Modelling and predicting the heating and cooling demand of a building with a simple fast data-driven model

Master thesis report

Sanne van Dorp (5891116 and s3741788)

MSc Industrial Ecology (TU Delft and Leiden University)

July 2024

First supervisor: Regina Bokel (TU Delft)

Second supervisor: Laure Itard (TU Delft)

External supervisors Deerns B.V.: Cristina Jurado López & Karthik Gunderi

Acknowledgements

When I was 9 years old, I drew a village called '*de goede dorp*', where i drew a few houses, a collective cable car and a collective garden and mill. For each house I drew solar panels and made a calculation of the electricity generation that was used for the houses and the cable car. I was always interested in renewable energy sources. I started with a bachelor Future Planet Studies at the University of Amsterdam, where I learned more about energy transitions, the multidisciplinary of the energy sector and made the first steps in modelling. By choosing the master Industrial Ecology at TU Delft and Leiden University, I wanted to learn more about the technical part of energy. Within this master I discovered my interest in modelling. When I contacted Laure Itard about a thesis object, and she introduced me to this project, I became enthusiastic and curious about the whole Brains4Building project. There, I had the honour to present my research at the yearly consortium day.

I did not have a strong background in the field of energy in Mechanical Engineering. I have learned so much within this project, where I kept a certainty about all the parts of this project. I want to thank Laure Itard and Regina Bokel for teaching me so much about this topic and providing me with useful feedback on how to improve this research. I also want to thank Regina Bokel for helping me with the measurements.

I want to thank Arie Taal and Baldiri Salcedo Rahola for providing me with the data from the Haagse Hogeschool in Delft and for offering me their time to explain to me how the building works and to help me with the calculations.

I want to express my warm thanks to the Deerns team for making my daily work at the Deerns Office enjoyable and for including me in the team, which made me feel very welcomed. These warm thanks go out especially to Cristina Jurado López and Karthik Gunderi, for supervising me daily, who were always happy to answer my questions and who were always bringing up great ideas on how to improve this research.

Sanne van Dorp

Winsum, July 2024

Summary

This research presents an innovative approach to modeling and predicting the heating and cooling demands of buildings using a simple and fast data-driven model. The research is set within the context of Model Predictive Control (MPC) systems, which optimize building energy use by predicting variables and adjusting control inputs accordingly, embodying the principles of industrial ecology by integrating technical, economic, and social dimensions of sustainability; enhancing energy efficiency, reducing CO_2 emissions, and improving occupant comfort, which are key goals of industrial ecology. From the available modelling methods: white, black, and grey box models, to model and predict the heating and cooling demand, a grey box model: multivariate linear regression model with as input actual data selected based on the thermal energy balance and Pearson correlation coefficients. A case study on two rooms: an office and a classroom in the Haagse Hogeschool in Delft serves as the practical application of the developed model. Several models are developed; static and dynamic models, and using different independent variables; indoor surface temperature, outdoor temperature, indoor air temperature, internal heat gains, wind speed and solar light intensity. A accuracy, expressed in the R^2 value between 22.52% and 78.57% is achieved with modelling the heating and cooling demand. The developed models are not able to predict the heating and cooling demand, due to multicollinearity between independent variables, overfitting and endogeneity.

Contents

1. Introduction

The biggest 21st-century challenge is reducing climate change. The weather patterns of the earth are changing due to the increase in greenhouse gas emissions since the Industrial Revolution in the 18th century. Greenhouse gases include CO_2, CH_4, N_2O, SF_6, and NF_3. HFK, CFK and PFK (Calvin et al., 2023). An increase in these greenhouse gasses in the atmosphere causes an increase in heat absorption in the atmosphere (Calvin et al., 2023). A big share of this heat is absorbed by the earth's surface. The temperature of the earth's surface increases, and a part of this heat is emitted back to the atmosphere, however, due to the greenhouse gases, a part of this heat is reflected to the earth's surface. The changing weather patterns have a negative effect on the environment which affects human society (Calvin et al., 2023). Effects of climate change are increased sea levels, higher temperatures, losses of biodiversity, more extreme weather conditions, melting of permafrost, pressure on ecosystems, and a threat to food supply security (Calvin et al., 2023).

To reduce these effects of climate change on the environment there are international treaties between countries. The front-runner was the Kyoto Protocol, signed in 1997 (Cary & Stephens, 2024) A follow-up treaty is the Paris Agreement in 2015, where the goal is set to reduce global warming to 1,5 degrees Celsius, with a maximum of 2 degrees Celsius (UNFCCC), 2015). To achieve these goals, each involved country must translate this into policies. In the Netherlands, an agreement (Klimaatakkoord) was reached between the Dutch government and companies in 2019 (Dutch Parliament, 2019). The Dutch climate goal is to reduce the national emissions of greenhouse gases to 55% by 2030, in comparison to 1990, the pursuance is to reduce to 60%. In 2050, the Netherlands should be climate neutral. Agreements are made for the sectors; of electricity, built environment, industry, mobility agriculture, and land use.

A key player in the energy transition to achieve the climate goals is the building sector. In the European Union, the building sector consumes 47% of the final energy demand to heat and cool buildings. Within households, 86% of final energy use is used for heating, commerce, services, and agriculture this share is 76% (Santamarta et al., 2021). Within the Klimaatakkoord the goal is set for the building sector to be CO_2 neutral by 2050. The building sector needs to increase energy efficiency, save energy, and integrate renewable energy sources into the supply of energy (Dutch Parliament, 2019). To integrate more renewable energy sources, besides the use of green electricity, more renewable energy sources are needed to be used to heat and cool a building.

The context of this research is shown in Figure 1. The assumption is made that the building relies solely on electricity, with no connection to the gas grid for heating and cooling. The current electricity production includes both green electricity from renewable sources like wind and solar, and grey electricity from fossil fuels. As the production of green electricity increases, there is a decrease in the use of fossil fuels, resulting in lower CO_2 emissions. However, the growing production of renewable energy sources, including prosumers, is putting pressure on the electricity grid, leading to imbalances and network congestion.

To address this issue, there is a need to optimize the electricity demand and supply using model predictive control (MPC). With MPC, data is collected from the electricity supply, the grid, and the building, and is used to predict electricity consumption for heating and cooling, as well as the production of renewable energy. This allows for the adjustment of electricity demand during times of high green electricity availability, leading to lower CO_2 emissions, increased usage of renewable energy sources, balanced electricity, reduced network congestion, and the possibility of reducing the electricity bill.

Figure 1 illustrates the relationship between renewable energy sources, fossil fuels, the electricity grid, and buildings. The green dashed arrows represent data collection and optimization flows as part of MPC, highlighting the feedback loop within the system.

In the context of industrial ecology, this research emphasizes the interaction between different components of energy production, distribution, and consumption. Industrial ecology focuses on optimizing resource use and minimizing environmental impact through a holistic approach. By integrating renewable energy sources by employing MPC, the system achieves higher efficiency and sustainability, embodying the principles of industrial ecology. This approach not only reduces CO_2 emissions but also enhances the resilience and reliability of the electricity grid, promoting a more sustainable and balanced industrial ecosystem.
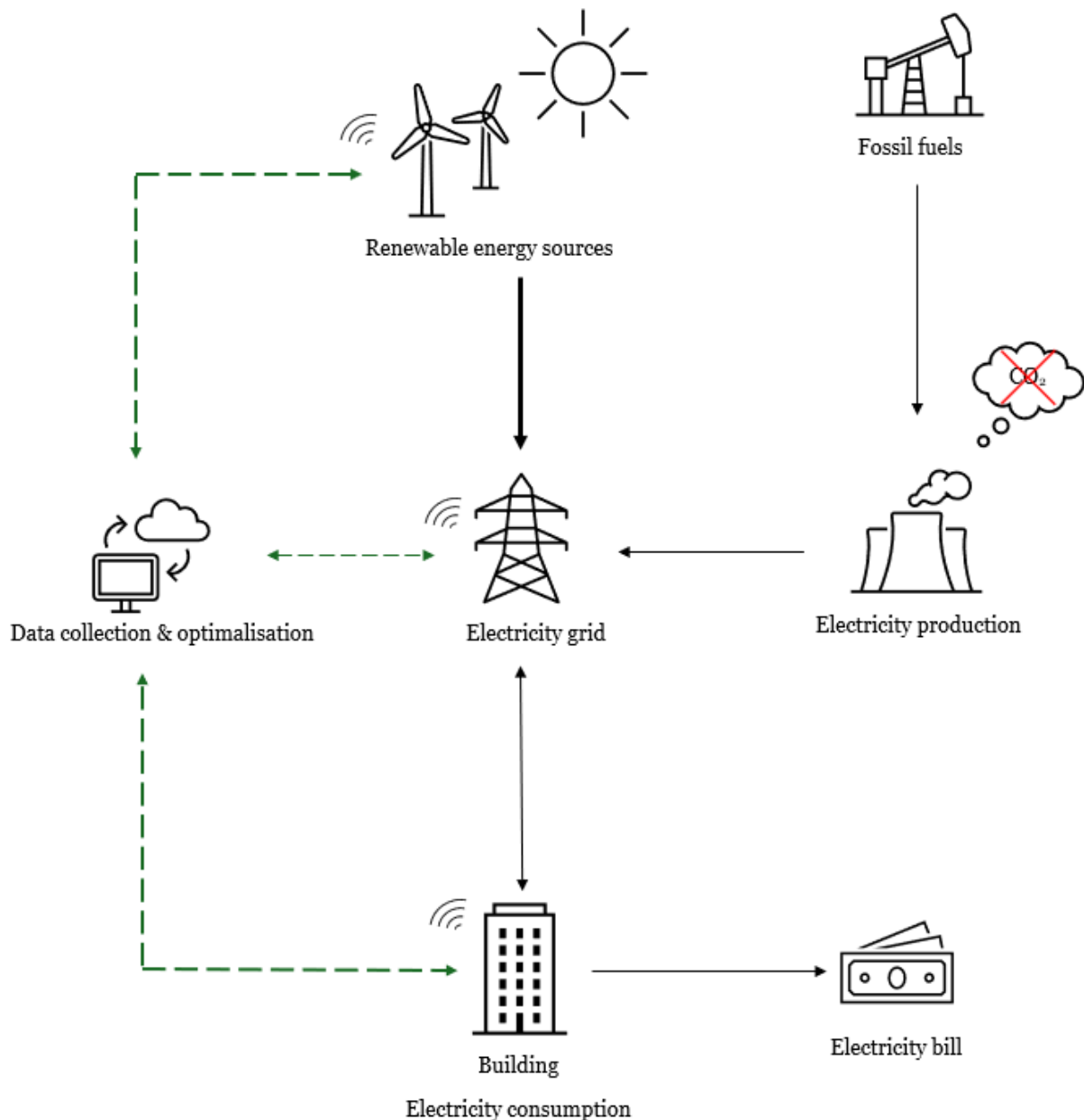
Figure 1: schematic overview of the context of this research

This research is part of the Brains4Building project, a multi-stakeholder and multi-year initiative aiming to research and develop a Model Predictive Control (MPC) framework for buildings. The goal is to reduce energy consumption, increase comfort, respond flexibly to user behaviour and local energy supply and demand, and save on installation maintenance costs (by fault detection). Part of that project is to develop a model to model and predict the heating and cooling demand.

This study focuses on modelling and predicting the heating and cooling demand of buildings as part of MPC, using a simple, fast, data-driven multivariate linear regression model based on the thermal energy balance and with the use of actual data (grey box modelling) as

part of Model Predictive Control (MPC). Traditional models: white box and black models, have limitations in terms of complexity and data requirements (Rasooli & Itard, 2020). This research aims to address these limitations by developing a model that leverages measurable data inputs like indoor surface temperature and outdoor temperature.

The primary objective of this study is to determine whether it is possible to model and predict the heating and cooling demand of a building for seven hours in advance using actual data. This involves developing a model that can be validated with real-world data and refined to improve accuracy by adding and removing variables. The case study for this research is the Haagse Hogeschool building in Delft, a highly energy-efficient building equipped with advanced heating and cooling systems, which enable the use of real-time data. The following research question is asked:

Is it possible to model and predict for 7 hours in advance the heating and cooling demand of a building during opening hours with actual data as input for a simple, fast and data-driven model?

The following structure will be followed. To answer the main research question is divided into sub-questions and report chapters. A research flow diagram, is shown in Figure 2. First, a literature review is conducted to further explain model predictive control and the potential of modelling and predicting the heating demand, to emphasise the social and academic relevance of this study. After this, the existing modelling and predicting models to model and predict the heating and cooling demand are explained, to identify the most fit to purpose modelling method. After the literature review, the model methodology including the required data inputs. Thereby the following sub-research question will answered:

What are the data inputs of the multivariate linear regression model and which statistical validation and search process are used to build the model?

This model is applied to a real case study: the Haagse Hogeschool in Delft, where the selected rooms are explained, data from these room is gathered. For each variable, it will be discussed how the data is collected and if the data is complete, whereby the following sub-research question is answered:

How is the required data collected for each variable, and is this data set complete?

To understand how the variables are correlated to the heating or cooling demand, a data analysis is performed with the use of the Pearson-correlation coefficient. This gives insight into the accuracy of the model. The following sub-research question is answered:

How are the variables and the heating and cooling demand correlated to each other, and which variables will significantly contribute to a higher model accuracy?

Models to model and predict the heating and cooling demand are developed. Different combinations of variables and time delays will be used. The following sub-research question will be answered:

What is the best combination of independent variables in the model to achieve the highest accuracy in modelling the heating and cooling demand?
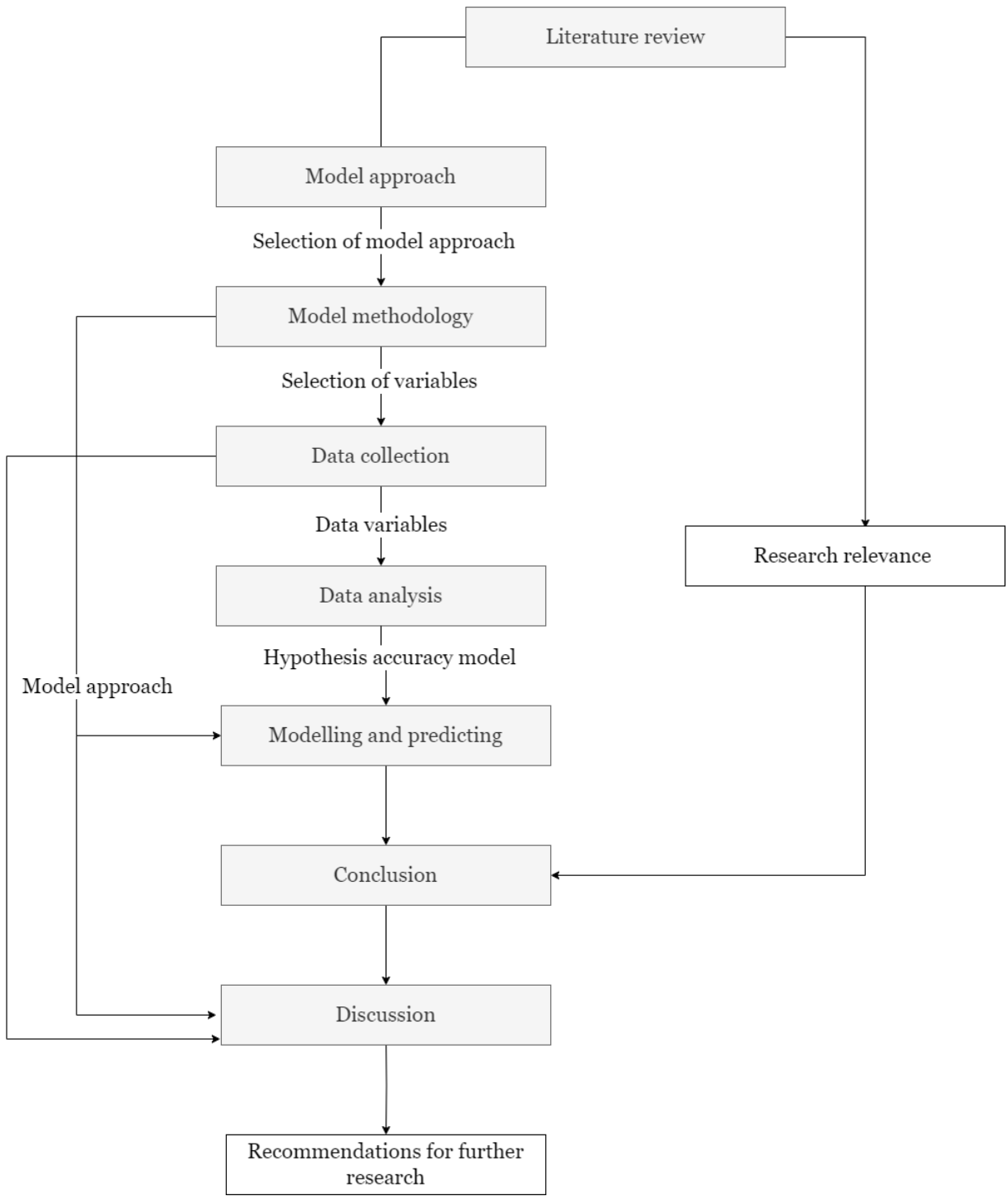
Figure 2: reseach flow diagram

2. Literature review

The literature review starts with an explanation of Model Predictive Control, to explain the context of the development of a model to model and predict the heating and cooling demand of the office. After that, the potential of modelling and predicting the heating and cooling demand within an MPC is explained, which was already shortly introduced in the introduction with the connection to the field of Industrial Ecology. The last part of the literature review will explain the different model approaches, and what results in a selection of the used model.

2.1. Model Predictive Control (MPC)

Model Predictive Control (MPC) is a control strategy used to improve the performance of HVAC and other (heating and cooling) building systems. It works by predicting the variables and making adjustments to control inputs accordingly. This method involves creating an optimization problem that aims for example to minimize a cost function, to use more renewable energy sources (RES) (during a high supply of RES) or to increase the thermal comfort (Chen & You, 2023) while considering system dynamics and constraints. By using MPC, buildings can effectively manage energy usage systematically and flexibly while ensuring occupant comfort and meeting operational goals. One of the main challenges for widespread application of MPC is the development of a user-friendly, control-oriented, accurate, and computationally efficient building modelling (Drgoňa et al., 2020).

Figure 3 shows the controlled physical system: the building. The room temperatures (y) are affected by (d), as weather conditions, they are predicted by weather forecasts ($\hat{d}$) and heat flows through the building (u). The estimator estimates the effect of the room temperatures on the buildings thermal mass temperatures (x). The predicted building thermal mass and the predicted disturbances ($\hat{d}$) are used as data input for the building model. The building model predicts the future building thermal mass temperatures for 1 timestep later (X_{k+1}). With these predictions, the cost function and constraints (for example thermal comfort ranges) are used to solve the optimization problem, which results in a cost-effective heat flow (u) within the thermal comforts and that the actions are feasible and safe. This is a constant loop, where prediction is done for the building surface temperatures and the disturbances, whereby a prediction horizon (N) is stated (for example a prediction of 7 hours).
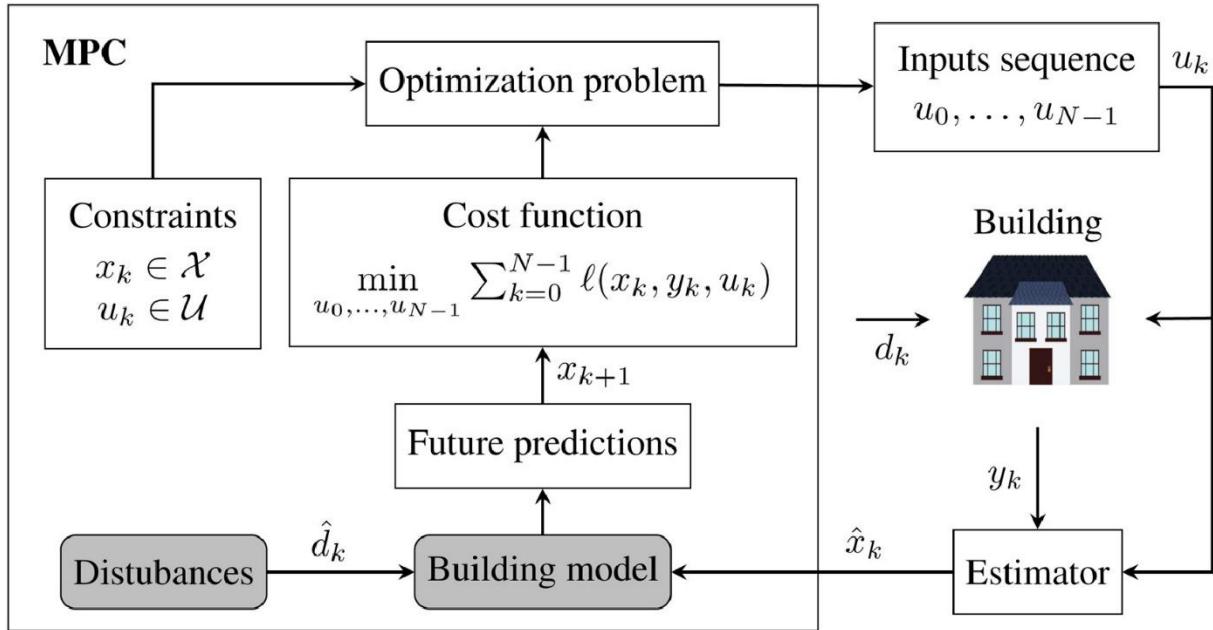
Figure 2: visualization of the parts of MPC k is timestep (Drgoňa et al., 2020)

Part of an MPC is the required building energy simulation tool. The development of these tools emerged around 50 years ago and can be divided into four generations (Kwak et al., 2015).

1. First Generation: Early models provided a basic performance indication for buildings. They were easy to use but difficult to apply to real-world scenarios due to limited computational power. These models were simplified and segmented, lacking integration and often hiding deficiencies(Kwak et al., 2015).

2. Second Generation (mid-1970s): These models included time dynamics, using the dynamic response of constructional elements to capture the time-delaying effect of a building's thermal mass. While still simplified, they began to address the temporal aspects of building dynamics(Kwak et al., 2015).

3. Third Generation (mid-1980s): With increased computing power, integrated modelling emerged, considering thermal, visual, and acoustic performance together. Space and time dimensions were treated as independent variables, meaning energy transfer processes could not be solved in isolation, enhancing the models' accuracy and real-world applicability(Kwak et al., 2015).

4. Fourth Generation and Beyond (mid-1990s): These models featured domain integration and program interoperability, closely matching reality. Intelligent, knowledge-based systems became fully integrated and network-compatible. These tools were easy to use and interpret, predictive, multi-variable, ubiquitous, and accessible, with transparent deficiencies(Kwak et al., 2015).

Around the third and fourth generation, around the late eighties, when the models became more integrated, the concept of MPC emerged for supervisory control of building energy

management systems, despite the limited computer power (Bianchini et al., 2016). The use of MPC to save load, electricity-market-related control, and control based on the building thermal capacity determined by the thermal mass is studied in several studies (Kwak et al., 2015). MPC research in buildings has evolved to include both optimization and probabilistic methods. Optimization studies have enhanced HVAC system (and other thermal supply systems) design and control, leading to significant energy savings and improved comfort. Probabilistic studies have introduced methods to handle uncertainties, making control systems more robust and adaptable. Together, these approaches are pushing the boundaries of building energy management, leading to smarter, more efficient buildings (Kwak et al., 2015).

An example of a probabilistic approach is the study of Chen & You (2023). They proposed a data-driven robust model predictive control (DDRMPC) framework to address building climate control with a renewable hybrid energy system (geothermal heat pump, photovoltaic panel, and an electricity storage battery) under weather forecast uncertainty. A data-driven model (with the use of machine learning) is used to investigate the forecast errors of the weather conditions. The optimization problem of the MPC was to ensure thermal comfort, enhance energy efficiency, and reduce costs. The research highlights the potential of advanced clustering techniques in managing the complexities of building energy systems, ensuring a balance between renewable energy use and maintaining indoor comfort levels while addressing the challenge of uncertainties of weather variables predictions.

The study of Bianchini et al. (2016) discusses the optimization of building heating systems using Model Predictive Control (MPC) within the framework of demand-response (DR) programs. Optimization was used to minimize energy costs and maximize efficiency by solving complex algorithms that predict future energy needs and adjust system parameters accordingly. Probabilistic methods, on the other hand, incorporate uncertainty and variability in predictions, allowing for more flexible and robust control strategies. The proposed heuristic optimization approach decouples the problem into smaller sub-problems, reducing computational complexity and making real-time application feasible. Testing with the EnergyPlus simulator (later further explained) on different building scales showed the heuristic approach provides near-optimal solutions efficiently. The research demonstrates how combining optimization and probabilistic methods in MPC can effectively balance energy savings, cost reduction, and occupant comfort while managing the uncertainties inherent in energy demand and supply.

2.2. Potential of modelling and predicting heating and cooling demand

The aim of the second part of the literature review is to explain the scientific and social relevance of the development of a fast simple data-driven model to predict the heating and cooling demand of a building as part of a model predictive control (MPC). The following sub-research question will be answered:

How could predicting the heating and cooling demand contribute to reducing the environmental impact of the building sector?

The energy demand of a building is determined by the outside temperature, physical building characteristics, socioeconomics, and occupant behaviour and preferences (Peplinski et al., 2024). The decarbonization of the energy system in the Netherlands requires a change towards smart grids and buildings (Lund et al., 2014; Marszal-Pomianowska et al., 2024; Norouzi et al., 2023). A prediction of the heating and cooling demand as part of Model Predictive Control (MPC) of a building is required to achieve demand flexibility (S. Yang et al., 2024). The potential of forecasting the heating and cooling demand of a building to reduce the environmental impact of the building sector is divided into three dimensions; technical, economic, and social.

Before the potential of predicting the heating and cooling demand can be described, it is essential to explain how the supply of heat or cold is connected to the electricity grid. The heating demand in the Netherlands is mostly supplied by the direct use of fossil fuels in buildings, this is mostly natural gas. In the building sector, there is of electrification, this is the transition from the use of fossil fuels to the use of electricity (from renewable energy sources). An example of this is the replacement of a natural gas boiler with an (electrical) heat pump to heat or cold a building. In this way, the heating and cooling demand is getting more connected to the electricity grid (Hoseinpoori, 2022).

2.1.1 Technical potential

Integration of renewable energy resources

More distributed energy sources: PV panels, wind turbines, heat pumps, or electrical vehicles are connected to the grid, to decarbonize the electricity grid (Hennig et al., 2023). Most of these renewable energy sources are dependent on weather conditions, for example, the electricity supply generated by solar panels is dependent on the amount of solar radiation. The integration of renewable energy sources into the electricity grid, requires balancing between demand and supply, ensuring supply security and avoiding network congestion (Norouzi et al.,

2023; S. Yang et al., 2024). The increasing amount of intermittent electricity generated by renewable energy sources requires more flexibility in the demand (Müller & Möst, 2018). Forecasting the heating and cooling demand contributes to this, which results in a reduction of imbalances on the electricity grid, to reduce network congestion, and to reduce CO_2 emissions. The reduction of these three problems is interdisciplinary as among others social costs are reduced (Hennig et al., 2023).

Imbalances

The amount of electricity supply and demand needs to be always in balance to secure stability, reliability, and efficiency of the electricity grid (Chaves-Ávila et al., 2013). When there is an imbalance in the electricity grid, this gap is filled with reserve power plants. In the Netherlands, the transmission system operator (TSO), TenneT is responsible for keeping the electricity grid in balance (Chaves-Ávila et al., 2013). The peak hours of demand and supply coincide less when there is a higher share of renewable energy sources integrated into the electricity grid as shown in Figure 4 (Goodarzi et al., 2019; Laugs et al., 2020).

Network congestion

The integration of renewable energy sources as part of the energy transition leads to an increase in electrification (Hennig et al., 2024; Norouzi et al., 2023; Peplinski et al., 2024). Another part of the energy transition is the decentralization of electricity generation. There is more pressure on the electricity grid, which is not designed for an increasing load, and the problem of network congestion has emerged (Hennig et al., 2023; Norouzi et al., 2023). Network congestion occurs on an electricity grid when the predicted demand load exceeds the capacity of the network (Hennig et al., 2023). In the Netherlands, more regions must challenge network congestion. For parts of the Netherlands, it is harder to get a new connection to deliver to or to consume from the electricity grid (Hennig et al., 2024) Network congestion is shown in Figure 5 for off-take capacity and input capacity in Figure 6 (Netbeheerder Nederland, 2024).

A strategy to alleviate network congestion involves shifting the peak hours of demand to align more closely with the peak hours of supply, a practice known as load shifting, a form of network congestion (Hennig et al., 2024; Rodrigues et al., 2022; S. Yang et al., 2024). This concept is illustrated in Figure 7, which demonstrates how load shifting can help mitigate network congestion. Through load shifting, the grid can better accommodate fluctuations in renewable energy generation, ensuring a more stable and efficient supply of electricity to consumers (Rodrigues et al., 2022). This load shifting can be done by several end-users of the electricity grid. For example, on the Dutch market, it is possible to smart charge an electrical vehicle

during peak hours of energy supply, consumers receive compensation for this, or the heating or cooling demand of well-insulated buildings could be flexible (Hennig et al., 2024). In Norway, research is done about the possibility of load shifting in an educational building. This research showed that a reduction of 50% of the load during peak hours by using a model-predictive control (MPC) was possible, whereby the daily energy use was not significantly increased (Clauß, 2024)
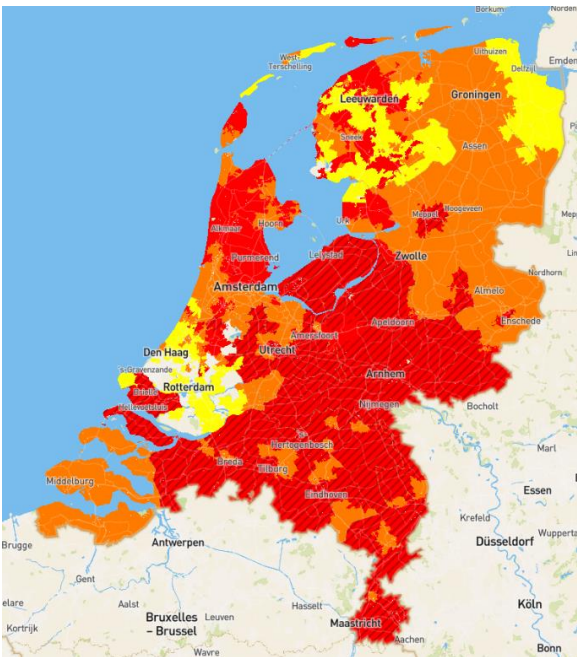
Figure 3: Offtake capacity map of the Netherlands on March 14, 2024 (Netbeheerder Nederland, 2024)
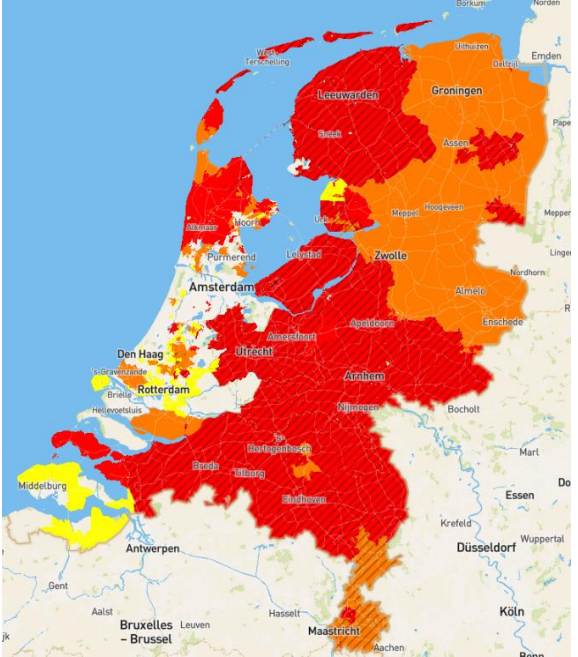
Figure 4: Input capacity map of the Netherlands on March 14, 2024 (Netbeheerder Nederland, 2024)

Legend: Within the red areas no transport capacity is available, and no congestion management is possible. Within the orange areas, no transport capacity is temporarily possible, and congestion management is investigated. Limited transport is available within the yellow areas. Transport capacity is available in the transparent areas
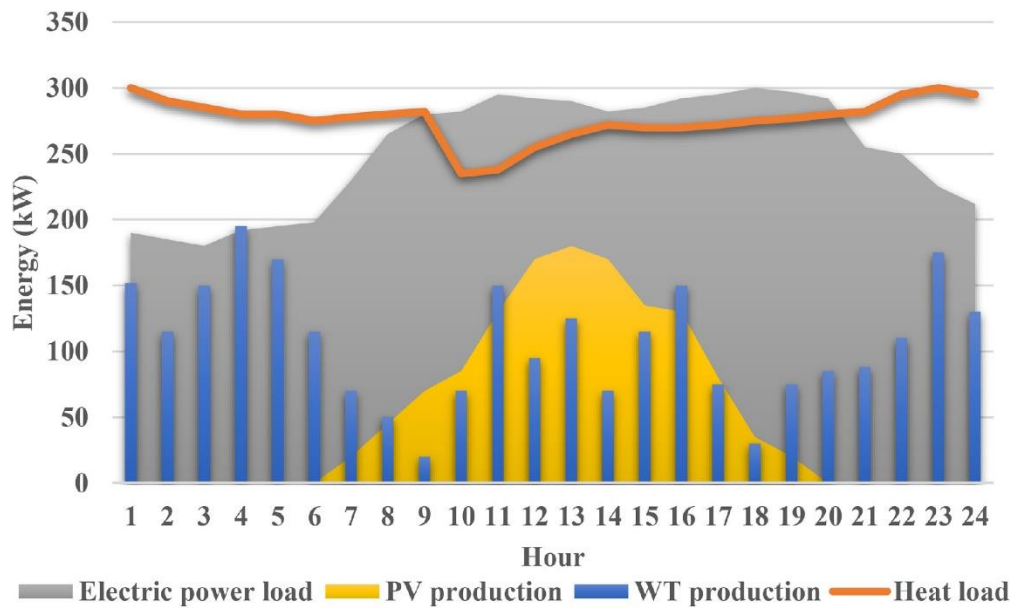
Figure 5: Daily profile of the renewable energy supply by photovoltaic and wind production, the electricity power and heat load (Yang & Jiang, 2024)

CO₂ emission reduction

The emission of CO_2 within a building depends on the building characteristics and the type of energy used to meet the energy demand (Pino-Mejías et al., 2017). Energy use within a building can be divided into; primary and secondary energy (Pino-Mejías et al., 2017; Pulido-Arcas et al., 2016). Primary energy is extracted or captured directly from natural energy sources. Examples are most of the fossil fuels than can be directly used in a building, for example, to heat a building with natural gas. The impact of these sources can be reduced by increasing the efficiency of using energy sources with lower CO_2 emissions. The second category of energy is energy generated somewhere else and consumed within the building, for example, electricity generated by a power plant. Each type of energy has a carbon dioxide equivalent emission factor (CO_2-eq) (Pino-Mejías et al., 2017; Pulido-Arcas et al., 2016).

Energy efficiency

With forecasting the thermal energy demand, the used energy resources can be used more efficiently (Kathirgamanathan et al., 2021; S. Yang et al., 2024). By accurate prediction of the heating and cooling demand, the building management systems can adjust heating and cooling schedules to meet the predicted demand. In this way, the building will only be heated and cooled when there is a demand. Energy waste will be reduced by preventing underheating or overheating of a building.

In the research of Merema et al. (2022), several model predictive control frameworks are compared on, among others, energy efficiency. The case study was a university building in Belgium. The highest heating demand (all-air installations) reduction of 55% was achieved by a grey box model: RC models (see Chapter 2.2.4) and occupancy-based temperature setpoints. A simplified calculation results in a reduction of 272.78 MWh on the yearly scale for the case study Haagse Hogeschool, when the yearly electricity consumption is estimated based on data from Centraal Bureau voor de Statistiek & Kadaster (n.d.) (40.82 kWh per m2), at 495.96 MWh.

2.1.2 Economic potential

Reducing energy costs

The challenges of imbalance on the electricity grid and network congestion emerged a growing interest in demand-responding mechanisms; market-based methods, network reconfiguration, network reinforcement, and dynamic prices (Hennig et al., 2024). Besides the reduction of energy costs through a higher energy efficiency, there is the potential for a lower energy bill by dynamic prices (S. Yang et al., 2024).

Nowadays, the electricity price for consumers is stated in energy contracts between the energy supplier and consumer for a fixed period (mostly 1 or 3 years) or 3 months. Energy for these types of contracts is purchased on the forward market, in the Netherlands (Pollitt et al., 2024). Instead of a fixed or flexible energy contract, there is also the option for a dynamic contract, whereby the energy is purchased on the same day (Miletić et al., 2022; Pollitt et al., 2024). The dynamic electricity price is determined by the hourly energy demand and supply curve. In the moments when there is more supply than demand, it is possible to get a negative price, so consumers are paid to use electricity at this moment. A higher price is paid in the moments when there is more demand than supply (Hofmann & Lindberg, 2024; Miletić et al., 2022). This financial incentive has the effect of demand shifting from periods with low supply and high demand to periods with low demand and high supply. The hourly maximal consumption will be reduced and distributed over the day. A study from Hofmann & Lindberg (2024) about dynamic electricity contracts in Norwegian concludes that these types of contracts result in savings in the hourly electricity demand and there is the potential to lower the energy bill for consumers when there is the possibility to shift the demand (Miletić et al., 2022).

However, the electricity prices are not connected to the CO_2 emissions, which results in not a direct reduction of the CO_2 emissions (Dahl Knudsen & Petersen, 2016). The other option is to couple the heating and cooling demand with the objective of CO_2 emission reduction. In this case, the load will be shifted to periods of low CO_2 emissions. The disadvantage of this approach

is the not direct connection to the load. To achieve the highest CO_2 reduction the heating and cooling demand needs to be connected to a combination of energy prices and CO_2 emissions (Dahl Knudsen & Petersen, 2016).

2.1.3 Social Potential

Environmental, Social and Governance (ESG) reporting

Stock market-listed companies are obliged to report on Environmental, Social and Governance (ESG). Within this framework, non-financial indicators are documented, and companies are allowed to choose their indicators within these three dimensions (Ferreira-Quilice et al., 2023). This needs to be done for the whole chain; from the *inside-out perspective* (regarding the output of a company) and the *outside-inperspective* (regarding the input of a company) (Ferreira-Quilice et al., 2023). The first dimension: environment includes the direct and indirect impact of the company on the natural environment. The circumstances for the employees are covered in the second dimension: social. The last dimension: governance, refers to the way a company is managed. Besides the fact that ESGs are mandatory for stock-market companies, it is also preferable to report these ESGs (which can be expressed in a score (A, B, etc.)) for the position on the market, as the output of the company could be the input of a company (Bonacorsi et al., 2024). Predicting the heating and cooling demand could be used to report the dimension environment, estimate the heating and cooling demand and the potential CO_2 reduction, and contribute to higher energy efficiency and lower energy costs. Forecasting the thermal demand could indirectly contribute also to higher thermal comfort, which satisfies the employees within the social part of ESG.

2.2. Modelling and predicting heating and cooling demand model methods

This research aims to develop a model and predict the heating or cooling demand as part of a Model Predictive Control system, therefore following sub-research question is be answered:

Which models are used within Model Predictive Control to model and predict the heating and cooling demand of a building?

The choice of the used model method is based on the available information and the prediction accuracy (Li et al., 2014). The types of data processing and analysing could be broadly classified based on the modelling method and the modelling problem into three models; white, black and grey box (Joseph Thaddeus et al., 2021). First, the thermal energy balance and the difference between static and dynamic modelling are explained.

2.2.1 Modelling background theory

Thermal energy balance

Analysing the thermo-physical properties of a building is an essential step for understanding the thermal performance of a building, leading to an estimation of the potential for energy-saving (Rasooli & Itard, 2020)

Laws of thermodynamics

There are many forms of energy; kinetic energy, mechanical energy, electricity and heat among others. The first law of thermodynamics states that no energy can be wasted or created (Itard, 2011). Energy could only be converted to another form of energy, for example, electrical energy to kinetic energy. In other words, the quantity of energy does not change. However, the second law of thermodynamics states that natural processes lead to both an increase in entropy and a decrease in the exergy of a system (Itard, 2011). Entropy is the degree of disorder, also called irreversibility. Exergy is the maximum amount of work available. High-quality energy forms (relatively high amount of exergy), such as mechanical work or electricity, can be fully converted to lower-quality energy forms (relativity low amount of exergy), such as heat. The reverse is not possible, heat cannot be converted to electricity with 100% efficiency (Itard, 2011).

Energy flows

Heat can be transferred through conduction (direct contact between materials), convection (movement of fluids or gases), and radiation (emission of electromagnetic waves) (Itard, 2011). The thermodynamic principle of conversion of energy is the base of determining the energy balance, by inventorying the entering and internal energy flows in a building. There is a heating demand when the energy balance is negative. The buildings need to be cooled when the energy balance is positive. The energy balance within a steady-state situation for every time interval is calculated with the following formula (Itard, 2011):

$$Q_{demand} = Q_{envelop} + Q_{ground} + Q_{ventilation} + Q_{infiltrations} + Q_{internal} + Q_{solar} + Q_{thermal,mass} \quad (1)$$

Within the energy balance of a building the three types of heat transfers are combined into four key energy flows: transmission, ventilation and infiltration, internal and, solar radiation. These energy flows are influenced by building characteristics, thermal comfort standards and outdoor climate (Itard, 2011). An overview of the terms of each variable can be found in Appendix A.

Transmission heat losses occur due to a difference in inside and outside temperature $(T_{outdoor} - T_{indoorair})$ through the envelope of the building and the ground $(T_{ground} - T_{indoorair})$. The heat flow through the building depends on the transfer coefficient (U), determined by the combined thermal conductivity of a material for convection and radiation, the wind speed, and the thermal resistance of the composed wall (Van Bueren et al., 2011). Heat transmission through the envelope is a sum of the four facades, dependent on the orientation and the roof.

$$Q_{envelop} = \sum_j U_{envelop}^j \cdot A_{envelop}^j \cdot (T_{outdoor} - T_{indoorair}) \quad (2)$$

$$Q_{ground} = U_{floor} \cdot A_{floor} \cdot (T_{ground} - T_{indoorair}) \quad (3)$$

Ventilation is the aimed natural or mechanical air flow from between inside and outside and infiltration of air through cracks of the building construction. Flowing air is required to meet the thermal comfort standards of among others air quality (measured by the CO_2 level) and for security reasons (Itard, 2011). The ventilation and infiltration heat gains are calculated by the mass flow rate of the entering air (m), multiplied by the heating capacity of the air ($C_{p,air}$) and the difference in the temperature of the air out of the AHU and the indoor air in the case of ventilation ($T_{out\ AHU} - T_{indoorair}$) or the temperature difference between the outdoor and indoor air for infiltrations ($T_{outdoor} - T_{indoorair}$).

$$Q_{ventilation} = m_{vent} \cdot C_{p,air} (T_{out\ AHU} - T_{indoorair}) \quad (4)$$
$$Q_{infiltrations} = (m_{opening} + m_{cracks}) \cdot C_{p,air} (T_{outdoor} - T_{indoorair}) \quad (5)$$

Internal heat gains are produced inside the building by people and electrical appliances (Itard, 2011). The heat gained from the human body depends on activity, worn clothes, air temperature, and humidity (Q_{body}). With these conditions and the occupancy (n_{people}) of a room or building the internal heat gains from people could be calculated. Another internal heat gain is electrical appliances, by using these appliances heat is released. This could be divided into heat gains from lighting (Q_{light}) and equipment ($Q_{equipment}$), where the released heat is determined by the type of lighting and equipment (Itard, 2011). The internal heat gains are calculated with the following formula:

$$Q_{internal} = n_{people} \cdot Q_{body} + A_{ceilings} \cdot Q_{light} + A_{floor} \cdot Q_{equipment} \quad (6)$$

Solar radiation (in the forms of direct radiation ($Q_{sol\ direct}$), reflected radiation ($Q_{relective}$), and diffuse radiation ($Q_{sol\ dif}$) influences the energy balance with the radiation entering the building through the windows and the absorption of the heat by the envelope. The impact of solar radiation is dependent on building characteristics such as orientation, size, and windows. The assumption could be made that the heat absorbed by the envelope of a building could be neglected in the case of a well-insulated building (Itard, 2011). Solar heat gains are calculated by the formula:

$$Q_{solar} = Q_{sol\ direct} + Q_{sol\ dif} + Q_{relective} \quad (7)$$

Heat is not absorbed by air; heat is accumulated in the thermal mass of a building. Thermal mass refers to the ability of a material to absorb, store, and release heat or cold over time. Materials with a high ability to store heat, that have a buffering effect on temperature fluctuations, have a high thermal mass (Itard, 2011). When the temperature of the surfaces is higher than the air temperature, the air is heated up by the thermal mass. This phenomenon occurs, for example, during the night, when the thermal mass is heated up by solar radiation during the day and released into the air during the night (shown in Figure 8) (Itard, 2011).
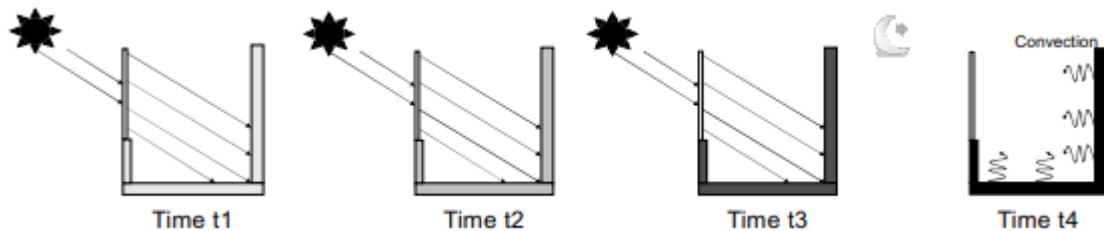
Figure 6: Absorption of heat by the thermal mass of a building. The heat from solar radiation accumulates over the day and is released during the night when the air temperature is lower than the surface temperature (Itard, 2011)

Heat gains from the accumulated heat within the thermal mass of a building are calculated with the formula (Itard, 2011):

$$Q_{thermal\ mass} = \sum_j \alpha_{i,j} \cdot A_j \cdot \left(T_{surface,j} - T_{indoorair}\right) \quad (8)$$

Equation 8 is a static model, where no time delay is included, in contrast to a dynamic model. The total heat gains from the thermal mass is determined by the sum of the heat transfer coefficient of the concerned surface ($\alpha_{i,j}$) multiplied by its surface (A_j) and the temperature difference between the surface and the indoor air ($T_{surface,j} - T_{indoorair}$).

Static and dynamic modelling

A model can be static or dynamic. In the case of the thermal energy balance, the heating or cooling demand is modelled at moment t, with the data of moment t, in other words there is no time dependency. In this formula there a no time delays used. An overview of the formulas used in the thermal energy balance, and the used variables with the units can be found in Appendix A.

$$Q_{demand}(t) = Q_{envelop}(t) + Q_{ground}(t) + Q_{ventilation}(t) + Q_{infiltrations}(t) + Q_{internal}(t) + Q_{solar}(t) + Q_{thermal,mass}(t) \quad (9)$$

Within these models, a single node is used (mostly the air temperature). A node in a thermal model represents a discrete point where temperature is measured or calculated. It is a location where thermal properties such as temperature, heat capacity, and thermal resistance are defined.

However, as previously described, there is a delay in the heat gains from solar radiation, absorbed by the thermal mass of a building, on the heating or cooling demand. In Figure 8, heat from solar radiation is absorbed by the thermal mass during the period between t1 and t3. At time t4, there is no solar radiation, and the air temperature is lower than the surface temperature, the accumulated heat is released to the indoor air in the room. So, in this case, the heat gains from solar radiation can be seen after 4 hours in the temperature of the indoor air. Other heat gains as internal heat gains and the heating supply are also accumulated by the thermal mass. Within dynamic models, this delay is included by including some variables from times before t. For example, the indoor surface of 4 hours before, the moment the heating or cooling demand is modelled.

There are different methods to cope with this time dependency.

The first possible method is making the whole model dynamic by introducing partial differential equations. Partial differential equations (PDEs) are a type of mathematical equation used to describe the distribution of heat (or other quantities) in a given space and time. They allow for a more detailed and precise representation of the physical processes involved in heat transfer, taking into account the spatial and temporal variations within a system. They can be solved with formulas in the case of simple differential equations. When these equations becomes complex, the Resitance-Capacitance model (Rastegarpour et al., 2020) or the Response Factors can be used

The first method is using *Resistance-Capacitance models* (Rc models). In this model, several nodes (for example, surface temperatures, outdoor temperatures and air temperatures) are modelled as part of a thermal network, where the time delay is captured in the heat capacity of

each surface or air (Bacher & Madsen, 2011). This model is further explained in Chapter 2.2.4. This model uses *Response Factors* (Itard, 2023). The thermal nodes within a room are connected by heat fluxes. Within the method for each node, an energy balance is described. An example is the room in Figure 9. With three thermal nodes; glass wall, indoor air and the (thick) back wall.
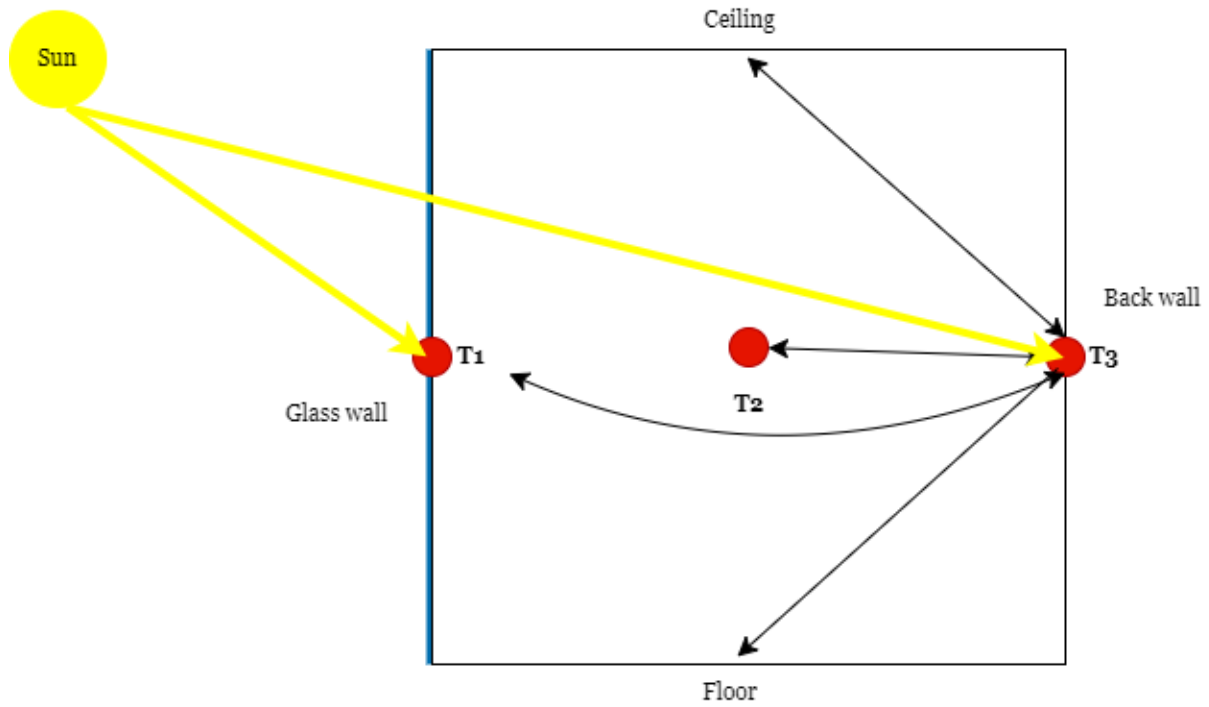
Figure 7: simplified drawing of a room with temperatures nodes, with the influence of the sun, from the perspective of the thermal node T_3. The heat is accumulated and released with a time delay (after Itard (2023)

The energy balance for the thermal node of the back wall is described in Equation 10.

$$\alpha_i \cdot A_w \cdot (T_2 - T_3) + \alpha_r \cdot A_{gl} \cdot (T_1 - T_3) + t \cdot A_{gl} \cdot P_{sol} + \lambda \frac{\partial^2 T_3}{\partial x^2} - \rho \cdot C_p \cdot \lambda \frac{\partial T_3}{\partial t} = 0 \text{ (10)}$$

With:

- α_i and α_r: heat transfer coefficient inside and outside
- A_w and A_{gl}: surface area of the back wall and the glass wall
- T: temperature
- P_{sol}: solar radiation
- λ: thermal conductivity
- C_p: specific heat capacity
- t: time
- ρ: density

The last part of the Equation 10: $\lambda \frac{\partial^2 T_3}{\partial x^2} - \rho \cdot C_p \cdot \lambda \frac{\partial T_3}{\partial t}$, described the thermal behaviour within in this case the back wall. This can be replaced by using response factors (X) and time dependency, as shown in Equation 11.

$$\alpha_i . A_w . (T_2 - T_3) + \alpha_r \cdot A_{gl} \cdot (T_1 - T_3) + t \cdot A_{gl} \cdot P_{sol} + X_o \cdot T_3 + X_1 \cdot T_3^{t-1} + X_2 \cdot T_3^{t-2} \approx$$
$$0 \ (11)$$

In this Equation, there is a time dependency of 2 hours before, this time dependency increases when the thermal mass of a building or room increases. These factors are calculated for each surface, describing the thermal performance of a wall and not describing what happened in the wall. These factors are calculated by describing the heat flux of both sides of the wall by a temperature change.

