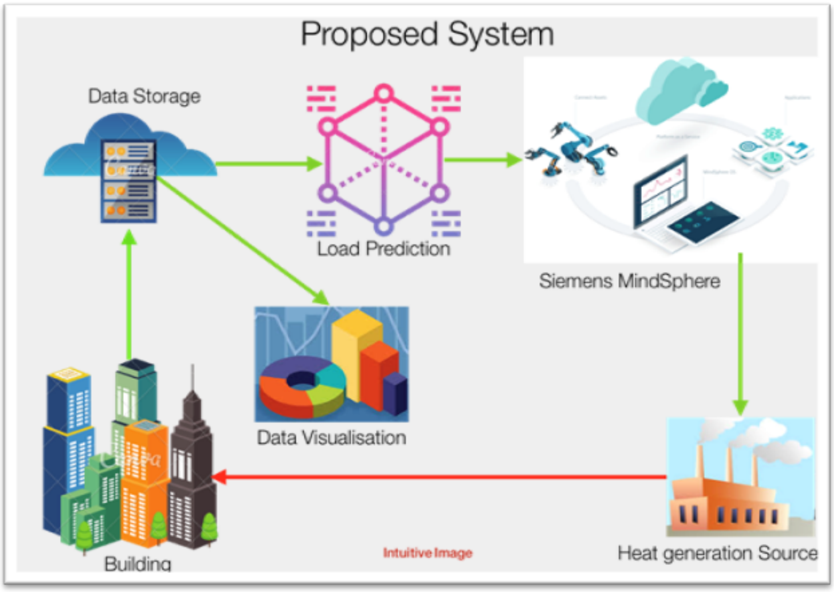


Fig. 2.9 Intuitive Image of Proposed System

including data visualisation, exploratory analysis and data-driven modelling for forecasting. The forecast data obtained from the models will then be used for planning, scheduling and controlling the heat generation process. The control technique will utilise the forecast data as an input for optimising and regulating the heat generation process.

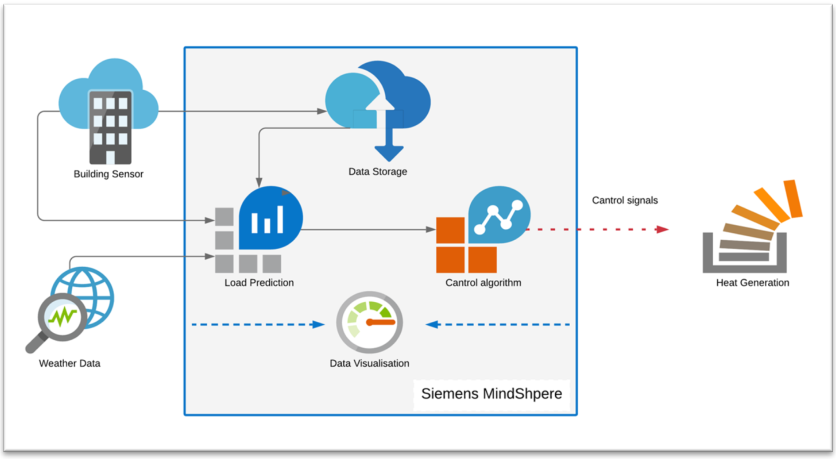


Fig. 2.10 Cloud Based Architecture

The proposed system envisions a building that is seamlessly connected to a heat generation source through the cloud, as depicted in Figure 2.10. The cloud serves as a repository for data storage, prediction of heat load, data visualization, and decision-making for controlling the heat generation. Sensor and smart meter data from the building are transmitted and stored on the cloud. Through the microservice architecture implemented on the cloud, data is visualized and heat load is predicted. This is achieved through a data-driven load prediction algorithm that utilises stored data and weather APIs. The predicted heat load value serves as an input for various control algorithms, including linear programming, integer programming, quadratic programming, nonlinear programming, stochastic programming, dynamic programming, combinatorial optimisation, infinite-dimensional optimisation, machine learning, and deep learning. The cloud-based approach allows for efficient and effective control of heat generation, making it a promising solution for achieving energy efficiency and carbon emission reduction in the district heating system.

To summarize, the proposed system is expected to yield three main research outputs. The first output is the development of a novel algorithm for load prediction that will improve the accuracy and reliability of the predictions. The second output is the implementation of data storage, visualization, load prediction, and control strategy on the cloud. This will enable leveraging the benefits of cloud computing, while also addressing the problem of optimising district heating. The third output is the integration of real-time data processing into the load prediction and control strategy. The decision-making process will occur on the cloud to ensure that the system is optimised, robust, and capable of processing real-time data. Overall, the proposed system aims to address the challenges associated with district heating optimisation and improve its energy efficiency while reducing carbon emissions.

Chapter 3

Methodology

This chapter delves into the method employed to accomplish the aim of the research. To attain the objectives, an experimental methodology was employed, given that the data was drawn from both real-life and simulated case studies. Furthermore, to avoid redundancy and ensure coherence, the methodology outlines the three distinct projects that underpin the research. Each of these projects is described in detail to provide a comprehensive account of the research process and outcomes. By employing this approach, the research seeks to achieve a holistic understanding of the subject matter and provide a valuable contribution to the existing body of knowledge.

3.1 Overview

The core of this thesis centers around two significant projects, namely, the REMOURBAN project and the Sharing Cities Project. In these projects, data was generated using a simulator in the Sharing Cities Project, while real-life data was acquired from the REMOURBAN Project sensors and smart meters. Due to the impact of the COVID-19 pandemic, a test rig was set up in the lab to contribute to research validation. The research approach and methodology adopted are depicted in Figure 3.1.

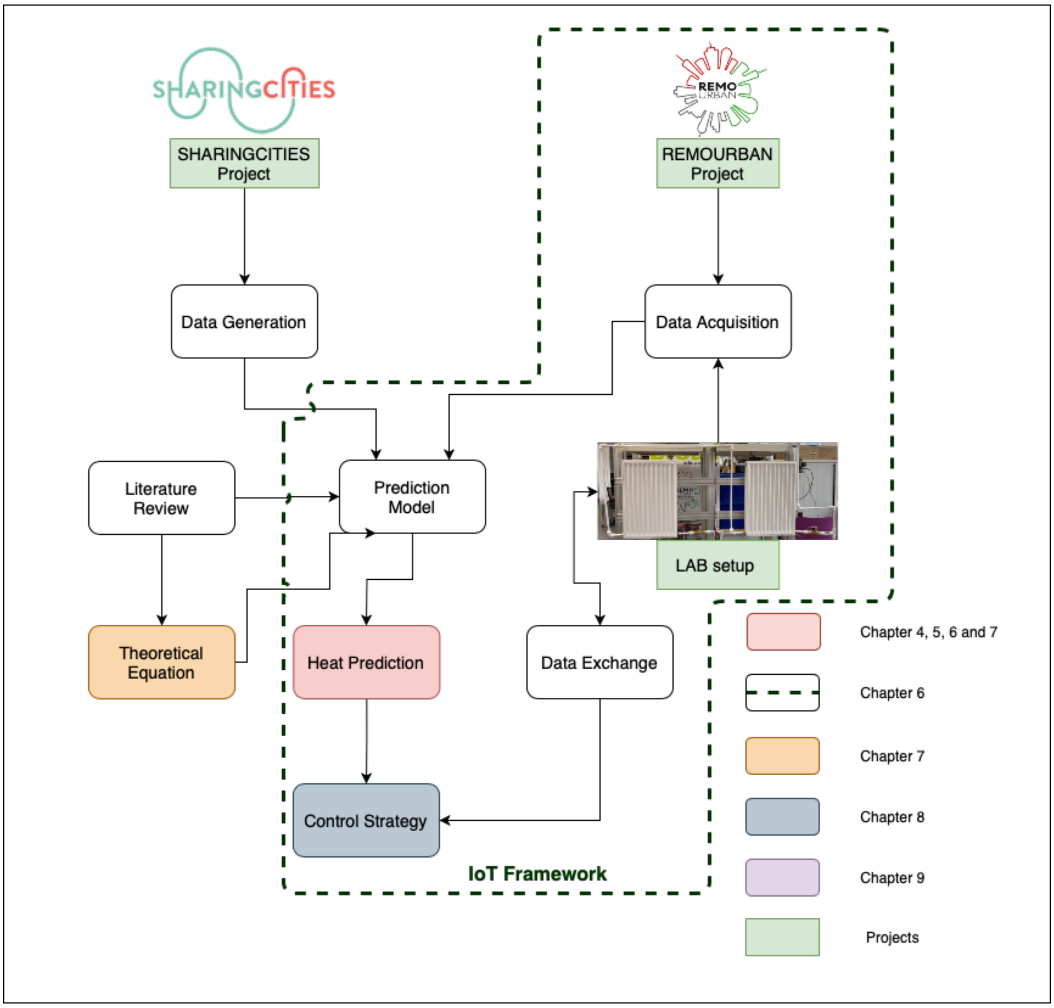


Fig. 3.1 Overview of Research Methodology

After the data was generated or acquired, prediction models were built based on the literature review conducted in Chapter 2. These models were subsequently used to predict heat demand, and the best models were determined based on their accuracy. The selected models were then used to develop a control strategy. Chapters 4 and 5 are primarily focused on determining the best model for heat demand prediction and comparing it with the type of data used.

Through this research, the aim is to provide valuable insights into the prediction of heat demand and contribute to the advancement of this field. By leveraging real-life and simulated data and building accurate prediction models, this thesis aims to develop a robust and effective control strategy for heat demand management.

In Chapter 6, the fundamentals established in Chapters 4 and 5 are utilised to develop a novel IoT framework tailored exclusively for district heating systems. The validation of this framework is achieved by developing a live data visualization tool, implementing multiple algorithms simultaneously, and controlling the system based on CO2 and fuel prices.

Chapter 7 focuses on deriving theoretical equations for heat estimation based on the principles of thermodynamics and heat transfer. Additionally, this chapter covers the development of user profiling based on the analysis of collected data. Moreover, as a byproduct of the developed framework, a simple yet effective control strategy is tested in the lab environment and discussed in Chapter 8.

By presenting the findings of these chapters, this thesis aims to provide valuable insights into the potential of novel IoT frameworks and simulation approaches in addressing the challenges facing district heating systems. Through a comprehensive analysis of theoretical equations, user profiling, and energy simulations, this research aims to offer practical solutions for achieving energy efficiency and sustainability in district heating systems.

3.2 Sharing Cities

3.2.1 Vision and goals

The vision and goals of the Sharing Cities project are rooted in the pioneering smart city strategy proposed by the Borough of Greenwich. As the first London borough to adopt such a strategy in October 2015, the Borough's plan is built around four primary components: changing neighbourhoods and communities, changing infrastructure, changing public services, and changing the Greenwich economy.

Energy management, building retrofit, e-mobility, smart parking, smart lampposts, and urban sharing platforms (USPs) are identified as the key focus areas of the smart city strategy. These focus areas are all encompassed within the Sharing Cities project, which seeks to leverage innovative technologies and sustainable solutions to enhance the quality of life for residents while promoting energy efficiency and sustainability. Through this project, the aim is to create a smarter, more connected, and sustainable urban environment, one that prioritizes the needs of its citizens and addresses the challenges of urbanization.

3.2.2 What is Sharing cities

The Sharing Cities project is more than just a testing ground for smart city technologies - it aims to provide an effective and unifying strategy for bringing smart cities to life. By promoting global industry and city collaboration, the project seeks to create cost-effective smart city solutions that can be integrated on a large scale, offering a huge market potential. The project partners will work closely with the European Innovation Partnership on Smart Cities and Communities, as well as other "lighthouse" consortia, to create a framework for municipal collaboration and citizen involvement, which will enhance trust between communities and their residents.

Supported by €24 million in funding from the EU, the Sharing Cities project aims to spur a total investment of €500 million, with more than 100 municipalities in Europe participating. The project will apply repeatable urban digital solutions and collaborative models in the "lighthouse" cities of Milan, London, and Lisbon. Buildings in the Royal Borough of Greenwich in London, Porta Romana/Vettabbia in Milan, and downtown Lisbon will be repurposed, and shared e-mobility services will be implemented. Energy management systems, smart lampposts, and an urban sharing platform will also be installed, with active interaction with the public.

In addition to these three "lighthouse" cities, Bordeaux, Burgas, and Warsaw will also work together to create, evaluate, or implement these ideas and models. Overall, the Sharing Cities project represents a significant step towards creating sustainable, smart cities that prioritize the needs of their citizens, while fostering innovation and collaboration on a global scale.

3.2.3 Project background

The Royal Borough of Greenwich, with its prime location along the river Thames and prestigious UNESCO World Heritage Site, serves as the home for the exciting Sharing Cities project. The borough's thriving tourism sector brings in £1.2 billion annually, and the four main areas of the Smart Urban plan are smart communities and localities, infrastructure supporting change, a creative and perceptive council, and wealth creation and higher-paying jobs. The plan includes initiatives to improve energy efficiency, reduce energy bills and CO2 emissions, improve safety and liveability, and increase comfort for residents.

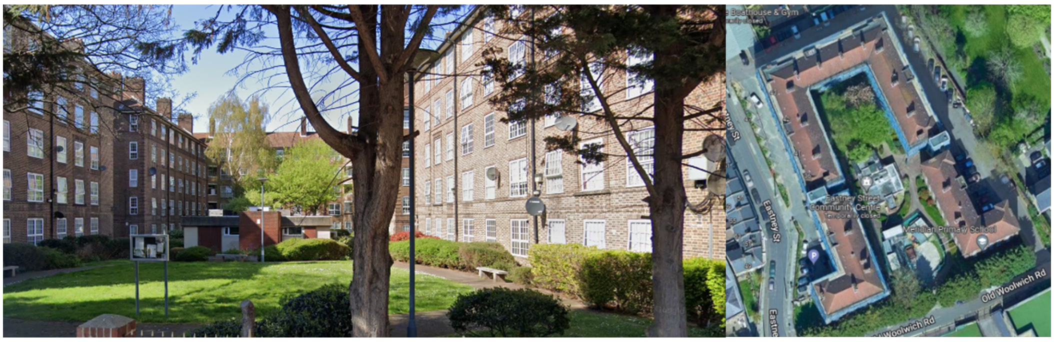


Fig. 3.2 Ernest Dence Estate

To engage citizens and encourage behavioural change, the plan seeks to understand engagement on digital services, sustainable mobility, and energy efficiency and to ensure locals can comment on proposed changes and their implementation. The Sustainable Energy Management System (SEMS) is a flagship initiative that involves partnerships among Siemens, Nottingham Trent University, Kiwi Power, Imperial College of London, and local government. SEMS aims to provide Advanced Process Control (APC) at the city level, enabling smart integration of city infrastructure and equipment to achieve optimised and predictive control.

As part of the Sharing Cities project, SEMS will be tested in the Royal Borough of Greenwich, with the Ernest Dence building serving as the test site for the heat generation and consumption. The system will use assets that generate heat from both gas and electricity, as well as thermal storage, to demonstrate multi-vector operation. The goal of SEMS is to provide a city-wide cloud-based platform for the implementation of energy-related APC, enabling stakeholders to benefit from a 'plug-and-play' facility. The expected benefits of SEMS include reduced operational costs, reduced energy consumption, better utilisation of existing city infrastructure capacity, and avoided infrastructure investment.

With advances in computing power and process control theory, APC principles can be effectively and efficiently implemented across cities, and SEMS is a promising approach to this. Chapters 4 and 5 of the PhD thesis will discuss the detailed results of SEMS implementation in the Ernest Dence building figure 3.2.

3.2.4 What is SEMS

Over the past few years, the Royal Borough of Greenwich has made significant progress in upgrading its energy infrastructure, with the installation of over 350kWp of photovoltaic solar panels and 18 electric vehicle charging stations. The Sustainable Energy Management System (SEMS) will ensure the effective use of these resources, creating a connected network of energy resources that are balanced and optimised to deliver efficient energy systems when required.

SEMS is aligned with the "Infrastructure" theme of the Greenwich Smart City Strategy, which embraces change and acknowledges the potential benefits of using sensors and Internet of Things (IoT) technology, smart neighbourhoods, sustainability, and energy efficiency. The project involves the creation of a sophisticated, data-rich management system that maximizes the use of refurbished buildings and shares energy data through an open platform to provide energy services that lower bills and reduce energy consumption. This will enable the creation and release of more advanced applications for individuals and authorities, utilising the mounted actuators and sensor layers.

3.2.5 Data Generated

The dataset used in this study was generated through the use of digital twin software, specifically EnergyPlus (EnergyPlus, 2021) and Ptolemy II (EnergyPlus, 2021). By inputting site details, energy demand was generated and used to optimise assets. Additionally, the generated dataset was used to predict future energy demand. The main focus of the research was on optimising district heating and predicting heat demand. Table 3-1 provides a list of the data variables that were generated and utilised.

Table 3.1 Feature for Sharing Project

Project	Feature
SHARING CITIES Project	Timestamp Occupancy Outdoor Temperature Heat Demand

3.3 REMOURBAN Project

The REMOURBAN project is a significant Smart Cities initiative that received support and funding from the European Union's Horizon 2020 research and innovation program. Nottingham Trent University played a pivotal role as a partner in the project, which emphasizes the need for creative, customized, and comprehensive regeneration models that maximize the convergence of ICTS, mobility, and energy. One of the primary goals of the REMOURBAN project is to reduce greenhouse gas emissions by retrofitting existing buildings to meet future GHG emission targets, and collaborating closely with local communities to achieve this objective.

Nottingham was one of the cities involved in the REMOURBAN H2020 project, which aims to develop and validate a sustainable urban regeneration model that takes advantage of the convergence of the energy, mobility, and ICT sectors in three lighthouse cities, including Valladolid, Spain, and Tepebasi/Eskisehir, Turkey. The city of Nottingham has already met its Energy

Strategy target, and it has set its sights on becoming the first carbon-neutral city in the United Kingdom by 2028.

The REMOURBAN project aims to create a sustainable urban regeneration model that maximizes the convergence of energy, mobility, and ICT sectors. Sneinton, a residential area in Nottingham, was chosen as the demonstration site for this project, where 27 residences were deep rehabilitated using both active and passive techniques to raise the living standards and energy efficiency of the homes. The retrofitting of homes is being made available to all households, regardless of ownership, and the REMOURBAN project grant is supporting over 400 social and private homes. Most of the properties in the region are social housing units owned by Nottingham City Council and managed by Nottingham City Homes. To achieve an elevated level of decarbonisation in residential homes, multi-faceted strategies such as retrofitting existing homes to reduce energy demand and decarbonising heat sources using energy systems that use renewable sources of energy are required. The ultimate goal is to reduce greenhouse gas emissions and create a sustainable living environment.

3.4 Experiment Overview

The experiment involves implementing the developed methodology in the real world on the pilot homes in the REMOURBAN project in Sneinton. The steps involved in this experiment are:

A. Install sensors and connect them to the local server.

B. Set up the Cloud infrastructure for collecting and storing data.

C. Implement the load prediction algorithm on the Cloud.

D. Streamline the collected data and use the predicted values as input for the control algorithm.

E. Generate the control variable from the control algorithm for heat generation.

F. Implement the control algorithm to regulate the heating system in the pilot homes.

G. Collect data on the performance of the system and use it to improve the methodology.

These experiments help in validating the methodology for collecting data from sensors and using it to predict and control the heating demand in residential homes. The implementation of the methodology in the REMOURBAN project pilot homes in Sneinton will help in achieving the project's objective of reducing greenhouse gas emissions and increasing energy efficiency in residential homes.

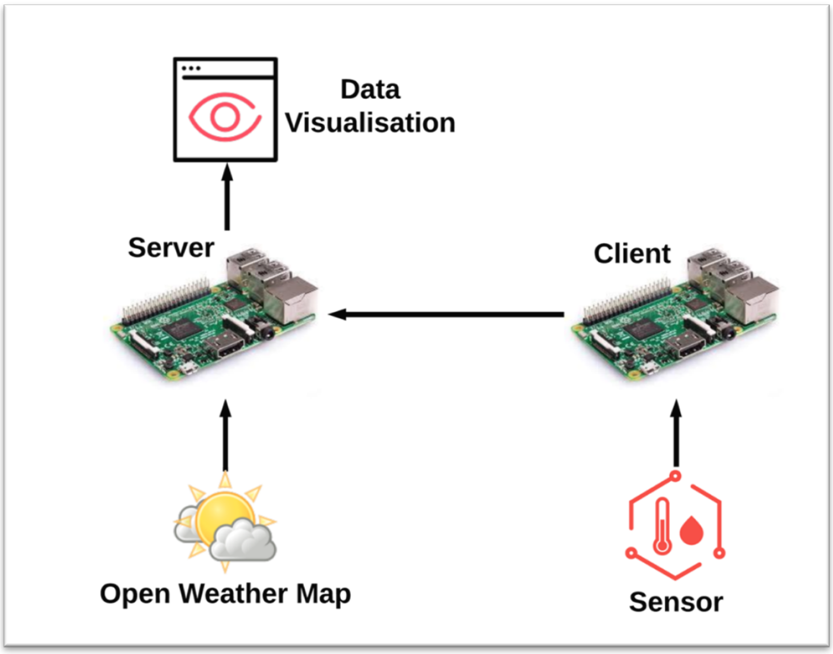


Fig. 3.3 Schematics of lab experiment

Figure 3.3 depicts the setup of the lab experiment, where one Raspberry Pi device is acting as a server and another as a client. The server is running a database to collect data, an Apache server for communication with the outer world, and HTML with Python as the backend for data visualization. The client device has temperature, humidity, CO2, and motion sensors attached to it to simulate real-world scenarios. The data collected from the client device would be used to explore the capabilities of Cloud, and once the validation is completed, data from the Nottingham buildings of the REMOURBAN project would be collected on Cloud. The testing of heat demand would be done on the Cloud using the collected data.

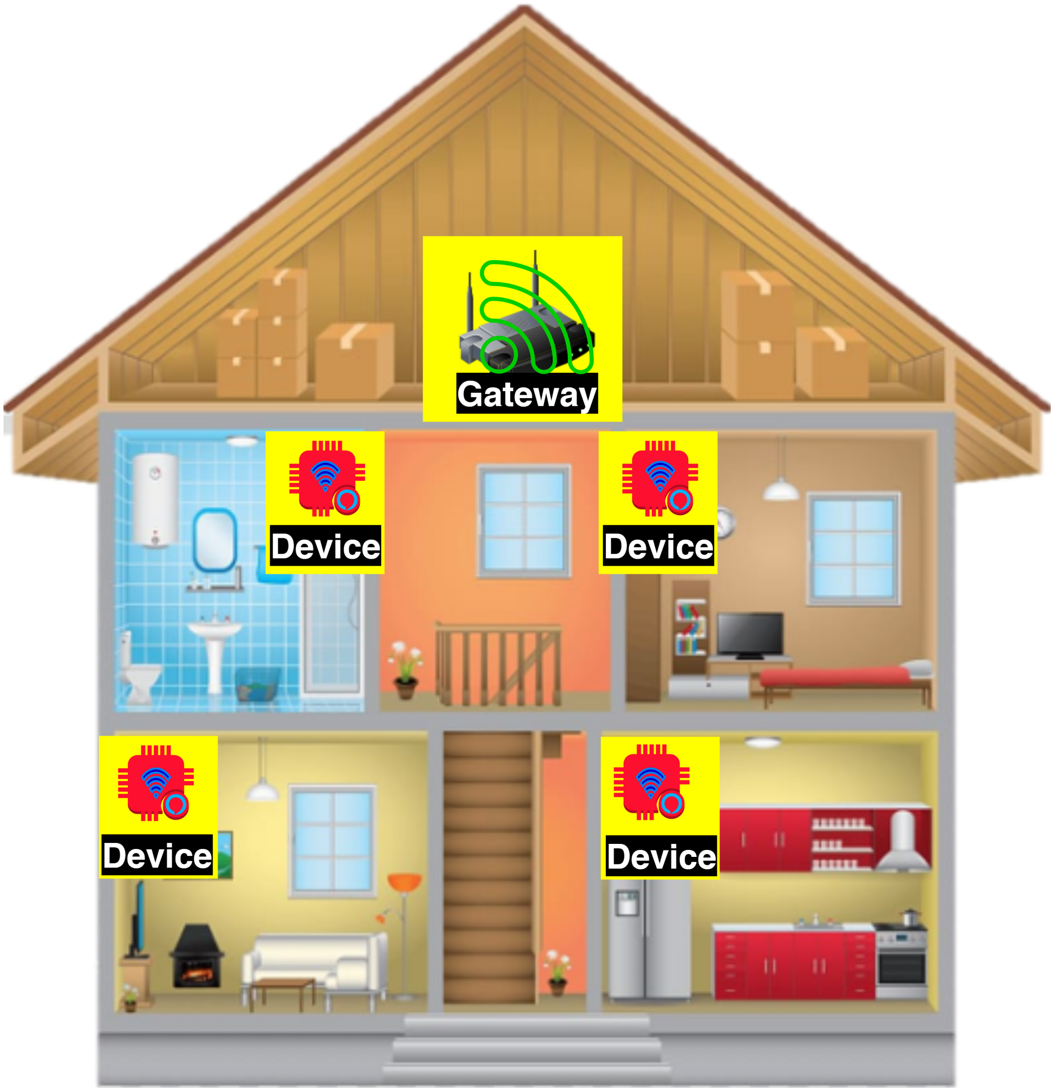


Fig. 3.4 Illustrative image of the site and sensor placing

In the REMOURBAN project, sensors were installed in the living room and bedroom, as shown in Figure 3.4. All the houses are connected via wireless

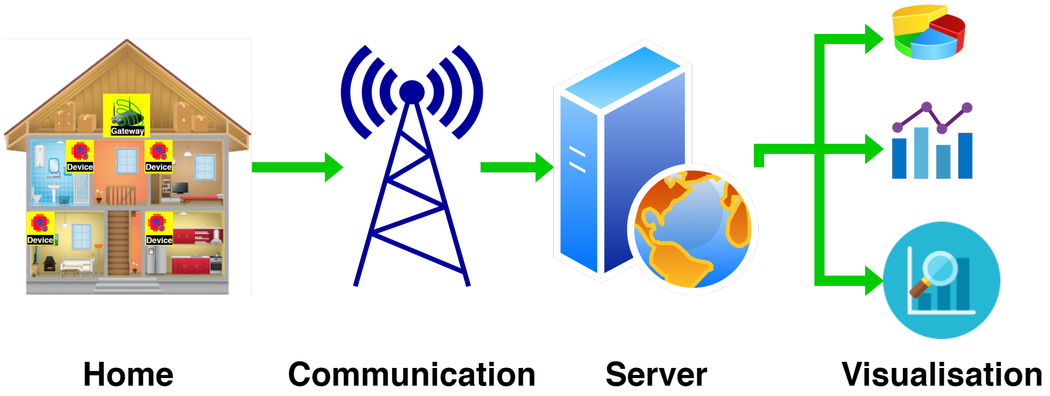


Fig. 3.5 Data flow

sensors to a gateway, which is then connected to the server on Cloud/University. The data analysis, control, and prediction algorithms are hosted on the server. The overall process is illustrated in Figure 3.5, showing how the sensors in the houses collect data, which is then analysed, and used to control the heating system to minimise energy consumption while maintaining a comfortable indoor temperature.

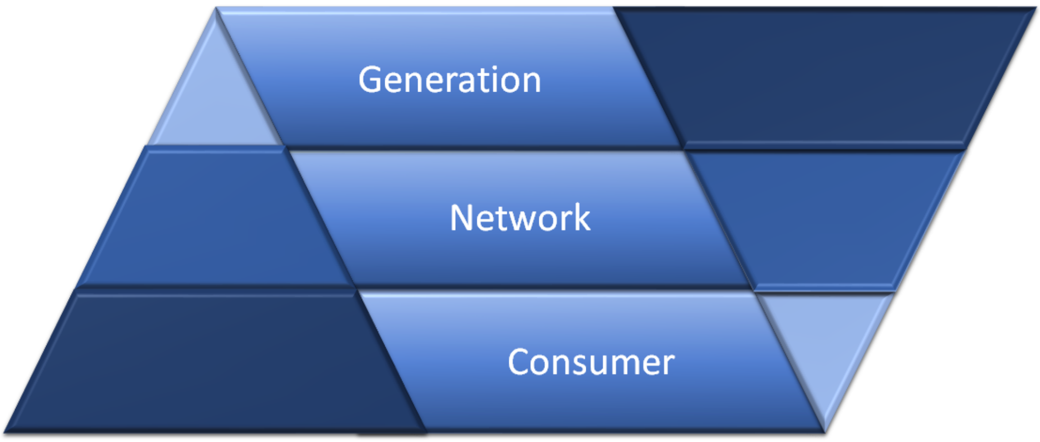


Fig. 3.6 Different Approach

The approach taken in this research work is different from the traditional top-down approach to district heating optimisation, which places a greater focus on heat generation and supply network rather than the customer. By taking a bottom-up approach, the research places the customer at the center of the optimisation process, allowing for a micro-level analysis of individual homes to optimise their heating requirements. The research aims to bridge the research gaps in heat estimation for individual homes, customer profiling and control strategyin a unique way. By taking a consumer-centric approach, the research aims to optimise district heating systems at the home level, which can improve energy efficiency and reduce energy waste. Figure 3.6 illustrates the bottom-up approach used in this research work.

This PhD thesis introduces a human-centric approach to heat estimation, which is based on a sound theoretical foundation and is supported by empirical evidence gathered from the REMOURBAN Project. To achieve accurate and personalised heat estimation, individual customer models are constructed using data collected from various sources, including sensors and weather APIs. To test the efficacy of the proposed approach, an experimental rig was set up in a laboratory to evaluate the real-time control strategy. Overall, the research work offers novel contributions to the field of district heating optimisation, demonstrating the importance of taking a customer-centric approach that empowers individual customers and helps to minimise energy waste.

3.5 Test Rig

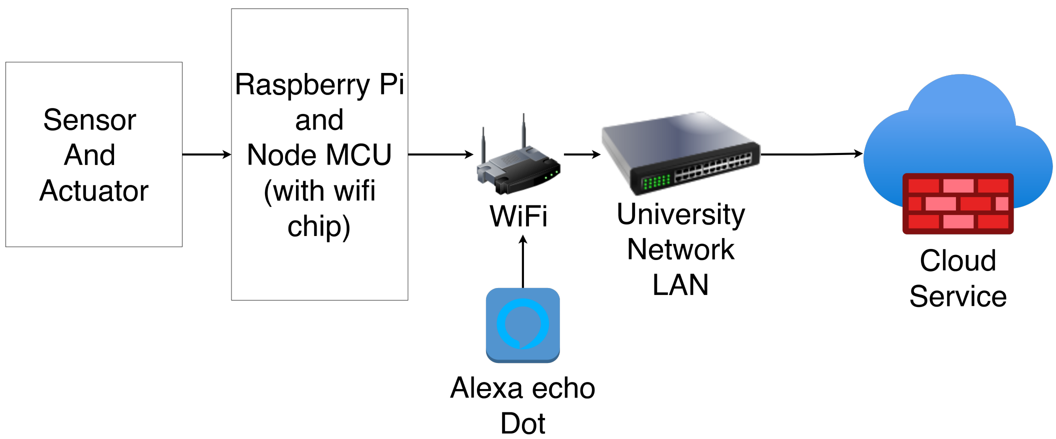


Fig. 3.7 Test-rig overview

The Covid-19 lockdown resulted in changes to the original research plan. To validate and extend the concept, a test rig was developed that emulates the heating system of a residential setup, specifically, radiator heating homes. The goal was to control heating in individual rooms based on pre-set schedules or automatically generated schedules. Additionally, users were given the flexibility to control the system using a mobile app, web app, or virtual assistant. The test rig was designed and built in the lab, and similar to the first experiment, different sensors were connected to a microcontroller or microprocessor with an actuator. The microprocessor was connected to the cloud via Wi-Fi for data collection and control. The virtual assistant, Alexa Eco Dot, was also connected to the same cloud as the sensors and actuators. Figure 3.7 depicts the block connection of the test rig setup.

3.6 Modelling process

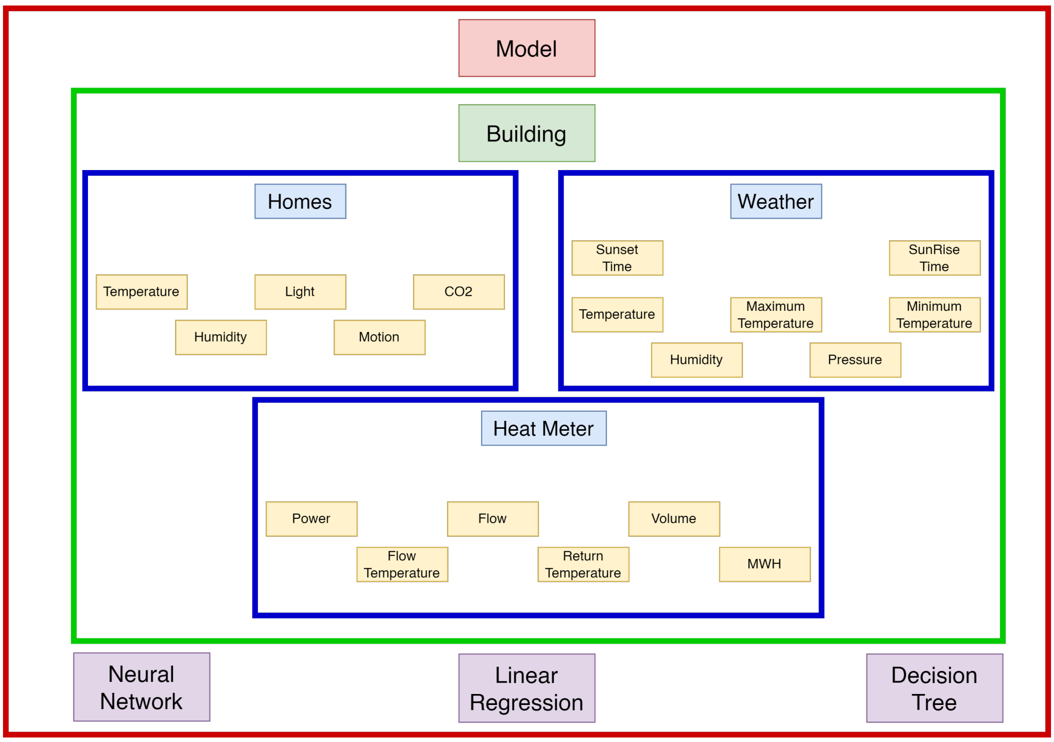


Fig. 3.8 Modelling overview

Figure 3.8 illustrates the variables that have been incorporated in the models along with the algorithms that have been employed for heat prediction. To create these models, three different datasets were collected from various sources. The first dataset includes data obtained from homes, such as temperature, light, motion, humidity, and CO2 levels. The second dataset includes data from smart heat meters, such as flow, power, flow/return temperature, volume, and heat consumption. The third dataset is based on weather APIs and includes data such as temperature, humidity, pressure, wind speed, maximum/minimum temperature, and sunrise/sunset times.

To create accurate models, three major algorithms were employed in this research, namely neutral network, decision tree, and linear regression. Models were constructed at different levels using datasets in different combinations to enhance the accuracy of the predictions.

Table 3.2 Feature of Dataset

Project	Feature
	Hour
	Minute
	Day
	Weekday
	Month
	Outdoor Temperature
	Indoor Temperature
REMOURBAN Project	Heat Demand
	Motion
	Pressure
	Humidity
	Wind Speed
	Sunset time
	Sunrise time
	Weather Type

3.7 Data

The collected data are from 40 homes part of the REMOURBAN project for the period from February 2020 to February 2021. Simultaneously, Accuweather API is used to collect the weather data. The variables used for this research work are shown in Table 3.1 and Table 3.2.

To improve data analysis and ease of use, some data sets are combined into groups for further use in this research. The timestamp feature is used to

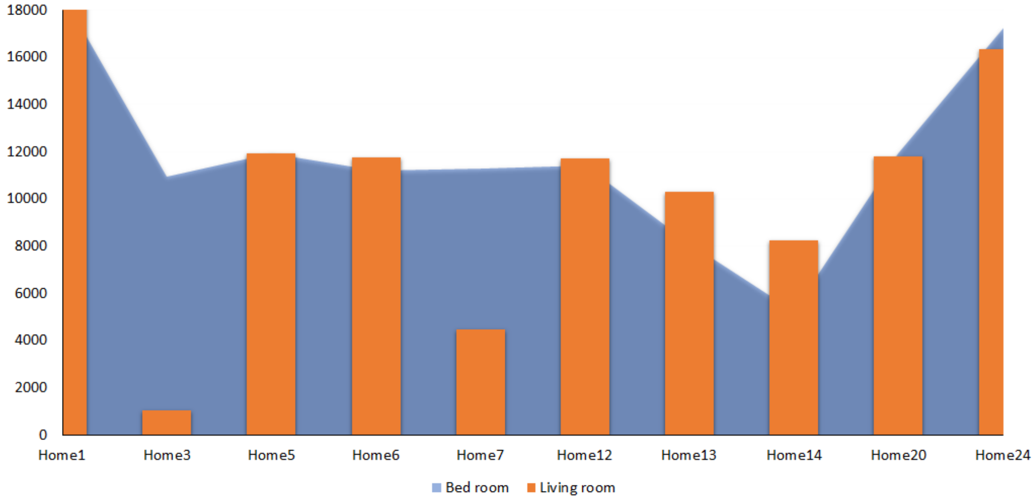


Fig. 3.9 Data Collected from each home

merge two datasets: the weather API and the home sensor data, creating a comprehensive dataset. The timestamp feature is divided into hourly, minute, daily, weekday, weekend, and monthly intervals. The weather API variables used in this research include temperature, humidity, pressure, sunset time, sunrise time, weather type, and wind speed. From the home sensor dataset, only temperature and motion variables are considered for this research, as they are most relevant to the study's objectives. After initial data analysis, several variables with less significance are removed, and the remaining variables are hour, minute, day, weekday, month, outdoor temperature, temperature, motion, heat demand, pressure, humidity, wind speed, sunset time, sunrise time, and weather type.

3.7.1 Secondary Data

$$Q_{\text{Home}} = \text{UA}\Delta T \tag{3.1}$$

This research work is a combination of real-life data calculation and simulation data. Government reports, literature, website, and book are used for secondary data. The secondary data are used in different capacities and at different levels. Equation 3.1 [186] is used to derive heat estimation equation. The U-values based on different wall type from [7, 8] are shown in Table 3.3. The areas exposed to the external environment drawn from [8]is displayed in Table 3.4. Table 3.3 and Table 3.4 can be used to simulate the heat estimation for different scenarios.

Table 3.3 U-Value based on wall types [7, 8]

Wall Types	U-value (W/m^2K)
Solid wall in very old buildings	2.30
Solid wall in old buildings	1.70
Unfilled cavity wall	1.50
Solid wall with 100 mm thick external insulation	0.32
Filled cavity wall with 100 mm thick external wall	0.25

Table 3.4 Area exposed based on home type [8]

Dwelling	Net area of external walls to be retrofitted (m^2)	Exposed Surface Area (m^2)
End-terrace	80	138
Mid-terrace	44	103
Semi-detached	91	159
Detached	141	257
Bungalow	64	152

Table 3.5 and Table 3.6 shows the price per unit across the United Kingdom. This data is used to calculate the cost and validate the control strategy.

Average Unit Price in UK [9]

Table 3.5 Average Unit Price in UK [9]

UK region	Average unit rate (pence per kWh)
North Scotland	15.60p
South Scotland	13.97p
Northeast	14.26p
Northwest	14.27p
Yorkshire	13.92p

Table 3.6 Average Unit Price in England [9]

UK region	Average unit rate (pence per kWh)
East Midlands	13.86p
West Midlands	14.25p
Merseyside and North Wales	15.18p
South Wales	15.07p
Southwest	15.54p
London	14.53p
Southeast	14.68p
Eastern	14.06p
Southern	14.29p

3.7.2 Model Parameter

The published paper [16] shows the procedure followed to select the variables for building the prediction model. Table 3.1 shows the list of the variables which are used for two purposes, firstly is for training and testing the simulation model using an appropriate machine learning algorithm, secondly for applying real-time control strategy.

As mentioned above to build a data-driven model for heat estimation, the input variables are hour, minute, day, weekday, outdoor temperature, pressure, humidity, wind speed, sunset, sunrise, and weather type, whereas output variable is the desired temperature by the customer. User profiling was carried out using the collected data from the home dataset and using outdoor temperature from the weather dataset. The control strategy uses the model built for heat estimation for heat generating schedules (detail can be found in the paper) as well as real-time temperature and motion data.

The variables related to timestamp can be described as an hour value from 0 to 23, minute values (15, 30, 45 and 60), days from 1 to 31, weekdays Monday to Friday, and months from 0 to 11. Moreover, variables derived from the weather API can be described in units that are temperatures in Celsius, pressure in pascals, humidity in percentage (0-100), wind speed in meter per second (m/s), and time in seconds (sunset and sunrise). The rest of the variables that are motion measured from 0 to 5 and the weather types are clear, clouds, drizzle, rain, fog, haze, mist, snow, thunderstorm, and tornado.

3.8 Equipment, Components and Software

The hardware and software used for building the test ring and conducting experiments are summarized in Table 3.7. The hardware components include a Raspberry Pi, which serves as the central controller, a Mac book and a Lenovo workstation. The system also includes various sensors such as the LCD display, the DHT11 temperature and humidity sensor, a solenoid valve, a water pump with 1 Mpa of pressure, and a water tank.

In addition to the hardware, the system relies on a suite of software applications. Python is used for programming the Raspberry Pi, while Apache Server is used for hosting web content. Salesforce Cloud is used for data management, while simulation tools such as Design Builder and EnergyPlus are used for building and energy simulation, respectively.

To analyse and visualize the data, the system utilises software tools such as Excel and Visio. Pynomo is also used for equation and model generation. The system runs on different operating systems such as Mac OS, Windows OS, and Linux OS.

All of these hardware and software components work together to facilitate the construction of the test ring, as well as the execution of experiments and the collection and analysis of data.