2.2.2 White box modelling method

The models built with this approach are based on the known physics of the buildings which are used in mathematical equations based on physical knowledge (Boodi et al., 2018; Ghiaus, 2014). These models are based on the principles of heat transfers and conservation of energy and mass (Drgoňa et al., 2020). The advantage of these models is that the variables and results of these models are interpreted in physical terms, which makes these models transparent. However, the input parameters are often unknown, besides that, these parameters are not often representative of the real building. Because these models include many parameters, which are often unknown, this complex model is not representative of the building (Rasooli & Itard, 2018; Yu et al., 2024). In the case that the parameters of the white box model are known and accurate, a higher accuracy can be achieved (Drgoňa et al., 2020). These models are used by different researchers (Coffey et al., 2010; Corbin et al., 2013)to couple them with an optimization problem with MPC, however, these schemes were computationally expensive (Drgoňa et al., 2020). Besides that, white models are not a easily scalable option to other buildings (Klanatsky et al., 2023).

An example of a white box model is the Low Energy Architecture model (LEA) developed by Deerns which calculates hourly energy needs for heating, cooling, humidification, lighting, ventilation, and equipment (Jurado López, 2017). It predicts thermal energy by calculating the energy demand to maintain indoor air temperature set points based on the thermal energy balance. This model uses dynamic equations based on a resistance network (RC model) for the time delay of the effect of the thermal mass by transferring the stored heat through convection to the other indoor surfaces, which are in contact with the indoor air (Jurado López, 2017).

Energy+ is another example of a model that follows the principles of white-box modelling (Del Ama Gonzalo et al., 2023). This model is a comprehensive building energy simulation program

that can simulate heating, cooling, lighting, ventilation, and other energy flows in buildings. Similar to LEA, Energy+ uses detailed physical equations to represent building components and systems (Del Ama Gonzalo et al., 2023). Energy+ is a dynamic modelling tool that performs detailed, time-step-based simulations of building energy use, capturing continuous changes and fluctuations over time (Del Ama Gonzalo et al., 2023). However, like other white box models, the accuracy of Energy+ depends on the availability and accuracy of input data, which can be a limitation when input parameters are unknown or not accurately representative of the real-world scenario (Del Ama Gonzalo et al., 2023).

2.2.3 Black box modelling method

The black box models are data-driven and statistical models based on the principle of machine learning (Kamel, Sheikh & Huang, 2020; Wei, 2018). These models use mathematical equations from statistics. These statistics are built on influential inputs to the outputs. A system is modelled as a box, whereas the parameters are unknown. The input of these models is data from sensors in the building. More buildings are equipped with smart meters. Data from these real-time measurements have an increasing potential for energy monitoring (Rasooli & Itard, 2020). The advantage of these models is that not all building parameters are needed, because of this it takes less time to develop. The biggest disadvantage of these models is that it is harder to interpret the results of this model to better manage the heating and cooling demand of the building in response to the energy supply (Jurado López, 2017). Black box models are easier to build than white-box models, however, a large dataset is needed to train this model (Klanatsky et al., 2023; Yu et al., 2024). The performance of model predictive control (MPC) relies heavily on the quality of models that describe the thermal behaviour of a building (Lin et al., 2024). Black-box models fall short in this aspect, as the results are hardly physically interpretable, making them less suitable for control and optimization applications (Klanatsky et al., 2023).

Gradient Boosting is a black box model that builds an ensemble of decision trees incrementally. Each new tree corrects the errors made by the previous trees, resulting in a highly accurate predictive model. Gradient Boosting is effective for capturing complex patterns in energy consumption data but can be computationally intensive and harder to interpret due to its complexity (Guo et al., 2023).

Random Forest is another black box model that creates an ensemble of decision trees. Unlike Gradient Boosting, it builds each tree independently using a random subset of data and features (Di Persio & Fraccarolo, 2023). This method is robust and less prone to overfitting compared to single decision trees. It is also easier to implement and understand than Gradient

Boosting, although still less interpretable than white-box models (Di Persio & Fraccarolo, 2023).

Another model built with the principles of black box modelling is Artificial Neural Networks (ANN) (Veljkovic et al., 2023). This model approach is inspired by the structure and function of biological neural networks in brains, where nodes are connected by edges. Via these edges, signals are given to the nodes, which results in an output described by a non-linear function based on the inputs. The benefits of these models to predict the heating and cooling demand are the accurate approach to include the non-linear relations between the parameters, accurate prediction with a minimum amount of input parameters, and easy implementation with a fast result (Veljkovic et al., 2023).

2.2.4 Grey box modelling method

For making a black box model a complete and large dataset is needed to achieve a high accuracy. Where for white box models the disadvantage is the lack of known building characteristics. The advantage of white box models is that the results can be physically interpreted, which is harder for black box models (Foucquier et al., 2013). A model in which the disadvantages of both the white and black models are resolved is the grey modelling approach (Kroll, 2000)(Bacher & Madsen, 2011; Raftery et al., 2011). These models are based on known physics of the building as on real-time data, which is obtained from sensors in the building (Rasooli & Itard, 2018; Wang & Xu, 2006). With this data, a set of equations is built by the model to determine the parameters (Rasooli & Itard, 2020).

In the research of (Veljkovic et al., 2023), a white box model (Energy+) is used to determine the inputs (for example the insulation value) of an ANN model, which makes it a grey box model. Besides combinations of black and white box models, there are models categorized as grey box models.

An example of a method of a grey box model is the use of the concept of (multivariate) linear regression where independent variables are chosen based on the energy thermal balance principles. This is a statistical method used to model the relationship between a dependent variable (for example, the heating and cooling demand) and one or more independent variables by fitting a linear equation to the observed data. This linear equation has the following mathematical form (Chatzithomas et al., 2015):

$$Y = a + b_1 X_1 + b_2 X_2 + \cdots + b_n X_n \quad (12)$$

where Y is the dependent variable (for example the heating demand), X_1, X_2 and X_n are the dependent values, b_1, b_2 and b_n are the coefficients of the variables, and a is a constant (so when Y is zero) (Chatzithomas et al., 2015). This method can be used in statistical software such as MATLAB. In several research papers, the multivariate linear regression model is categorised as a black box model (Drgoňa et al., 2020), however, as the parameters are based on the thermal energy balance (with among others building characteristics), this model method is categorized as a grey-box model.

Another grey box model is the Resistance-Capacitance model (Rc model) (Bacher & Madsen, 2011). In the case of the LEA model, the RC model is used as a white box modelling method, with as data input physical building characteristics (Jurado López, 2017). This model approach is a grey-box model when some parameters are known and based on physical laws, while others are empirically determined (Drgoňa et al., 2020). This model is built with the modelling of the heat fluxes through a thermal network (based on electronic circuits) of a building. In Figure 10 a thermal network is shown, where the heat flow flows through (shown with an arrow in the circuit). The heat flux is influenced by the thermal resistors and capacities of several temperature nodes. In Figure 10 the temperature nodes are; Ti (temperature of the interior), Te (temperature of the envelope) and Ta (temperature of ambient (outdoor air)). The heat gains from the heater (ϕ_h) and the solar radiation (ϕ_s). The heat flows through the circuit and so through the temperature nodes, are determined by the thermal resistance (R) between the nodes and the heat capacity (C) of the temperature nodes ((Bacher & Madsen, 2011). The driving force in this circuit (comparable to the voltage in an electrical circuit) is the temperature difference between the different temperature nodes. The parameters are estimated by the maximum likelihood. This gives the advantage that the used building characteristics are limited (Bacher & Madsen, 2011).
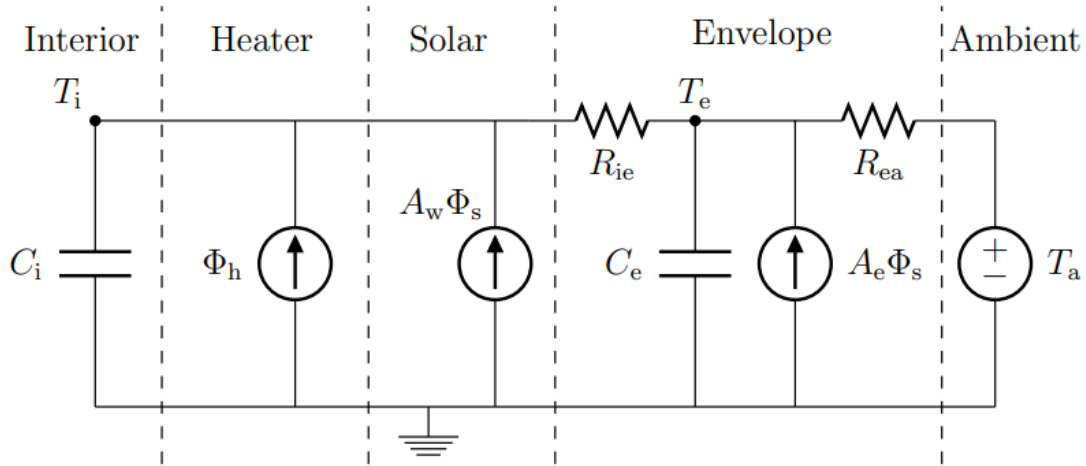
Figure 8: a simple thermal network as part of a RC-model, divided into different parts of the building(Bacher & Madsen, 2011)

The advantage of the implementation of a grey box model in MPC over the other model approaches is that the equation used in these models van be easily adapted to other similar buildings, this results in that only a few models can be used to represent a majority of the building (Drgoňa et al., 2020). Besides that these models can be easily adapted to the needs of MPC solver (Drgoňa et al., 2020):

1. Ensuring continuity: the equations do not have abrupt changes)
2. Linearity: equations are simple and fast to solve
3. Differentiability: differentiability ensures that the equations can be differentiated, which is required for calculating gradients and performing optimization in MPC.

2.2.5 Choice of model

Figure 11 summarizes the implementation of white, grey, and black box models in MPC. Based on the descriptions of each model approach, the choice has been made to use a grey box model (Drgoňa et al., 2020).

The specific grey-box model selected for this study is a multivariate linear regression model. A RC models can be effective for describing physical thermal processes in buildings (Drgoňa et al., 2020). However multivariate linear regression (MLR) model offers significant advantages in terms of simplicity, computational efficiency, flexibility and robustness ((Korolija et al. , 2013). These features are essential to be suitable for use in MPC applications, where fast and reliable optimization of control actions is crucial for achieving energy efficiency and comfort goals (Drgoňa et al., 2020).
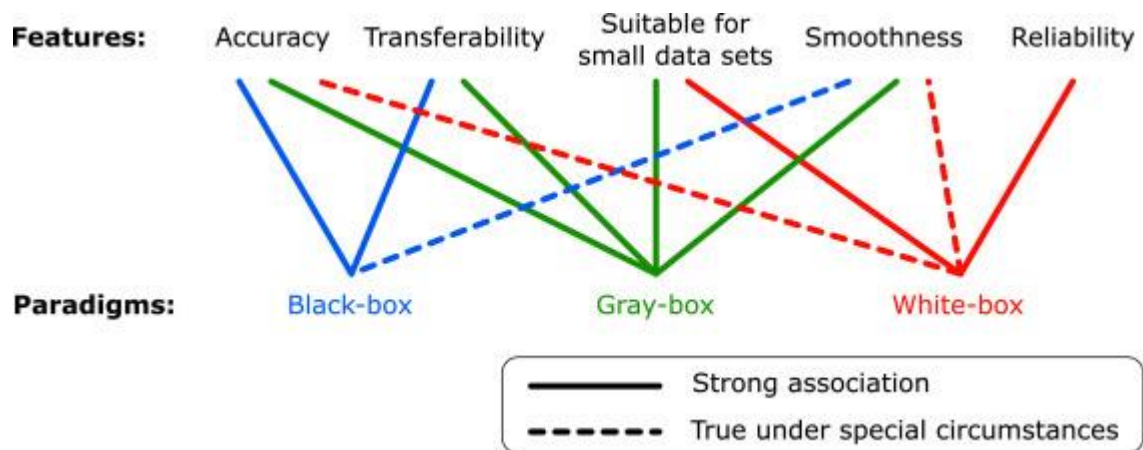
Figure 9: the features that can achieve with white, black or grey models within MPC (Drgoňa et al., 2020)

3. Research methodology

3.1 Research gap

Korolija et al. (2013) used several linear regression models to predict the annual heating, cooling and auxiliary energy requirements for HVAC systems. (Catalina et al., 2013) were able to predict the monthly heating demand for single-family residential sector in temperate climate. The input of the regression model were the building shape factor, the building envelope U-value, the window to floor area ratio, the building time constant and the climate. The data input of the multivariate linear regression model of (Catalina et al., 2013) were the building global heat loss coefficient, the south equivalent surface and the difference between the indoor heating set point and the sol-air temperature.

No study has been conducted to use a multivariate linear regression model with the indoor surface temperature as a data input, except the research of Jurado López (2017). Heat is stored within the thermal mass and released when the indoor temperature is lower than the temperature of the thermal mass. The indoor surface temperature has a damping effect on the heat transfers within the building. By including the indoor surface temperature as one of the data inputs of a linear regression model, the amount of variables could be reduced. Jurado López (2017) has developed a multivariate linear model with only inputs from the outdoor temperature and the indoor temperature. The heating demand is predicted during the opening hours of three TU Delft buildings. However, there was no actual data available, a physical-based model was used to simulate the data inputs. The research gap is that this model needs to be validated with actual data, and when needed to improve this model and when possible to predict the heating demand for 7 hours. This prediction time is set by the Brains4Building project, with taking in considering the horizon of the other parts of MPC, for example the electricity market and the power plants.

This research is building further on the model developed by Jurado López (2017). In this chapter the developed multivariate linear model by Jurado López (2017) to predict the heating and cooling demand is described.

3.2. Multivariate linear regression model method

The key objective of building a mathematical model with the linear regression model is to combine variables, to define a correlation to predict the heating or cooling demand of a building.

Equation 10 shows the analytical form of a multivariate linear regression model to model the heating or cooling demand. The dependent variable (in this case the heating or cooling demand) is predicted with the use of independent variables and coefficients. Examples of used independent variables are weather conditions and measured temperatures. These independent variables can be measured, which makes this model relatively simple to use. In the case there are multiple independent variables, a multivariate regression model is developed. The building and system characteristics are captured in the coefficients, which are obtained from training the model with historical data. As explained in Chapter 2.2.2. the building and system characteristics are not always known, for example, in a white box modelling method, however, with this type of model, a grey box modelling method, these building and system characteristics do not need to be exactly known.

$$Q_i = C_o + \sum_{j=1}^{m} C_j \cdot X_{i,j} \quad (13)$$

- Q: the dependent variable, the (hourly) heating or cooling demand
- C_0: the constant, the expected value of the dependent value when all the independent variables are set to zero. This constant is in most case positive (when the dependent variable is positive), however, it can be negative when the model is inaccurate, which results in a compensating constant.
- C_i: the coefficient corresponding to the impact of the independent variable on the dependent variable
- X_i: the independent variable, with M parameters

The independent variables change over the used time step, therefore it is needed to express this formula in a matrix form (shown in Formula 11). This has also the advantage of making it more efficient to use this approach in the software MATLAB.

$$\begin{bmatrix} Q_1 \\ \vdots \\ Q_N \end{bmatrix} = \begin{bmatrix} 1 & X_{1,1} & \cdots & X_{1,M} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{N,1} & \cdots & X_{N,M} \end{bmatrix} \begin{bmatrix} C_0 \\ \vdots \\ C_0 \end{bmatrix} \quad (14)$$

$$\begin{cases} Q_1 = C_0 + C_1 \cdot X_{t=1,1} + \ldots + C_p \cdot X_{t=1,p} \\ \vdots = C_0 + C_1 \cdot \vdots + \ldots + C_p \cdot \vdots \\ Q_N = C_0 + C_1 \cdot X_{t=x,1} + \ldots + C_p \cdot X_{t=x,p} \end{cases} \quad (15)$$

- t: time step, for example, hour 1 until hour 8760 for a whole year, then N = 8760
- C_0: constant

Formula 14 is expressed by a system of equations in Formula 15. Per time step the heating or cooling demand is calculated with this formula, where the independent variables are changing

per time step, which results in a different heating or cooling demand per time step. The dependent and independent variables are known, and with this data, the model is trained to determine the best-fitted coefficients (what corresponds to the relation between the dependent and independent variables) and the constant.

With this approach, the most significant parameters are chosen to determine the heating or cooling demand with the highest accuracy possible. This accuracy is determined by the pre-selection of the parameters, to be candidate to use in the model. The selection of parameters will be based on the thermal energy balance (described in Chapter 2.2.1). With these parameters, all different possible linear equations are tested based by the stepwiselm function in MATLAB. The accuracy and the selection of the parameters are established with a search procedure and statistical criteria described in Chapter 3.4. With this MATLAB function it is possible to use *'linear'* or *'interactions'* in the stepwiselm code. The first option includes only the independent variables as a single data input, for example, the outdoor temperature ($T_{outdoor}$). The second option, gives the possibility to use a combination of two independent variables, which are correlated to the heating or cooling demand, for example the outdoor temperature ($T_{outdoor}$) and the indoor temperature ($T_{indoorair}$). There are two options of combing these two independent variables, it could be multiplied by each other or divided by each other, the function determines which option has the most influence on the heating or cooling demand.

3.3 Data input model and formulas

In Chapter 2.2.1 the thermal energy balance is described. As the multivariate linear regression model is categorized as a grey box model, the independent variables are selected on this thermal energy balance. Equation 1 derived from the thermal energy balance is rewritten to a multivariate linear regression equation, shown in Equation 16.

$$Q_{demand} [W] = C_o + C_1 (T_{outdoor} - T_{indoorair}) + C_2 (T_{floor} - T_{indoorair}) + C_3 (T_{out\,AHU} - T_{indoorair}) + C_4 (V_{wind} \cdot (T_{outdoor} - T_{indoorair})) + C_5 (Q_{solar}) + C_6 (Q_{internal}) + C_7 (T_{surface} - T_{indoorair}) \quad (16)$$

In Table 1 an overview is given of the physical significance of the coefficients. The coefficients are calculated by the multivariate linear regression model, therefore for the associate variables in Table 1 do not need to be included in the model.

Table 1: physical significance of Equation 16

Coefficient	Physical significance
C_1	$C_1 \sim \sum_i U^i \cdot A^i$
C_2	$C_2 \sim U_{floor} \cdot A_{floor}$
C_3	$C_3 \sim M_{vent} \cdot C_{p\ air}$
C_4	$C_4 \sim (m_{openings} + m_{cracks}) \cdot C_{p\ air}$
C_5	$C_5 \sim effect\ of\ solar\ radation$ $on\ thermal\ demand$
C_6	$C_6 \sim effect\ of\ internal\ heat\ gains\ on\ thermal\ demand$
C_7	$C_7 \sim \alpha_i \cdot A_{indoorsurface}$

The variables that could be included in the model to follow the thermal energy balance, based on Equation 16 are:

- $T_{outdoor}$: outdoor temperature
- $T_{indoorair}$: indoor air temperature
- T_{floor}: floor temperature
- $T_{out\ AHU}$: temperature of ventilation air entering the room
- V_{wind}: wind speed
- Q_{solar}: solar heat gains
- $Q_{internal}$: internal heat gains
- $T_{surface}$: indoor surface temperature

The multivariate linear regression model will calculate the coefficients of each variable and will determine if a variable needs to be included based on the statistical validation and search procedures as described in Chapter 3.4. In this research, the choice is made to exclude $\boldsymbol{T_{out\ AHU}}$ as an independent variable, to avoid double counting as this variable is used to calculate the heating and cooling supply of the ventilation (see Chapter 5.5). Jurado López (2017) developed two models based on these principles.

Model 1: dynamic model excluding the indoor surface temperature

The first model of (Jurado López, 2017) was built with the independent variables of measurable temperatures and variables, however, in this model the indoor surface temperatures are excluded, because these temperatures are often not measured. The general equation of this model is Equation 17 within Table 2 the physical significance.

$$Q_{demand\,(t)}\,[W] = constant + \; C_a\,(T_{outdoor}\,(t)) + \; C_b\,(T_{indoorair}\,(t)) + \; C_c\,(V_{wind}\,(t)) +$$

$$\cancel{C_e\,(T_{out\,AHU}\,(t))} + \; C_{f,int,1a}\,(Q_{internal,1a}(t-1)) + \; C_{f,solar3a}\,(Q_{solar,3a}\,(t-3)) +$$

$$C_g\,(Q_{solar}\,(t)) + C_h(Q_{internal}\,(t)) \;\;(17)$$

Table 2: detailed overview of the physical significance coefficients of the coefficients used in Equation 17

Coefficient	Physical significance
C_a	$C_a \sim (U_{floor} \cdot A_{floor}) + (\dot{V}_{openings} \cdot \rho_{air} + V_{building} \cdot {}^{0.15}/_{53} \cdot \rho_{air})$
C_b	$C_b \sim (U_{floor} \cdot A_{floor}$ $\quad - \; (\sum_i U^j_{envelop} \cdot A^j_{envelop}) - (\dot{V}_{openings} \cdot \rho_{air} + V_{building} \cdot {}^{0.15}/_{53}$ $\quad \cdot \rho_{air}) - (m_{vent} \cdot C_{pair}) - (\alpha_i \cdot A_{indoor\,surfaces})$
C_c	$C_c \sim \dot{V}_{openings} \cdot \rho_{air} + V_{building} \cdot {}^{0.15}/_{53} \cdot \rho_{air}$
$\cancel{C_e}$	$\cancel{C_e \sim m_{vent} \cdot C_{pair}}$
$C_{f,int,1a}$	$C_{f,int,1a} \sim \alpha_i \cdot A_{indoor\,surfaces}$
$C_{f,solar3a}$	$C_{f,solar,3a} \sim effect\;variation\;total\;horizontal\;solar\;radiation\;on\;the\;heating\;demand$
C_g	$C_g \sim effect\;of\;solar\;radation\;on\;thermal\;demand$
C_h	$C_h \sim effect\;of\;internal\;heat\;gains\;on\;thermal\;demand$

Within this model, the indoor surface temperatures are replaced by making the model dynamic, by including the internal heat gains of 1 hour before ($C_{f,int,1a}$) and the solar heat gains by 3 hours ($C_{f,solar3a}$).

The accuracy of the model's predictions for the dependent variable varied from 73.5% to 90.7% in the case studies of Jurado López (2017)

Model 2: static model including the indoor surface temperatures

In the second model of (Jurado López, 2017), the indoor surface temperature is included along with the independent variable outdoor temperature, resulting in an impressive R_2 value range of 96% to 99% With excluding the indoor surface temperatures it was possible to exclude the other variables; T_{indoor}, $T_{out\,AHU}$, V_{wind}, Q_{solar} & $Q_{internal}$. This could be done because the influences of these variables on the heating or cooling demand are captured by the thermal mass of the building. The general equation of the second model is:

$$Q_{demand}\ (t)[W] = constant +\ C_a\ (T_{outdoor}\ (t)) +\ C_f\ (T_{indoorsurfaces}^{t}\ (t))\ (18)$$

Table 3: detailed overview of the physical significance coefficients of the coefficients used in Equation 18

Coefficient	Physical significance
C_a	$C_a \sim (U_{floor} \cdot A_{floor}) + (\dot{V}_{openings} \cdot \rho_{air} + V_{building} \cdot {0.15}/{53} \cdot \rho_{air})$
C_f	$C_f \sim \alpha_i \cdot A_{indoor\ surfaces}$

Both models are developed by simulated data (by the white box model: LEA). Within this research, the models will be tested with actual data and, when necessary, improved. The used function of modelling and predicting the heating and cooling demand is *stepwiselm* Mathworks (n.d.)with the discussed in the next section statistical validation and search procedure.

3.4 Statistical validation and search procedure

To build a multivariate linear regression model with dependent variables and independent variables, a statistical validation and search procedure is used, to select the significant variables, wherefore an hypothesis is described in Chapter 6.2.

The data for the opening hours is split up into two parts:

1. **Training data set**: the first part (80% of the data set) including the heating or cooling demand is used to train the model, whereby the modelled heating or cooling demand is compared to the actual heating and demand based on the described statistical concepts in this chapter.
2. **Test data set**: The other 20% of the data set is used with the developed model to predict the heating or cooling demand. The accuracy of the prediction is determined by the comparison between the heating or cooling demand of the dataset and the predicted heating or cooling demand.

The model is created using the MATLAB statistical toolbox's stepwise regression function. This function helps to develop a regression model for the dependent variable - in this case, either the heating or cooling demand - based on different independent variables. By testing various independent variables, the function determines the most effective combination of variables to accurately predict and model the thermal power. The statistical validation and search procedure consists of three major analysis concepts:

1. *Residuals of the data set*

Residuals refer to the discrepancy between the predicted or modelled dataset and the actual observed data. The ultimate objective of utilizing a model is to ensure that the predictions are as accurate as possible. Therefore, the mean of the residuals must be (almost) zero for each value of **X**, the variance is approximately constant for all **X** values, and the distribution conforms to a normal distribution.

2. *Individual significant level of the variables' coefficients estimated*

The significance of variable coefficients is assessed using p-values and t-statistics to test hypotheses and determine the significance of independent variables to determine which variables need to be included in the linear regression model (Chatzithomas et al., 2015).

- The **p-value** measures the probability that the observed data would occur under the null hypothesis, indicating the statistical significance of the results. When the p-value is less than 0.05, a significant level for rejecting the null hypothesis is indicated. A p-value below 0.05 is indicative of statistical significance and allows for the rejection of the null hypothesis. This means that there is a less than 5% chance that the observed results occurred purely by chance (Chatzithomas et al., 2015). The parameters with a p-value of less than 0.05 will be included in the model while parameters with a p-value greater than 0.1 will be removed. These thresholds are based on the stepwise regression function.

  Despite the standard exclusion of the independent variables with a p-value > 0.05, the results showed that independent variables with a p-value <0.05 were included, therefore an for-loop is added, to adjust the model by extracting the insignificant independent variables. In some cases, the constant has a p-value above 0.05, the constant is not checked with this for-loop.

- The **t-statistics** provides information on the significance of a variable's contribution to the model. A higher t-statistic value indicates a greater contribution to the fit of the curve. This value is calculated by dividing the estimated coefficient of a variable in a regression model by the standard error of that coefficient (Chatzithomas et al., 2015). The t-statistics value can be negative or positive, what indicates a negative or positive relation between the independent and dependent variable.
-

With each introduction of an independent variable in the model, the p-values and the t-statistics will be calculated. The variables with a p-value between the range will be included in the model and their priority will be determined based on their t-statistics.

3. *Significance of the model*

Once the independent variables have been chosen, it is crucial to evaluate the model's significance. There is a risk of overfitting the model if too many independent variables are selected, which can lead to poor predictive performance. The model may become too sensitive to minor changes in the training dataset, resulting in overreacting to these changes. To avoid this, another concept in statistics is used: the coefficient of determination (R^2), the adjusted coefficient of determination (R^2 adjusted), and the root mean square error (RMSE).

The coefficient of determination (**R^2**) is a statistical measure that determines the accuracy of a model to predict the dependent variable. The value of R^2 lies between 0 and 1, where 0 signifies that the linear regression model cannot predict the dependent variable and 1 indicates that the model can accurately predict the dependent variable. Values between 0 and 1 indicate that the model can predict the independent variable to some extent. In most cases the R^2 value is positive, in some extraordinary cases, this value can be negative, when the model performs very poorly. The goal is to achieve the highest R^2 possible, but this could result in overfitting (Everitt & Skrondal, 2020). To avoid this, the adjusted R^2 and RMSE are used.

- The adjusted coefficient of determination (**R^2 adjusted**) is a measure of how well the data points fit on a curve or line, which takes into account the number of terms used in the model. The R^2 adjusted value will increase when significant parameters are added to the model but will decrease when less significant or non-significant parameters are added (Everitt & Skrondal, 2020). Also the R^2 adjusted value is normally positive, but can be in the case of a very inaccurate model negative, what can be caused by insignificant variables, what results in adding noise instead of information.

- The Root Mean Square Error (**RMSE**) indicates the standard deviation of the actual data points from the regression line. A lower RMSE indicates a relatively good prediction of the dependent value (Everitt & Skrondal, 2020). Another option is the Mean Absolute Error (MAE). This is a measure of prediction accuracy that calculates the average absolute difference between the predicted and actual values, giving equal weight to all errors (Everitt & Skrondal, 2020). The advantage of RMSE over MAE is the sensitivity to outliers. This means that larger errors have a disproportionately large impact on the RMSE, which is useful for this application where accuracy for each time step is critical (Everitt & Skrondal, 2020).

3.5 Data analysis with Pearson-correlation coefficients

To understand the correlation between the chosen parameters (based on the thermal energy balance (as described in Chapter 2.2.1 and Appendix A), x in Equation 19, and the heating or

cooling demand, y in Equation 19, correlation matrixes will be made, presented in Chapter 6.2. The values within these matrixes are calculated in MATLAB with the Pearson-correlation coefficient principle, shown in Equation 19:

$$r_{xy} = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{[n \sum_i x_i^2 - (\sum_i x_i)^2][n \sum_i y_i^2 - (\sum_i y_i)^2]}} \quad (19)$$

The coefficients could be positive or negative, which indicates the relationship. In the case of a positive relation [> 0, max 1], the heating or cooling demand increases when the parameters increase, for the negative relations [< 0, max -1] this is the inverse, when the value of the parameter increases, the heating or cooling demand decreases. The correlations are categorized with the ranges in Table 4, after the categories in Zou et al. (2003).

Table 4: categorisation of the Pearson correlation values, based on Zou et al. (2003)

Pearson correlation coefficient (*r*) value	Strength	Direction
> 0.5	Strong	Positive
0.3 to 0.5	Moderate	Positive
0 to .3	Weak	Positive
0	None	None
0 to − 0.3	Weak	Negative
−0.3 to − 0.5	Moderate	Negative
< − 0.5	Strong	Negative

3.6 Calculation of the delay (n)

Within the research of Jurado López (2017), the time delay (n) of the internal heat gains and the solar heat gains are determined with the use of a graph of the normalised independent variables and the dependent variable: heating or cooling demand. This method is applied on the data set of this research, shown in Appendix J. It is hard to conclude from this graph what the time delay between the cooling demand the solar light intensity and the internal heat gains is, as the patterns are not following each other.

Therefore the choice is made to determine the delay by the cross-correlation function in MATLAB. This function calculates based on the Pearson-correlation coefficient at which time delay the correlation between the independent variable and the dependent variables is the strongest. Thereby the maximum possible time delay is set on 1 day (144 timesteps with timesteps of 10 minutes). In contrast to the scope of this research: the opening hours, the time delay is calculated with the data set of all the hours. With this the independent variable values of before the opening hours are also included. For the static models the data set for only the

opening hours is used. In the case of the dynamic model the dataset of all the hours is used, whereby the condition is set that only the heating or cooling demand is included during opening hours, whereby it is possible use the values of the delayed independent variables from before the opening hours. The independent variables with a delay are included twice, first with the current value, so at moment t, second with the delayed value, at moment t − n. It was not possible to use the values between these two points, because the modelling time was exceeding the practical use.

3.7 Dealing with NaN and interpolation

Within the gathered data, there are some points with no data, expressed with NaN. This can be due a fault in the meter for a short period. In the whole code it is included that these rows will be skipped. Another option is to change NaN to 0, however, this gives a distorted picture of this value, therefore the choice is made to keep these values as NaN. There are independent variables with another timestep, these data points are linearly interpolated with the '*interp1*' function in MATLAB. The independent variables wherefore this method is used will be explained in Chapter 5.

4. Description and analysis case study

The building and operational characteristics of the case study are described to gain knowledge to analyse the data set and to determine the available parameters of the predicting model, whereby the following sub-question will be answered:

What are the main building and operating characteristics and how is the building heated and cooled?

4.1 Overview building

In 2009 the location of the Haagse Hogeschool, Academy for Engineering (around 1200 students) was built on the TU Delft campus (shown in figure 12). The building was the first educational building with an energy label of A++ (Bouw Wereld, n.d.). The building automation system within the building can be controlled on a demand-driven basis. This demand is primarily determined at room level by the presence of users and their wishes for the indoor climate. The building is designed to facilitate study material for student to research, which makes this building a suitable case study. Plans are made to install wind turbines and make use of a fuel cell when these technologies are further developed (Syb van Breda & Co, n.d.). There are four building layers, including a basement used for the installations and as bicycle storage (Syb van Breda & Co, n.d.). Cars can park on the roof, where also solar panels and collectors, and an asphalt collector are installed. These solar panels are used as study material, for research, the solar collectors are used for hot tap water. The inside walls (from now called system walls) can be replaced, which makes the building flexible.
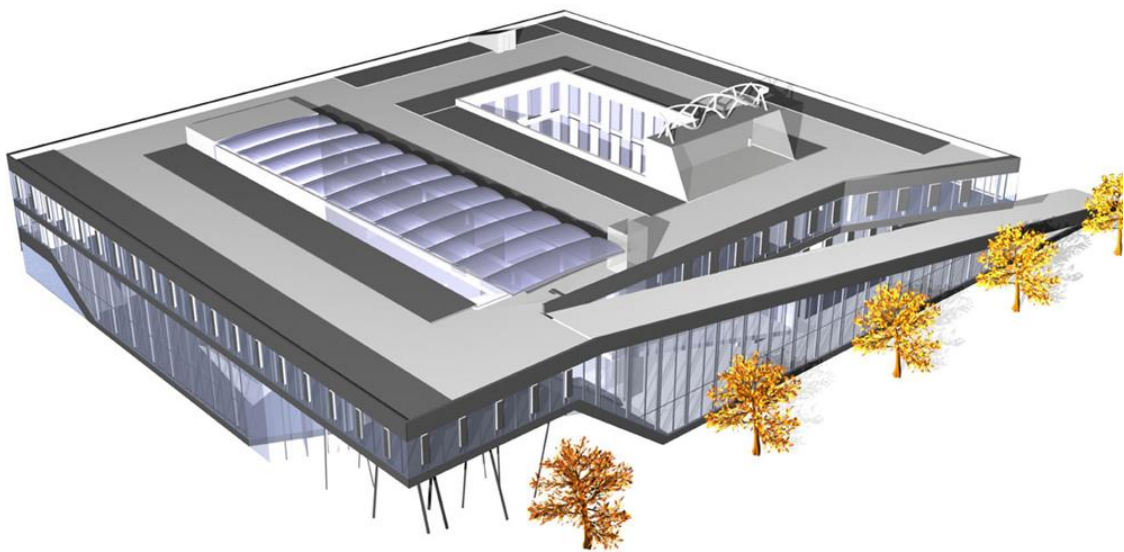
Figure 10: Drawing of Haagse Hogeschool in Delft (Syb van Breda & Co, n.d.)

4.2   Building characteristics

The building characteristics are based on a non-public report by Peutz.

An overview of the building characteristics is given in Table 5. The building has a relatively high envelope/floor ratio (= 0.76), which in general indicates a higher heat gain in the summer and heat loss during winter. The building has a relatively high window/façade area ratio, which indicates a high influence of solar heat gains on the thermal demand (the highest at the North-West façade and the lowest at the North-East façade indicated by the window/façade ratio of 0.60 and 0.25, respectively). These effects are reduced by the relatively high insulation value of the façade (Rc = 4 m²K/W and U = 0.24 W/m²K) and the low g-value (= 0.25) of the windows. This indicates that there will be a small correlation between the thermal demand and the outdoor parameters. The floors with floor heating and cooling and concrete construction columns are expected to have a relatively higher thermal mass (= 400 kg/m²) than the inner system walls (thermal mass unknown). However, the time-delaying and dampening effect on the air temperature described in Chapter 2.2.1 is expected to be decreased by the high envelope/floor area ratio.

Table 5: Overview of the building characteristics of the case study

Parameter	Value	Unit
Floor area	12150	m²
Educational function	4140	m²
Meeting function	4987	m²
Office function	1993	m²
Common function	985	m²
Ratio envelop area (walls + windows + roof)/floor area*	0.76	-
Insulation value (Rc)	4	m²K/W
U-value façade	0.24	W/m²K
U-value windows	1.7	W/m²K
g-value windows	0.25	-
U-value doors	3.2	W/m²K
Specific thermal mass of building (floor)	400	kg/m²
Ratio window area/façade area		
South-East	0.29	-
North-East	0.25	-
North-West	0.60	-
South-West	0.58	-
Roof	0.12	-

* excluding the basement

4.3 Thermal systems

Information provided about the thermal systems is based on a non-public report by Peutz and the technical description of the 'naregelingen' of DWA.

The building is equipped with heat and cold storage in the ground in combination with an electrical heat pump (ground water source), an asphalt collector (500 m²) on the roof, and HR boilers (to supply heat during extremely high demand peak moments). A more detailed view of the building including the installation can be found in Appendix B. Under the building, there is an aquifer thermal storage, where warm and cold water is separately stored for one season.

In the summer, the building is cooled with cold water, which is heated up after cooling and injected back into the warm source. During the winter, warm water is extracted from this warm source, and with a heat pump, this heat is brought up to the desired temperature to heat the building. The cooled-down water will be injected into the cold sources. In the Netherlands, it is obligatory to keep the heat and cold storage in balance; therefore, the asphalt collector is installed on the roof. Water flows through the pipelines through the roof and driving lane. Additionally, there is also an air heat exchanger for additionally loading cold or warmth, which helps in balancing the thermal energy storage more efficiently.

As stated earlier, on the room level there is more cooling than heating demand; however, on the building level, this is reversed. This can be explained by the fact that the asphalt on the roof is heated by the collector to prevent frost on the roadway for the cars on the roof.

Connected to the aquifer thermal storage is a heat pump. This heat pump provides warm water to the floor heating system and the air handling unit. In the summer, the heat pump works as a chiller and delivers cold water to cool the rooms with the floor system, climate ceiling, additional active ceiling panels and air handling unit. This heat pump can deliver heat and cold at the same time. Cold can also be directly delivered by the cold storage of the aquifer thermal storage when the cold load is not too high, the heat pump is not used as a chiller

The floor is heated or cooled by a floor heating and cooling system, per floor there are two groups; North and South. The heating and cooling pipes go to the floor groups from the heat pump and boiler. Per group heating or cooling is possible, but it is not possible to heat or cool at the same time with this installation. Each room has its valves, which can be opened or closed to receive heat and cold (for the ventilation system this is shown in Appendix D), which is determined by the occupancy and temperature sensors in the room. The floor system supplies the base heat or cold, however, the disadvantage of this system is the time delay due to the

thermal mass of the system. Additional heat can be supplied by the climate system where mechanical ventilation is installed. This system is also used for additional cold, besides that, in some rooms additional active ceiling panels (shown in Figure 13). These heating and cooling systems are summarized in Table 6.

Besides additional heating and cooling, the mechanical ventilation system is used to secure good air quality. This is an exhaust & supply system with heat recovery (exhaust air ducts shown in Figure 14). The amount of fresh air supplied by this system is determined by occupancy and temperature sensors in the rooms. Outside air is cleaned and preheated by the recovery (in the air handling unit (AHU), and there is additional heating and cooling through two coils connected to the heat pump/boiler or chiller. Humidity and temperature are regulated by extracting humidity from the extracted air (only during summer by deep cooling) and by recycling the heat.

In addition to the heat pump, HR boilers are installed. During high-demand peaks of heat, these boilers can deliver additional heat besides the heat pump. On the roof, 21 m² of solar collectors are installed to deliver together with the HR boiler, heat for hot tap water.

Figure 11: Active ceiling panels used for additional cooling in several rooms (own picture)

Figure 12: Exhaust air ducts of the exhaust & supply system to supply fresh air and for additional heating or cooling (own picture)

Table 6: Overview of installations for heating and cooling in the Haagse Hogeschool in Delft

Demand	Supplied by
Heating	Floor heating or cooling
	Air heating and cooling by mechanical ventilation
Cooling	Floor heating or cooling
	Air heating and cooling by mechanical ventilation
	Additional active ceiling panels

4.4 Operating characteristics

Information provided about the operating characteristics is based on a non-public technical description of the 'naregelingen' of DWA.

At room level, the desired air temperature is set, in general for offices at 21,5 °C and lecture rooms at 21 °C. In each room, there is a thermostat, which can be used to change the temperature set point. However, this option is not used now, the temperature setpoints are changed by the building managers, where occupants of the room can give their preference. In practice, each room has a different temperature set point. The room is cooled or heated when the temperature set point is higher or lower than the air temperature of the room. As the heating and cooling demand is connected to the air temperature (influenced by the occupancy, solar gains, infiltrations and the indoor surface temperature), a correlation between the air temperature and the thermal demand is expected.

To optimise the energy costs and to shave the peaks for heating and cooling, the installations are released or blocked according to time switches *('tijdschakelklokken'),* which are correlated to the opening hours:

- Opening hours: Monday – Thursday 8 - 23 and Friday 8 – 19
- Closing hours: Monday – Thursday 23 - 8, Friday 19 - 8 and weekend

Due to the thermal inertia of the floor heating and cooling, the time switch of this installation is during the early morning. When this installation is switched off, the ventilation and panels can switched on, according to the time switches. The ventilation is switched on when there is a release from the time switch, occupancy is measured for a minimum of 10 minutes or a high CO_2 content is measured. The ceiling panels are switched off when the windows are closed. During closing hours there is no heating or cooling, except when there is occupancy measured.

4.5 Selection of rooms

The Haagse Hogeschool Building in Delft is designed to be used as study material. The whole building is equipped with sensors. In the so-called test zone, there are more sensors installed. In Figure 15, a map of the test zone can be found. Rooms within this test zone are selected to use as a case study. Within this test zone, two rooms are selected based on the different usages of the room, orientation, area and thermal mass (as shown in Table 7). The rooms are supplied with heat and cold by the same type of installations; floor heating and cooling, ventilation and panels. The two rooms are opposite to each other on the first floor of the building, with a corridor in between, as shown in Figure 15.

The first room is an office for 6 teachers, at the south façade, with a floor area of 50.82 m². The office has a glass wall to the corridor and four windows in the south façade. There are six desks in the room. The other room is a classroom for 46 persons, at the courtyard, facing north, with a floor area of 78.11 m². The classroom has two glass doors to the corridor because the system walls are replaced, to get a bigger classroom. The four windows are facing the courtyard. The thermal mass of the classroom is higher than the office, as the office has a glass wall instead of a system wall. It is expected that the classroom is less exposed to weather conditions than the office, as the classroom is adjacent to the courtyard. These rooms are within the test zone with more sensors than the rest of the building, however, the indoor surface temperature is not measured by the existing sensors. Therefore, the indoor surface temperature is measured with self-placed sensors. In Figure 16 a simplified visualization of the two rooms is shown, with the placed indoor surface temperature sensors. In this Figure, only the used sensors for data collection are shown (further explanation in Chapter 5.7).

View from courtyard faced north

Façade south

Room 1.071
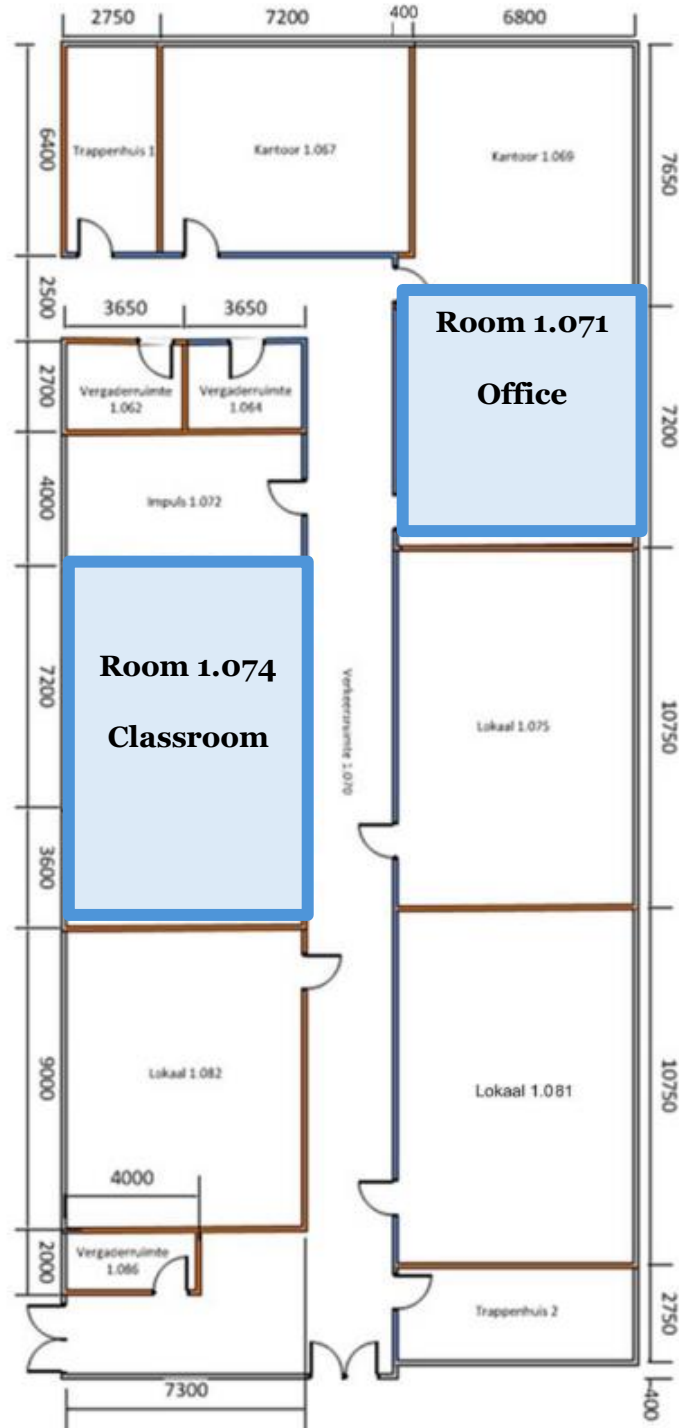
Office

Room 1.074

Classroom

Figure 13: map of the test zone with the selected rooms and pictures of the corresponding facades

Table 7: characteristics overview of the selected rooms

	Office: 1.071	**Classroom: 1.074**
Usage type	Office room teachers	Lecture room
Adjacent to	Outside façade	Courtyard
Oriented to	South(-East)	North(-East)
Maximum persons	6	46
Measurement period	03/08/2024 - 03/22/2024 (11 working days)	03/22/2024 - 04/11/2024 (15 working days)
Data points opening hours	851	1173
Area [m²]	50.82	78.11
Volume [m³]	152.46	234.33
Photo rooms		

Office

Classroom
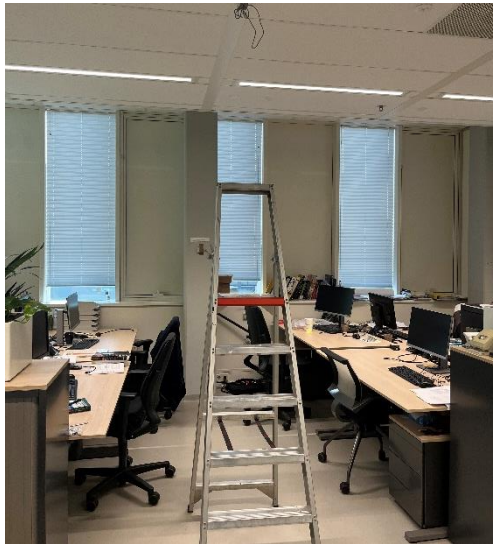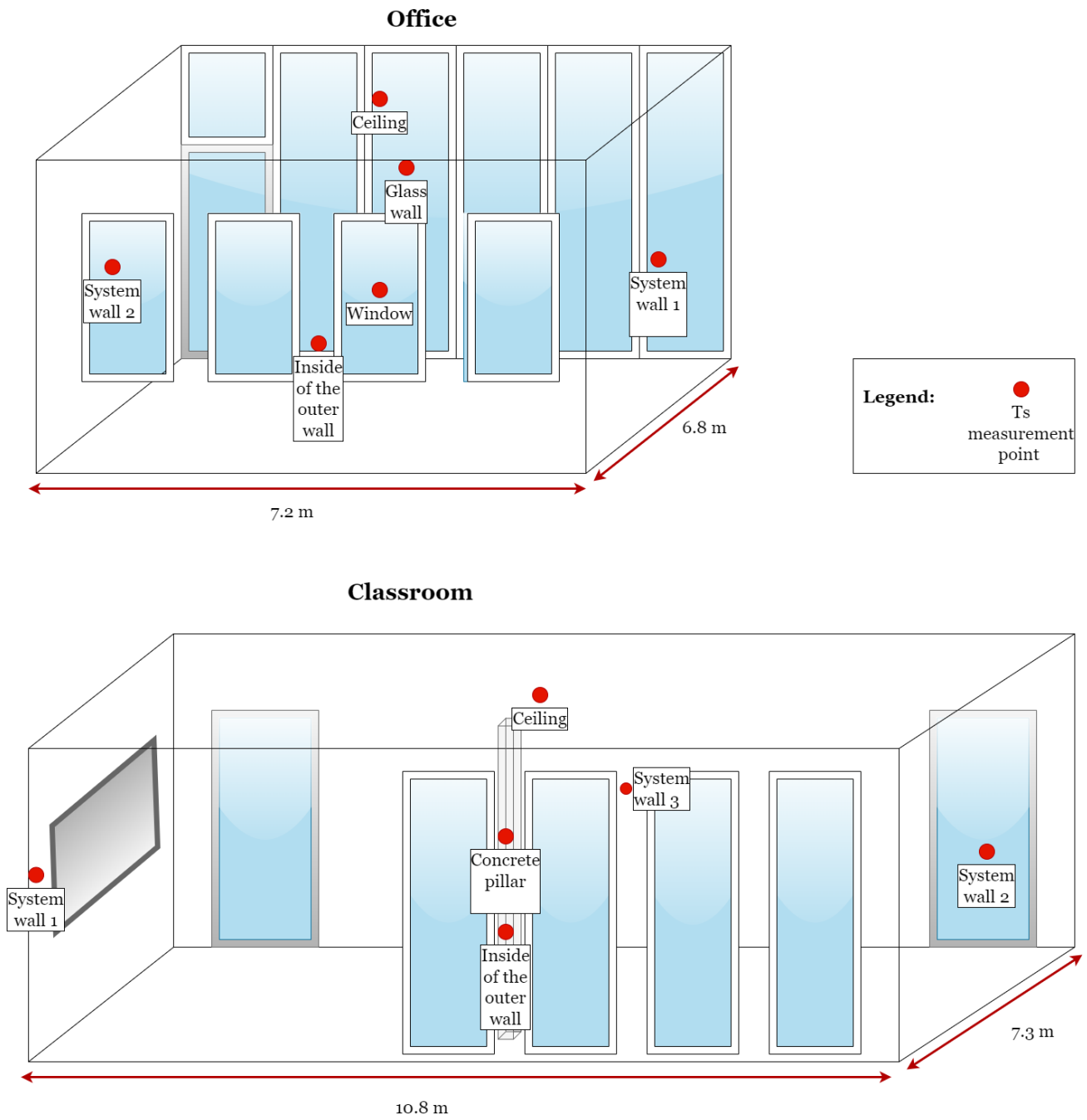
Figure 14: schematic overview of the rooms with the position of the windows and doors, with the sizes of the room. The used indoor surface temperatures are shown.