Table 3.7 Details of hardware and software used

Hardware	Software
Raspberry Pi	Python
Mac book i5, 8 GB, 256 GB	Apache Server
Lenovo Intel Xeon 8 core, 16 GB, 256 GB SSD	Salesforce Cloud
LCD 16 x 4	Design Builder
DHT 11 (Temperature & Humidity)	EnergyPlus
Solenoid Valve	Excel, Visio
Water Pump 1 Mpa	Pynomo
Water Tank	Mac OS, Windows OS, and Linux OS

1. Raspberry Pi - A small, single-board computer that can be used for a variety of applications. It's commonly used for prototyping and DIY

projects. In this case, it is likely being used as a central controller for the other components.

2. Mac book i5, 8 GB, 256 GB - A laptop computer made by Apple. It likely serves as a development and testing machine for the project.

3. Lenovo Intel Xeon 8 core, 16 GB, 256 GB SSD - A powerful desktop computer with a high-end Intel processor. It may be used for running simulations and other computationally intensive tasks.

4. LCD 16 x 4 - A type of display that can show 16 characters per line, with 4 lines. It may be used to display information from sensors or other components in the system.

5. DHT 11 (Temperature & Humidity) - A sensor that can measure temperature and humidity. It likely provides data to the system that is used for control and monitoring purposes.

6. Solenoid Valve - A valve that is controlled by an electric current. It may be used to regulate the flow of water or other fluids in the system.

7. Water Pump 1 Mpa - A pump that is used to move water or other fluids through the system.

8. Water Tank - A container that holds water or other fluids.

9. Python - A programming language commonly used for scientific and engineering applications. It's likely being used to write code for the Raspberry Pi or other components in the system.

10. Apache Server - A web server that can be used to serve web pages or other content. It may be used to provide a user interface for the system.

11. Design Builder - A building simulation software that can be used to model and analyse building performance. It may be used to optimise the performance of the system.

12. EnergyPlus - A building energy simulation software that can be used to model and analyse building energy use. It may be used to optimise the energy efficiency of the system.

13. Excel, Visio - Microsoft Excel is a spreadsheet program that can be used for data analysis and manipulation, while Microsoft Visio is a diagramming program that can be used to create flowcharts and other visual representations. They may be used for data analysis and visualization purposes.

14. Pynomo - A software package for generating equations and models. It may be used for modelling and optimisation of the system.

15. Mac OS, Windows OS, and Linux OS - Three popular operating systems that may be used to run various software applications in the system.

Chapter 4

Heat Prediction for SEMS

4.1 Introduction

The high demand for energy and limited space for renewable generation in urban areas present significant challenges to achieving decarbonisation goals. To address this challenge, the electrification of heating and transportation through the adoption of heat pumps (HP) and electric vehicles (EV) is a promising solution with the potential to significantly reduce the environmental impact of the energy sector [187]. However, to ensure the effective management of demand across various energy vectors and prevent excessive strain on grid infrastructure [188], suitable energy management strategies must be implemented. Additionally, it is crucial to effectively utilise renewable generation sources to supply power to the grid while minimising the cost burden for end-users [189].

To achieve this, flexible demand management techniques must be implemented to shift and store demand across different energy vectors. However, the complexity of this multi-vector problem requires appropriate tools to address it. In particular, models and forecasting methods are essential to predict future system states, while optimal control approaches can determine the best operational solutions to multi-objective control problems based on these predictions.

Thus, it is crucial to develop and employ suitable models, forecasting methods, and control approaches to tackle this complex problem effectively. This will enable the efficient and cost-effective management of energy demand, the integration of renewable energy sources, and the achievement of decarbonisation goals in urban settings.

As urban data infrastructures become more widely established, opportunities emerge to make use of techniques from machine learning and data science for forecasting and decision-making in energy management strategies. In [190] for example, a long short-term memory learning model is applied for district-level building energy modelling, while for more general district level energy forecasting, [191], compare the performance of three machine learning algorithms. Aside from forecasting, optimal control methods have been applied for real-time operational decision-making in the energy domain, showing clear advantages compared to more traditional approaches [192], with a review of different strategies provided in [193]. Achieving the goals of forecasting and optimisation in a transferable and scalable manner remains a challenge however, with components, data architectures and system hierarchies potentially varying greatly in different contexts. Software tools are needed that can successfully deploy optimisation and forecasting methods in a consistent, replicable manner.

This chapter explores the application of various machine learning algorithms to predict heat demand using simulated data generated from a digital twin. The prediction of heat demand is crucial in today's energy system due to the increasing use of district heating (DH) systems, which are more efficient and environmentally friendly than traditional heating systems.

The data used for heat demand forecasting was generated by a simulator. To generate the predictions, the K-Means cluster, KNN, Decision Tree, Neural Network, and Linear Regression algorithms were applied and tested on the simulated data. Figure 4.1 below illustrates the general process followed in this study.

First, historical data generated from the simulator was used to train the machine learning model. Then, weather data from an API was passed to the trained model to generate heat demand predictions. By applying different machine learning algorithms to the simulated data, this study provides valuable insights into the performance of these algorithms in predicting heat demand, which can inform the development of more accurate and efficient forecasting methods for district heating systems.

4.2 Input Data

The input data for the heat demand prediction model consists of two data streams. The first stream is historical data, which is used to train the model. It includes two variable values: outdoor temperature and DH. The second stream is weather API data, which provides 12 hours of forecasted temperature values.

Before passing the input data to the model, it needs to be cleaned to ensure accuracy and consistency in the predictions. The cleaning process involves dropping any null fields from the data. Additionally, the DH values are rounded to improve the accuracy of the model and reduce the number of attributes for data management.

In contrast, the second data stream, which is the weather API data, does not require any cleaning as it is generated from a reliable source and contains accurate forecasted temperature values for the upcoming 12 hours. By ensuring that the input data is clean and accurate, the heat demand prediction model can produce more reliable and accurate results, which can inform the development of effective demand management strategies for district heating systems.

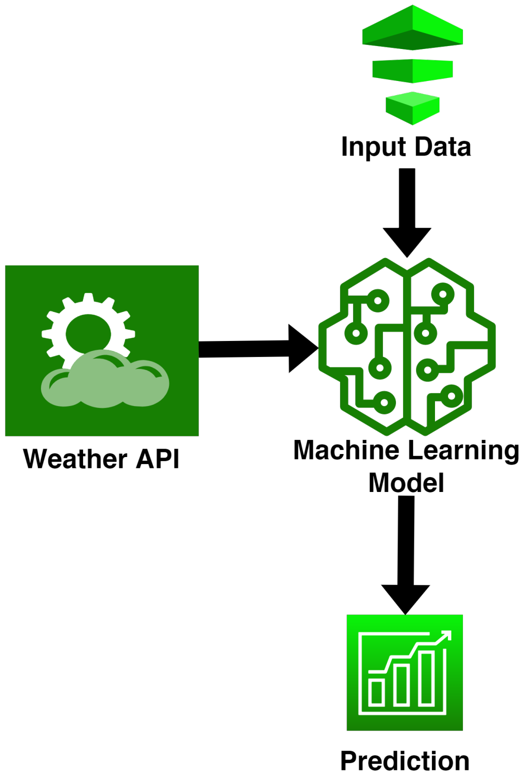


Fig. 4.1 Prediction process model diagram

Table 4.1 Data generated for Sharing Project

Outdoor Temperature (^{0}C)	Occupancy	DH
7.2725	1.0	220
6.0650	1.0	220
4.8575	1.0	240
3.6500	1.0	307
3.8375	1.0	250

The input data used for training the models and generating the output has been modified to improve the accuracy of the predictions. In real-life situations, occupancy is determined using motion sensors, but since they are not available in this study, occupancy is determined based on time. Additionally, the results are generated using three different data sets, which include the data generated from energy plus, pre-retrofit, and post-retrofit.

The table 4.2 shows a sample of the data with the predicted value. In addition to occupancy, the Outdoor temperature is another data field used to generate the model. The model is trained using Day, Month, Hour, and Outdoor Temperature as input and Pre-retro or Post-retro or Extra heat demand as the output to be predicted.

By using different data sets and input variables, this study provides valuable insights into the factors that influence heat demand and can help inform the development of more accurate and efficient heat demand prediction models.

Table 4.2 Simulated Second Dataset

DateTime	Pre retro (kW)	Post retro (kW)	Extra heat demand data (kW)	Day	Month	Hour	Predict Post Retro	Predict Pre Retro	Predict Extra Heat Demand
2018-01-01 00:00:00	180.79	143.26	154.13	1	1	0	216.69	255.92	247.68
2018-01-01 00:15:00	187.89	144.56	156.02	1	1	0	217.0	256.57	248.19
2018-01-01 00:30:00	209.93	175.43	196.02	1	1	0	217.32	257.22	248.71

Table 4.2 Simulated Second Dataset

DateTime	Pre retro (kW)	Post retro (kW)	Extra heat demand data (kW)	Day	Month	Hour	Predict Post Retro	Predict Pre Retro	Predict Extra Heat Demand
2018-01-01 00:45:00	223.09	186.31	210.58	1	1	0	217.63	257.87	249.22

4.3 Building Model

The model is built using the following procedure:

1. The data set is read from a CSV file.
2. The date, time, day, month, hour, and minute are extracted from the data.
3. The data set is divided into an 80-20 ratio for training and testing the model.
4. The machine learning model is trained using 80% of the data.
5. The day, month, hour, and outdoor temperature of the remaining 20% test data are passed into the trained model to predict the DH.
6. The accuracy of the model is derived based on the predicted value and the known value from the test set.

This procedure is applied with different combinations to generate the results using the three different data sets mentioned earlier. By training the machine learning model on a subset of the data and testing it on another subset, the accuracy of the model can be evaluated and improved. This approach helps to ensure that the model can accurately predict DH under various conditions and inform the development of effective demand management strategies.

4.4 Models

In this project, various machine learning models are tested and the outputs are compared to determine the best algorithm for further development. By evaluating and comparing the performance of different models, it is possible to identify the strengths and weaknesses of each algorithm and select the one that is best suited to the specific requirements of the project. This process of testing and evaluation is crucial in ensuring that the final model is accurate, efficient, and reliable, and can be used to inform decision-making and strategy development.

4.4.1 K-Means Clustering

In the clustering algorithm, the first important step is to determine the optimal number of clusters based on the data collected. This is typically done using a method such as the elbow method, which helps to identify the point at which the addition of more clusters does not significantly improve the model's performance. Figure 4.2 shows the optimal number of clusters identified using the elbow method. As can be seen in the figure 4.2, the optimal number of clusters is reached at 5, as increasing the number of clusters beyond this point does not result in a significant improvement in performance. Based on the

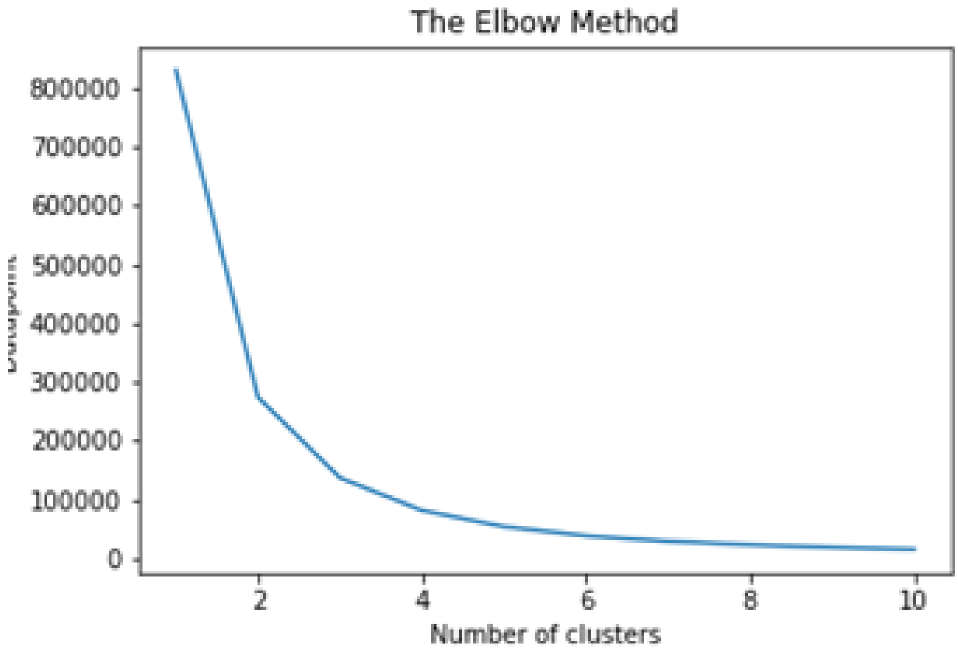


Fig. 4.2 Elbow Method

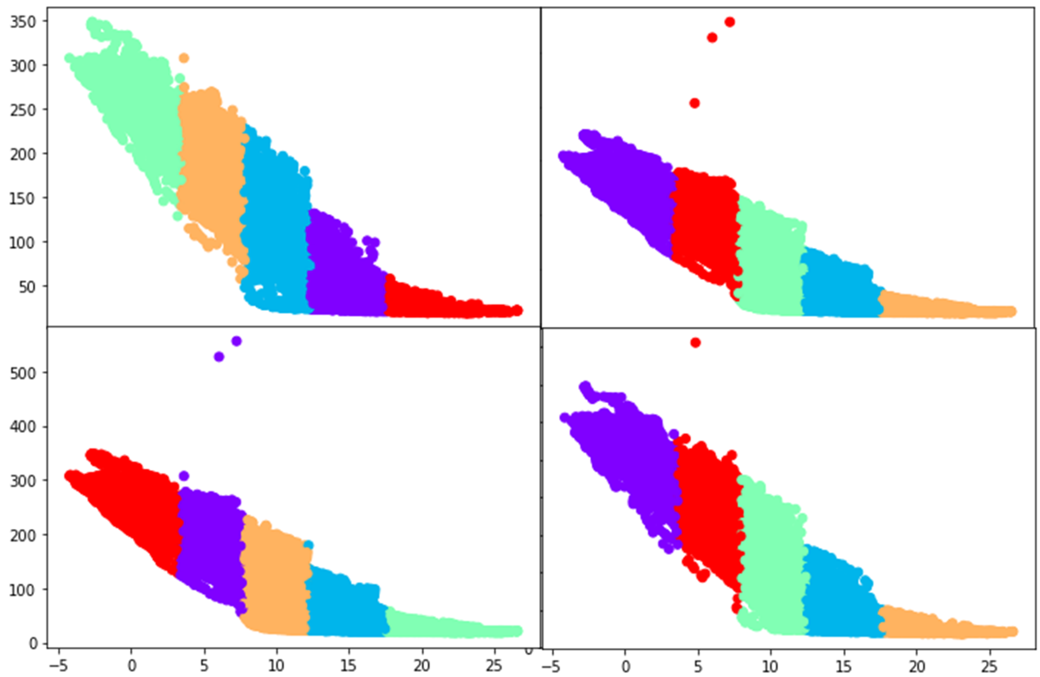


Fig. 4.3 K-Means Clustering output

results of the elbow method, a K-means model is trained with 5 clusters to further analyse and understand the data.

The top two charts show the results of applying the K-Means model to both the training and testing sets using both variables, while the bottom two charts show the results using only a single variable (temperature). It is clear that there is not a significant difference in the results between the top and bottom charts (as shown in Figure 4.3). However, it is important to note that the outdoor temperature appears to be the dominant factor in determining the heating demand, which can result in a wide range of heating demand values for a given temperature. This can make accurate prediction more challenging and can introduce bias in the model. Therefore, further analysis and refinement of the model may be needed to improve its accuracy and reduce the impact of this bias.

4.4.2 K-Nearest Neighbour

In this section, the K-Nearest Neighbour (KNN) algorithm was applied to the dataset, and the resulting output was plotted in Figure 4.4. The KNN algorithm is a simple yet effective classification algorithm used for both regression and classification tasks. It operates by identifying the k-nearest neighbours to a particular data point in the training set and then predicting the label of the data point based on the labels of its K-nearest neighbours.

The output in Figure 4.4 clearly shows that all the clusters are overlapping with each other, which suggests that the KNN model is inaccurate. This could be due to the fact that the KNN algorithm is highly dependent on the choice of K, and selecting an optimal value of K can be a challenging task. In addition, the KNN algorithm assumes that the data is uniformly distributed across the feature space, which may not always be the case in real-world scenarios.

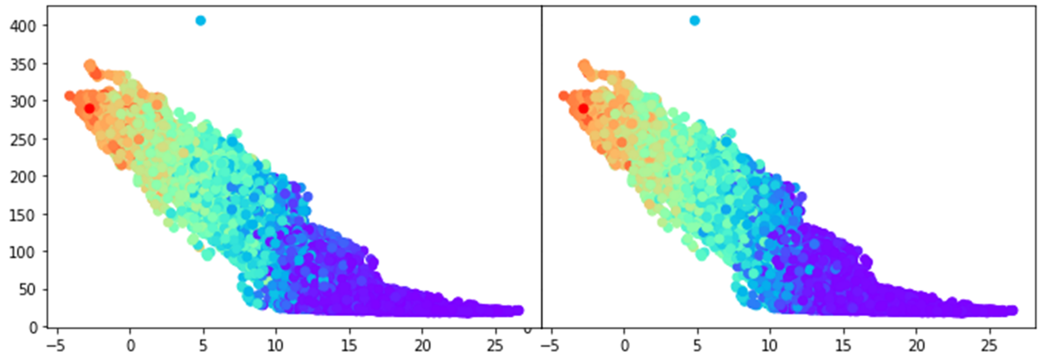


Fig. 4.4 K-Nearest Neighbour output

Therefore, it is important to consider the limitations of the KNN algorithm before selecting it as the final model for the task at hand.

4.4.3 Decision Tree

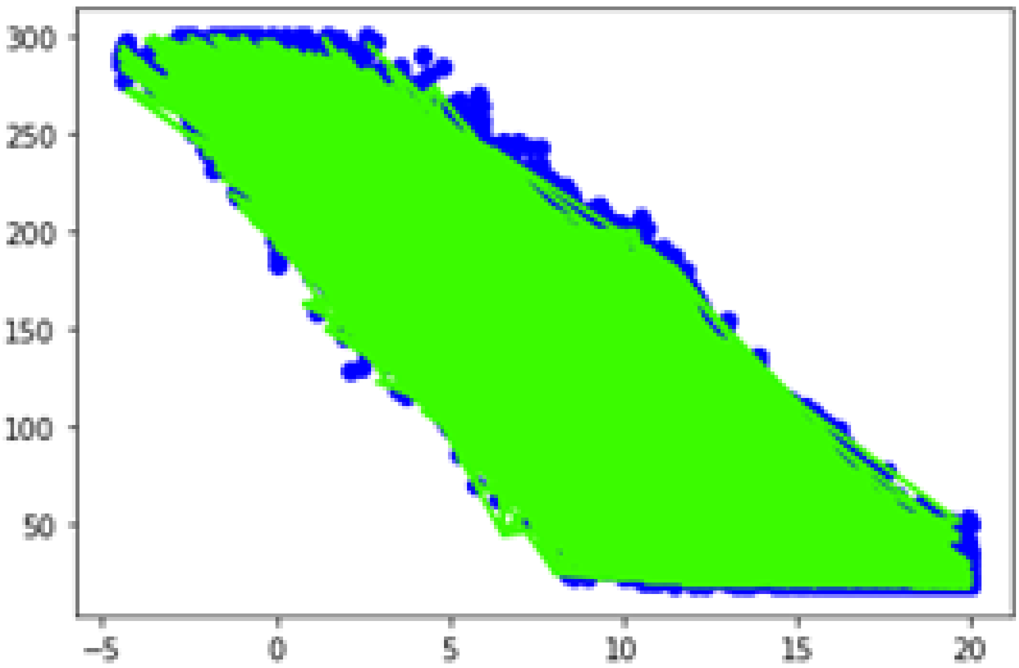


Fig. 4.5 Decision Tree output

To expand on this, the decision tree algorithm was not able to effectively capture the relationship between the input variables and the output variable. This can be seen in Figure 4.5, where the predicted values are significantly different from the actual values, resulting in a high error rate. This indicates that the model is not accurate and cannot be relied upon for heat demand prediction.

The decision tree algorithm works by recursively splitting the data into subsets based on the value of an input variable, and then making predictions based on the majority class or average value of the output variable within each subset. However, in this case, the algorithm was not able to identify a set of input variables that would effectively split the data into meaningful subsets.

Therefore, it is clear that the decision tree algorithm is not suitable for this particular dataset and cannot provide accurate predictions of heat demand. Other machine learning algorithms need to be explored to identify the best approach for this particular problem.

4.4.4 Neural Network

In this case, it seems that the neural network model did not perform well even after trying different combinations of hyperparameters such as hidden layer sizes and activation functions. This may indicate that the neural network is not well-suited for this particular dataset or that the data may not have enough information for the neural network to learn and make accurate predictions. In such cases, it is important to consider alternative machine learning algorithms or collect more relevant data that may better inform the model.

4.4.5 Linear Regression

The results of the study showed that Linear Regression was the most effective algorithm among the three tested, due to the fact that the dataset contained only two variables: outdoor temperature and desired DH. The outdoor temperature was found to have a significant impact on the required temperature for the desired DH, as indicated by the home temperature desired. A summary of the accuracies of all tested models is presented in Table 4.3.

The K-Means model could not produce a measurable level of accuracy, while the KNN and Neural Network models achieved accuracies of 3.46% and 3.06% respectively for the two-variable input, and 2.81% and 4.00% for the single-variable input. The Decision Tree model produced an accuracy of 2.75% for the single-variable input. On the other hand, the Linear Regression model achieved a high R score of 85.23% for the two-variable input and 83.87% for the single-variable input.

Based on these results, it was concluded that Linear Regression would be the most effective algorithm for this problem. Although the input variables of temperature and motion were considered in the project, the accuracy difference between the two-variable and single-variable inputs was not significant. Moreover, the acquisition and generation of motion data were found to be challenging for the project. Therefore, Linear Regression was deemed the best-performing model for this chapter. Further refinement of the Linear Regression model was conducted by building various scenarios to finalise the model.

Table 4.3 Summary of accuracy

Models	Two Variables	Single Variable	Remarks
K-Means	N/A	N/A	It is difficult to find accuracy.
KNN	3.46 %	2.81 %	Accuracy
Neural Network	3.06 %	4.00 %	Accuracy
Decision Tree	N/A	2.75 %	Accuracy
Linear Regression	85.23 %	83.87 %	R Score

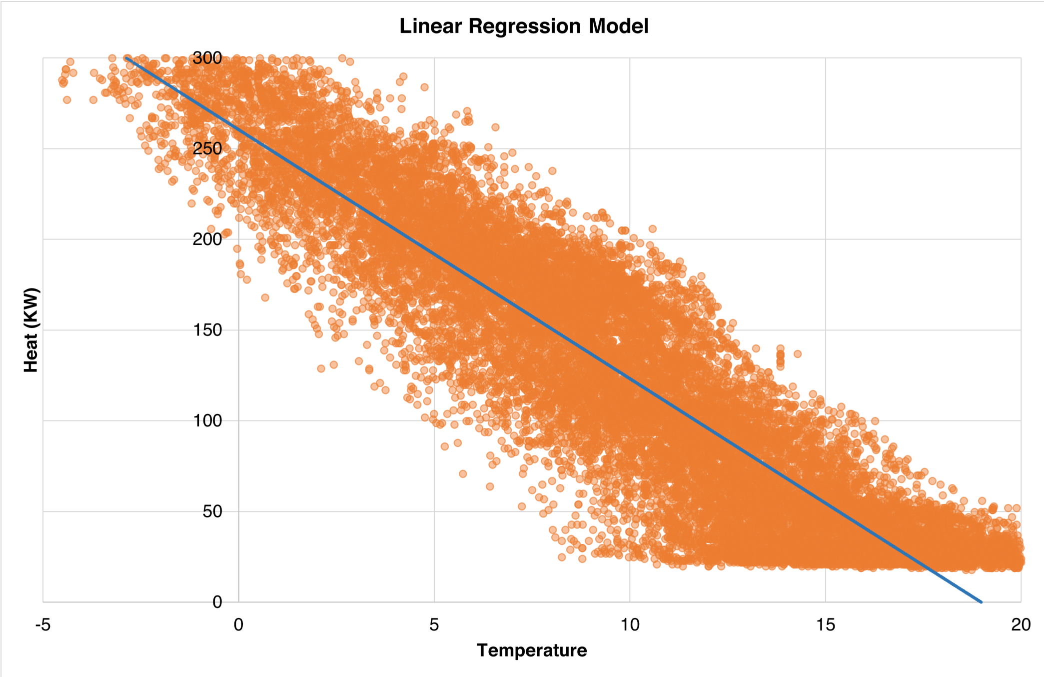


Fig. 4.6 Linear regression output

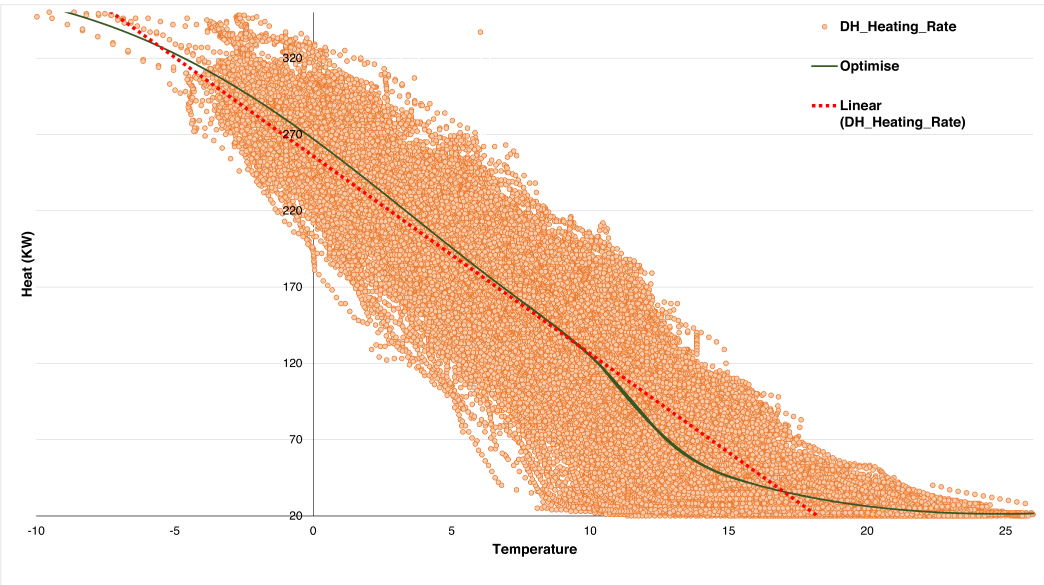


Fig. 4.7 Curve Fitting output

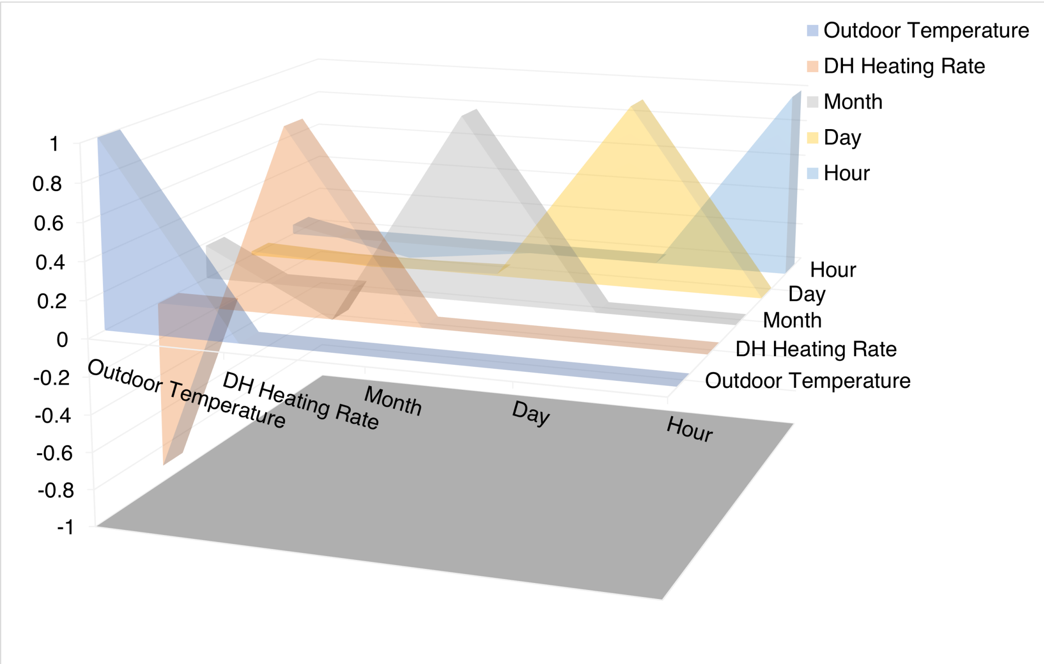


Fig. 4.8 Correlation Between Features

Figure 4.6 depicts the output of the linear regression model, while Figure 4.7 shows the results of the curve fitting model. Comparing the two results, there is only a 2% difference in accuracy, which is not significant. Thus, sacrificing accuracy can be considered to leverage the computational cost. The final linear regression model that was constructed will be discussed in a later section of this chapter.

4.5 Scenarios

The primary result obtained from the dataset was the correlation between the input variables and the output variable. As shown in Figure 4.8, it can be concluded that outdoor temperature is a critical determinant of heating rate. Linear regression was employed with different scenarios to determine the optimal solution for the problem.

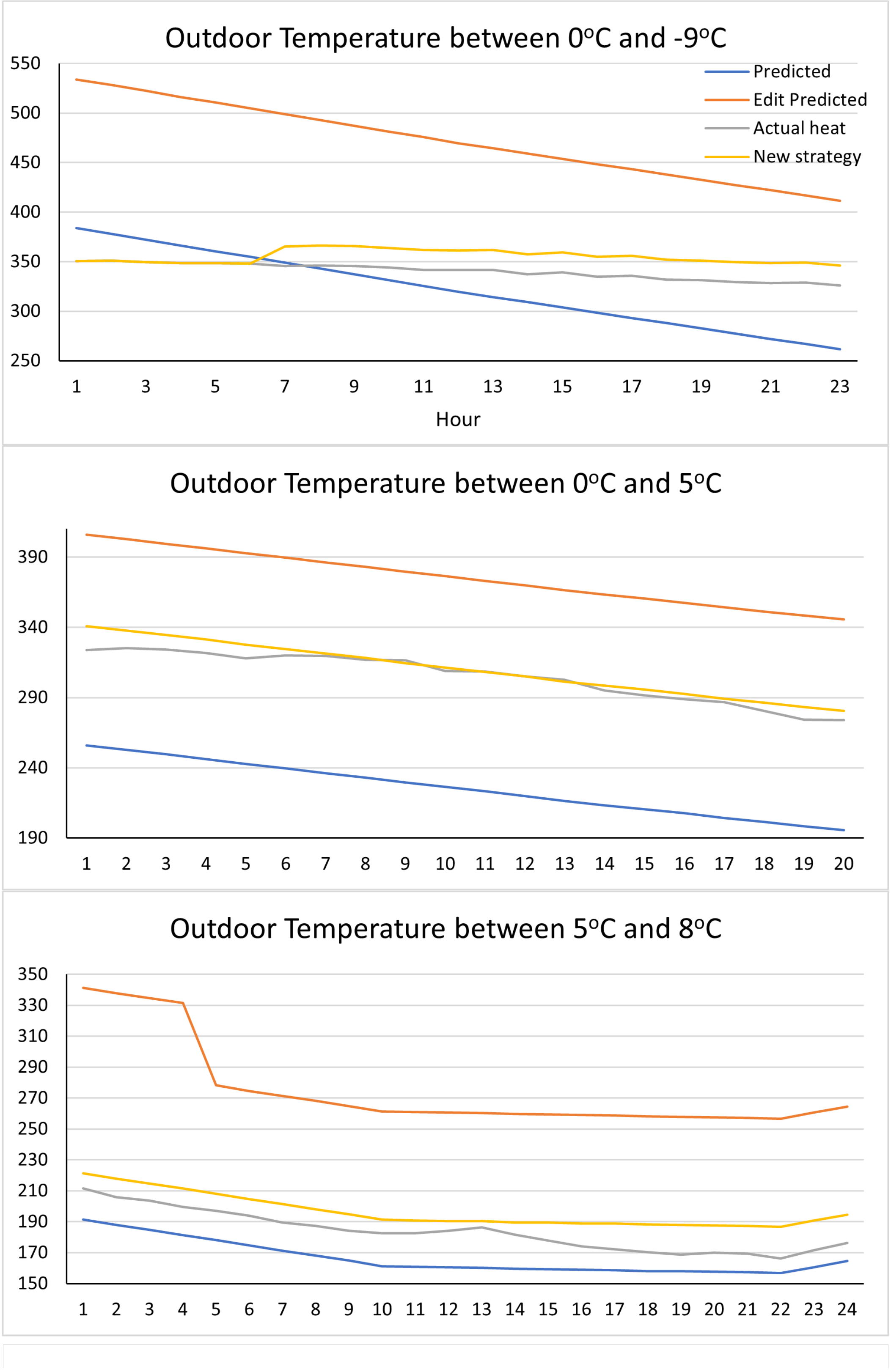


Fig. 4.9 Outdoor Temperature based division output part 1

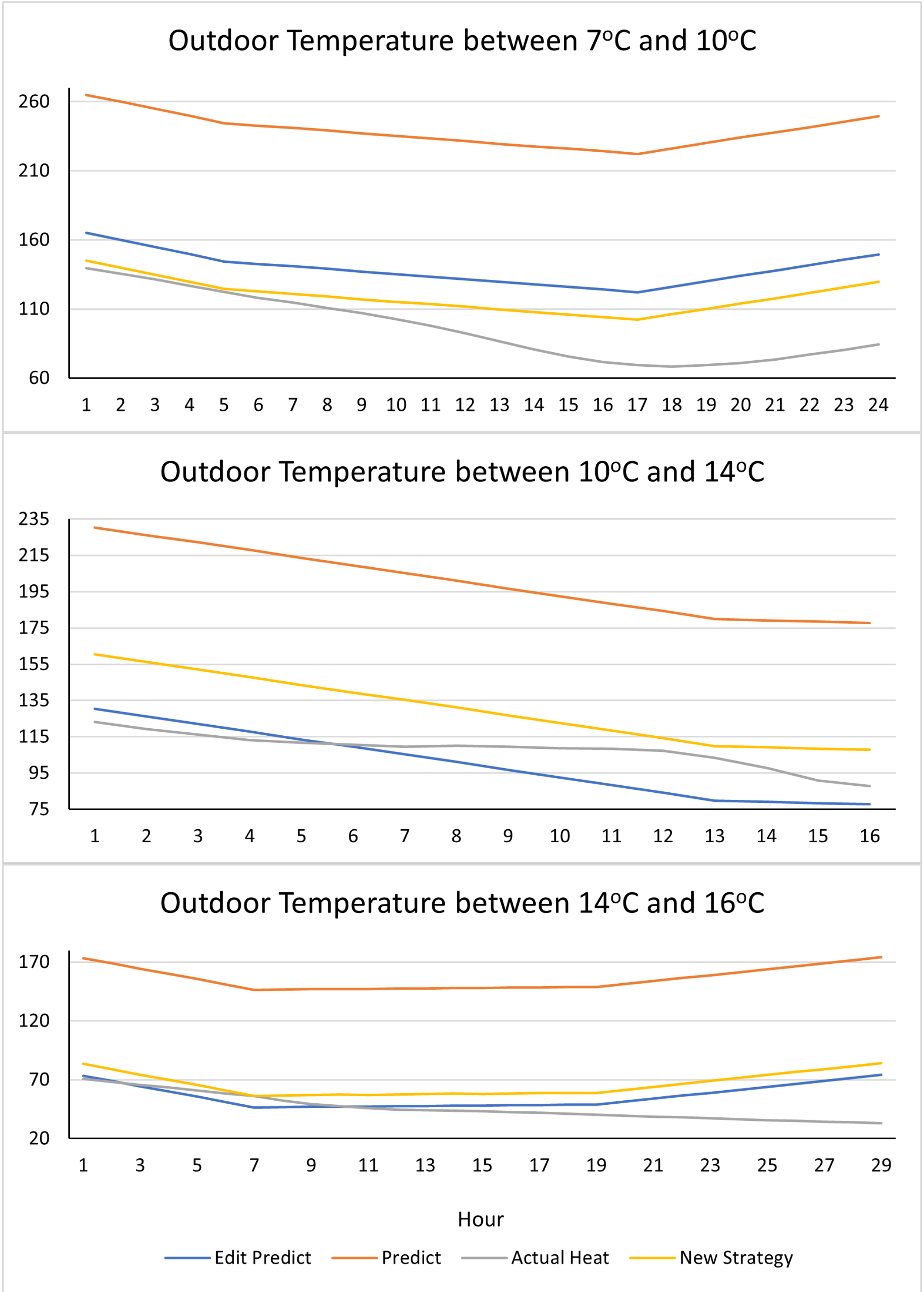


Fig. 4.10 Outdoor Temperature based division output part 2

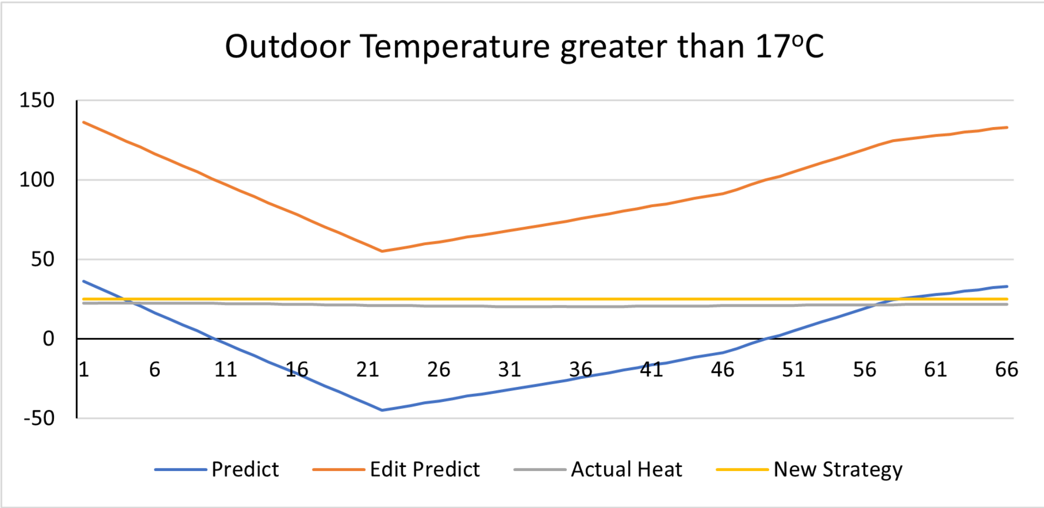


Figure 4-11

Fig. 4.11 Outdoor Temperature based division output part 3

4.5.1 Scenario 1

In the first scenario, the dataset was partitioned into various datasets based on the range of outdoor temperature. Subsequently, each divided dataset was processed through the model building process, and the results obtained were plotted against the corresponding outdoor temperature range. However, the results generated from this approach were not satisfactory. Therefore, this approach was deemed unnecessary for further investigation.

Table 4.4 R-Score for Scenerio1

Outdoor Temperature	R-Score
$<$ -5	18.53%
0-5	31.54%
5-10	23.34%

4.5.2 Scenario 2

In the second scenario, the data is divided into individual home datasets. Each home has a unique set of data, including temperature and heating rate, that can be used to train a linear regression model. The model building process is carried out on each individual home dataset.

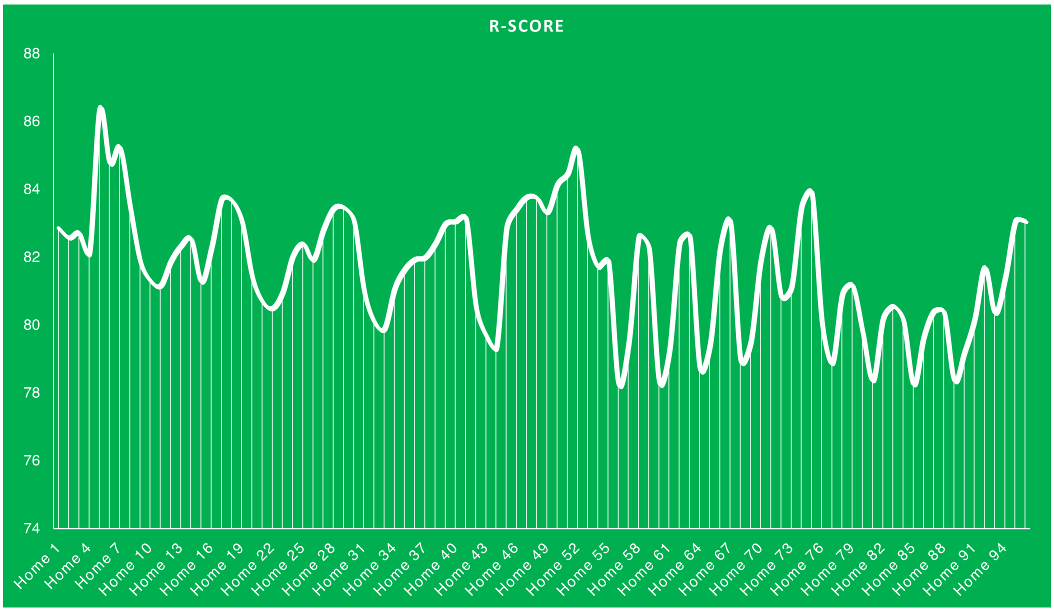


Fig. 4.12 R-Score for Individual Home

Figure 4.12 shows the R-score for each individual home dataset. The R-score is a statistical measure of how well the linear regression model fits the data. A higher R-score indicates a better fit, while a lower R-score indicates a poorer fit. However, in this case, the individual home R-scores do not reflect a consistent score, indicating that this approach is not successful.

To further investigate this approach, one of the predicted values is plotted for Home 1 in Figure 4.13. The line graph shows the heating rate predicted vs the hour of the day. However, the results from this approach are not satisfactory, as the R-scores are inconsistent for each individual home dataset, and therefore, this approach is not successful in achieving the desired accuracy.

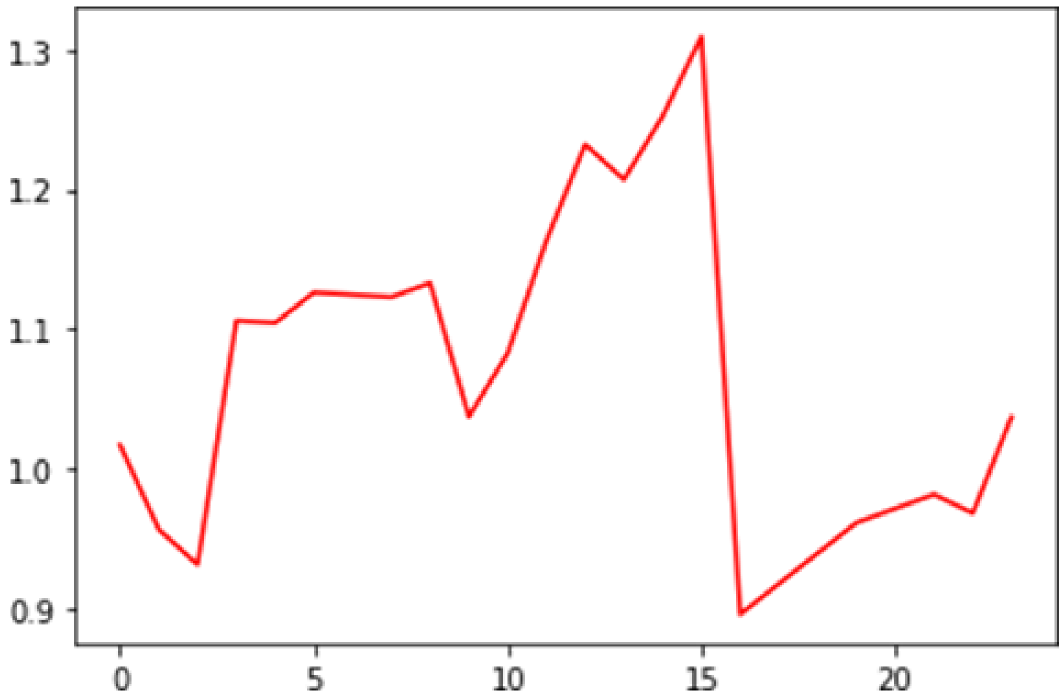


Fig. 4.13 Home 1 heat prediction for a day

4.5.3 Scenario 3

In the third scenario, a different approach was taken where the aggregated heat demand was used to train and test the model. A sample day data was created, and the model was used to generate predictions. The days were sampled based on the coldest, mildest, and hottest day. The output was plotted in Figure 4.14.

In addition to that, other sampled data was generated by using various temperature ranges. Figure 4.15 shows the data plotted based on the 24-hour with different temperature ranges.

A simple control strategy was applied to the data where the heating was switched off at night or when the demand was minimum. The results predicted by the model were plotted in Figure 4.16.

Further processing was done to smoothen the curves and obtain better peaks of the day. The demand was driven based on the time of the day, so it

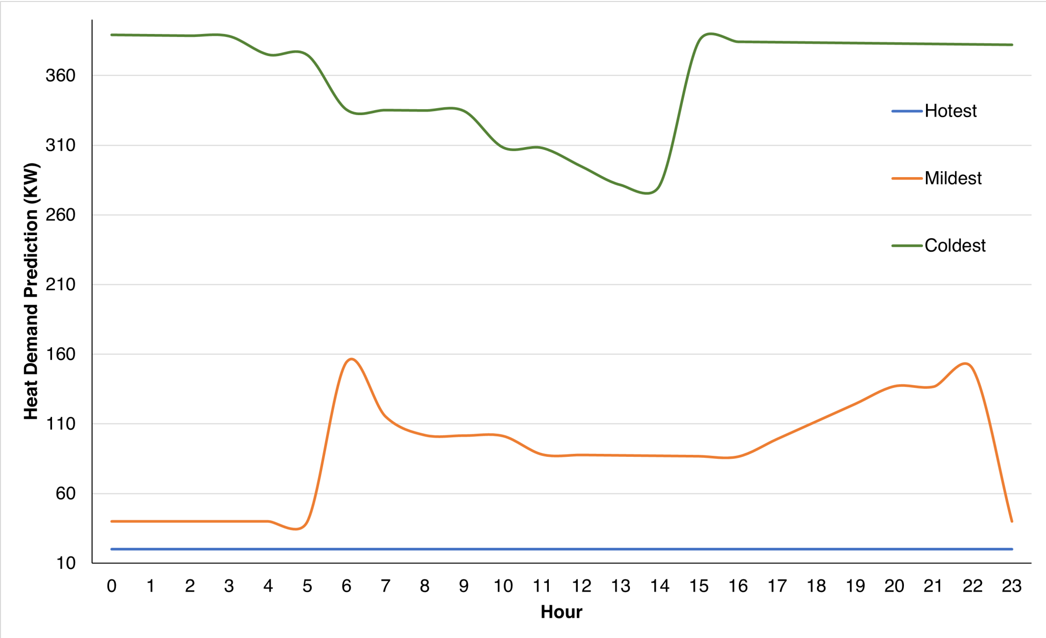


Fig. 4.14 Categorisation of Day

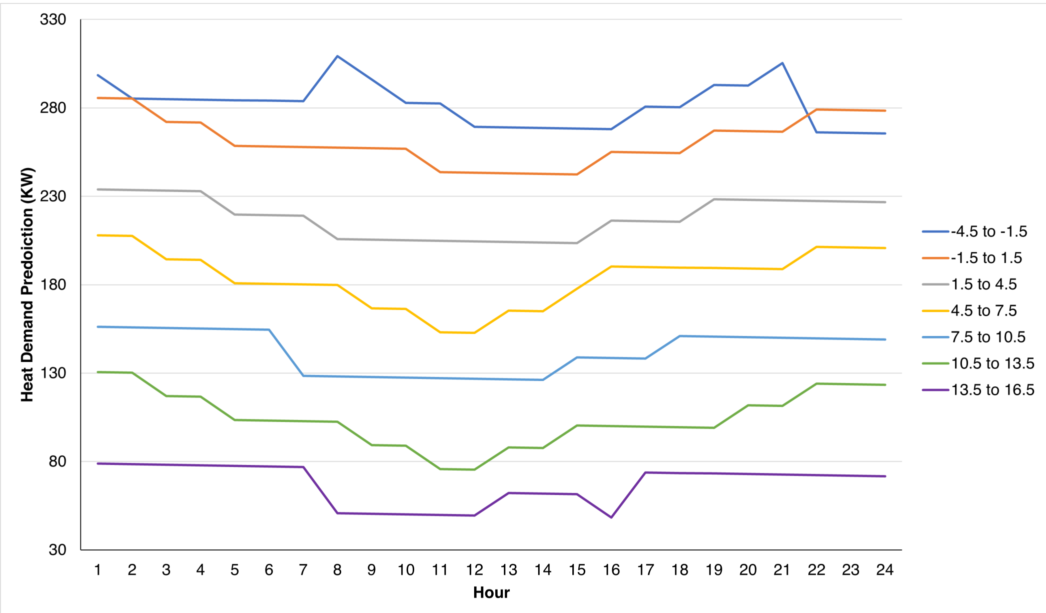


Fig. 4.15 Outdoor Temperature categorising plus division of Day

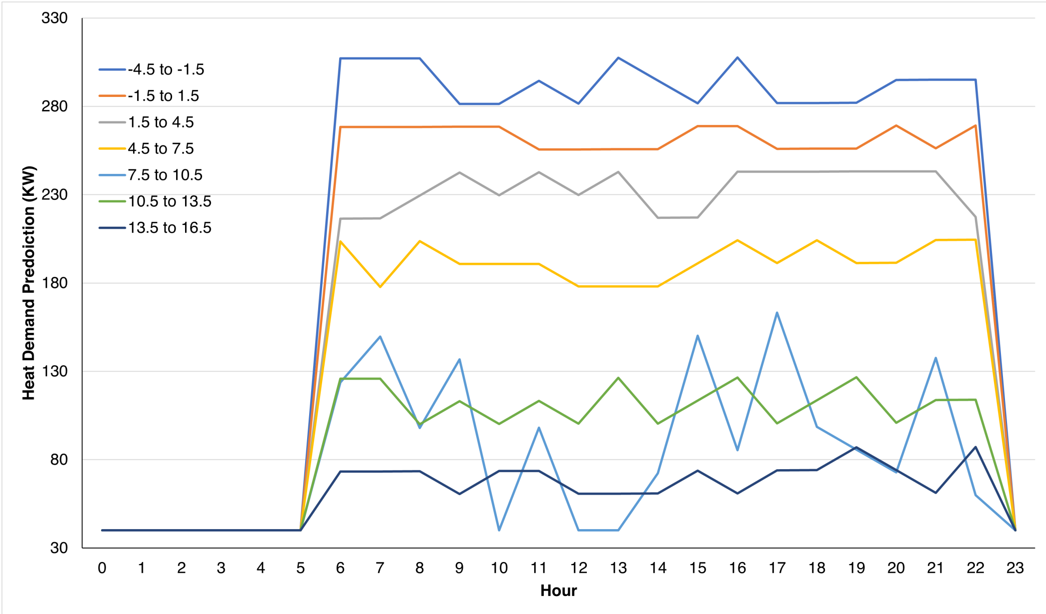


Fig. 4.16 Ridig control Stragegy

was assumed that the demand in the noon and midnight was closer to zero. The output was displayed in Figure 4.17.

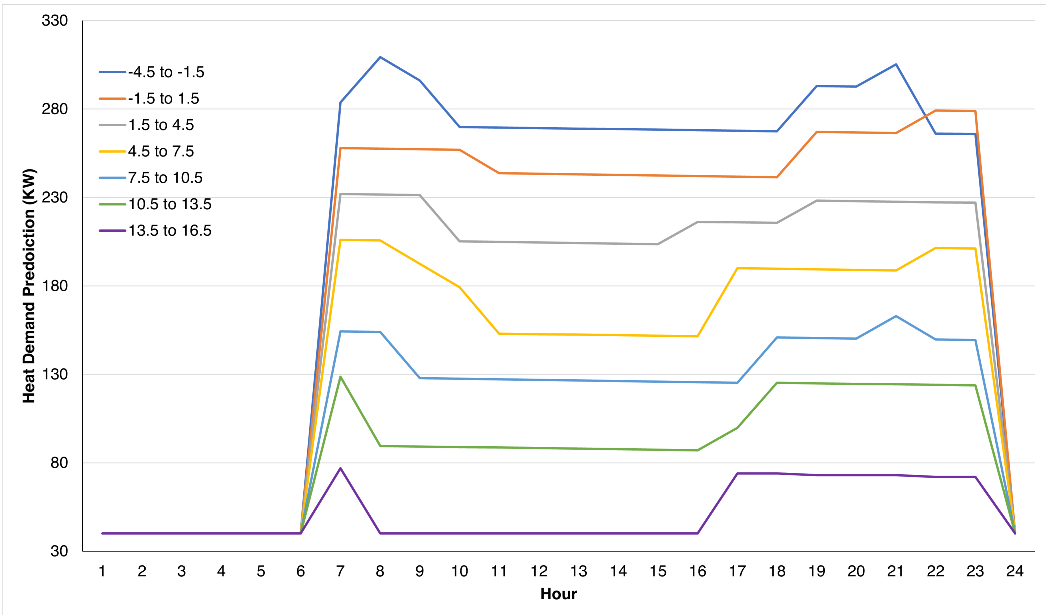


Fig. 4.17 Smoothening Peaks Ridig control Stragegy

4.5.4 Scenario 4

In this scenario, the model is built using the entire dataset without dividing it into training and testing data. After training the model, random daily data is selected with different temperature ranges. The model is used to generate predictions for each day, and the results are plotted. This approach is known as the "whole dataset" approach.

Using the whole dataset for model training and prediction can be beneficial in some cases, as it can provide a more comprehensive understanding of the relationship between the input variables and the output variable. However, it can also lead to overfitting and poor generalization if the model is not carefully designed and evaluated.

Figure 4.18 displays the results obtained using the whole dataset approach. The predicted heating rate is plotted against the actual heating rate for each day in the sample. The results show a good correlation between the predicted and actual values, indicating that the model is performing well.

Overall, the whole dataset approach can be a viable option for building predictive models, but it is essential to carefully evaluate the model's performance and avoid overfitting.

4.5.5 Scenario 5

In this scenario, a new dataset is used to generate predictions (Table 4.2 Simulated Second Dataset). To compare the predicted data with the actual data, a line chart is plotted in Figure 4.19. Additionally, line charts of individual datasets with predicted and actual data are also plotted for better understanding. The datasets used in this scenario are pre-retrofit, post-retrofit, and extra data generated using a simulator.

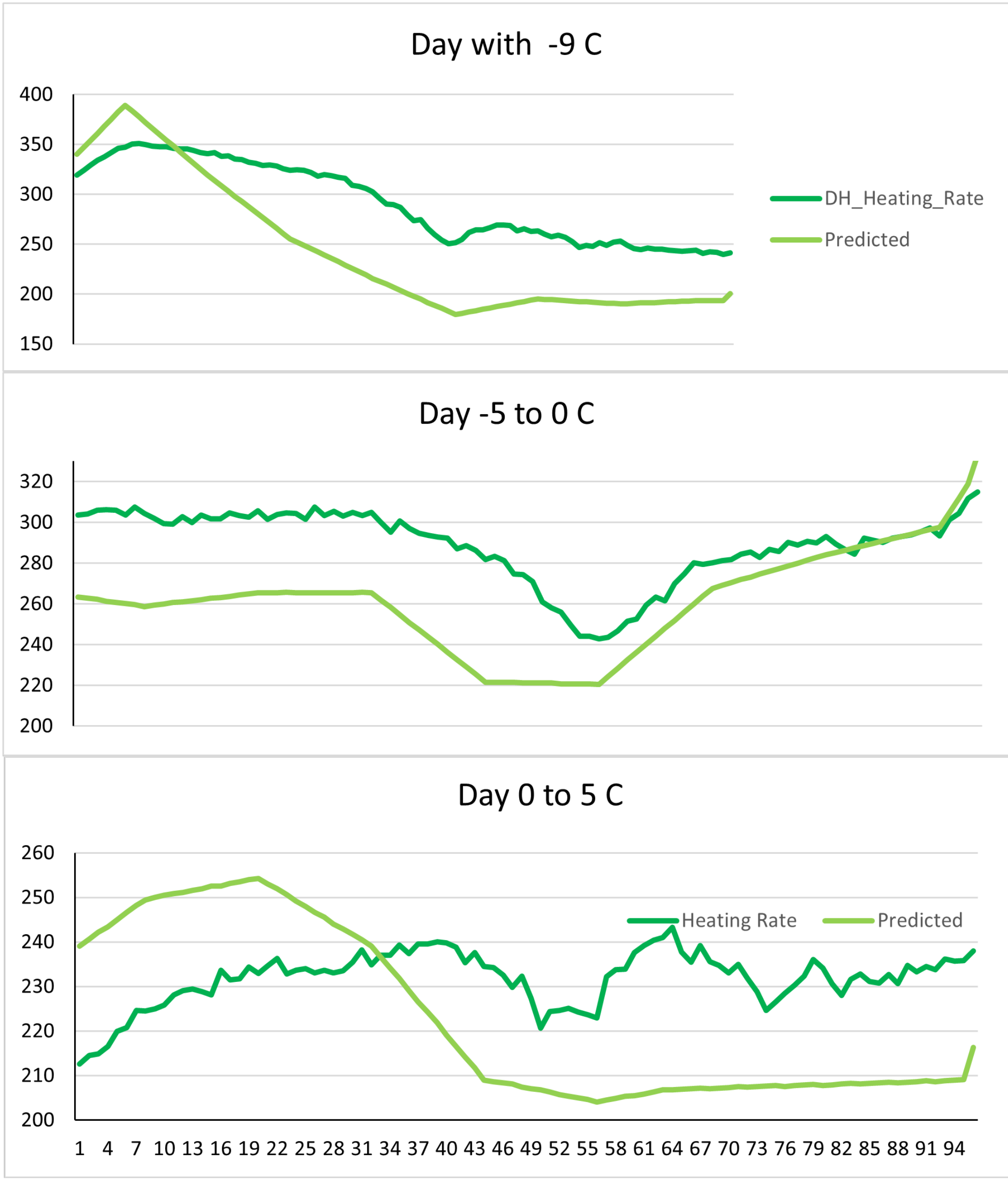


Fig. 4.18 Whole Data set modelling

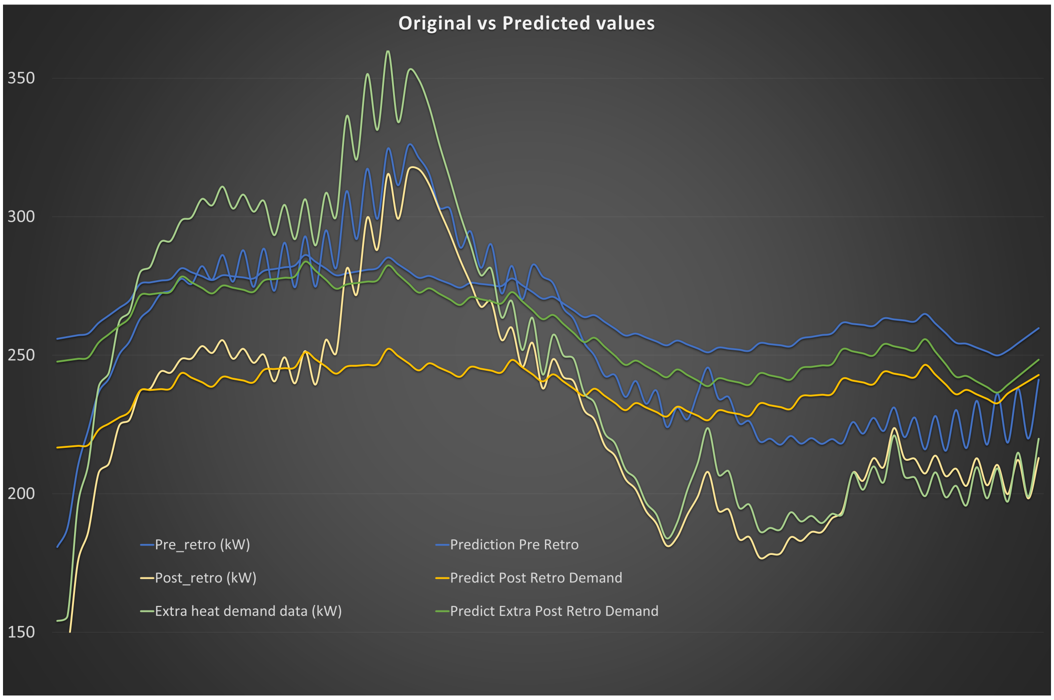


Fig. 4.19 Second dataset Analysis

4.6 Conclusion of Scenario

Based on the analysis presented in the scenarios, it can be concluded that linear regression is a suitable method for predicting heat demand. However, the accuracy can be improved by using individual home data in scenario 2 and processing the data in scenario 3 to show the expected peaks in the demand curve.

In scenario 4, it was observed that the predicted heat demand was always below the actual data. This suggests that there may be factors affecting the demand that are not captured by the model. By shifting the curve on top, it may be possible to improve the accuracy of the model.

In scenario 5, it was demonstrated that by applying some constraints and processing the data, the post-retrofit and extra heat data can be used to improve the accuracy of the model.

Overall, the scenarios presented in this analysis provide insights into how linear regression can be used to predict heat demand and the factors that can affect the accuracy of the model.

4.7 Forecasted data

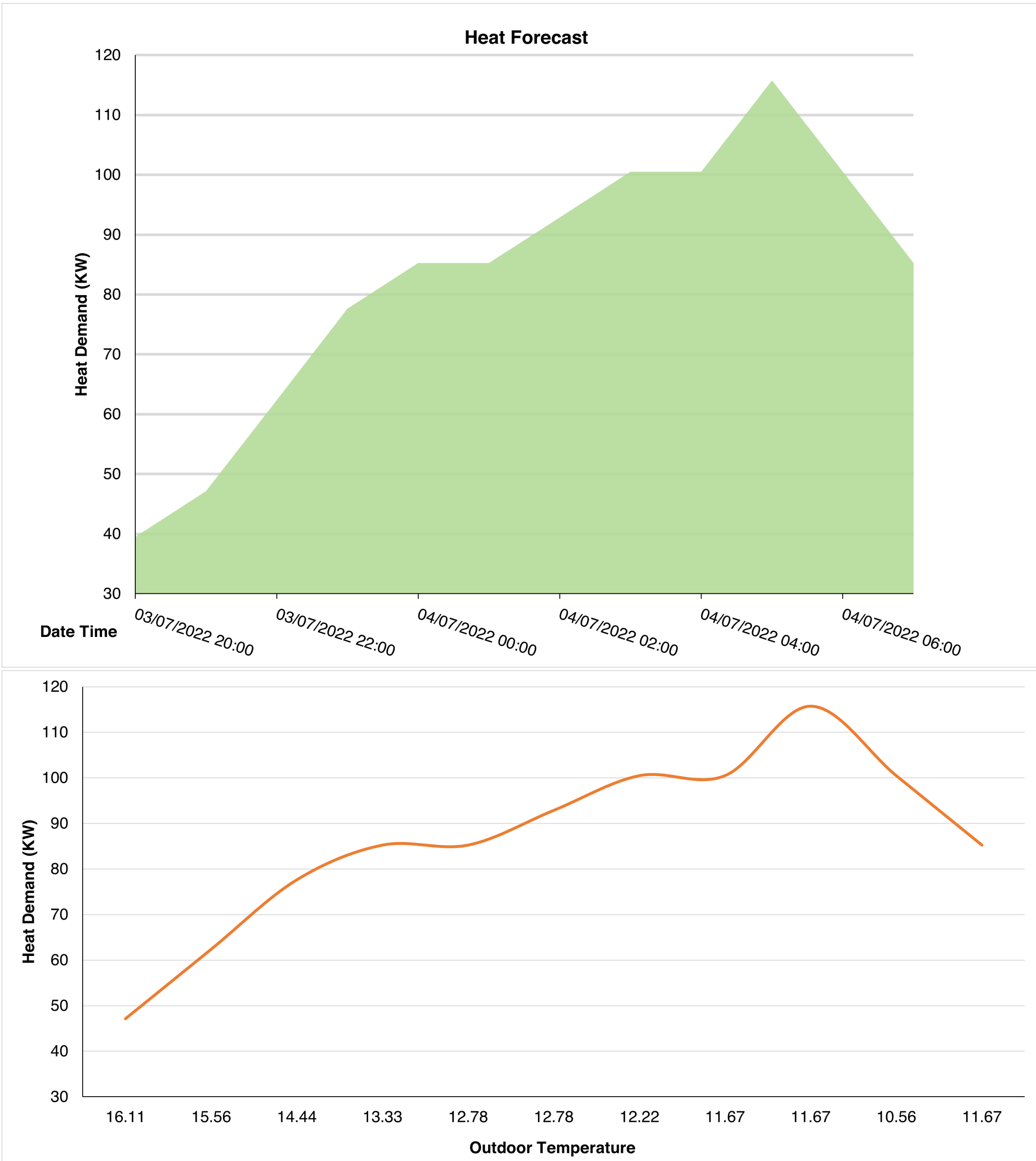


Fig. 4.20 Prediction

Once the model is trained which in this project is linear regression. The weather API is called. The weather API used is Accuweather. The temperature of the next 12 hours is called from API than passed to the model. The output from the model is the forecasted data. The sample chart is shown below (Figure 4.20).

4.8 Sustainable Energy Management System (SEMS)

This section presents the architecture and components of a multi-energy system optimisation software that has been developed by a collaborative team from Siemens, the University of Nottingham, Imperial College London, and Nottingham Trent University. This software is an extension of the work by [194, 13], and addresses the challenge of scalability and transferability through the use of a Hypernetworks Theory (HT) based approach to choreograph the various modular data collection, forecasting, and optimisation components, thereby ensuring consistent orchestration.

The software represents a novel solution to the complex problem of optimising multi-energy systems, which involves coordinating the generation, storage, and consumption of multiple energy sources, as well as their associated infrastructures, such as electrical grids, thermal networks, and gas networks. The HT-based approach is particularly suited to this task, as it enables the modular components to be designed independently of each other, and then integrated in a flexible and scalable way.

In addition to describing the software architecture, this chapter also provides details of the test site where the software is being applied, along with a summary of the results from two microservice predictive algorithms. The test site is a real-world campus that has been equipped with a variety of renewable energy sources, such as solar panels, wind turbines, and biomass boilers, as well as a battery storage system and a microgrid controller. The microservice predictive algorithms are used to forecast the energy demand and supply of the campus, and to optimise the operation of the energy systems in real-time.

Overall, this software represents a significant contribution to the field of multi-energy system optimisation, and has the potential to improve the efficiency, reliability, and sustainability of energy systems on a large scale. The detailed description of the software architecture and the results of the test site demonstrate the practical application of the HT-based approach, and provide valuable insights for researchers and practitioners in the field.

4.9 Test site

The control methodology was specifically designed for testing at the Ernest Dence estate in Greenwich, London, which serves as a demonstrator for the H2020 Sharing Cities project [195]. The estate consists of three buildings, namely Aylmer House, Gifford House, and Jennings House, which together house a total of 95 social housing apartments ranging from one to four bedrooms.

4.9.1 Energy system

Prior to the retrofit, the space heating and domestic hot water needs of the Ernest Dence estate in Greenwich, London, were met by three centralised gas boilers. The heat was distributed through the use of control valves regulating the flow of water through radiators. Domestic hot water was supplied through individual hot water tanks within each dwelling.

Following the retrofit, an energy centre was installed which includes a groundwater source heat pump (HP) and back-up gas boilers connected to a 10 m3 thermal store. The retrofit also involved building fabric energy efficiency measures such as loft insulation, enabling low-temperature heating for greater comfort. Apartment heat interface units replaced water tanks and control valves to provide direct connection for space heating and instantaneous hot water via flat plate heat exchangers. High-temperature radiators were removed and replaced with those suitable for a flow and return temperature of 60/30 °C. Heat distribution pipes were also rerouted to minimise thermal length and insulated to prevent heat losses.

In the future, the installation of a 100 kWp solar photovoltaic (PV) array and ten 7 kW onsite EV chargers is planned. These additions are also considered in the software optimisation process.

4.9.2 Heat and electricity demand

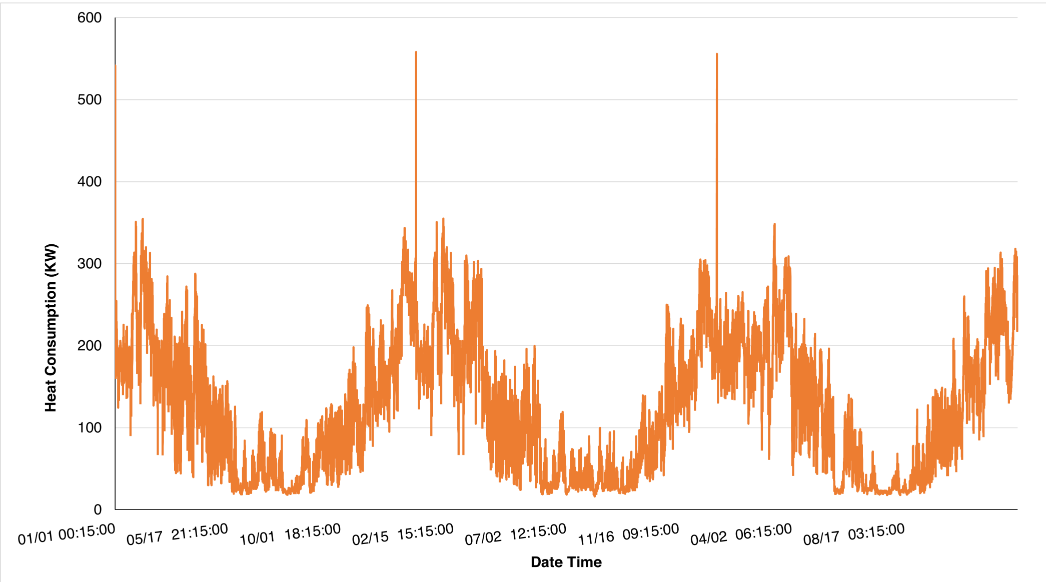


Fig. 4.21 Heat demand profile Ernest Dence estate – Greenwich, London

Figure 4.21 presents the whole-building heat demand data over a period of three years at the Ernest Dence Estate, which was simulated using EnergyPlus (EnergyPlus, 2021) and Ptolemy II [196] digital twin software. This data was utilised to develop heat demand forecasting algorithms that inform the energy system optimisation. Similarly, four years of energy demand data were simulated using a Matlab reconstruction of the CREST demand model

(Loughborough University, 2020) to develop forecasting algorithms that inform the system energy optimisation. Figure 4.22 represents a typical week's building electrical load data for the Ernest Dence Estate, which was also simulated using the CREST demand model.

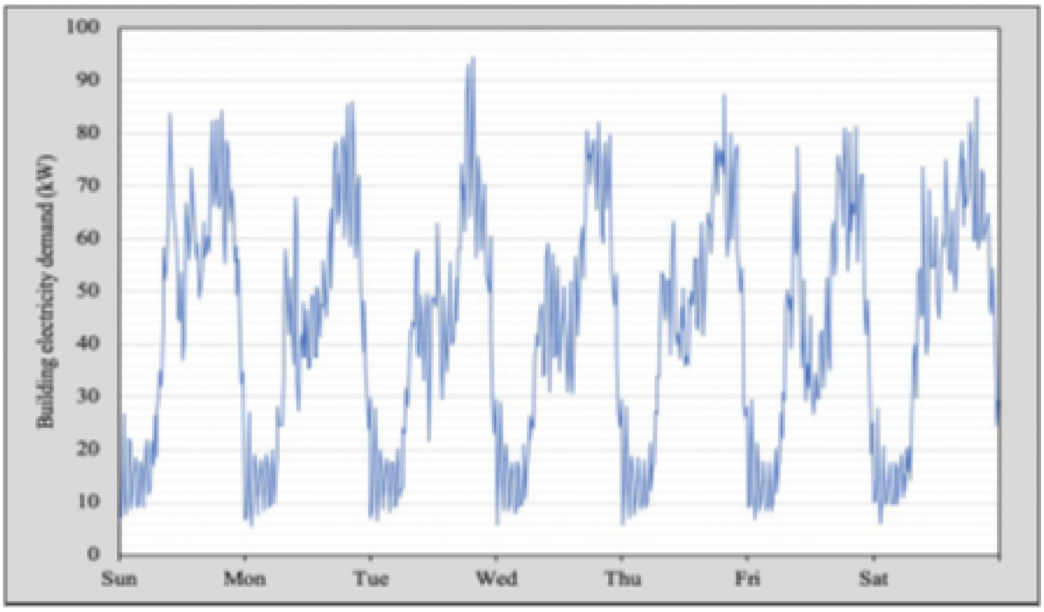


Fig. 4.22 Electricial demand profile Ernest Dence estate – Greenwich, London

4.10 System architecture

All energy assets in the system have direct low-level control internally, as shown in Layer 1 of Figure 4.23. For example, the gas boilers will regulate their own firing and pumps. Medium-level control, with a system view, is managed by a Building Management System (BMS), as seen in Layer 2 of Figure 4.23. On the other hand, high-level control, which takes into account the entire system and external factors such as variable prices, carbon factors, and demand forecasts, is performed by a cloud-hosted SEMS, as illustrated in Layer 3 of Figure 4.23. The SEMS optimises energy usage and suggests how to best operate energy

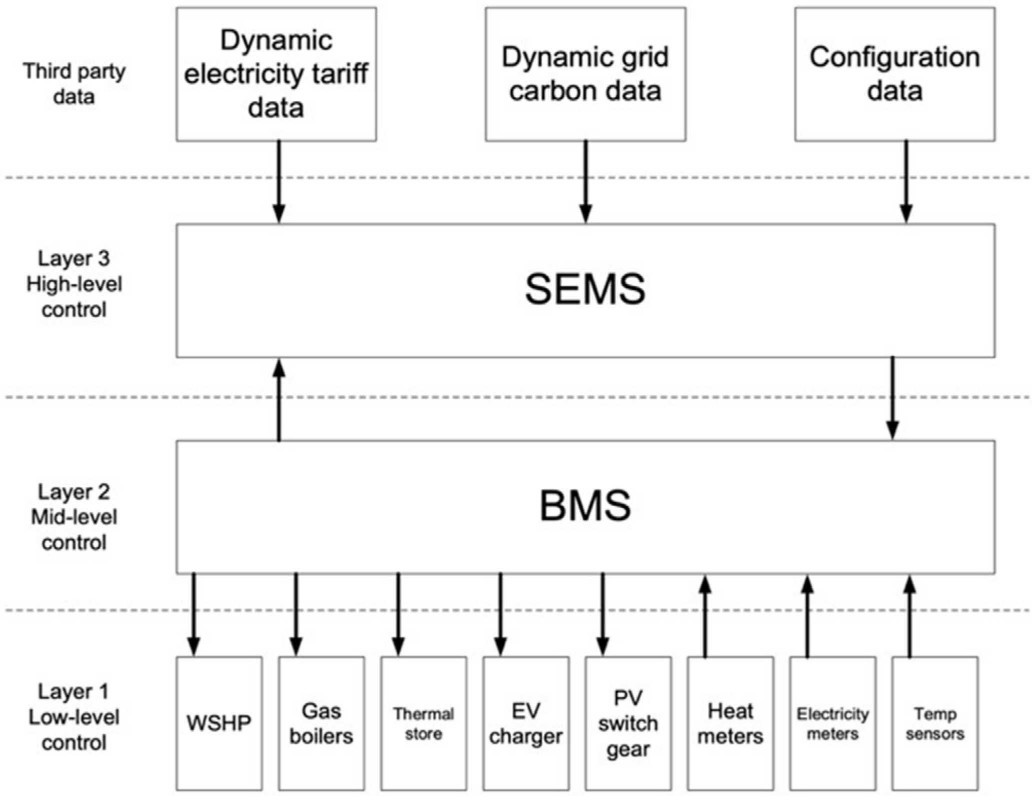


Fig. 4.23 Control layering at the Ernest Dence Estate, Greenwich, London.

assets. All safety-critical scenarios are covered by leaving direct control to local controllers.

4.10.1 Application architecture

The SEMS application architecture is depicted in Figure 4.25, and it consists of four main components: SEMS Engine, services, in-memory database (DB), and web. In addition to these components, there are two forms of configuration: the Hn model and the metadata used to specify the parameters. The Hn model is used to describe the site, its assets (such as PV, EV, HP, etc.), and the traffic required to perform actions such as forecasting and optimisation. The metadata, on the other hand, supports the Hn model by providing the specific parameters for each physical asset or logical vertex used in the optimisation.

Fig. 4.24 Application architecture used for SEMS

Fig. 4.25 SEMS software architecture illustrating HT

Fig. 4.26 SEMS microservices/vertices, simplices and data flows

Fig. 4.27 Relationship Network (Copyright Siemens)

Each vertex in the architecture represents a specific asset on the site and is named accordingly. Additionally, each vertex can include a function that performs the specific action required. This allows the SEMS to reuse common functions by tailoring them to the specifics of each asset. To implement this architecture, each function was implemented as a web service, thus separating the choreography from the data and individual function.

The SEMS engine executes the model periodically by calling each service concurrently based on the Hn model. The first level of vertices executes when the SEMS is started, but vertices at subsequent levels wait for all their parts, specified as a simplex, to complete before executing, as shown by the bar in Figure 4.27. This allows all vertices to execute in parallel, except when they are waiting for all parts to complete. Once all parts have completed, the waiting vertex collects the data produced by the parts from the in-memory DB where it is stored. Upon completion, the vertex stores its data in the in-memory DB, releases, and returns a completed status back to the SEMS engine, allowing it to continue with the next levels until there are no more. This architecture provides an efficient and scalable way of performing energy optimisation and forecasting, with the ability to reuse common functions across different assets.

4.10.2 Microservice structure

The Sharing Cities project has developed a multi-energy system software consisting of a multitude of microservices or vertices, as depicted in Figure 4.26. These services are responsible for acquiring and integrating data, generating predictive forecasts, executing optimisations, or interpreting control signals to achieve the optimal system operation. The software's compartmentalisation of subtasks (see Figure 4.26) offers greater flexibility to modify, remove, or add functionalities, thereby enhancing its scalability. The distributed architecture of the software facilitates easy testing and more flexible execution, making it a powerful tool for managing complex energy systems.

4.10.3 Optimisation & Interpretation

The developed multi-energy system software utilises the Pyomo package for the formulation, solution, and analysis of a structured optimisation model, with the GLPK [197] and Gurobi [198] solvers as stand-alone options. The primary objective of the model is to minimise scaled, weighted values of total operational greenhouse gas emissions and/or cost, while adhering to constraints based on equipment operating limits and heat-power system balancing. An additional constraint ensures that substation capacity limits are not exceeded, thus avoiding the need for costly upgrades to electricity grid infrastructure.

The outputs of the optimisation model consist of 24-hour profiles for energy asset power inputs and/or outputs, as well as energy storage levels. These profiles are subsequently translated into control signals for the BMS (see Figure 4.26) to schedule equipment and execute storage charge/discharge/bypass commands using setpoints. To promote software flexibility and ease of deployment, the software components, including the SEMS engine, services, and in-memory

DB (as depicted in Figure 4.24), are containerised and deployed using Docker on an AWS EC2 server.

4.11 Result

4.11.1 Heat demand forecast

The process of forecasting heat demand is a crucial step in the development of an efficient and optimised energy system. A variety of algorithms are available to perform this task, including artificial neural networks and support vector machines, but the accuracy of these methods can vary. After careful testing, it was found that a multivariate regression model was the most effective method for forecasting heat demand in the specific context of the Sharing Cities project.

Although multivariate regression is generally regarded as the more effective method, for the purpose of this research, linear regression is chosen instead. While multivariate regression offers greater analytical power and flexibility by considering multiple independent variables simultaneously, linear regression is deemed more appropriate due to its simplicity and better alignment with the research objectives. Although linear regression may have certain limitations, such as assuming linearity and independence of variables, it is still deemed suitable for addressing the research questions at hand. In addition, it should be noted that the difference in the R2 score between multivariate regression and linear regression is only 2%. Despite this small difference, linear regression is still chosen for this research due to its practicality and ease of interpretation, while still providing a reasonable level of accuracy in capturing the relationship between the variables.

The regression model achieved an R2 score of 0.80 - 0.85, which indicates that the predicted heat demand values were close to the actual values. Ad-

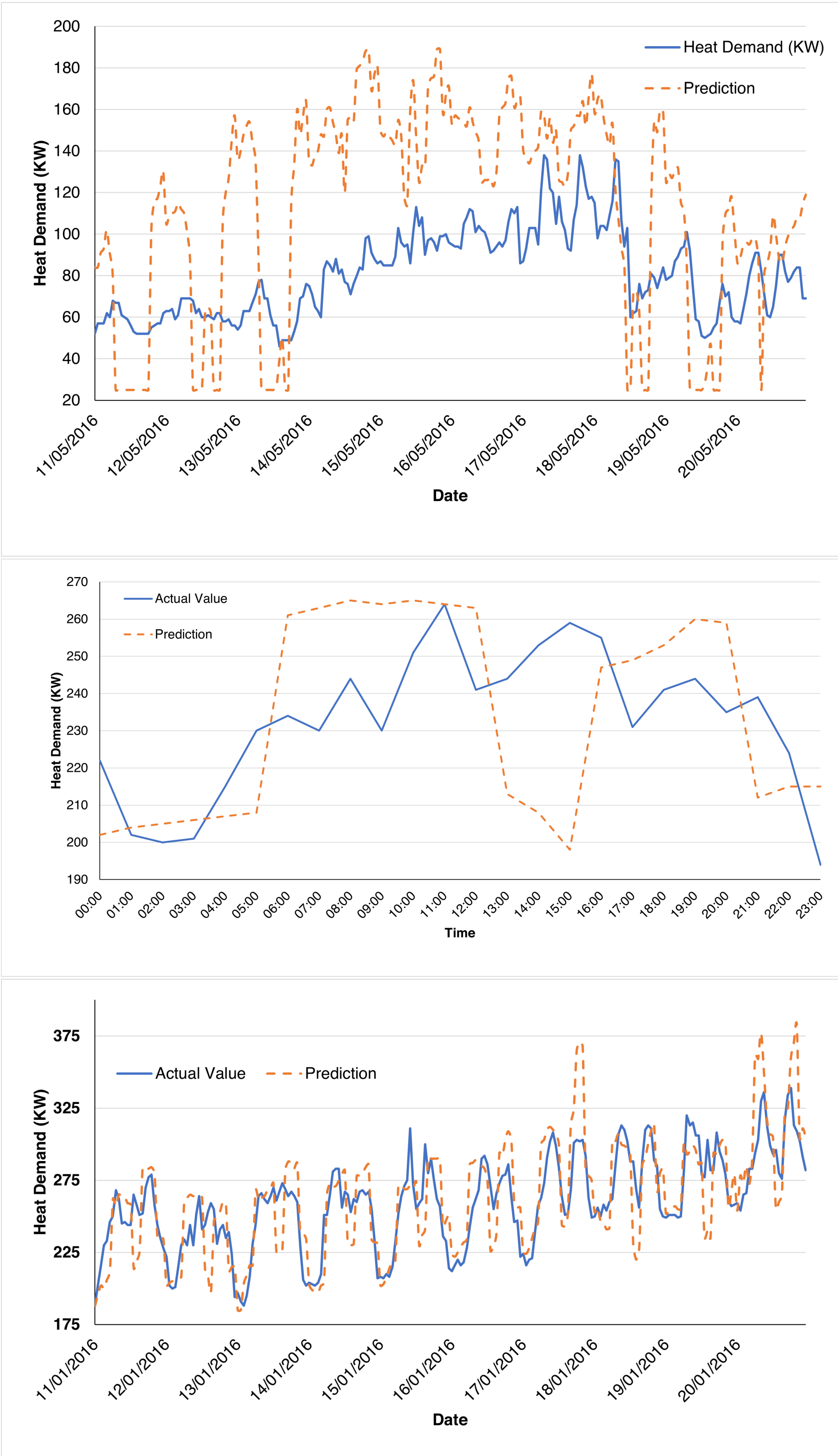


Fig. 4.28 Example heat demand forecast values (red dashed) vs actual demand / validation dataset (blue) for 10 days in winter

ditionally, the Root Mean Square Error (RMSE) of the model was found to be 33.61, meaning that the predicted values differed from the actual values by approximately 34 kW. This level of accuracy is important for ensuring that energy resources are used efficiently and effectively, as it allows energy system operators to better anticipate the heat demand and allocate resources accordingly (Figure 4.28).