5. Data collection and preparation

In this chapter, the collection of the required data of the variables as determined in Chapter 3.3, for the multivariate linear regression model is described. Two datasets are made, one for opening hours (see Chapter 3.4) and one for all the hours. Both datasets have a timestep of 10 minutes.

5.1 Historical data

Historical data is available for the rooms within the test zone for the years 2014 and 2015, per hour. This dataset comprises significant parameters such as valve positions for the floor and the air system, temperature outflow of the floor system, supply air temperature, inside wall temperature, air temperature, and supply airflow. However, it is unclear where all the sensors were placed and what they exactly measured. Per room there are 2 or 3 heating and cooling demands calculated, however, it is not clear how these heating and cooling demands are calculated. Therefore the choice is made to use actual data.

5.2 Weather data $(T_{outdoor}, Q_{solar} \& V_{wind})$

Sensors are placed outside the building to measure weather parameters. The relevant parameters are wind speed, light intensity at different orientations and outdoor temperature. The light intensity is measured in lux, however with a factor it is possible to calculate the horizontal global solar radiation in W/m² (Michael et al., 2020), the influence of the sun will remain the same, and because of this, the choice is made to use the measured light intensity. In contrast to the data per 10 minutes from the parameters inside the building, the weather parameters are measured per 8 minutes. Therefore, the weather data is interpolated in MATLAB to convert the time step from 8 to 10 minutes.

5.3 Internal heat gains $(Q_{internal})$

To calculate the internal heat gains within the rooms, the number of persons in the room is needed. Unfortunately, the occupancy is not measured in the rooms. In the rooms there is a sensor, to measure if someone is there, this is shown as 1 (there are persons in the room) or 0 (no one in the room), however, this data is not stored. For the classroom, a schedule was available, with each time slot indicating the number of persons.

Equation 20 is used to calculate the occupancy based on the air change rate and the CO_2 emission per person, outside CO_2 concentrations and the inside CO_2 concentrations (Bokel, 2021).

$$n_{persons}(t) = \frac{\left(C_{CO_2}(t-dt) - C_{CO_2outside}\right) \cdot n \cdot V}{P_{CO2person}} \quad (20)$$

With:

- $n_{persons}$: persons in the room
- C_{CO_2}: CO_2 concentration in the room [ppm]
- $C_{CO2\_outside}$: CO_2 concentration outside
- $P_{CO2person}$: CO_2 emissions per person [ppm/hr]
- n : air change rate [h⁻¹]
- V: the volume of the room [m³]
- t: time [hr]
- dt: time delay [hr]

The time delay is calculated with Equation 21, with this Formula the time delay is calculated when the CO_2 concentration reached 95% of the constant value after a long time:

$$dt = \log(0.05) / -n \quad (21)$$

The CO_2 concentration is measured in each room. As outside CO_2 concentration, the minimum CO_2 concentration measured is used. The air change rate (ACH) was unavailable but is calculated with Equation 22.

$$n = \frac{Q}{V} \quad (22)$$

With:

- Q: ventilation airflow [m³/h]
- V: the volume of the room [m³]

This method was initially implemented in the classroom to compare the scheduled occupancy with the calculated occupancy. The results are depicted in Figure 17. The calculated occupancy corresponds to the CO_2 levels in the room, and the ACH is relatively low as shown in Figure 18. Equation 20 does not account for the effect of opening the windows, although the windows were seldom opened, as indicated in Appendix C. It can be seen that there is a lower ACH when

the windows are opened because the ventilation should stop when the windows open. There is no available data on the position of the doors. During high peaks of occupancy, the ACH is low, possibly due to low pressure from the ventilation system during those periods. There are multiple periods where the calculated occupancy differs from the scheduled occupancy, suggesting the unexpected presence of additional people or the absence of scheduled individuals. The calculated occupancy, derived from real-time CO_2 measurements and ventilation rates, likely offers a more accurate representation of actual occupancy. Scheduled occupancy data can be unreliable as it does not accommodate variations and anomalies in actual room usage.
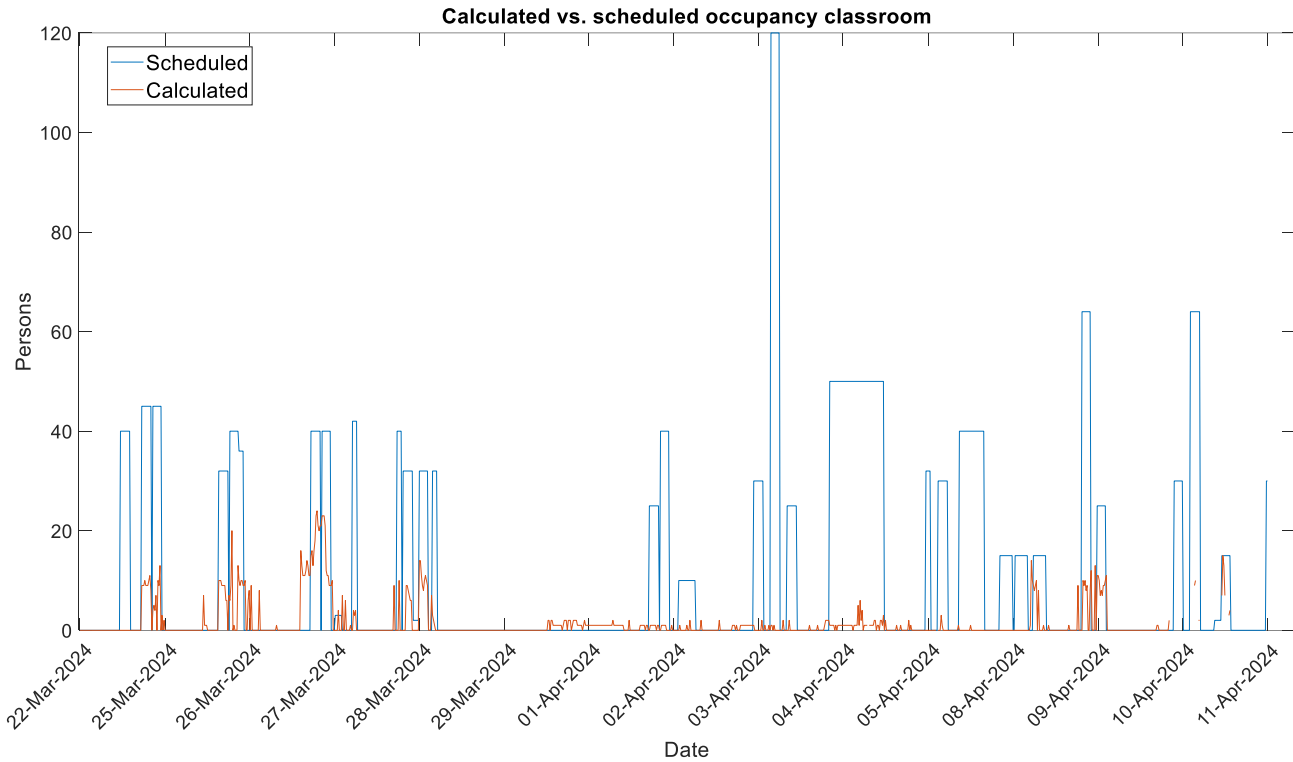
Figure 15: scheduled versus calculated occupancy in the classroom
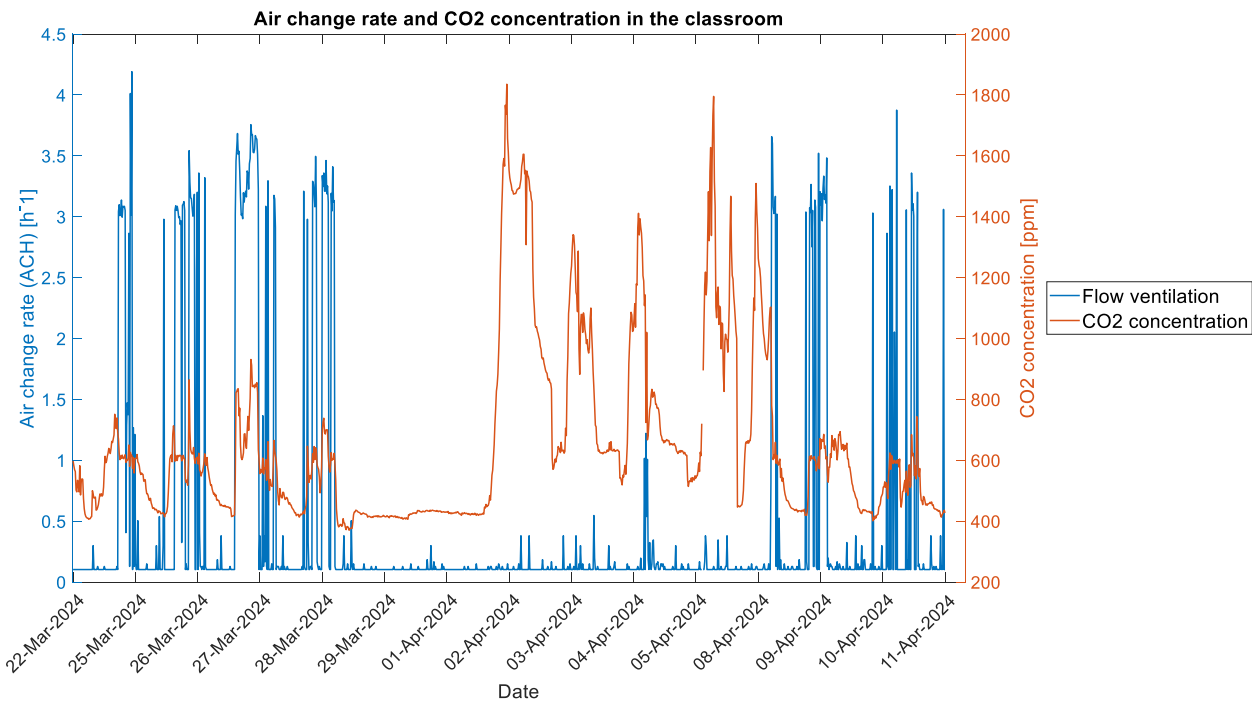
Figure 16: air change rate and CO₂ concentration of the classroom

The method is applied to the office where no reference data is available. There results of this calculation are shown in Figure 19. In the office room, there are places for six teachers. The calculated number of people is a maximum of three persons in the room. This seems plausible as the offices are mostly used by the teacher between giving the lectures. The ACH of the office

is lower than the classroom, while the volume of the classroom is 1.5 times bigger and the maximum CO_2 ppm level in the classroom is much higher than that of the office. Thereby in the office, there are no periods of lower ACH.
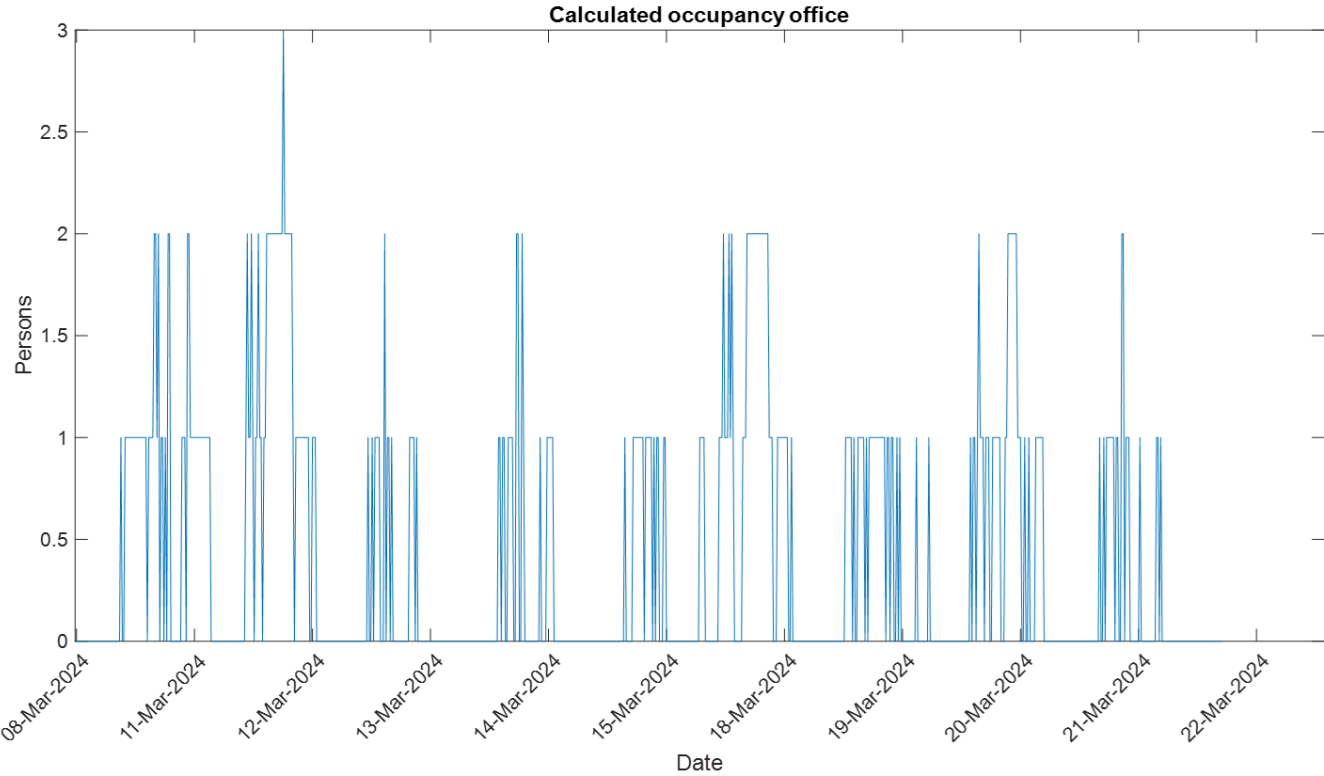
Figure 17: calculated occupancy in the office
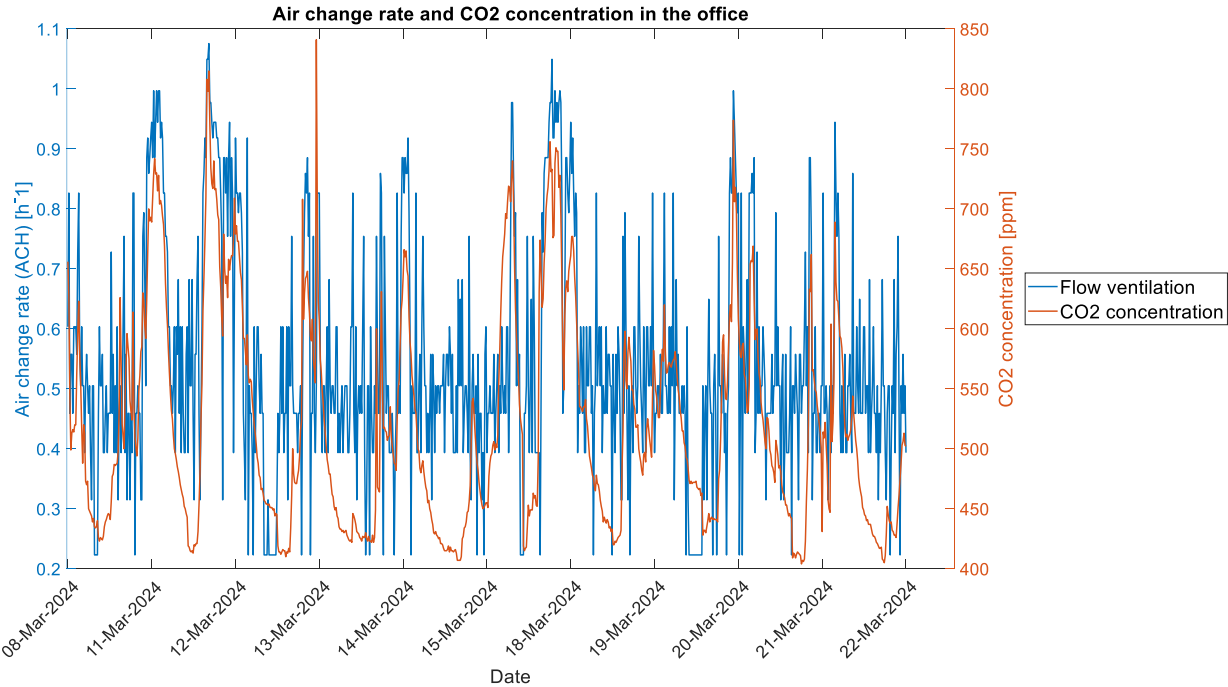
Figure 18: the air change rate corresponding to the CO_2 concentration in the office

With the calculated occupancy for the office room and the classroom, the internal heat gains are calculated with Equation 23 from the thermal energy balance as described in Chapter 2.2.1 and Appendix A:

$$Q_{internal} = n_{people} \cdot (Q_{body} + Q_{applications}) + A_{ceilings} \cdot Q_{light} + Q_{smartboard} \text{ (23)}$$

With:

- n_{people} : number of people
- Q_{body} : heat gain per person [W] (office = 117, classroom = 117)
- $Q_{applications}$: heat gain applications [W] (office = 80, classroom = 40)
- $A_{ceilings}$: total area of all the ceilings [m²] (office = 50.82, classroom = 78.11)
- Q_{light} : heat gain artificial light [W/m²] (office = 14, classroom = 14)
- $Q_{smartboard}$: heat gain smartboard [W] (office = 0, classroom = 175)

Equation 23 is used in a code in MATLAB to calculate the internal heat gains. In the code, it is added that there can be only an internal heat gain when there are people in the room, otherwise, the lightning is not switched on (as they are occupancy regulated) and the smartboard is not used. The values of $Q_{body}, Q_{applications}$ and $Q_{smartboard}$ are based on the values in (Itard, 2011). $Q_{applications}$ is the heat gain of applications, in this case, laptops. The value is higher for the office as there are desktops used.

5.5 Heating and cooling demand (Q_{demand})

The Haagse Hogeschool building in Delft is heated or cooled with floor heating and cooling, additional panels on the ceiling, and mechanical ventilation (as described in Chapter 4). The total heating and cooling is calculated with Equation 24. When Q_{demand} is negative (< 0 W), there is a cooling supply, when Q_{demand} is positive (> 0 W), there is a heating supply.

$$Q_{demand} = Q_{demand,floor} + Q_{demand,panel} + Q_{demand,ventilation} \text{ [W] (24)}$$

In Figures 21 and 22 the total thermal demand during opening hours is shown for each room.
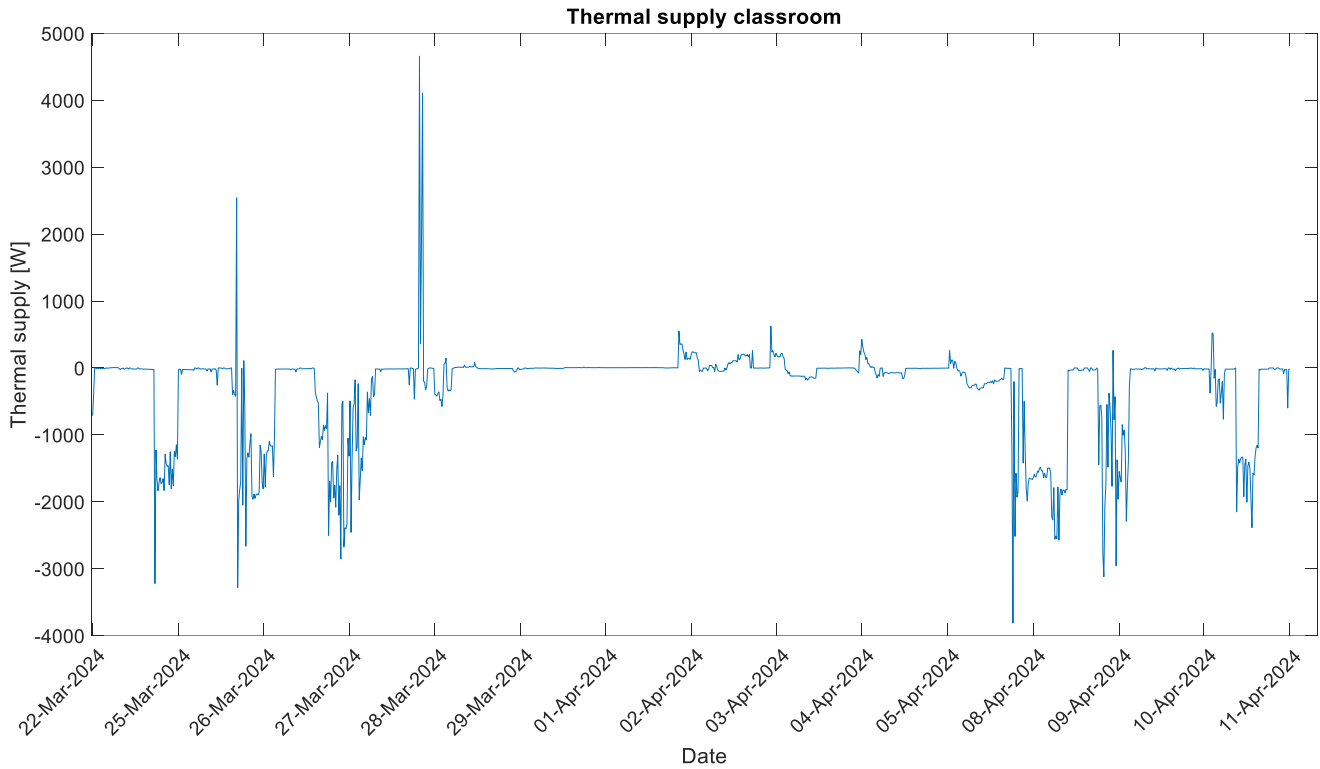
Figure 19: the total thermal energy supply in the classroom, max heat = 4659.37 W, max cold = 3811.79 W
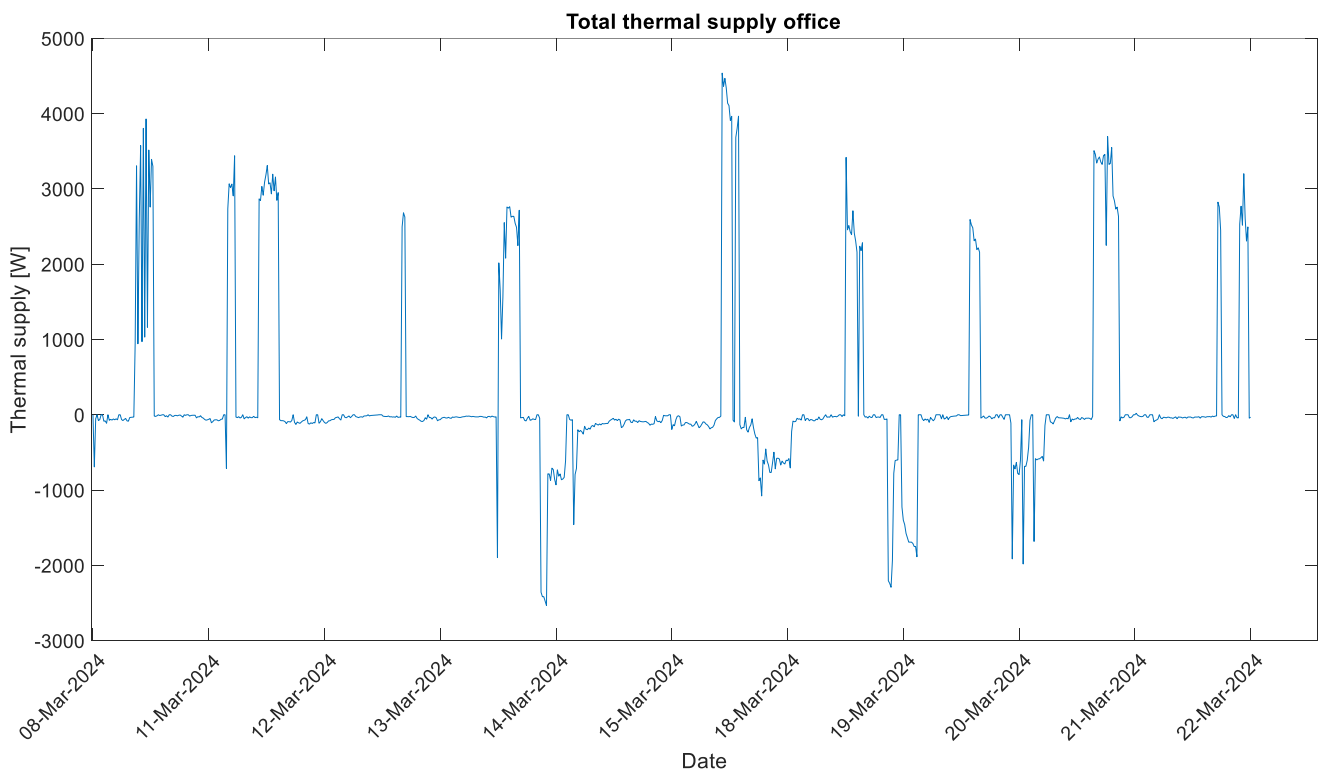
Figure 20: the total thermal energy balance in the office, max heating = 4538.18 W, max cooling = 2534.33 W

Floor heating and cooling ($Q_{demand,floor}$)

The heat or cold supply of the floor heating or cooling system is calculated with Equation 25. The water flow entering the room in the system is not measured. This missing data is replaced by the measured position of the valve (θ) in the room and the known maximum flows within the ground groups of the rooms ($\Phi_v, \_max$). Each room is connected to two groups, maximum flow of each group is summed up, to get the maximum flow per room. The valve position is measured in percentages, where for both rooms, a valve position was possible of 0% (closed), 99% (almost fully open) and 100% (fully open). This percentage is divided by 100, to get a factor of 0, 0.99 or 1. When the valves are closed, $\theta = 0$, so there is no heating or cooling supply. The density (ρ) and specific heat (c) of water is used, these are constant values. As the last step, the difference between the temperature of the water flow entering the room (T_{supply}) is subtracted from the temperature of the leaving water flow (T_{return}). When this ΔT is positive, heat is given off in the room (heating supply), when ΔT ($T_{supply} - T_{return}$) is negative, cooling happens the room (cooling supply).

$$Q_{demand,floor} = \theta \cdot \Phi_{v,max} \cdot \rho \cdot c \cdot \Delta T \quad (25)$$

With:

- $Q_{demand,floor}$: thermal power floor [W]
- θ: position of the valve (0 for closed and 1 for zero)
- Φ_v, max: max volume flow [l/s] (office: 0.071 + 0.070 l/s , classroom: 0.071 + 0.072 l/s)
- ρ: density (= 1000 kg/m³ for water)
- c: specific heat (= 4180 J/kgK for water)
- $\Delta T = T_{supply} - T_{return}$: the temperature of flow entering the room – temperature of flow leaving the room

Panels (Q_{panel}*)*
With flowing water through panels cold is supplied by panels in the rooms. The volume flow (Φ_v) is measured in the room, together with the constants of density (ρ) and specific heat (c) and the measured temperatures (measured in the room) of the flow entering (T_{supply}) and leaving the room T_{return}, the heat or cold supplied by the panels is calculated. In the office, there is one panel group. In the classroom, there are two panel groups, for each panel the heat or cold is calculated, and summed up to one value for the classroom.

$$Q_{demand,panel} = \Phi_v \cdot \rho \cdot c \cdot \Delta T \quad (26)$$

With:

- $Q_{demand,panel}$: thermal power panel [W]
- Φ_v: volume flow [m³/s]
- ρ: density (= 1000 kg/m³ for water)
- c: specific heat (= 4180 J/kgK for water)
- $\Delta T = T_{supply} - T_{return}$: temperature of flow entering the room – temperature of flow leaving the room

Ventilation ($Q_{demand,ventilation}$)

The same principle as for the panels is used for the ventilation. The difference is that air flows through the system instead of water. Thereby the supply and return temperature is not measured in the room. The leaving temperature of the system is used as the supply temperature (T_{supply}) and the air temperature as the return temperature (T_{return}). In comparison to the other independent variables, a lot of data points of the supply temperature were missing and the timestep was per 8 minutes instead of 10 minutes, therefore the choice is made to interpolate this data. In Appendix E a graph can be found to show the interpolated supply temperature.

$$Q_{demand,ventilation} = \Phi_v \cdot \rho \cdot c \cdot \Delta T \quad (27)$$

With:

- $Q_{ventilation}$: thermal power [W]
- Φ_v: volume flow [m³/s]
- ρ: density (= 1.2 kg/m³ for air)
- c: specific heat (= 1000 J/kgK for air)
- $\Delta T = T_{supply} - T_{return}$: the temperature of flow entering the room – air temperature

5.7 Indoor surface temperature ($T_{surface}$)

In the two rooms the indoor surface temperature is measured for around two weeks, the exact measurement periods are shown in Table 7. The placed sensors are shown in Figure 16. Unfortunately, not all the data points were possible to use, because the measured temperatures were very high or low and/or the sensors had fallen off the wall. The data set is analysed and the measurements that were in line with the air temperatures are selected to use in the model, as shown in Table 8 and Figure 16. Unfortunately, no indoor surface temperature is measured on the floor (with a relatively high thermal mass), because the sensors were replaced or extreme temperatures were measured, however, in the classroom the surface temperature of the concrete pillar, with a high thermal mass, is measured.

Table 8: Overview of the used indoor surface temperature measurements, where X indicates the not used measurements, ✓ indicates the used measurements and N/A indicates that this measurement does not exist

Indoor surface temperature	Office	Classroom
Window	✓	X
Floor	X	X
Ceiling	✓	✓
System walls (2/3)	✓	✓
Concrete pillar	X	✓
Glass wall	✓	N/A
Inside of the outer wall	✓	✓

5.8 Indoor air temperature ($T_{indoorair}$)

The air temperature is measured in the room with the existing sensors. During the measurement of the indoor surface temperature, the indoor air temperature is also measured with an extra sensor. As the indoor surface temperatures are not measured by the existing sensors, it is not possible to validate the operation system and the measurements with each other. Therefore the indoor air temperature measurements in the office are used as validation. The scatterplot in Figure 23 shows a strong correlation between the measured air temperatures. However, in Figure 24, the air temperature measured by the placed sensors is most of the time a little bit higher than the measurement by the already existing sensors. This existing sensor is placed on a system wall, so it could be influenced by the temperatures of the wall surfaces. The extra sensor is placed in the middle of the room. However, based on these graphs the conclusion is made that the own sensor and the existing sensors are measuring the

air temperature in the same magnitude with a difference which could be explained by the different places, so the data from the measurements and the existing sensors can be used. In the model, the air temperature measured by the already existing sensors is used as the heating and cooling demand is driven by the difference between the temperature setpoint and the measured air temperature.
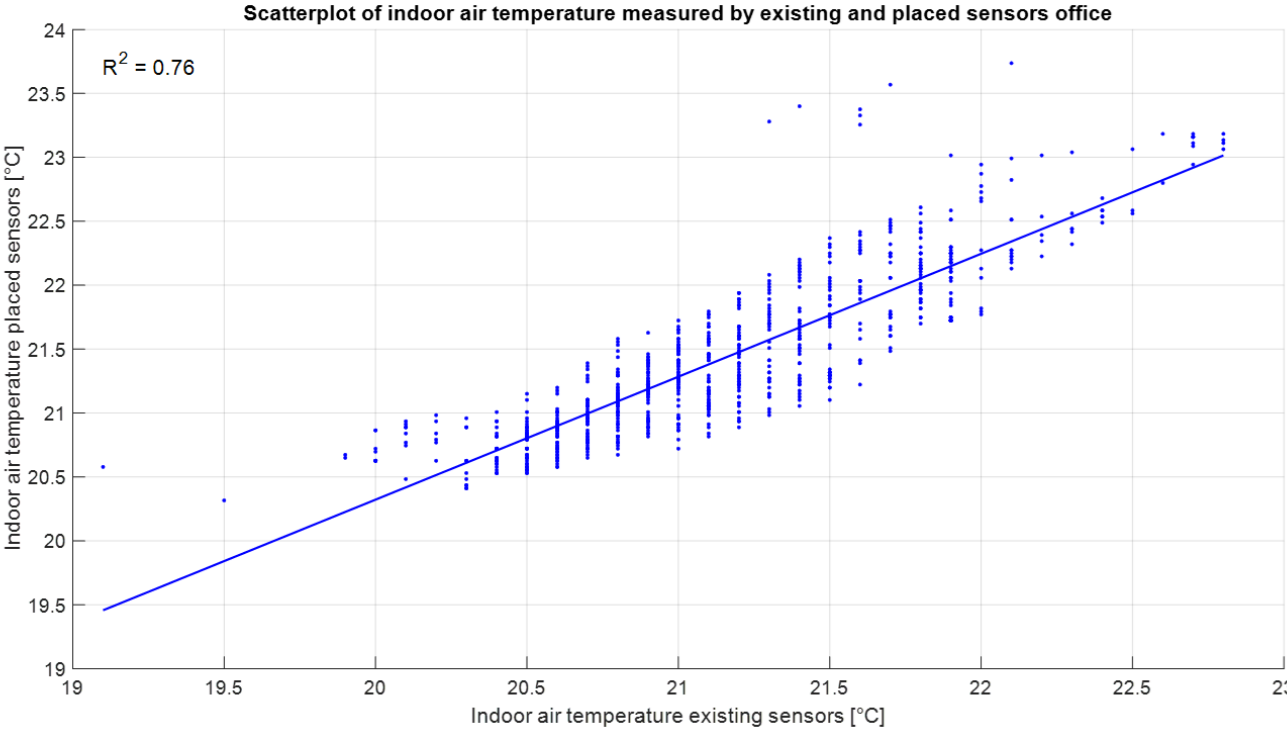
Figure 21: comparison of the measured indoor air temperature by existing and placed sensors office, the polyfit function in MATLAB is used to calculate a linear regression line
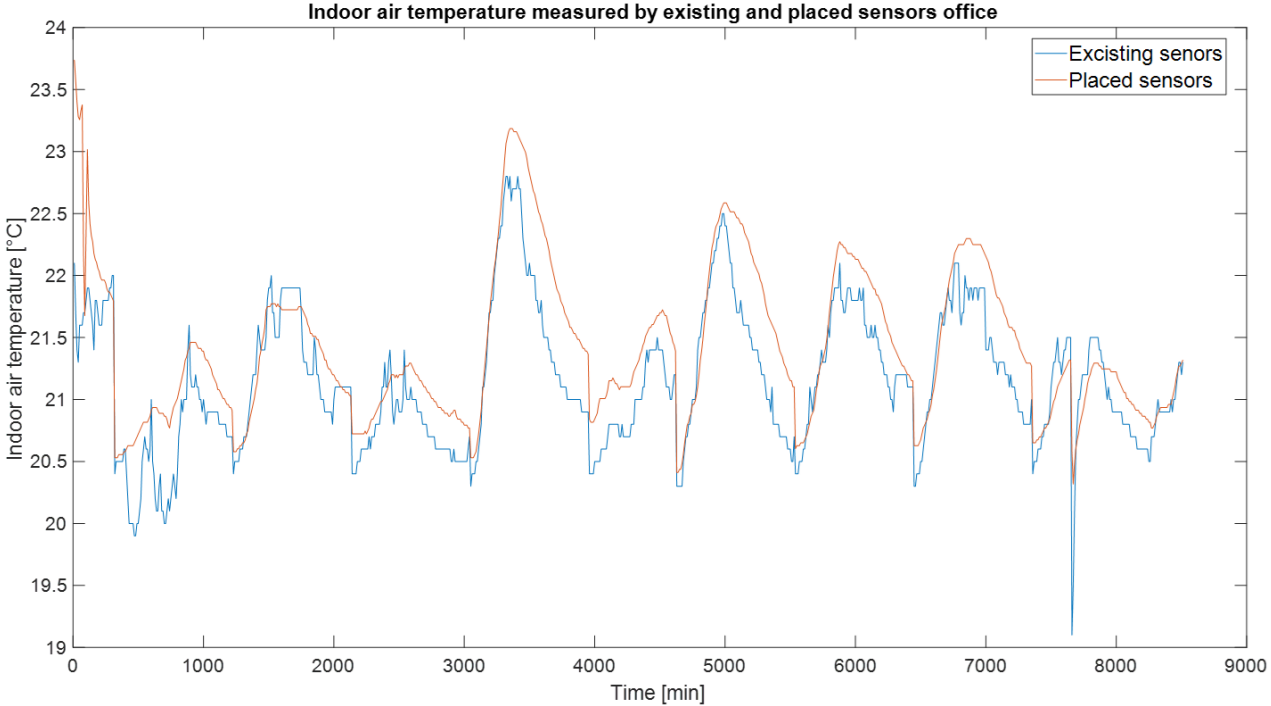
Figure 22: graph showing the difference over time between the measured indoor air temperature of the existing and placed sensors

6. Data analysis

In this Chapter, the gathered data in Chapter 5 is analysed to get an understanding of the correlations between the chosen parameters for the development of the model to model and predict the heating and cooling demand.

6.1 Heating and cooling demand

First, the calculated heating and cooling demand in Chapter 5.5 is analysed. The heat and cold delivered by the three systems are shown in Table 9 and Figure 25. In the office, there is a much higher heating demand than cooling demand, which is reversed for the classroom. This could be explained by the lower temperature set point of the classroom of 21 °C, and 21.5 °C for the office. Also, the internal heat gains are expected to be higher for the classroom than for the office, because of the fact the maximum number of people is higher in the classroom, with less space per person.

The heat and cold in the office are mostly supplied by the floor system, followed by the ventilation and the panel of the office. In the classroom, the biggest supplier of heat and cold is also the floor system, followed by the panels and the ventilation.

Table 9: total heating and cooling demand, delivered by the floor, panel and ventilation system for both rooms during opening hours

	Office (3/8/2024 – 3/22/2024)				Classroom (3/22/2024 – 4/11/2024)			
	Heat [kWh]	*Share [%]*	Cold [kWh]	*Share [%]*	Heat [kWh]	*Share [%]*	Cold [kWh]	*Share [%]*
Floor	54.72	*0.99*	12.48	*57.34*	6.17	*91.20*	34.61	*52.66*
Panels	$2.00e^{-3}$	*$3.05e^{-3}$*	3.75	*17.23*	0.43	*6.36*	15.29	*23.26*
Ventilation	$3.00e^{-3}$	*$6.10e^{-3}$*	5.54	*25.42*	0.167	*2.44*	12.50	*19.01*
Total	54.73	*100%*	21.77	*100%*	6.77	*100*	62.50	*100*
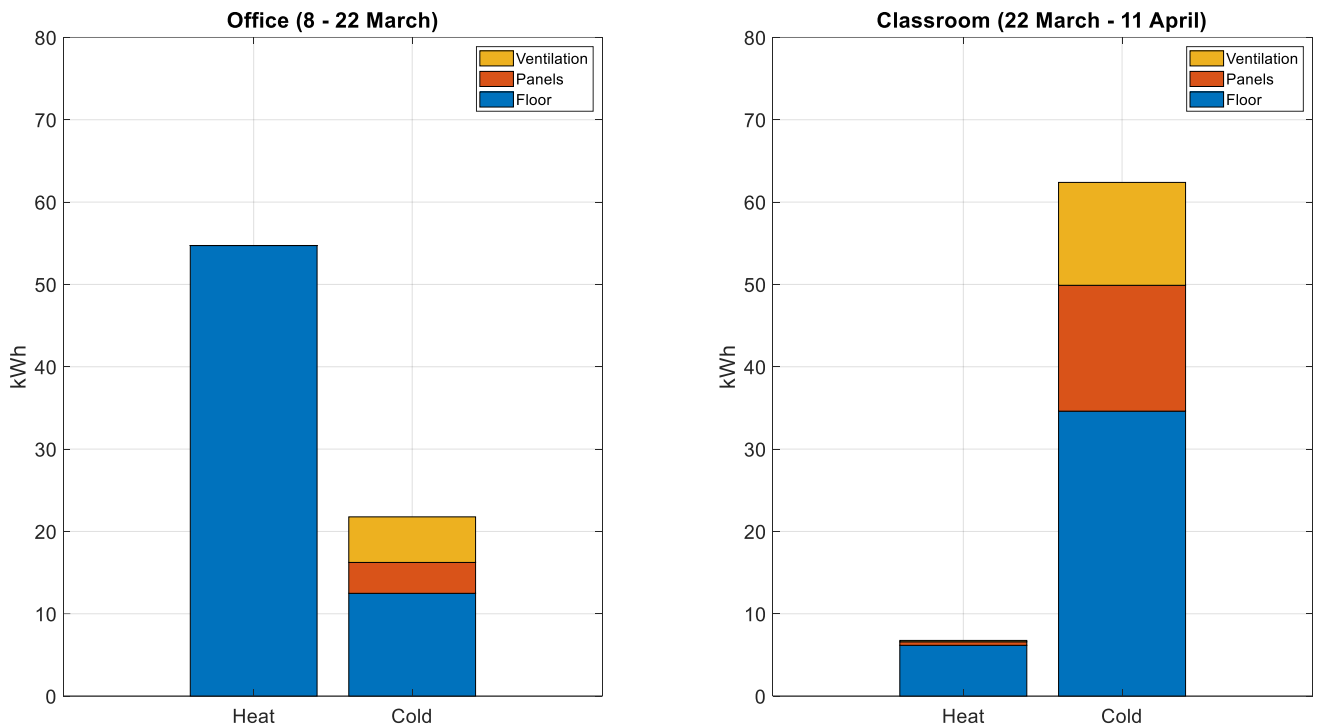
Figure 23 the total energy demand for each room delivered by the floor, panel and ventilation system

To understand how the installations are used to supply the demand of heat and cold, Figure 26 – 31 is made to show the thermal supply of each installation during opening hours.

In the office room, the floor systems supply heat or cold during longer periods of hours, during the midday. The panel follows a different pattern, with three longer periods of cold. The first period corresponds for the first couple of hours to a cold supply by the floor system. Then the floor system is switched off, and the room is further cold by the panels. After that, the floor starts to heat. The other two periods of cooling by the panels correspond to periods of cooling by the floor system, the panels are used to supply extra cold, to achieve the desired air room temperature. The supply curve of the ventilation systems fluctuates more, this could be explained by that this system could be used for shorter periods of demand, where the floor heating and cooling system and the panels cannot correspond so quickly as the ventilation to the demand. Besides that the ventilation system is mostly used to get a higher air quality, after that the system can be used to heat or cold. In Appendix E graphs can be found of the supply and return temperatures of the installations.
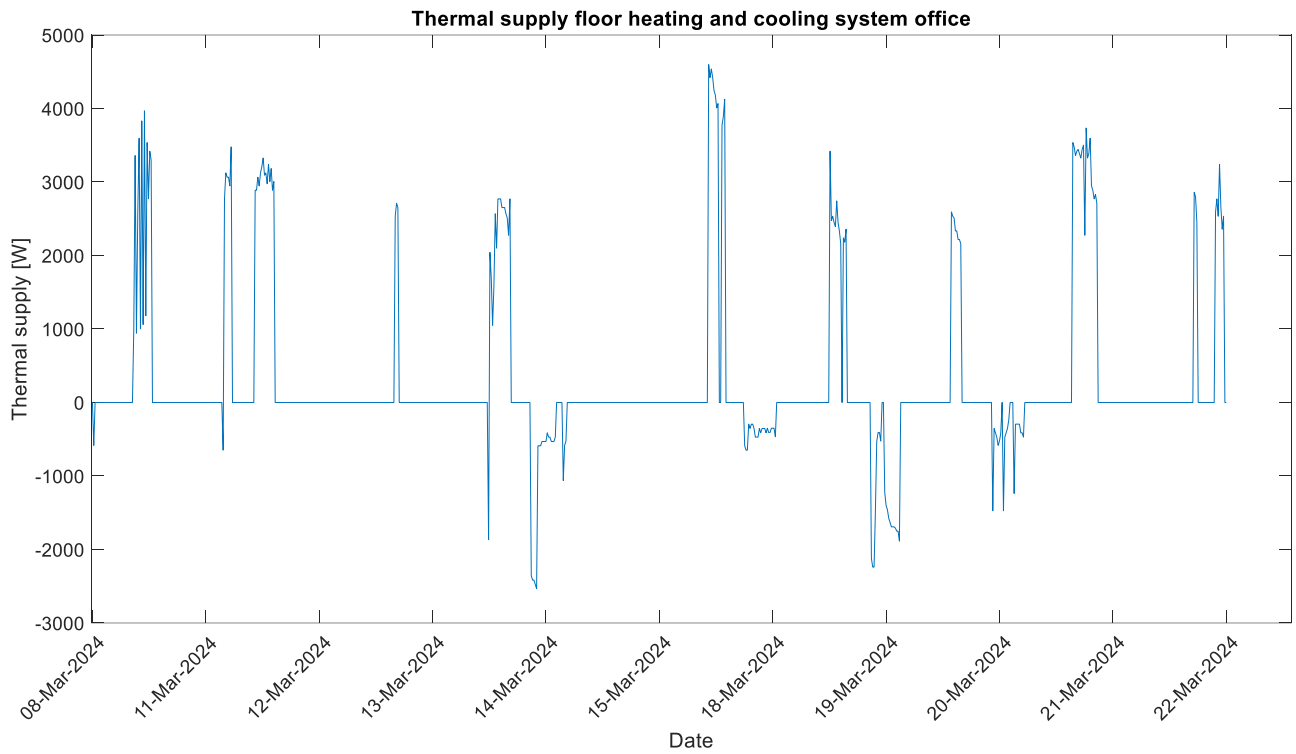
Figure 24: thermal energy supply by the floor heating and cooling system in the office
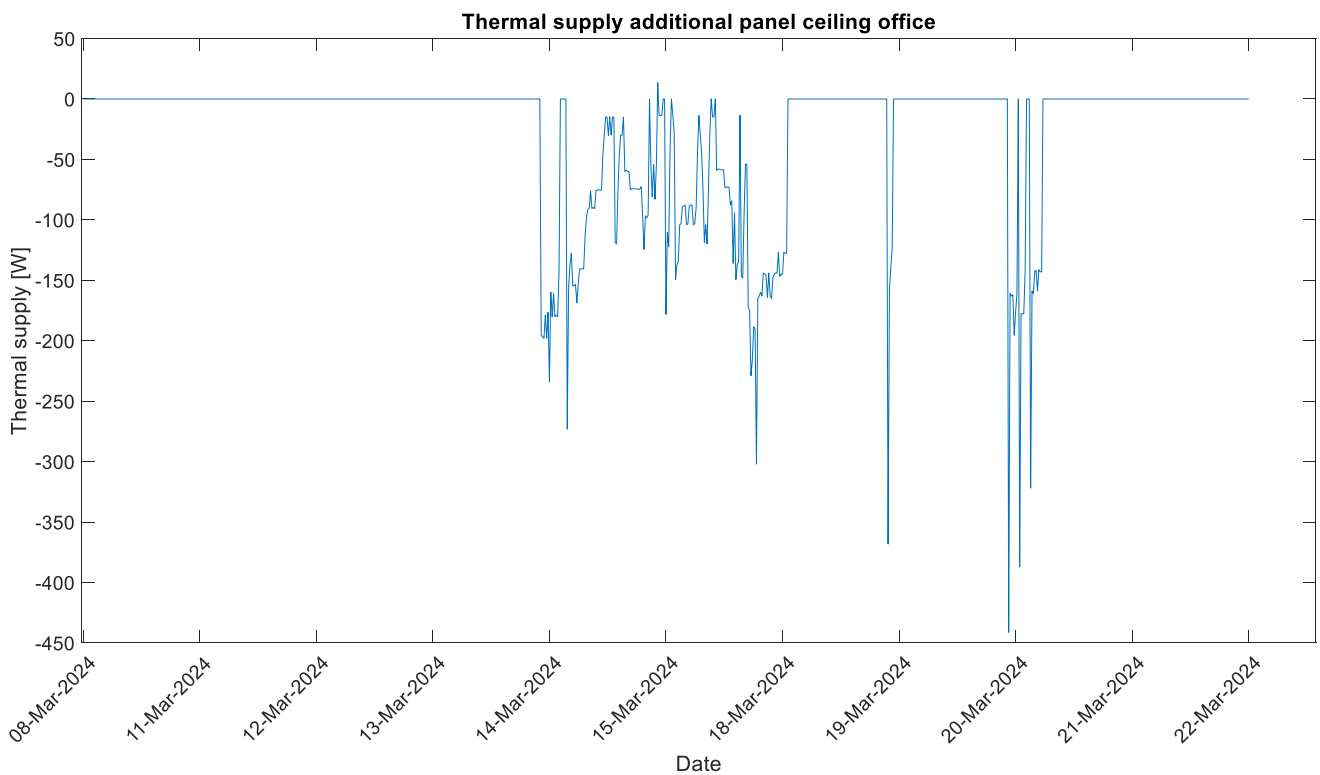
Figure 25: thermal energy supply by the additional panels in the office
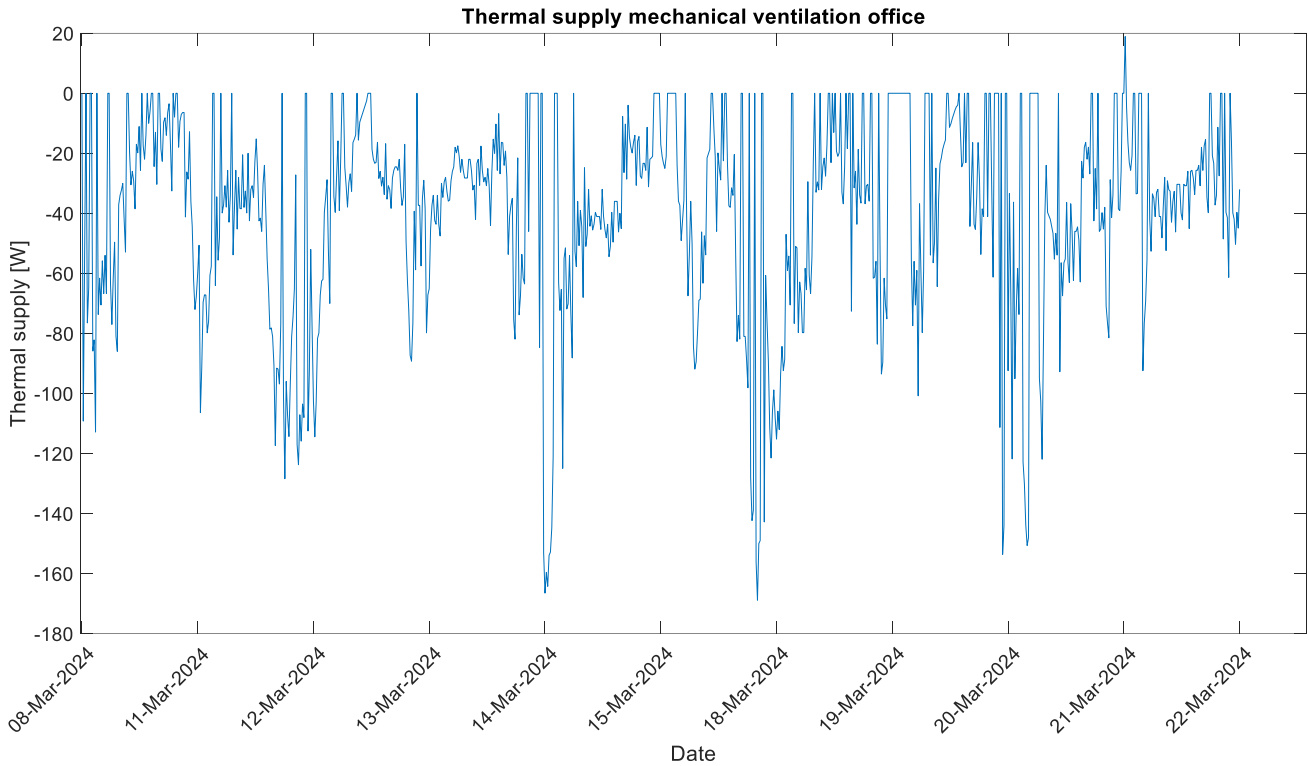
Figure 26: thermal energy supply by the ventilation in the office