To develop and validate the model, a three-year dataset was used, which was divided into two years for training and one year for validation. This process allowed for the model to be refined and optimised, ensuring that it was as accurate and effective as possible in predicting heat demand in the context of the Sharing Cities project. Overall, the use of a multivariate regression model for heat demand forecasting represents a key element in the successful development of an efficient and optimised energy system.

4.11.2 Electricity demand forecast

The forecasting of whole building electricity demand is a crucial component of the SEMS developed as part of the Sharing Cities project. To achieve accurate and reliable predictions, a Seasonal Autoregressive Integrated Moving Average (SARIMA) model was selected. This model is a popular choice for forecasting univariate time series data with trend and seasonality [199, 94]. The SARIMA model was trained on three years of historical data and then validated against a further year of data to ensure its accuracy.

The accuracy of the SARIMA model was evaluated using the Root Mean Square Error (RMSE), which is a standard measure of the difference between the predicted and actual values. The model achieved an RMSE of 10.652, indicating that on average, the model was wrong by approximately 10.7 kW for each prediction made. While this may seem significant, it is considered a reasonable error given the complexity of electricity demand forecasting and can be accounted for in the energy optimisation logic.

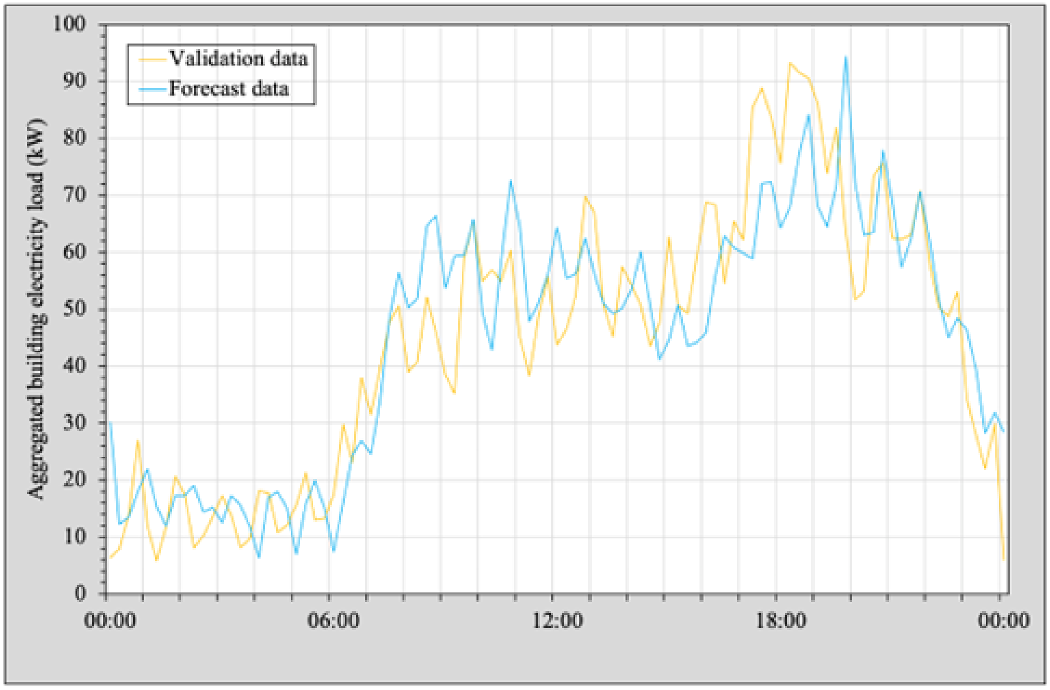


Fig. 4.29 Example aggregated domestic electricity demand profile for a typical summers day

To provide a visual representation of the model's performance, Figure 4.29 displays typical winter day demand profiles for both the validation and forecast data. As can be seen from the figure, both datasets are relatively well matched, demonstrating the consistency of daily residential electrical load trends and seasonality. Overall, the SARIMA model proved to be an effective solution for forecasting whole building electricity demand in the SEMS, providing a valuable tool for efficient energy management and optimisation.

4.12 Discussion

This study details the creation of a minimum viable software product that was specifically designed for implementation at the Ernest Dence Estate in Greenwich, London. This estate was selected as a case study for a system that requires high-level control due to the complexities of interlinking energy vectors and the associated constraints. The software has been developed to offer several commercial and environmental benefits, including the reduction of operational emissions and costs, optimal energy balancing, and the prevention or deferral of reinforcement costs of network infrastructure. While the initial focus of the software design was on a heat and power system, the project specifications prioritised scalability, flexibility, and replicability to ensure that future software applications and developments are not limited.

The Hn offers a straightforward means to establish a hierarchical structure that organises inputs and parameters into discrete calculations with corresponding outputs. These calculations can be used for subsequent levels of analysis or as constituent parts of various other calculations. The microservices architecture allows for a clear separation of responsibilities among the controller, data, and calculator components, while simultaneously providing scalable functional units that are distinct and self-contained.

There is an overarching objective to apply the SEMS framework to a range of diverse energy optimisation problems. For instance, one potential application involves utilising SEMS to manage an eBus smart charging depot, which entails considering battery capacity, real-time bus route tracking, and scheduling. Additional case studies and algorithm development will be undertaken on a project-by-project basis. Nonetheless, efforts to enhance the existing 22 services in the current version are also underway. Specifically, for the two case studies featured in this chapter — heat and electricity demand forecasting — alternative machine learning algorithms, such as Long Short-Term Memory, will be evaluated in order to improve accuracy and adaptability.

The present design establishes a solid foundation upon which to expand the capability and versatility of the SEMS software. Further efforts can be focused on diversifying its functions and applications. A forthcoming paper will detail the outcomes of integrating the software with the physical system and evaluating its operational performance at the Ernest Dence estate.

4.13 SEMS Algorithms

Final algorithms used in the SEMS project is listed in the Table 4.5.

Table 4.5 Summary of SEMS algorithms

Module	Algorithm
Building Heat Forecast	Linear regression model (temperature)
Building Electric Forecast	Seasonal autoregressive integrated moving average (SARMINA)
EV Forecast	Timeseries event probability distribution
PV Forecast	Power generation equation
Ground Temperature Forecast	Autoregressive moving average (ARMA)
Stored thermal energy	Heat equation
COP Forecast	Search COP matrix based on inlet and outlet temperature
HP Power Forecast	Power = heat demand / Coefficient of Performance
Optimisation	Pyomo ‘gplk’ solver

4.14 Summary of Chapter

In chapter 4 of the report, the procedure for building the model is explained, which involves reading the data from a CSV file, extracting date, time, day, month, hour, and minute, and dividing the data into an 80-20 ratio for training and testing the model. Various machine learning models were tested, including K-Means Clustering, K-Nearest Neighbour, Decision Tree, Neural Network, and Linear Regression. The results showed that Linear Regression was the most effective algorithm among the three tested, due to the fact that the dataset contained only two variables: outdoor temperature and desired DH. The K-Means Clustering model was trained with 5 clusters, while the KNN algorithm produced inaccurate results due to the overlapping of all the clusters. The Decision Tree algorithm was not able to capture the relationship between the input variables and the output variable. The Neural Network model did not perform well, even after trying different combinations of hyperparameters such as hidden layer sizes and activation functions. The report suggests that further analysis and refinement of the model may be needed to improve its accuracy and reduce the impact of bias.

The chapter discusses an optimisation process for energy centers that considers both energy prices and CO2 emissions. The process involves switching between a boiler and heat pump for heat generation using pynomo in a virtual environment. The optimisation is based on different scenarios, including fuel price, carbon emissions, and trade-off between price and emissions. The goal is to find the best combination of energy source and heat generation method that achieves a balance between cost and emissions. The results show that effective optimisation can be achieved if heat prediction works well. The switching of heating assets on different types of days, where the heat pump and thermal storage provide the best COP for optimising in terms of emissions and prices.

Chapter 5

A Comparative Analysis of Heat Demand Prediction Using Real-Life Data Model and Simulated Data Model

5.1 Introduction

The District Heating (DH) system is a critical component in achieving the goal of reducing carbon emissions to a low level by 2050. In order to facilitate its expansion, the UK government and private sector have invested GBP 1351M [200]. However, it is important to note that as of 2018, only 2000 DH networks were operational, despite there being a recorded 14,000 heat networks in the UK. Additionally, the estimated annual turnover generated from heat networks is approximately 300 million pounds. Despite this, there is an expectation that the number of heat networks will increase significantly over the next ten years [201].

An overview of the DH system reveals that significant investment is being made towards research and development to enhance its efficiency and sustainability. By 2050, the UK aims to connect half of its buildings to District Heating Networks. Moreover, a cost evaluation of DH versus traditional heating indicates a saving of 4 pence per kWh per customer. The use of DH has demonstrated a reduction in CO2 emissions from 240 kg/MWh to 60 kg/MWh [202, 203]. With DH's potential to grow over the next 30 years, the scope for technological advancements is significant.

This chapter focuses on the 4th Generation Low Temperature District Heating (LTDH), which is the most efficient DH system. For precise control of heat generation, it is essential to have accurate and realistic heat demand predictions. LTDH systems can achieve a higher degree of independence by utilising renewable energy sources and heat storage. Due to the lower feeding temperature in LTDH, sources of heat supply can be diversified by using renewables and recycling of local excess heat. The application of LTDH will require realistic and more accurate designs with respect to heat demand, costs, and operating conditions. It will also require heating systems with low temperature demands and no short-circuit flows in the distribution networks.

The use of technology has facilitated the application of various machine learning algorithms for heat demand prediction, as demonstrated in various literature sources. Three of the most commonly used machine learning algorithms are Artificial Neural Network, Decision Tree, and Linear Regression. This paper compares the performance of several machine learning algorithms applied to two different datasets [20].

In summary, the DH system plays a crucial role in achieving low carbon emissions in the UK. With the government and private sector investing heavily in its development, the installation of more DH networks is expected. The LTDH system is the most efficient DH system and can achieve a higher degree of independence through the use of renewable energy sources and heat storage. By accurately predicting heat demand, the application of LTDH can be optimised for greater efficiency and sustainability. The use of machine learning algorithms for heat demand prediction is a promising approach that warrants further research.

5.2 Datasets

This section discusses the datasets from two distinct projects: the SHARING CITIES project and the REMOURBAN project, both of which collected data from residential homes. The SHARING CITIES project, which aims to develop open-source smart city solutions, generated data by utilising EnergyPlus and Ptolemy II energy simulation software. The data were obtained based on building characteristics and environmental variables and collected from 95 homes located in Ernest State, London. The data generated by the simulators consisted of three variables: timestamp, outdoor temperature, and heat demand for each home. The data were collected over a period of three years at a frequency of 15 minutes.

The author compared two datasets, which were converted to a frequency of 15 minutes for the purpose of comparison. Aggregate heat demand was used for prediction in both datasets. The first dataset was generated using simulators in the SHARING CITIES project, which aimed to develop open-source smart city solutions. Specifically, the data was generated by using EnergyPlus and Ptolemy II energy simulation software based on building characteristics and environmental variables. The simulated data was obtained from 95 homes located in Ernest State, London, and consists of three variables: timestamp, outdoor temperature, and heat demand for each home. The data was generated over a three-year period at a frequency of 15 minutes.

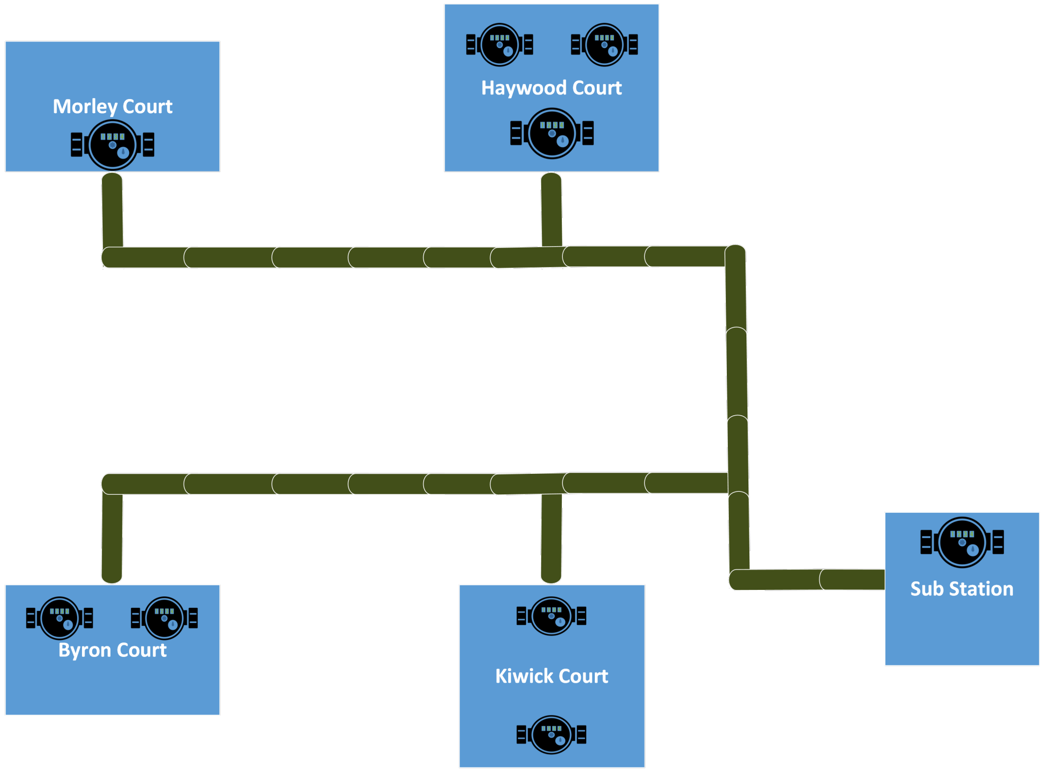


Fig. 5.1 HeatMeter location of REMOURBAN Project

In contrast, the second dataset was collected in real-life settings by the REMOURBAN project. This project aimed to leverage Information Communication Technology, Energy, Society, Mobility, and Sustainability. The REMOURBAN project collected data from 30 homes in Sneinton, Nottingham, UK, for a duration of one year at a frequency of one minute. A weather station was installed in parallel with the data acquisition process, and weather data were also collected for the same duration at a frequency of 5 minutes. The heat demand data in this dataset was obtained from the main meter and submeters, and was aggregated to the 15-minute frequency used in this thesis.

Both datasets show a high correlation with outdoor temperature, which is a key factor in determining heat demand. The scatter plots of the overall datasets, as shown in Figure 5.2 (for Dataset 1) and Figure 5.3 (for Dataset 2), confirm this relationship between outdoor temperature and heat demand.

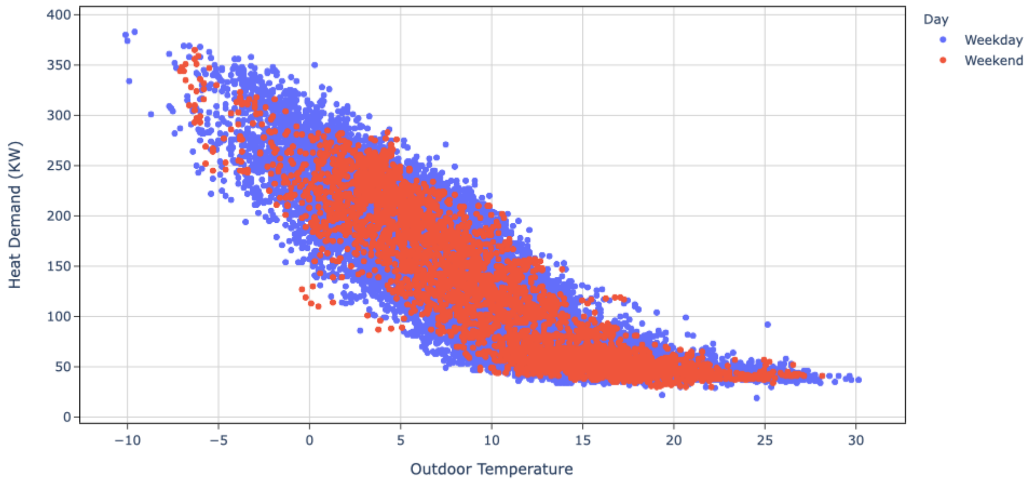


Fig. 5.2 Visualising the Relationship between Simulated Heat Demand and Outdoor Temperature

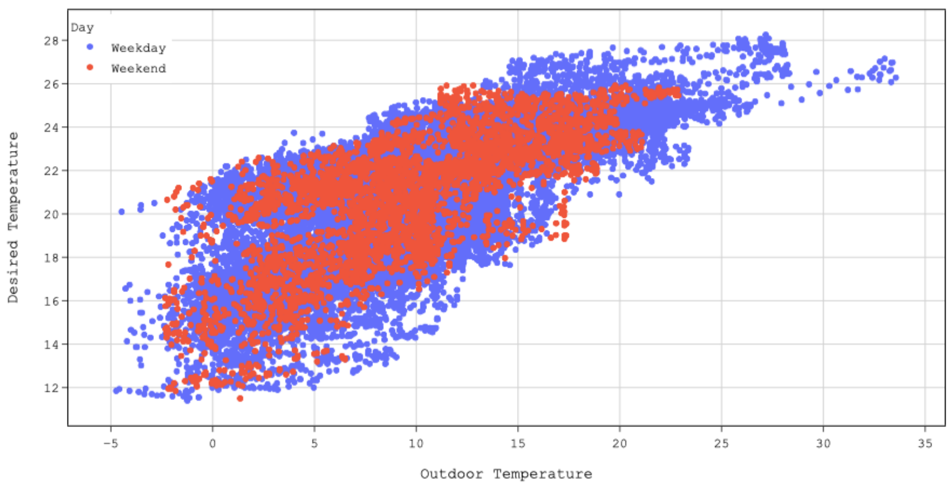


Fig. 5.3 Visualization of the Relationship between Desired Indoor Temperature and Outdoor Temperature using Collected Data

Overall, this comparison of the two datasets sets the stage for the machine learning algorithms that were applied to the data to predict heat demand in the context of 4th Generation Low Temperature District Heating (LTDH).

Table 5.1 Key Features for Building a Heat Demand Prediction Model

Project	Feature	Value Units
REMOURBAN and SHARING CITIES Project	Hour	0 to 23
	Minute	15,30,45,60
	Day	1 to 31
	Weekday	Monday to Sunday
	Month	1 to 12
	Outdoor Temperature	- to + in degrees
	Heat Demand	KW
REMOURBAN	Pressure	hPa
	Humidity	0 to 100 %
	Wind Speed	m/s
	Sunset time	In seconds
	Sunrise time	In seconds
	Weather Type	Clear, Clouds, Drizzle, Fog, Haze, Mist, Rain, Snow, Thunderstorm, Tornado

Prior to delving into the analysis section, it is crucial to comprehend the features utilised for constructing heat demand prediction models. One significant feature that is transformed from a single value to multiple values is the timestamp. The timestamp is converted to the hour, minute, day, weekday, weekend, and month. In addition to the timestamp, the SHARING CITIES project comprises outdoor temperature and heat demand as its features. On the other hand, the REMOURBAN project encompasses features such as temperature, humidity, pressure, sunset time, sunrise time, weather type, and wind speed, in addition to the timestamp. The features and their respective values are presented in a summarised manner in Table 5.1.

5.3 Analysis

This section outlines the analysis of the collected data, as well as its practical application. Once the data is obtained or generated, the first step involves data cleaning and removal of outliers in both datasets. The simulated dataset, covering a period of three years, is then divided into two separate datasets: the first two years are used for training the models, and the remaining one-year dataset is used for testing.

For the real-life dataset (REMOURBAN project), the weather data and data collected from the monitored residential homes are merged using timestamp, and randomly split into 2/3 data for training models and 1/3 data for testing. Three models, including Linear Regression, Decision Tree, and Artificial Neural Network, are then defined and compared. The configuration of all models is kept the same, including the hidden layer, solver, iteration, leaf size, and other parameters.

Before training the models, a correlation matrix is used for feature selection to avoid overfitting or underfitting of the data. The training dataset is then used to build the models, while the test dataset is used to predict the value. In evaluating the models, their performance is measured based on mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R^2) on the testing dataset. The model with the lowest MAE and RMSE and the highest R^2 is considered the best model.

Furthermore, feature importance is analysed for each model to determine the significance of each feature in predicting heat demand. This helps identify the most critical features for predicting heat demand.

Finally, the analysis section also includes a comparison of the performance of the models trained on the simulated dataset and the real-life dataset. This provides insight into the effectiveness of simulation models in predicting heat demand compared to models trained on real-life data.

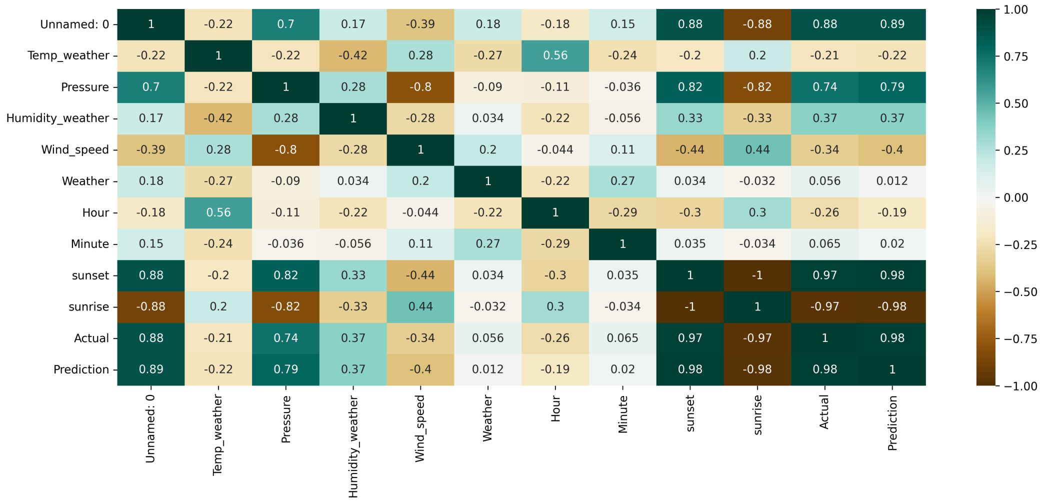


Fig. 5.4 Visualizing the Main Meter Correlation Matrix

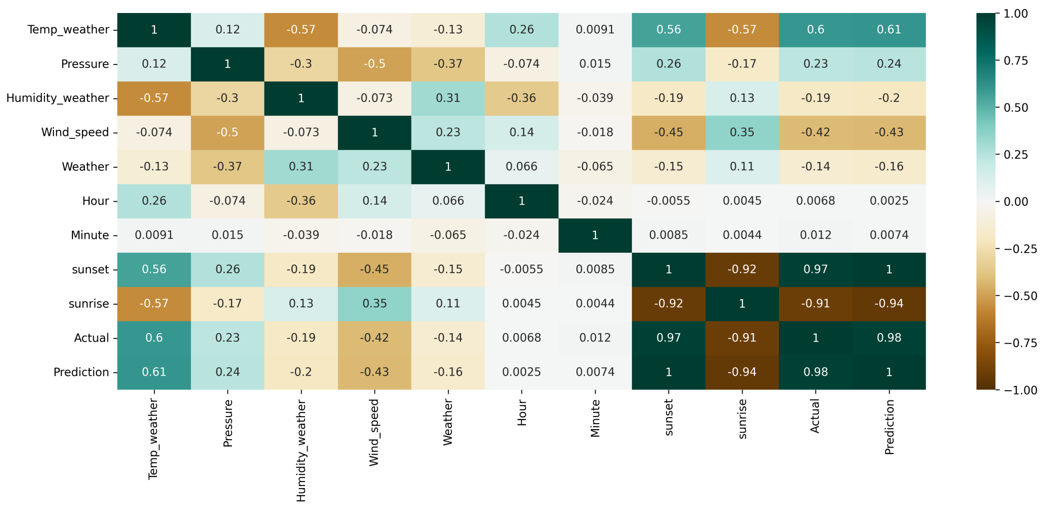


Fig. 5.5 Visualizing the Sub Meter Correlation Matrix

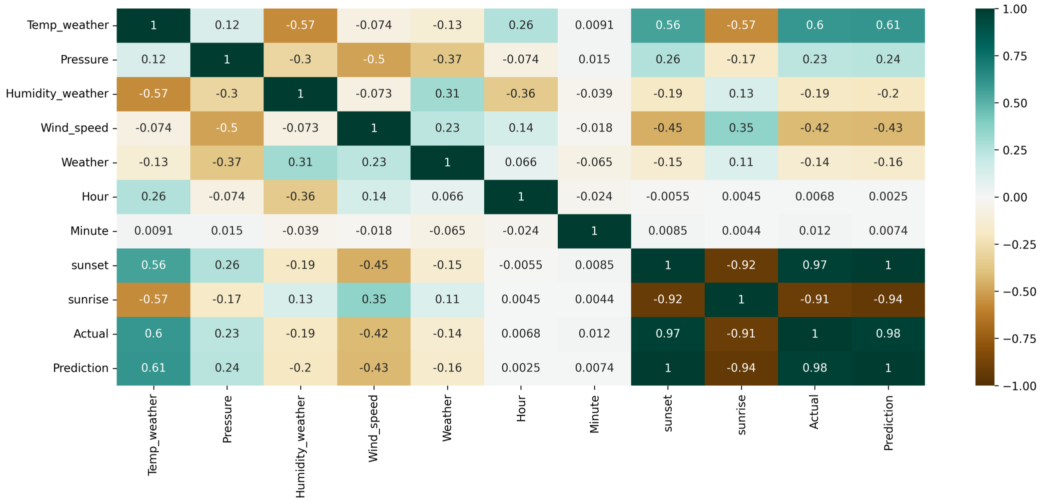


Fig. 5.6 Sub-Meter2 Correlation Matrix

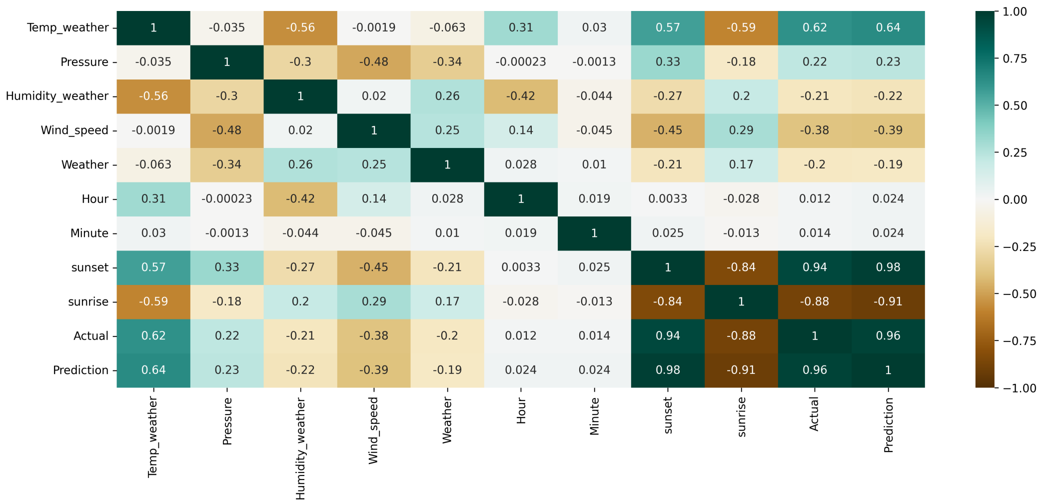


Fig. 5.7 Sub-Meter3 Correlation Matrix

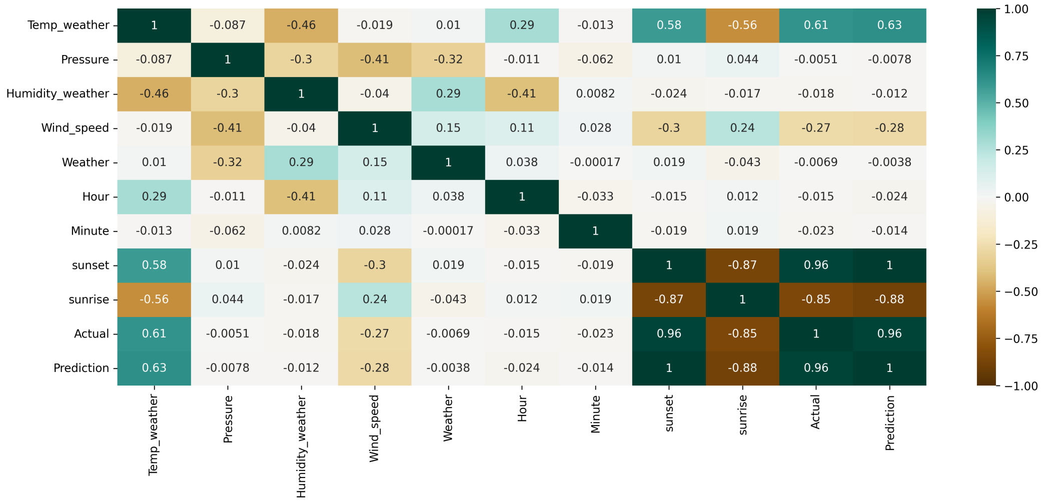


Fig. 5.8 Sub-Meter4 Correlation Matrix

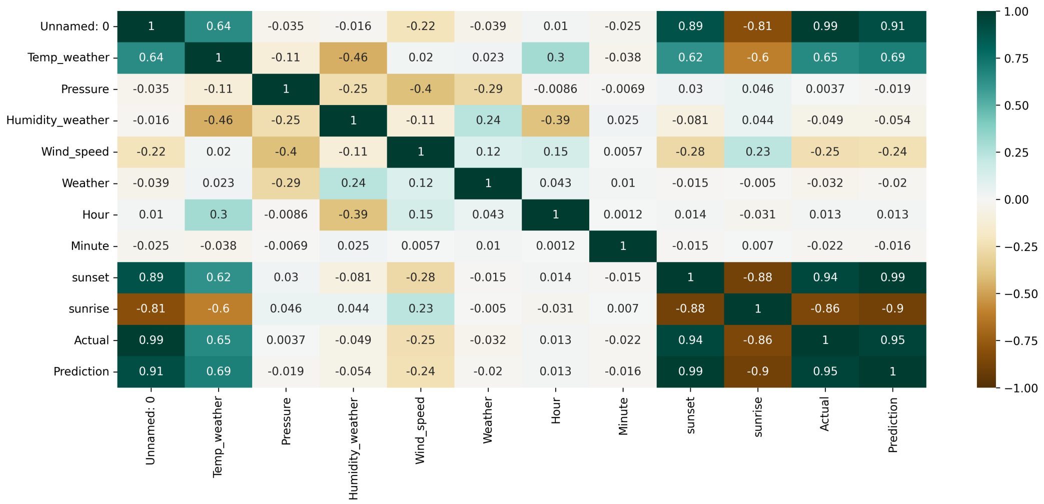


Fig. 5.9 Sub-Meter5 Correlation Matrix

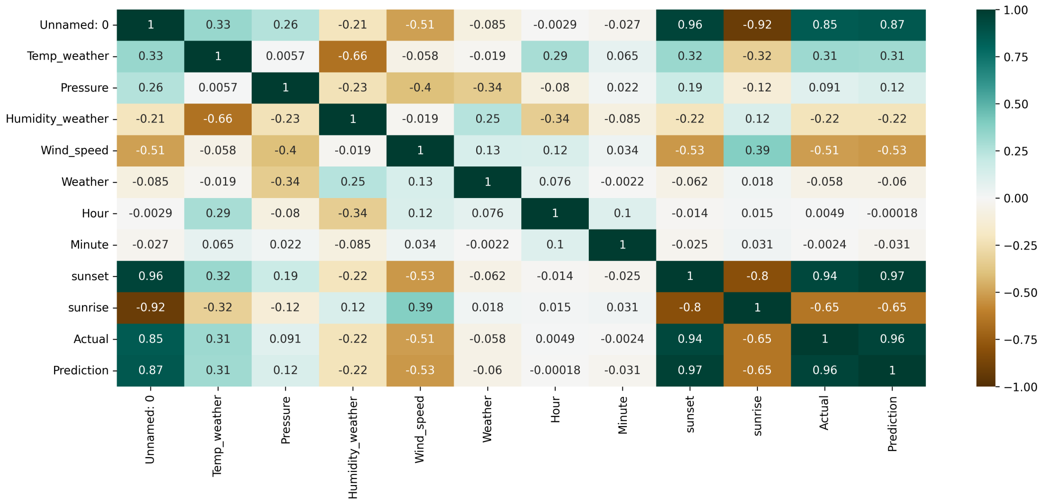


Fig. 5.10 Sub-Meter6 Correlation Matrix

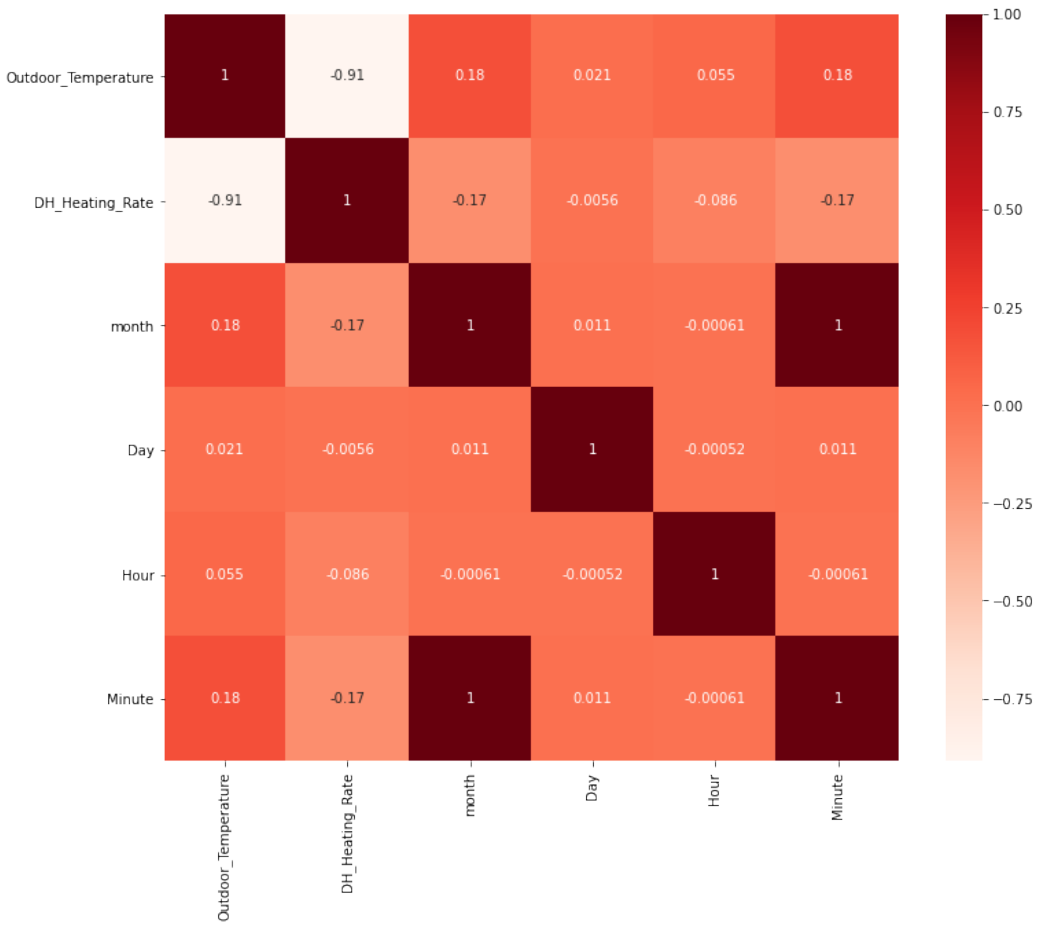


Fig. 5.11 Correlation Matrix representation of Simulated Data

Figures 5.4 and 5.11 present graphical representations of the correlations calculated for both datasets, with values expressed as percentages ranging from -100 to +100. The sign '+' or '-' denotes whether the correlation is positive or negative, and the decimal component is omitted for ease of interpretation. The correlation analysis reveals that outdoor temperature has a substantial impact on heat demand, with month indicating seasonal variation. In contrast, the correlation value is low for minute, day, and weekday, suggesting that these features may be excluded from the building model. Although the correlation percentage is small for hour and weekday, these features are retained since they reveal variations in consumption resulting from differences between weekdays and weekends, and working hours versus non-working hours, respectively, as shown in Figure 5.2.

To evaluate the performance of the models on the REMOURBAN project data, a similar analysis is conducted, as shown in Figure 5.4. The correlation between hour, minute, pressure, wind speed, and weekday is found to be close to zero. However, as previously reasoned, hour and weekday are retained in the analysis. Other variables exhibit significant correlations, which can be utilised in the Decision Tree and Artificial Neural Network models.

The performance of the models is assessed using different evaluation metrics. For the Linear Regression model, the R2 score is utilised, whereas the confusion matrix is used for the other models.

Figures 5.12 and 5.13 illustrate the overall predictions for the SHARING CITIES and REMOURBAN projects using Linear Regression, Artificial Neural Network, and Decision Tree. In the REMOURBAN project, the data are randomly split into a 70:30 ratio for training and testing the models, so the X-axis of the plot does not represent a timestamp but rather a numerical value. The trend in the REMOURBAN prediction plot is not clearly visible, but a more detailed analysis is provided in the results section. On the other hand,

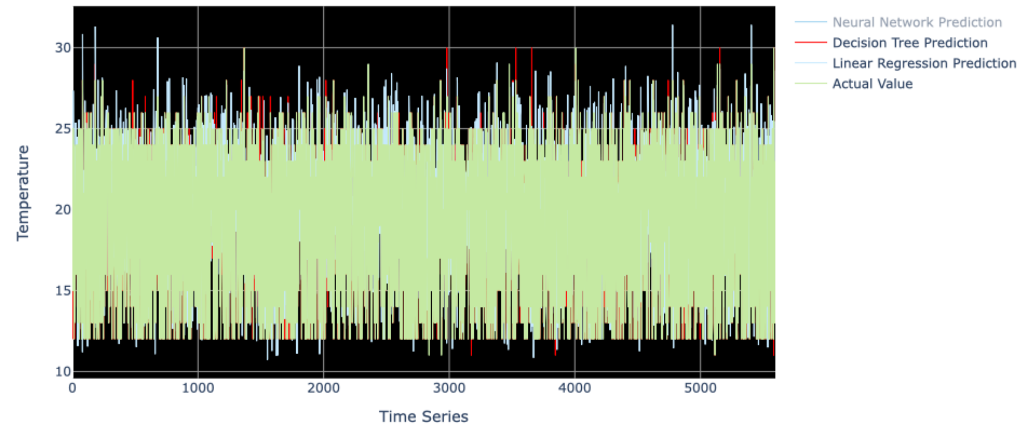


Fig. 5.12 Prediction of REMOURBAN Project

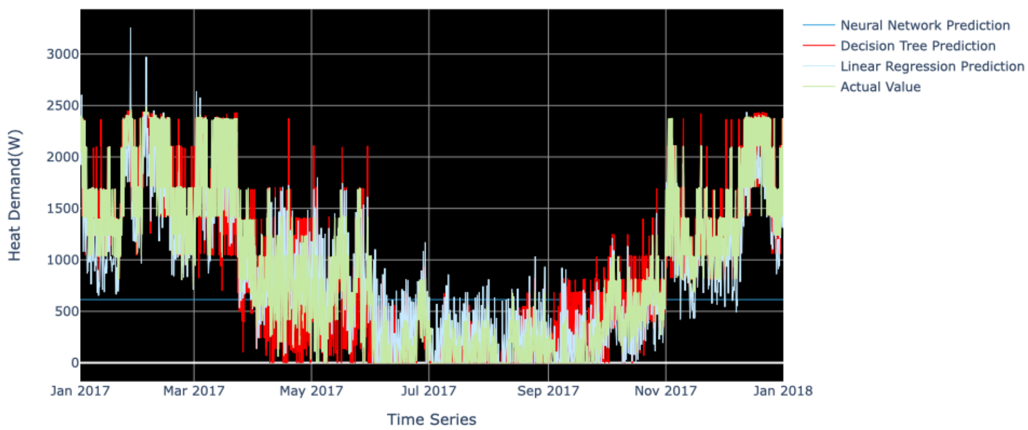


Fig. 5.13 Prediction of Sharing Cities

the SHARING CITIES prediction plot (Figure 5.13) shows that the predicted heat demand closely follows the trend of the actual heat demand.

5.4 Discussion and Results

The results of the performance evaluation indicate that all three models, Linear Regression, Decision Tree, and Artificial Neural Network, can be used for predicting heat demand in both the SHARING CITIES and REMOURBAN projects. The Linear Regression model shows good performance with high R^2 score on the SHARING CITIES dataset, while the Decision Tree and Artificial Neural Network models perform better on the REMOURBAN project dataset.

Furthermore, the feature importance analysis indicates that outdoor temperature has the highest impact on heat demand, followed by month, hour, and weekday. The minute, day, pressure, wind speed, and weekday features have a low correlation with heat demand and can be dropped from the building model.

In conclusion, the analysis shows that machine learning models can effectively predict heat demand in smart cities, and the choice of model and features should be tailored to the specific project and dataset characteristics. The results of this study can be used to inform the development and implementation of smart city energy management systems.

The superior performance of Linear Regression on the simulated data can be attributed to the fact that the data generation process is based on the physical characteristics of buildings and weather data, which can be well-represented by a linear model. In contrast, real-life data includes the influence of customer behaviour, which is a significant factor to consider, but introduces complexity in the analysis. Customer behaviour is influenced by non-tangible factors such as mood and affordability, making it difficult to capture in a model. However, over a longer period, consumption patterns tend to become more consistent, which can be leveraged for predictive modelling.

Table 5.2 Comparing Results: Summary of Model Performance Evaluation

Models	SHARING CITIES Project	REMOURBAN Project	Evaluation
Linear Regression	81.53%	54.77%	R Score
Decision Tree	24.75%	87.22%	Accuracy
Artificial Neural Network	21.00%	73.12%	Accuracy

Table 5.2 provides important insights into the performance of different algorithms on both the SHARING CITIES and REMOURBAN projects. The high R2 Score of 81.53% for Linear Regression suggests a close correlation between the test dataset and the predicted values, indicating that this algorithm is suitable for the SHARING CITIES project. In contrast, the accuracy of Decision Tree and Artificial Neural Network is less than 25%, indicating that they are not effective in predicting heat demand for this project.

However, for the REMOURBAN project, the situation is different. Decision Tree algorithm demonstrated the highest accuracy of 87.22%, outperforming both Artificial Neural Network and Linear Regression. This result suggests that Decision Tree is the most appropriate algorithm for predicting heat demand in the REMOURBAN project. The superior performance of Decision Tree in this context could be attributed to the complexity of the data generated by the REMOURBAN project, which contains a larger number of features and is derived from different processes.

The high correlation of 91% between outdoor temperature and heat demand is the dominant factor driving the prediction for the SHARING CITIES dataset.

Insufficient features are available for Artificial Neural Network to develop a learning pattern and improve performance. The performance of Linear Regression varies for different datasets, indicating that there cannot be a single method of heat prediction suitable for all district heating networks. The simulated and real-life collected data have different parameters, which can affect performance.

Decision Tree outperformed in accuracy for the REMOURBAN dataset but showed a low level of accuracy for the SHARING CITIES project due to the smaller number of features and the output's dependence on the dominant feature - outdoor temperature. Artificial Neural Network also showed a low level of accuracy for the SHARING CITIES project, with the algorithm failing to converge after 500 iterations due to a smaller number of features. However, the larger number of features had different effects in the case of the REMOURBAN project, where the Artificial Neural Network converged before 100 iterations and showed decent accuracy that can further improve with an increase in collected data.

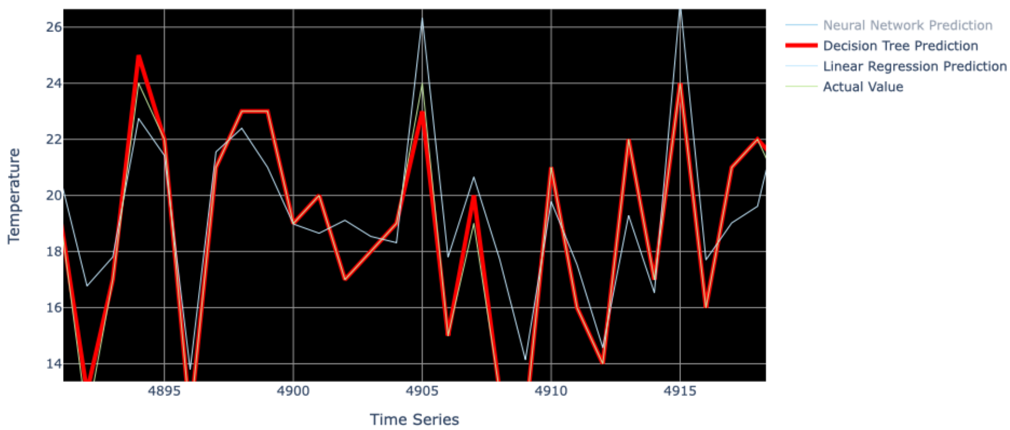


Fig. 5.14 Examining Heat Demand Prediction

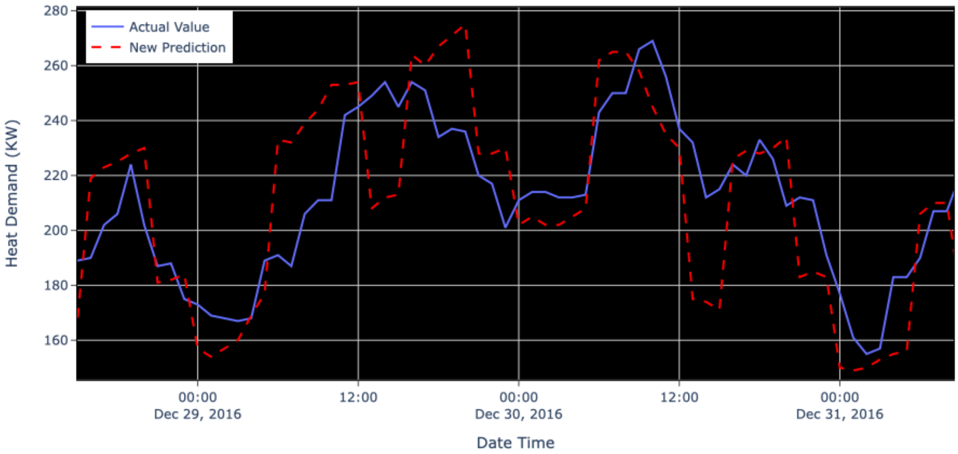


Fig. 5.15 Predicting Heat Demand for a Sample Day in the Sharing Cities Project

Figure 5.14 compares the prediction of Linear Regression and Decision Tree with actual values, with Decision Tree's prediction closely matching the actual value in red for the REMOURBAN data. Figure 5.15 compares Linear Regression's prediction with actual heat demand for the SHARING CITIES project, where the prediction is close to the actual value, showing two peaks per day. The accuracy of Artificial Neural Network will improve as data points with an increased number of features are collected. It can also be concluded that for datasets with a smaller amount of data, Decision Tree is highly accurate. This approach can also be used to predict short-term or long-term heat demand using weather forecast data.

5.5 Finalisation

The first model uses real-life data from a building, including temperature, humidity, and heat demand readings, to predict future heat demand. The second model uses simulated data generated from a computer model of a building, which takes into account various factors such as building materials, insulation, and heating systems.

The study found that both models were able to accurately predict heat demand, with the simulated data model performing slightly better in some cases. However, the authors note that the simulated data model requires more information and is more time-consuming to set up, while the real-life data model is easier to implement but may require more data cleaning and preprocessing.

The author conclude that both models have their advantages and disadvantages and can be useful in different contexts. The choice between the two models depends on the availability of data, the level of accuracy required, and the resources available for modelling.

5.6 Summary of Chapter

Chapter 5 discusses the datasets and analysis methods used in the study of predicting heat demand in the context of 4th Generation Low Temperature District Heating. The chapter compares two datasets from the SHARING CITIES project and the REMOURBAN project, which collected data from residential homes. It also presents the features used for constructing heat demand prediction models and outlines the analysis process, including data cleaning, feature selection, and model training and evaluation. The chapter concludes with a comparison of the performance of models trained on simulated and real-life datasets.

Chapter 6

Internet of Things Framework for District heating

6.1 Introduction

District Heating can be linked with ideas of robotised house framework, home automation, IoT based smart security and smart home, to predict heat demand and optimise the heat generation. Based on the literature review, IoT based framework is proposed, designed, and physically implemented as proof of concept in real life. Moreover, three machine learning algorithms are implemented, and a maximum-minimum algorithm is used to select best prediction out of three. The details of Machine learning algorithm and methodology implemented can be read in [2]. The results of framework are discussed in the later section.

6.2 An Internet of Things based Framework for District Heating

In this section, the existing framework is described, the summary of IoT framework is drawn and represented in graphics form. Also, the general framework with district heating as a demo case is discussed.

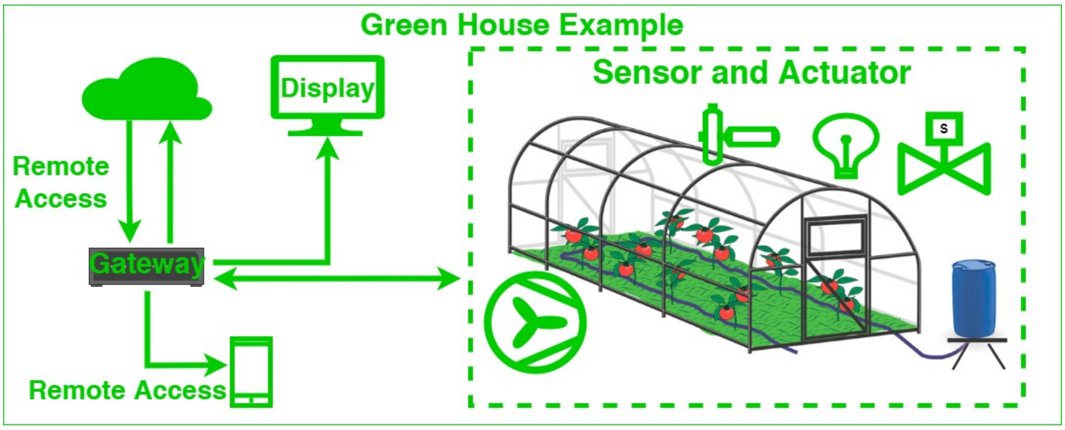


Fig. 6.1 IoT example of Green House [5]

In the thesis, [5, 166–168] the framework is described and explained in separate ways. As the thesis is not review thesis of IoT frameworks, it is out of scope to cover all frameworks. Overall, there are two major approaches in developing IoT frameworks, first is the way data flows and second is the process involved. To summarise the first approach an example of green house is shown in the [5] wherein Figure 6.1 the signal is sent from the remote-control device via gateway to the Actuator. The Sensor reading can be displayed via gateway on display panel. The second approach can be data acquisition from the sensor, data visualisation of sensor data and control the actuator.

Figure 6.2 shows the overview of the proposed IoT framework for District Heating. Sensors are connected to the Cloud computing or Webserver. Actuator receives the signal from the webserver. Before looking into the architecture, four

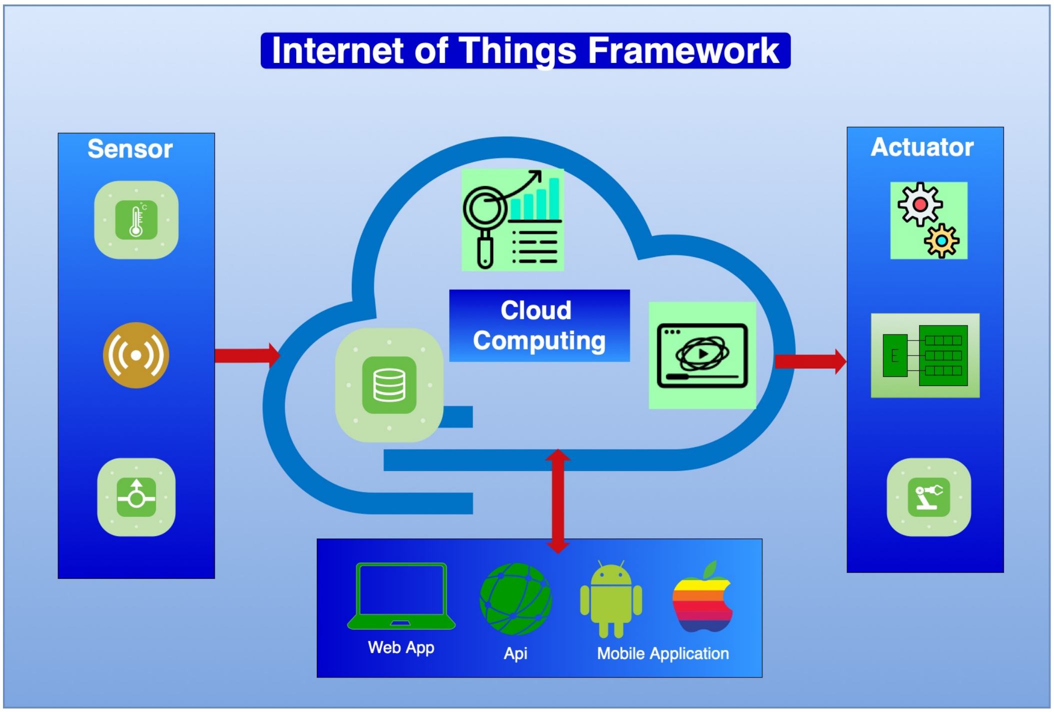


Fig. 6.2 Generic IoT Framework

important components such sensor, actuator, cloud computing and external devices/services needs to be defined for easy understanding.

6.2.1 Sensor

Sensor is physical device which can sense the change in the environment. Sensor can detect change in the environment depending on the use of data. The sensed data can be used locally or also can be stored for later use. The sensor for district heating system are smart heat meters. In the implementation section several types of sensors apart from smart heat meter are described.

6.2.2 Actuator

Actuator is physical device which converts energy to motion. The example of actuator is simple DC (Direct Current) motor, light, or bulb. The actuator produces the output based on the input. The typical actuator in district heating systems are boilers, valve, and heat pumps. Different District heating system may have actuators based on the system scale.

6.2.3 Cloud Computing

To understand cloud computing in simple manner, it is a computer which cannot be seen physical by user but can be accessed from anywhere in the world. Companies like Amazon, Microsoft, Google, IBM etc are famous providers. The proposed framework is using cloud computer for IoT. Three main aims of cloud computer are to store data, analysis data and visualise the data.

6.2.4 External Device/Service

Apart from sensors and actuators, other devices and services are connected to the cloud. In the district heating systems users of energy plays key role, so the framework allows them to connect to cloud with limited access via computer, laptop, and mobile phone. Moreover, sending notification via email, SMS, and phone, are the external or paid services required to be used by the cloud.

6.3 IoT Architecture

The framework is shown in Figure 6.2 and it is simple for understanding, but lot more is involved technologically. IoT Architecture is illustrated in the Figure 6.3. Sensor/Actuator needs to communicate with the cloud to send and receive signals. So, microprocessor or microcontroller is needed to communicate with the cloud. While communication is done via internet different communication network and technologies can be used. Moreover, local sensor and actuator network can be built which are accessed by single microprocessor. Communication can be done wireless and wired, using long- or short-range

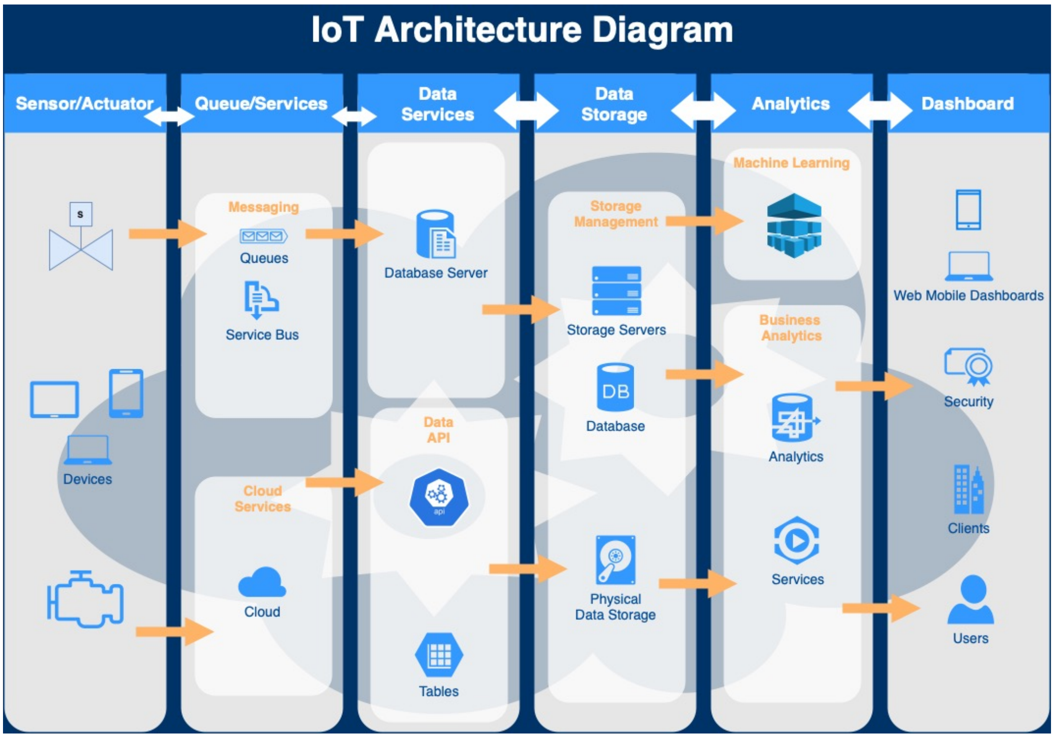


Fig. 6.3 IoT Architecture for proposed framework

protocol. The messaging in other words data transfer over internet can be HTTPS or MQTT which are industry standard. Custom microservices can be deployed to interact with multiple cloud providers/platforms, to leverage the advantages of cloud. Data in the framework is not only about sensor data but it can be files, images etc. So, data services are required where the database server is hosted. The data server will be connected to the data storage. The data server will handle the data coming from the sensors, users, and external services, store it at correct location.

Data storage should host storage server for all time data availability. Database where structured data is stored on data storage. Data storage would also store files. Data analytics would be connected to the storage server to access the data and analyse it.

The machine learning algorithms would predict the heat demand using the data gathered from sensors. The prediction is not the only thing done by data analytics; business insights can be derived from the data as well. Last but not the least dashboard is where users and clients can securely interact with the different data visualisations and can access control panel. In the next section implementation of the framework is discussed.

6.4 Implementation

In this section, the demonstration of the Framework and IoT architecture is discussed. REMOURBAN a major Smart Cities demonstrator project funded through the Lighthouse project scheme of the European Union's Horizon 2020 research and innovation programme is used as case for the implementation. It aims to support the design, testing and validation of new models of urban regeneration in the 3 lighthouse cities two follower cities. The project places an emphasis on the need to develop innovative and holistic regeneration models that maximise the convergence of energy, mobility and ICTS.

As part of the REMOURBAN project, deep retrofitting and integrated local energy system was introduced to existing buildings (27 terrace houses in Sneinton, Nottingham). This research aims to introduce economically sustainable intervention aiming to maximum utilise the energy generated on site and achieve maximum offset of the energy consumed from the network. The heating system has been completely changed, the gas boilers has been removed and replaced by completely new, low temperature heating system. The cluster of houses now is configured as a Micro Low Temperature District Heating (LTDH) network. Retrofitting funded by REMOURBAN project connects Nottingham existing extensive district heating to a secondary network of 94 dwellings via the return pipe of the existing system. This new secondary system has a lower feeding temperature of 60 C to 65 C, developing in this way a new 4th generation Low Temperature District Heating (LTDH) for the first time in such scale in UK. This new lower cost, low carbon alternative of district heating replaced the expensive and inefficient gas heating within the 94 dwellings. The LTDH system extracts unused heat from the existing system making it more efficient. The LTDH system takes water in at 60 C to 65 C, extracts the heat and returns it at about 35 C. As well as providing heating for the four blocks, a high-efficiency plate heat exchanger converts mains cold water into instantaneous hot water for each property. Because it operates at a lower temperature, costs are reduced, heat losses in the system are lower and efficiency higher. The system successfully provides low-carbon space and water heating to a significant number of properties, taking advantage of heat which would otherwise will be unused.

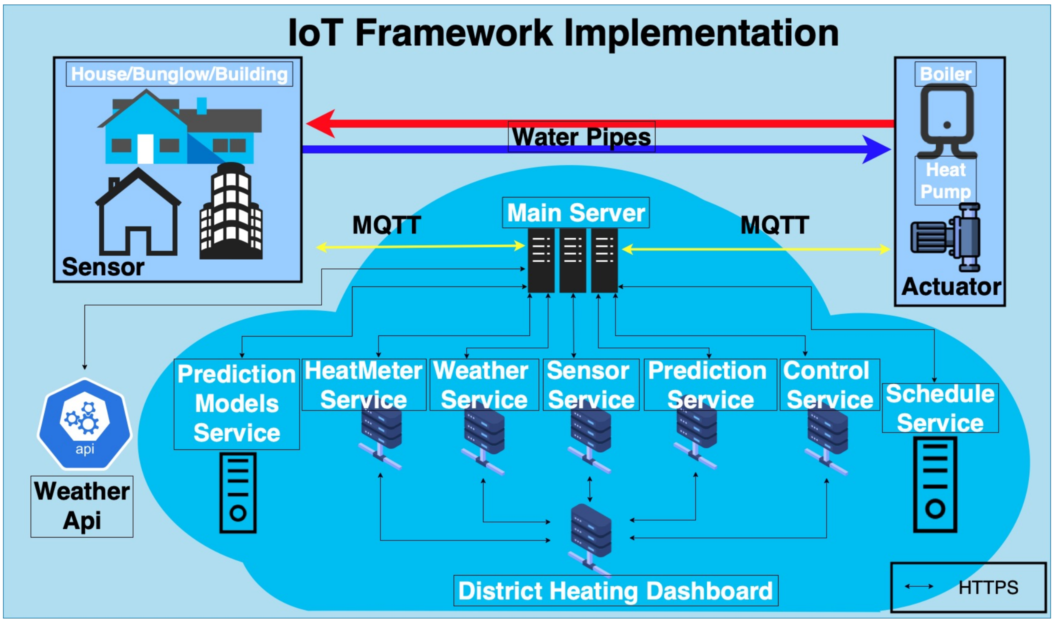


Fig. 6.4 The IoT Framework Setup

The overview of the implementation is shown in the Figure 6.4. Homes and actuators are connected via pipes, sensor and actuator are connected via cloud in the district heating. Weather data is vital for heat prediction which in return can be used to make decisions on the actuators. Weather data is acquired from weather API. On the cloud various process needs to be carried out. First process is to store the data, second is to analyse data and last but not the least data visualisation. All the details and break down of each part are discussed below.

The sensors are installed in the homes as the part of REMOURBAN project. The smart meters are also installed at different courts and at substations. The sensors installed in 20 Homes and 10 Heat meters are installed across residential homes part of the REMOURBAN project. These sensors and smart meters are connected using Long Range (LoRa) Gateway to cloud storage. LoRa is one of the open standards used for IoT which is been adopted by over 500 companies in the world [204]. MQTT (Message Queuing Telemetry Transport) protocol is implemented for data transfer of the devices. The data transferred by standard text-based format JSON (JavaScript Object Notation) as data are generally in structured format. Sensor senses temperature, CO2, humidity, light, and motion whereas smart heat meters send heat consumed, flow temperature, return temperature, flow, power, volume.

For the implementation purposes open-source and limited call weather APIs (Application Programming Interface) were used. The weather APIs used are OpenWeatherMap and AccuWeather. The weather APIs are used for two purposes, first current weather and second forecasted weather. The HTTPS (Hyper Text Transfer Protocol Secure) protocol is implemented for calling the API. The weather API can return in XML (eXtensible Markup Language) or JSON, but to keep consistency JSON response is called. AccuWeather variables are Weather Text, Temperature, Real Feel Temperature, Real Feel Temperature Shade, Relative Humidity, Indoor Relative Humidity, Dew Point, Wind angle, Wind direction, Wind speed, Wind Gust, UV Index, Visibility, Obstructions Visibility, Cloud cover, Ceiling Pressure, Pressure Tendency, Past 24-Hour Temperature Departure, Apparent temp, Wind Chill Temp, Wet Bulb Temp and Precip1hr. OpenWeatherMap variables are Temperature, Maximum

Temperature, Minimum Temperature, Pressure, Humidity, Wind speed, Sunset time, Sunrise time, Weather, Latitude and Longitude.

The data coming from sensors and API must be stored. The data can only be stored if data is coming from the authentic sources. Main server is built on the cloud to accept the data. Main Server consist of LAMP (Linux, Apache, MySQL, PHP/Perl/Python) server. The data coming from sensors and API are structured, so MySQL is used for storing the data. Python based services are written on the main server to store data on MySQL. MySQL is database used to store the data in which separate tables are stored. Main server has main database and runs storage server which runs 24*365 as well as to orchestrate the process on the cloud. The cloud architecture implemented is micro-service based. The architecture implemented is inspired from the previous [6]. Eight micro-servers are implemented with 8 different services namely prediction models service, schedule service, heat meter service, weather service, sensor service, prediction service, control service and district heating dashboard. The implementation of such micro-servers helps in reducing load on single server as well as reduce the running cost due to pay as you use model of cloud. The services communicate over HTTPS protocol using JSON as messaging and files over FTPS (File Transfer Protocol Secure) protocol. All the services are written using python as coding language.

Out of 8 micro-servers, 2 micro-servers run the algorithms that are prediction and control algorithm. Prediction Models service is connected to the main server for data of sensor and weather. The prediction models service receives the data and generates different model files for prediction. The generated files are stored on the main server. For the implementation, 3 machine learning algorithms are used which were Linear Regression, Decision Tree and Neural Network. The prediction models service will generate at-least 3 model files.

Schedule Service is used for generating the schedule for controlling actuator. The schedule duration is for 24 hours. This service uses predicted heat demand from the main server to generate the schedule to turn on/off the assets of heat generation plant. As this service is currently configured to run only once in 24 hours unless configured otherwise. The schedule service runs prediction service to use latest prediction. The result - generated schedule and predicted heat demand are stored in the database of the main server.

Three micro-servers out of remaining six micro-servers, hosts the services which are used for data visualisation. The services hosted are heat meter service, sensor service and weather service. All three services are used to prepare data for visualisation which helps better analyse and understand data. The process of preparing data involves cleaning of data, variables names and providing data privacy, security thereby displaying only data which the user is allowed access. Example smart heat meters in REMOURBAN project are not in the individual homes, so user will not have access to smart heat meter data. But user still be able to visualise heat consumed. Also, in the example user will have only one variable that is heat consumed. Moreover, each service will only prepare the data for respectively service meaning sensor will prepare only sensor data. Prediction service is the service which forecasts an estimate of user heat demand. Current implementation of this service is set to predict heat demand for 5 days. This service requests main server for weather forecasted data and ML model files to predict the heat demand. Again, the prediction can be configured based on the forecasted weather data available.

The heat generation is dependent on the assets of district system. Assuming that only one boiler and one heat pump is installed, control is simulated. The idea of the thesis is to run the assets automatically based on the schedule. In certain use cases the asset needs to be controlled manually. Such requests are handled by the control service. Control service is written to show the status of the actuators and manually control it. This service is directly used via mobile application or dashboard.

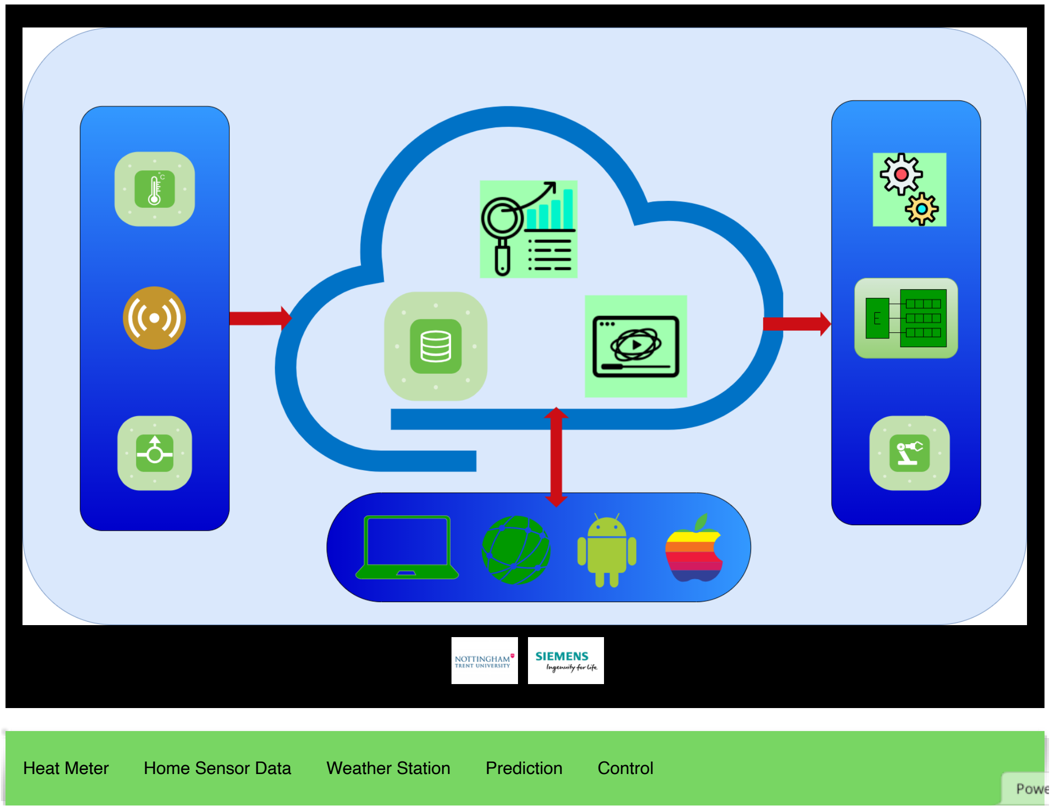


Fig. 6.5 Home Screen

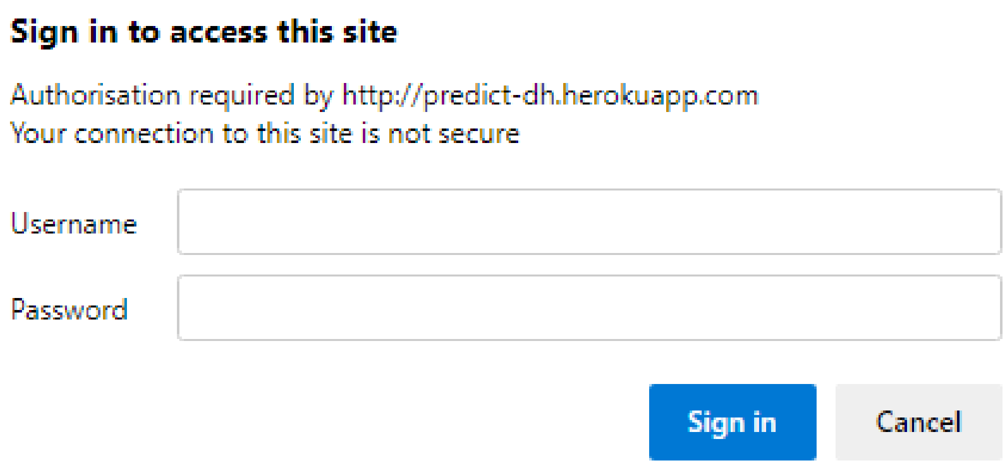


Fig. 6.6 Login Screen

Last but not the least District heating dashboard is hosted on webserver using PHP (Hypertext Pre-processor) and HTML (HyperText Markup Language) pages. District heating dashboard allows user to communicate via web browser. Dashboard visualises different data such as heat meter, sensor, weather, prediction, and control. Dashboard is connected to the five services namely; heat meter service, home sensor service, weather service, prediction service and control service. Based on the input from the user the service is called, and data visualisation is displayed. When district heating dashboard called from computer or mobile, the homepage is opened. The homepage is shown below in the Figure 6.5. The five tabs options are displayed. Before reaching to the homepage authentication is passed by using username and password as shown in the Figure 6.6.

The implementation is across the 2 different cloud providers and on-premises server to illustrate the interoperability. Sensors sends data using Azure cloud services to on-premises server. The micro-servers are deployed on the salesforce cloud. The micro-servers are deployed on 512 MB (Mega Byte) RAM and 2 process cloud computers. Main server which is on-premises server, is Xeon processor, 16 GB RAM, 256 SSD and 1 TB HDD system. Moreover, for algorithms i5 processor, 8 GB RAM and 256 HDD system is used.

6.4.1 Model Definition

This thesis works and on three machine learning algorithms which are commonly used for heat prediction. Also, the thesis takes a novel approach to use all three machine learning algorithms to predict heat demand. Optimisation of the three predicted values is carried by maximum - minimum algorithm to generate the optimised schedule for heat generation plants. To simplify the equations, they are not expanded with all features, weight, probability and co-efficient.

6.4.1.1 Neural Network

Neural Networks are often used called black box, give input, and generate output. But the network built can be mathematically formulated. Equation 6.1 shows that summation of Weighted matrix of neuron times feature. The weighted matrix for the model is built with 12 hidden layers. So, the weighted matrix size is 12 x 12.

$$Q_{NN} = \sum_{i=1}^{n} W_i X_i \tag{6.1}$$

Where,

Q_{NN} = Heat Demand Predicted by Neural Network method

W_i = Weight matrix of neuron

X_i = Features like Temperature, Pressure

6.4.1.2 Decision Tree

Decision Tree algorithm is probability driven. Each combination probabilities are generated, and tree structure is used. The probabilities are calculated using historical data. The search algorithm to find the output from structured tree. The equation of each branch is given equation 6.2. All branches probability is added to find the final probability.

$$Q_{DT} = \sum_{i=1}^{n} P_i X_i \tag{6.2}$$

Where,

Q_{DT} = Heat Demand Predicted by Decision Tree method

P_i = Probability of branch

6.4.1.3 Linear Regression

This algorithm calculates coefficient for each feature based on the historical data. The prediction is generated by inputting data into equation 6.3.

$$Q_{LR} = \sum_{i=1}^{n} B_i X_i + C \tag{6.3}$$

Where,

Q_{LR} = Heat Demand Predicted by Linear Regression method

B_i = Co-efficient of regression

C = Intercept constant

6.4.1.4 Minimum Maximum

The algorithm is used to optimise the output of the heat demand predicted by the three machine learning algorithms discussed above. The algorithm also works as filter to eliminate the odd prediction. The equation 6.4 represents the optimisation equation. a,b,c is the constant derived from the data collected in the past.

$$Q_{Predict} = a * Q_{NN} + b * Q_{DT} + c * Q_{LR} \tag{6.4}$$

Where,

$Q_{Predict}$ = Heat Demand Predicted by Maximum Minimum method

a, b, c = constant calculate using historical data

The equation and the condition used for the optimisation is simple but effective. The condition used to optimise is given by equation 6.5.

$$Min(Q_{NN}, Q_{DT}, Q_{LR}) \leq Q_{Predict} \leq Max(Q_{NN}, Q_{DT}, Q_{LR}) \tag{6.5}$$

6.5 Results & Discussion

6.5.1 Analysis

In this section, the implementation performance is analysed. The data flows from the sensor to the data storage. The raw data stored in the database from each home for a year is summarised in Figure 10. This figure indicates that data collected from home 23 is less than 300 which suggests that data from that sensor is not transmitted regularly after installation. The service can be written to check regularly if data has been received. In case of any faults detected maintenance can be carried out. As per Figure 6.7, apart from one home, all the other homes transmitted data as expected and it was stored in the database.

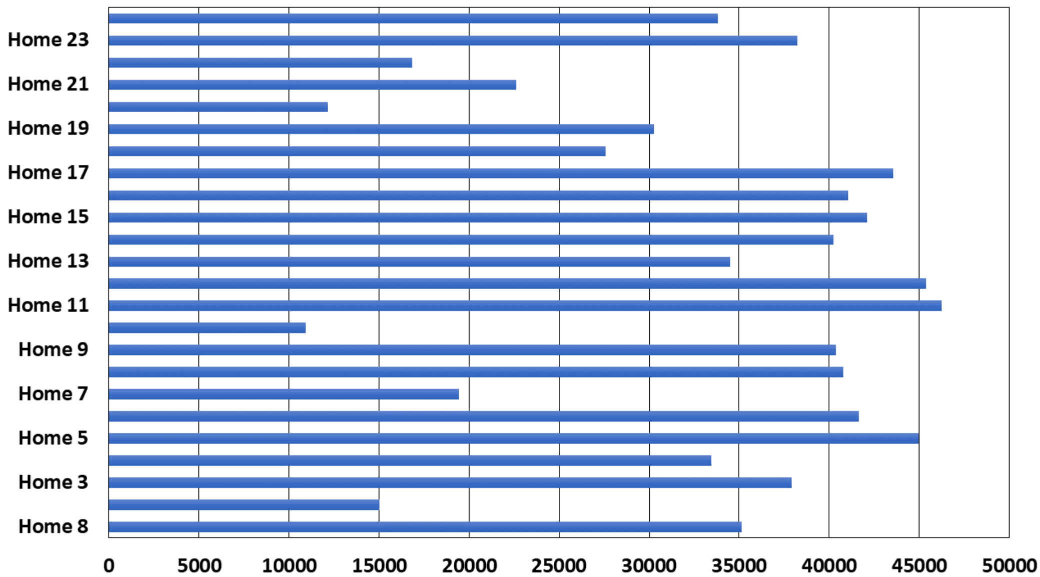


Fig. 6.7 Summary of Data Collected from Each Home

Figure 6.8 shows the performance of the prediction models. The Decision Tree algorithm shows consistent performance with accuracy over 85% to predict heat demand. Green trendline displays that the average accuracy is more than 90%. Neural Network algorithm accuracy ranges from 30 – 90 %. The

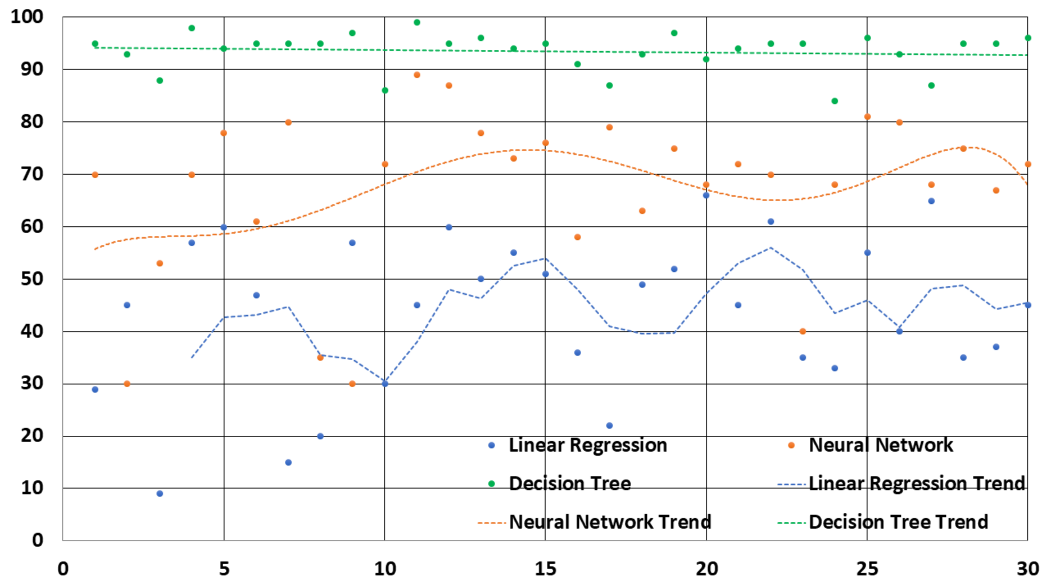


Fig. 6.8 Summary of Prediction Model Accuracy

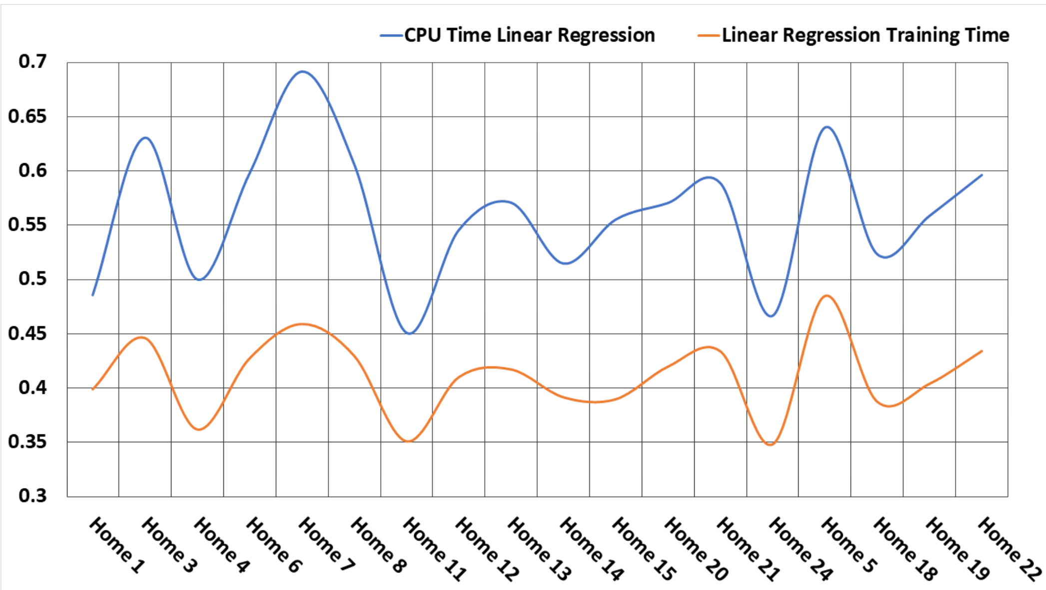


Fig. 6.9 Comparison of CPU Time vs Training Time

prediction using Neural Network can improve over the time as the data increases. Linear Regression shows lower accuracy than Neural Network and Decision Tree. As the primary idea is to use the cloud which works on the pay as you use model, the usage time plays a vital role.

Figure 6.9 illustrates time required for training and overall process for Linear Regression. The difference in the CPU time and training model can be seen. The charges are applied based on the CPU time because entire process that is receiving data, building model, and transmitting the model file utilises the processing power. For testing purposes different combinations were tried, and maximum CPU time is considered. Linear Regression takes less than 1 second of processing time while Decision Tree takes maximum 25 seconds for each home. Maximum time is utilised by the Neural Network which is 40 mins for each home. For all homes linear regression takes 10 seconds, Decision Tree takes 4 mins and Neural Network takes 6 hours. In a day, the prediction model service can run 4 times a day.