In the classroom, the pattern of the floor system and panels are quite the same, there are periods for a couple of hours of heating or cooling or none of them. During the peaks and valleys of the heating and cooling demand supplied by the floor system and the panels the heat or cold supplied by the ventilation system increases. During periods of no heating or cooling, there is still heat or cold supplied by the ventilation system. As explained in Chapter 4.3, the ventilation system is primarily used for air purification, whereby heating or cooling is possible. In Appendix E graphs can be found of the supply and return temperatures of the installations.
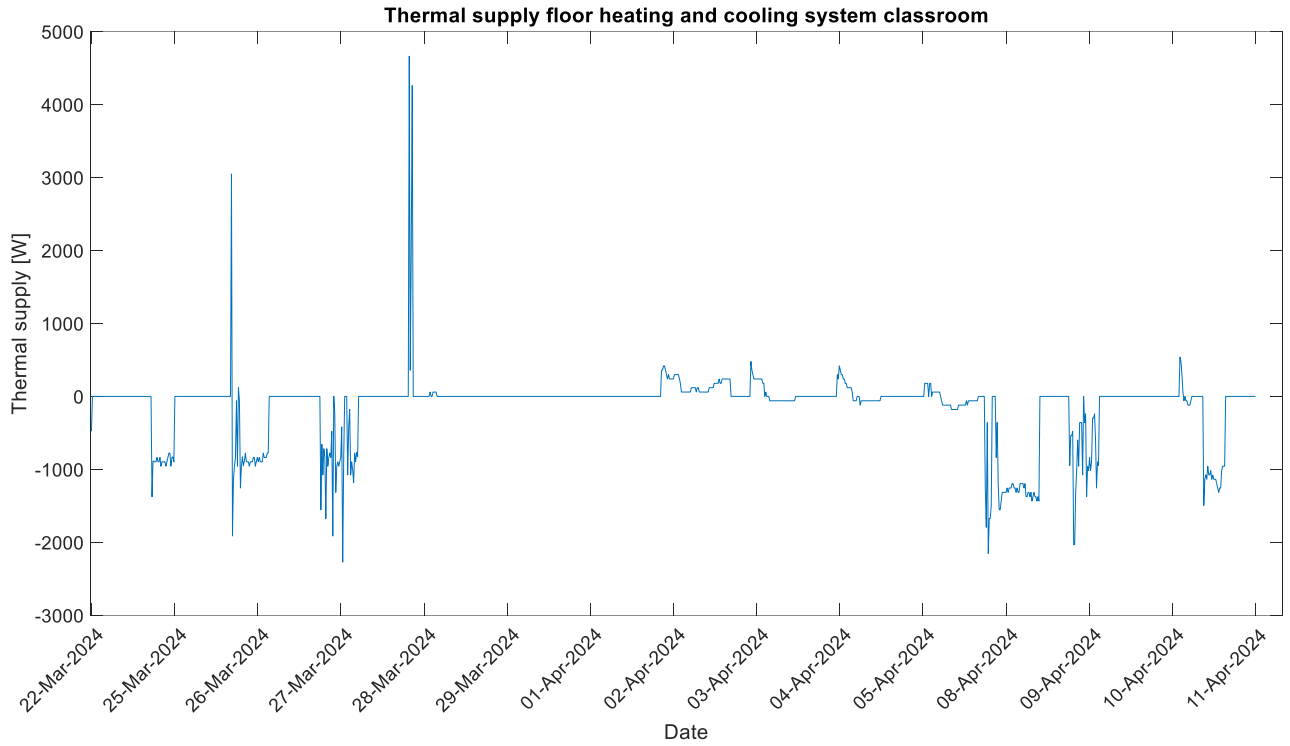
Figure 27: thermal energy supply by the floor in the classroom
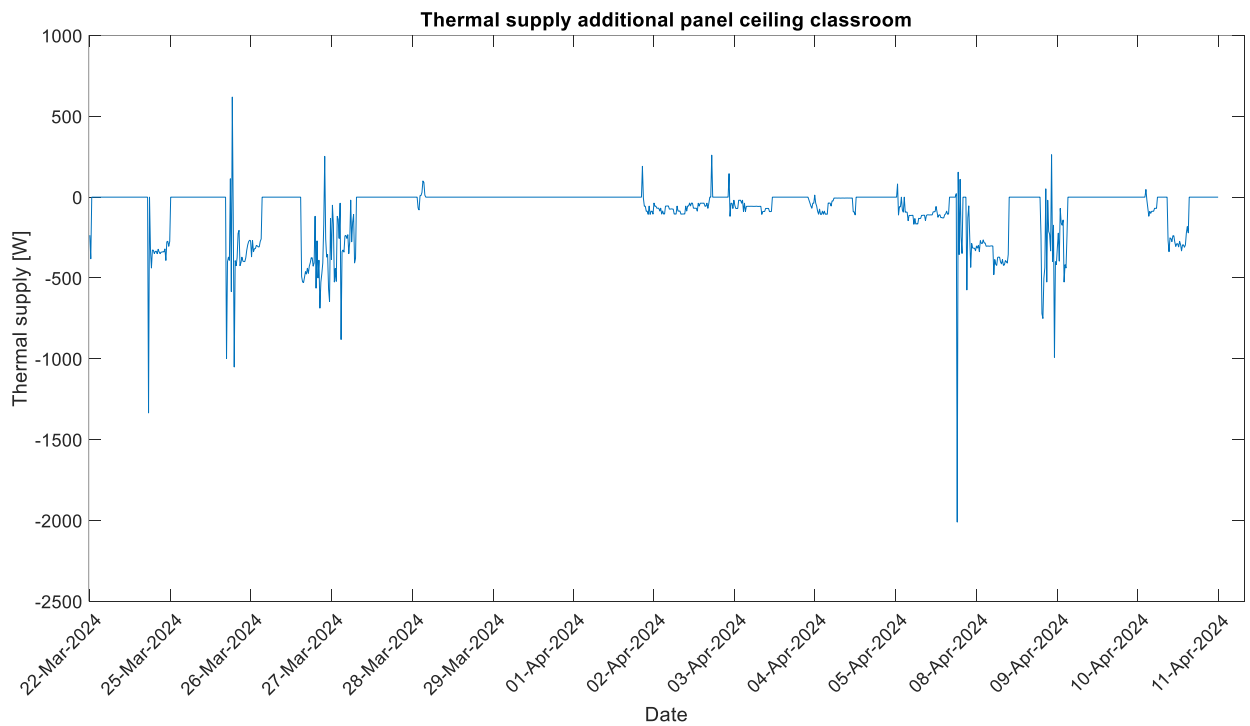
Figure 28: thermal energy supply by the additional panels in the classroom
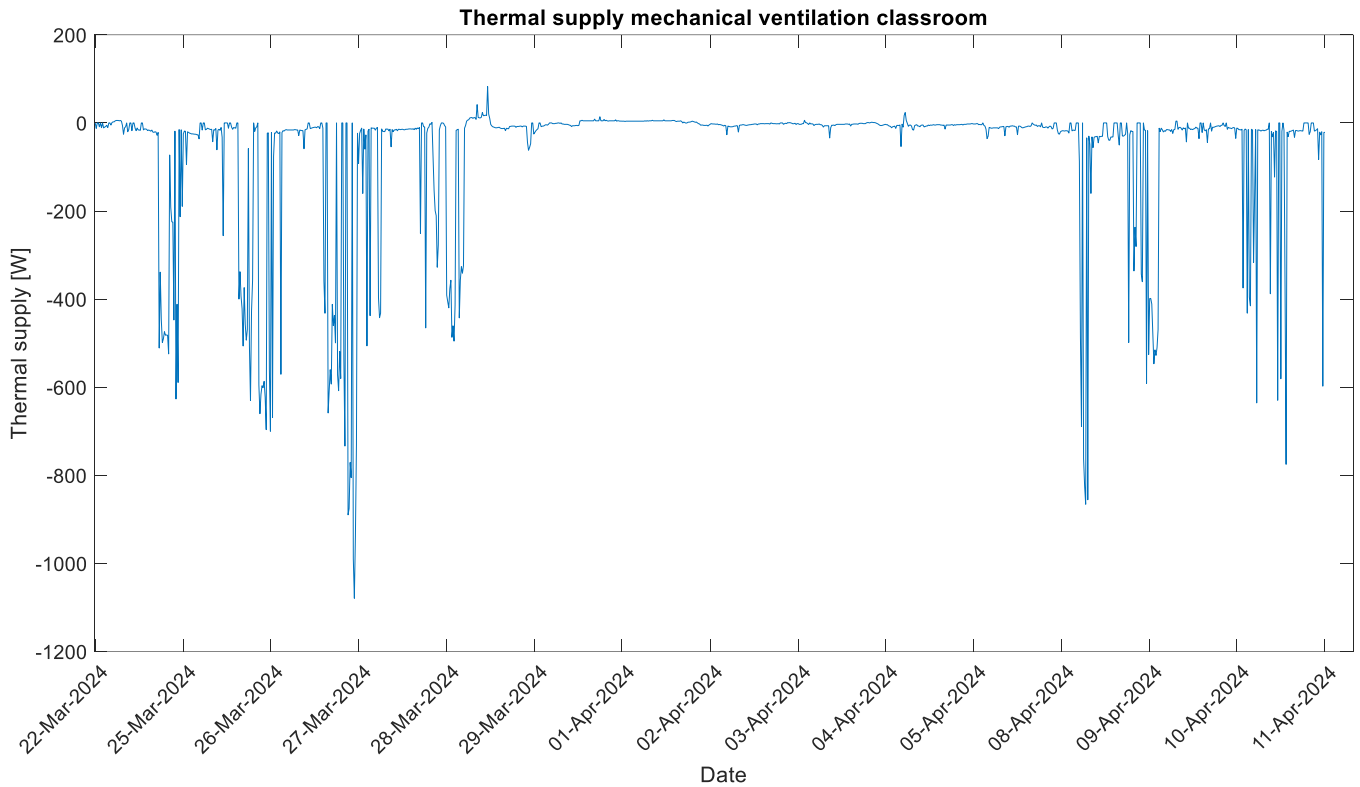
Figure 29: thermal energy supply by the ventilation in the classroom

In Figures 32 – 33 the heating or cooling demand is shown per hour of the day for one week for each room, which gives insight into the distribution of the heating or cooling demand over the day. In contrast to the other analysis, the thermal demand during closing and opening hours is included. In the office the heating demand is much higher than the cooling demand, this can also been seen in Figure 25. There is almost a constant supply of cold, which is from the ventilation, there is a peak during the Midday of Thursday. Heating starts during the early hours of the day, but remarkable is the heating demand on Monday and Wednesday evening. In the classroom there is more cooling demand than heating demand, which is also shown in Figure 25. The peaks of cooling demand occurs during opening hours on Monday until Wednesday. On Thursday there is a peak of heating. Also in the office there is almost a constant supply of cold, which is supplied by the ventilation system.
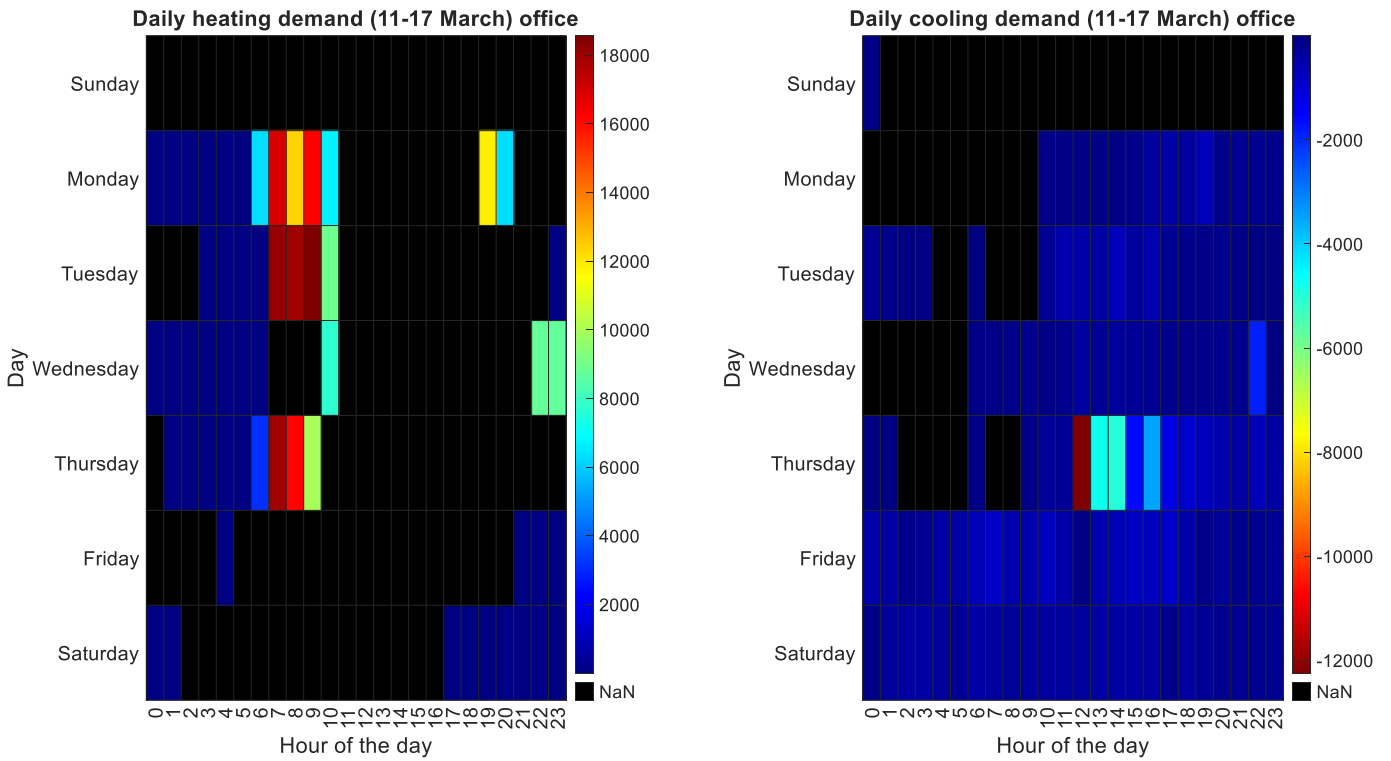
Figure 30: representation of the daily pattern on the heating demand, NaN = 0 W
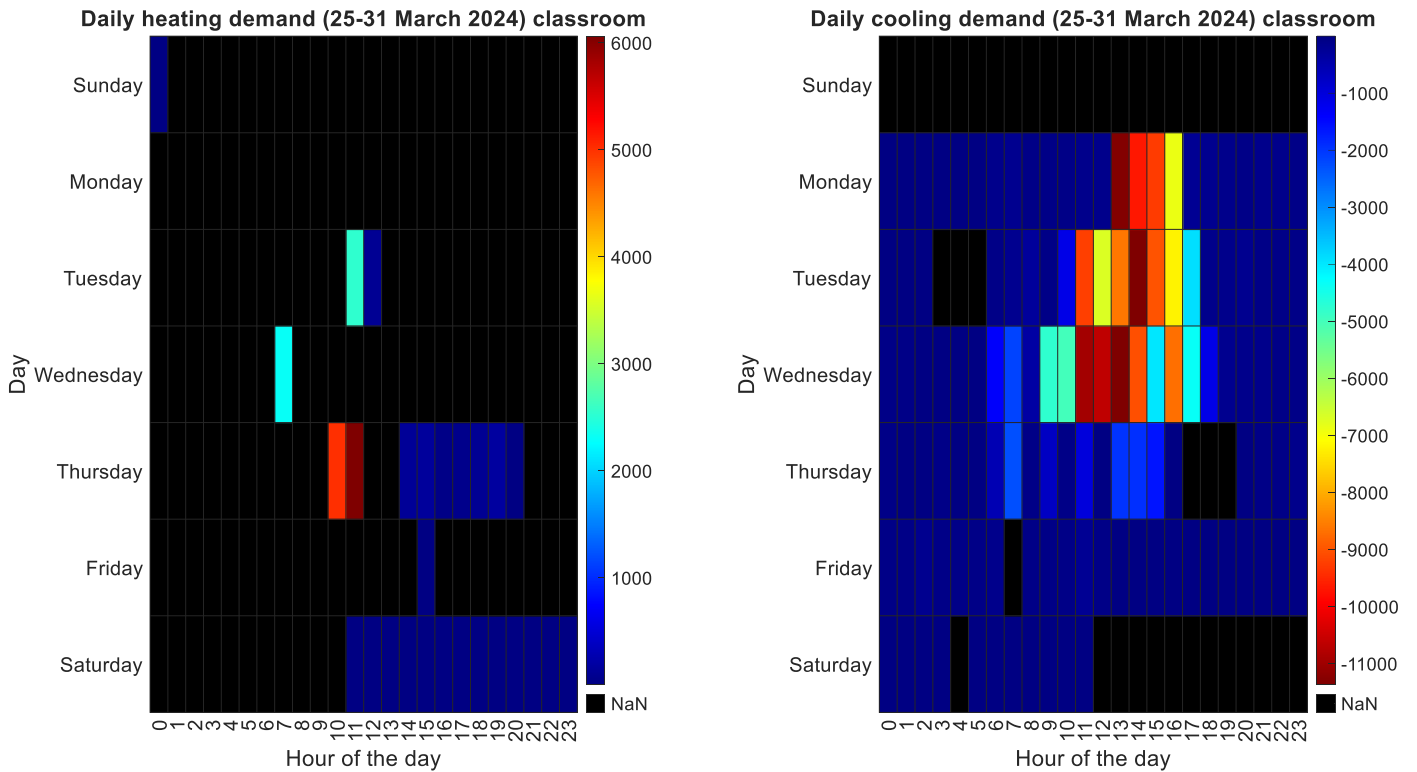
Figure 31: representation of the daily pattern on the cooling demand, NaN = 0 W

6.2 Correlation matrix parameters

After a short analysis of the correlation matrix, the correlation between the independent variables and the heating or cooling demand is further explained per variable.

6.2.1 Heating demand office

In Figure 34 the correlation matrix is shown for the heating demand of the office. Except the light intensity east and the internal heat gains, all the independent variables have a negative correlation with the heating. The absolute value of the coefficients is between 0.13 (solar light intensity east) and 0.39 (surface temperature system wall 1). Most of the correlations are weak to moderate, will results in an expected lower accuracy of the model.

Correlation matrix heating demand office

	Heating Demand	Ts Window	Ts System wall 1	Ts Inside of the outer wall	Ts Ceiling	Ts Glass wall	Ts System wall 2	Indoor air Temperature	Outdoor temperature	Wind Speed	Solar light Intensity West	Solar light Intensity East	Solar light Intensity South	Internal Heat Gains
Heating Demand		-0.25	-0.39	-0.17	-0.32	-0.37	-0.28	-0.34	-0.29	-0.19	-0.17	0.13	-0.11	0.21
Ts Window			0.75	0.31	0.66	0.71	0.62	0.75	0.69	0.20	0.49	0.53	0.83	0.20
Ts System wall 1				0.40	0.89	0.98	0.76	0.93	0.54	0.19	0.62	0.13	0.61	0.01
Ts Inside of the outer wall					0.31	0.36	0.26	0.47	0.36	-0.01	0.17	-0.03	0.12	-0.12
Ts Ceiling						0.91	0.79	0.85	0.43	0.24	0.55	0.20	0.60	0.03
Ts Glass wall							0.77	0.91	0.46	0.20	0.62	0.12	0.61	0.03
Ts System wall 2								0.70	0.33	0.08	0.53	0.28	0.54	0.02
Indoor air Temperature									0.52	0.16	0.53	0.18	0.60	0.06
Outdoor temperature										0.27	0.40	0.26	0.45	-0.04
Wind Speed											0.25	0.17	0.22	-0.09
Solar light Intensity West												0.22	0.55	0.00
Solar light Intensity East													0.57	0.32
Solar light Intensity South														0.24
Internal Heat Gains														

Figure 32: correlation matrix of the heating demand in the office

6.2.2 Heating demand classroom

In Figure 35 the correlation matrix is shown for the heating of the classroom. In contrast to the heating demand in the office, the direction of the correlations for the classroom are distributed in positive and negative. The absolute value of the coefficients is between 0.00 (light intensity south) and 0.11 (surface temperature system wall 1). The heating demand as shown in Figure 25 is much lower in the classroom than in the office. All these correlations in the classroom are weak, and a low accuracy of the heating model for the classroom is expected.

71

Figure 33: correlation matrix of the heating demand in the classroom

6.2.3 Cooling demand office

In Figure 36 the correlation matrix is shown for the cooling demand of the office. Except the return temperature of the floor and panel, all the independent variables have a positive correlation with the heating. The range of the absolute value of the correlation coefficient is between 0.08 (internal heat gains and solar light intensity east) and 0.53 (indoor air temperature and surface temperature of system wall 1). In contrast to the heating demand, there are more variables with a strong relation (surface temperatures of system wall 1 and glass wall and the indoor air temperature). The correlations between the independent variables and the cooling demand are stronger than the heating demand in the office, and a high accuracy is expected.

Figure 34: correlation matrix of the cooling demand in the office

6.2.4 Cooling demand classroom

In Figure 37 the correlation matrix is shown for the cooling heating of the classroom, with a maximum of 0.50 and a minimum of 0.01. The strongest correlation is between the surface temperature system wall 3 and the cooling demand. Most of the correlations have a positive correlation, except the wind speed and the surface temperature of system wall 1. There are 5 moderate and strong correlations, which give the expectation that the accuracy of the cooling model will be higher than the accuracy of the heating demand of the classroom.

Correlation matrix cooling demand classroom

	Cooling Demand	Ts Inside of the outer wall	Ts System wall 1	Ts System wall 2	Ts Concrete pillar	Ts System wall 3	Ts Ceiling	Indoor air temperature	Outdoor temperature	Wind Speed	Solar light Intensity West	Solar light Intensity East	Solar light Intensity South	Internal Heat Gains
Cooling Demand		0.49	-0.01	0.41	0.44	0.50	0.29	0.29	0.21	-0.09	0.26	0.05	0.16	0.49
Ts Inside of the outer wall			-0.05	0.97	0.95	0.97	0.91	0.85	0.35	-0.07	0.16	0.02	0.13	0.25
Ts System wall 1				-0.09	-0.11	-0.01	0.00	-0.11	-0.02	0.09	0.11	-0.09	-0.00	0.03
Ts System wall 2					0.96	0.97	0.95	0.88	0.43	0.02	0.16	0.04	0.16	0.21
Ts Concrete pillar						0.93	0.92	0.84	0.43	-0.03	0.11	-0.01	0.06	0.17
Ts System wall 3							0.94	0.82	0.40	0.04	0.21	0.05	0.20	0.29
Ts Ceiling								0.84	0.40	0.02	0.11	0.04	0.14	0.11
Indoor air temperature									0.38	-0.00	0.09	0.03	0.13	0.08
Outdoor temperature										0.04	0.42	-0.14	0.17	-0.12
Wind Speed											0.22	-0.04	0.18	0.09
Solar light Intensity West												0.09	0.49	0.14
Solar light Intensity East													0.50	0.06
Solar light Intensity South														0.16
Internal Heat Gains														

Figure 35: correlation matrix of the cooling demand in the classroom

6.3 Correlation of heating and cooling demand and independent variables

Based on the correlation matrixes in Figure 34 – 37, an analysis is made per independent variable.

To understand how the heating and cooling demand is disturbed, and why there are periods of zero thermal demand, Table 10 and 11 are made to understand if indoor or outdoor air temperature ranges could be determined for the three different nodes; heating, cooling or no thermal demand, to predict when there is a heating or cooling demand

6.3.1 Outdoor temperature ($T_{outdoor}$)

An overview of the outdoor temperature ranges, for the three thermal demands is shown in Table 10, this is visualized in Appendix G. The average outdoor temperature has increased over the measurement periods, which results in a higher average outdoor temperature of 12.72 °C (for the office) and 12.90 °C (for the classroom). The outdoor temperature ranges for the three different thermal demands are broad, except for the heating demand. When the outdoor temperature is above 14.33 °C for the office and 15.33 °C for the classroom, there is no heating, only cooling or no cooling. As for both rooms, the outdoor temperature ranges of the periods of no heating or cooling and cooling are broad, it is not possible to predict when there are heat or cold periods.

Table 10: outdoor temperature ranges per thermal energy demand

	Outdoor temperature (°C)					
	Office			Classroom		
Thermal demand	Average	Min.	Max.	Average	Min.	Max.
Heating: > 0 W	10.74	8.40	14.33	12.31	8.90	15.33
Cooling: < 0 W	13.07	8.40	19.03	13.12	7.30	21.05
No heating or cooling: = 0 W	12.78	8.40	19.15	12.45	7.55	20.60

The Pearson correlation coefficients in Table 11 show that there is a negative correlation between outdoor temperature and heating demand and a positive correlation between outdoor temperature and cooling demand. This means that as the outdoor temperature increases, the heating demand decreases and the cooling demand increases. The correlation with the cooling demand is stronger than the correlation with the heating demand. Furthermore, the correlation is stronger for both the heating and cooling demand for the office compared to the classroom (refer to Figures 34 – 37). This difference might be due to the office's location next to the outside of the building, while the classroom is next to the courtyard. Additionally, the windows in the office were opened more frequently and for longer periods during the measurement period compared to those in the classroom. This difference in window usage could explain the higher correlation with the heating or cooling demand, as there is no difference in insulation between the two rooms.

Table 11: Pearson correlation between thermal demand and the outdoor temperature

Independent variable: Outdoor temperature	Pearson-correlation coefficient [0-1]	Strength	Direction
Heating office	-0.29	Weak	Negative
Heating classroom	-0.09	Weak	Negative
Cooling office	0.42	Moderate	Positive
Cooling classroom	0.21	Weak	Positive

6.3.2 Indoor air temperature ($T_{indoorair}$)

The average indoor air temperature during opening hours is in the office 21.15 °C, for the classroom this is 21.47 °C. In Table 12 the air temperature ranges are shown for thermal demands, in Appendix F a graph of these ranges can be found.

Table 12: indoor air temperature ranges per thermal energy demand type

| | Indoor air temperature (°C) | | | | | |
| | Office | | | Classroom | | |
Thermal demand	Avg.	Min.	Max.	Avg.	Min.	Max.
Heating: > 0 W	20.63	19.50	21.50	20.96	17.50	22.50
Cooling: < 0 W	21.22	19.90	22.80	21.63	19.70	23.00
No heating or cooling: = 0 W	21.27	19.10	22.40	21.32	20.20	22.20

For the office room, there is a heating demand between 19.5 °C and 21.5 °C. The range of the air temperature is higher for the cooling demand with an air temperature between 19.9 °C and 22.8 °C. When the air temperature is higher than 21.5 °C, the systems in the room start cooling instead of heating. The air temperature range for periods of no heating or cooling is broadest with 19.1 °C and 22.4 °C. Between an indoor air temperature of 22.4 °C and 22.8 °C there are only periods of cooling. However, as shown in Table 12, the ranges of the indoor air temperature for the three different nodes; heating, cooling or no thermal demand are overlapping.

The indoor air temperature range is broader for the classroom. The ranges for cooling and no heating or cooling are almost the same. The range for the heating demand is smaller for the classroom, where there is only heating from an air temperature of 17.50 °C to 19.70 °C when the air temperature is higher than 19.70 °C the room can be heated, cooled or none of them until an air temperature of 22.50 °C, when a higher air temperature is measured there are only periods of cooling.

In Table 13 the Pearson-correlation coefficients for each thermal demand and per room are shown to get insight what the influence of the air temperature on the heating or cooling demand. For both heating demands the direction is negative, in other words, when the air temperature increases the heating demand decreases, which is logical, as the air needs to be less heated. The inverse is the case for the cooling demand, an increasing air temperature, results in a higher cooling demand.

Table 13: Pearson correlation between thermal demand and the indoor air temperature

Independent variable: Indoor air temperature	Pearson-correlation coefficient [0-1]	Strength	Direction
Heating office	-0.34	Moderate	Negative
Heating classroom	-0.06	Weak	Negative
Cooling office	0.53	Strong	Positive
Cooling classroom	0.29	Weak	Positive

Comparing the Pearson correlation coefficients and the scatterplots in Appendix H, the correlation between the air temperature and the thermal demand in the office is stronger than in the classroom. It could be that in the classroom other independent variables have a stronger correlation with the thermal demand, for example, the occupancy rate (which is higher and shorter for the classroom in comparison to the office).

6.3.4 Wind speed (V_{wind})

The correlations between the windspeed and the heating or cooling demand are weak, which means that there is not much infiltration (see Chapter 2.2.1). The correlations in the classroom are negligible. For the office, the correlations are stronger with the wind speed. As earlier explained, the office is adjacent to the outer space, whereas the classroom is adjacent to the courtyard, whereby the assumption can be made that the office is facing more wind. What is remarkable is the direction of the correlations within the office. The correlation between the heating demand and the windspeed is negative and the inverse for the cooling demand. This could be explained that the influence of the windspeed is also determined by the temperature of the outdoor air.

Table 14: Pearson correlation between thermal demand and the windspeed

Independent variable: Wind speed	Pearson-correlation coefficient [0-1]	Strength	Direction
Heating office	-0.19	Weak	Negative
Heating classroom	0.01	Weak	Positive
Cooling office	0.12	Weak	Positive
Cooling classroom	-0.09	Weak	Negative

6.3.5 Solar light intensity ($\sim Q_{solar}$)

In Chapter 5.2, the data is collected from all available solar light intensity directions: south, east, and west. It is expected that the solar light intensity with the same orientation as the concerned room has the strongest correlation with the heat or cold demand. For the office, the strongest correlation with the cooling demand is with the solar light from the south, followed by the west and east. However, for the heating demand, the significant order is west, east, and south, although the difference between these orientations is not significant. The room is oriented to the southeast, however, from Table 15 and Appendix H, it can be concluded that the solar intensity from the West has a significant impact of the same order as the other orientations, possibly due to the glass wall adjacent to the corridor. Therefore, all three orientations will be included in the model.

The classroom is oriented to the northeast, but unfortunately, no data is available for the north orientation. The correlations between the orientations and the heating and cooling demand for the classroom are less strong compared to the office. This could be explained by the windows in the classroom being oriented to the courtyard, whereas the windows of the office are oriented to the open space around the building. Remarkable is that the correlation between the solar light from the east has a positive relation, however, it could be that the sun blinds are used for example.

Table 15: Pearson correlation between thermal demand and the light intensity south

Independent variable: Solar light intensity south	Pearson-correlation coefficient [0-1]	Strength	Direction
Heating office	-0.11	Weak	Negative
Heating classroom	0.00	Weak	Positive
Cooling office	0.44	Moderate	Positive
Cooling classroom	0.16	Weak	Positive
Solar light intensity east			
Heating office	0.13	Weak	Positive
Heating classroom	0.03	Weak	Positive
Cooling office	0.08	Weak	Positive
Cooling classroom	0.05	Weak	Positive
Solar light intensity west			
Heating office	-0.17	Weak	Negative
Heating classroom	-0.04	Weak	Negative
Cooling office	0.35	Moderate	Positive
Cooling classroom	0.26	Weak	Positive

6.3.6 Internal heat gains ($Q_{internal}$)

The Pearson-correlation coefficient of the thermal demand and the internal heat gains in the office and the class room have all a positive correlation (see Table 16 and Appendix H). There is a moderate correlation between cooling in the classroom and the internal heat gains, this could be explained by the fact that the occupancy is higher in the classroom than in the office. All the correlations are positive, what is remarkable for the heating demand, in other words when the internal heat gains increases, the heating demand also increases. For heating, for both rooms the correlations are weak. These correlations can be explained by the occupant regulation of the heating and cooling demand.

Table 16: Pearson correlation between thermal demand and the internal heat gains

Independent variable: Internal heat gains	Pearson-correlation coefficient [0-1]	Strength	Direction
Heating office	0.21	Weak	Positive
Heating classroom	0.06	Weak	Positive
Cooling office	0.08	Weak	Positive
Cooling classroom	0.49	Moderate	Positive

6.3.7 Indoor surface temperatures ($T_{surface}$)

In Chapter 5.7, the measured surface temperatures for each room are discussed. The Pearson correlation coefficients for all the indoor surface temperatures can be found in Table 17 and Appendix H. These temperatures are categorized by surface within a room, such as system wall 1 being different for each room.

The cooling demand in the classroom shows multiple moderate and strong correlations with the indoor surface temperatures of system wall 2, the inside of the outer wall, the window, system wall 3, and the concrete pillar. In contrast, the heating demand in the classroom shows only weak correlations.

The heating demand in the office has multiple moderate correlations, unlike the classroom. These include system wall 1, the ceiling, and the glass wall. The cooling demand in the office has a strong correlation with the surface temperatures of system wall 1 and the glass wall, and a moderate correlation with the surface temperatures of the ceiling and the window.

Despite the surface temperature of system wall 1 in the classroom, all the correlations between the surface temperatures and the heating demand are negative. Conversely, the correlations with the cooling demand are positive for each room, except for system wall 1 in the office. The temperature-damping effect, as explained in Chapter 2.2.1, is evident in this data. This effect has a stronger impact on the correlation with the cooling demand, as almost all the Pearson correlation coefficients for cooling are higher than those for heating. An exception is system wall 1 and 2 in the classroom; however, for system wall 2, the difference between the coefficients of heating and cooling is small. System wall 1 has a digital smart board that emits heat, and behind system wall 1, there is a small room for a student association with a higher occupancy rate.

The cooling demand in the office has a moderate correlation with the solar light intensity from the south and west. This heat gain is captured in the thermal mass of the room. System wall 2 in the office (see Appendix H) has a moderate correlation with the cooling demand, in contrast

to the weak correlation with system wall 1. This difference is due to the room's orientation (southeast), where system wall 2 receives more sunlight. Consequently, the temperature of system wall 2 is somewhat higher than that of system wall 1 (see Figure in Appendix H).

Another notable correlation is between the heating (moderate) and cooling demand (strong) and the glass wall within the office. As the U-value is much lower for the glass than for the walls (see Chapter 4.2), there is a strong correlation with the variables of the corridor on the other side of the glass wall.

Overall, most indoor surface temperatures show a moderate or strong correlation with the heating or cooling demand.

Table 17: Pearson correlation between thermal demand and the indoor surface temperatures

Independent variable: T_s system wall 1	Pearson-correlation coefficient [0-1]	Strength	Direction
Heating office	-0.39	Moderate	Negative
Heating classroom	0.11	Weak	Positive
Cooling office	0.53	Strong	Positive
Cooling classroom	-0.01	Weak	Negative
T_s system wall 2			
Heating office	-0.28	Weak	Negative
Heating classroom	-0.05	Weak	Negative
Cooling office	0.30	Moderate	Positive
Cooling classroom	0.41	Moderate	Positive
T_s ceiling			
Heating office	-0.32	Moderate	Negative
Heating classroom	-0.01	Weak	Negative
Cooling office	0.45	Moderate	Positive
Cooling classroom	0.29	Weak	Positive
T_s inside of outer wall			
Heating office	-0.17	Weak	Negative
Heating classroom	-0.04	Weak	Negative
Cooling office	0.25	Weak	Positive
Cooling classroom	0.49	Moderate	Positive

Tₛ glass wall

Heating office	-0.37	Moderate	Negative
Cooling office	0.52	Strong	Positive

Tₛ window

Heating office	-0.25	Weak	Negative
Cooling office	0.50	Strong	Positive

Tₛ system wall 3

Heating classroom	-0.02	Weak	Negative
Cooling classroom	0.50	Strong	Positive

Tₛ concrete pillar

Heating classroom	-0.07	Weak	Negative
Cooling classroom	0.44	Moderate	Positive

6.4 Correlation between independent variables

Besides the correlations between the independent variables and the heating or cooling demand, the correlation matrixes (Figures 34 – 37) shows also strong correlations between independent variables. There is an option in MATLAB to include the possibility of including the correlation between two independent variables as an independent variable. For this correlation is a coefficient calculated, in the same way as for a single independent variable. The phenomenon of high correlation between independent variables is called multicollinearity (May-Ostendorp et al., 2011). There is a strong correlation between the different indoor surface temperatures, which are in contact with each other and influenced by the same variables. In this chapter, the (moderate to strong) correlations between the indoor surface temperatures and the air temperature and outdoor temperature are discussed.

6.4.1 Indoor surface temperatures and air temperature

As explained in Chapter 2.2.1. is the indoor surface temperature an indicator of a room's historical heat gains or losses over time, heat gains or losses are captured in the indoor surface temperature. The air temperature generally fluctuates more, as its heat capacity is lower than the materials in a building. By analysing the behaviour of the different measured indoor surface temperatures compared to the air temperature, it is possible to get insight into how the thermal mass of a building influences the heating and cooling demand. In general, the pattern of the air temperature changes would be reflected in the indoor surface temperatures with a time delay.

In Figures 38 and 39 the different indoor surface temperatures and the air temperature during opening hours for the office are shown. There is a strong positive correlation between the indoor surface temperatures and the air temperature within the office (see Appendix H). The correlation is weaker with the inside of the outer wall.

In Figure 38, the bold blue line is the air temperature. Most of the indoor surface temperatures follow the same pattern and are in the same magnitude as the air temperature, however, the air temperature fluctuates more, with short peaks and troughs, where the indoor surface temperatures are more flatted. Remarkable are the indoor surface temperatures of the inside of the outer wall (the purple line) and the window (orange line). This surface temperature has the weakest correlation with the air temperature. Most of the time the temperature of the inside of the outer wall is lower than the air temperature, this could be explained by relatively higher thermal mass. On 19 and 22 March 2024, there are peaks of the indoor surface temperature of the inner of the outer wall, in contrast to the decreasing air temperature. On 19 March 2024, this could be explained by a higher outdoor temperature or solar radiation, as the surface temperature of the window (in orange) was high before this peak. It is possible that on 22 March 2024, there was another heat influencing this wall. The most fluctuating surface temperature is seen at the window, with the highest range of temperature. The peaks of the indoor surface temperature of the window are correlated to the peaks of the air temperature, there is still a strong correlation.

Compared to the office, the air temperature fluctuates more for short periods in the classroom, which further explains the more fluctuating thermal power in this room as concluded in Chapter 6.1. Where the air temperature in the office is most of the time between the indoor surface temperatures (see Figure 39), in the classroom, the air temperature is most of the time higher than the indoor surface temperatures, with the surface temperature of the system wall 2 as an exception (see Figure 39). Overall there are strong positive correlations between the indoor surface temperature and the air temperature, however, they are weaker than within the office, this could be explained by the larger volume of the classroom. The only negative correlation with the air temperature is with system wall 1

The surface temperature of system wall 1 is showing a high peak from 27 March until 2 April 2024. This system wall is between the classroom and an office. An explanation could be the influence of solar radiation, however, this could not be verified with the temperature of the window as the data on its surface temperature is missing. On this wall, there is a digital smart

board, that is emitting heat, which is influencing the air temperature by convection, and the temperature of the wall by conduction what is a slower process.
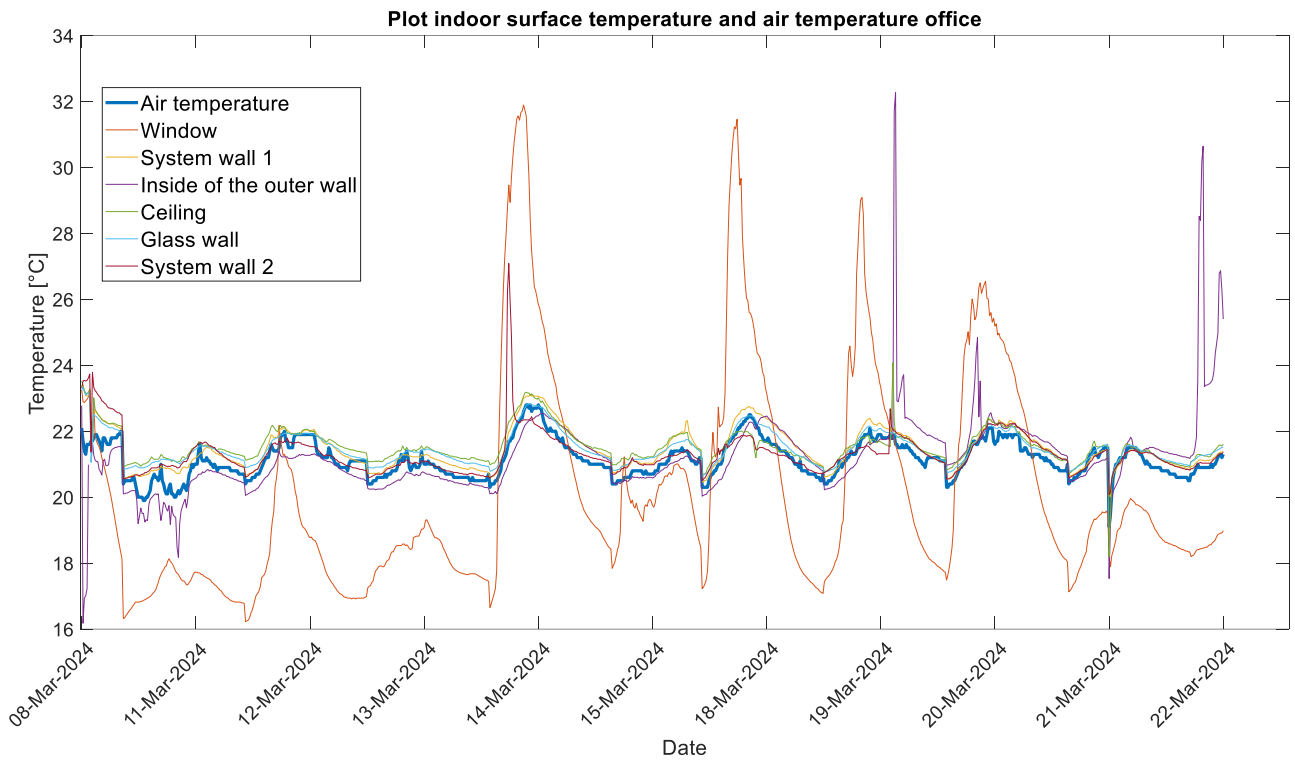
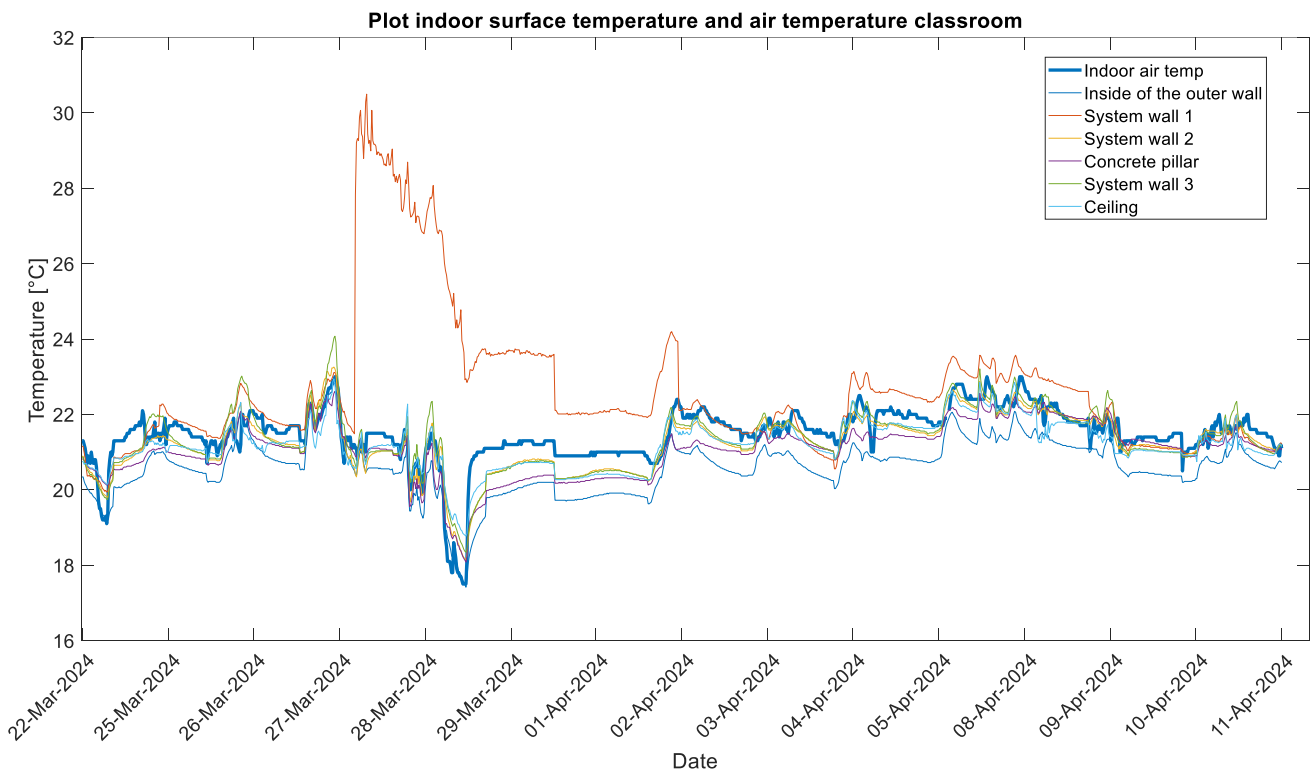*Figure 36: indoor air temperature (in bold blue) with the indoor surface temperatures office*

Figure 37: indoor air temperature (in bold blue) with the indoor surface temperatures classroom

6.4.2 Indoor surface temperatures and outdoor temperature

The correlation matrixes (Figures 34-37) show a moderate to strong correlation between the outdoor temperatures and the indoor surface temperatures for both rooms, the correlations in the classroom are weaker. Overall the correlations with the indoor surface temperatures and the indoor air temperature are stronger than with the outdoor temperatures.

As seen in Chapter 6.4.1, the indoor surface temperature of the window in the classroom shows high peaks and troughs, which could correlated to the outdoor parameters. For the outdoor temperature, this is confirmed by a relatively strong correlation between the surface temperature of the window and the outdoor temperature, in Appendix I and the correlation matrixes (Figure 34-37). The influence of the outdoor parameters, such as solar radiation and outdoor temperature (as shown in Figure 40), is the highest at the window as its U-value is lower than the façade's (see Chapter 4.2.) which results in the other indoor surface temperatures show a relatively weak correlation with the outdoor temperature. The window shows, for example, also a very strong correlation with the light intensity south. The weakest correlation is with the inside of the outer wall, which is remarkable as this is beside the window, the only surface which is in direct contact with the outdoor temperature, however, the difference is small with the other indoor surface temperatures.

In the classroom, a wider range of indoor surface temperatures is shown compared to the office. The weakest correlation is between the outdoor temperature and surface temperature of system wall 1, where the digital board is placed, this small correlation could be explained by the higher influence of the digital board on this temperature than by the outdoor temperature. The strongest correlation is with the concrete pillar in the classroom. There is no data available from the window, however, this pillar is at the side of the room that faces the outdoor.
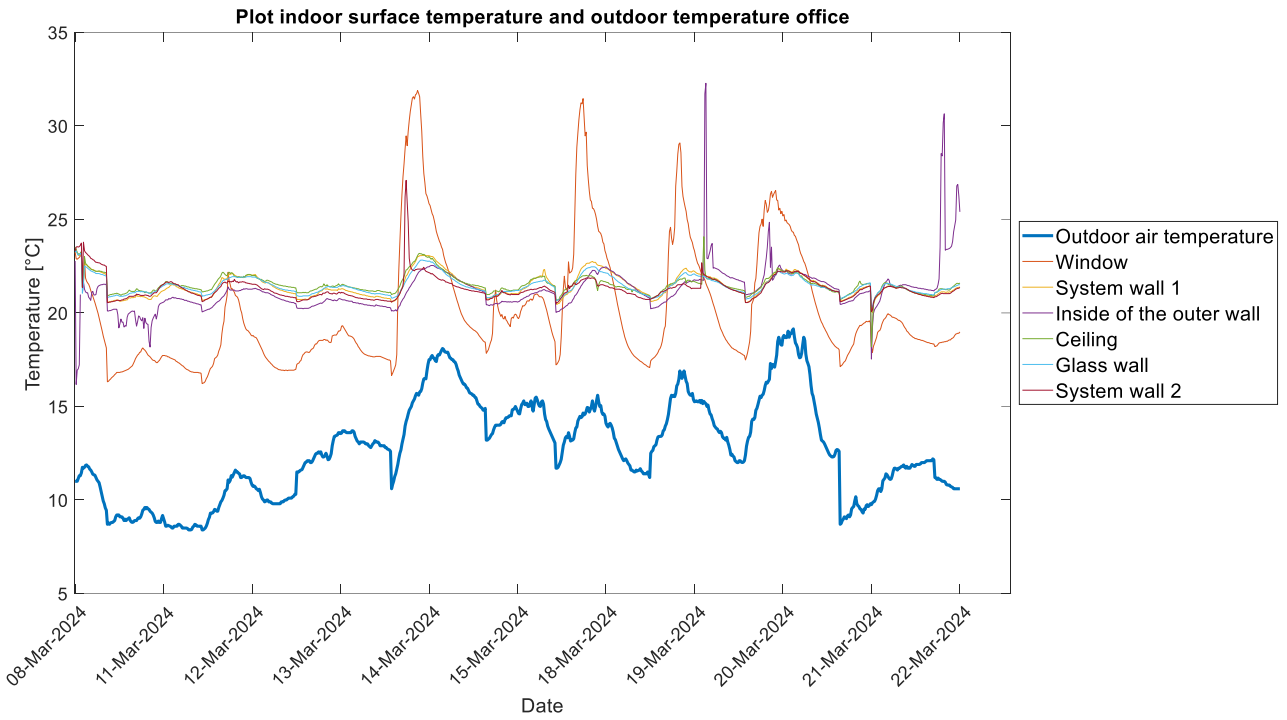
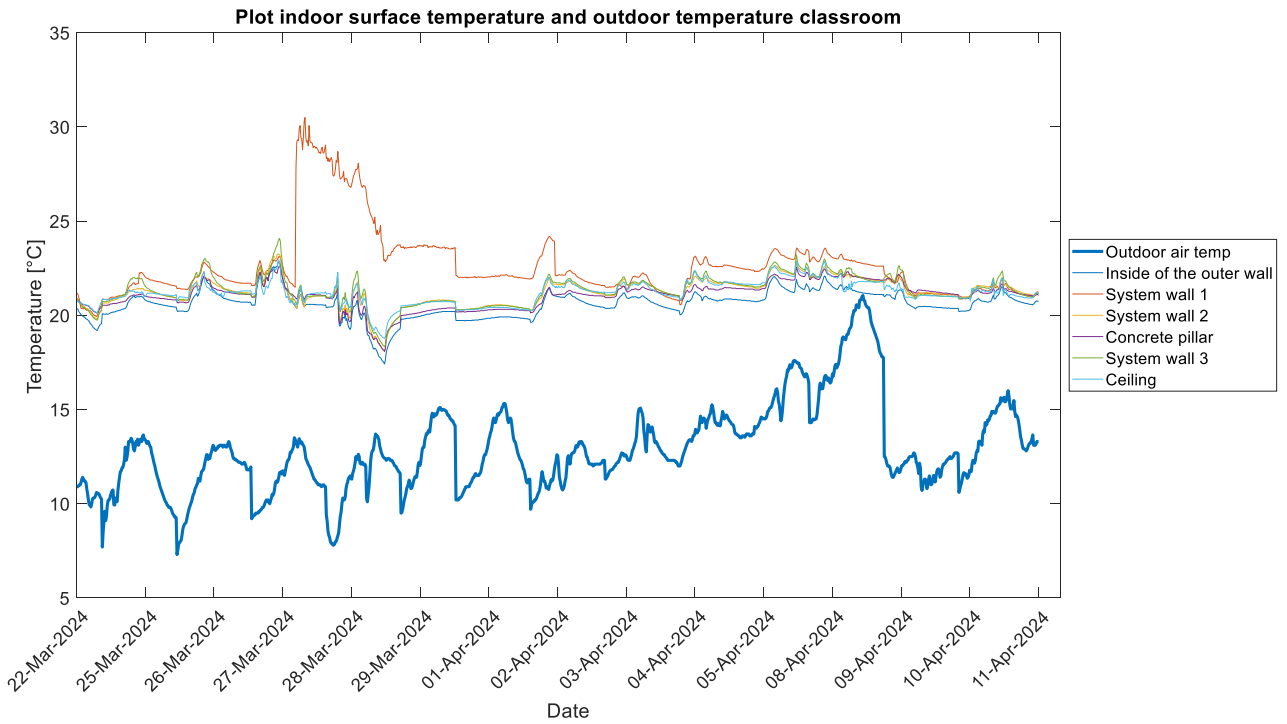*Figure 38: outdoor temperature (in bold blue) with the indoor surface temperatures office*

Figure 39: outdoor temperature (in bold blue) with the indoor surface temperatures classroom

7. Multivariate linear regression model

In the study of Jurado López (2017) two models are developed to model the heating demand during opening hours, whereby simulated data is used. The first model is a static model that includes only as independent variables the outdoor temperature and the indoor surface temperatures. In most cases, the indoor surface temperature is not known, a second model has developed with independent variables: the outdoor temperature, the indoor air temperature, internal heat gains, heat gains from solar radiation, and wind speed. This model is dynamic, the internal heat gains and the heat gains from solar radiation are included from time steps before.

7.1 Model A: Static model with outdoor and indoor surface temperatures

The first model is trained with the following independent variables; outdoor temperature and indoor surface temperatures.

Office

The Equations from the model derived in MATLAB for the office are 28 for heating and Equation 29 for cooling. In Table 18 the corresponding values, T-statistics, and P-values for the coefficients in Equation 28 and 29 are shown. The independent variables with a P-value > 0.05 are excluded. For both thermal modes, the outdoor temperature and the surface temperature of the window are included. The surface temperature of the window corresponds in contrast to the other indoor surface temperatures the most with the weather conditions due to the lower Rc-value than the walls in the room and the direct contact with the outdoor. For the heating, the surface temperature of system wall 1 is also included. With these independent variables, an accuracy of 17.76% is achieved. In Figure 42 an underestimation of the heating demand is shown, where the pattern of the modeled thermal demand is smooth in comparison to the more fluctuating pattern of the actual thermal demand. Besides the overlapping included independent variables with the heating demand, the surface temperatures of the glass wall and system wall 2 are included in the cooling model, where an accuracy of 34.70% is achieved. In Table 19, the RMSE for both rooms is shown, with a percentage of the maximum heating or cooling demand. The RMSE for the heating model is higher than the cooling demand. This is explained by the fact that there are more cooling data points (649) than heating data points (116) which makes the cooling model more trained. Together with the stronger correlations with the independent variables for the cooling model as concluded from the correlation matrixes, a higher accuracy of the cooling model in comparison to the heating model is achieved.

With this model and the corresponding independent variables, a prediction of the heating and cooling demand is made 7 hours in advance. In the case of the office, this is Thursday 21 March 2024 from 8 A.M until 3 P.M. During this period there is almost no cooling demand. The model is not able to model and predict a low or zero cooling demand, which results in an R^2 value of -416.13%. Until 11 AM there is a heating demand. The predicted heating demand is underestimated during the heating demand period. After the actual heating demand, the model predicts a heating demand, with a high peak at 1 P.M. This peak is due to a surface temperature decrease of 1.59 °C at 1:10 P.M. of the window, in comparison to the timesteps before. The model overacts to this temperature decrease, which results in an R^2 value of -10.98%.

$$Q_{demand,heat,office}(t) = C_0 + C_1\left(T_{s,outdoor}(t)\right) + C_2\left(T_{s,window}(t)\right) + C_3\left(T_{s,systemwall1}(t)\right) \quad (28)$$

$$Q_{demand,cold,office}(t) = C_0 + C_1\left(T_{s,outdoor}(t)\right) + C_2\left(T_{s,window}(t)\right) + C_6\left(T_{s,glasswall}(t)\right) + C_7\left(T_{s,systemwall2}(t)\right) \quad (29)$$

Table 18: coefficient values of the independent variables, with the corresponding T-statistics and P-value of the heating and cooling Equation 28 and 29 derived from model A for the office

Coefficient	Independent variable	Heating			Cooling		
		Value	T-statistics	P-value	Value	T-statistics	P-value
C_0		14917.00	10.74	4.40e-25	-6175.00	-9.26	2.32e-19
C_1	$T_{outdoor}(t)$	-54.53	-3.44	6.19e-4	17.84	2.92	3.60e-3
C_2	$T_{s,window}(t)$	58.55	3.71	2.20e-4	21.98	3.57	3.83e-4
C_3	$T_{s,systemwall1}(t)$	-700.55	-9.54	2.06e-20			
C_4	$T_{s,insideoftheouterwall}(t)$						
C_5	$T_{s,ceiling}(t)$						
C_6	$T_{s,glasswall}(t)$				427.27	10.43	8.12e-24
C_7	$T_{s,systemwall2}(t)$				-164.44	-5.87	6.58e-9

Table 19: R^2, R^2 adjusted, and RMSE (with a percentage of the maximum heating or cooling demand to the corresponding period) of the modelling and prediction of the heating and cooling demand in the office with model A

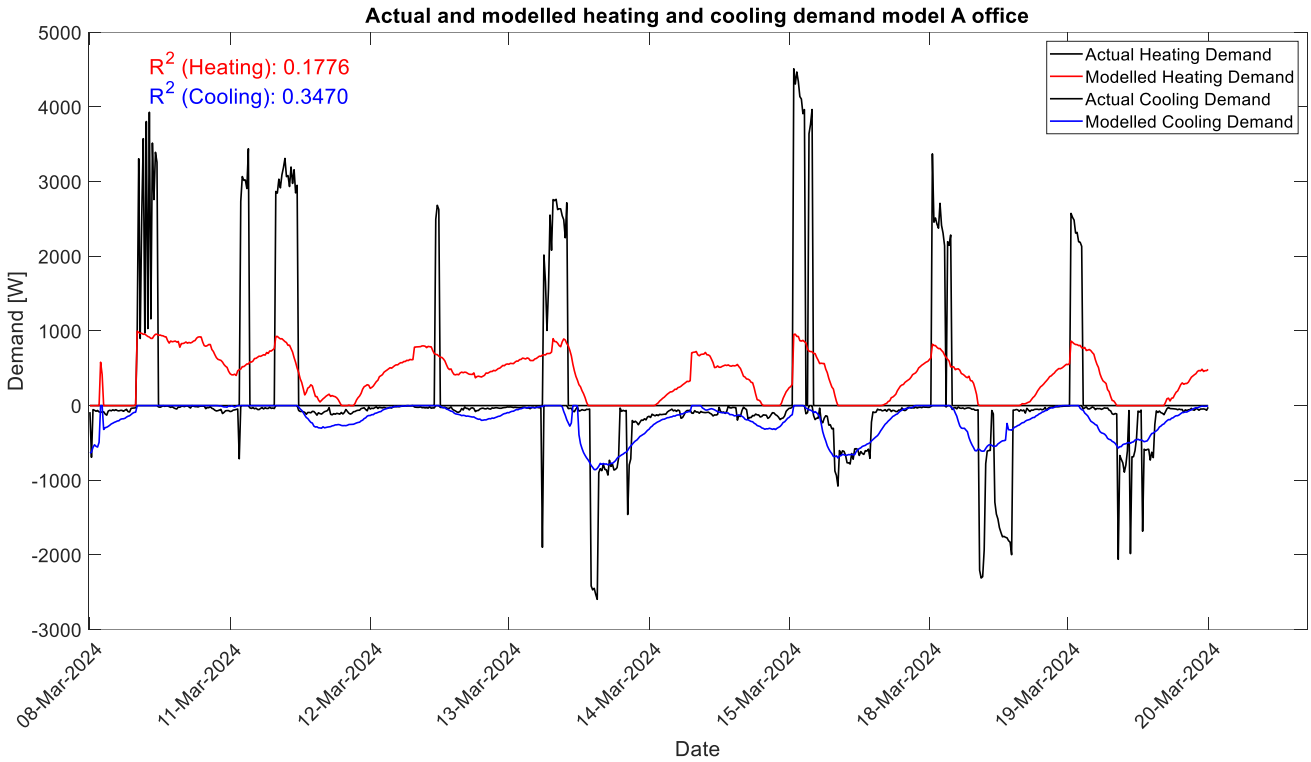|	Heating		Cooling	
	Modelling	*Prediction*	*Modelling*	*Prediction*
R^2	**17.76%**	**-10.98%**	**34.70%**	**-416.13%**
R^2 adjusted	17.40%	-33.83%	34.30%	-522.39%
RMSE [W]	847 (18.66%)	1693.80 (45.83%)	321 (12.66%)	43.90 (58.02%)

Figure 40: actual and modelled heating and cooling demand by model A for the office from 8 until 20 March 2024

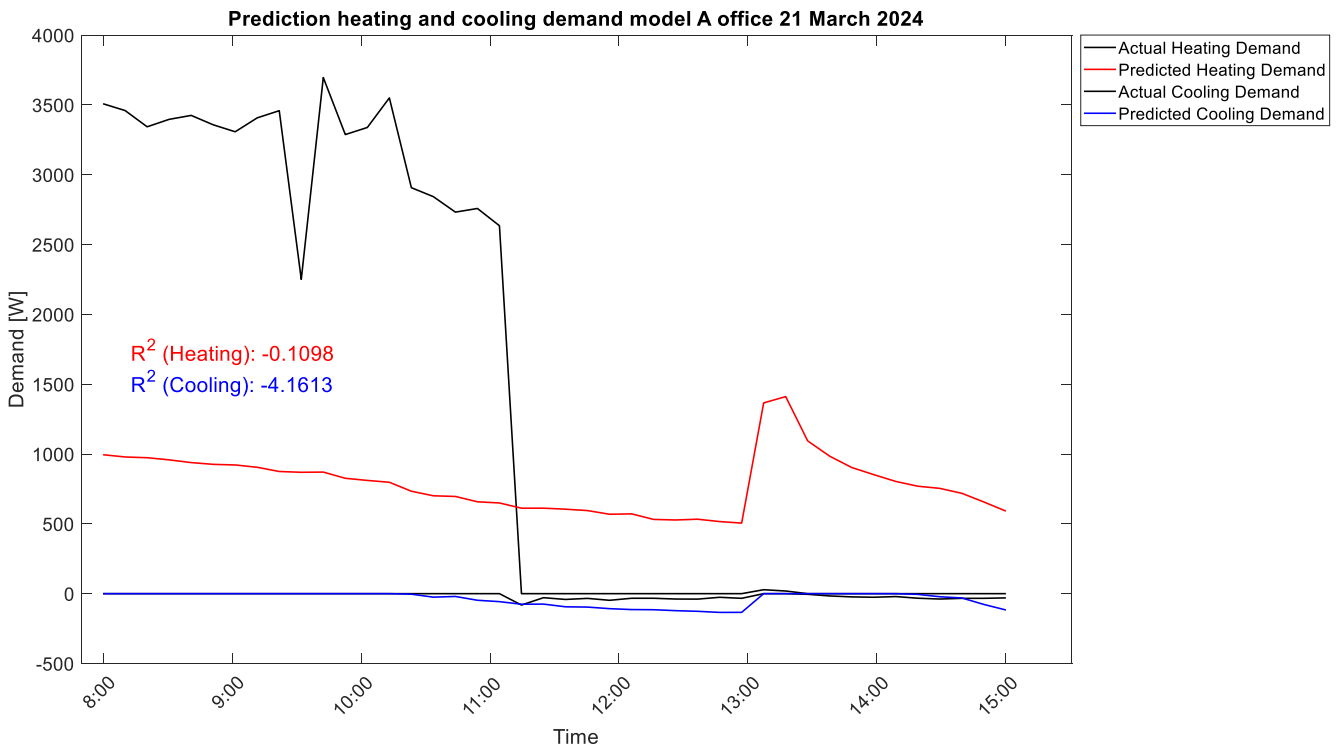*Figure 41: actual and predicted heating and cooling demand by model A for the office on 21 March 2024 8 A.M. - 1 P.M.*

Classroom

Remarkable for the model for the classroom is that the outdoor temperature is not included in the heating demand. This independent variable is included in the cooling demand Equation 30, however, the coefficient of the outdoor temperature has the lowest value in comparison to

the coefficient values of the other independent variables. This corresponds with the very weak correlation shown in the correlation matrix between the heating demand and the outdoor temperature in the classroom. Another remarkable exclusion of an independent variable is the indoor surface temperature of system wall 1, on this wall a smart board is placed, however, from Equations 30 and 31, it is concluded that this does not significant influence on the heating and cooling demand. The heating demand is modeled by model A with a very low accuracy of 3.98%, where an accuracy of 58% is achieved for the cooling demand. This was expected due to the relatively small amount of heating data points: 225 over 842 cooling data points. Thereby there is in sum a much smaller amount of heat supplied than supplied cold. This means the model is less trained to model heat demand. Thereby there are no moderate or strong correlations with the independent variables in the heating correlation matrix. The RMSE of the heating demand is proportionally lower than for the cooling demand, however, there are only two peaks of heating demand in comparison to more and longer periods of cooling demand, shown in Figure 44. The model is not able to reach the peaks of the heating demand, also the cooling demand is underestimated, but follows the pattern of the actual cooling demand.

With model A, a prediction of the heating and cooling demand is made for 9 April from 8 A.M. until 3 P.M. In contrast to the office, there is almost no heating demand and a high (fluctuating) cooling demand. The model can predict a heating demand of 0 W, however, the model is not able to predict a heating demand at 10:30 A.M. A higher heating demand at 11:50 A.M. is predicted. There are no big changes in the included indoor surface temperatures, that could cause this heating demand. An R^2 value of -8.83% is achieved. A higher R^2 value of 26.95% is reached by the prediction of the cooling demand. Model A is not able to reach the peaks of cooling demand but follows slightly the same pattern.

$$Q_{demand,heat,classroom} = C_0 + C_4 \left(T_{s,systemwall2}(t)\right) + C_5 \left(T_{s,concretepillar}(t)\right) + C_6\left(T_{s,systemwall3}(t)\right) + C_7 \left(T_{s,ceiling}(t)\right) (30)$$

$$Q_{demand,cold,classroom} = C_0 + C_1\left(T_{outdoor}(t)\right) + C_2 \left(T_{s.insideoftheouterwall}(t)\right) + C_4 \left(T_{s,systemwall2}(t)\right) + C_5 \left(T_{s,cocnretepillar}(t)\right) + C_6 \left(T_{s,systemwall3}(t)\right) + C_7 \left(T_{s,ceiling}(t)\right) (31)$$

Table 20: coefficient values of the independent variables, with the corresponding T-statistics and P-value of the heating and cooling Equation 30 and 31 derived from model A for the classroom

Coefficient	Independent variable	Heating			Cooling		
		Value	T-statistics	P-value	Value	T-statistics	P-value
C_0		113.01	0.46	0.64555	-669.94	-1.46	0.14
C_1	$T_{outdoor}(t)$				22.73	3.30	9.88e-4

C_2	$T_{s,inside\,of\,the\,outer\,wall}(t)$				354.44	3.45	5.81e-4
C_3	$T_{s,system\,wall1}(t)$						
C_4	$T_{s,system\,wall2}(t)$	-215.78	-3.53	4.28e-4	-1022.10	-8.55	4.60e-17
C_5	$T_{s,concrete\,pillar}(t)$	-91.78	-2.44	0.01	544.41	6.51	1.17e-10
C_6	$T_{s,system\,wall3}(t)$	172.32	4.19	3.00e-5	1547.30	18.26	2.43-64
C_7	$T_{s,ceiling}(t)$	129.67	3.39	7.19e-4	-1385.30	-18.26	2.73e-64

Table 21: R^2, R^2 adjusted, and RMSE (with a percentage of the maximum heating or cooling demand to the corresponding period) of the modelling and prediction of the heating and cooling demand in the classroom with model A

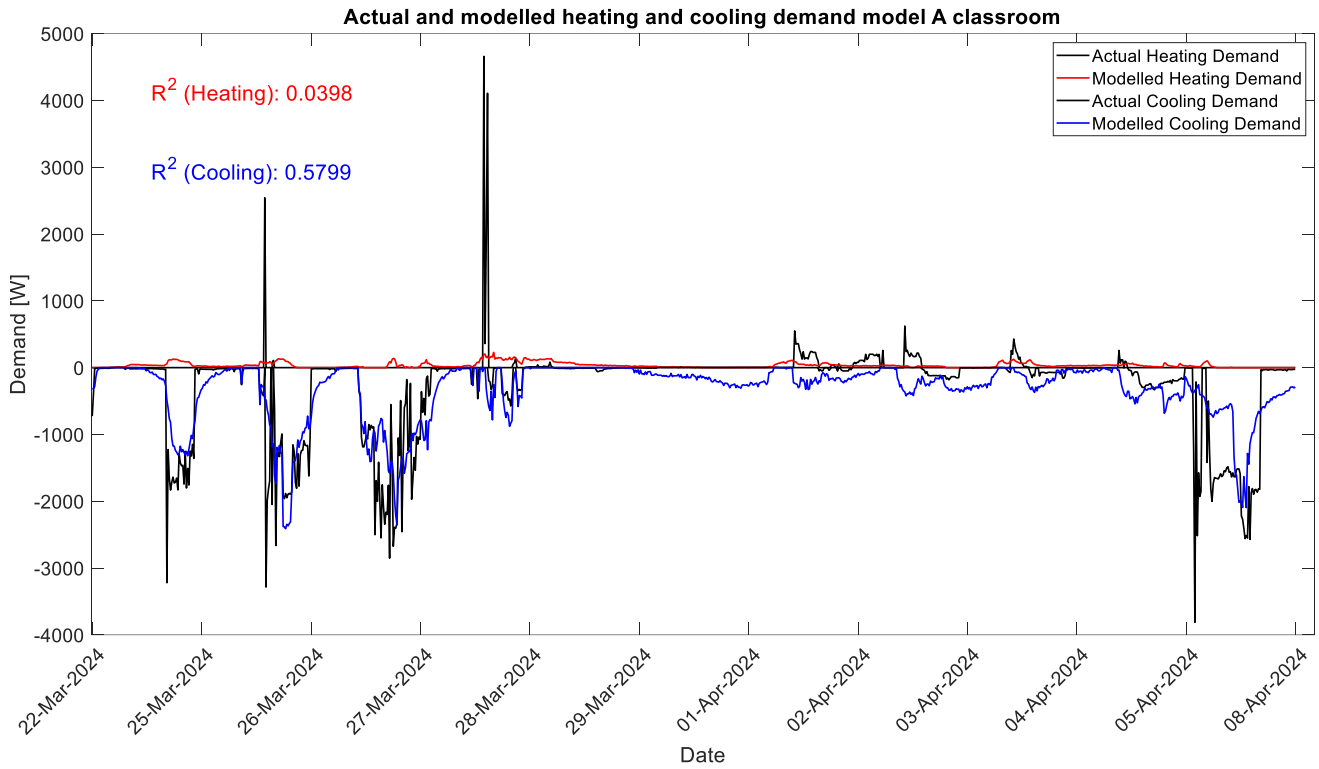|	Heating		Cooling	
	Modelling	*Prediction*	*Modelling*	*Prediction*
R²	**3.98%**	**-8.83%**	**58.00%**	**26.95%**
R² adjusted	3.59%	30.62%	57.70%	11.91%
RMSE [W]	225 (4.83%)	38.94 (15.87%)	407 (10.68%)	755.66 (24.23%)

Figure 42: actual and modelled heating and cooling demand by model A for the classroom from 22 March until 8 April 2024
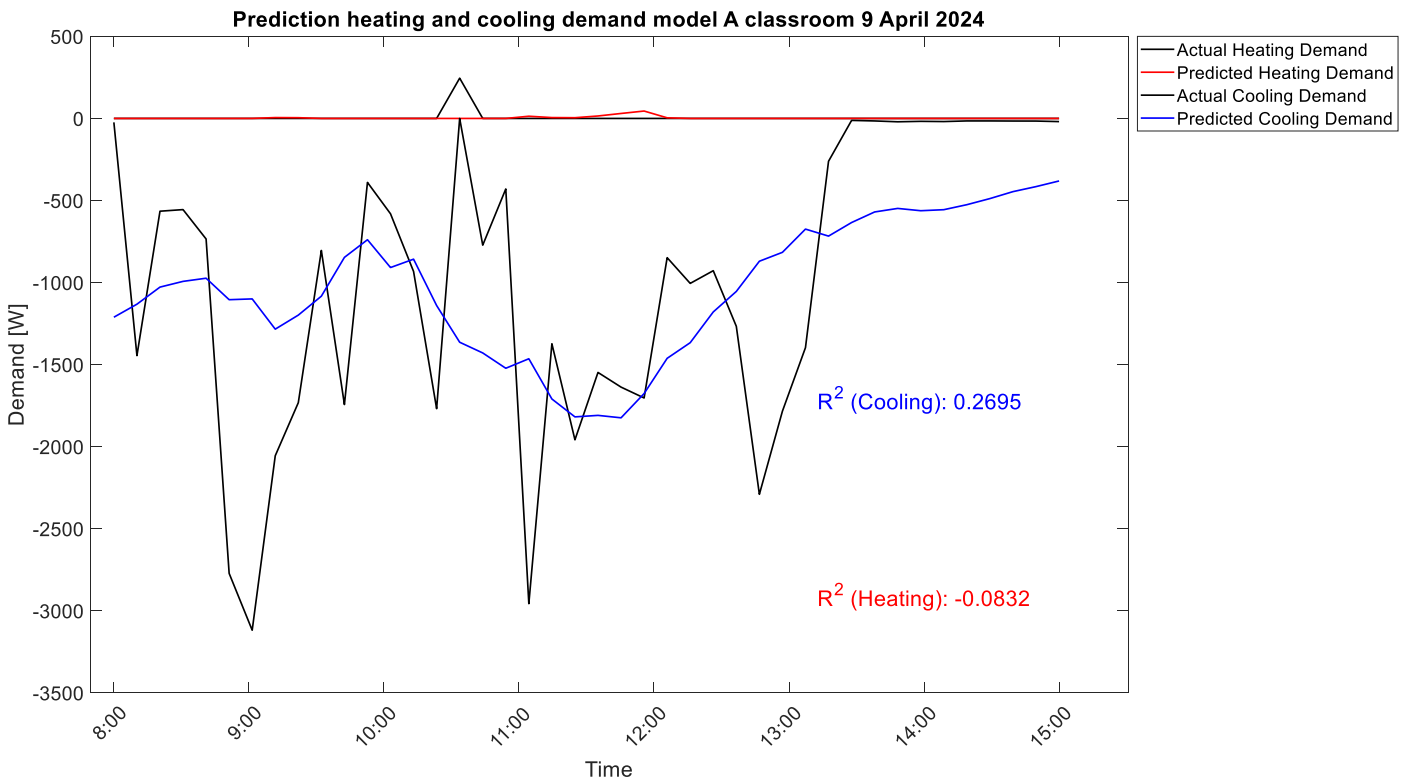

Figure 43: actual and predicted heating and cooling demand by model A for the classroom on 9 April 2024 8 A.M. - 1 P.M.

7.2 Model B: Dynamic model with outdoor and indoor air temperature, wind speed, solar heat gains, and internal heat gains

The second model was developed by Jurado López (2017), is a dynamic model whereby the indoor surface temperatures are excluded and replaced by time-delayed internal heat gains and heat gains from solar radiation. The following independent variables are used; outdoor temperature, indoor air temperature, windspeed, solar light intensities, and internal heat gains.

Office

The R^2 value of model B in the case of the office increased by 6.96% to 24.72% for heating and by 10.83% to 43.53% for cooling. The RMSE values decreased with 27 W and 17 W, for heat and cold, respectively. For the heating demand, there are no time delays calculated by the cross-correlation function, despite that with the introduction of the outdoor and indoor air temperature, wind speed, solar light intensity east, and internal heat gains a higher accuracy is achieved in comparison to Model A. Remarkable is that a higher accuracy with modelling the cooling demand is achieved with only solar light intensities. Where the pattern of the modeled heating and cooling demand by Model A was smooth, the modeled heating and cooling demand by Model B followed the fluctuating pattern of the actual heating and cooling demand (shown in Figure 46).

However, when model B is used to predict the heating demand, the R^2 value of the prediction decreases by 5.19%. In Figure 47 an underestimation of the heating demand is shown. Also with this model, there is a peak after 1 P.M. There is no big change in the values of the independent variables during this timestep. A cause could be that there are other independent variables needed to accurately predict low or zero heating demand. The R^2 value of predicting the cooling model has increased by 206.85%, however, this value is still indicating a very poor modelling ability of -209.28%.

$$Q_{demand,heat,office} = C_0 + C_1(T_{outdoor}(t)) + C_2(T_{indoorair}(t)) + C_3(Windspeed(t)) + $$
$$C_5(Solar\ light\ intensity\ east(t)) + C_7(Q_{internal}(t))\ (32)$$

$$Q_{demand,cold,office} = C_0 + C_4(Solar\ light\ intensity\ south\ (t)) + $$
$$C_5(Solar\ light\ intensity\ east\ (t)) + C_8(Solar\ light\ intensity\ south\ (t-n)) + $$
$$C_9\ (Solar\ light\ intensity\ east\ (t-n))\ (33)$$

Table 22: coefficient values of the independent variables, with the corresponding T-statistics and P-value of the heating and cooling Equation 32 and 33 derived from model B for the office

Coefficient	Independent variable	Heating			Cooling		
		Value	T-statistics	P-value	Value	T-statistics	P-value
C_0		12443	10.06	2.28e-22	-9.23	-0.56	0.57
C_1	$T_{outdoor}\ (t)$	-32.85	-2.35	0.02			
C_2	$T_{indoorair}\ (t)$	-553.59	-8.86	6.29e-18			
C_3	$Windspeed\ (t)$	-62.97	-3.94	8.92e-05			
C_4	Solar light intensity south (t)				10.91	8.44	1.84e-16
C_5	Solar light intensity east (t)	29.61	6.45	2.03e-10	-7.83	-3.84	1.32e-4
C_6	Solar light intensity west (t)						
C_7	$Q_{internal}\ (t)$	0.29	4.14	3.88e-05			
C_8	Solar light intensity south $(t-n)$ n for cooling: 17				9.90	7.96	6.87e-15
C_9	Solar light intensity east $(t-n)$ n for cooling: 25				9.33	3.96	8.22e-05
C_{10}	Solar light intensity west $(t-n)$ n for cooling: 1						
C_{11}	$Q_{internal}\ (t-n)$ n for cooling: 20						

Table 23: R^2, R^2 adjusted, and RMSE (with a percentage of the maximum heating or cooling demand to the corresponding period) of the modelling and prediction of the heating and cooling demand in the office with model B

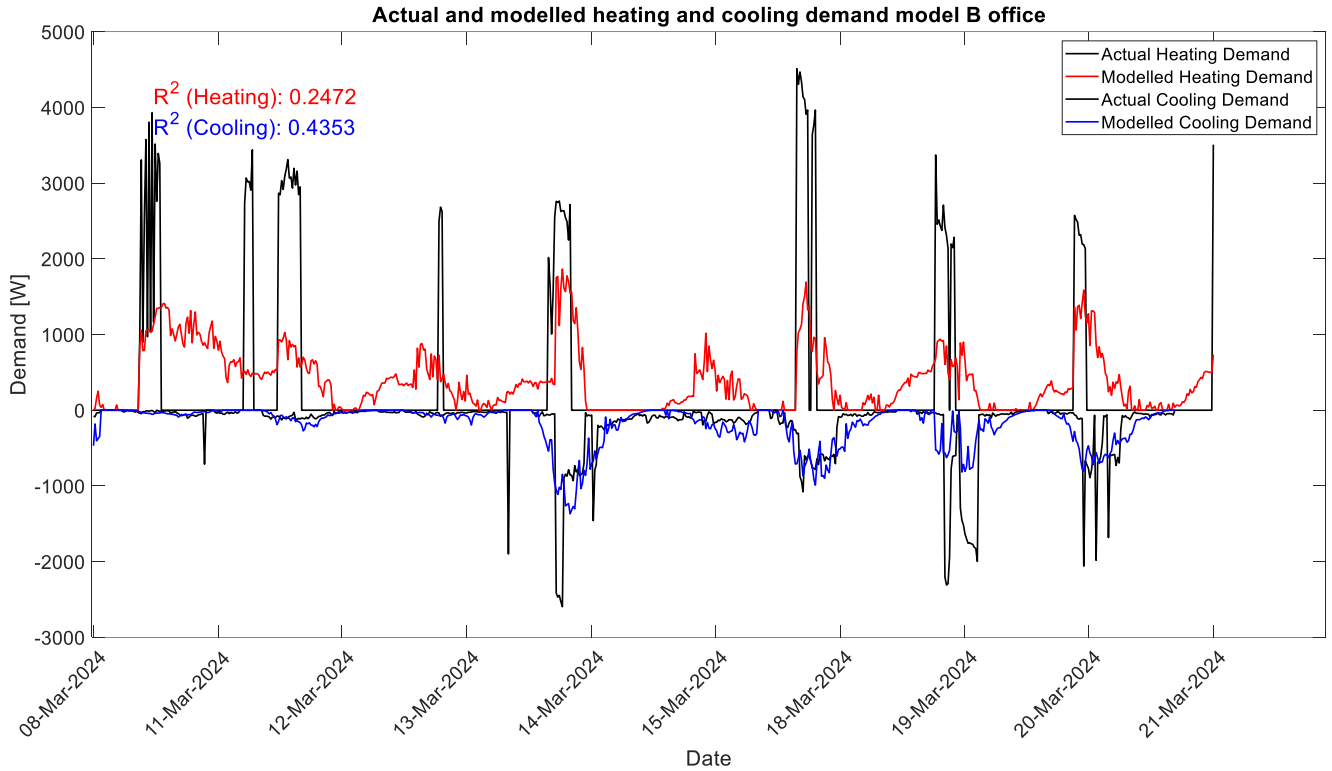|	**Heating**		**Cooling**	
	Modelling	*Prediction*	*Modelling*	*Prediction*
R²	**24.72%**	**-16.17%**	**43.53%**	**-209.28%**
R² adjusted	24.20%	-40.80%	43.20%	-322.68%
RMSE [W]	820 (18.07%)	1732.90 (46.89%)	304 (12%)	33.98 (41.71%)

Figure 44: actual and modelled heating and cooling demand by model B for the office from 8 until 20 March 2024
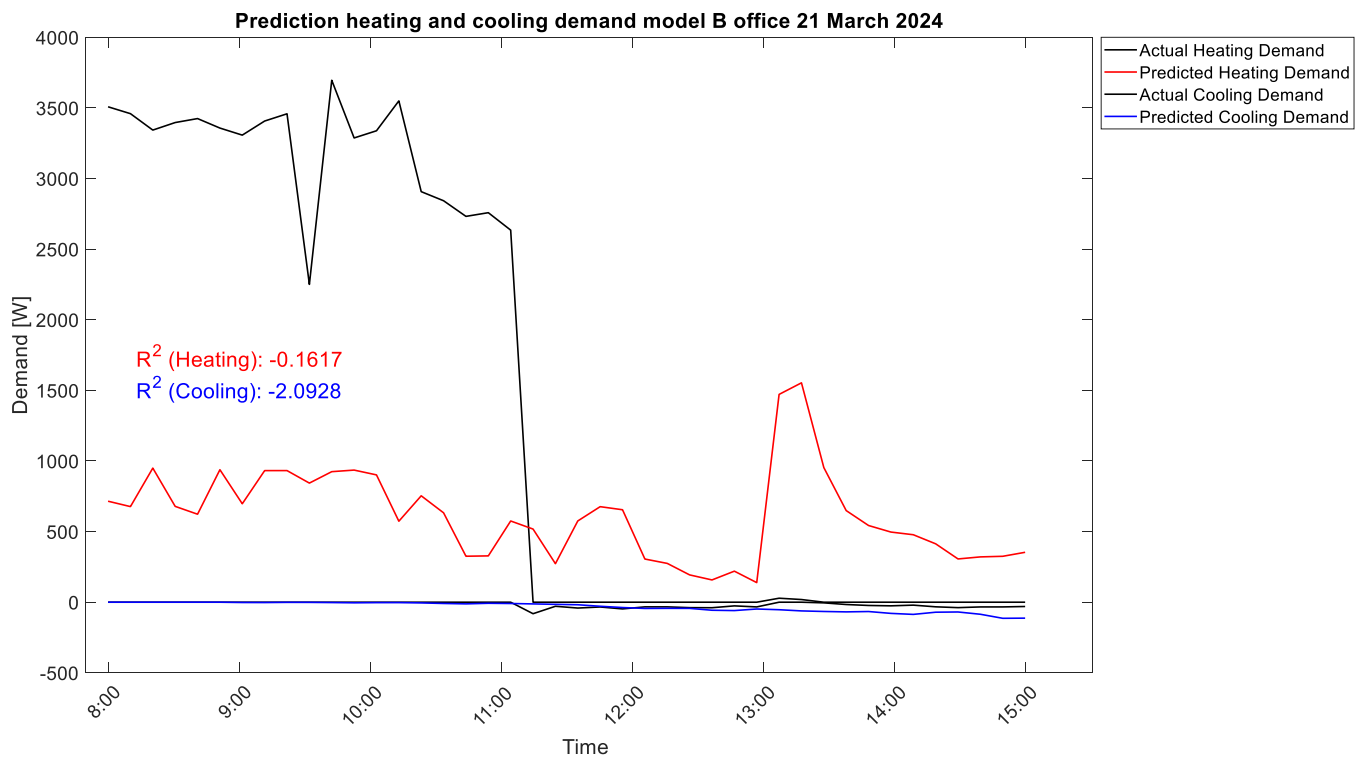

Figure 45: actual and predicted heating and cooling demand by model B for the office on 21 March 2024 8 A.M. - 1 P.M.

Classroom

By leaving out the indoor surface temperatures the R^2 value for modelling the heat dropped by 3.05% to an accuracy of 0.93% for heating in the classroom. The only included independent variable is the outdoor temperature. This corresponds to the Pearson correlation matrix, where the strongest correlation between the heating demand is the outdoor temperature, however, all the correlations are very weak. For the cooling demand, the model achieved an R^2 value of 44.91%, in comparison to Model A, a reduction of 13.90%. Only the internal heat gains and the solar light intensity east are included by the model as delayed independent variables. The current internal heat gains have a higher coefficient than the delayed internal heat gains. This is in contrast to the lower coefficient value of the current solar light intensity east, than of the delayed solar light intensity east, thereby the coefficients are low. Despite the decreased R^2 value, the RMSE value for modelling the cooling demand is slightly increased with 70 W. The cooling demand is more underestimated with model B than model A for periods with high cooling demand, however, the periods of no cooling demand are more fluctuating modeled by model B than the smooth modeled cooling demand by Model A.

The prediction of the heating and cooling demand by model B for the classroom results in a decrease of accuracy of 60.44% and 71.74%, for heating and cooling respectively. An almost constant heating demand of $35 - 40$ W is predicted. This corresponds to an almost constant outdoor temperature. In the case of cooling, the demand is more underestimated with model B than model A.

$$Q_{demand,heat,classroom} = C_0 + C_1\big(T_{outdoor}(t)\big) \ (34)$$

$$Q_{demand,cold,classroom} = C_0 + C_1\big(T_{outdoor}(t)\big) + C_2\big(T_{indoorair}(t)\big) + C_3\big(Windspeed(t)\big) + \\ C_5\big(Solar\ light\ intensity\ east\ (t)\big) + C_7\big(Q_{internal}(t)\big) + C_9\big(Solar\ light\ intensity\ east(t-n)\big) + \\ C_{10}\big(Q_{internal}(t-n)\big) \ (35)$$

Table 24: coefficient values of the independent variables, with the corresponding T-statistics and P-value of the heating and cooling Equations 34 and 35 derived from model B for the classroom

Coefficient	Independent variable	Heating			Cooling		
		Value	T-statistics	P-value	Value	T-statistics	P-value
C_0		144.29	3.73	2.02e-4	-3720.60	-8.32	3.22e-16
C_1	$T_{outdoor}\ (t)$	-8.86	-3.02	2.63e-3	45.74	6.41	2.37e-10
C_2	$T_{indoorair}\ (t)$				153.66	6.98	5.45e-12
C_3	$Windspeed\ (t)$				-54.67	-8.18	8.95e-16
C_4	$Solar\ light\ intensity$						

	south (t)							
C_5	Solar light intensity east (t)					8.87	4.29	2.02e-05
C_6	Solar light intensity west (t)							
C_7	$Q_{internal}(t)$					0.20	9.27	1.22e-19
C_8	Solar light intensity south $(t-n)$ n for cooling: 15							
C_9	Solar light intesnity east $(t-n)$ n for cooling: 30					20.08	9.88	5.81e-22
C_{10}	$Q_{internal}(t-n)$ n for cooling: 1					0.10	4.75	2.37e-06

Table 25: R^2, R^2 adjusted, and RMSE (with a percentage of the maximum heating or cooling demand to the corresponding period) of the modelling and prediction of the heating and cooling demand in the classroom with model B.

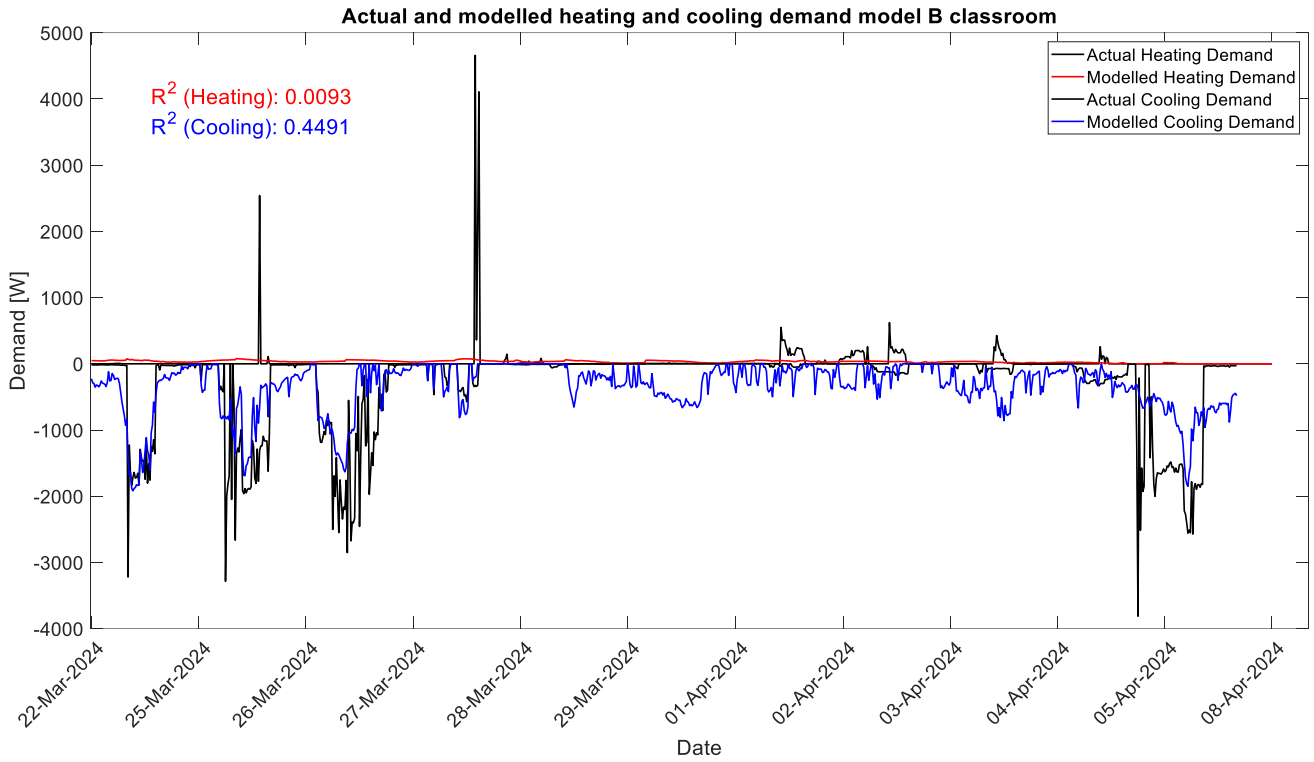|	Heating		Cooling	
	Modelling	Prediction	Modelling	Prediction
R²	**0.93%**	**-69.27%**	**44.91%**	**-44.79%**
R² adjusted	0.83%	-104.12%	44.50%	-91.50%
RMSE [W]	230 (4.94%)	48.68 (19.84%)	474 (12.44%)	1069.30 (34.29%)

Figure 46: actual and modelled heating and cooling demand by model B for the classroom from 22 March 2024 until 8 April 2024
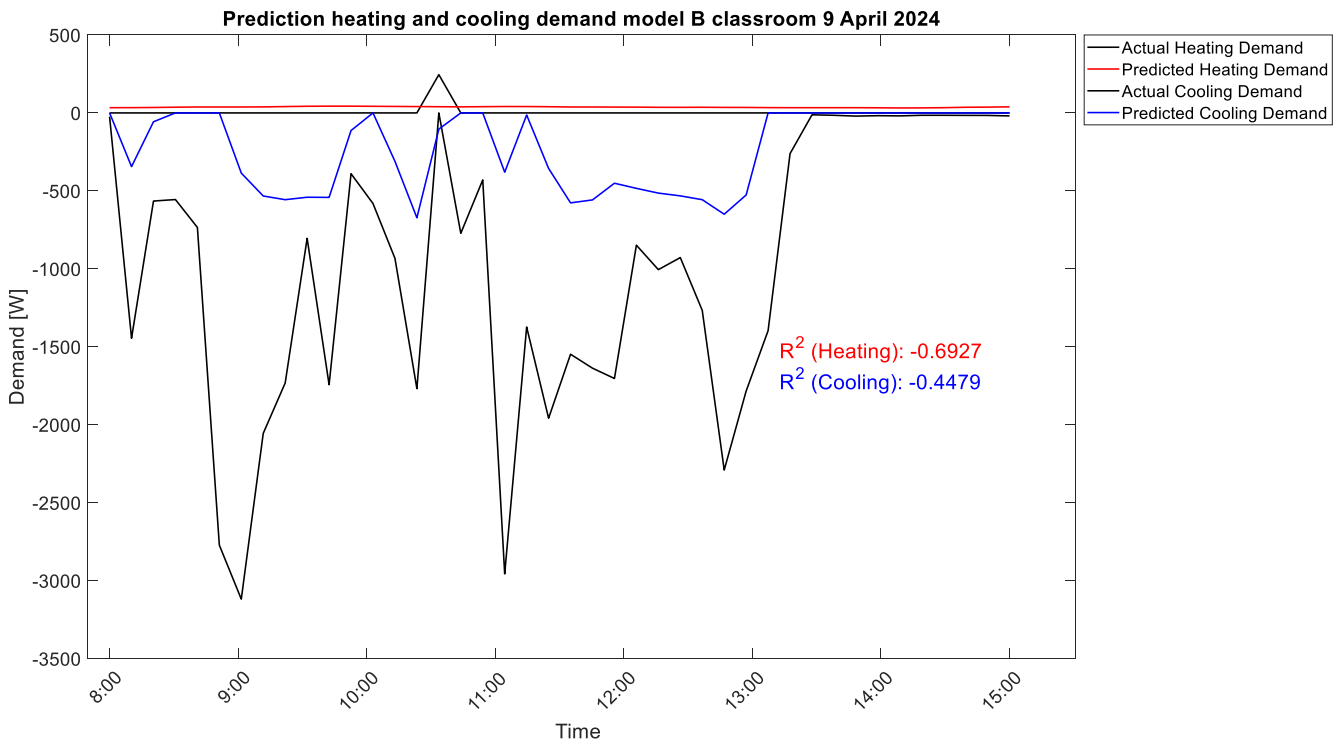

Figure 47: actual and predicted heating and cooling demand by model B for the classroom on 9 April 2024 8 A.M. - 1 P.M.

7.3 Model (A & B) analysis and selection of improvements

The R^2 values of modelling the heating demand with model A are 17.86% and 3.98%, for the office and the classroom, respectively. A higher accuracy of 24.72% is achieved by model B for the heating demand in the office, in contrast to a lower accuracy of 0.93% for the classroom. Higher accuracy is achieved by both models and rooms by modelling the cooling demand in comparison to the heating demand. Model A achieved an accuracy of 34.70% and 58.00%, for the office and classroom, respectively. The accuracy increased also within the office by using model B, with an accuracy of 24.72%. For the classroom, the accuracy decreased to 44.91%, by using model B.

By using both models to predict the heating and cooling demand for 7 hours in advance, almost all the R^2 values dropped to a negative value. The only positive R^2 value (= 26.95%) is achieved by predicting the cooling demand in the classroom by model A. Expect from the heating demand in the classroom, model B is less capable of predicting the heating or cooling demand in comparison to model A. This indicates that the indoor surface temperatures are essential to include in the model to predict the heating and cooling demand.

The two developed models of Jurado López (2017) are in the case of actual data, not able to model and predict the heating and cooling demand. Therefore the following improvements will be tested and analyzed in Chapter 8:

1. **Using the total thermal demand** *(Model C)*
   For both rooms, models A and B are better at modelling the cooling demand than the heating demand. The cooling demand has stronger correlations with the independent variables than the heating demand, thereby for both rooms there are more cooling data points than heating data points. Making one model for the total thermal demand per room could solve these model problems.

2. **Use of temperature differences** *(Model D)*
   In the study of Jurado López (2017), the choice is made to use the different temperatures separately instead of the temperature differences as shown in Equation 16, which is based on the thermal energy balance.

3. **Using all the independent variables based on the thermal energy balance** *(Model E)*
   There is a need for more (or other) independent variables as the heating and cooling demand is not accurately modeled and predicted by models A and B. A higher accuracy would be tried to achieve by including all the independent

variables based on the thermal energy balance. This makes it also possible to compare the results of using temperature differences instead of the single independent variables as input based on the thermal energy balance.

4. **Selecting independent variables based on correlation matrix** *(Model F₁)*

   From models A and B, the inclusion and exclusion of the independent variables correspond to the calculated Pearson correlation coefficients shown in the correlation matrix. Thereby the heating and cooling demand is underestimated, which can be caused by missing independent variables. By selecting the independent variables based on their Pearson correlation coefficients, a high accuracy is expected.

5. **Allowing interaction between independent variables** *(Model F₂)*

   As shown in the correlation matrixes, there are several strong correlations between independent variables and the heating or cooling demand. In the *stepwiselm* function of MATLAB there is the possibility to allow the inclusion of including the correlation between independent variables.

8. Improvements of the model

8.1 Model C: Using the total thermal demand

Instead of making an Equation per heating or cooling demand, it is tried to increase the accuracy of the model by making an Equation for the total thermal demand, where a negative value, corresponds to a cooling demand and a positive value to a positive demand. The same independent variable inputs are used as model A and model B. A static model (model C_1) is developed with the indoor surface temperatures and the outdoor temperature. The same independent variables as the dynamic model B are used to develop model C_2, however, there are no significant delays calculated by the cross-correlation function in MATLAB, which makes the model not dynamic anymore. In this model, the indoor surface temperatures are replaced by the windspeed, indoor air temperature, internal heat gains, and solar light intensities.

8.1.1 Model C_1

Office

By comparing the Equations derived from model A and C_1 for the office Equation 36 of model C_1 is different from the heating and cooling Equations 28 and 29 derived from model A. In both Equations from model A the outdoor temperature and the surface temperature of the window are included. The surface temperature of the inside of the outer wall was not included in both Equations from model A. The last included surface temperature of the glass wall is also included in the cooling Equation 29 from model A. The constant of 34 from model C_1 has increased in comparison to model A, which shows that the model is compensating the independent variables with a relatively high constant. The coefficients of the independent variables of model C_1 are in the same magnitude as model A (see Appendix K). The accuracy of model C_1 is 31.50%, which is higher than for heating and lower than for cooling in model A.

To conclude if modelling the total thermal demand instead of the heating and cooling demand separately, Figure 44 and Figure 50 are compared. The modelled thermal demand is higher for both heating and cooling, which means that model C_1 is less underestimating the thermal demand.

$$Q_{demand,office}(t) = constant + C_1\left(T_{s,outdoor}(t)\right) + C_2\left(T_{s,window}(t)\right) +$$
$$C_4\left(T_{s,insideoftheouterwall}(t)\right) + C_6\left(T_{s,glasswall}(t)\right) (36)$$

Table 26: R2, R2 adjusted, and RMSE (with a percentage of the maximum thermal demand to the corresponding period) of the modelling and prediction of the thermal demand in the office with model C_1

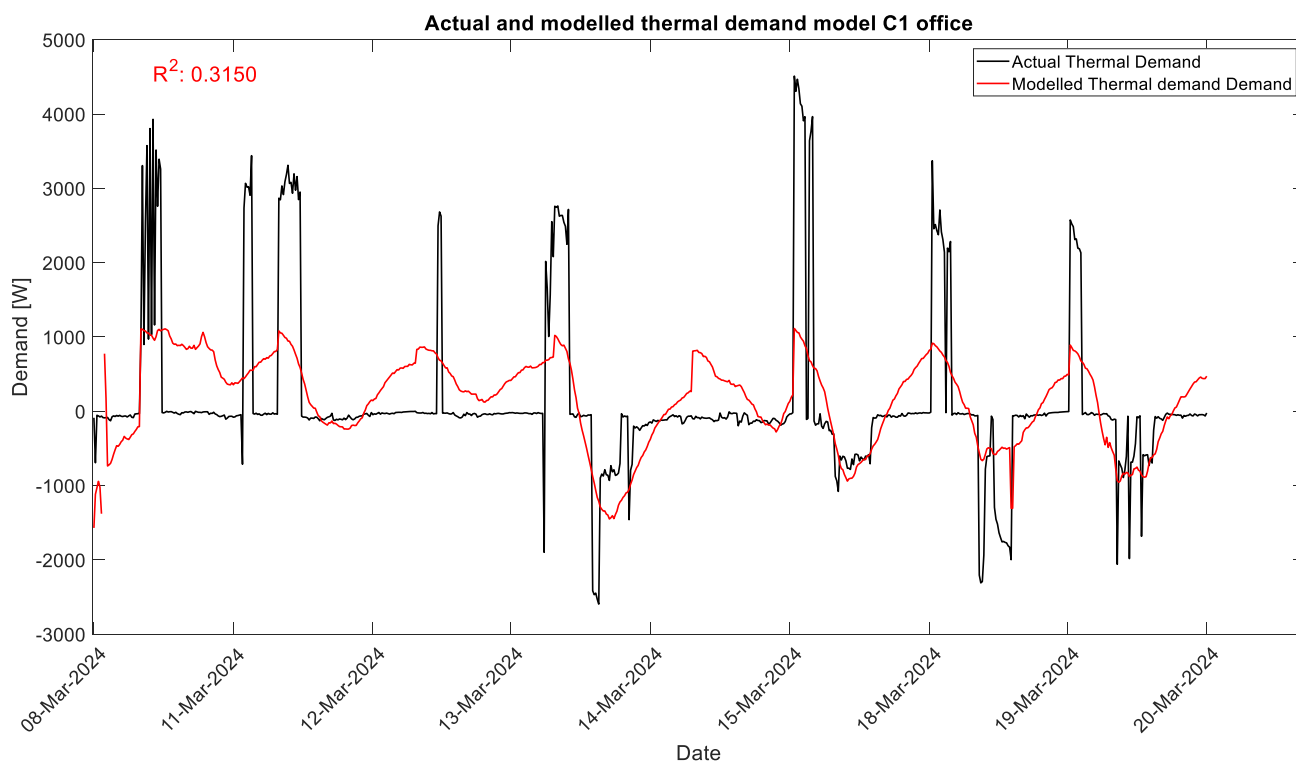|	Thermal demand	
	Modelling	*Prediction*
R²	**31.50%**	**-5.73%**
R² adjusted	31.10%	-2.75%
RMSE [W]	887.00 (19.55%)	1668.60 (45.14%)

Figure 48: actual and modelled thermal demand by model C_1 for the office from 8 until 20 March 2024
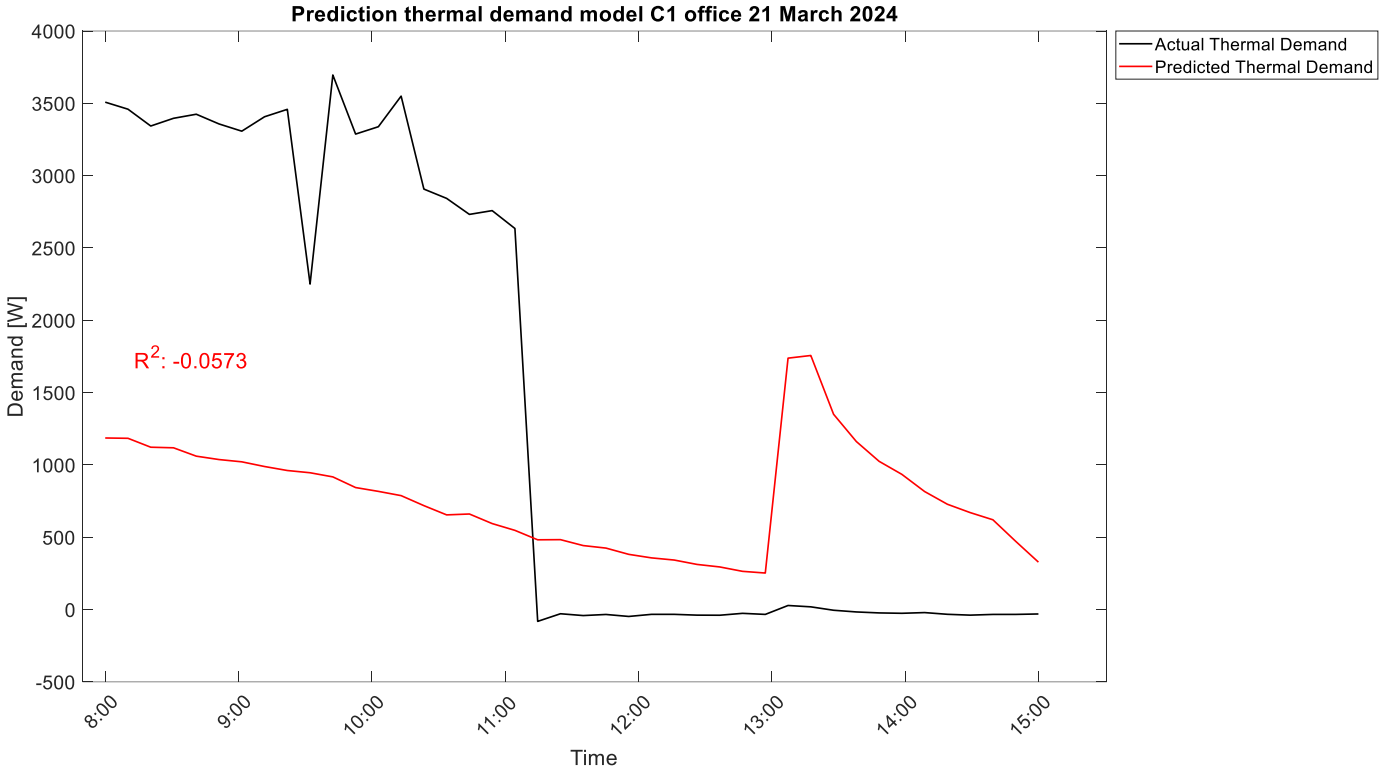
Figure 49: actual and predicted heating and cooling demand by model C_1 for the office on 21 March 2024 8 A.M. - 1 P.M.