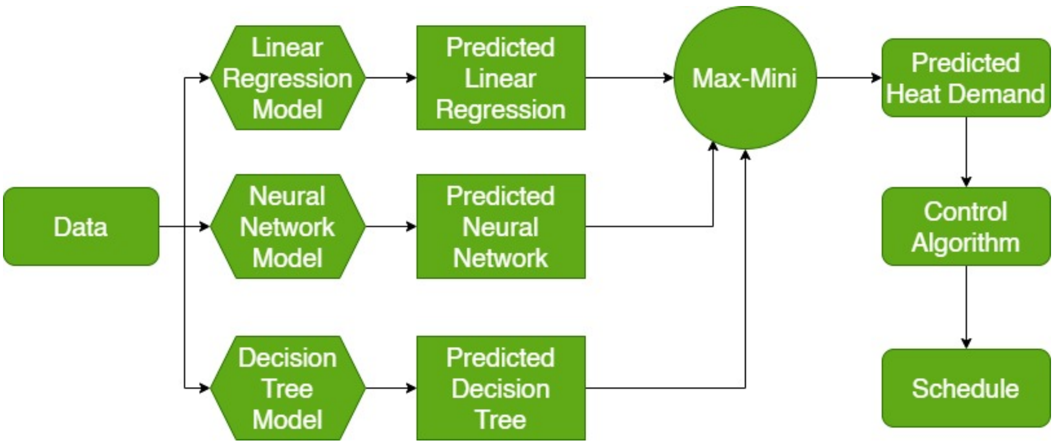


Fig. 6.10 Data flow of Prediction process

Prediction process flow is shown in Figure 6.10. As a starting step, Linear Regression model, Neural Network model and Decision Tree model, predict the heat demand simultaneously. The sample prediction of all three algorithms is presented in Figure 6.11. All three data vectors are passed to activation

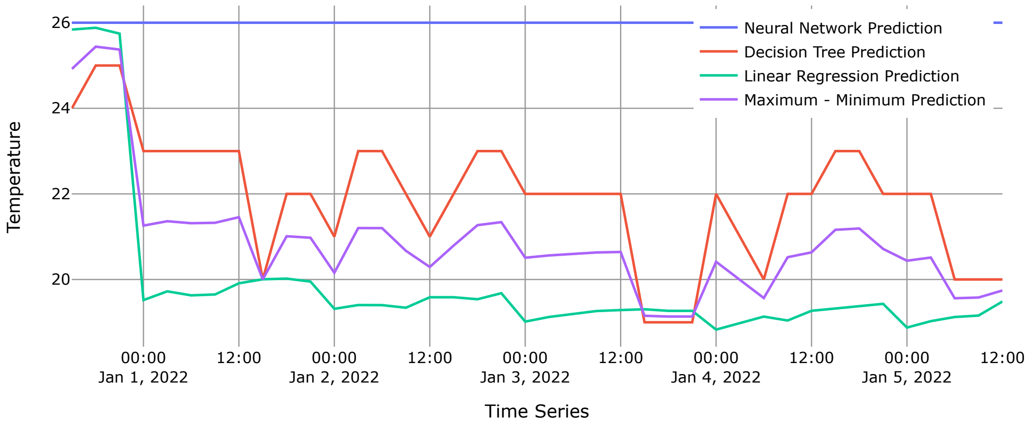


Fig. 6.11 Sample of Home 1 Prediction

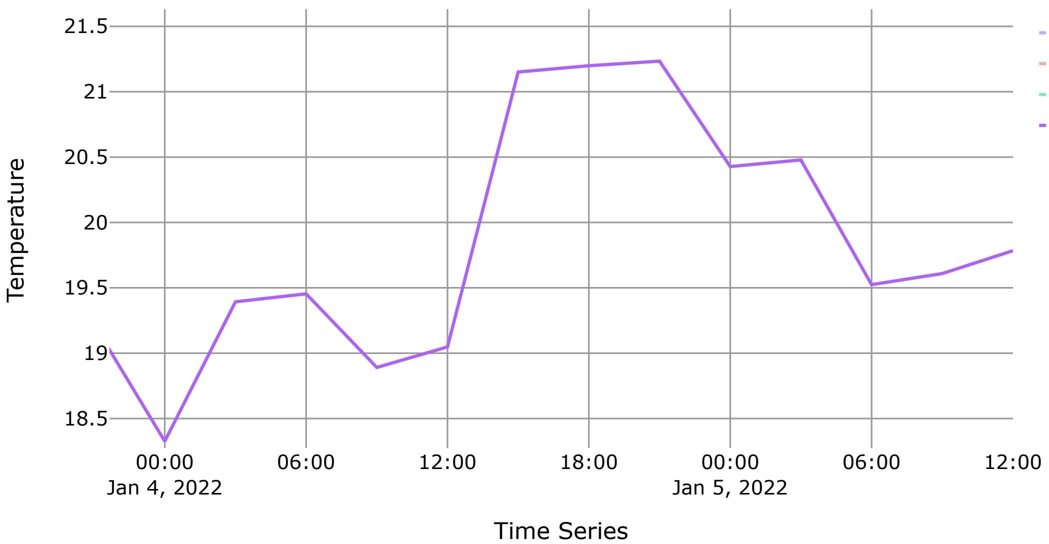


Fig. 6.12 Minimum - Maximum Algorithm Output

function i.e., minimum maximum algorithm, which optimises the prediction value. This predicted value works as input for control algorithm to generate schedule. Figure 6.12 illustrates the output of a day using minimum maximum algorithm. The output predicts the two clear peaks, which is significant for heat generation.

A dashboard has been built on the server for easy navigation between various tabs. The heat meter data visualisation includes data from all the heat meters and can only be accessed by master user. Heat meter dashboard has three dynamic charts where different variables can be selected for each chart. The charts are i) a simple line chart of the variable selected by the user, ii) a scatter chart to visualise data, and iii) a scatter chart of all accumulated data.

By default, variables for line chart are flow and date time which make time series analysis easy. Another feature of the first graph is that the window can be selected, and slider moved across to visualise data for other time periods in the selected period window. E.g., if a one-month window is selected, on using the slider the graph would update to appropriate one-month window. In the Second chart the same variables as the first chart are selected by default. The special feature here is that data can be zoomed by dragging across the area of interest on the chart. The screenshot of the dashboard is shown in Figure 6.13. The third chart illustrates the correlation between flow temperature and return temperature. Different correlation can be explored using the dashboard by user as well as the energy provider.

In Figure 6.14, screenshot of the home sensor dashboard is shown. Like heat meter dashboard this dashboard has 3 charts. In the heat meter dashboard, the variables that can be selected are power, flow, volume, flow temperature, return temperature and heat consumption while the variables for home sensor dashboard are different sensor selection, temperature, humidity, light, CO2, and motion. On both dashboards, the first chart, that visualises time series,

Fig. 6.13 Data Visualisation of Heat Meter

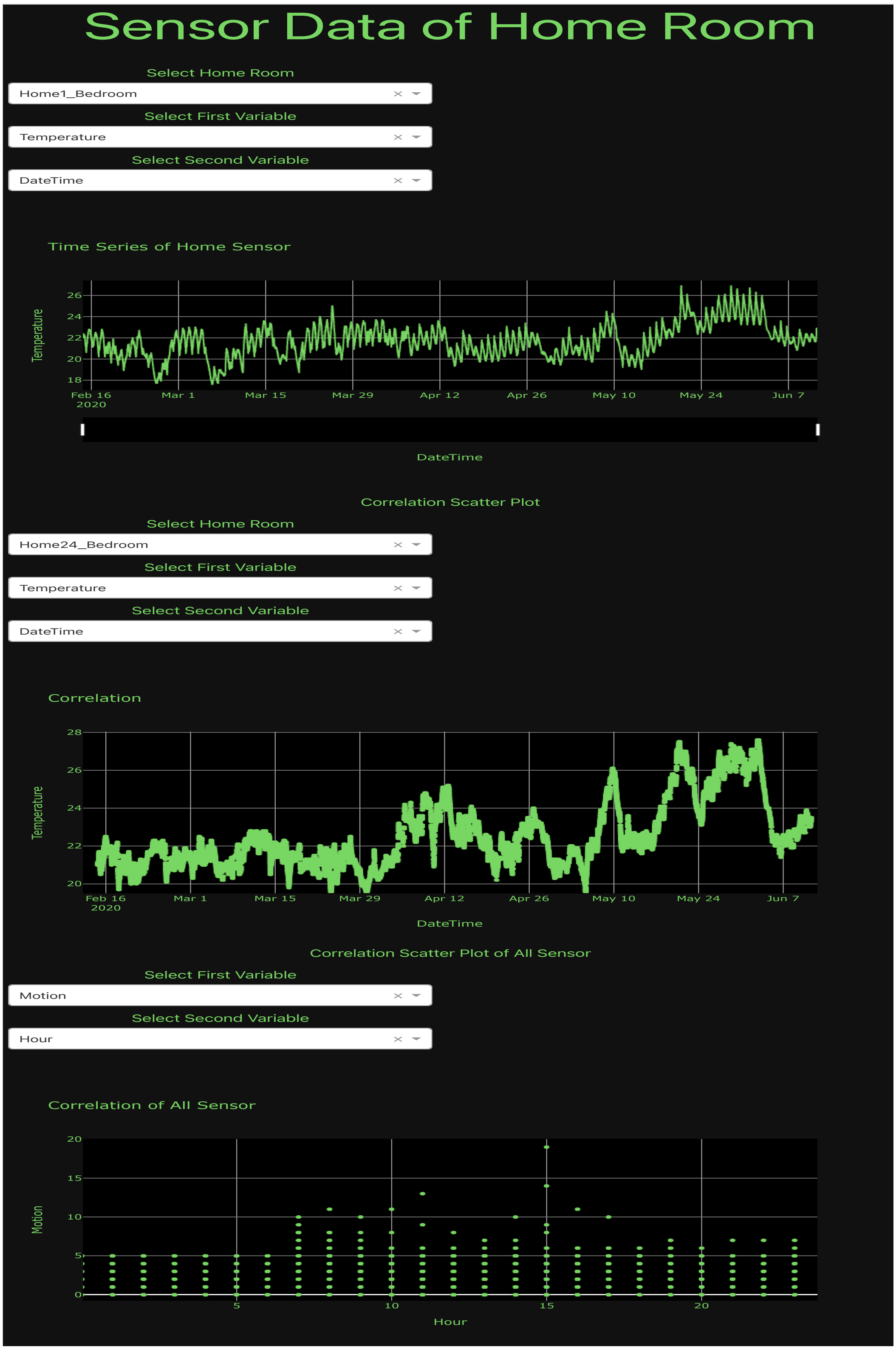


Fig. 6.14 Data Visualisation of Home Sensor

the variables that can be selected for y-axis are date time, minute, hour, year, day, and month. On the other charts any variable can be selected as x-axis and y-axis based on the heat meter and sensor to which the user has access. The third chart displays the correlation of the accumulated data. Both these dashboards can be used as tools to explore data and learn the pattern.

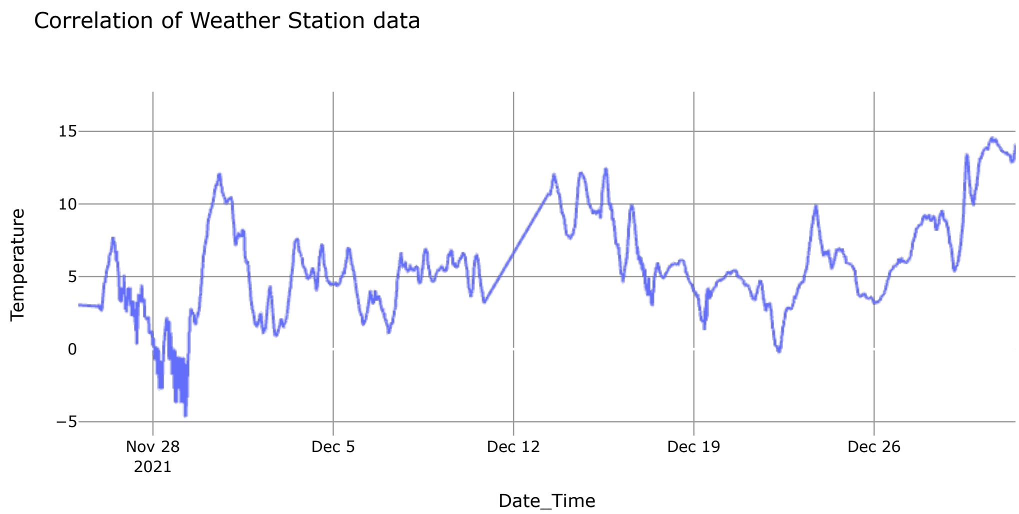


Fig. 6.15 Correlation chart from weather dashboard

The next dashboard is the weather station dashboard which will request weather data from the main server based on the details shared by the user. The weather station dashboard has two dynamic charts. The first chart displays temperature vs date time of the previous week by default and when zoomed out, historical temperature of up to previous year can be seen. Similarly, the second chart by default displays overall temperature vs date time data as shown in Figure 6.15. This chart allows the user to play more than usual with the weather app by selecting first and second variable as well as zooming in/out and selecting area of interest.

The important part of the dashboard is prediction tab in which accuracy of the models and prediction value are displayed. In this tab user input or selections are limited. For a normal user, the performance of the prediction models is displayed in the table. Moreover, a chart displaying prediction for the next 5 days is shown. The screenshot of the sample is illustrated in Figure 6.16. The user can still interact with the chart using zoom in/out and selecting the area of interest. Another feature is by clicking the method name, the prediction value can be turned on/off. Last but not the least, the amount of data collected from the house is displayed as the last chart.

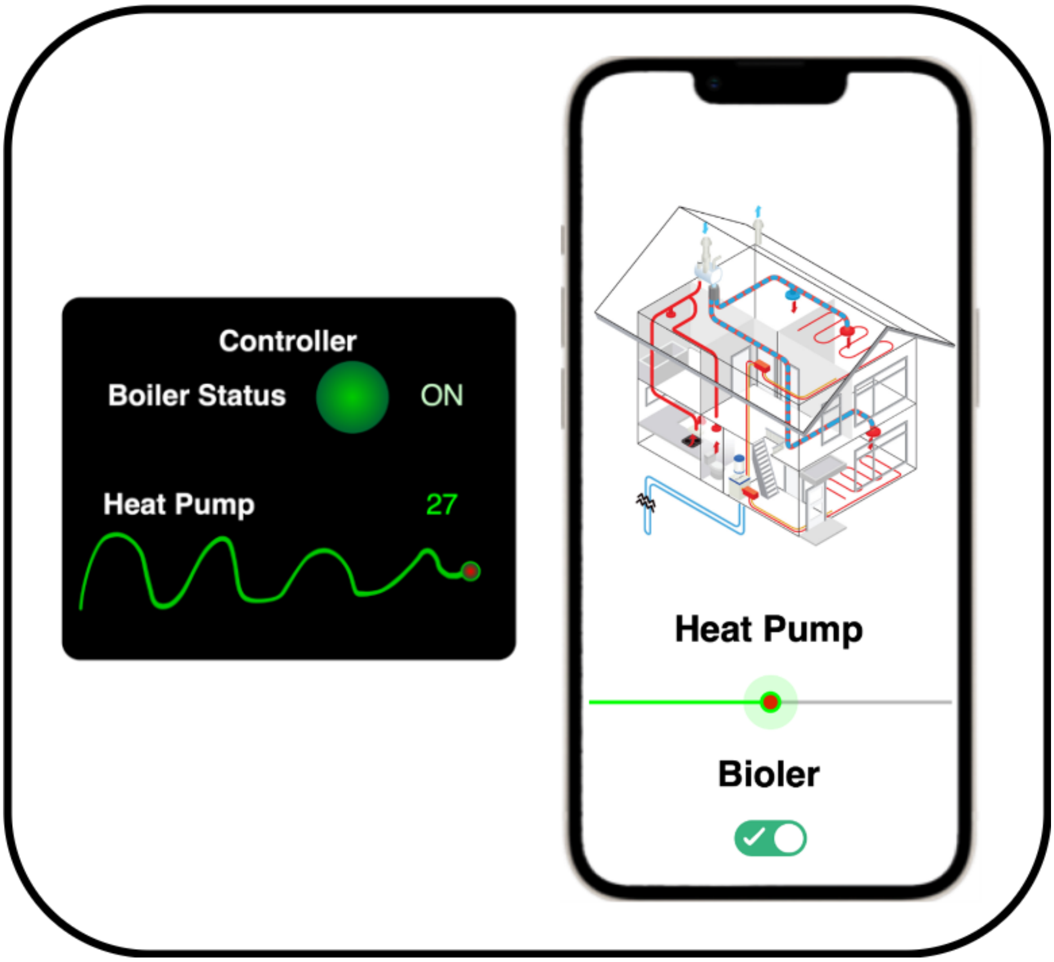


Fig. 6.16 Screenshot of Freeboard.io dashboard (left) and Mobile App (right)

The control of heat generation plant is an important part of the dashboard. A separate app has been developed for the research. The web app and mobile app both have the same user interface (UI). As shown in Figure 6.16, the slider is used to control the heat pump while a simple toggle button is used to control the boiler. However, the thesis is currently in testing phase and so control signals cannot be sent to the heat generation plant. But for validation purposes the freedboard.io dashboard is used which is not a part of the framework. An LED (Light Emitting Diode) and graph with text are configured on freeboard.io to represent the actuators and validate the receipt of signals.

Overall, the proposed framework and the implementation gives validation to the research gap. The framework implementation is active and running since February 2020. The data collected creates the new avenues to explore in district heating. Also, the tool developed while implementing the framework can be used to derive the data insight and data exploration process. Moreover, the framework allows the researchers to use different heat prediction algorithms which are studied in the literature review, to be implemented in real-world scenario. An improved performance of heat prediction algorithms can be achieved by evaluating the implemented algorithms. Furthermore, different control algorithms can also be easily implemented, tested, and evaluated. This helps in the optimising district heating system as a whole. Finally, the framework can monitor and control the assets from anywhere via the dashboard. The dashboard is live, and it generates charts using real-time data and the prediction carried out is also real-time.

The district heating framework can incorporate thermal storage, valves and networks which can help optimise district heating system. As the framework has used the cloud, the advantage of cloud technologies can be leveraged easily. For example, cloud provides the SLA (Service Level Agreement) of 99.9% which means the reliability or availability of the servers is all time, Scalable which means when user base increases, services can cope with requests and last but not the least, pay as you use will allow the servers to be charged only when in use. The system is also cost-effective since next to pay as you use, the client will only be charged when the dashboard is used, and not at all times. The framework can be extended to incorporate the building management system as well as automated billing system.

6.6 Summary of Chapter

The chapter discusses the proposed IoT framework for district heating, which includes sensors, actuators, cloud computing, and external devices/services. The IoT architecture involves communication between sensors/actuators and the cloud, using microprocessors or microcontrollers, various communication networks and technologies, and industry-standard messaging protocols. The data gathered from the sensors is stored in a data server and analysed using machine learning algorithms to predict heat demand and derive business insights. The implementation of the framework is demonstrated using the REMOURBAN project, which involves installing sensors in homes and smart meters at different courts and substations. The data is transferred in a standard text-based format and analysed using machine learning algorithms such as Linear Regression, Decision Tree, and Neural Network. The prediction models service generates different model files for prediction, and the schedule service generates a schedule for controlling the actuator. The chapter also discusses data visualization and the services used to prepare data for visualization, such as the heat meter service, sensor service, and weather service.

Chapter 7

Micro-level of heat prediction with user centric approach

7.1 Introduction

The research in this chapter focuses on the implementation of modelling approach based on emerging in IT technologies. For this research we have a group of residential homes equipped with individual smart meters, the heat is provided by an energy centre and all homes in the group are connected into a local district heating system and are also part of the IoT framework. The technique - how the data from the smart meters can be collected in the central database is already proven. The research is based on collected data from sensors installed in the pilot group of homes, part of EU REMOURBAN project. In addition, a smart radiator rig is designed to test the control strategy. Figure 7.1 shows a schematic overview of a home with sensors, smart meters, smart radiators and data storage. The data storage is cloud based and is connected to the prediction algorithm for generating the heating schedule. This schedule is used to generate the control signals for the energy centre.

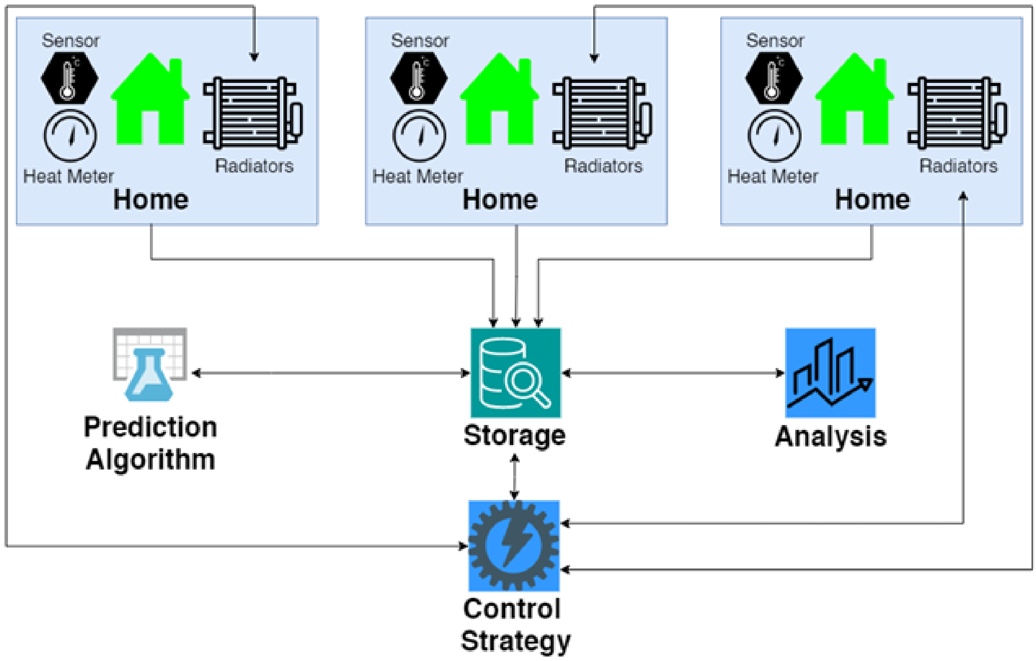


Fig. 7.1 Novel Framework based on IoT framework

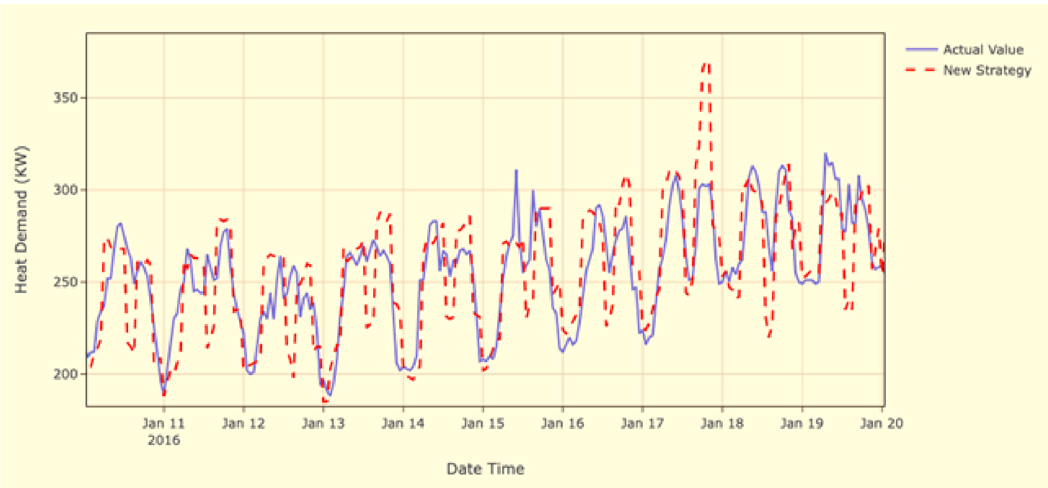


Fig. 7.2 Regenerated heat demand prediction of a month [6]

7.2 Heat Estimation Derivation

$$Q_{\text{Total}} = Q_{\text{Losses}} + \sum_{i=1}^{n} Q_{\text{Home}} \tag{7.1}$$

The normal approach to predict heat demand is to use the heat consumption data of the past and built the associated prediction model. The regeneration of the prediction from previous work ([6]) is shown in figure 7.2. The novel approach proposed here is to introduce the laws of thermodynamics for space heating. To optimise district heating systems, the heat generated should be close to the heat consumed. The total heat generation can be calculated by the heat consumed, and the heat lost during the transfer of energy. The heat lost is consistent in the network, whereas heat consumed for space heating depends on the customer's behaviour as well as the type of space needed to be heated. For this paper, only residential properties are considered. The total heat required for the residential property scheme in District Heating System can be given by equation 7.1.

$$Q_{\text{Home}} = \text{UA}\left(T_{\text{Indoor}} - T_{\text{Outdoor}}\right) \tag{7.2}$$

where $Q_{\textbf{Total}}$, is the heat required to be generated which is dependent on the heat losses occurring in the pipe while transferring from source to destination and the heat consumed by individual homes. Heat losses are given by $Q_{\textbf{Losses}}$ which is straight forward to calculate and it is consistent. Therefore, losses are not the focus of this research. The main focus is the heat estimation of $Q_{\textbf{Home}}$ which is the heat consumed by each residential home. The heat consumed by an individual home can be considered as the heat loss in the actual home which is written in the form of equation 7.2.

where **U** stands for U-value, which is dependent on many factors like the wall insulation, wall type etc. Different U-values are shared in the background and the related work section. **A** represents the surface area in contact with the external environment. Temperature inside the residential home is represented by $T_{\textbf{Indoor}}$ whereas the outdoor temperature is $T_{\textbf{Outdoor}}$.

Heat demand estimation can be carried out by using equation 7.2. The **UA** values are considered to be known. The indoor temperature can be forecasted using the data-driven models while outdoor temperature can be forecasted using methods like Monte- Carlo simulation. This research work uses the weather API, developed in this research for forecasted values. Considering the above discussion, the equation 7.2 is transformed into equation 7.3, where the forecasted indoor temperature is represented by $T_{\textbf{Predict}}$ and outdoor temperature by $T_{\textbf{Weather(predict)}}$.

$$Q_{\mathrm{Home}} = \mathrm{UA}\left(T_{\mathrm{Predict}} - T_{\mathrm{Weather(predict)}}\right) \tag{7.3}$$

The research novelty is that by using micro-level details to estimate heat demand. In other words, the actual home can be divided into the individual rooms that can be given by an equation, describing the heat consumed in each room.

$$Q_{\mathrm{Home}} = \sum_{j=1}^{n} Q_{\mathrm{Room}} \tag{7.4}$$

The typical heat consumption in sample home is shown in equation 7.5. The residence property has two main elements, a living room and a bedroom. $Q_{\textbf{Home}_{\textbf{Example}}}$ is the heat consumed by the sample home, which is equal to the summation of the heat consumed in living room ($Q_{\textbf{Livingroom}}$) and the heat consumed in the bedroom ($Q_{\textbf{Bedroom}}$).

$$Q_{\mathrm{Home}_{\mathrm{Example}}} = Q_{\mathrm{Livingroom}} + Q_{\mathrm{Bedroom}} \tag{7.5}$$

The individual room would have different **U**-value as **A** is the area exposed to the external environment. Moreover, the indoor temperature would vary based on the type of room. Data-driven model can be built for any room. The extended equation for sample home is written in equation 7.6.

$$\begin{aligned} Q_{\text{Home}_{\text{Example}}} = U_{\text{Livingroom}} A_{\text{Livingroom}} \\ \left(T_{\text{Livingroom(Predict)}} - T_{\text{Weather(Predict)}}\right) \\ + U_{\text{Bedroom}} A_{\text{Bedroom}} \\ \left(T_{\text{Bedroom(Predict)}} - T_{\text{Weather(Predict)}}\right) \end{aligned} \tag{7.6}$$

where $U_{\textbf{Livingroom}}$ and $U_{\textbf{Bedroom}}$ are U-value of livingroom and bedroom respectively. $A_{\textbf{Livingroom}}$ and $A_{\textbf{Bedroom}}$ represents the area exposed to the external environment. Forecasted temperature for living room and bedroom is given by $T_{\textbf{Livingroom(Predict)}}$ and $T_{\textbf{Bedroom(Predict)}}$ respectively.

Equation 7.7 shows the tranformation of the sample home energy equation into a generalised form for the whole system. This equation can be used to forecast the heat required by the customer. Data-driven model accuracy and output are discussed in the results and discussion section.

$$Q_{\text{Home}} = \sum_{i=1}^{n} U_i A_i \left(T_{i\text{(Predict)}} - T_{\text{Weather(Predict)}}\right) \tag{7.7}$$

7.3 User Profiling

The traditional approach is to classify customers based on the annual heat consumption. Figure 7.3 replicates the consumption of users by using data from previous work. User can be classified as very low, low, medium and high energy user. The problem with this approach is that the estimation of expense is not accurate and is dependent on the customer's sending the meter reading regularly. Moreover, the classification is a board, which does not allow the company to predict the customer requirement accurately.

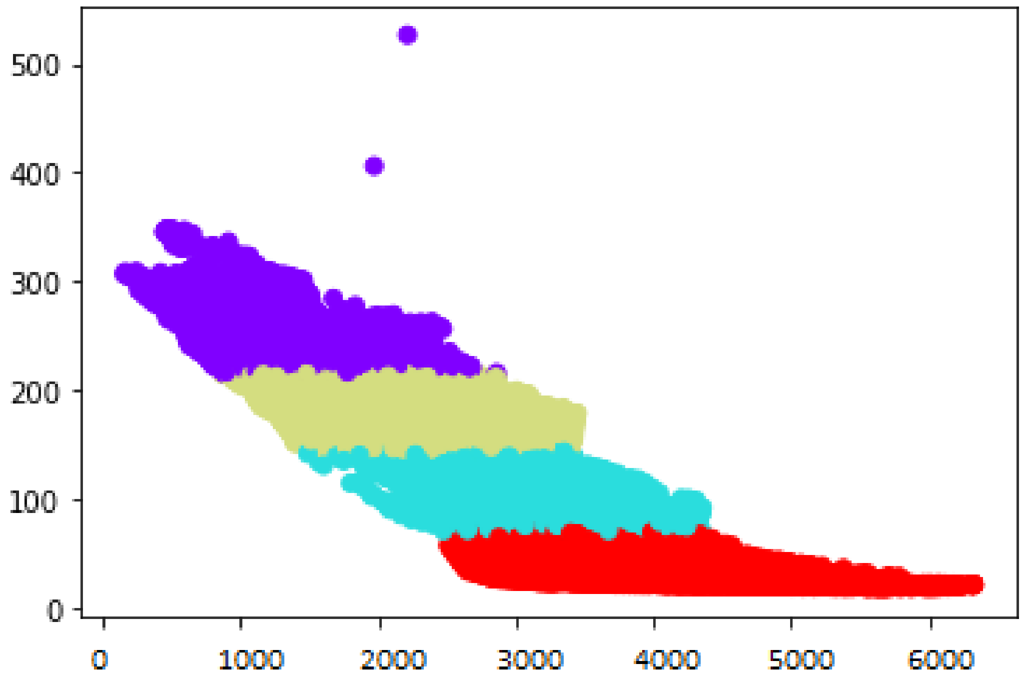


Fig. 7.3 Classification of user based on the consumption per Annum

The novel approach taken for this research work is with the assumption that each user is unique in his/her consumption. The customer socio-economical condition more or less drives the use of energy in his/her home. Also, it is a fact, that the customer socio-economic condition could change over time. User profiling is carried out using a machine-learning algorithm to build the model of the heat, using the parameters collected from each home. The model is built on the individual characteristics. The model adapts to the changes of the customer behaviour as the data collected in real-time are used to update the model regularly.

7.4 Results and Discussion

7.4.1 Data Analysis

As mentioned before the data used for the research is from the REMOURBAN project. The data collected from the smart meter are not used for analysis for this work but is used for proof of concept later. Multiple sensors are installed in each home, out of which all are not used. Moreover, 10 homes data out of 40 homes are selected because sensors were installed in the bedroom and the living room, which plays a vital role in this research.

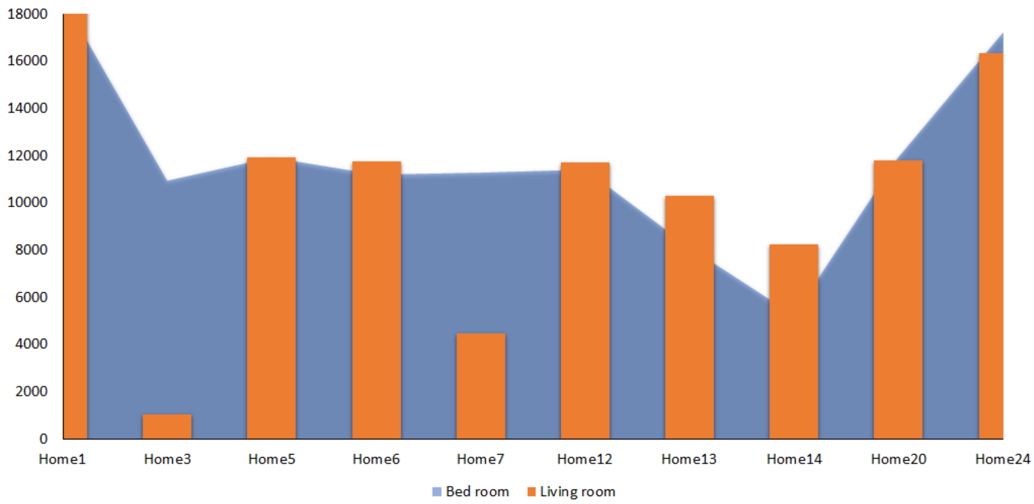


Fig. 7.4 Amount of data collected from living room and bedroom

Figure 7.4 displays the data collected from the 10 homes. Data analysis shows that the data collected are not consistent. So, the data from home 3, home 7, are ignored due to the large discrepancy with the values of the other homes. Home 13 is removed due to the lower level accuracy as we aimed to achieve at least 80 % accuracy of the trained model. This means data of home 1, home 5, home 6, home 12, home 20 and home 24 are used to test the proposed theory.

7.4.2 Evaluation of Heat Estimation

The derivation of heat estimation shows that the prediction of temperature is vital. The U-value of REMOURBAN project is available in the project report as well as in the [205]. Moreover, the area can be calculated from the floor plan. Both U-value([205]) and Area are constant.

The background study shows commonly used prediction models are built using machine learning algorithms. The decision tree algorithm is used to train the model using [16] as this algorithm achieved the best results.

The average accuracy of all the models trained is more than 80 %. The accuracies are plotted in figure 7.5. Living room model accuracy and bedroom model accuracy are calculated and plotted. As mentioned in the data section only 6 Homes are plotted.

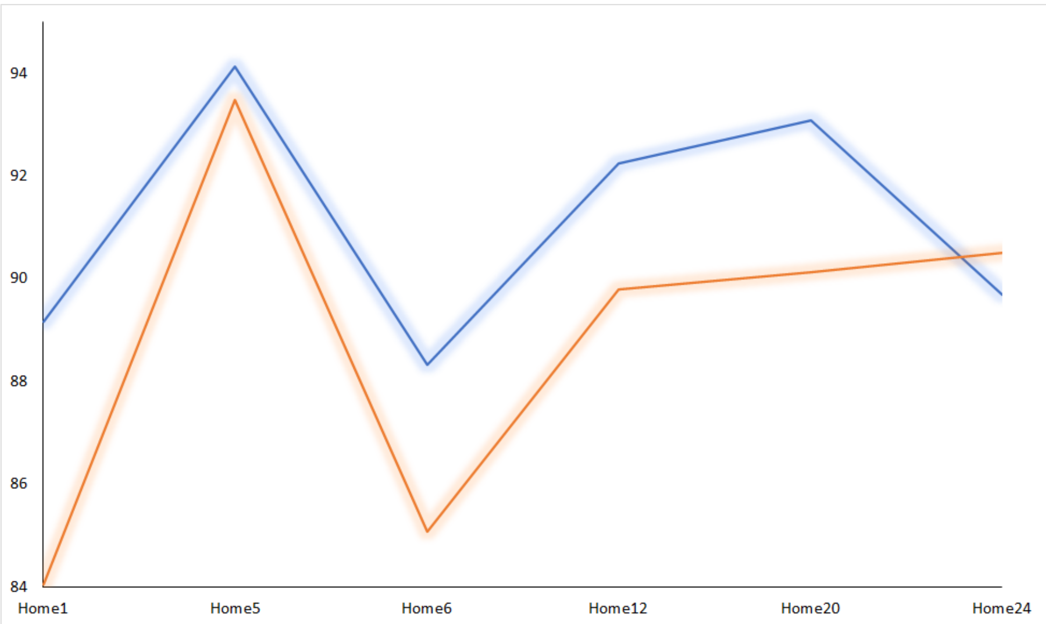


Fig. 7.5 Accuracy of individual rooms

The output of a typical room is displayed in 7.6. The blue line rep- resents the actual temperature in the room of a particular home whereas the green line represents forecasted temperature. Figure 7.6 shows clearly that the actual temperature is close to the forecasted temperature as fewer blue lines can be seen.

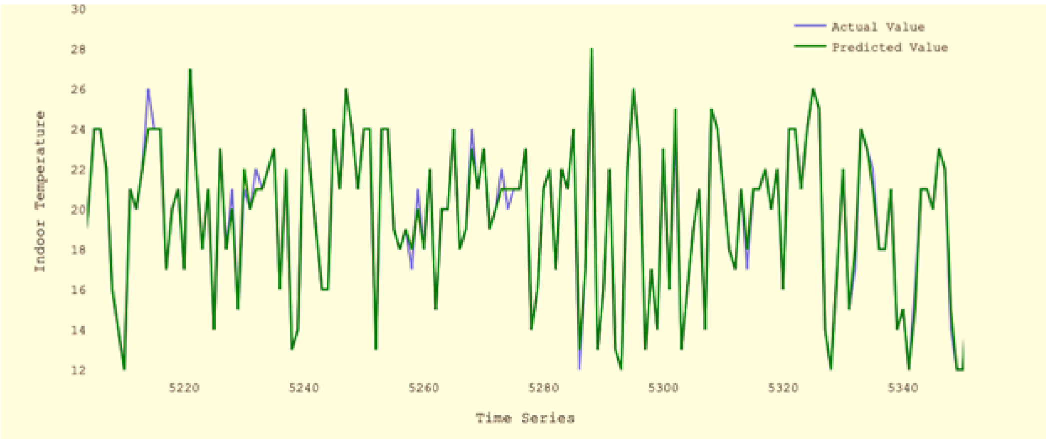


Fig. 7.6 Temperature prediction comparison with actual temperature of single room

From figure 7.5 and figure 7.6, it is clear that the temperature which plays a vital role for heat estimation is calculated quite good accuracy. This figure also, validates the estimated heat.

7.4.3 Evaluation of User Profiling

Heat consumption is dependent on the external environment and the social behaviour of the customer. The model was built using external weather data and internal environment data collected. Combining the data gives the profile for each home. The temperature profile of all the individual rooms are shown in figure 7.7. It is easy to see different temperature profiles despite kiosks which illustrate that each customer is unique.

For discussion two home data in winter are selected as shown in figure 7.8. The different temperature profiles can be clearly identified. In home 24 the desired temperature is higher than home 1. In home 24 the desired temperature is typical for a residential home while in home 1 the temperature is clearly too low. This means that in home 1 is the problem is either in the data collected

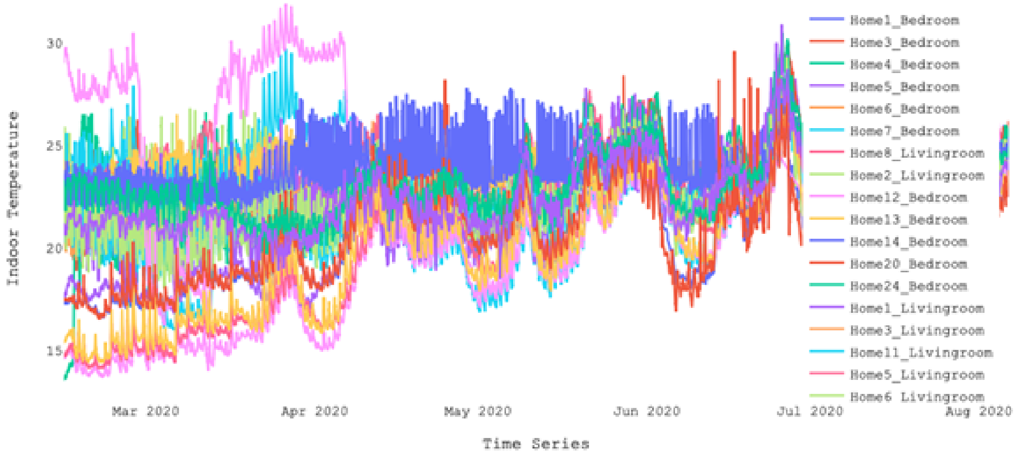


Fig. 7.7 Overview of temperature profile of user

or in the customer behaviour. The overall check shows that the sensors were working fine, which means the customer was not turning the heating on due to some reasons, mostly economic. This is an example of fuel poverty.

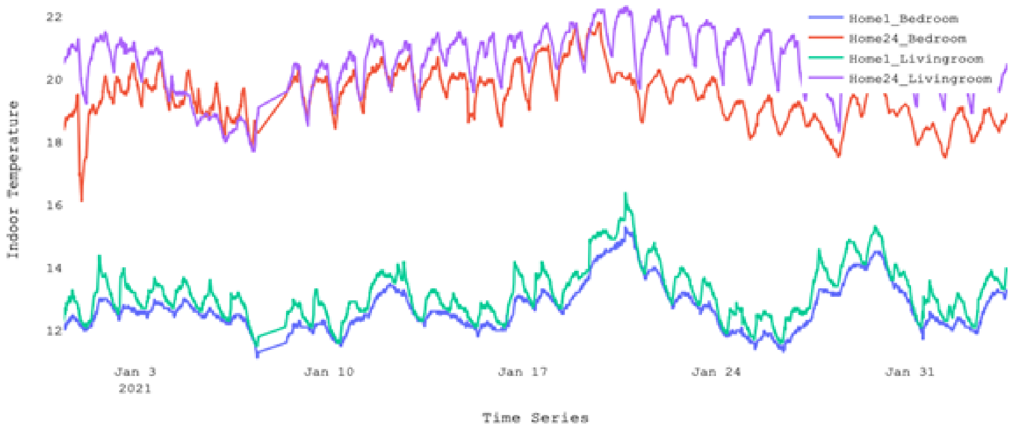
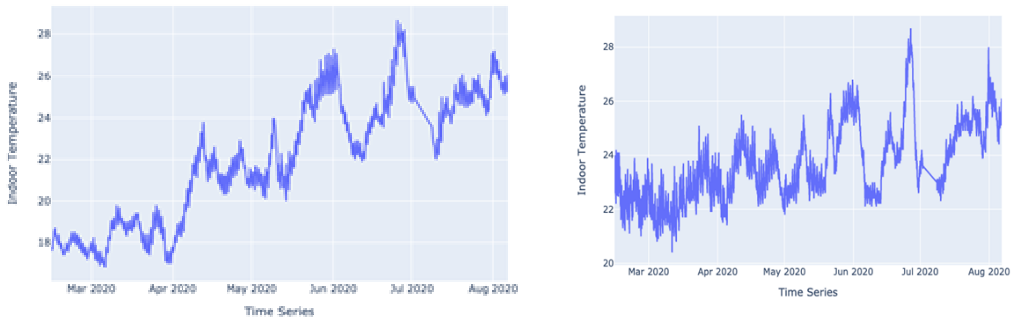


Fig. 7.8 Home 24 and Home 1 Temperature profile

Customer behaviour is not driven only by economic conditions but is also dependent on personal comfort. To display the effect of personal comfort by the desired temperature, two homes' living room temperatures are plotted in figure 7.9.

The temperature desired by the first home in May is about 18 °C whereas in the second home is 22 °C. During winter both home desired temperatures are the



(a) Example 1 living room home (b) Example 2 living room home

Fig. 7.9 Comparison of two home

same. This analysis shows that desired comfort for the second home customer is at a higher temperature than the first home. Moreover, the conclusion drawn is that second homes may use heating during the spring.

7.5 Summary of Chapter

This chapter delves into two methods for forecasting heat demand. The conventional approach involves utilising past heat usage data to construct a prediction model, whereas the innovative method proposed in this study involves applying the laws of thermodynamics with power of AI to space heating. In order to optimise district heating systems, it is imperative that the total heat generation closely aligns with the heat consumed. The new approach takes into account micro-level particulars to gauge heat demand, recognizing that each user has a unique consumption pattern that is influenced by their socio-economic status. The heat consumed in each room can be estimated through a mathematical equation, which accounts for individual differences in heating requirements. The traditional method categorises customers according to their annual heat consumption, while the new approach uses socio-economic factors to create customer profiles, resulting in more precise energy requirement predictions.

Chapter 8

Smart Radiator

The test rig described in this chapter is a tool for developing and validating control strategies for the systems. It includes an electrical circuit with a microcontroller or microprocessor based on the Arduino or Raspberry Pi platform to control valves, and other components, as well as sensors to measure humidity and temperature. A data acquisition system is included to collect data and a user interface is provided for monitoring and modifying the control strategy. The test rig also includes a simulator for testing the control strategy in a virtual environment. The valve system consists of a solenoid valve and a pump to control fluid flow. The test rig is designed to be easy to use, reliable, and flexible, and comes with multiple software packages for customising control strategies. Safety features are also included for user protection.

8.1 Introduction

The test rig was added to the thesis project as an unexpected result of the alterations to the research plan brought about by the Covid-19 pandemic. The rig was designed to mimic a real-life home environment and was equipped with

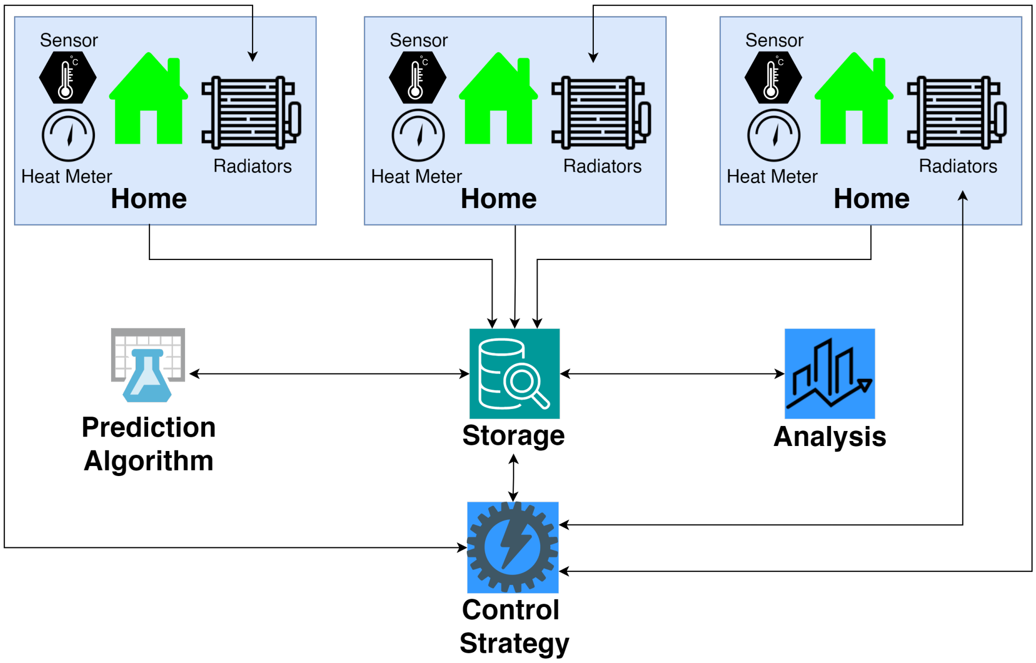


Fig. 8.1 Framework implementation in Lab Setup

the latest technology to support the research, including temperature, humidity, occupancy, and Total Volatile Organic Compound (TVOC) sensors.

One of the innovative aspects of the research was the installation of a smart heat meter within the home, instead of at a substation. This allowed for direct access to all the necessary parameters, which could be used for prediction or actual calculation. To further validate the results, commercially available sensors were also selected and installed alongside the in-house developed circuit drivers for the solenoid valve and temperature-humidity sensors.

The collected data and algorithms used for analysis were stored in a cloud-based data storage system. This provided easy and secure access to the data from any location and allowed for visualisation and control of the data through a user-friendly interface. The larger system, as depicted in Figure 8.1, involved the use of sensors, smart meters, and smart radiators to gather data, which was analysed by a prediction algorithm to generate a heating schedule for the home. The smart radiator rig played a crucial role in testing the effectiveness and efficiency of the control strategy for regulating home heating.

To summarise, the smart radiator rig was an integral component of a larger system for testing control strategies for home heating systems. Its design and implementation allowed for accurate data collection and analysis, leading to improved and more efficient regulation of home heating. The use of cutting-edge technology, including cloud-based data storage and a user-friendly interface, made the research both convenient and accessible.

8.2 Test rig Setup

The test rig operates using both local and cloud-based intelligence. The local intelligence uses simple algorithms on microcontrollers such as the Arduino or Raspberry Pi to collect data and generate control signals. Manual control rules are processed locally while more complex tasks requiring more computational power are hosted on the cloud. The schedule generation is done on the cloud, based on data collected from the weather station, local site sensors, and demand predictions. The generated control signals are then sent to the local intelligence hub for implementation (Figure 8.2).

The test rig involves converting normal home radiators into smart radiators. The design and development of these smart radiators is explained in a later section. The radiators have both temperature and humidity sensors, as well as valves. The data from the sensors is collected locally and sent to the cloud regularly. The valve's status and control are also communicated with the cloud on a regular basis, and all actions are logged in a database for record-keeping purposes.

Automatic control using a schedule is straightforward, but manual or local control (Figure 8.9) is more complex as it involves both local networking and

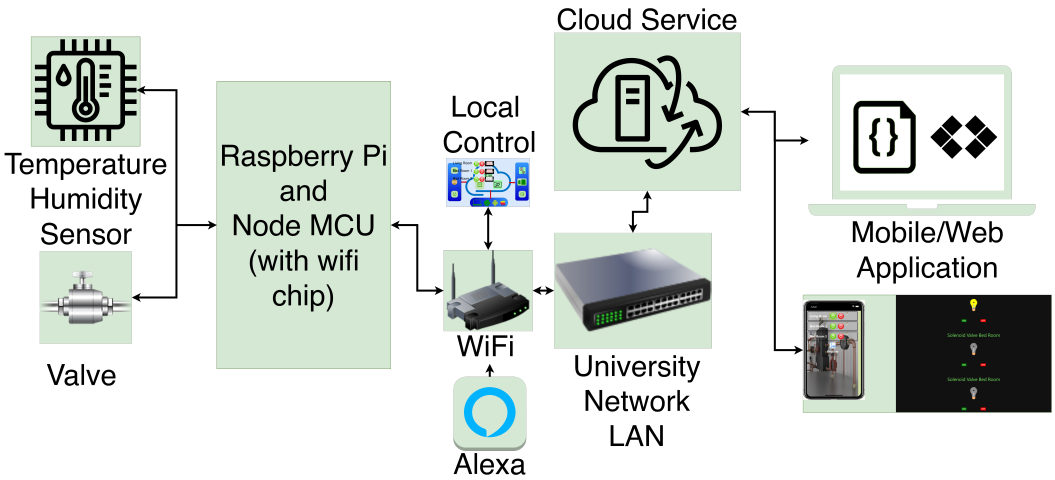


Fig. 8.2 Block View of Lab Setup

internet connectivity to control from a remote location. The test rig offers four different methods of controlling the valve, two of which are local and two are global. The first method is a local display with a touch screen, similar to Nest or a programmable logic controller (PLC), with a simple user interface. The second method is using a local speech assistant, such as Alexa, although this can be any other speech assistant. Although it is considered local, the control signals are sent through the cloud, so it is considered a global method.

The chapter later examines and assesses the control strategy, while the block diagram below provides an overview of data exchange. The microcontroller or microprocessor connects the sensors and valve, which then connect to a cloud service through the local network via the internet. Alexa and local control are connected to the local network, while a mobile app can be downloaded from the Apple Store or Play Store to communicate with the test rig through the cloud service using the internet. Additionally, the web app can be used similarly to a standard website.

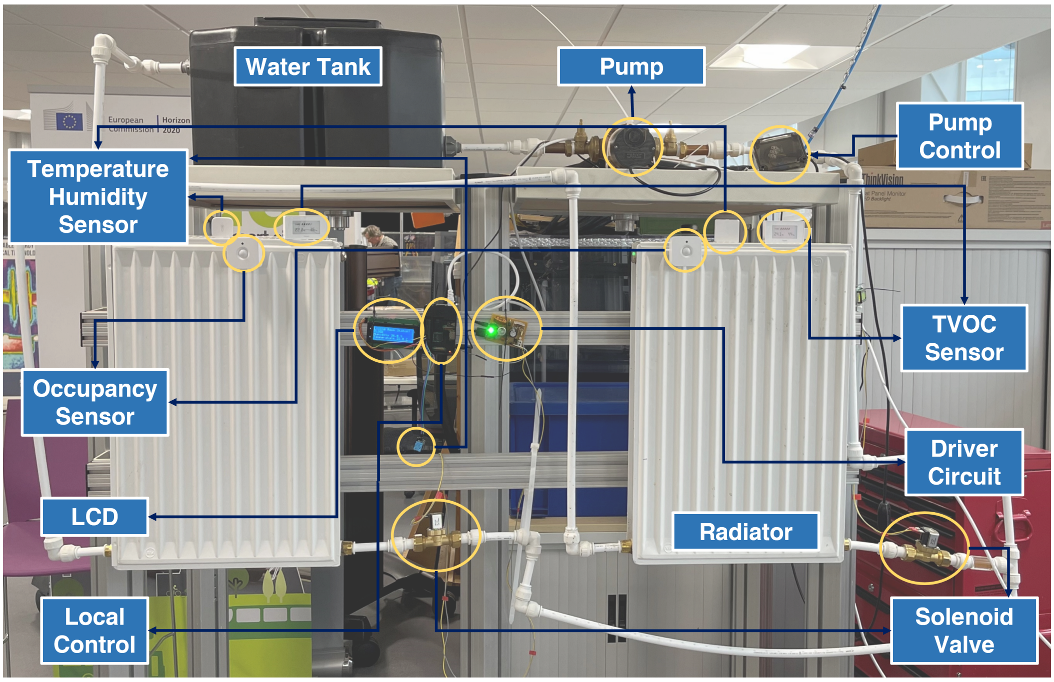


Fig. 8.3 Tag components of Test-rig

8.2.1 Test rig Design

The given text describes a test rig that replicates the residential heating system of a home. The test rig has been designed to accommodate two radiators, simulating those usually installed in two separate bedrooms of a house. In order to transform the traditional radiators into smart radiators, electro-mechanical solenoid valves have replaced the old radiators (Figure 8.3). The functioning of the solenoid valves is based on their ability to control the flow of hot water to the radiator. To operate the valves at the desired level, a separate driver circuit was designed, which would provide the necessary current and voltage. The driver circuit has been designed to turn the valve on/off, depending on the heating requirements of the room, thereby maintaining the desired temperature of the bedroom.

The process of designing the driver circuit began with the creation of a prototype that was tested iteratively to identify and rectify any errors. Once the prototype was successful, a miniature version was created that could be

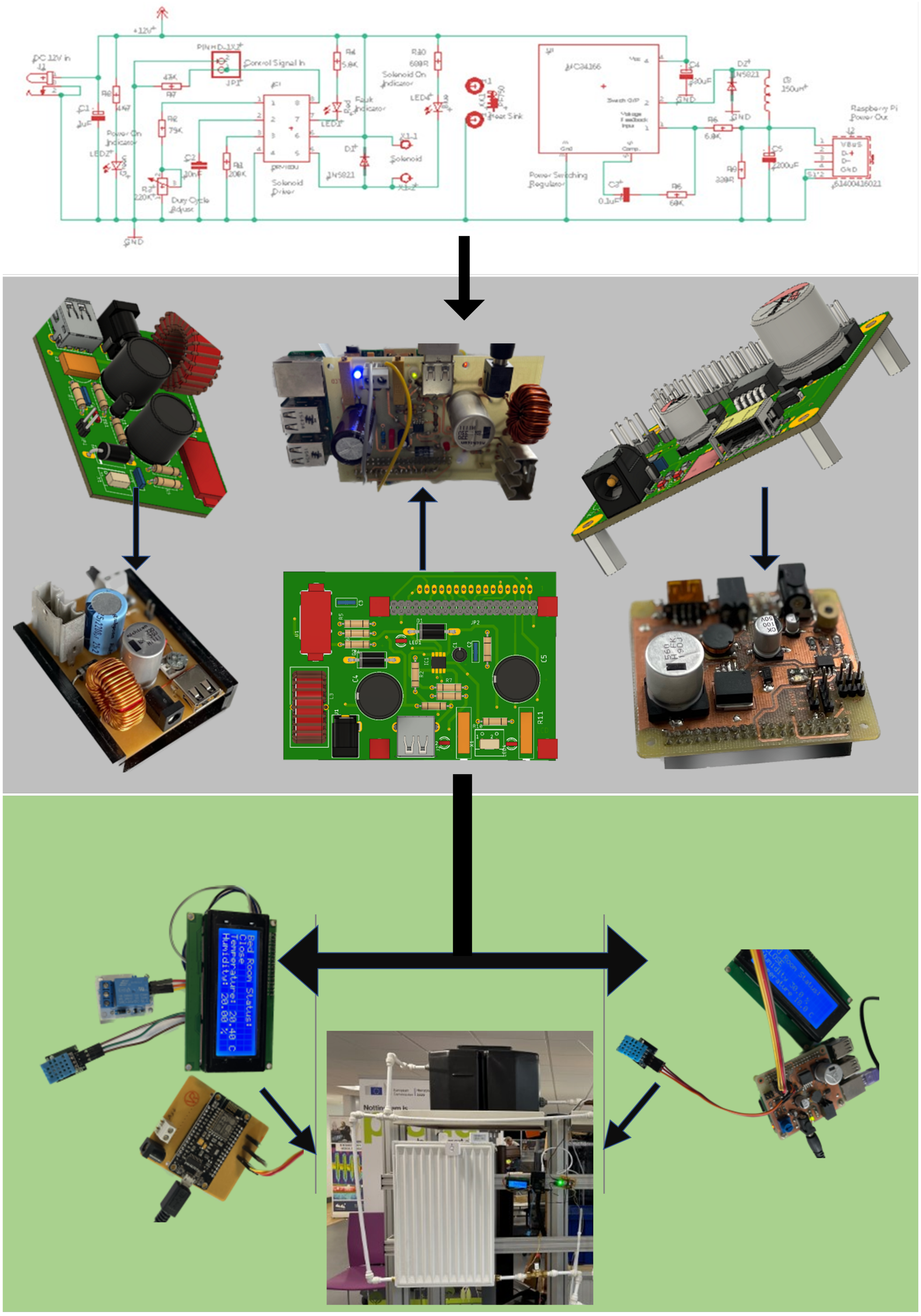


Fig. 8.4 Journey of Test-rig

easily mounted on the valve. The transition from computer-aided design (CAD) to the actual driver circuit has been shown in the accompanying figure 8.4.

The design and development of such a test rig, which replicates the actual conditions of a residential heating system, can be helpful in studying the performance and efficiency of smart heating systems. The rig can be used to test and evaluate the behavior of the system in different scenarios, and any necessary improvements or modifications can be made based on the results obtained.

The final version of the driver circuit is capable of powering both the microprocessor and microcontroller, thereby enabling the use of a single power source. Additionally, it can power three different sensors, which can operate on 5 volts and 1 ampere. For the purpose of the test rig, a temperature and humidity sensor from Adafruit was chosen, along with a PIR sensor for occupancy detection and an LDR sensor connected to the circuit board.

Although the driver circuit is compatible with both the microprocessor and microcontroller, the initial version of the circuit board was designed to work with the microcontroller and relay, with the primary goal of controlling the radiator without sensors. However, as the iterations continued and improvements were made to the driver circuit, a more advanced version was developed, which incorporated the microcontroller and other additional capabilities, resulting in a single circuit board design.

The ability to power the microcontroller and microprocessor using a single power source, along with the ability to integrate multiple sensors into the driver circuit, demonstrates the versatility of the design and its potential for use in various smart heating systems.

The development of the driver circuit has resulted in the creation of two final circuits. The first circuit uses a microcontroller, while the second uses a microprocessor. To achieve wireless connectivity, off-the-shelf components were used, such as the NodeMCU for the microcontroller and Raspberry Pi for the microprocessor.

The Raspberry Pi has greater capabilities than the NodeMCU, and thus provides a higher level of local intelligence. In the event of lost internet connectivity, the Raspberry Pi can store data, whereas the NodeMCU cannot due to limited memory capacity.

In addition, the LCD screen has been incorporated into the design to display information such as the open/closed status of the valve, temperature, and humidity. All circuits have been built using open source platforms and codes, making them highly reproducible and easy to customise. Furthermore, cost-effective components have been used in the development process.

The use of off-the-shelf components and open-source platforms has led to the creation of highly adaptable and customisable systems. The incorporation of wireless connectivity and data storage capabilities has made the smart heating system more versatile and reliable, while the use of cost-effective components makes the technology more accessible to a wider audience.

8.3 Control Strategy

The control strategy is inspired by the smart home considered in [2]. The idea is to control the radiators like bulbs. An experimental rig is set up in the lab by designing our own smart valve like a smart thermostatic valve, for building and developing open-source technology. The reason for using a smart valve is to take advantage of data collected from the various sensors set in the monitored homes.

Control strategy has two modes, automatic and manual mode. Any mode can be selected by the user from the user-built interface (Figure 8.5).

In the automatic mode, the scheduler services built in the Chapter 6 is called to generate the heating schedule. This schedule generates control signals to operate the smart radiator. This schedule then takes into consideration the motion in the room and depending on that it turns on/off the radiator.

In the manual mode, the decision is in the hand of the user. The decision can be made using a mobile application or web application. Users in manual mode can control the radiators of all rooms individually. Manual mode bypasses the automatic schedule and sensor reading to turn the radiator on/off.

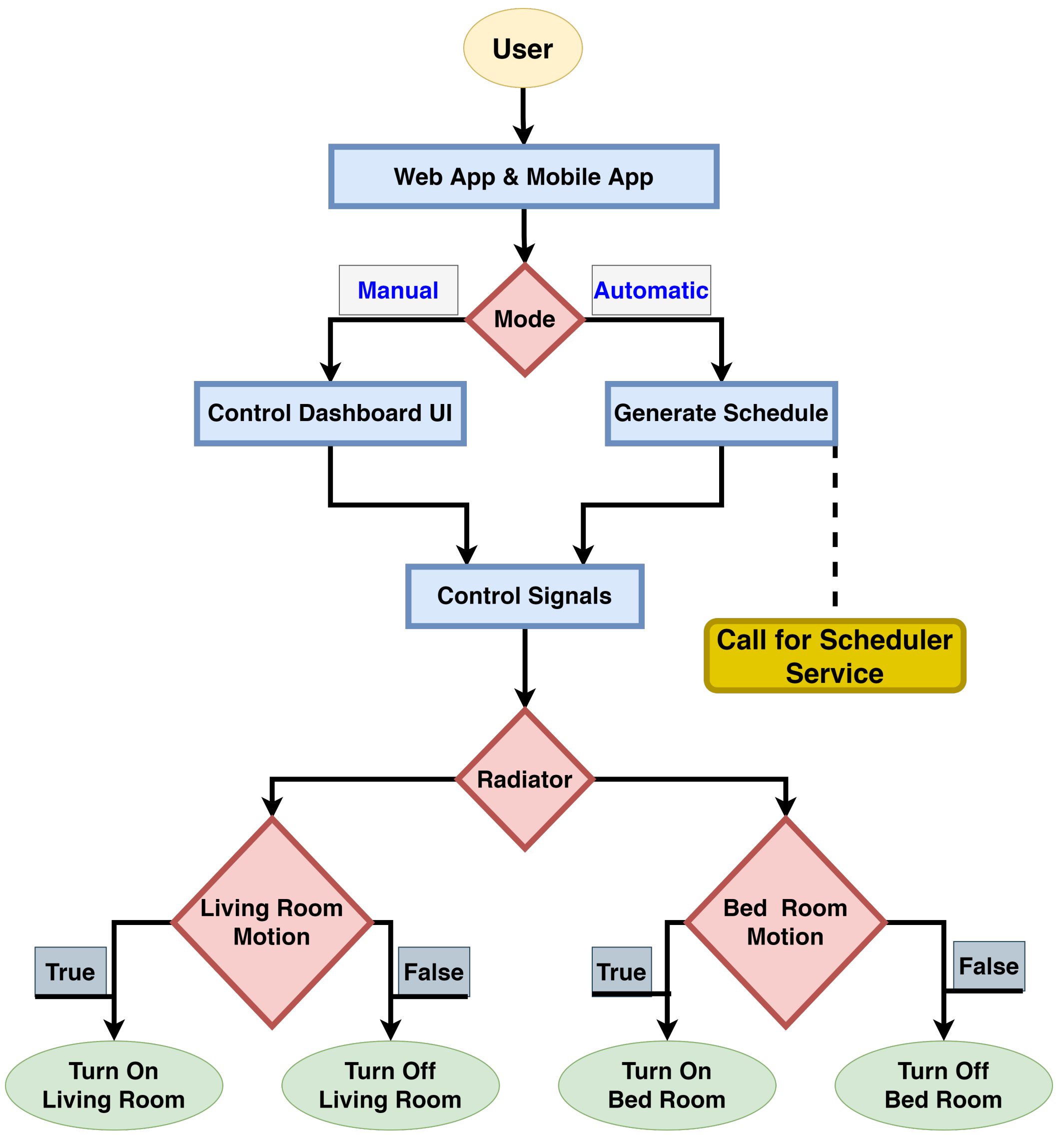


Fig. 8.5 Flow chart for control strategy for individual home

8.4 Evaluation of Control Strategy

In this section, the reason for control strategy theory is discussed with dashboard built. The possible cost-saving after implementing the strategy is simulated.

The control strategy is outlined in the theory section. In figure 8.6 the motion sensor data and temperature data versus timestamp for January month are plotted. The area of interest is highlighted to understand the logic of control strategy by black outline in the figure. The example shown has the same temperature in the room to show the effect of the motion. The heating is turned on in both rooms even though there is no motion in the living room. This means that the living room is heated without anyone using it. These can be avoided by using the control strategy derived in the theory section.

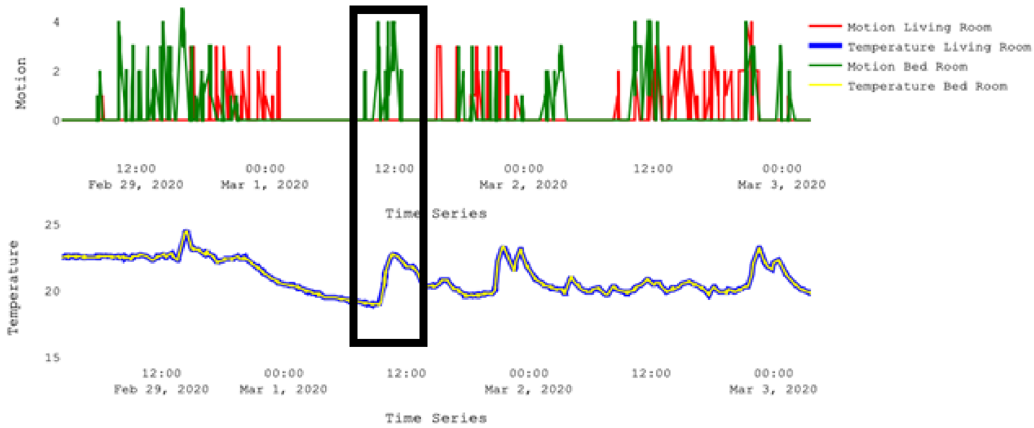


Fig. 8.6 Winter Month temperature and motion

To evaluate the effect of the control strategy the cost-saving is calculated. The price required for the calculation is drawn from the secondary data (Chapter 2 secondary data section). The temperature is converted in the heat losses as U-value and the area is known as part of the REMOURBAN project. The calculation of a winter day with and without control strategy is illustrated in figure 8.7. The constraint of the calculation is that the mode used is automatic.

Analysis of the chart shows that the cost after using control strategy is roughly 50 % on the typical winter day. Even if the heat losses increase due to the thermal bridge with the neighboring room the saving of cost expected is 25 %.

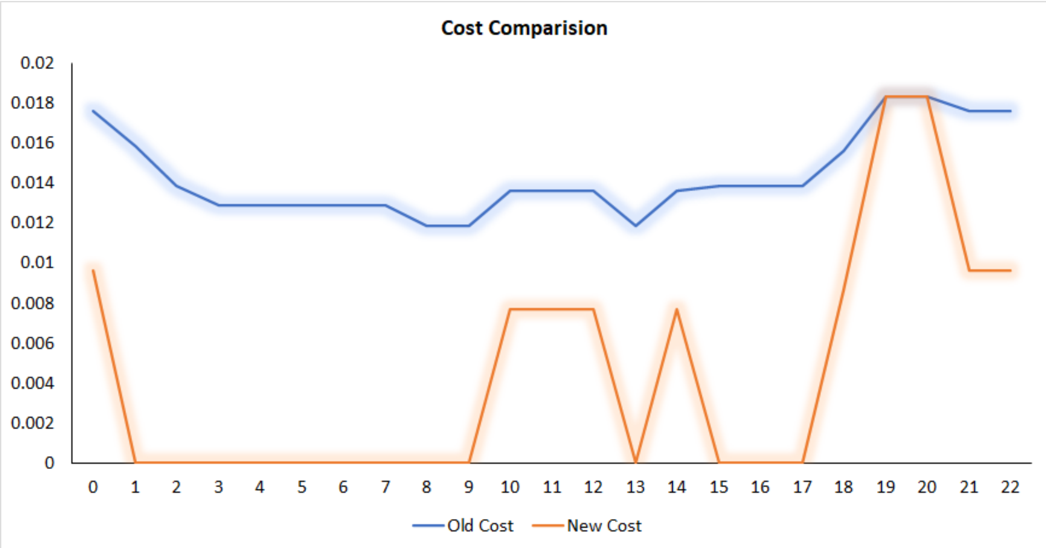


Fig. 8.7 Comparison with and without control strategy

Mobile application and web application screen for a smart home are shown in the figure 8.8. Currently, only On/off functionality is available on both applications. To keep the user interface similar same buttons and names are used. In the mobile application (Figure 8.10) speech recognition is available for the customer to control the radiator. The dashboard is available anywhere and anytime in the world.

8.5 Summary of Chapter

The chapter describes a test rig designed to develop and validate control strategies for home heating systems. The rig includes sensors to measure humidity and temperature, a data acquisition system, and a user interface to monitor and modify the control strategy. The rig is designed to be easy to use, reliable, and flexible, and comes with multiple software packages for

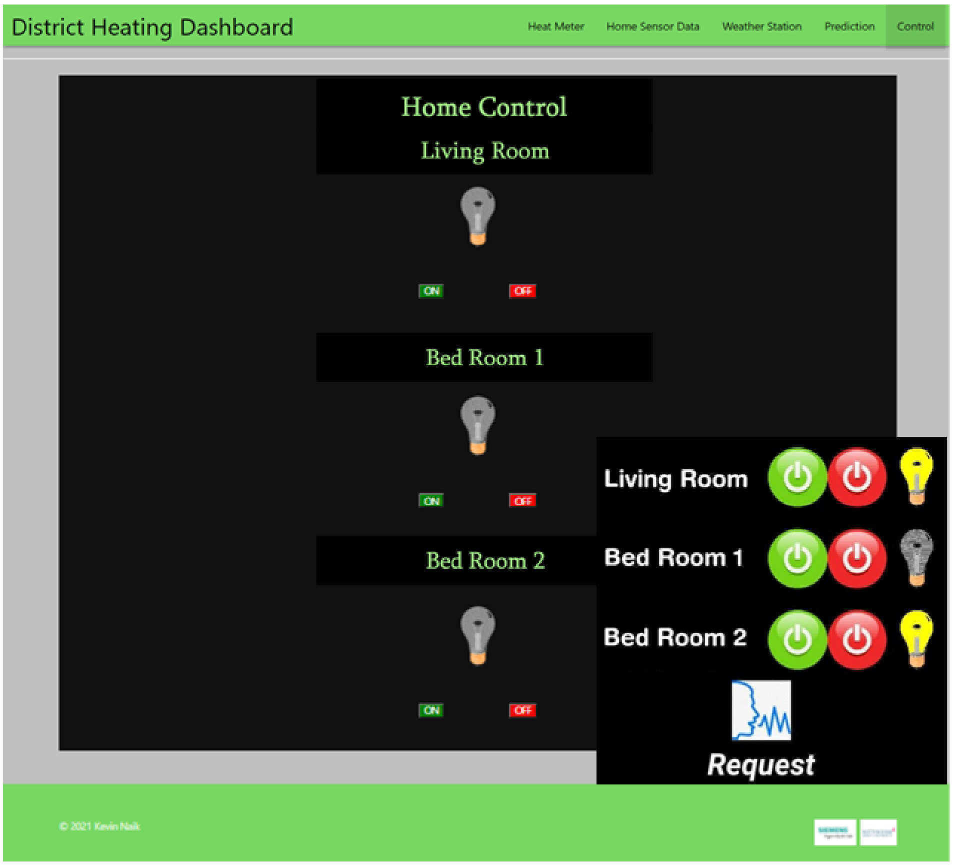


Fig. 8.8 Web application and mobile application

Fig. 8.9 Local Controller

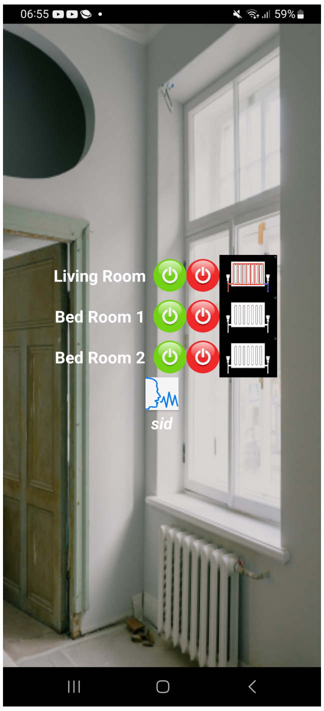


Fig. 8.10 Updated mobile update

customizing control strategies. The test rig also includes a smart radiator system that can be controlled locally or globally through a touch screen, speech assistant, mobile app, or web app. The chapter details the design of the smart radiator system, which involves replacing traditional radiators with electro-mechanical solenoid valves controlled by a driver circuit designed to maintain the desired temperature of a room. The test rig is an integral component of a larger system for testing control strategies for home heating systems, and its use of cutting-edge technology, including cloud-based data storage and a user-friendly interface, makes the research both convenient and accessible.

Chapter 9

Discussions and Conclusion

9.1 Discussions

This thesis focuses on the development of data-driven heat prediction algorithms for the "Sustainable Energy Management System" (SEMS), which is a multi-energy system optimisation software. SEMS was created through a collaboration between Siemens, the Greater London Authority, and the University of Nottingham, Nottingham Trent University, and Imperial College London. The journey of this thesis begins with a detailed exploration of SEMS and its associated algorithms. The software was developed for application at a social housing estate in Greenwich, London, as part of the Greater London Authority's efforts to retrofit the energy systems and building fabric of its housing stock. Its purpose is to balance energy across vectors and networks through day-ahead forecasting and optimisations that can be interpreted as control outputs for energy plant such as a water source heat pump, district heating pumps and values, power switchgear, gas boilers, a thermal store, electric vehicle chargers and a photovoltaic array. The optimisation objectives are to minimise greenhouse gas emissions and operational cost.

The tool uses Hypernetwork Theory based orchestration coupled with a microservice architecture. The distributed nature of the design ensures flexibility and scalability. Currently, microservices have been programmed to forecast domestic heating demand, domestic electricity demand, electric vehicle demand, solar photovoltaic generation, ground temperature, and to run a day-ahead energy balance optimisation. This thesis presents the results from both domestic heat and electricity demand forecasting, as well as the overall design and integration of the software with a physical system.

The works build on that of O'Dwyer, et al. (2020) who developed a preliminary energy management software and digital twin. Their work acts as a foundation for this real-world commercialisation-ready program that integrates with physical assets.

The Fourth Generation Low Temperature District Heating (LTDH) is highly efficient, but it demands precise heat generation control, thus accurate heat demand prediction is critical.

This thesis presents a data-driven approach to heat demand prediction for LTDH systems, but it can also be beneficial for older district heating systems. The heat demand prediction for residential buildings was studied using both simulated and collected data. The results showed that different algorithms performed better for the two different types of data. Linear Regression was better for simulated data, while Decision Tree was more accurate for real collected data. Outdoor temperature was found to have a major impact on heat demand prediction. The use of weather features could enhance the performance of the algorithms and improve heat demand prediction. Although Artificial Neural Network has performed well in previous studies, it showed disappointing results with the small amount of data collected in this study. Its accuracy could be improved with more data points and features. This study also confirmed that Decision Tree is highly accurate for datasets with smaller amounts of data.

This approach can also be applied to predict short-term or long-term heat demand using weather forecast data.

The heat prediction algorithms, which were rigorously tried and tested in the SEMS and REMOURBAN projects, served as the foundation for the development of an advanced IoT framework. This framework allows multiple heat prediction algorithms to run concurrently, enabling the system to choose the best algorithm based on real-time data analysis and optimisation.

This thesis aims to propose an IoT framework for district heating. The optimisation of district heating assets like boilers and heat pumps is crucial to reducing carbon emissions. The framework is driven by heat demand forecasting, with the use of real-time data-driven heat prediction to adjust control strategies as needed. The challenge in controlling district heating systems lies in their complexity due to a long history of control algorithms. To tackle this issue, the proposed framework tests and evaluates multiple machine learning algorithms for heat demand forecasting.

The results of the case study in the thesis shows that the developed Internet of Things Framework for district heating systems is effective in optimising the assets and controlling the heat demand. The framework successfully collects real-time data, uses historical data to make accurate heat demand predictions, and runs multiple machine learning algorithms simultaneously to achieve the best forecasted value. The implementation of the framework has shown good accuracy in the prediction of real-time heat demand for a range of one to five days. The real-time heat demand is displayed on a dashboard and can be used to make informed decisions about the control of district heating assets. Overall, the Internet of Things Framework provides a smart and efficient solution for optimising district heating systems and reducing their carbon footprint.

In conclusion, the implementation of the Internet of Things Framework for district heating can significantly improve the efficiency and sustainability of district heating systems. The real-time data-driven heat demand prediction and multiple machine learning algorithms running simultaneously provide reliable, flexible, and scalable solutions for the optimisation of district heating assets. The ability to collect and analyse customer usage patterns allows for the development of new control strategies and improved heat generation optimisation. Furthermore, being an open-source framework, it opens the opportunities for further research and innovation.

This thesis presents a novel Internet of Things Framework for district heating, which provides a real-time heat demand prediction. By using a micro-level and user-centric data-driven model, the framework has shown to have the potential to improve the efficiency of heat generation and provide more accurate predictions. The implementation of multiple machine learning algorithms at the same time ensures reliability and scalability. The open-source nature of the framework opens up several opportunities for future research and provides a solid foundation for further development and optimisation.

The heat prediction algorithms, which were rigorously tried and tested in the SEMS and REMOURBAN projects, served as the foundation for the development of an advanced IoT framework. This framework allows multiple heat prediction algorithms to run concurrently, enabling the system to choose the best algorithm based on real-time data analysis and optimisation.

Overall, the micro-level and user-centric data-driven approach introduced in this research has great potential to improve the efficiency of district heating systems. The results of the case study in REMOURBAN project demonstrate the effectiveness of the framework in predicting heat demand and optimising heat generation.

The results of the lab tests show that the smart radiator system can accurately control the temperature of a room, thus helping to reduce energy consumption and increase energy efficiency. The ability to profile user behaviour and adapt to changing customer behaviour is a key feature of the smart radiator system and can lead to further improvements in heat prediction and energy efficiency in the future. Overall, the research carried out in this thesis highlights the potential of the micro-level, user-centric approach to improve the efficiency of district heating systems and provides a valuable contribution to the field of energy management and sustainability.

The results of the implementation and testing show that the IoT framework for district heating is a successful solution to improve the efficiency and cost-effectiveness of the heat generation. The real-time heat demand prediction, user profiling, and smart control strategies have all contributed to the reduction in energy consumption and cost savings. The dashboard effectively displays the results and allows for further analysis and optimisation. The open-source and flexible nature of the framework opens up new opportunities for future research and development in the field of district heating.

The IoT framework developed in this study has great potential to improve the efficiency and cost savings of district heating systems by using real-time data and user-centric data-driven models. The implementation of the framework has shown promising results with a cost saving of 25-30% per year. The micro-level heat demand prediction using a fresh approach and user profiling has great potential to optimise heat consumption at individual homes. Overall, the human-centric approach in district heating systems can lead to significant improvement in the overall efficiency and cost savings of these systems.

9.2 Conclusion

This research will claim the following innovative research and contribution to knowledge:

- This thesis has made significant contributions to the field of district heating by developing a data-driven approach to predict heat demand in LTDH systems. The study of heat demand prediction for residential buildings using both simulated and collected data has provided new insights and paved the way for further advancements in the field. The findings of this research have the potential to inform future research and practical applications for improving the efficiency and sustainability of district heating systems.

- This thesis contributes a novel Internet of Things Framework for district heating, which enables real-time heat demand prediction and adjustment of control strategies. The framework leverages heat demand forecasting and real-time data-driven heat prediction to optimise the performance of district heating systems. Through this innovative approach, the framework offers reliable, flexible, and scalable solutions to improve the efficiency and sustainability of district heating assets.

- The case study conducted in the context of the REMOURBAN project has demonstrated that the implementation of the Internet of Things Framework for district heating can lead to significant improvements in the efficiency and sustainability of district heating systems. The findings indicate that the implementation of the framework has yielded promising results with a cost reduction of 25-30

- The lab tests conducted in this research demonstrate that the implementation of a smart radiator system can effectively regulate the temperature of a room and thus help to decrease energy usage and enhance energy efficiency.

- This thesis presents a human-centric approach to utilise real-life data for the creation of individual heat profiles of district heating users, resulting in a realistic individual heat demand for the heating system, and the establishment of an optimum heat generation mode. This approach is innovative and provides a practical solution for reducing energy waste and promoting energy efficiency.

9.3 Future work

Based on the lab tests, the smart radiator system has demonstrated its ability to effectively control the temperature of a room, resulting in reduced energy consumption and increased energy efficiency. Moving forward, efforts will focus on leveraging data from meters and sensors to improve heat prediction and energy efficiency.

One potential area for improvement is collecting more data to enhance heat prediction accuracy. This could involve utilising live data to dynamically parameterize predictive algorithms, enabling them to adapt to changing conditions and anomalies.

Moreover, the IoT framework tested in the lab can be extended to other areas such as control algorithms and prediction algorithms, as well as new architectures. This will introduce flexibility, scalability, and reliability to district heating systems, enabling plug-and-play algorithms that can be adjusted in response to consumer demand.

By analysing real-time consumer demand, it becomes possible to optimise heat generation and design new control strategies based on usage patterns. Heat estimation derivations can be thoroughly tested and evaluated on different architectural types and utility buildings, leading to further improvements in energy efficiency.

In summary, the future scope of the smart radiator system is broad and includes leveraging data, and extending the IoT framework to enhance energy efficiency in district heating systems.

References

[1] Henrik Lund, Sven Werner, Robin Wiltshire, Svend Svendsen, Jan Eric Thorsen, Frede Hvelplund, and Brian Vad Mathiesen. 4th generation district heating (4gdh): Integrating smart thermal grids into future sustainable energy systems. *Energy*, 68:1–11, 2014.

[2] Kevin Naik and Supriya Patel. An open source smart home management system based on iot. *Wireless Networks*, Nov 2018.