Classroom

Equation 37 for the thermal demand for the classroom derived from model C_1 is the same as Equation 31 derived for cooling from model A, the independent variables from the Equation derived for heating from model A, are also included. The coefficients of Equation 37 (see Appendix K) are of the same magnitude as the cooling demand Equation 31 from model A. Also in the case of the classroom, the R^2 of model C_1 is higher than for heating and lower than for cooling with model A. The modelled thermal demand by model C_1, is almost the same for cooling as model A, however, for heating there are more periods of modelled heating demand, where there is no actual heating demand.

The model is not able to predict the heating demand, only a cooling demand is predicted. This prediction has a 2.54% lower accuracy for cooling.

$$Q_{demand,classroom} = constant + C_1\big(T_{outdoor}(t)\big) + C_2\Big(T_{s.insideoftheouterwall}(t)\Big) +$$

$$C_4\Big(T_{s,systemwall2}(t)\Big) + C_5\Big(T_{s,concretepillar}(t)\Big) + C_6\Big(T_{s,systemwall3}(t)\Big) + C_7\Big(T_{s,ceiling}(t)\Big) (37)$$

Table 27: R_2, R_2 adjusted, and RMSE (with a percentage of the maximum thermal demand to the corresponding period) of the modelling and prediction of the thermal demand in the classroom with model C_1

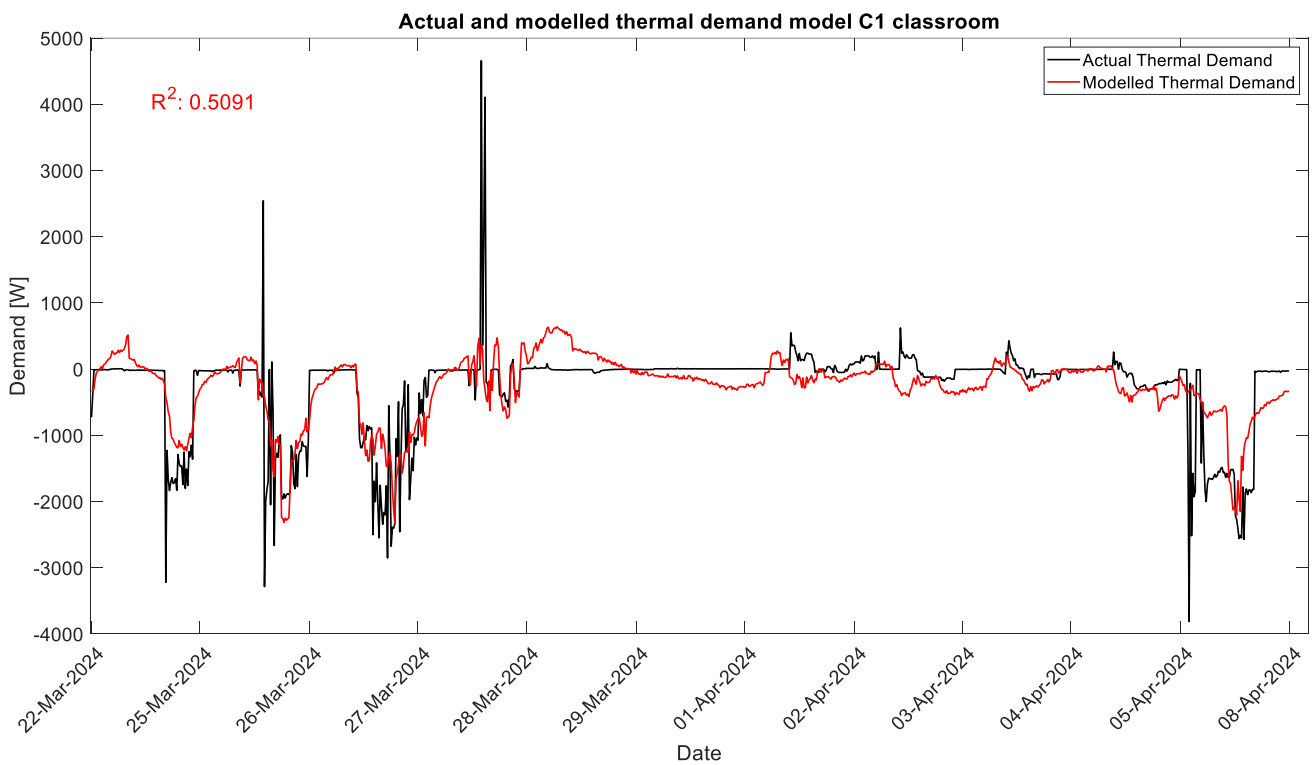|	**Thermal demand**	
	Modelling	*Prediction*
R²	**50.91%**	**24.41%**
R² adjusted	50.60%	8.85%
RMSE [W]	477 (10.24%)	775.04 (24.86%)

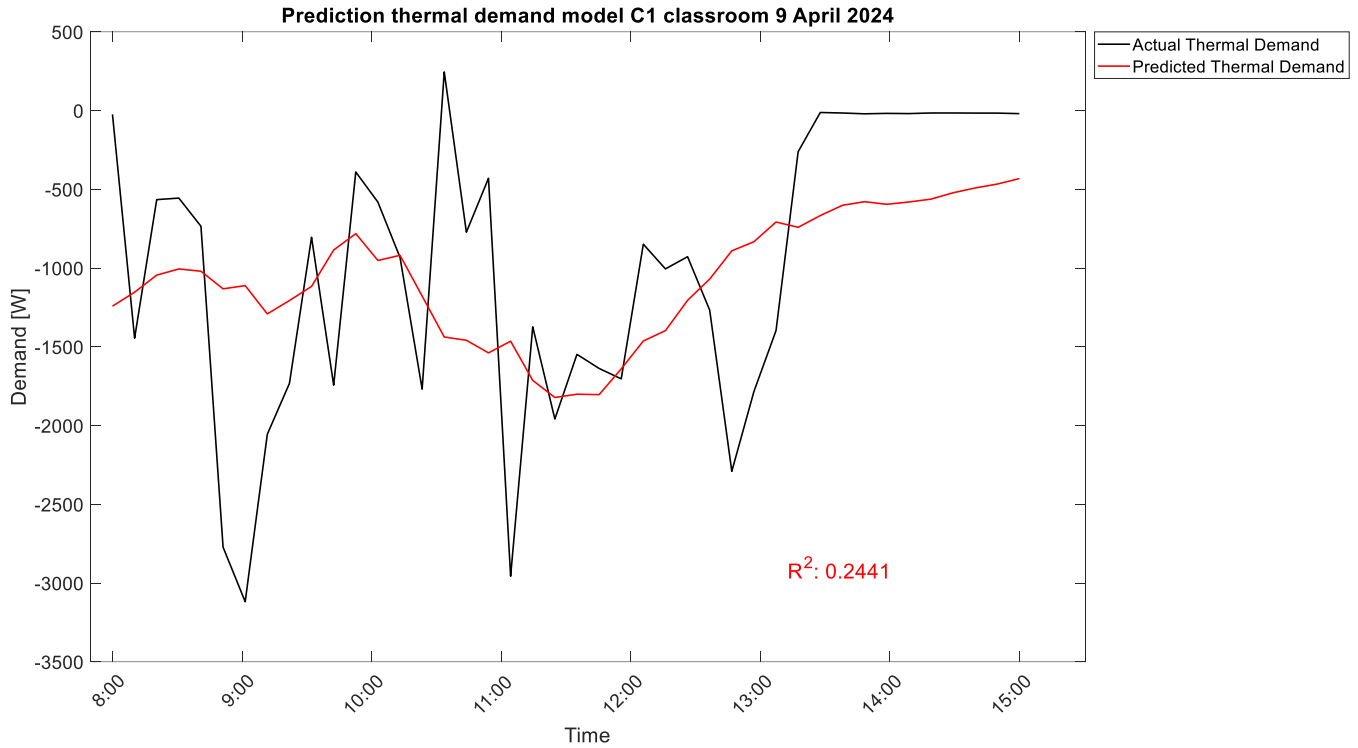*Figure 50: actual and modelled thermal demand by model C_1 for the classroom from 22 March until 8 April 2024*

Figure 51: actual and predicted heating and cooling demand by model C₁ for the classroom on 9 April 2024 8 A.M. - 1 P.M.

8.1.2 Model C₂

Office

The derived Equation 38 from model C₂ for the office is built up with the same independent variables as the heating Equation 32 derived from model B. Also in this case the coefficient values (see Appendix K) are of the same magnitude as the heating Equation 32 from model B. The R^2 value of model C₂ is of between the R^2 values of modelling heating and cooling demand with model B. The modelled thermal demand is higher in comparison to the modelled heating and cooling demand by model B. In other words, in the case of the classroom, modelling the thermal demand results in less underestimation of the thermal demand.

There is a small increase in the accuracy of predicting the heating demand with model C₂. No cooling demand is predicted.

$$Q_{demand,office} = C_1\big(T_{outdoor}(t)\big) + C_2\big(T_{indoorair}(t)\big) + C_3\big(Windspeed(t)\big) + C_5\big(Solar\ light\ intensity\ east(t)\big) + C_7\big(Q_{internal}(t)\big)\ (38)$$

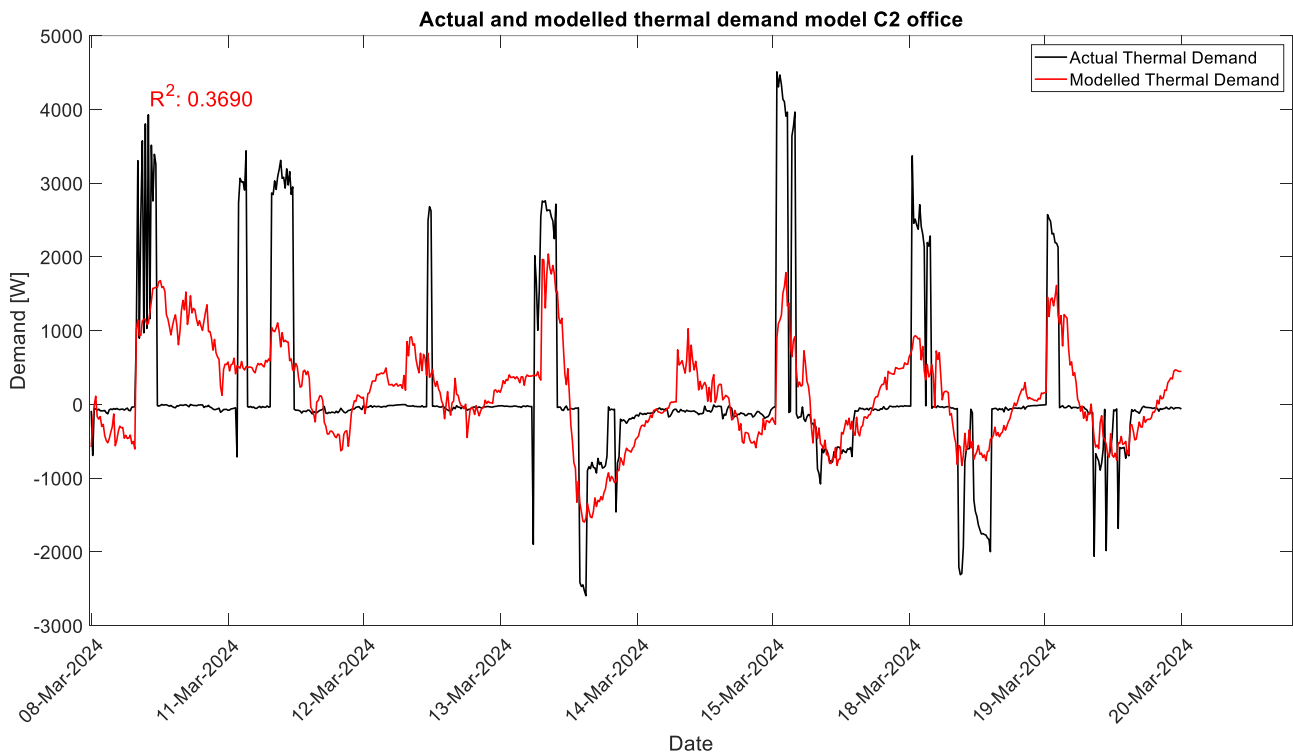|	Thermal demand	
	Modelling	*Prediction*
R²	**36.90%**	**-14.97%**
R² adjusted	36.50%	-38.64%
RMSE [W]	854 (18.82%)	1740.00 (47.08%)

Figure 52: actual and modelled thermal demand by model C₂ for the office from 8 until 20 March 2024
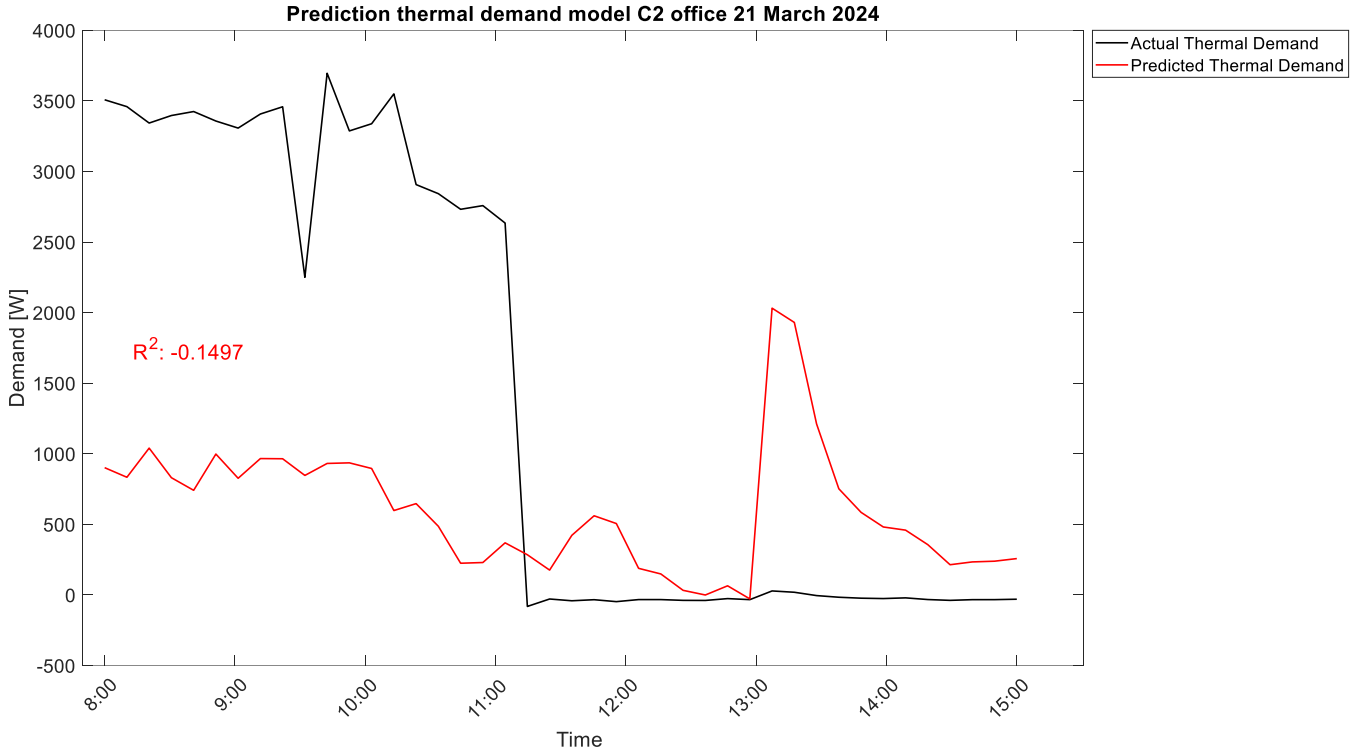
Figure 53: actual and predicted heating and cooling demand by model C_2 for the office on 21 March 2024 8 A.M. - 1 P.M.

Classroom

Equation 39 (Coefficient values can be found in Appendix C) derived from model C_2 for the classroom excludes the current and delayed solar light intensity east and the delayed internal heat gains in comparison to Equations 34 and 35 derived from model B. New included independent variables are the solar light intensities in the orientations south and west. The coefficient values are of the same magnitude as in model B, except the coefficient value of the air temperature. The indoor air temperature was only included in model B for modelling the heating demand with a coefficient value of -553.59, whereas this was in model C_2 -151.20. Modelling the thermal demand by model C_2, results in an overestimation of the heating demand, during periods of no actual heating demand, where the peaks are not reached. The prediction of the thermal demand is reaching the peak of the actual heating demand, however also in this case a heating demand is predicted when there is no actual heating demand and the predicted cooling demand is much lower.

$$Q_{demand,cold,classroom} = constant + C_1\big(T_{outdoor}(t)\big) + C_2\big(T_{indoorair}(t)\big) + C_3\big(Windspeed(t)\big) +$$
$$C_4\big(Solar\ light\ intensity\ south\ (t)\big) + C_6\big(Solar\ light\ intensity\ west(t)\big) + C_7\big(Q_{internal}(t)\big)$$

$$(39)$$

Table 29: R_2, R_2 adjusted, and RMSE (with a percentage of the maximum thermal demand to the corresponding period) of the modelling and prediction of the thermal demand in the classroom with model C_2

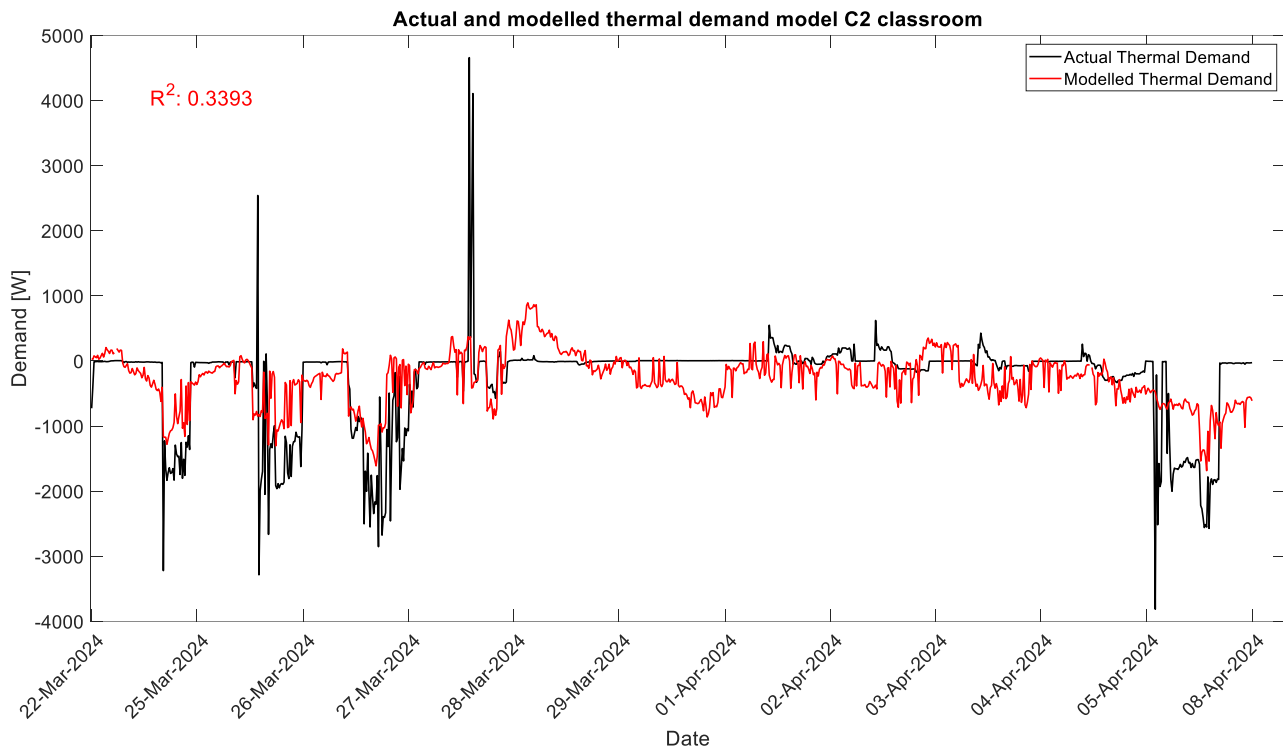|	**Thermal demand**	
	Modelling	*Prediction*
R²	**33.93%**	**-68.81%**
R² adjusted	33.50%	-103.56%
RMSE [W]	557.00 (11.95%)	1158.2 (37.14%)

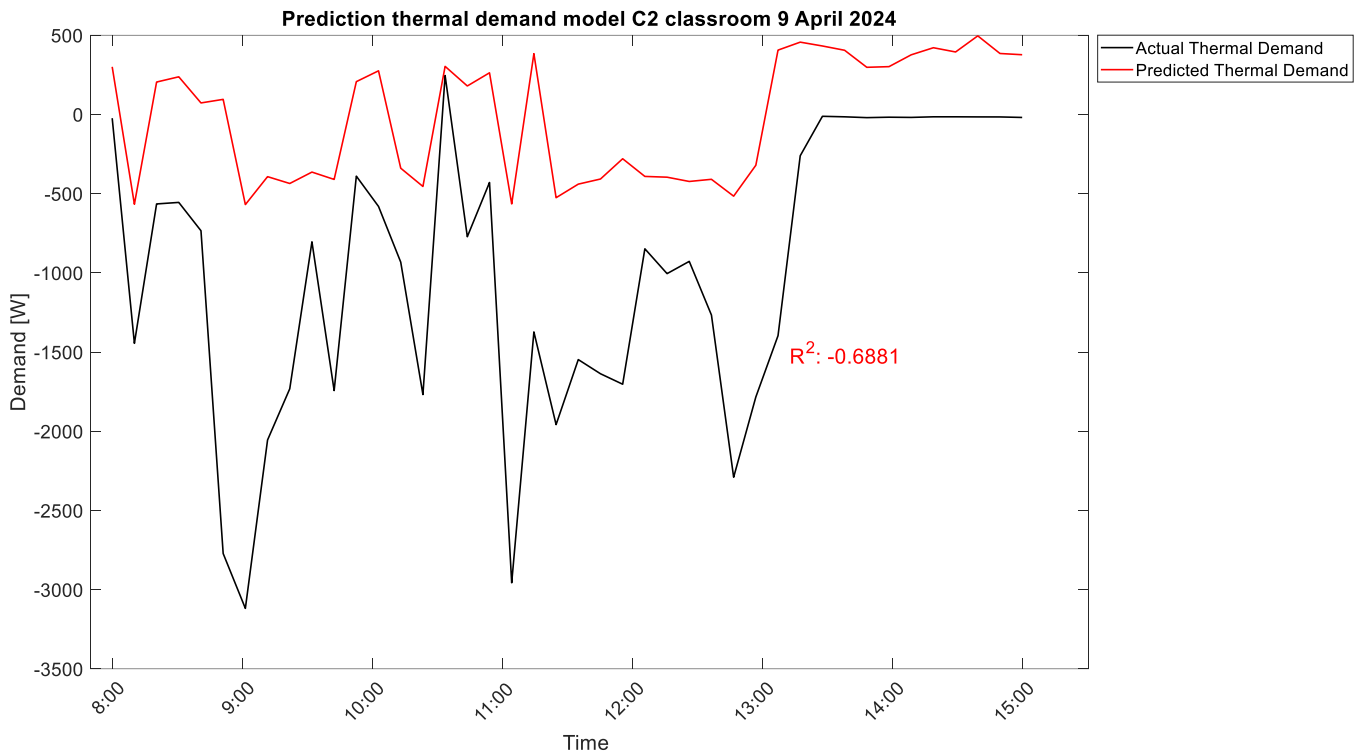*Figure 54: actual and modelled thermal demand by model C_2 for the classroom from 22 March until 8 April 2024*

Figure 55: actual and predicted heating and cooling demand by model C₂ for the classroom on 9 April 2024 8 A.M. - 1 P.M.

8.1.3 Model C analysis

Modelling the thermal demand instead of separately the heating and cooling demand, is applied to test if this would result in a higher accuracy for the heating demands in both rooms with less data points than the cooling demand, with a higher modelling accuracy. In the case of the office, the modelled heating and cooling demand is higher, there is less underestimation of the thermal demand. For the classroom modelling the thermal demand, results in an overestimation of the modelled heating demand when there is no actual heating demand. During the periods of actual heating demand, the model is not able to model this heating demand. The models are still not able to accurately predict the heating and cooling demand, especially, in the cases where there is almost none or zero demand. The predicted heating demand is overestimated by model C_2 for the classroom and underestimated by model C_1 for both rooms. For both rooms, the cooling demand is underestimated. The different model behaviors, explained by the different correlations with the independent variables, show that separate models need to be developed for heating and cooling demand.

8.2 Model D: use of temperature differences

Models A, B and C use the independent variables, as a single data input. This is in contrast to Equation 16 derived from the thermal energy balance. In this Equation, the temperature differences are used. Thereby the windspeed is multiplied by the difference between the outdoor and indoor air temperature. The wind speed does not have a big influence on itself (see also the correlation matrixes). The impact of the windspeed on the heating or cooling demand is determined by the temperature difference of the indoor and outdoor air. With the available data Equation 16 is rewritten to Equation 40.

$$Q_{demand}(t) = constant + C_1((T_{outdoor} - T_{indoorair})(t)) + C_2((V_{wind} \cdot (T_{outdoor} - T_{indoorair}))(t)) + C_{3-5}(Q_{solar}(t)) + C_6(Q_{internal}(t)) + C_{7-12}((T_{indoorair} - T_{surface})(t)) + C_{13}((T_{outdoor} - T_{indoorair})(t-n)) + C_{14}((V_{wind} \cdot (T_{outdoor} - T_{indoorair}))(t-n)) + C_{15-17}(Q_{solar}(t-n)) + C_{18}(Q_{internal}(t-n)) + C_{19-24}((T_{indoorair} - T_{surface})(t-n)) \ (40)$$

For each independent variable, a delay in the heating or cooling demand is calculated, so for example, for $(T_{outdoor} - T_{indoorair})$, shown in Appendix L, with the coefficient values and the T-statistics and P-values

Office

The Equation derived from model D for the heating and cooling demand, are Equations 41 and 42. There are calculated delays for the solar light intensities, internal heat gains and the difference between the indoor air temperature and the surface temperature of the window, however, they are not included by the model, which makes the model static for the office. The cooling demand is modelled without using the outdoor temperature and the wind speed. In contrast to model A where the outdoor temperature and the indoor surface temperatures were used as data input, for heating and cooling, two indoor surface temperatures are included in model B in comparison to four in model A. The model was expected to be complex with relatively a lot of independent variables, in comparison to the other models where fewer independent variables were included, however, only a few independent variables are included. With these variables, an R^2 value of 25.39% and 43.94%, for modelling the heating and cooling, respectively is achieved. The accuracy of model D for the classroom is slightly higher than model B. For model B the indoor surface temperatures are not needed, which makes the model easier to implement in an MPC, as these temperatures are often not measured.

Also, for the prediction, the difference in accuracy is not very significant compared to model B. What is remarkable is that model D is more robust than models A-C for (big) changes in the values of the independent variables as the peak at 1:10 of heating demand which was predicted

by models A-C, is not predicted by model D. However, also this model is not able to predict the heating and cooling demand accurately.

$$Q_{demand,heat,office}(t) = constant + C_2((V_{wind} \cdot (T_{outdoor} - T_{indoorair}))(t)) +$$

$$C_3(Solar\ light\ intensity\ south\ (t)) + C_4(Solar\ light\ intensity\ east\ (t)) + C_6(Q_{internal}(t)) +$$

$$C_7\left((T_{indoorair} - T_{s,window})(t)\right) + C_8\left((T_{indoorair} - T_{s,systemwall1})(t)\right)(41)$$

$$Q_{demand,cold,office}(t) = constant + C_3(Solar\ light\ intensity\ south(t)) +$$

$$C_4(Solar\ light\ intensity\ east\ (t)) + C_8\left((T_{indoor} - T_{s,systemwall1})(t)\right) + C_9((T_{indoorair} -$$

$$T_{s,insideoftheouterwall})(t))(42)$$

Table 30: R_2, R_2 adjusted, and RMSE (with a percentage of the maximum heating or cooling demand to the corresponding period) of the modelling and prediction of the heating and cooling demand in the office with model D

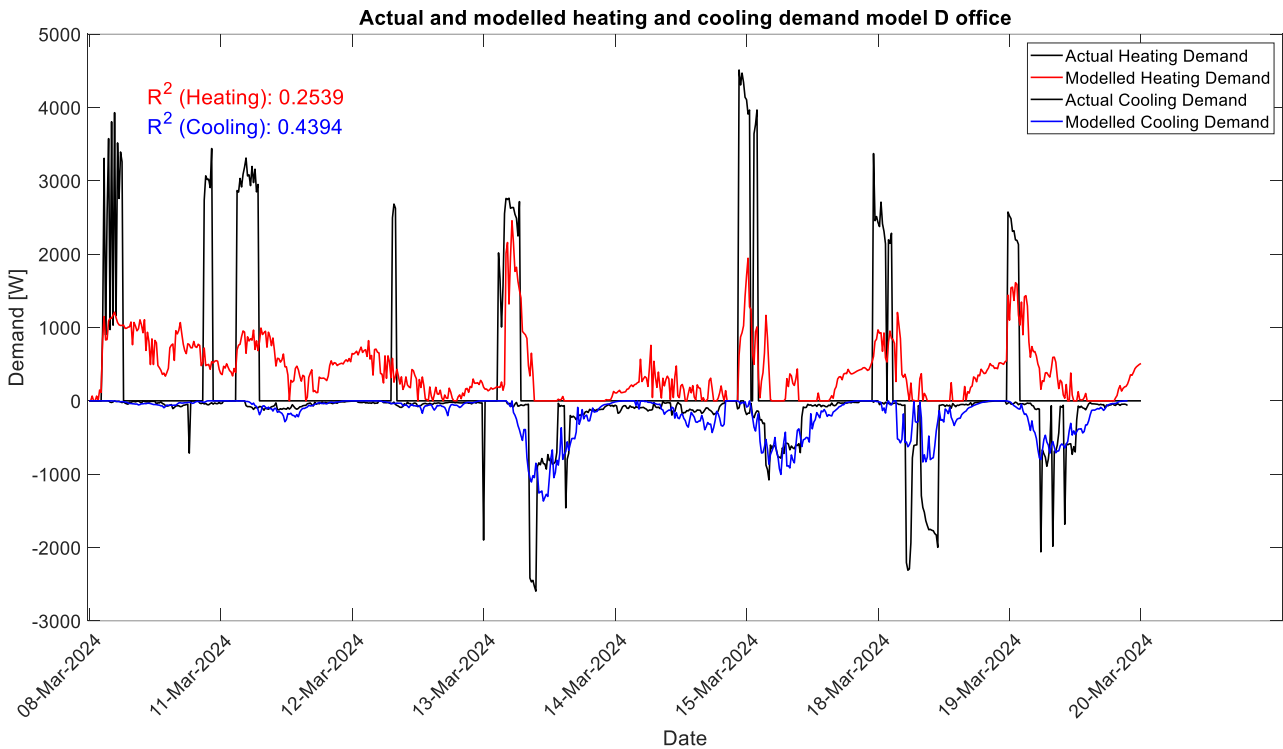|	**Heating**		**Cooling**	
	Modelling	*Prediction*	*Modelling*	*Prediction*
R²	**25.39%**	**-15.47%**	**43.94%**	**-216.03%**
R² adjusted	24.70%	-43.47%	43.60%	-346.80%
RMSE [W]	822.00 (18.11%)	1727.74 (46.75%)	304.00 (12.00%)	34.35 (42.17%)

v

Figure 56: actual and modelled heating and cooling demand by model D for the office from 8 until 20 March 2024
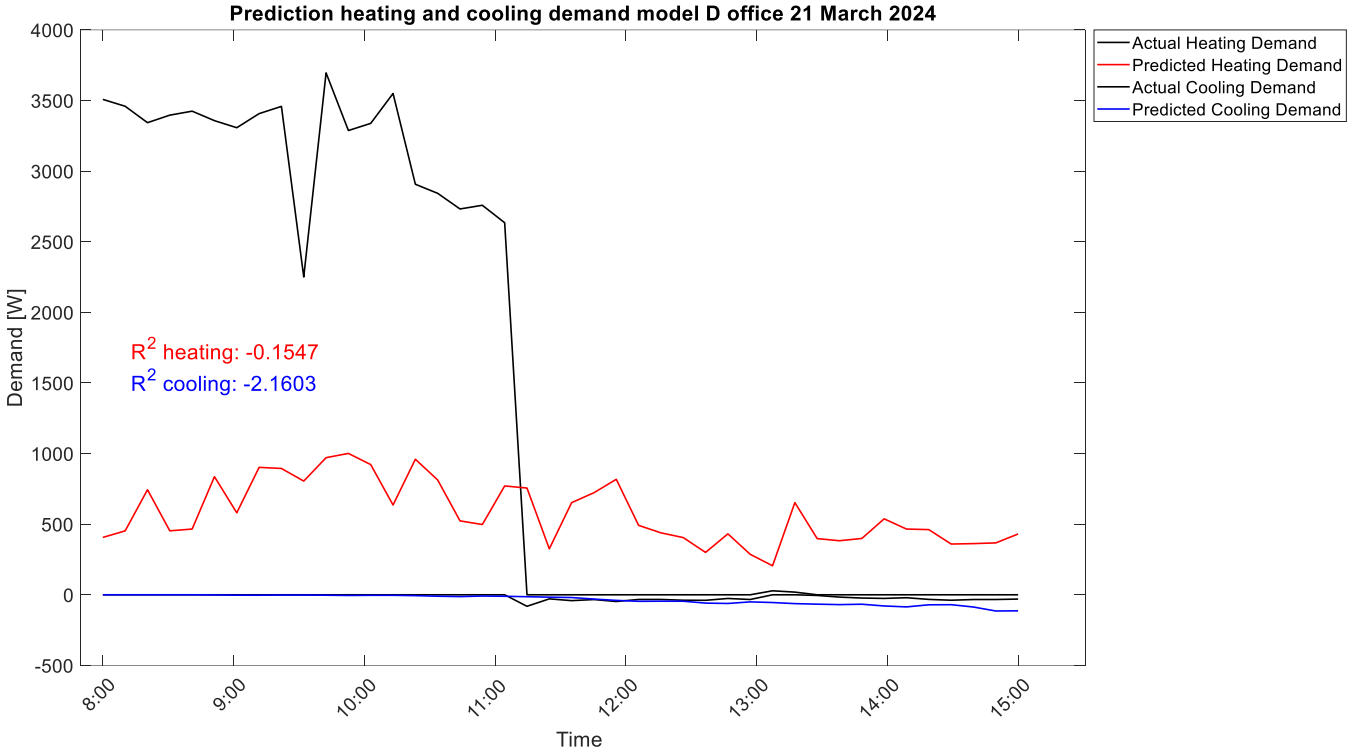
Figure 57: actual and predicted heating and cooling demand by model D for the office on 21 March 2024 8 A.M. - 1 P.M.

Classroom

In comparison to Equations 41 and 42 for heating and cooling, respectively, for the office, Equations 43 and 44 derived from model D are more complex for modelling the heating and cooling demand in the classroom. However, with these independent variables, an R^2 value of 4.40% and 62.84% is achieved, for heating and cooling, respectively. This is slightly higher than the R^2 values of model A. From Figure 46 and Figure 60 for model A, the modelled heating and cooling demand follow the same pattern, however, with model D, the cooling demand is further overestimated on 26 March 2024.

There is a significant difference in the prediction of the cooling demand by model D in comparison to model A, the R_2 value has increased by 7.77%. The predicted cooling demand by model D is lower than model A, which results in achieving the troughs of cooling demand, however, the model is still not able to predict the highest cooling demand.

$$Q_{demand,heat,classroom} = constant + C_6(Q_{internal}(t)) + C_9\left(\left(T_{indoorair} - T_{s,ceiling}\right)(t)\right) +$$
$$C_{10}\left(\left(T_{indoorair} - T_{s,systemwall2}\right)(t)\right) + C_{11}\left(\left(T_{indoorair} - T_{s,concretepillar}\right)(t)\right) +$$
$$C_{12}\left(\left(T_{indoorair} - T_{s,systemwall3}\right)(t)\right) \text{ (43)}$$

112

$$Q_{demand,cold,classroom} = constant + \ C_1\big((T_{outdoor} - T_{indoorair})(t)\big) + C_2\Big(\big(V_{wind} \cdot (T_{outdoor} -$$

$$T_{indoorair})(t)\big) + C_4(Solar\ light\ intensity\ east\ (t)) + \ C_5(Solar\ light\ intensity\ west\ (t)) +$$

$$C_6\ (Q_{internal}(t) + C_8\big((T_{indoorair} - \ T_{s,inside\,of\,the\,outerwall})(t)\big) + C_9((T_{indoorair} -$$

$$T_{s,ceiling})(t)) \ + C_{10}((T_{indoorair} - T_{s,systemwall2})(t)) + C_{11}((T_{indoorair} - T_{s,concretepillar})(t)) +$$

$$C_{12}((T_{indoorair} - T_{s,systemwall3})(t)) \ + \ C_{13}\big(Solar\ light\ intensity\ south(t-n)\big) \ +$$

$$C_{14}\big(Solar\ light\ intensity\ east\ (t-n)\big)\ (44)$$

Table 31: R₂, R₂ adjusted, and RMSE (with a percentage of the maximum heating or cooling demand to the corresponding period) of the modelling and prediction of the heating and cooling demand in the classroom with model D

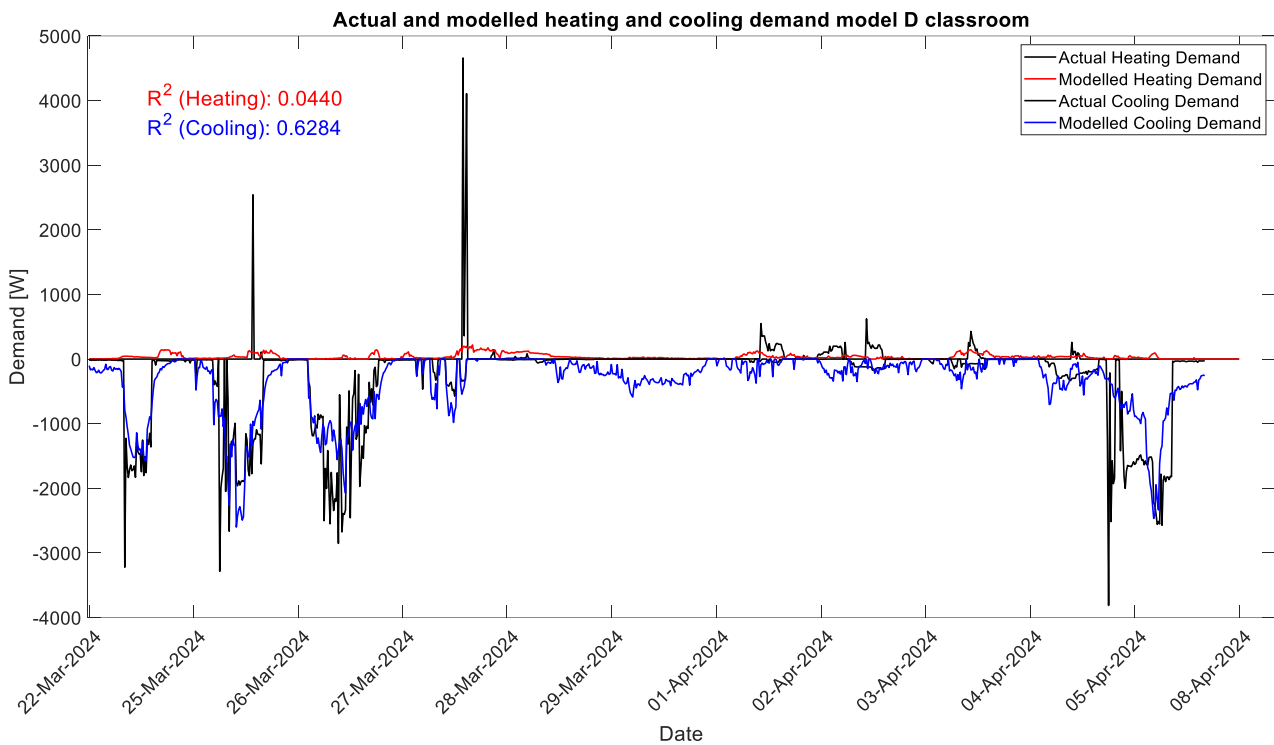|	Heating		Cooling	
	Modelling	*Prediction*	*Modelling*	*Prediction*
R^2	4.40%	-17.46%	62.84%	34.72%
R^2 adjusted	3.09%	-66.06%	62.40%	-2.94%
RMSE [W]	226.00 (4.85%)	40.55 (16.52%)	391.00 (10.26%)	714.36 (22.91%)

Figure 58: actual and modelled heating and cooling demand by model D for the classroom from 22 March until 8 April 2024
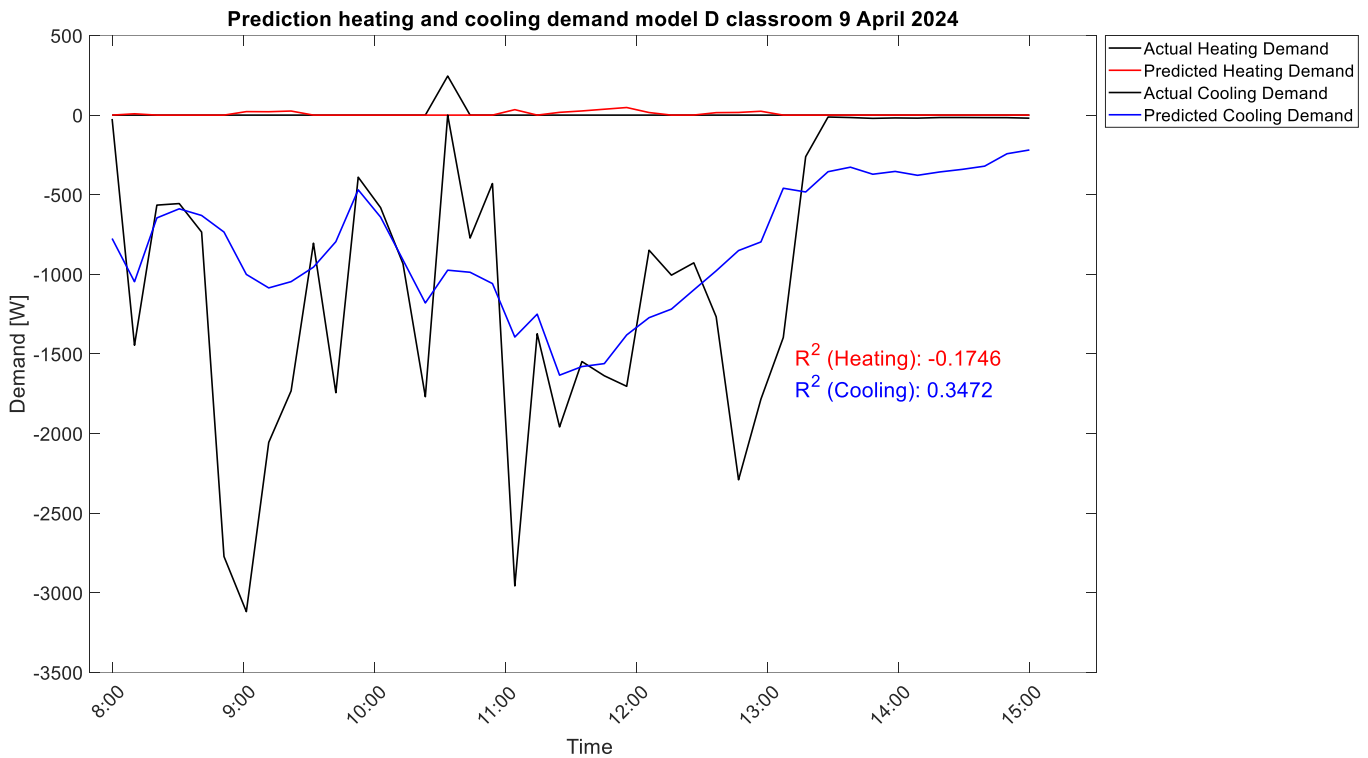
113

Figure 59: actual and predicted heating and cooling demand by model D for the classroom on 9 April 2024 8 A.M. - 1 P.M.

8.2.1 Model D analysis

In this model is in comparison to the other models the temperature difference used based on the Equation 40. A slightly higher accuracy is achieved. However, this was also the first model where all the independent variables based on the thermal energy balance were able to be included in the model. To conclude if the use of temperature difference is needed to achieve higher accuracy, model E is built, with all the independent variables from the thermal energy balance as single data input.

8.3 Model E: using all the independent variables from the thermal energy balance

By only using the outdoor temperature and the indoor surface temperatures (model A), replacing the indoor surface temperature with windspeed, solar light intensities (current and delayed) and the internal heat gains (current and delayed)(model B) and modelling the total thermal demand (model C), the models were not able to model and predict the heating and cooling demand accurately. Model D uses all the available independent variables based on the thermal energy balance, in the form of temperature differences, a slight increase in model accuracy is achieved. To estimate the use of temperature difference instead of using directly the independent surfaces, model E is developed.

114

Office

By using all the available data inputs in model E and the calculated time delays, Equations 45 and 45 are derived (coefficient values can be found in Appendix M), for modelling the heating and cooling demand, respectively. In both Equations the windspeed is not included, this confirms the earlier explanation of the effect of the windspeed together with the outdoor and the indoor temperature as used in model D. In both Equations the solar light intensity is included. The internal heat gains are only included in the heating demand equation. In the office the internal heat gains does not have a significant influence on the cooling demand. The delayed effect of the thermal mass on the heating and cooling demand, has a significant influence on the heating and cooling demand, as the most indoor surface temperatures are included with a delay.

The accuracy of modelling the heating and cooling demand has increased in comparison to models D, for the office, with 14.97 % and 6.41%, respectively. This higher accuracy can be seen in Figure 62 by a higher estimation of the heating and cooling demand, most of the peaks of actual demand are reached. However, this higher accuracy of modelling the heating and cooling demand, does not result in a higher accuracy of predicting the heating and cooling demand. The peak of predicted heating and cooling demand after 13:00, due a significant increase/decrease indoor surface temperatures. By using the temperatures as a single data input, the model is overreacting on these changes in indoor surface temperatures.

$$
\begin{aligned}
Q_{demand,heat,office} = {}& constant + C_1\big(T_{outdoor}(t)\big) + C_3\left(T_{s,window}(t)\right) + C_4\left(T_{s,systemwall1}(t)\right) + \\
& C_6\left(T_{s,ceiling}(t)\right) + C_7\left(T_{s,glasswall}(t)\right) + C_{10}\big(Solar\ light\ intensity\ south\ (t)\big) + \\
& C_{11}\big(Solar\ light\ intensity\ east\ (t)\big) + C_{13}\big(Q_{internal}(t)\big) + C_{15}\big(T_{indoorair}(t-n)\big) + \\
& C_{16}\left(T_{s,window}(t-n)\right) + C_{17}\left(T_{s,systemwall1}(t-n)\right) + C_{18}\left(T_{s,ceiling}(t-n)\right) + \\
& C_{19}\left(T_{s,glasswall}(t-n)\right) + C_{20}\,(T_{s,systemwall2}\,(t-n))\ (45)
\end{aligned}
$$

$$
\begin{aligned}
Q_{demand,cold,office} = {}& constant + C_2\big(T_{indoorair}(t)\big) + C_3\left(T_{s,window}(t)\right) + C_7\left(T_{s,glasswall}(t)\right) + \\
& C_{10}\big(Solar\ light\ intensity\ south\ (t)\big) + C_{12}\big(Solar\ light\ intensity\ west\ (t)\big) + C_{16}\left(T_{s,window}(t-n)\right) + \\
& C_{19}\left(T_{s,glasswall}(t-n)\right) + C_{20}\left(T_{s,systemwall2}\,(t-n)\right) + \\
& C_{22}\big(Solar\ light\ intensity\ south\ (t-n)\big)\ (46)
\end{aligned}
$$

Table 32: R_2, R_2 adjusted, and RMSE (with a percentage of the maximum heating or cooling demand to the corresponding period) of the modelling and prediction of the heating and cooling demand in the office with model E

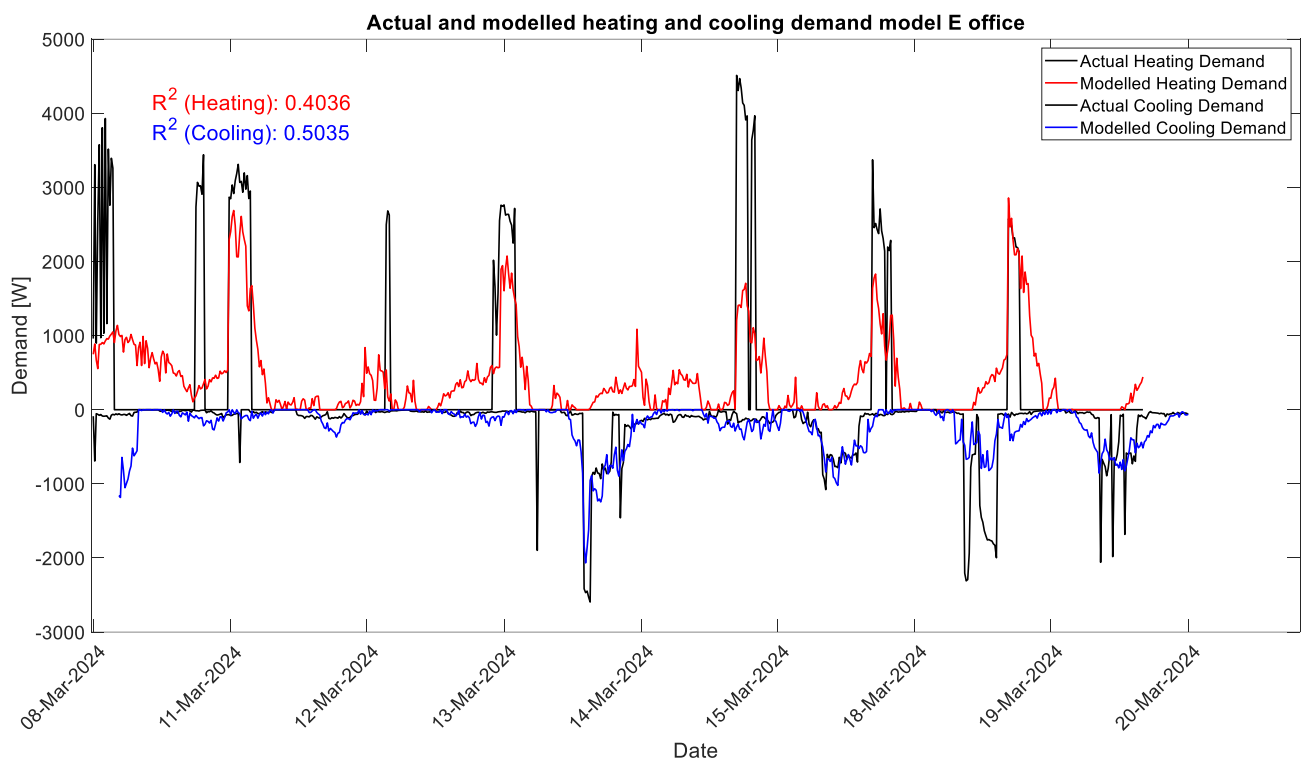|	Heating		Cooling	
	Modelling	*Prediction*	*Modelling*	*Prediction*
R^2	40.36%	-13.49%	50.35%	-14642.67%
R^2 adjusted	39.10%	-121.57%	49.7%	-37678.08%
RMSE [W]	743 (16.37%)	1712.80 (46.34%)	287 (11.32%)	234.69 (288.10%)

Figure 60: actual and modelled heating and cooling demand by model E for the office from 8 until 20 March 2024
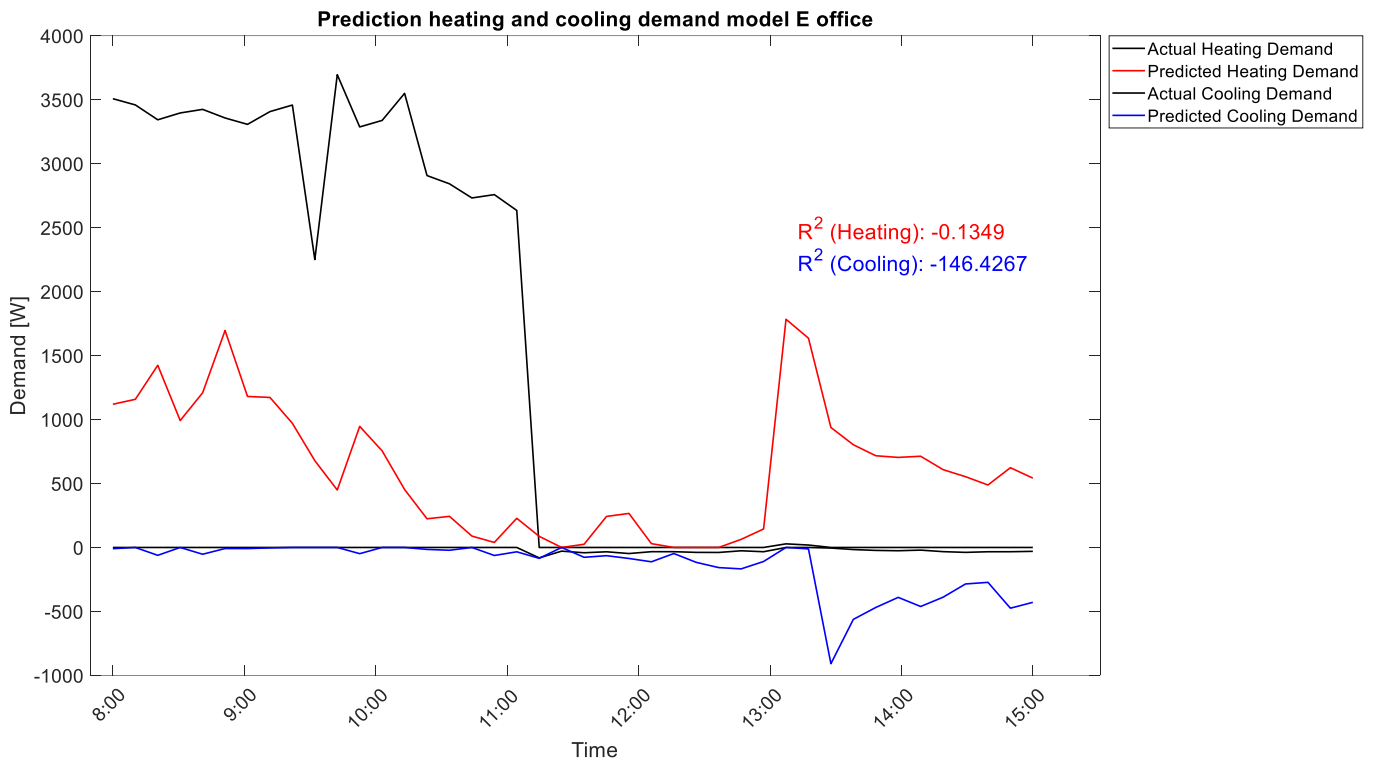
Figure 61: actual and predicted heating and cooling demand by model E for the office on 21 March 2024 8 A.M. - 1 P.M.

Classroom

For the heating demand in the classroom there are no calculated time delays of the independent variables, this makes Equation 47 static. There were calculated delays for the cooling demand, see Appendix M, however no time delayed indoor surface temperature is included in the Equation 48 derived from model E for the classroom, this was also the case for model D. As seen in Chapter 6.1, in the classroom there is relatively compared to the cooling demand, almost no heating demand expect one peak. In other words there is less heat to be stored in the thermal mass of the room. By including the internal heat gains, the solar light intensities, and the windspeed and making the model dynamic for the cooling demand (Equation 48) the accuracy of modelling the heating and cooling demand, has increased with 0.41% and 6.08%, for heating and cooling, respectively, compared to model A, where only the outdoor and indoor surface temperatures were included. The model does not included the indoor air temperature, despite that this was available to use. In comparison to the office, the effect of including these extra independent variables, the increase in accuracy is small. This is due the overall relatively small coefficients (see Appendix M) of these independent variables in Equation 47. Due to the orientation to the courtyard, the outdoor parameters have less influence on the heating and cooling demand than the classroom. In comparison to model D, the accuracy has slightly increased with 0.01% and 1.24%, for heating and cooling respectively.

In this case is model E is better in predicting the heating and cooling demand than model D, with an increase of 0.39% and 0.62% of the R_2 value, for heating and cooling, respectively.

$$Q_{demand,heat,classroom} = constant + C_6\left(T_{s,concretepillar}\left(t\right)\right) + C_7\left(T_{s,systemwall3}(t)\right) +$$

$$C_8\left(T_{s,ceiling}\left(t\right)\right) + C_9\left(V_{wind}(t)\right) + C_{13}(Q_{internal}(t)) \text{ (47)}$$

$$Q_{demand,heat,classroom} = constant + C_3\left(T_{s,insideoftheouterwall}(t)\right) + C_4\left(T_{s,systemwall1}(t)\right) +$$

$$C_6\left(T_{s,concretepillar}\left(t\right)\right) + C_7\left(T_{s,systemwall3}(t)\right) + C_8\left(T_{s,ceiling}\left(t\right)\right) + C_9\left(V_{wind}(t)\right) +$$

$$C_{11}(Solar\ light\ intensity\ east\ (t)) + C_{13}(Q_{internal}(t)) + C_{14}(T_{outdoor}(t-n)) +$$

$$C_{19}(Solar\ light\ intensity\ south\ (t-n)) + C_{20}(Solar\ light\ intensity\ east\ (t-n)) \text{ (48)}$$

Table 33: R_2, R_2 adjusted, and RMSE (with a percentage of the maximum heating or cooling demand to the corresponding period) of the modelling and prediction of the heating and cooling demand in the classroom with model E

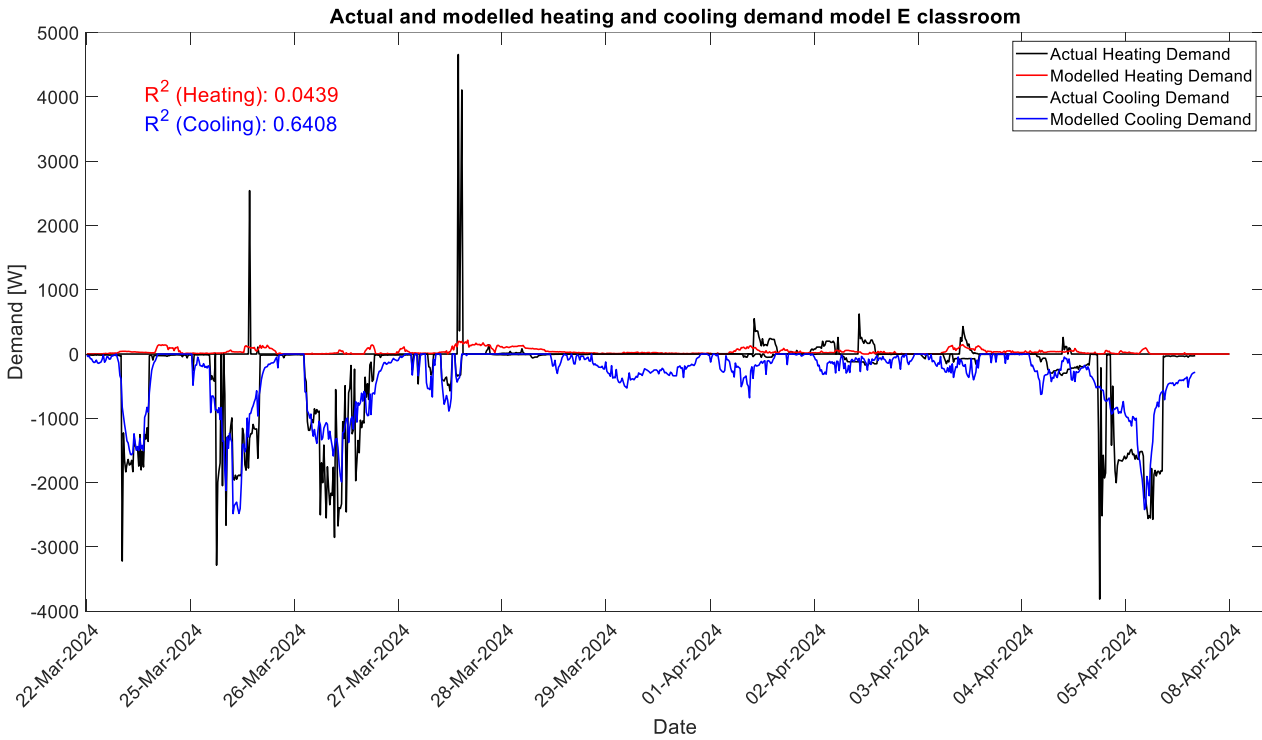|	**Heating**		**Cooling**	
	Modelling	*Prediction*	*Modelling*	*Prediction*
R²	**4.39%**	**-17.85%**	**64.08%**	**35.34%**
R² adjusted	3.90%	-72.57%	63.70%	-32.56%
RMSE [W]	226 (4.85%)	40.62 (16.55%)	384 (10.07%)	710.96 (22.79%)

Figure 62: actual and modelled heating and cooling demand by model E for the classroom from 22 March until 8 April 2024
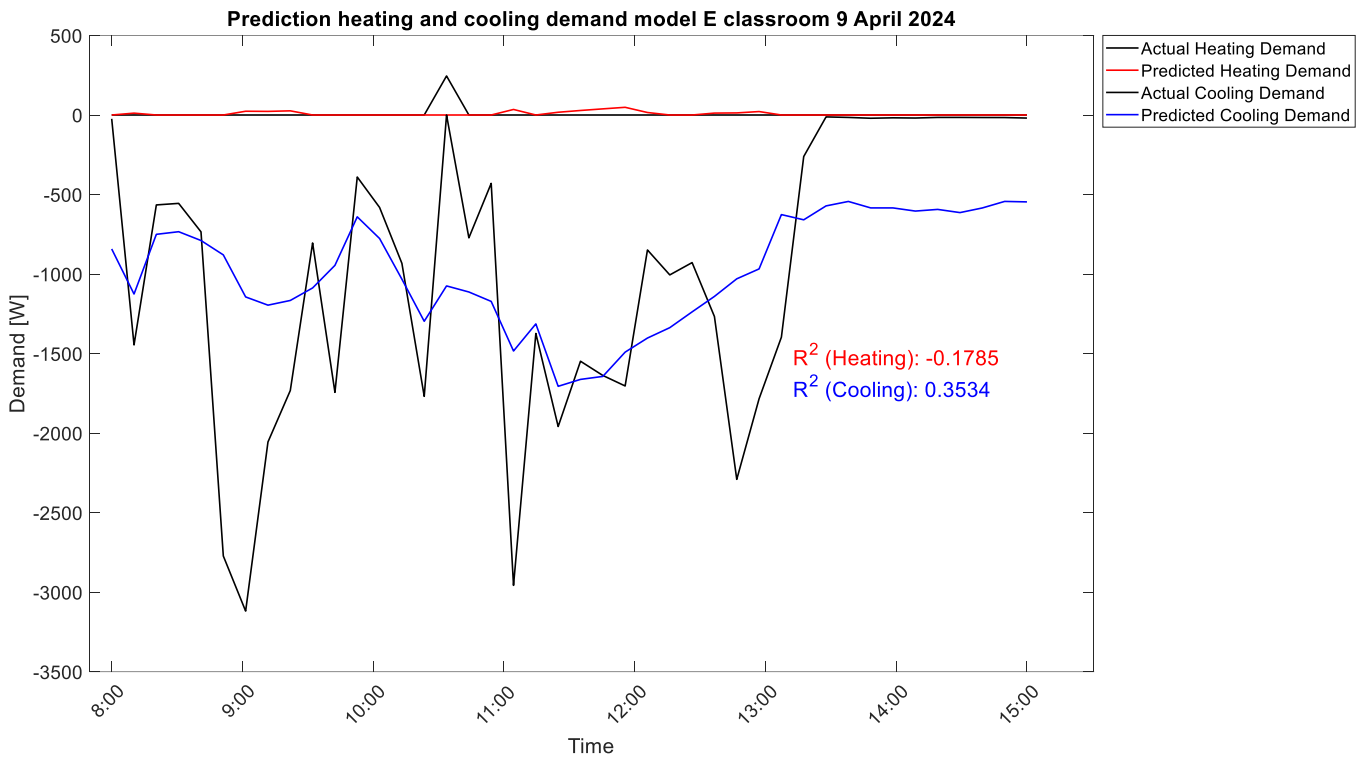
Figure 63: actual and predicted heating and cooling demand by model E for the classroom on 9 April 2024 8 A.M. - 1 P.M

8.3.1 Model E analysis

Within model E all the independent variables based on the thermal energy balance are used as a single data input. The accuracy has increased for the heating and cooling demand for both rooms in comparison to model D where temperature differences were used. This effect is the strongest for the office, where also the inclusion of more independent variables: internal heat gains, windspeed, solar light intensity and indoor air temperature has a significantly increasing effect on the accuracy of modelling compared to a small increase in the accuracy of modelling for the classroom. Overall, for both rooms, the coefficients of the indoor surface temperatures are the highest, which shows that the indoor surface temperatures have a significant influence on the heating and cooling demand. In the classroom there are no delays included in these indoor surface temperatures, this is explained by the fact that there is almost no heating demand in this room, in other words, less heat is stored and released with a time delay by the thermal mass. Thereby is the classroom less exposed to outdoor weather conditions than the office, as the classroom is oriented to the courtyard.

Using temperature differences in model D results in a lower accuracy of modelling than using the independent variables as a single input. This reduction in accuracy can be attributed to several factors. First, there is the risk of information losses, by using temperature differences, as several independent variables are captured in a few independent variables. Thereby the

multivariate linear regression model assumes a linear relationship between the dependent variables and the independent variables. This relation can be weakened when the differences between these independent variables are used. For practical use, it is easier to use individual independent variables, as it is easier to implement and understand.

However, the use of the temperature differences has in the case of the office an advantage for predicting the thermal demand. Using the temperature differences is more robust to process relatively high increases or decreases in the values of the independent variables. In the office, there is a significant increase in several indoor surface temperatures which results in an overreacting model when using the individual independent variables. This is caused by multicollinearity between the independent variables, this effect is weakened by using temperature differences as the indoor air temperature and the indoor surface temperature are combined. This is confirmed by using more independent variables, the use of all the independent variables based on the thermal energy balance, the predicted peak during the prediction period (where there is no actual thermal demand), is higher when introducing more independent variables.

8.4 Model F: selecting independent variables based on the correlation matrix

8.4.1 Model F_1: using most influencing independent variables

By including more independent variables, the model becomes more complex and more data points are needed, which makes the model harder to implement. A way to select the significant independent variables is using the correlation matrixes. Before developing the model, a study can be conducted in a building to measure for a period (during different thermal periods; heating/cooling, heating and cooling), the different independent variables, to select which independent variables are needed to model and predict the heating or cooling demand instead of on the thermal energy balance. However, in this study, the independent variables are selected based on the thermal energy balance, which makes this a grey-box modelling method and not a black-box model. The independent variables are selected when the Pearson correlation coefficient is above 0.3, as this is the minimal value for a moderate correlation with the heating or cooling demand. For these independent variables, a time delay is calculated with the cross-correlation function in MATLAB, see Appendix N for the coefficient values and the calculated delays.

Office

The selected independent variables to model and predict the heating demand are the indoor air temperature and surface temperatures of system wall 1, ceiling and glass wall, see Chapter 6 for a further explanation of the correlations between the independent variables and the heating demand. All the independent variables are included in Equation 48. Only the current indoor surface temperature of the ceiling is included. The Equation is further built up with delayed independent variables. For all the independent variables there was a calculated delay.

The R_2 value of modelling the heating demand in the office is 29.90%. This is a decrease in accuracy compared to model E of 10.46%, and an increase in accuracy compared to models A (+ 12.14%), B (+ 5.18%) and D (+4.51%). In comparison to model E, the modelled heating demand underestimation the actual heating demand. With using only the moderate and strongly correlated independent variables to the heating demand, there is a need to include more independent variables to increase the accuracy of modelling the heating demand. However, the expectation is that the accuracy of the prediction of the heating demand will increase as there are fewer independent variables, which makes the model less complex and sensitive to multicollinearity, which results in an overfitting model. Model F1 is still overestimating the heating demand after 1 P.M., however, this peak is for a shorter period.

$$Q_{demand,heat,office} = C_0 + C_2\left(T_{s,ceiling}(t)\right) + C_9\left(T_{s,systemwall1}(t-n)\right) + C_{10}\left(T_{s,ceiling}(t-n)\right) + C_{11}\left(T_{s,glasswall}(t-n)\right) + C_{13}(T_{indoorair}(t-n)) \ (49)$$

The selected independent variables to model and predict the cooling demand are the solar light intensities in the orientations south and west, indoor air temperature, outdoor temperature, surface temperatures of the window, system wall 1, ceiling and glass wall. In the case of the cooling demand, not all selected independent variables are included in Equation 49. The outdoor and indoor air temperatures are not included, despite their moderate correlation with the cooling demand. The R_2 value of modelling the cooling demand is 39.39%. This is a decrease in accuracy compared to model B (- 4.14%), D (- 4.55%) and E (- 10.96%), and an increase in accuracy compared to model A (+ 4.69%). Also in the case of cooling demand, it is not sufficient to include only the moderate and strongly correlated independent variables. However, also in this case, including fewer independent variables, compared to model E, results in better-predicting results, however, this model is not able to predict the demand accurately

$$Q_{demand,cold,office} = constant + C_3\left(T_{s,glasswall}(t)\right) + C_4\left(T_{s,window}(t)\right) +$$

$$C_8(Solar\ light\ intensity\ south\ (t)) + C_{10}\left(T_{s,ceiling}(t-n)\right) + C_{12}\left(T_{s,window}(t-n)\right) +$$

$$C_{14}(T_{outdoor}(t-n)) + C_{15}(Solar\ light\ intensity\ west\ (t-n))\ (50)$$

Table 34: R_2, R_2 adjusted, and RMSE (with a percentage of the maximum heating or cooling demand to the corresponding period) of the modelling and prediction of the heating and cooling demand in the office with model F_1

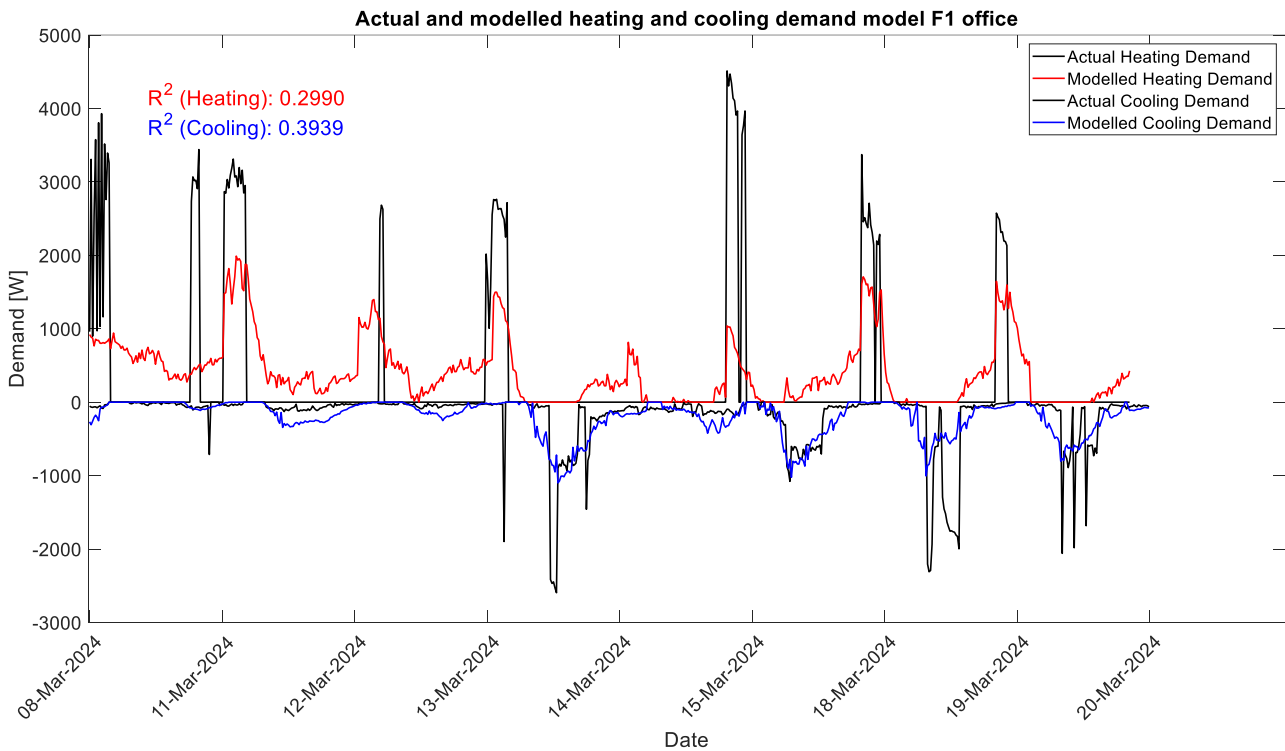|	**Heating**		**Cooling**	
	Modelling	*Prediction*	*Modelling*	*Prediction*
R²	**29.90%**	**-20.95%**	**39.39%**	**-345.05%**
R² adjusted	29.40%	-50.28%	38.80%	-629.88%
RMSE [W]	800.00 (17.63%)	1768.30 (47.85%)	314.00 (12.39%)	40.76 (50.03%)

Figure 64: actual and modelled heating and cooling demand by model F_1 for the office from 8 until 20 March 2024
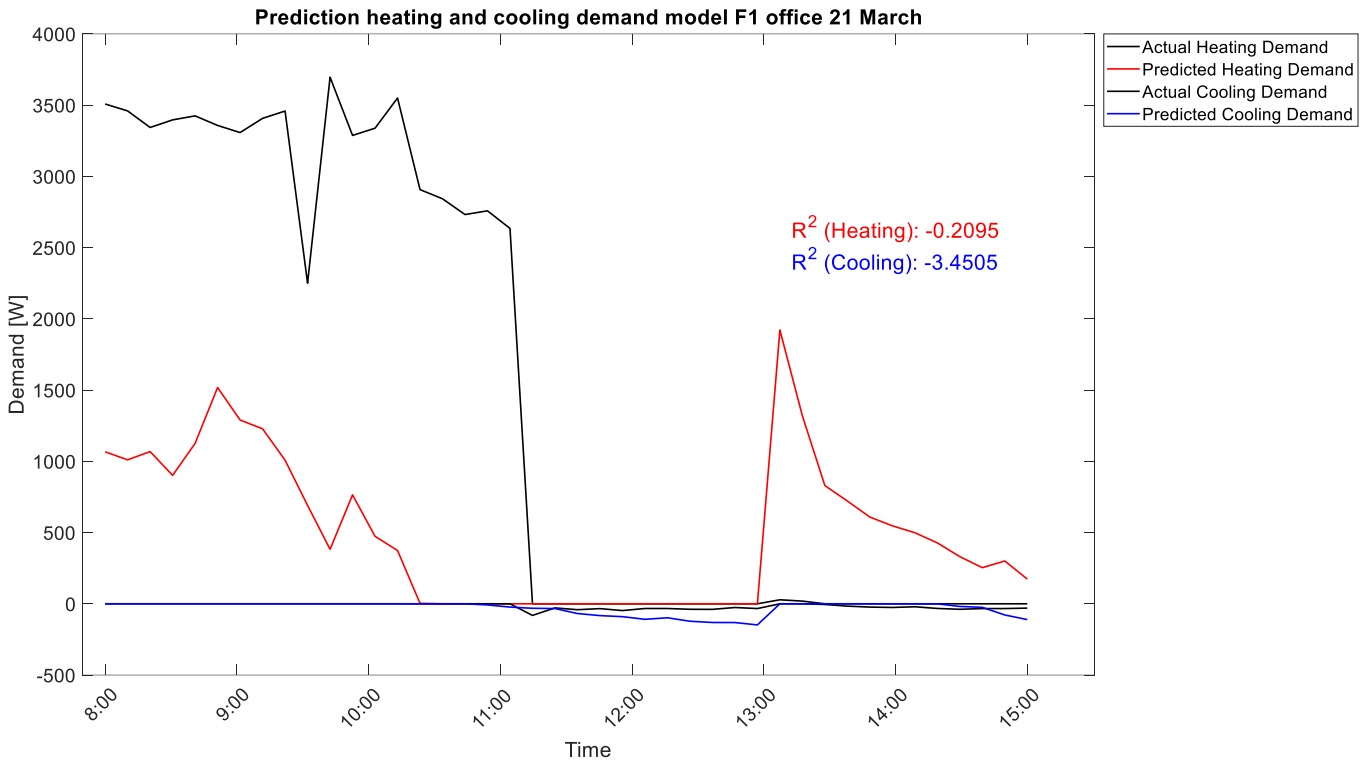
Figure 65: actual and predicted heating and cooling demand by model F₁ for the office on 21 March 2024 8 A.M. - 1 P.M.

Classroom

There are no moderate or strong correlations between independent variables and the heating demand in the classroom, no model is built. The selected independent variables for the cooling demand in the classroom are the internal heat gains and the indoor surface temperatures of the inside of the outer wall, system wall 2, concrete pillar and system wall 3. All the independent variables are included in Equation 51 to model the cooling demand.

The R_2 value for modelling the cooling demand in the classroom is 48.36%. This represents a decrease in accuracy compared to models B (-9.64%), D (-14.48%), and E (-15.72%), and an increase in accuracy compared to model B (+3.45%). The improvement in accuracy compared to model B, where indoor surface temperature is excluded, can be connected to the significant influence of indoor surface temperatures on heating and cooling demand. However, with the inclusion of moderately and strongly correlated independent variables, the highest accuracy in predicting cooling demand, with an R_2 value of 36.60% is achieved.

$$Q_{demand,cold,classroom} = C_0 + C_2\left(T_{s,systemwall2}(t)\right) + C_3\left(T_{s,concretepillar}(t)\right) +$$
$$C_5(Q_{internal}(t)) + C_6\left(T_{s,insideoftheouterwall}(t-n)\right) + C_8\left(T_{s,concretepillar}(t-n)\right) +$$
$$C_9(T_{s,systemwall3}(t-n))\ (51)$$

123

Table 35: R_2, R_2 adjusted, and RMSE (with a percentage of the maximum heating or cooling demand to the corresponding period) of the modelling and prediction of the heating and cooling demand in the classroom with model F_1

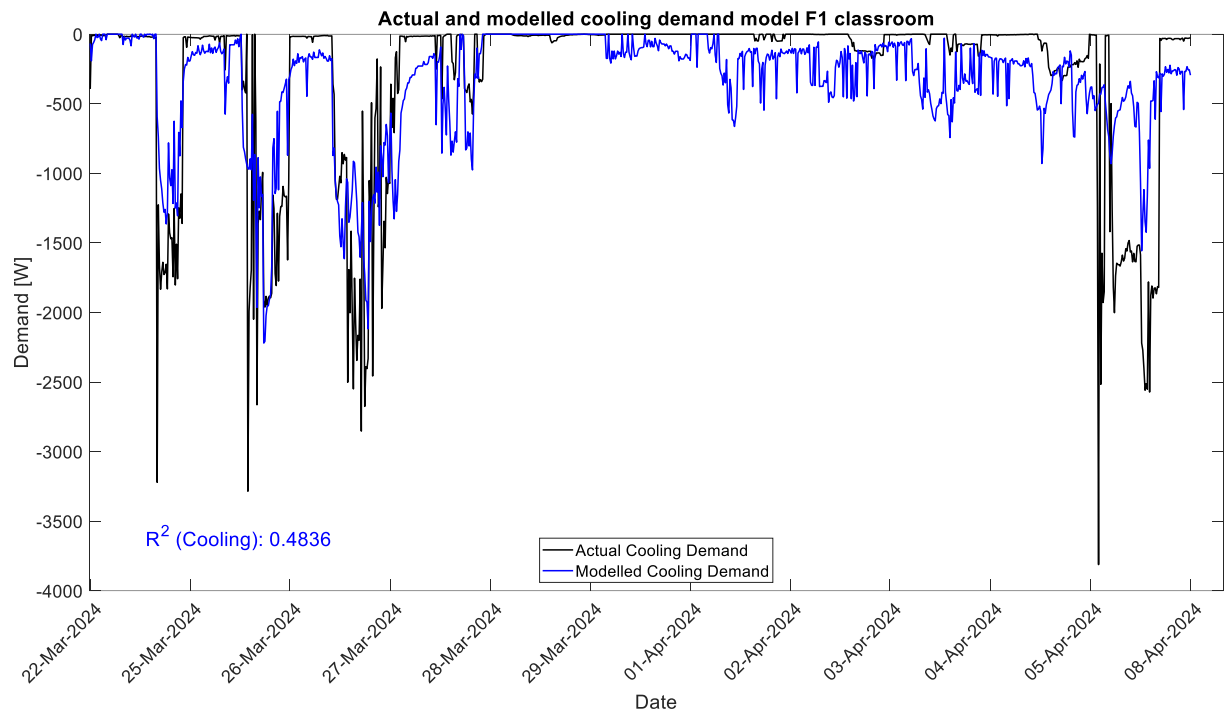|	Heating		Cooling	
	Modelling	*Prediction*	*Modelling*	*Prediction*
R²	**-**	**-**	**48.36%**	**36.60%**
R² adjusted	-	-	48.00%	16.15%
RMSE [W]	-	-	453.00 (11.88%)	703.99 (22.58%)

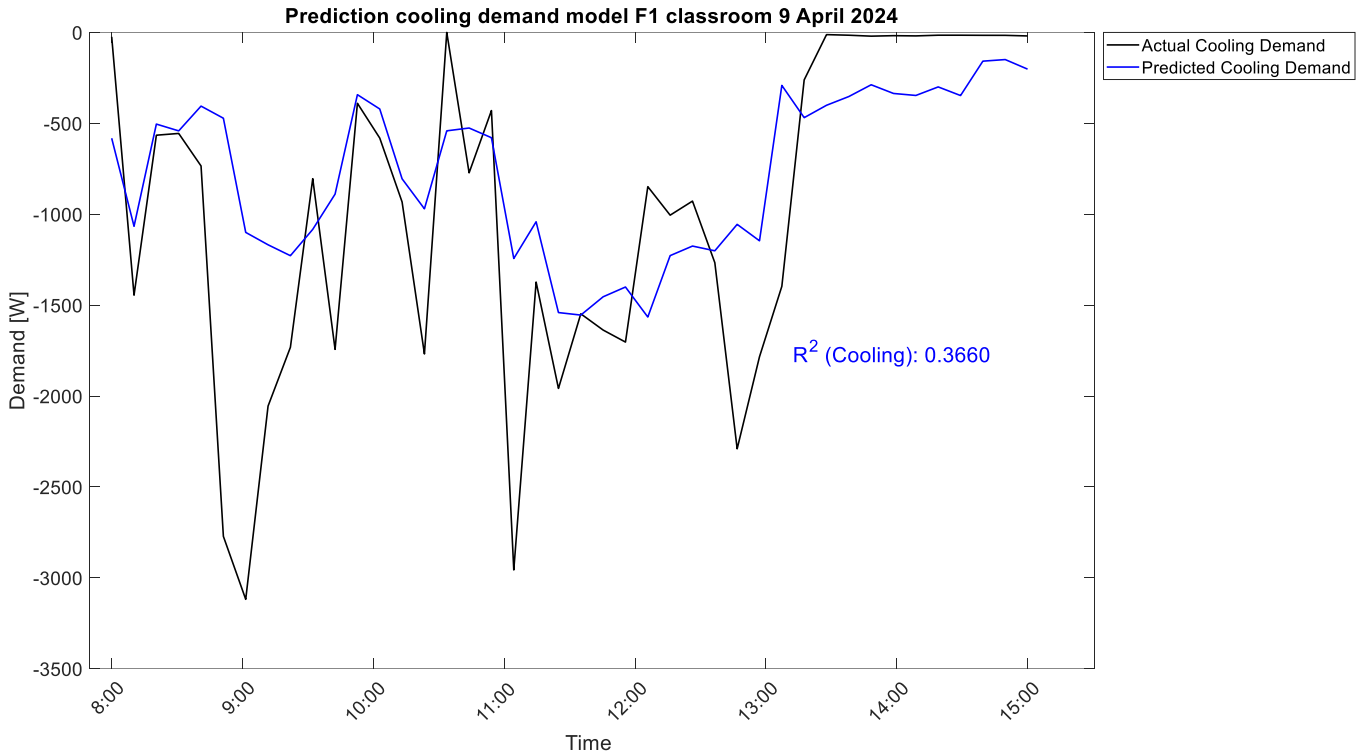*Figure 66: actual and modelled cooling demand by model F_1 for the classroom from 22 March until 8 April 2024*

Figure 67: actual and predicted cooling demand by model F_1 for the classroom on 9 April 2024 8 A.M. - 1 P.M

8.4.2 Model F_2: allowing interactions between independent variables

From the correlation matrixes, it is concluded that there is also a strong correlation between independent variables and the heating and cooling demand. As independent variables input for this model the independent variables are selected, who have a moderate and strong correlation (Person-correlation coefficient > 0.3) with another independent variable and the heating or cooling demand. In Appendix O the coefficient values, the corresponding statistical values and the delays are shown.

Office

The selected independent variables for the heating demand in the office for model F_2 are; indoor air temperature, outdoor temperature, solar light intensities in orientations west, east and south and the indoor surface temperatures of the window, system wall 1, inside of the outer wall, ceiling and glass wall. Equation 52 is derived from model F_2 for modelling the heating demand in the office. All the selected independent variables are included in the Equation. By allowing these interactions and selecting the independent variables strongly correlated with heating demand, the R_2 value increased to 69.71%, providing the highest accuracy for heating demand modelling. However, while the model effectively captures heating demand peaks, it sometimes overestimates the demand. Model F_2 performs better in capturing periods of no

125

actual heating demand compared to other models, but still shows some demand during these periods. Despite the improved R_2 value, the increased accuracy in modelling the heating demand does not translate to better prediction. Model F_2 predicts no heating demand after 9 A.M., whereas the other models predict a too-low demand in the morning, and it shows a very high peak after 1 P.M. Including interactions between independent variables results in complex Equation 52, leading to overfitting due to multicollinearity. This causes the model to overreact to small changes in independent variable values and makes it unable to predict heating demand accurately with new data.

$$Q_{demand,heat,office} = C_0 + C_1(Solar\ light\ intensity\ west\ (t)) + C_2\big(T_{s,ceiling}(t-n)\big) +$$
$$C_3\big(T_{s,insideoftheouterwall}(t) \cdot T_{s,glasswall}(t)\big) + C_4\big(T_{s,window}(t)/T_{indoorair}(t)\big) + C_5\Big(T_{s,systemwall1}(t)/$$
$$T_{indoorair}(t)\Big) + C_6\big(T_{s,insideoftheouterwall}(t)/T_{outdoor}(t)\big) + C_7\big(T_{s,ceiling}(t)/T_{outdoor}(t)\big) +$$
$$C_8\big(T_{s,systemwall2}(t)/Solar\ light\ intensity\ west\ (t)\big) + C_9\big(T_{s,window}(t) \cdot T_{s,window}(t-n)\big) +$$
$$C_{10}\big(T_{s,insideoftheouterwall}(t) \cdot T_{s,window}(t-n)\big) + C_{11}\big(T_{s,window}(t) \cdot T_{s,systemwall1}(t-n)\big) +$$
$$C_{12}\big(T_{s,systemwall2}(t)/T_{s,systemwall1}(t-n)\big) + C_{13}\big(T_{s,systemwall1}(t)/T_{s,ceiling}(t-n)\big) + C_{14}\big(T_{s,glasswall}(t) \cdot$$
$$T_{s,glasswall}(t-n)\big) + C_{15}\Big(Solar\ light\ intensity\ east\ (t)/T_{s,glasswall}(t-n)\Big) + C_{16}\big(T_{outdoor}(t)/$$
$$T_{s,systemwall2}(t-n)\big) + C_{17}\big(T_{s,window}(t-n) \cdot T_{s,systemwall2}(t-n)\big) + C_{18}\big(T_{s,systemwall1}(t-n) \cdot$$
$$T_{s,systemwall2}(t-n)\big) + C_{19}(T_{outdoor}(t)/T_{indoor}(t-n)) + C_{20}(Solar\ light\ intensity\ east\ (t)/T_{indoor}(t-n))$$
$$n)) + C_{21}\big(T_{s,window}(t) \cdot T_{outdoor}(t-n)\big) + C_{22}\big(T_{s,ceiling}(t) \cdot T_{outdoor}(t-n)\big) + C_{23}\big(T_{s,systemwall2}(t)/$$
$$T_{outdoor}(t-n)\big) + C_{24}(T_{outdoor}(t)/T_{outdoor}(t-n)) + C_{25}\big(T_{s,systemwall2}(t-n) \cdot T_{outdoor}(t-n)\big)\ (52)$$

The selected independent variables for the cooling demand in the office for model F_2 are; indoor air temperature, outdoor temperature, solar light intensities in orientations west, east and south, internal heat gains and the indoor surface temperatures of the window, system wall 1, inside of the outer wall, ceiling, glass wall and system wall 2. Equation 53 derived from model F_2 to represent the cooling demand in the office is more complex than the heating Equation 52 in this room. It involves highly intercorrelated independent variables. Despite its complexity, it achieved a high R_2 value of 78.57% and a low RMSE value. The Equation accurately models the peaks and troughs of actual cooling demand. However, similar to the heating demand, model F2 is unable to predict the cooling demand due to the same reasons.

$$Q_{demand,cold,office} = constant + C_1\Big(T_{s,glasswall}(t)\Big) + C_2\big(Solar\ light\ intensity\ east(t)\big) +$$
$$C_3\big(Solar\ light\ intensity\ west(t-n)\big) + C_4\Big(T_{s,window}(t)/Solar\ light\ intensity\ west(t)\Big) +$$
$$C_5\Big(T_{s,systemwall1}(t) \cdot Solar\ light\ intensity\ west(t)\Big) + C_6\Big(T_{s,insideoftheouterwall}(t) \cdot$$
$$Solar\ light\ intensity\ west(t)\Big) + C_7\Big(T_{s,ceiling}(t)/Solar\ light\ intensity\ west(t)\Big) + C_8\Big(T_{s,window}(t)/$$
$$Solar\ light\ intensity\ east(t)\Big) + C_9\Big(T_{s,window}(t)/Solar\ light\ intensity\ south(t)\Big) +$$
$$C_{10}\Big(T_{s,systemwall1}(t) \cdot Solar\ light\ intensity\ south(t)\Big) + C_{11}\big(Solar\ light\ intensity\ west(t) \cdot$$
$$Solar\ light\ intensity\ south(t)\big) + C_{12}\big(T_{s,systemwall2}(t)/Q_{internal}(t)\big) + C_{13}\Big(T_{s,systemwall2}(t)/T_{s,window}(t-$$
$$n)\Big) + C_{14}\Big(Solar\ light\ intensity\ south(t) \cdot T_{s,window}(t-n)\Big) + C_{15}\Big(T_{s,insideoftheouterwall}(t) \cdot$$
$$T_{s,ceiling}(t-n)\Big) + C_{16}\Big(Solar\ light\ intensity\ west(t) \cdot T_{s,ceiling}(t-n)\Big) + C_{17}\Big(T_{s,systemwall2}(t)/$$

$$T_{s,systemwall2}\left(t-n\right)) + C_{18}\left(Q_{internal}(t)/T_{s,systemwall2}\left(t-n\right)\right) + C_{19}(Solar\ light\ intensity\ west(t)/$$

$$T_{indoor}(t-n)) + C_{20}(Solar\ light\ intensity\ west(t) \cdot T_{outdoor}(t-n)) +$$

$$C_{21}(Solar\ light\ intensity\ south(t) \cdot T_{outdoor}(t-n)) + C_{22}\left(T_{s,ceiling}\ (t)/Solar\ light\ intensity\ west(t-n)\right) + C_{23}\left(Q_{internal}(t)/Solar\ light\ intensity\ west(t-n)\right) + C_{24}\left(T_{s,window}\ (t)/\right.$$

$$Solar\ light\ intensity\ south(t-n)) + C_{25}\left(T_{s,systemwall2}\ (t)/Solar\ light\ intensity\ south(t-n)\right) +$$

$$C_{26}(T_{indoor}(t)/Solar\ light\ intensity\ south(t-n)) + C_{27}(Solar\ light\ intensity\ west(t) \cdot Solar\ light\ intensity\ south(t-n)) + C_{28}(Solar\ light\ intensity\ south(t) \cdot$$

$$Solar\ light\ intensity\ south(t-n)) + C_{29}\left(T_{s,systemwall2}\ (t-n) \cdot Solar\ light\ intensity\ south(t-n)\right) +$$

$$C_{30}(T_{indoorair}(t-n)/Solar\ light\ intensity\ south(t-n)) + C_{31}\left(T_{s,window}\ (t)/\right.$$

$$Solar\ light\ intensity\ south(t-n)) + C_{32}\left(T_{s,systemwall1}\ (t)/Solar\ light\ intensity\ south(t-n)\right) +$$

$$C_{33}\left(T_{s,inside\ of\ the\ outerwall}\ (t)/Solar\ light\ intensity\ south(t-n)\right) + C_{34}\left(T_{s,glasswall}\ (t)\ /\right.$$

$$Solar\ light\ intensity\ south(t-n)) + C_{35}\left(T_{s,systemwall2}\ (t)/Solar\ light\ intensity\ south(t-n)\right) +$$

$$C_{36}(Solar\ light\ intensity\ west(t)/Solar\ light\ intensity\ south(t-n)) +$$

$$C_{37}(Solar\ light\ intensity\ east(t)/Solar\ light\ intensity\ south(t-n)) + C_{38}\left(T_{s,window}\ (t)/Q_{internal}(t-n)\right) + C_{39}\left(T_{s,ceiling}\ (t)/Q_{internal}(t-n)\right) + C_{40}(Solar\ light\ intensity\ east(t)/Q_{internal}(t-n)) +$$

$$C_{41}(Q_{internal}(t)/Q_{internal}(t-n)) + C_{42}(T_{s,window}\ (t-n)/Q_{internal}(t-n)) + C_{43}(T_{s,glasswall}\ (t-n)/Q_{internal}(t-n)) + C_{44}(Solar\ light\ intensity\ south(t)/Q_{internal}(t-n))\ (53)$$

Table 36: R_2, R_2 adjusted, and RMSE (with a percentage of the maximum heating or cooling demand to the corresponding period) of the modelling and prediction of the heating and cooling demand in the office with model F_2

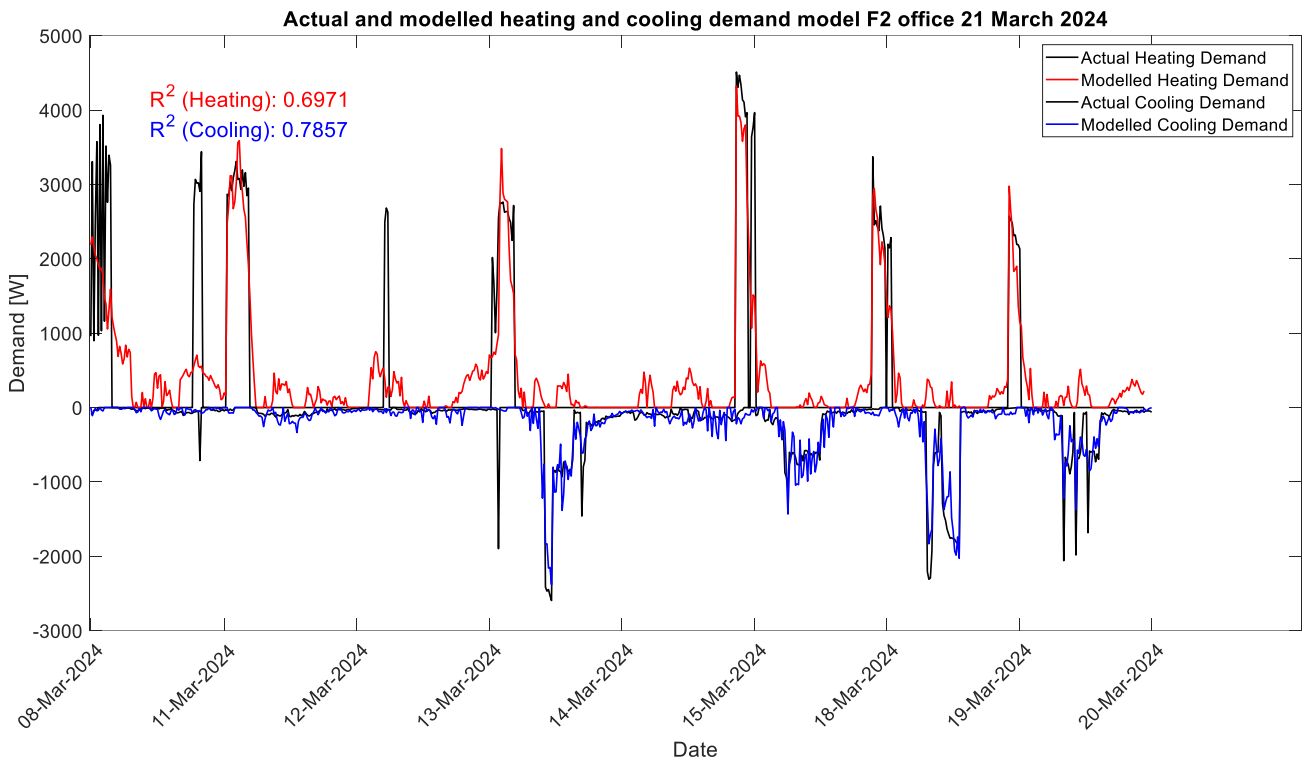|	Heating		Cooling	
	Modelling	*Prediction*	*Modelling*	*Prediction*
R²	**69.71%**	**-190.07%**	**78.57%**	**-70238.06%**
R² adjusted	68.10%	-417.08%	76.80%	-106.11e-5
RMSE [W]	538.00 (11.85%)	2738.40 (74.09%)	194.00 (7.65%)	512.47 (629.18%)

Figure 68: actual and modelled heating and cooling demand by model F_2 for the office from 8 until 20 March 2024
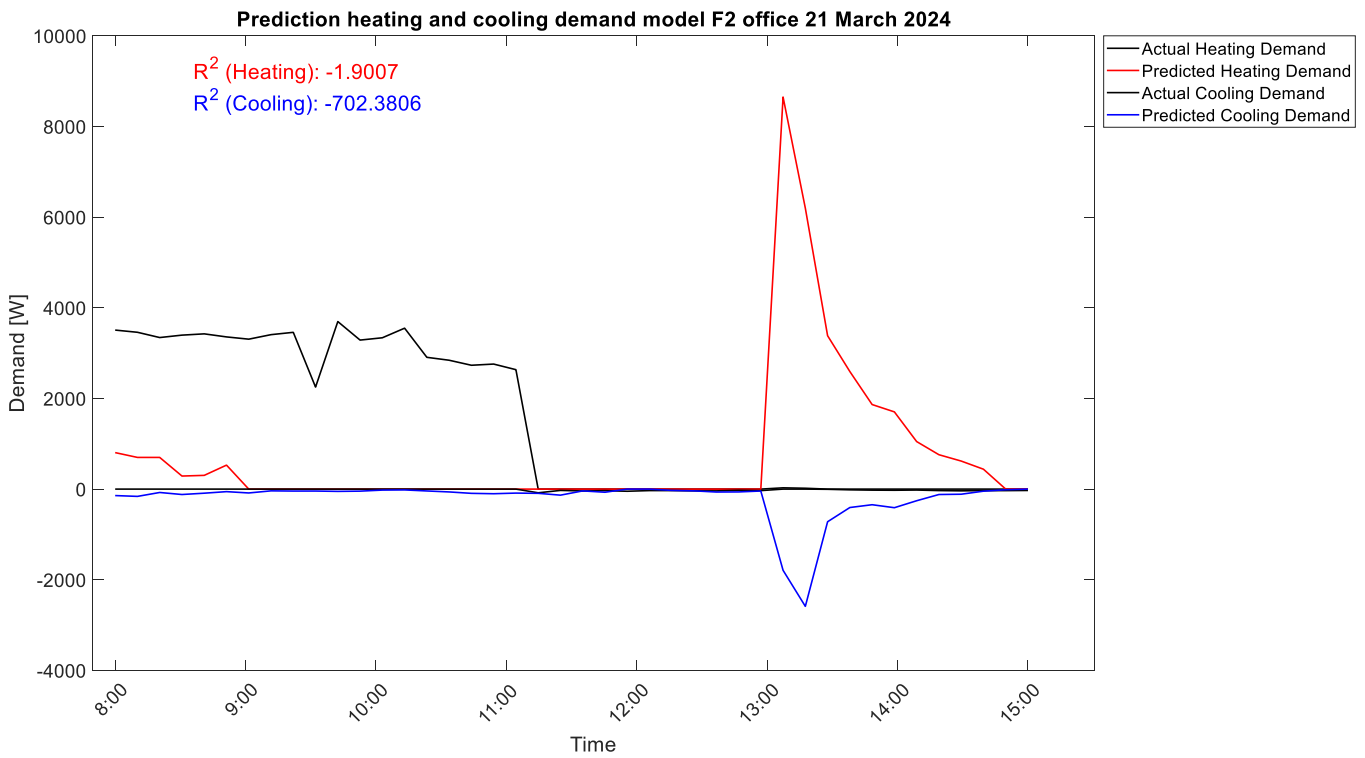

Figure 69: actual and predicted heating and cooling demand by model F_2 for the office on 21 March 2024 8 A.M. - 1 P.M.

Classroom

The selected independent variables for the heating and demand in the classroom for model F_2 are; indoor air temperature, outdoor temperature, solar light intensities in orientations west, east and south and the indoor surface temperatures of the inside of the outer wall, system wall 2, concrete pillar, system wall 3 and ceiling. There are no calculated delays of the independent variables for the heating demand in the classroom.