[3] Hugh Boyes, Bil Hallaq, Joe Cunningham, and Tim Watson. The industrial internet of things (iiot): An analysis framework. *Computers in Industry*, 101:1–12, 2018.

[4] Liz Waters. Special feature -experimental statistics on heat networks experimental statistics on heat networks introduction, 2021.

[5] Sachin Kumar, Prayag Tiwari, and Mikhail Zymbler. Internet of things is a revolutionary approach for future technology enhancement: a review. *Journal of Big Data*, 6(1):111, Dec 2019.

[6] Sean Jones, Richard Charlesworth, Kevin Naik, Thomas Charlesworth, Edward O'Dwyer, Anton Ianakiev, Jeff Johnson, Rabah Boukhanouf, Mark Gillott, Victor Sellwood, and Joy Aloor. A multi-energy system optimisation software for advanced process control using hypernetworks and a micro-service architecture. *Energy Reports*, 7:167–175, 2021. The 17th International Symposium on District Heating and Cooling.

[7] Arijit Sen and Amin Al-Habaibeh. The effect of covid-19 pandemic on household budget, energy demand and carbon emission of households in england due to working from home. 12 2020.

[8] Constantinos Loucari, Jonathon Taylor, Rokia Raslan, Eleni Oikonomou, and Anna Mavrogianni. Retrofit solutions for solid wall dwellings in england: The impact of uncertainty upon the energy performance gap. *Building Services Engineering Research and Technology*, 37(5):614–634, 2016.

[9] UKpower. Compare energy prices per kwh. https://www.ukpower.co.uk/home_energy/tariffs-per-unit-kwh, 2023. [Accessed 28-Feb-2023].

[10] Michael-Allan Millar, Neil M. Burnside, and Zhibin Yu. District heating challenges for the uk. *Energies*, 12(2), 2019.

[11] Fabian Levihn. Chp and heat pumps to balance renewable power production: Lessons from the district heating network in stockholm. *Energy*, 137:670–678, 2017.

[12] Christoph Molitor, Stephan Groß, Jakob Zeitz, and Antonello Monti. Mescos—a multienergy system cosimulator for city district energy systems. *IEEE Transactions on Industrial Informatics*, 10(4):2247–2256, 2014.

[13] O'DwyerE., PanI., AchaS., GibbonsS., and ShahN. *Modelling and Evaluation of Multi-Vector Energy Networks in Smart Cities*, pages 161–168. 2019.

[14] Henrik Lund, Poul Alberg Østergaard, Miguel Chang, Sven Werner, Svend Svendsen, Peter Sorknæs, Jan Eric Thorsen, Frede Hvelplund, Bent Ole Gram Mortensen, Brian Vad Mathiesen, Carsten Bojesen, Neven Duic, Xiliang Zhang, and Bernd Möller. The status of 4th generation district heating: Research and results. *Energy*, 164:147–159, 2018.

[15] Simone Buffa, Marco Cozzini, Matteo D'Antoni, Marco Baratieri, and Roberto Fedrizzi. 5th generation district heating and cooling systems: A review of existing cases in europe. *Renewable and Sustainable Energy Reviews*, 104:504–522, 2019.

[16] Kevin Naik and Anton Ianakiev. Heat demand prediction: A real-life data model vs simulated data model comparison. *Energy Reports*, 7:380–388, 2021. The 17th International Symposium on District Heating and Cooling.

[17] Microsoft. Azure iot hub. https://azure.microsoft.com/en-gb/services/iot-hub/, February 2023.

[18] IBM. Internet of things on ibm cloud. https://www.ibm.com/cloud/internet-of-things, February 2023.

[19] Google. Google cloud iot solutions. https://cloud.google.com/solutions/iot/, February 2023.

[20] Erik Dotzauer. Simple model for prediction of loads in district-heating systems. *Applied Energy*, 73(3):277–284, 2002.

[21] Bronislav Chramcov, Petr Dostál, and Jaroslav Balátě. Forecast model of heat demand. 10 2021.

[22] Christian Johansson, Markus Bergkvist, Davy Geysen, Oscar De Somer, Niklas Lavesson, and Dirk Vanhoudt. Operational demand forecasting in district heating systems using ensembles of online machine learning algorithms. *Energy Procedia*, 116:208–216, 2017. 15th International Symposium on District Heating and Cooling, DHC15-2016, 4-7 September 2016, Seoul, South Korea.

[23] Spyridon Provatas, Niklas Lavesson, and Christian Johansson. An online machine learning algorithm for heat load forecasting in district heating systems. volume 12, 09 2014.

[24] Behrang Talebi, Parham A. Mirzaei, Arash Bastani, and Fariborz Haghighat. A review of district heating systems: Modeling and optimization. *Frontiers in Built Environment*, 2:22, 2016.

[25] Itzal del Hoyo Arce, Saioa Herrero López, Susana López Perez, Miika Rämä, Krzysztof Klobut, and Jesus A. Febres. Models for fast modelling of district heating and cooling networks. *Renewable and Sustainable Energy Reviews*, 82:1863–1873, 2018.

[26] Davy Geysen, Oscar De Somer, Christian Johansson, Jens Brage, and Dirk Vanhoudt. Operational thermal load forecasting in district heating networks using machine learning and expert advice. *Energy and Buildings*, 162:144–153, 2018.

[27] Michael Short, Tracey Crosbie, Muneeb Dawood, and Nashwan Dawood. Load forecasting and dispatch optimisation for decentralised co-generation plant with dual energy storage. *Applied Energy*, 186:304–320, 2017. Sustainable Thermal Energy Management (SusTEM2015).

[28] R. Bavière and M. Vallée. Optimal temperature control of large scale district heating networks. *Energy Procedia*, 149:69–78, 2018. 16th International Symposium on District Heating and Cooling, DHC2018, 9–12 September 2018, Hamburg, Germany.

[29] Etienne Saloux and José A. Candanedo. Forecasting district heating demand using machine learning algorithms. *Energy Procedia*, 149:59–68, 2018. 16th International Symposium on District Heating and Cooling, DHC2018, 9–12 September 2018, Hamburg, Germany.

[30] Ashreeta Prasanna, Viktor Dorer, and Nadège Vetterli. Optimisation of a district energy system with a low temperature network. *Energy*, 137:632–648, 2017.

[31] Gowri Suryanarayana, Jesus Lago, Davy Geysen, Piotr Aleksiejuk, and Christian Johansson. Thermal load forecasting in district heating networks using deep learning and advanced feature selection methods. *Energy*, 157:141–149, 2018.

[32] Xu Xu, Youwei Jia, Yan Xu, Zhao Xu, Songjian Chai, and Chun Sing Lai. A multi-agent reinforcement learning-based data-driven method for home energy management. *IEEE Transactions on Smart Grid*, 11(4):3201–3211, 2020.

[33] Milan Zdravković, Ivan Ćirić, and Marko Ignjatović. Explainable heat demand forecasting for the novel control strategies of district heating systems. *Annual Reviews in Control*, 53:405–413, 2022.

[34] Guixiang Xue, Yu Pan, Tao Lin, Jiancai Song, Chengying Qi, and Zhipan Wang. District heating load prediction algorithm based on feature fusion lstm model. *Energies*, 12(11), 2019.

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

[36] Masatoshi Sakawa and Takeshi Matsui. Heat load prediction in district heating and cooling systems through recurrent neural networks. *International Journal of Operational Research*, 23(3):284–300, 2015.

[37] Kosuke Kato, Masatoshi Sakawa, Keiichi Ishimaru, Satoshi Ushiro, and Toshihiro Shibano. Heat load prediction through recurrent neural network in district heating and cooling systems. In *2008 IEEE International Conference on Systems, Man and Cybernetics*, pages 1401–1406, 2008.

[38] Puning Xue, Yi Jiang, Zhigang Zhou, Xin Chen, Xiumu Fang, and Jing Liu. Multi-step ahead forecasting of heat load in district heating systems using machine learning algorithms. *Energy*, 188:116085, 2019.

[39] Tanveer Ahmad and Huanxin Chen. Short and medium-term forecasting of cooling and heating load demand in building environment with data-mining based approaches. *Energy and Buildings*, 166:460–476, 2018.

[40] Cheng Fan, Fu Xiao, and Yang Zhao. A short-term building cooling load prediction method using deep learning algorithms. *Applied Energy*, 195:222–233, 2017.

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

[42] Cheng Fan, Yongjun Sun, Yang Zhao, Mengjie Song, and Jiayuan Wang. Deep learning-based feature engineering methods for improved building energy prediction. *Applied Energy*, 240:35–45, 2019.

[43] S. Sholahudin and Hwataik Han. Simplified dynamic neural network model to predict heating load of a building using taguchi method. *Energy*, 115:1672–1678, 2016. Sustainable Development of Energy, Water and Environment Systems.

[44] Kyungtae Yun, Rogelio Luck, Pedro J. Mago, and Heejin Cho. Building hourly thermal load prediction using an indexed arx model. *Energy and Buildings*, 54:225–233, 2012.

[45] Alexander Sandberg, Fredrik Wallin, Hailong Li, and Maher Azaza. An analyze of long-term hourly district heat demand forecasting of a commercial building using neural networks. *Energy Procedia*, 105:3784–3790, 2017. 8th International Conference on Applied Energy, ICAE2016, 8-11 October 2016, Beijing, China.

[46] Subodh Paudel, Mohamed Elmitri, Stéphane Couturier, Phuong H. Nguyen, René Kamphuis, Bruno Lacarrière, and Olivier Le Corre. A relevant data selection method for energy consumption prediction of low energy building based on support vector machine. *Energy and Buildings*, 138:240–256, 2017.

[47] Riasat Sarwar, Heejin Cho, Sam J. Cox, Pedro J. Mago, and Rogelio Luck. Field validation study of a time and temperature indexed autoregressive with exogenous (arx) model for building thermal load prediction. *Energy*, 119:483–496, 2017.

[48] Miguel Martínez Comesaña, Lara Febrero-Garrido, Francisco Troncoso-Pastoriza, and Javier Martínez-Torres. Prediction of building's thermal performance using lstm and mlp neural networks. *Applied Sciences*, 10(21), 2020.

[49] Qiong Li, Qinglin Meng, Jiejin Cai, Hiroshi Yoshino, and Akashi Mochida. Predicting hourly cooling load in the building: A comparison of support vector machine and different artificial neural networks. *Energy Conversion and Management*, 50(1):90–96, 2009.

[50] Abdullatif E. Ben-Nakhi and Mohamed A. Mahmoud. Cooling load prediction for buildings using general regression neural networks. *Energy Conversion and Management*, 45(13):2127–2141, 2004.

[51] Ryohei Yokoyama, Tetsuya Wakui, and Ryoichi Satake. Prediction of energy demands using neural network with model identification by global optimization. *Energy Conversion and Management*, 50(2):319–327, 2009.

[52] Chaobo Zhang, Yang Zhao, Xuejun Zhang, Cheng Fan, and Tingting Li. An improved cooling load prediction method for buildings with the estimation of prediction intervals. *Procedia Engineering*, 205:2422–2428, 2017. 10th International Symposium on Heating, Ventilation and Air Conditioning, ISHVAC2017, 19-22 October 2017, Jinan, China.

[53] Min Hee Chung, Young Kwon Yang, Kwang Ho Lee, Je Hyeon Lee, and Jin Woo Moon. Application of artificial neural networks for determining energy-efficient operating set-points of the vrf cooling system. *Building and Environment*, 125:77–87, 2017.

[54] Jianjuan Yuan, Zhihua Zhou, Huajie Tang, Chendong Wang, Shilei Lu, Zhao Han, Ji Zhang, and Ying Sheng. Identification heat user behavior for improving the accuracy of heating load prediction model based on wireless on-off control system. *Energy*, 199:117454, 2020.

[55] Qiong Li, Qinglin Meng, Jiejin Cai, Hiroshi Yoshino, and Akashi Mochida. Applying support vector machine to predict hourly cooling load in the building. *Applied Energy*, 86(10):2249–2256, 2009.

[56] Deyin Zhao, Ming Zhong, Xu Zhang, and Xing Su. Energy consumption predicting model of vrv (variable refrigerant volume) system in office buildings based on data mining. *Energy*, 102:660–668, 2016.

[57] Chengliang Fan, Yunfei Ding, and Yundan Liao. Analysis of hourly cooling load prediction accuracy with data-mining approaches on different training time scales. *Sustainable Cities and Society*, 51:101717, 2019.

[58] Chengliang Fan, Yundan Liao, and Yunfei Ding. Development of a cooling load prediction model for air-conditioning system control of office buildings. *International Journal of Low-Carbon Technologies*, 14(1):70–75, 01 2019.

[59] Chengliang Fan and Yunfei Ding. Cooling load prediction and optimal operation of hvac systems using a multiple nonlinear regression model. *Energy and Buildings*, 197:7–17, 2019.

[60] Xi Fang, Guangcai Gong, Guannan Li, Liang Chun, Wenqiang Li, and Pei Peng. A hybrid deep transfer learning strategy for short term cross-building energy prediction. *Energy*, 215:119208, 2021.

[61] Ran Wang, Shilei Lu, and Wei Feng. A novel improved model for building energy consumption prediction based on model integration. *Applied Energy*, 262:114561, 2020.

[62] Yan Ding, Qiang Zhang, Tianhao Yuan, and Kun Yang. Model input selection for building heating load prediction: A case study for an office building in tianjin. *Energy and Buildings*, 159:254–270, 2018.

[63] P.J.C. Vogler–Finck, P. Bacher, and H. Madsen. Online short-term forecast of greenhouse heat load using a weather forecast service. *Applied Energy*, 205:1298–1310, 2017.

[64] Yan Ding, Qiang Zhang, and Tianhao Yuan. Research on short-term and ultra-short-term cooling load prediction models for office buildings. *Energy and Buildings*, 154:254–267, 2017.

[65] Jing Zhao and Xiaojuan Liu. A hybrid method of dynamic cooling and heating load forecasting for office buildings based on artificial intelligence and regression analysis. *Energy and Buildings*, 174:293–308, 2018.

[66] Jihao Gu, Jin Wang, Chengying Qi, Chunhua Min, and Bengt Sundén. Medium-term heat load prediction for an existing residential building based on a wireless on-off control system. *Energy*, 152:709–718, 2018.

[67] Vincent Bakker, Albert Molderink, Johann L. Hurink, and Gerard J. M. Smit. Domestic heat demand prediction using neural networks. In *2008 19th International Conference on Systems Engineering*, pages 189–194, 2008.

[68] Aowabin Rahman and Amanda D. Smith. Predicting heating demand and sizing a stratified thermal storage tank using deep learning algorithms. *Applied Energy*, 228:108–121, 2018.

[69] Zheng O'Neill and Charles O'Neill. Development of a probabilistic graphical model for predicting building energy performance. *Applied Energy*, 164:650–658, 2016.

[70] Cristian Ghiaus. Experimental estimation of building energy performance by robust regression. *Energy and Buildings*, 38(6):582–587, 2006.

[71] G Mihalakakou, M Santamouris, and A Tsangrassoulis. On the energy consumption in residential buildings. *Energy and Buildings*, 34(7):727–736, 2002.

[72] Zhanyu Ma, Hailong Li, Qie Sun, Chao Wang, Aibin Yan, and Fredrik Starfelt. Statistical analysis of energy consumption patterns on the heat demand of buildings in district heating systems. *Energy and Buildings*, 85:464–472, 2014.

[73] F. Martellotta, U. Ayr, P. Stefanizzi, A. Sacchetti, and G. Riganti. On the use of artificial neural networks to model household energy consumptions. *Energy Procedia*, 126:250–257, 2017. ATI 2017 - 72nd Conference of the Italian Thermal Machines Engineering Association.

[74] D. Koschwitz, J. Frisch, and C. van Treeck. Data-driven heating and cooling load predictions for non-residential buildings based on support vector machine regression and narx recurrent neural network: A comparative study on district scale. *Energy*, 165:134–142, 2018.

[75] Jihoon Jang, Joosang Lee, Eunjo Son, Kyungyong Park, Gahee Kim, Jee Hang Lee, and Seung-Bok Leigh. Development of an improved model to predict building thermal energy consumption by utilizing feature selection. *Energies*, 12(21), 2019.

[76] Simon S.K. Kwok, Richard K.K. Yuen, and Eric W.M. Lee. An intelligent approach to assessing the effect of building occupancy on building cooling load prediction. *Building and Environment*, 46(8):1681–1690, 2011.

[77] Simon S.K. Kwok and Eric W.M. Lee. A study of the importance of occupancy to building cooling load in prediction by intelligent approach. *Energy Conversion and Management*, 52(7):2555–2564, 2011.

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

[79] S.L. Wong, Kevin K.W. Wan, and Tony N.T. Lam. Artificial neural networks for energy analysis of office buildings with daylighting. *Applied Energy*, 87(2):551–557, 2010.

[80] Joseph C. Lam, Kevin K.W. Wan, S.L. Wong, and Tony N.T. Lam. Principal component analysis and long-term building energy simulation correlation. *Energy Conversion and Management*, 51(1):135–139, 2010.

[81] Tiberiu Catalina, Vlad Iordache, and Bogdan Caracaleanu. Multiple regression model for fast prediction of the heating energy demand. *Energy and Buildings*, 57:302–312, 2013.

[82] Fazel Khayatian, Luca Sarto, and Giuliano Dall'O'. Application of neural networks for evaluating energy performance certificates of residential buildings. *Energy and Buildings*, 125:45–54, 2016.

[83] Tiberiu Catalina, Joseph Virgone, and Eric Blanco. Development and validation of regression models to predict monthly heating demand for residential buildings. *Energy and Buildings*, 40(10):1825–1832, 2008.

[84] Saleh Seyedzadeh, Farzad Pour Rahimian, Parag Rastogi, and Ivan Glesk. Tuning machine learning models for prediction of building energy loads. *Sustainable Cities and Society*, 47:101484, 2019.

[85] Mingju Gong, Jin Wang, Yin Bai, Bo Li, and Lei Zhang. Heat load prediction of residential buildings based on discrete wavelet transform and tree-based ensemble learning. *Journal of Building Engineering*, 32:101455, 2020.

[86] Jiyang Xie, Jiaxin Guo, Zhanyu Ma, Jing-Hao Xue, Qie Sun, Hailong Li, and Jun Guo. Sea: A combined model for heat demand prediction. In *2018 International Conference on Network Infrastructure and Digital Content (IC-NIDC)*, pages 71–75, 2018.

[87] Szabolcs Kováč, German Micha'čonok, Igor Halenár, and Pavel Važan. Comparison of heat demand prediction using wavelet analysis and neural network for a district heating network. *Energies*, 14(6), 2021.

[88] Zhanyu Ma, Jiyang Xie, Hailong Li, Qie Sun, Fredrik Wallin, Zhongwei Si, and Jun Guo. Deep neural network-based impacts analysis of multimodal factors on heat demand prediction. *IEEE Transactions on Big Data*, 6(3):594–605, 2020.

[89] Krzysztof Wojdyga. Predicting heat demand for a district heating systems. *International Journal of Energy and Power Engineering*, 3:237, 01 2014.

[90] Jiyang Xie, Hailong Li, Zhanyu Ma, Qie Sun, Fredrik Wallin, Zhongwei Si, and Jun Guo. Analysis of key factors in heat demand prediction with neural networks. *Energy Procedia*, 105:2965–2970, 2017. 8th International Conference on Applied Energy, ICAE2016, 8-11 October 2016, Beijing, China.

[91] Kody M. Powell, Akshay Sriprasad, Wesley J. Cole, and Thomas F. Edgar. Heating, cooling, and electrical load forecasting for a large-scale district energy system. *Energy*, 74:877–885, 2014.

[92] Mingju Gong, Yin Zhao, Jiawang Sun, Cuitian Han, Guannan Sun, and Bo Yan. Load forecasting of district heating system based on informer. *Energy*, 253:124179, 2022.

[93] Hongying Yang, Shuanglong Jin, Shuanglei Feng, Bo Wang, Fei Zhang, and Jianfeng Che. Heat load forecasting of district heating system based on numerical weather prediction model. 01 2016.

[94] Tingting Fang and Risto Lahdelma. Evaluation of a multiple linear regression model and sarima model in forecasting heat demand for district heating system. *Applied Energy*, 179:544–552, 2016.

[95] Aleksandra A. Sretenović Radiša Ž. Jovanović and Branislav D. Živković. Ensemble of various neural networks for prediction of heating energy consumption. *Energy and Buildings*, 94:189–199, 2015.

[96] Thomas Nigitz and Markus Gölles. A generally applicable, simple and adaptive forecasting method for the short-term heat load of consumers. *Applied Energy*, 241:73–81, 2019.

[97] O.A Olanrewaju, A.A Jimoh, and P.A Kholopane. Energy efficiency assessment using artificial neural network combined with data envelopment analysis. In *IEEE Africon '11*, pages 1–5, 2011.

[98] Chujie Lu, Sihui Li, and Zhengjun Lu. Building energy prediction using artificial neural networks: A literature survey. *Energy and Buildings*, 262:111718, 2022.

[99] Li Ming-hai and Liu Min. Prediction of public building energy consumption by using artificial fish-swarm algorithm. In *2015 Chinese Automation Congress (CAC)*, pages 2043–2047, 2015.

[100] M.A. Ancona, L. Branchini, A. De Pascale, and F. Melino. Smart district heating: Distributed generation systems' effects on the network. *Energy Procedia*, 75:1208–1213, 2015. Clean, Efficient and Affordable Energy for a Sustainable Future: The 7th International Conference on Applied Energy (ICAE2015).

[101] Atli Benonysson, Benny Bøhm, and Hans F. Ravn. Operational optimization in a district heating system. *Energy Conversion and Management*, 36(5):297–314, 1995.

[102] Diana Böttger, Mario Götz, Nelly Lehr, Hendrik Kondziella, and Thomas Bruckner. Potential of the power-to-heat technology in district heating grids in germany. *Energy Procedia*, 46:246–253, 2014. 8th International Renewable Energy Storage Conference and Exhibition (IRES 2013).

[103] Dario Buoro, Piero Pinamonti, and Mauro Reini. Optimization of a distributed cogeneration system with solar district heating. *Applied Energy*, 124:298–308, 2014.

[104] E. Carpaneto, P. Lazzeroni, and M. Repetto. Optimal integration of solar energy in a district heating network. *Renewable Energy*, 75:714–721, 2015.

[105] Katinka Johansen and Sven Werner. Something is sustainable in the state of denmark: A review of the danish district heating sector. *Renewable and Sustainable Energy Reviews*, 158:112117, 2022.

[106] Henrik Lund, Neven Duic, Poul Alberg Østergaard, and Brian Vad Mathiesen. Future district heating systems and technologies: On the role of smart energy systems and 4th generation district heating. *Energy*, 165:614–619, 2018.

[107] D. Connolly, H. Lund, B.V. Mathiesen, and M. Leahy. A review of computer tools for analysing the integration of renewable energy into various energy systems. *Applied Energy*, 87(4):1059–1082, 2010.

[108] Mei Gong and Sven Werner. An assessment of district heating research in china. *Renewable Energy*, 84:97–105, 2015. Sustainable energy utilization in cold climate zone (Part I).

[109] Aira Hast, Sanna Syri, Vidas Lekavičius, and Arvydas Galinis. District heating in cities as a part of low-carbon energy system. *Energy*, 152:627–639, 2018.

[110] H. Lund, B. Möller, B.V. Mathiesen, and A. Dyrelund. The role of district heating in future renewable energy systems. *Energy*, 35(3):1381–1390, 2010.

[111] Behnaz Rezaie and Marc A. Rosen. District heating and cooling: Review of technology and potential enhancements. *Applied Energy*, 93:2–10, 2012. (1) Green Energy; (2)Special Section from papers presented at the 2nd International Enery 2030 Conf.

[112] George Xydis. Wind energy integration through district heating. a wind resource based approach. *Resources*, 4(1):110–127, 2015.

[113] Henrik Lund, Neven Duić, Goran Krajac˘ić, and Maria da Graça Carvalho. Two energy system analysis models: A comparison of methodologies and results. *Energy*, 32(6):948–954, 2007. Third Dubrovnik Conference on Sustainable Development of Energy, Water and Environment Systems.

[114] H Lund and E Münster. Modelling of energy systems with a high percentage of chp and wind power. *Renewable Energy*, 28(14):2179–2193, 2003.

[115] Henrik Lund, Poul Alberg Østergaard, David Connolly, and Brian Vad Mathiesen. Smart energy and smart energy systems. *Energy*, 137:556–565, 2017.

[116] K. Sartor, S. Quoilin, and P. Dewallef. Simulation and optimization of a chp biomass plant and district heating network. *Applied Energy*, 130:474–483, 2014.

[117] Janelle S. Hygh, Joseph F. DeCarolis, David B. Hill, and S. Ranji Ranjithan. Multivariate regression as an energy assessment tool in early building design. *Building and Environment*, 57:165–175, 2012.

[118] Haichao Wang, Wusong Yin, Elnaz Abdollahi, Risto Lahdelma, and Wenling Jiao. Modelling and optimization of chp based district heating system with renewable energy production and energy storage. *Applied Energy*, 159:401–421, 2015.

[119] Cihan Turhan, Tugce Kazanasmaz, Ilknur Erlalelitepe Uygun, Kenan Evren Ekmen, and Gulden Gokcen Akkurt. Comparative study of a building energy performance software (kep-iyte-ess) and ann-based building heat load estimation. *Energy and Buildings*, 85:115–125, 2014.

[120] Sungjin Lee, Soo Cho, Seo-Hoon Kim, Jonghun Kim, Suyong Chae, Hakgeun Jeong, and Taeyeon Kim. Deep neural network approach for prediction of heating energy consumption in old houses. *Energies*, 14(1), 2021.

[121] Alaoui Sosse Jihad and Mohamed Tahiri. Forecasting the heating and cooling load of residential buildings by using a learning algorithm "gradient descent", morocco. *Case Studies in Thermal Engineering*, 12:85–93, 2018.

[122] Sachin Kumar, Saibal K. Pal, and Ram Pal Singh. A novel method based on extreme learning machine to predict heating and cooling load through design and structural attributes. *Energy and Buildings*, 176:275–286, 2018.

[123] Min-Yuan Cheng and Minh-Tu Cao. Accurately predicting building energy performance using evolutionary multivariate adaptive regression splines. *Applied Soft Computing*, 22:178–188, 2014.

[124] Athanasios Tsanas and Angeliki Xifara. Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. *Energy and Buildings*, 49:560–567, 2012.

[125] Jui-Sheng Chou and Dac-Khuong Bui. Modeling heating and cooling loads by artificial intelligence for energy-efficient building design. *Energy and Buildings*, 82:437–446, 2014.

[126] Mehrbakhsh Nilashi, Mohammad Dalvi-Esfahani, Othman Ibrahim, Karamollah Bagherifard, Abbas Mardani, and Norhayati Zakuan. A soft computing method for the prediction of energy performance of residential buildings. *Measurement*, 109:268–280, 2017.

[127] Sanjiban Sekhar Roy, Reetika Roy, and Valentina E. Balas. Estimating heating load in buildings using multivariate adaptive regression splines, extreme learning machine, a hybrid model of mars and elm. *Renewable and Sustainable Energy Reviews*, 82:4256–4268, 2018.

[128] Sokratis Papadopoulos, Elie Azar, Wei-Lee Woon, and Constantine E. Kontokosta. Evaluation of tree-based ensemble learning algorithms for building energy performance estimation. *Journal of Building Performance Simulation*, 11(3):322–332, 2018.

[129] Ngoc-Tri Ngo, Anh-Duc Pham, Thi Thu Ha Truong, Ngoc-Son Truong, Nhat-To Huynh, and Tuan Minh Pham. An ensemble machine learning model for enhancing the prediction accuracy of energy consumption in buildings. *Arabian Journal for Science and Engineering*, 47(4):4105–4117, Apr 2022.

[130] Mauro Castelli, Leonardo Trujillo, Leonardo Vanneschi, and Aleš Popovič. Prediction of energy performance of residential buildings: A genetic programming approach. *Energy and Buildings*, 102:67–74, 2015.

[131] Sanjiban Sekhar Roy, Pijush Samui, Ishan Nagtode, Hemant Jain, Vishal Shivaramakrishnan, and Behnam Mohammadi-ivatloo. Forecasting heating and cooling loads of buildings: a comparative performance analysis. *Journal of Ambient Intelligence and Humanized Computing*, 11(3):1253–1264, Mar 2020.

[132] Azadeh Sadeghi, Roohollah Younes Sinaki, William A. Young, and Gary R. Weckman. An intelligent model to predict energy performances of residential buildings based on deep neural networks. *Energies*, 13(3), 2020.

[133] Mabrook Al-Rakhami, Abdu Gumaei, Ahmed Alsanad, Atif Alamri, and Mohammad Mehedi Hassan. An ensemble learning approach for accurate energy load prediction in residential buildings. *IEEE Access*, 7:48328–48338, 2019.

[134] Le Thi Le, Hoang Nguyen, Jie Dou, and Jian Zhou. A comparative study of pso-ann, ga-ann, ica-ann, and abc-ann in estimating the heating load of buildings' energy efficiency for smart city planning. *Applied Sciences*, 9(13), 2019.

[135] Muhammad Sajjad, Samee Ullah Khan, Noman Khan, Ijaz Ul Haq, Amin Ullah, Mi Young Lee, and Sung Wook Baik. Towards efficient building designing: Heating and cooling load prediction via multi-output model. *Sensors*, 20(22), 2020.

[136] Arash Moradzadeh, Amin Mansour-Saatloo, Behnam Mohammadi-Ivatloo, and Amjad Anvari-Moghaddam. Performance evaluation of two machine learning techniques in heating and cooling loads forecasting of residential buildings. *Applied Sciences*, 10(11), 2020.

[137] Xuan-Nam Bui, Hossein Moayedi, and Ahmad Safuan A. Rashid. Developing a predictive method based on optimized m5rules–ga predicting heating load of an energy-efficient building system. *Engineering with Computers*, 36(3):931–940, Jul 2020.

[138] Sachin Kumar, Saibal K. Pal, and Ram Pal Singh. Intra elm variants ensemble based model to predict energy performance in residential buildings. *Sustainable Energy, Grids and Networks*, 16:177–187, 2018.

[139] Issa Jaffal, Christian Inard, and Christian Ghiaus. Fast method to predict building heating demand based on the design of experiments. *Energy and Buildings*, 41(6):669–677, 2009.

[140] Betul Bektas Ekici and U. Teoman Aksoy. Prediction of building energy needs in early stage of design by using anfis. *Expert Systems with Applications*, 38(5):5352–5358, 2011.

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

[142] İsmail Yabanova and Ali Keçebaş. Development of ann model for geothermal district heating system and a novel pid-based control strategy. *Applied Thermal Engineering*, 51(1):908–916, 2013.

[143] L Lianzhong and M Zaheeruddin. Hybrid fuzzy logic control strategies for hot water district heating systems. *Building Services Engineering Research and Technology*, 28(1):35–53, 2007.

[144] Loïc Giraud, Massinissa Merabet, R. Bavière, and Mathieu Vallée. Optimal control of district heating systems using dynamic simulation and mixed integer linear programming. 05 2017.

[145] Jelger Jansen, Filip Jorissen, and Lieve Helsen. Optimal control of a fourth generation district heating network using an integrated non-linear model predictive controller. *Applied Thermal Engineering*, 223:120030, 2023.

[146] Siemens. MindSphere: Enabling the world's industries to drive their digital transformations. *Https://Www.Plm.Automation.Siemens.Com/Global/En/Products/Mindsphere/Internet-of-Things-Iot.Html*, page 24, 2018.

[147] Robert Haafst. Ge vs siemens - a strategic analysis. 11 2016.

[148] T. Macaulay S. Carey and H. Williams. Best cloud iot platforms. https://www.computerworld.com/article/3430236/best-cloud-iot-platforms.html#slide1, August 2017.

[149] RAHUL BADNAKHE. Evaluating top iot platforms to kick start your iot development, Apr 2020.

[150] BENTLEY. District energy network planning and design solution. https://www.bentley.com/en/solutions/district-energy-network-planning-and-design, February 2023.

[151] AVEVA. District energy management. https://www.aveva.com/en/solutions/operations/continuous-process-operations/, February 2023.

[152] GLOBEMA. District heating industry. https://www.globema.com/industries/district-heating/, February 2023.

[153] PassivUK. Network load management. https://passivuk.com/, February 2023.

[154] Essential. Communal and district heating. https://www.essentialcontrol.co.uk/industries/network-heat/, February 2021.

[155] Jeffrey H. Johnson and Pejman Iravani. The multilevel hypernetwork dynamics of complex systems of robot soccer agents. *ACM Trans. Auton. Adapt. Syst.*, 2(2):5–es, jun 2007.

[156] Jeffrey Johnson. Hypernetworks for reconstructing the dynamics of multilevel systems. *European Conference on Complex Systems (25-29 September 2006)*, 2(September):17, 2006.

[157] Jeffrey Johnson. *The Multidimensional Networks of Complex Systems*, pages 49–79. 01 1995.

[158] Jeffrey H. Johnson, Cristian Jimenez-Romero, David M. S. Rodrigues, and Jane Bromley. Hypernetwork-based peer marking for scalable certificated mass education. 2014.

[159] Jeffrey Johnson. Hypernetworks for policy design in systems of systems of systems. In Kristin Glass, Richard Colbaugh, Paul Ormerod, and Jeffrey Tsao, editors, *Complex Sciences*, pages 179–189, Cham, 2013. Springer International Publishing.

[160] Jeffrey H. Johnson. Multidimensional multilevel networks in the science of the design of complex systems. 2006.

[161] Jeffrey Johnson. Hypernetworks of complex systems. In Jie Zhou, editor, *Complex Sciences*, pages 364–375, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

[162] J. H. Johnson. Hypernetworks: Multidimensional relationships in multilevel systems. *The European Physical Journal Special Topics*, 225(6):1037–1052, Sep 2016.

[163] R. Atkin. *Mathematical Structure in Human Affairs*. Heinemann educational books. Heinemann Educational, 1974.

[164] R. Atkin. *Multidimensional Man*. Penguin books. Penguin Books, 1981.

[165] Satish Palaniappan, Naveen Hariharan, Naren T Kesh, Vidhyalakshimi S, and Angel Deborah S. Article: Home automation systems - a study. *International Journal of Computer Applications*, 116(11):11–18, April 2015. Full text available.

[166] Mohammed Faisal Elrawy, Ali Ismail Awad, and Hesham F. A. Hamed. Intrusion detection systems for iot-based smart environments: a survey. *Journal of Cloud Computing*, 7:1–20, 2018.

[167] Giovanny Mondragón-Ruíz, Alonso Tenorio-Trigoso, Manuel Castillo-Cara, Blanca Caminero, and Carmen Carrión. An experimental study of fog and cloud computing in cep-based real-time iot applications. *Journal of Cloud Computing*, 10:1–17, 2021.

[168] Antonio Iyda Paganelli, Pedro Elkind Velmovitsky, Pedro Miranda, Adriano Branco, Paulo Alencar, Donald Cowan, Markus Endler, and Plinio Pelegrini Morita. A conceptual iot-based early-warning architecture for remote monitoring of covid-19 patients in wards and at home. *Internet of Things*, 2021.

[169] D. Halliday, R. Resnick, and J. Walker. *Fundamentals of Physics*. John Wiley & Sons, 2010.

[170] G Cucca. *Innovative co-simulation tool for assessment and optimisation of energy consumption in deep retrofitted buildings*. PhD thesis, Nottingham Trent University, July 2021.

[171] R. McMullan. *Environmental Science in Building*. Bloomsbury Publishing, 2017.

[172] Iman Asadi, Payam Shafigh, Zahiruddin Fitri Bin Abu Hassan, and Norhayati Binti Mahyuddin. Green. *Journal of Building Engineering*, 20:81–93, 2018.

[173] C.P. Kothandaraman. *Fundamentals of Heat and Mass Transfer*. New Age International (P) Limited, Publishers, 2012.

[174] Elisabeth Green, Tristram Hope, and Alan Yates. *Sustainable Infrastructure: Sustainable Buildings*. ICE Publishing, 2015.

[175] Energy & Industrial Strategy Department for Business. Heat networks pipeline. https://www.gov.uk/government/publications/hndu-pipeline, October 2021.

[176] Amin Al-Habaibeh, Arijit Sen, and John Chilton. Evaluation tool for the thermal performance of retrofitted buildings using an integrated approach of deep learning artificial neural networks and infrared thermography. *Energy and Built Environment*, 2(4):345–365, 2021.

[177] Alireza Bahadori, Sohrab Zendehboudi, and Kamel Hooman. Estimation of characteristic temperature ratios for panel radiator and high temperature radiant strip systems to calculate heat loss within a room space. *Energy and Buildings*, 55:508–514, 2012. Cool Roofs, Cool Pavements, Cool Cities, and Cool World.

[178] Tomasz Kisilewicz, Małgorzata Fedorczak-Cisak, and Tamas Barkanyi. Active thermal insulation as an element limiting heat loss through external walls. *Energy and Buildings*, 205:109541, 2019.

[179] Qian Jin, Fabio Favoino, and Mauro Overend. Design and control optimisation of adaptive insulation systems for office buildings. part 2: A parametric study for a temperate climate. *Energy*, 127:634–649, 2017.

[180] R. O'Hegarty, O. Kinnane, D. Lennon, and S. Colclough. In-situ u-value monitoring of highly insulated building envelopes: Review and experimental investigation. *Energy and Buildings*, 252:111447, 2021.

[181] Blanca Tejedor, Miquel Casals, and Marta Gangolells. Assessing the influence of operating conditions and thermophysical properties on the accuracy of in-situ measured u-values using quantitative internal infrared thermography. *Energy and Buildings*, 171:64–75, 2018.

[182] Tania Berger, Christof Amann, Herbert Formayer, Azra Korjenic, Bernhard Pospichal, Christoph Neururer, and Roman Smutny. Impacts of external insulation and reduced internal heat loads upon energy demand of offices in the context of climate change in vienna, austria. *Journal of Building Engineering*, 5, 11 2015.

[183] Paul Baker. U-values and traditional buildings in situ measurements and their comparison to calculated values, 01 2011.

[184] Lin Gao, Xuyang Cui, Jiaxin Ni, Wanning Lei, Tao Huang, Chao Bai, and Junhong Yang. Technologies in smart district heating system. *Energy Procedia*, 142:1829–1834, 2017. Proceedings of the 9th International Conference on Applied Energy.

[185] Nikolay N. Novitsky, Zoya I. Shalaginova, Aleksandr A. Alekseev, Vyacheslav V. Tokarev, Oksana A. Grebneva, Aleksandr V. Lutsenko, Olga V. Vanteeva, Egor A. Mikhailovsky, Roman Pop, Petr Vorobev, and Michael Chertkov. Smarter smart district heating. *Proceedings of the IEEE*, 108(9):1596–1611, 2020.

[186] M.J. Moran, H.N. Shapiro, B.R. Munson, and D.P. DeWitt. *Introduction to Thermal Systems Engineering: Thermodynamics, Fluid Mechanics, and Heat Transfer*. Wiley, 2002.

[187] Jacques A de Chalendar, Peter W Glynn, and Sally M Benson. City-scale decarbonization experiments with integrated energy systems. *Energy & Environmental Science*, 12(5):1695–1707, 2019.

[188] Daniel Quiggin and Richard Buswell. The implications of heat electrification on national electrical supply-demand balance under published 2050 energy scenarios. *Energy*, 98:253–270, 2016.

[189] Peter D Lund, Juuso Lindgren, Jani Mikkola, and Jyri Salpakari. Review of energy system flexibility measures to enable high levels of variable renewable electricity. *Renewable and sustainable energy reviews*, 45:785–807, 2015.

[190] Wei Wang, Tianzhen Hong, Xiaodong Xu, Jiayu Chen, Ziang Liu, and Ning Xu. Forecasting district-scale energy dynamics through integrating building network and long short-term memory learning algorithm. *Applied Energy*, 248:217–230, 2019.

[191] Tanveer Ahmad and Huanxin Chen. Potential of three variant machine-learning models for forecasting district level medium-term and long-term energy demand in smart grid environment. *Energy*, 160:1008–1020, 2018.

[192] Michaela Killian and Martin Kozek. Ten questions concerning model predictive control for energy efficient buildings. *Building and Environment*, 105:403–412, 2016.

[193] Mohammad Mohammadi, Younes Noorollahi, Behnam Mohammadi-ivatloo, Mehdi Hosseinzadeh, Hossein Yousefi, and Sasan Torabzadeh Khorasani. Optimal management of energy hubs and smart energy hubs–a review. *Renewable and Sustainable Energy Reviews*, 89:33–50, 2018.

[194] Edward O'Dwyer, Indranil Pan, Salvador Acha, and Nilay Shah. Smart energy systems for sustainable smart cities: Current developments, trends and future directions. *Applied Energy*, 237:581–597, 2019.

[195] sharingcities.eu. Sharing cities consortium, Nov 2022.

[196] UC Berkley. ptolemy. berkeley. edu.[online] available at: h ttps. *ptolemy. berkeley. edu/ptolemyII/index. htm*, 2021.

[197] AMAI Makhorin. Glpk-gnu project-free software foundation (fsf). *Free Softw. Found., Boston, MA, USA, Tech. Rep*, 2012.

[198] Gurobi Optimization. gurobi. com.[online] available at: h ttps, 2021.

[199] Ane Blázquez-García, Angel Conde, Aitor Milo, Roberto Sánchez, and Irantzu Barrio. Short-term office building elevator energy consumption forecast using sarima. *Journal of Building Performance Simulation*, 13(1):69–78, 2020.

[200] Department for Business. Energy & Industrial Strategy. Heat networks investment project, introduction to the scheme, 2018.

[201] David Hawkey and Janette Webb. Comments on the competition and markets authority's "heat networks market study–statement of scope". 2018.

[202] Sven Werner. International review of district heating and cooling. *Energy*, 137:617–631, 2017.

[203] Diane Orme. 5 simple statistics that prove district heating is the future of uk heat distribution, Sep 2022.

[204] LORA Alliance®. About lora alliance®. https://lora-alliance.org, October 2021.

[205] Giorgio Cucca and Anton Ianakiev. Assessment and optimisation of energy consumption in building communities using an innovative co-simulation tool. *Journal of Building Engineering*, 32:101681, 2020.