Also in the case of the classroom, the highest R_2 values of 22.52% and 77.39%, for heating and cooling, respectively, are achieved by model F_2. The modelled cooling demand is almost reaching the peaks of the cooling demand, where the periods of no actual cooling demand are well-modelled. In contrast to the other heating Equations derived from the other models, there is a modelled peak of 28 March 2024, however, the model is still not able to reach the actual heating demand peak. Within the training data, the actual heating demand is not consistent enough to train the model well, thereby there are fewer highly intercorrelated independent variables. This means that the heating demand is not accurately predicted. In the case of the cooling demand, there are more data points of cooling demand, and several highly intercorrelated independent variables, which in this case results in an R_2 value of 35.61% for the prediction of the cooling demand, however, still the model is not able to predict the peaks and troughs of the cooling demand.

$$Q_{demand,heat,classroom} = constant + C_1\big(T_{s,concretepillar}(t)\big) + C_2(Solar\ light\ intensity\ south\ (t)) + C_3\big(T_{s,insideoftheouterwall}\ (t) \cdot T_{s,systemwall3}\ (t)\big) + C_4\big(T_{s,systemwall2}\ (t) \cdot T_{s,ceiling}\ (t)\big) + C_5\big(T_{s,insideoftheouterwall}\ (t) \cdot T_{indoorair}(t)\big) + C_6\big(T_{s,systemwall2}\ (t) \cdot T_{indoorair}(t)\big) + C_7\big(T_{s,ceiling}\ (t) \cdot T_{indoorair}(t)\big) + C_8\big(T_{s,systemwall2}\ (t)/T_{outdoor}(t)\big) + C_9\big(T_{s,systemwall3}\ (t)/T_{outdoor}(t)\big) + C_{10}\big(T_{s,systemwall2}\ (t)/ Solar\ light\ intensity\ west\ (t)\big) + C_{11}\big(T_{s,systemwall3}\ (t)/\ Solar\ light\ intensity\ west\ (t)\big) + C_{12}\big(T_{s,ceiling}\ (t)/\ Solar\ light\ intensity\ west\ (t)\big) + C_{13}\big(T_{s,systemwall2}\ (t) \cdot Solar\ light\ intensity\ south\ (t)\big) + C_{14}\big(T_{s,systemwall3}\ (t) \cdot T_{s,insideoftheouterwall}\ (t)\big) + C_{15}(T_{outdoor}(t)/ Solar\ light\ intensity\ west\ (t))\ (54)$$

$$Q_{demand,cold,classroom} = constant + C_1\big(T_{s,systemwall3}\ (t)\big) + C_2\big(T_{s,systemwall2}\ (t)/\ T_{s,systemwall3}\ (t)\big) + C_3\big(T_{s,systemwall2}\ (t)/T_{indoorair}(t)\big) + C_4\big(T_{s,systemwall2}\ (t)/T_{outdoor}(t)\big) + C_5(T_{indoorair}(t)/T_{outdoor}(t)) + C_6(T_{indoorair}(t)/Solar\ light\ intensity\ west\ (t)) + C_7\big(T_{s,insideoftheouterwall}\ (t)/ Solar\ light\ intensity\ south\ (t)\big) + C_8\big(T_{s,ceiling}\ (t) \cdot T_{s,insideoftheouterwall}\ (t-n)\big) + C_9\big(T_{indoorair}(t)/ T_{s,insideoftheouterwall}\ (t-n)\big) + C_{10}\big(T_{s,ceiling}\ (t) \cdot T_{s,concretepillar}\ (t-n)\big) + C_{11}\big(T_{outdoor}(t) \cdot T_{s,concretepillar}\ (t-n)\big) + C_{12}\big(T_{indoorair}(t)/T_{s,systemwall3}\ (t-n)\big) + C_{13}\big(T_{outdoor}(t) \cdot T_{s,systemwall3}\ (t-n)\big) + C_{14}\big(Solar\ light\ intensity\ south\ (t) \cdot T_{s,systemwall3}\ (t-n)\big) + C_{15}(T_{indoorair}(t)/T_{outdoor}(t-n)) + C_{16}(T_{outdoor}(t) \cdot T_{outdoor}(t-n)) + C_{17}(Solar\ light\ intensity\ west\ (t)/T_{outdoor}(t-n)) + C_{18}\big(T_{s,ceiling}\ (t) \cdot Solar\ light\ intensity\ east(t-n)\big) + C_{19}\big(T_{outdoor}(t) \cdot Solar\ light\ intensity\ east(t-n)\big) + C_{20}\big(Solar\ light\ intensity\ south\ (t) \cdot Solar\ light\ intensity\ east(t-n)\big) + C_{21}\big(T_{outdoor}(t-n) \cdot$$

$$Solar\ light\ intensity\ east(t-n)\big) + C_{22}\big(Solar\ light\ intensity\ south\ (t) \cdot$$
$$Solar\ light\ intensity\ south(t-n)\big)\ (55)$$

Table 37: R_2, R_2 adjusted, and RMSE (with a percentage of the maximum heating or cooling demand to the corresponding period) of the modelling and prediction of the heating and cooling demand in the classroom with model F

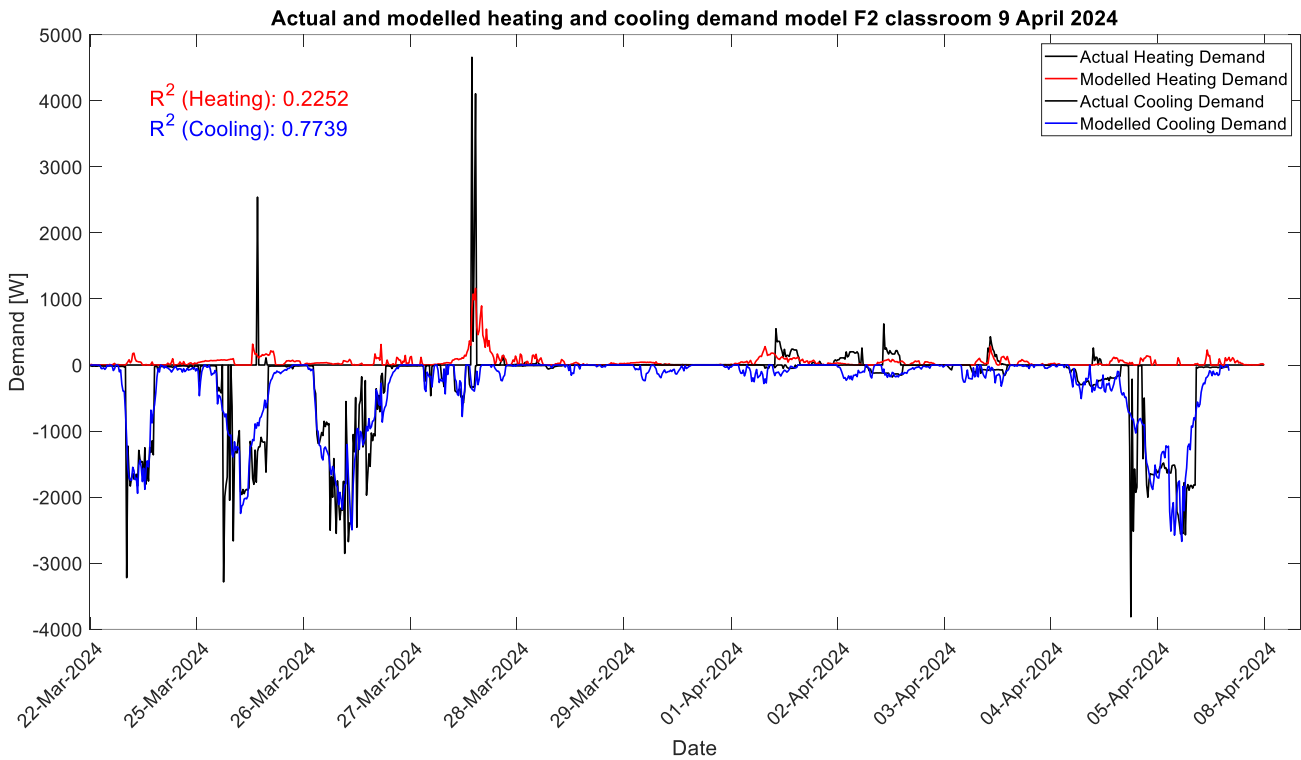|	Heating		Cooling	
	Modelling	*Prediction*	*Modelling*	*Prediction*
R²	**22.52%**	**-61.82%**	**77.39%**	**35.61%**
R² adjusted	20.80%	-114.02%	76.60%	- 10.00%
RMSE [W]	205 (4.40%)	47.59 (19.39%)	308 (8.08%)	709.46 (22.75%)

Figure 70: actual and modelled heating and cooling demand by model F_2 for the classroom from 22 March until 8 April 2024
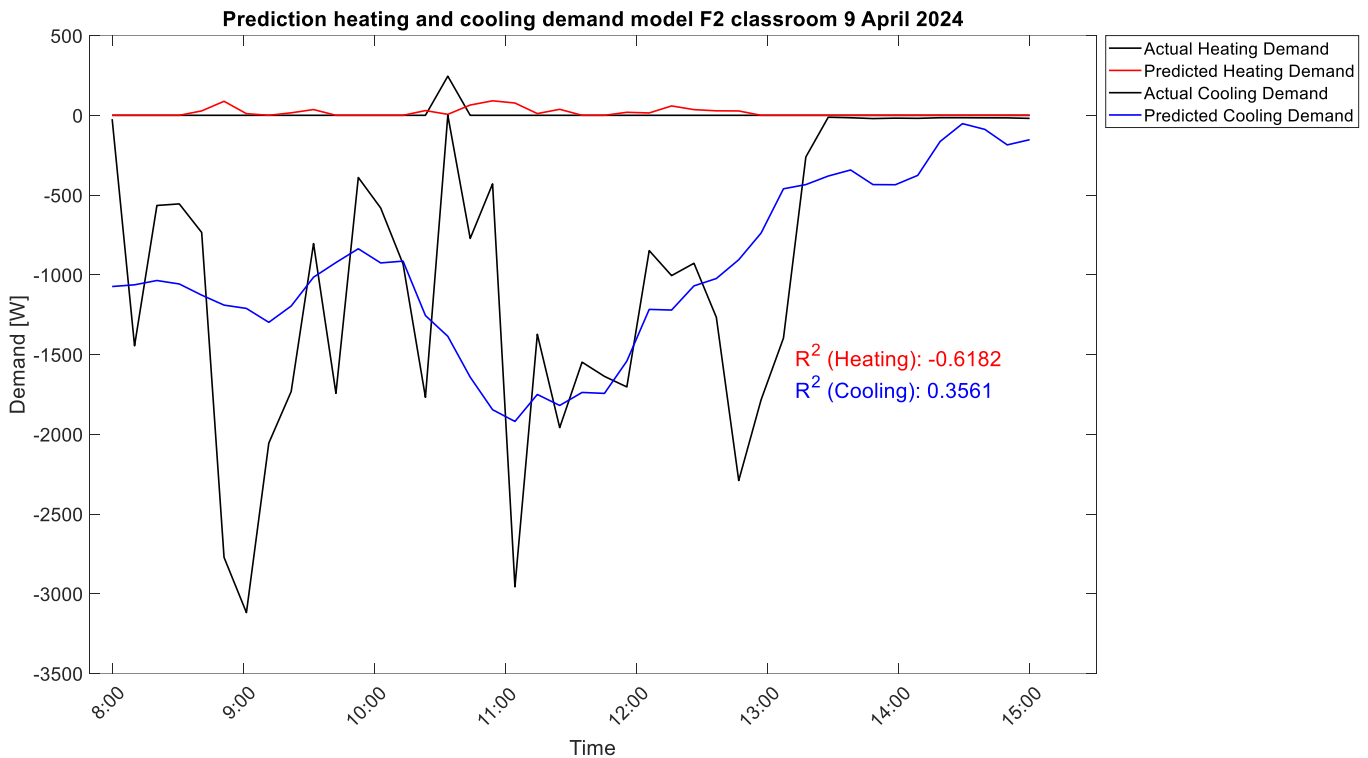
Figure 71: actual and predicted heating and cooling demand by model F_2 for the classroom on 9 April 2024 8 A.M. - 1 P.M

8.4.3 Model F analysis

For model F, the independent variables are selected based on the Pearson correlation matrices of each room for heating and cooling. For model F_1, the independent variables are selected if they have a moderate to strong correlation (Pearson correlation coefficient < 0.3) with the heating or cooling demand. No moderate to strong correlations were found between the independent variables and the heating demand in the classroom, so no model was developed for this demand in the room. Using only these independent variables for each room results in lower accuracy than model E, which includes all the independent variables. This shows that even weakly correlated independent variables are needed to achieve higher accuracy in modelling the heating and cooling demand. However, including only the moderately and strongly correlated independent variables results in better prediction of the cooling demand for 7 hours in advance compared to models A and E. The highest accuracy for predicting the cooling demand in the classroom is achieved. Using fewer independent variables makes the models less complex and less sensitive to multicollinearity, and the model is robust for using new data to predict the heating and cooling demand.

In the case of the office, models A, B, and E predict a high peak in heating and cooling demand after 1 P.M., but there is no actual demand. In contrast, model F_2 predicts shorter peaks.

There are also independent variables that, when combined with another independent variable, have a strong correlation with the heating or cooling demand. An example is the solar light intensity, which was only included in model F_1 for the cooling demand in the office. Solar light intensities have strong correlations with the heating and cooling demand in both rooms, explained by the heat capturing of the thermal mass that is released later. In model F_2, independent variables are selected that have a moderate or strong correlation with the heating or cooling demand together with another independent variable, resulting in the highest modelling accuracies for the heating and cooling demand in both rooms. The model is better at modelling low or zero heating and cooling demand. The achieved equations are complex, and not all correlations between the independent variables and the heating or cooling demand can be easily explained physically. The increased computer load made the model harder to implement. Most independent variables have several moderate or strong correlations with other independent variables, causing high multicollinearity, and resulting in an overfitting model. The results of modelling the heating and cooling demand are very high, but when the model is introduced to new data, it overreacts to small changes in the value of the independent variables.

9. Conclusion

This study aimed to develop a simple, fast and data-driven model to predict the heating and cooling demand of a building 7 hours in advance as part of model predictive control (MPC).

To emphasize the social, economic and technical potential connected to the field of Industrial Ecology of predicting the heating and cooling demand, it is possible to be demand flexible as a building. This results in the potential of reducing CO_2 emissions by using less fossil fuels, increasing energy efficiency what results in resource use reduction, reporting for ESG, and reducing network imbalances and network congestion.

Different models within model predictive control (MPC) to model and predict the heating and cooling demand are discussed. Therefore the choice is made to select a grey box model: a multivariate linear regression model. In theory, this is a black box model, however, the independent variables are selected on the thermal energy balance. The main advantages of this model are that a relatively small dataset is required, and the results are physically interpretable.

To model the heating and cooling demand, the thermal energy balance is analysed, and the measurable variables are selected, whereby the following sub-question is answered:

What are the data inputs of the multivariate linear regression model and which statistical validation and search process are used to build the model?

This resulted in the following chosen independent variables; indoor surface temperature, indoor air temperature, outdoor temperature, internal heat gains, solar light intensity and wind speed. The data is collected from the case study: the Haagse Hogeschool in Delft. In this building, two rooms are selected, based on orientation, occupancy rate and thermal mass. With the data collection, the following sub-question is answered:

How is the required data collected for each variable, and is this data set complete?

The building is equipped with a lot of sensors, which makes it possible to use actual data. The indoor surface temperatures, which are related to the thermal mass of the rooms are not measured in the room, this data is collected with external sensors. The other missing data points are discussed within the report and the discussion.

After the data collection, it was essential to analyse the data to get a hypothesis of the accuracy of the models, to answer the following sub-question:

How are the variables and the heating and cooling demand correlated to each other, and which variables will significantly contribute to a higher model accuracy?

First, the heating and cooling demand is analysed. It is concluded that the heating and cooling demand does not follow a certain pattern. As expected the indoor surface temperatures are highly correlated with the heating and cooling demand. Where the internal heat gains and the windspeed were not highly correlated. These correlations are visualized with correlation matrixes, where also the correlation between the independent variables where shown. It is concluded that there is a strong correlation between the independent variables, indoor surface temperature and also with the outdoor temperature.

The results of the developed models are summarised in Table 38. The first two developed models (A and B) followed the approach of Jurado López (2017). Model A is a static model which uses only the indoor surface temperatures and the outdoor temperature. R_2 values between 3.98% and 58% were reached for modelling the heating and cooling demand. Only for the prediction of the cooling demand a positive R_2 was achieved. Model B is a dynamic model were the indoor surface temperatures were replaced by current and delayed (calculated by cross-correlation) internal heat gains and solar light intensities, an R^2 value of 0.93% - 44.53%. This model is not able to predict the heating or cooling demand with a positive R^2 value. To overcome the problem of limited heating demand data points, model C is developed to model the thermal demand with the Approach of Model A, resulting in model C_1 and with the Approach of Model B, resulting in Model C_2. The R^2 value of these models is between the R^2 values of the heating and cooling demand in models A and B, not resulting in a better modelling and predicting result. After that, model D is developed, where the temperature difference where used instead of single data inputs. To determine if this increased the accuracy, a comparison is made with model E, where all the independent variables based on the thermal energy balance, were used. Higher R^2 values (4.39% - 64.08%) were achieved by model E than model D. Lastly model F is developed to reduce the independent variables, to overcome the problem of overfitting and complexity. Independent variables were selected based on the Pearson-correlation matrixes, where in model F_1 the directly moderate or strong correlated independent variables were selected and in model F_2 the independent variables were selected that had a moderate or strong correlation with the heating or cooling demand with another independent variable. The highest R^2 values of the model are achieved by model F_2, however, this was only the case for modelling the heating and cooling demand, the model is not able to predict the heating or cooling demand, due to overfitting. To answer the main research question:

Is it possible to model and predict for 7 hours in advance the heating and cooling demand of a building during opening hours with actual data as input for a simple, fast and data-driven model?

It is possible to mode the heating and cooling demand, however it is not possible to predict the heating and cooling demand with a multivariate linear regression model.

Tabel 38: overview of the used models for heating and cooling for each model, with the used independent variables and the R² values

Model	Static/dynamic	Independent variables	Room	Demand	R² modelling	R² prediction
A	Static	$T_{s,surfaces}(t), T_{outdoor}(t)$	Office	Heating	17.76%	-10.98%
				Cooling	34.70%	-416.13%
			Classroom	Heating	3.98%	-8.83%
				Cooling	58.00%	26.95%
B	Dynamic	$T_{outdoor}(t), T_{indoorair}(t), Windspeed(t)$, solar light intensity$(t), Q_{internal}(t)$, solar light intensity $(t-n), Q_{internal}(t-n)$	Office	Heating	24.72%	-16.17%
				Cooling	43.53%	-209.28%
			Classroom	Heating	0.93%	-69.27%
				Cooling	44.91%	-44.79%
C₁	Static	$T_{s,surfaces}(t), T_{outdoor}(t)$	Office	Thermal	31.50%	-5.73%
			Classroom		50.91%	24.41%
C₂	Static	$T_{outdoor}(t), T_{indoorair}(t), Windspeed(t)$, solar light intensity$(t), Q_{internal}(t)$, solar light intensity $(t-n), Q_{internal}(t-n)$	Office	Thermal	36.90%	-14.97%
			Classroom		33.93%	-68.81%
D	Static	$T_{outdoor} - T_{indoor}(t), V_{wind} \cdot (T_{outdoor} - T_{indoor})(t), Q_{solar}(t), Q_{internal}(t), (T_{indoor} - T_{surface}(t), T_{outdoor} - T_{indoor}(t-n), (V_{wind} \cdot (T_{outdoor} - T_{indoor})(t-n), Q_{solar}(t-n), Q_{internal}(t-n), (T_{indoor} - T_{surface}(t-n)$	Office	Heating	25.39%	-15.47%
				Cooling	43.94%	-216.03%
	Dynamic		Classroom	Heating	4.40%	-17.46%
				Cooling	62.84%	34.72%
E	Dynamic		Office	Heating	40.36%	-13.49%

		$T_{indoor}(t), T_{outdoor}(t), T_{surface}(t), V_{wind}(t),$		Cooling	50.35%	-14642.67%
	Static	$Q_{solar}(t), Q_{internal}(t), T_{indoor}(t-n),$	Classroom	Heating	4.39%	-17.85%
	Dynamic	$T_{outdoor}(t-n), T_{surface}(t-n), V_{wind}(t-n),$		Cooling	64.08%	35.34%
		$Q_{solar}(t-n), Q_{internal}(t-n)$				
F_1	Dynamic	$T_{surface}(t), T_{surface}(t-n), T_{indoor}(t-n)$	Office	Heating	29.90%	-20.95%
		$T_{surface}(t), Q_{solar}(t), T_{surface}(t-n),$		Cooling	39.39%	-345.05%
		$Q_{solar}(t-n)$				
			Classroom	Heating	-	-
		$T_{surface}(t), Q_{internal}(t), T_{surface}(t-n)$		Cooling	48.36%	36.60%
F_2	Dynamic	$Q_{solar}(t), T_{surface}(t), T_{outdoor}(t), T_{indoorair}(t),$	Office	Heating	69.71%	-190.07%
		$T_{surface}(t-n), T_{outdoor}(t-n)$				
		$T_{surface}(t), Q_{solar}(t), Q_{internal}(t), T_{indoorair}(t),$		Cooling	78.57%	-70238.06%
		$T_{surface}(t-n), Q_{solar}(t-n), Q_{internal}(t-n),$				
		$T_{indoorair}(t-n), T_{outdoor}(t-n)$				
	Static	$T_{surface}(t), Q_{solar}(t), T_{indoorair}(t), T_{outdoor}(t)$	Classroom	Heating	22.52%	-61.82%
	Dynamic	$T_{surface}(t), Q_{solar}(t), T_{indoorair}(t), T_{outdoor}(t),$		Cooling	77.39%	35.61%
		$T_{surface}(t-n), Q_{solar}(t-n), T_{outdoor}(t-n)$				

10. Discussion

This research, to determine if it is possible to predict the heating and cooling demand with a fast simple data-driven model: a multivariate linear regression model has several limitations.

Incomplete data

This research aimed to use actual data of the test the models Jurado López (2017) has developed with simulated data. The case study was chosen on the availability of data, however, the data set was not complete. There was no data available about the solar radiation on the building. Therefore the solar light intensity is used. Methods are searched to calculate the solar radiation with the solar light intensity by using a conversion factor, however, because this would become a matter of a factor, the choice is made to use the solar light intensity. This solar light intensity is not measured in the orientation North, where the windows of the classroom is faced. Therefore the choice is made to include the light intensities of the south, east and west.

The flow of the floor system is not measured within the operation system. To calculate the heating and cooling demand, the measured valve position is used, with the maximum flow possible. This method was possible to use as the valve position where fully closed or fully opened (99-100%), therefore there was not the issue of resistance of the valve, what could result in a in reality more opened valve.

Unfortunately, there was no data available on the occupancy. In theory, the internal heat gains are a significant independent variable, this is not the case for the developed models. With the calculation of the occupancy, the accuracy of the model has not much increased with the implementation of it. The calculation of the occupancy is compared in the classroom as there was a schedule available for this room, however, this schedule seems not to be accurate as the maximum amount of persons, whom the room is designed for is exceeded. Thereby was no occasion for the number of people available, in this case, the occupancy is estimated based on the type of occasion.

From the theory and the model, it is concluded that the thermal mass has a significant influence on the heating or cooling demand. Therefore indoor surface temperature sensors are placed. Unfortunately, some sensors have been fallen or have been replaced. Thereby some sensors were giving a high or low temperature. The choice is made to exclude this data, with the temperature of all the building surfaces in the rooms not measured, which makes it harder

to include the thermal processes in the room. The most significant missing indoor surface temperature is from the floor. The importance of using this temperature in the thermal energy balance is described in Chapter 3.2.1, the biggest thermal mass of this building is within the floor, which makes the models less accurate. All the other used data points were available and accurate, which makes the data set of this study reliable.

In the model of Jurado López (2017), the input temperature of the AHU was used. Within this study to choice is made to exclude this because this temperature was also used to calculate the heating and cooling demand. By including these input temperatures, there would be double counting. The first models are developed with these input temperatures, which increases the accuracy, however, because of the argument of double counting the choice is made to exclude them. However, the indoor air temperature is used in the model, which is also the return temperature of the ventilation, which is used to calculate the heating and cooling supply. The choice is made to include the indoor air temperature as the heating and cooling demand is regulated based on the air temperature.

Multifunctionality ventilation system

As earlier explained the ventilation system is also used to guarantee a good air quality, this is the main function of this system. Thereby is the option to heat or cold the room. Multiple rooms are connected to this system, whereby an on room level regulation to heat or cold the room is not possible. For example it is possible that there is demand for fresh air, the flow entering the room will be increased, it could be that without a heating or cooling demand, heat or cold is supplied. This makes the model less accurate as the heating or cooling supply is not the same as the heating or cooling demand.

Temperature set points

The temperature set points are regulated by the building managers, whereby the occupants of the room can give their preferences. In this research this influence is neglected. The temperature set points are shown in the operation system of the building, however this data is not stored, which made it not possible to determine how these temperature set points changes over time.

The choice for the multivariate linear regression model

The accuracy of the dynamic model with independent variables: indoor surface temperature, indoor air temperature, outdoor temperature, light intensity and windspeed has high accuracy between 76.5 % and 86.1%. However, as seen in the equations there is a strong correlation between the independent variables in the model, which makes the model harder to interpret

physically. The correlation between independent variables results in some cases in a negative coefficient, where a positive coefficient is expected. Thereby was the constant negative, in cases where the dependent variable was positive, which indicates that the model was compensating the independent variables. Thereby was the constant in a lot of cases high, which indicates that the heating or cooling demand is heavily dependent on the training data set of heating and cooling demand.

The expectation was that with this accuracy of modelling the heating and cooling demand prediction was possible, however, the expectation was also that the accuracy would become lower when the heating or cooling demand was predicted. An example of this is the high achieved R^2 values of model F_2, but high negative R^2 values of predicting. However, the model was not able to predict, with very inaccurate results. A reason for this is multicollinearity. Multicollinearity occurs when two or more predictors in a regression model are highly correlated. This leads to the estimates of the regression coefficients becoming unreliable and very sensitive to small changes in the model. The standard errors of the coefficients become larger, which can result in wider confidence intervals and reduced significance of the predictors. The model because very complex and there is overfitting(May-Ostendorp et al., 2011), what was seen in model F_2. Within this model, it was able to get high accuracies of the modelling of the heating and cooling demand, but not result in predicting the heating and cooling demand. Another reason for the low accuracies of the models to predict the heating and cooling demand is endogeneity. This is statistical fault in the model, when the model indicates a cause and effect, where in reality the independent and dependent variables are influences each other in both directions (Cooper et al., 2020). In other words the independent variables are in reality not independent of the heating or cooling demand. This could be in the case of the indoor surface temperatures and the indoor air temperature, which are influenced by the heating or cooling demand.

Another reason that the R_2 values of the prediction is very low, is that the heating and cooling demand is hardly predictable, not a constant daily pattern in followed. In the case of the rooms in the Haagse Hogeschool, the heating and cooling demand is not constant or there is not a robust pattern. This makes it harder to use the multivariate linear regression model. The use of a linear model for a non-linear system would not be in theory a problem.

11. Recommendations for further research

Based on the founding of this research the following recommendation for further research are proposed:

1. **Coping with missing data**

   Within this research data points were missing (as explained in the Discussion). The biggest uncertainty was in the calculation of the occupancy, there was only data from the classroom to validate the calculation, however, this validation data was also not complete. It is advised to besides the measurements of the indoor surface temperature, to measure to occupancy, a higher influence of the internal heat gains is expected.

2. **Excluding highly intercorrelated independent variables or combing them.**

   A solution for coping with multicollinearity and endogeneity (as cause of the bad prediction results) is to exclude the highly correlated variables. Shown in the correlation matrixes this would mean excluding the indoor surface temperatures, however, this would result in very low accuracy. Another solution is to combine the highly correlated variables, this would result in one indoor surface temperature for the room. This could be done for example with the RC-model.

3. **Using a RC-model to overcome the non-linearity of the heating and cooling demand.**

   In different MPC frameworks, linear models are used, as the other parts of MPC are also often developed with linear models (Drgoňa et al., 2020). In this research the RC-model by Bacher & Madsen (2011) is discussed, the advice is to use this model for further research. The RC-model is also a grey box model, what is accurate approach to develop a model for in MPC (Drgoňa et al., 2020).

12. References

Bacher, P., & Madsen, H. (2011). Identifying suitable models for the heat dynamics of buildings. *Energy and Buildings*, *43*(7), 1511–1522. https://doi.org/10.1016/j.enbuild.2011.02.005

Bianchini, G., Casini, M., Vicino, A., & Zarrilli, D. (2016). Demand-response in building heating systems: A Model Predictive Control approach. *Applied Energy*, *168*, 159–170. https://doi.org/10.1016/j.apenergy.2016.01.088

Bokel, R. M. J. (2021). Building Physics Energy Natural Ventilation. In *Delft University of Technology*.

Bonacorsi, L., Cerasi, V., Galfrascoli, P., & Manera, M. (2024). ESG Factors and Firms' Credit Risk. *Journal of Climate Finance*, *6*, 100032. https://doi.org/10.1016/j.jclimf.2024.100032

Boodi, A., Beddiar, K., Benamour, M., Amirat, Y., & Benbouzid, M. (2018). Intelligent Systems for Building Energy and Occupant Comfort Optimization: A State of the Art Review and Recommendations. *Energies*, *11*(10), 2604. https://doi.org/10.3390/en11102604

Bouw Wereld. (n.d.). *Energielabel A++ voor de Haagse Hogeschool* . Https://Www.Bouwwereld.Nl/Bouwtechniek/Energielabel-a-Voor-de-Haagse-Hogeschool/.

Calvin, K., Dasgupta, D., Krinner, G., Mukherji, A., Thorne, P. W., Trisos, C., Romero, J., Aldunce, P., Barrett, K., Blanco, G., Cheung, W. W. L., Connors, S., Denton, F., Diongue-Niang, A., Dodman, D., Garschagen, M., Geden, O., Hayward, B., Jones, C., … Ha, M. (2023). *IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland.* https://doi.org/10.59327/IPCC/AR6-9789291691647

Cary, M., & Stephens, H. M. (2024). Economic, environmental, and technical gains from the Kyoto Protocol: Evidence from cement manufacturing. *Resources Policy*, *91*, 104926. https://doi.org/10.1016/j.resourpol.2024.104926

Catalina, T., Iordache, V., & Caracaleanu, B. (2013). Multiple regression model for fast prediction of the heating energy demand. *Energy and Buildings*, *57*, 302–312. https://doi.org/10.1016/j.enbuild.2012.11.010

Centraal Bureau voor de Statistiek, & Kadaster. (n.d.). *Energieverbruik vastgoed funderend onderwijs* . Https://Dashboards.Cbs.Nl/v2/Energieverbruik_vastgoed_funderend_onderwijs/.

Chatzithomas, C., Alexandris, S., & Karavitis, C. (2015). Multivariate linear relation for precipitation: A new simple empirical formula. *Studia Geophysica et Geodaetica*, *59*(2), 325–344. https://doi.org/10.1007/s11200-013-1162-6

Chaves-Ávila, J. P., Hakvoort, R. A., & Ramos, A. (2013). Short-term strategies for Dutch wind power producers to reduce imbalance costs. *Energy Policy*, *52*, 573–582. https://doi.org/10.1016/j.enpol.2012.10.011

Chen, W.-H., & You, F. (2023). Data-Driven Robust Model Predictive Control on Building Climate Control with Renewable Energy. *33rd European Symposium on Computer Aided Process Engineering*, 2847–2852. https://doi.org/10.1016/B978-0-443-15274-0.50453-4

Coffey, B., Haghighat, F., Morofsky, E., & Kutrowski, E. (2010). A software framework for model predictive control with GenOpt. *Energy and Buildings*, *42*(7), 1084–1092. https://doi.org/10.1016/j.enbuild.2010.01.022

Cooper, B., Eva, N., Zarea Fazlelahi, F., Newman, A., Lee, A., & Obschonka, M. (2020). Addressing common method variance and endogeneity in vocational behavior research: A review of the literature and suggestions for future research. *Journal of Vocational Behavior*, *121*, 103472. https://doi.org/10.1016/j.jvb.2020.103472

Corbin, C. D., Henze, G. P., & May-Ostendorp, P. (2013). A model predictive control optimization environment for real-time commercial building application. *Journal of Building Performance Simulation*, *6*(3), 159–174. https://doi.org/10.1080/19401493.2011.648343

Dahl Knudsen, M., & Petersen, S. (2016). Demand response potential of model predictive control of space heating based on price and carbon dioxide intensity signals. *Energy and Buildings*, *125*, 196–204. https://doi.org/10.1016/j.enbuild.2016.04.053

Del Ama Gonzalo, F., Moreno Santamaría, B., & Montero Burgos, M. J. (2023). Assessment of Building Energy Simulation Tools to Predict Heating and Cooling Energy Consumption at Early Design Stages. *Sustainability*, *15*(3), 1920. https://doi.org/10.3390/su15031920

Di Persio, L., & Fraccarolo, N. (2023). Energy Consumption Forecasts by Gradient Boosting Regression Trees. *Mathematics*, *11*(5), 1068. https://doi.org/10.3390/math11051068

Drgoňa, J., Arroyo, J., Cupeiro Figueroa, I., Blum, D., Arendt, K., Kim, D., Ollé, E. P., Oravec, J., Wetter, M., Vrabie, D. L., & Helsen, L. (2020). All you need to know about model predictive control for buildings. *Annual Reviews in Control*, *50*, 190–232. https://doi.org/10.1016/j.arcontrol.2020.09.001

Dutch Parliament. (2019). *Klimaatakkoord*.

Everitt, B. S., & Skrondal, A. (2020). *The Cambridge Dictionary of Statistics* (4th ed.). Cambridge University Press.

Ferreira-Quilice, T., Hernández-Maestro, R. M., & Gonzalez Duarte, R. (2023). Corporate sustainability transitions: Are there differences between what companies say and do and what ESG ratings say companies do? *Journal of Cleaner Production*, *414*, 137520. https://doi.org/10.1016/j.jclepro.2023.137520

Foucquier, A., Robert, S., Suard, F., Stéphan, L., & Jay, A. (2013). State of the art in building modelling and energy performances prediction: A review. *Renewable and Sustainable Energy Reviews*, *23*, 272–288. https://doi.org/10.1016/j.rser.2013.03.004

Ghiaus, C. (2014). Linear algebra solution to psychometric analysis of air-conditioning systems. *Energy*, *74*, 555–566. https://doi.org/10.1016/j.energy.2014.07.021

Goodarzi, S., Perera, H. N., & Bunn, D. (2019). The impact of renewable energy forecast errors on imbalance volumes and electricity spot prices. *Energy Policy*, *134*, 110827. https://doi.org/10.1016/j.enpol.2019.06.035

Guo, J., Yun, S., Meng, Y., He, N., Ye, D., Zhao, Z., Jia, L., & Yang, L. (2023). Prediction of heating and cooling loads based on light gradient boosting machine algorithms. *Building and Environment*, *236*, 110252. https://doi.org/10.1016/j.buildenv.2023.110252

Hennig, R. J., de Vries, L. J., & Tindemans, S. H. (2023). Congestion management in electricity distribution networks: Smart tariffs, local markets and direct control. *Utilities Policy*, *85*, 101660. https://doi.org/10.1016/j.jup.2023.101660

Hennig, R. J., de Vries, L. J., & Tindemans, S. H. (2024). Risk vs. restriction—An investigation of capacity-limitation based congestion management in electric distribution grids. *Energy Policy*, *186*, 113976. https://doi.org/10.1016/j.enpol.2023.113976

Hofmann, M., & Lindberg, K. B. (2024). Residential demand response and dynamic electricity contracts with hourly prices: A study of Norwegian households during the 2021/22 energy crisis. *Smart Energy*, *13*, 100126. https://doi.org/10.1016/j.segy.2023.100126

Itard, L. (2011). Energy in the Built Environment. In *Sustainable Urban Environments* (pp. 113–175). Springer Netherlands. https://doi.org/10.1007/978-94-007-1294-2_5

Itard, L. (2023). Lecture: Set up of a dynamic model (like Energy +). In *TU Delft*.

Joseph Thaddeus, A. I., van den Brom, P. I., & Itard, L. C. M. (2021). *Introduction to digital twins, models and parameter estimation.*

Jurado López, C. (2017). *Data-driven Predictive Control for Heating Demand in Buildings*. Delft University of Technology.

Kathirgamanathan, A., De Rosa, M., Mangina, E., & Finn, D. P. (2021). Data-driven predictive control for unlocking building energy flexibility: A review. *Renewable and Sustainable Energy Reviews*, *135*, 110120. https://doi.org/10.1016/j.rser.2020.110120

Klanatsky, P., Veynandt, F., & Heschl, C. (2023). Grey-box model for model predictive control of buildings. *Energy and Buildings*, *300*, 113624. https://doi.org/10.1016/j.enbuild.2023.113624

Korolija, I., Zhang, Y., Marjanovic-Halburd, L., & Hanby, V. I. (2013). Regression models for predicting UK office building energy consumption from heating and cooling demands. *Energy and Buildings*, *59*, 214–227. https://doi.org/10.1016/j.enbuild.2012.12.005

Kroll, A. (2000). Grey box models: concepts and application. In *New frontiers in computer intelligence and its application* (pp. 42–51). IO press.

Kwak, Y., Huh, J.-H., & Jang, C. (2015). Development of a model predictive control framework through real-time building energy management system data. *Applied Energy*, *155*, 1–13. https://doi.org/10.1016/j.apenergy.2015.05.096

Laugs, G. A. H., Benders, R. M. J., & Moll, H. C. (2020). Balancing responsibilities: Effects of growth of variable renewable energy, storage, and undue grid interaction. *Energy Policy*, *139*, 111203. https://doi.org/10.1016/j.enpol.2019.111203

Li, Z., Han, Y., & Xu, P. (2014). Methods for benchmarking building energy consumption against its past or intended performance: An overview. *Applied Energy*, *124*, 325–334. https://doi.org/10.1016/j.apenergy.2014.03.020

Lin, X., Tian, Z., Song, W., Lu, Y., Niu, J., Sun, Q., & Wang, Y. (2024). Grey-box modeling for thermal dynamics of buildings under the presence of unmeasured internal heat gains. *Energy and Buildings*, *314*, 114229. https://doi.org/10.1016/j.enbuild.2024.114229

Lund, H., Hvelplund, F., Østergaard, P., Möller, B., Mathiesen, B. V., Connolly, D., & Andersen, A. N. (2014). Analysis. In *Renewable Energy Systems* (pp. 131–184). Elsevier. https://doi.org/10.1016/B978-0-12-410423-5.00006-7

Marszal-Pomianowska, A., Motoasca, E., Pothof, I., Felsmann, C., Heiselberg, P., Cadenbach, A., Leusbrock, I., O'Donovan, K., Petersen, S., & Schaffer, M. (2024). Strengths, weaknesses, opportunities and threats of demand response in district heating and cooling systems. From passive customers to valuable assets. *Smart Energy*, 100135. https://doi.org/10.1016/j.segy.2024.100135

Mathworks. (n.d.). *stepwiselm*. Https://Nl.Mathworks.Com/Help/Stats/Stepwiselm.Html.

May-Ostendorp, P., Henze, G. P., Corbin, C. D., Rajagopalan, B., & Felsmann, C. (2011). Model-predictive control of mixed-mode buildings with rule extraction. *Building and Environment*, *46*(2), 428–437. https://doi.org/10.1016/j.buildenv.2010.08.004

Merema, B., Saelens, D., & Breesch, H. (2022). Demonstration of an MPC framework for all-air systems in non-residential buildings. *Building and Environment*, *217*, 109053. https://doi.org/10.1016/j.buildenv.2022.109053

Michael, P. R., Johnston, D. E., & Moreno, W. (2020). A conversion guide: solar irradiance and lux illuminance. *Journal of Measurements in Engineering*, *8*(4), 153–166. https://doi.org/10.21595/jme.2020.21667

Miletić, M., Gržanić, M., Pavić, I., Pandžić, H., & Capuder, T. (2022). The effects of household automation and dynamic electricity pricing on consumers and suppliers. *Sustainable Energy, Grids and Networks*, *32*, 100931. https://doi.org/10.1016/j.segan.2022.100931

Müller, T., & Möst, D. (2018). Demand Response Potential: Available when Needed? *Energy Policy*, *115*, 181–198. https://doi.org/10.1016/j.enpol.2017.12.025

Netbeheerder Nederland. (2024, March 14). *Capaciteitskaart invoeding elektriciteitsnet*. https://capaciteitskaart.netbeheernederland.nl/

Norouzi, F., Hoppe, T., Kamp, L. M., Manktelow, C., & Bauer, P. (2023). Diagnosis of the implementation of smart grid innovation in The Netherlands and corrective actions. *Renewable and Sustainable Energy Reviews*, *175*, 113185. https://doi.org/10.1016/j.rser.2023.113185

Peplinski, M., Dilkina, B., Chen, M., Silva, S. J., Ban-Weiss, G. A., & Sanders, K. T. (2024). A machine learning framework to estimate residential electricity demand based on smart meter electricity, climate, building characteristics, and socioeconomic datasets. *Applied Energy, 357,* 122413. https://doi.org/10.1016/j.apenergy.2023.122413

Pino-Mejías, R., Pérez-Fargallo, A., Rubio-Bellido, C., & Pulido-Arcas, J. A. (2017). Comparison of linear regression and artificial neural networks models to predict heating and cooling energy demand, energy consumption and CO 2 emissions. *Energy*, *118*, 24–36. https://doi.org/10.1016/j.energy.2016.12.022

Pollitt, M. G., von der Fehr, N.-H. M., Willems, B., Banet, C., Le Coq, C., & Chyong, C. K. (2024). Recommendations for a future-proof electricity market design in Europe in light of the 2021-23 energy crisis. *Energy Policy*, *188*, 114051. https://doi.org/10.1016/j.enpol.2024.114051

Pulido-Arcas, J. A., Pérez-Fargallo, A., & Rubio-Bellido, C. (2016). Multivariable regression analysis to assess energy consumption and CO2 emissions in the early stages of offices design in Chile. *Energy and Buildings*, *133*, 738–753. https://doi.org/10.1016/j.enbuild.2016.10.031

Raftery, P., Keane, M., & O'Donnell, J. (2011). Calibrating whole building energy models: An evidence-based methodology. *Energy and Buildings*, *43*(9), 2356–2364. https://doi.org/10.1016/j.enbuild.2011.05.020

Rasooli, A., & Itard, L. (2018). In-situ characterization of walls' thermal resistance: An extension to the ISO 9869 standard method. *Energy and Buildings*, *179*, 374–383. https://doi.org/10.1016/j.enbuild.2018.09.004

Rasooli, A., & Itard, L. (2020). Automated in-situ determination of buildings' global thermo-physical characteristics and air change rates through inverse modelling of smart meter and air temperature data. *Energy and Buildings*, *229*, 110484. https://doi.org/10.1016/j.enbuild.2020.110484

Rodrigues, L. S., Marques, D. L., Ferreira, J. A., Costa, V. A. F., Martins, N. D., & Neto Da Silva, F. J. (2022). The Load Shifting Potential of Domestic Refrigerators in Smart Grids: A Comprehensive Review. *Energies*, *15*(20), 7666. https://doi.org/10.3390/en15207666

Santamarta, J. C., García-Gil, A., Expósito, M. del C., Casañas, E., Cruz-Pérez, N., Rodríguez-Martín, J., Mejías-Moreno, M., Götzl, G., & Gemeni, V. (2021). The clean energy transition of heating and cooling in touristic infrastructures using shallow geothermal energy in the Canary Islands. *Renewable Energy*, *171*, 505–515. https://doi.org/10.1016/j.renene.2021.02.105

Syb van Breda & Co. (n.d.). *Haagse Hogeschool gebouw Delft* . Https://Sybvanbreda.Com/Nl/Project/Haagse-Hogeschool-Gebouw-Delft/.

United Nations Framework Convention on Climate Change (UNFCCC). (2015). The Paris Agreement. *Conference of the Parties (COP) 21*.

Veljkovic, A., Pohoryles, D. A., & Bournas, D. A. (2023). Heating energy demand estimation of the EU building stock: Combining building physics and artificial neural networks. *Energy and Buildings*, *298*, 113474. https://doi.org/10.1016/j.enbuild.2023.113474

Wang, S., & Xu, X. (2006). Simplified building model for transient thermal performance estimation using GA-based parameter identification. *International Journal of Thermal Sciences*, *45*(4), 419–432. https://doi.org/10.1016/j.ijthermalsci.2005.06.009

Yang, S., Gao, H. O., & You, F. (2024). Demand flexibility and cost-saving potentials via smart building energy management: Opportunities in residential space heating across the US. *Advances in Applied Energy*, *14*, 100171. https://doi.org/10.1016/j.adapen.2024.100171

Yang, Z., & Jiang, Y. (2024). Quantifying resilient urban energy systems: Statistical analysis of climate adaptation, renewable integration, and socioeconomic dynamics. *Sustainable Cities and Society*, *101*, 105153. https://doi.org/10.1016/j.scs.2023.105153

Yu, X., Ren, Z., Liu, P., Imsland, L., & Georges, L. (2024). Comparison of time-invariant and adaptive linear grey-box models for model predictive control of residential buildings. *Building and Environment*, *254*, 111391. https://doi.org/10.1016/j.buildenv.2024.111391

Zou, K. H., Tuncali, K., & Silverman, S. G. (2003). Correlation and Simple Linear Regression. *Radiology*, *227*(3), 617–628. https://doi.org/10.1148/radiol.2273011499

Appendix A: List of variables: Thermal energy balance

Table 39: overview of the equations used in the thermal energy balance (Itard, 2011; Jurado López, 2017)

Variable	Description	Equation
$Q_{envelop}$	Heat transmission through the facades and roof.	$Q_{envelop} = \sum_i U^j_{envelop} \cdot A^j_{envelop} \cdot (T_{outdoor} - T_{indoorair})[W]$ j: for each façade or roof of orientation $U^j_{envelop}$: heat transfer coefficient of the envelop [W/m²/K] $A^j_{envelop}$: the surface area of the façade or roof [m²] $T_{outdoor}$: outdoor temperature [°C] $T_{indoorair}$: indoor air temperature [°C]
Q_{ground}	Heat transmission through the ground.	$Q_{ground} = U_{floor} \cdot A_{floor} \cdot (T_{ground} - T_{indoorair})[W]$ U_{floor}: heat transfer coefficient of the floor [W/m²/K] A_{floor}: surface area floor [m²] T_{ground}: surface temperature of the ground [°C]
$Q_{internal}$	Internal heat gains by people, light and/or applications.	$Q_{internal} = n_{people} \cdot Q_{body} + A_{ceilings} \cdot Q_{light} + A_{floor} \cdot Q_{equipment} [W]$ n_{people}: number of people $A_{ceilings}$: total area of all the ceilings [m²] Q_{body}, Q_{light} and $Q_{applications}$: heat gain per person, artificial light source and applications [W]
Q_{solar}	Solar heat gains by direct radiation, reflected radiation and diffuse radiation.	$Q_{solar} = Q_{sol\,direct} + Q_{sol\,dif} + Q_{relective} [W]$ $Q_{sol\,direct}$: direct solar radiation, depends on the reflection properties of the windows[W] $Q_{sol\,dif}$: diffused solar radiation (for example by clouds) [W] $Q_{reflective}$: solar radiation reflected on the ground[W]
$Q_{ventilation}$	Heat losses by natural ventilation, mechanical supply ventilation or mechanical exhaust ventilation or, balanced ventilation.	$Q_{ventilation} = m_{vent} \cdot C_{p,air} (T_{out\,AHU} - T_{indoorair})[W]$ m_{vent}: mass flow rate of the ventilation air [m³/s] $C_{p,air}$: heating capacity of air (J/kg·K) $T_{out\,AHU}$: temperature of ventilation air from air handling unit (AHU) [°C]
$Q_{infiltrations}$	Infiltration losses, for example through cracks in the construction	$Q_{infiltrations} = (m_{opening} + m_{cracks}) \cdot C_{pair} (T_{outdoor} - T_{indoorair})[W]$ $m_{opening}$: mass flow rate through openings [m³/s]

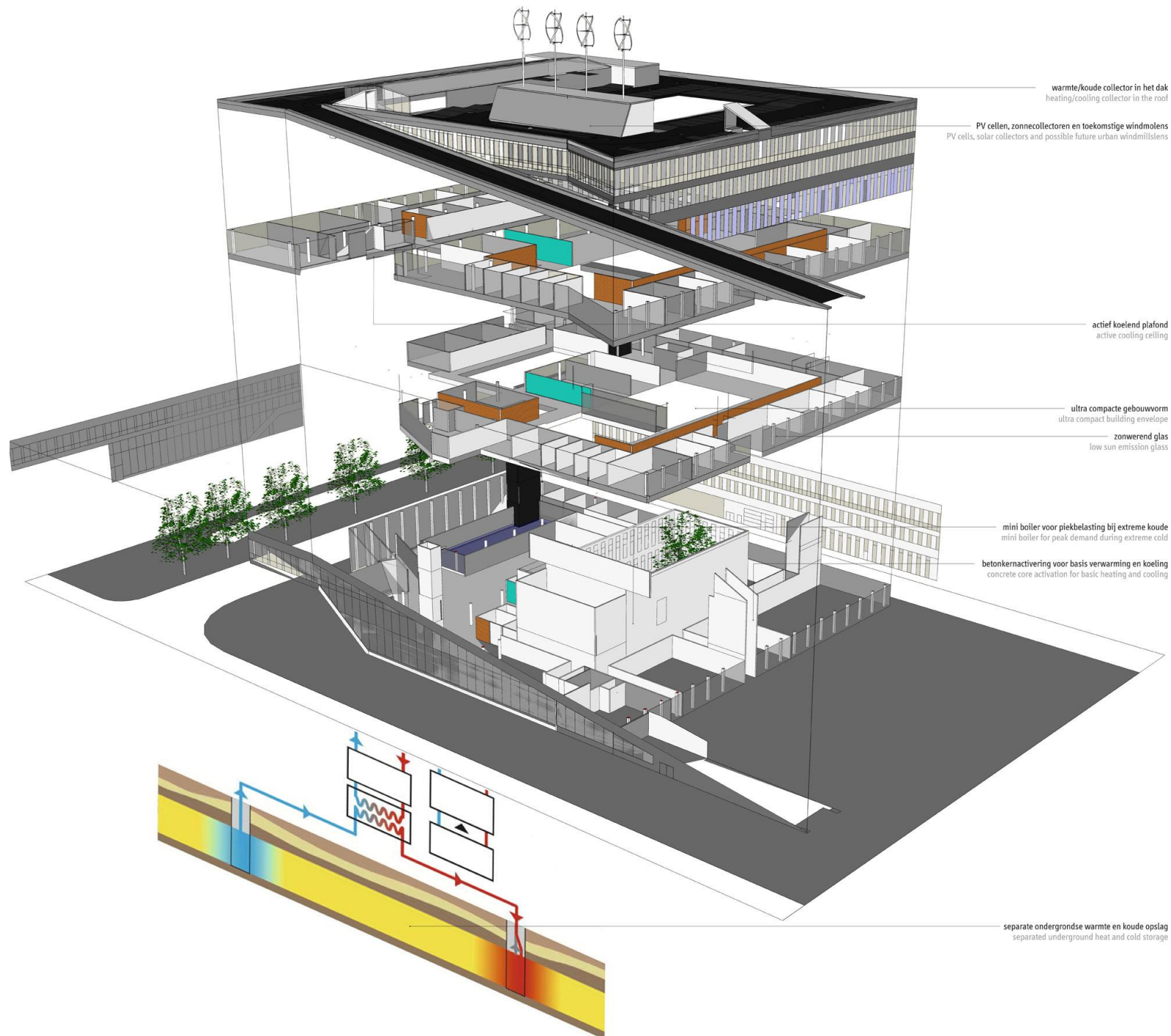|		m_{cracks}: mass flow rate through cracks [m³/s]
$Q_{thermal.mass}$	Heat is stored within the thermal mass of a building. Heavy buildings are storing more heat in the building than lighter buildings.	$Q_{thermal\ mass} = \alpha_i \cdot A_{indoor\ surfaces} \cdot \left(T^t_{surface} - T_{indoorair}\right)[W]$ α_i: indoor combined heat transfer coefficients for convection and radiation [W/m²/K] $A_{indoor\ surfaces}$: total area of indoor surfaces in contact with the indoor air [m²] T^t_s: indoor surface temperature [°C]

Appendix B: Building overview Haagse Hogeschool Delft

warmte/koude collector in het dak
heating/cooling collector in the roof

PV cellen, zonnecollectoren en toekomstige windmolens
PV cells, solar collectors and possible future urban windmillsens

actief koelend plafond
active cooling ceiling

ultra compacte gebouwvorm
ultra compact building envelope

zonwerend glas
low sun emission glass

mini boiler voor piekbelasting bij extreme koude
mini boiler for peak demand during extreme cold

betonkernactivering voor basis verwarming en koeling
concrete core activation for basic heating and cooling

separate ondergrondse warmte en koude opslag
separated underground heat and cold storage

Figure 74: schematic overview of the case building: The Haagse Hogeschool in Delft (Syb van Breda & Co, n.d.)

Appendix C: Window position
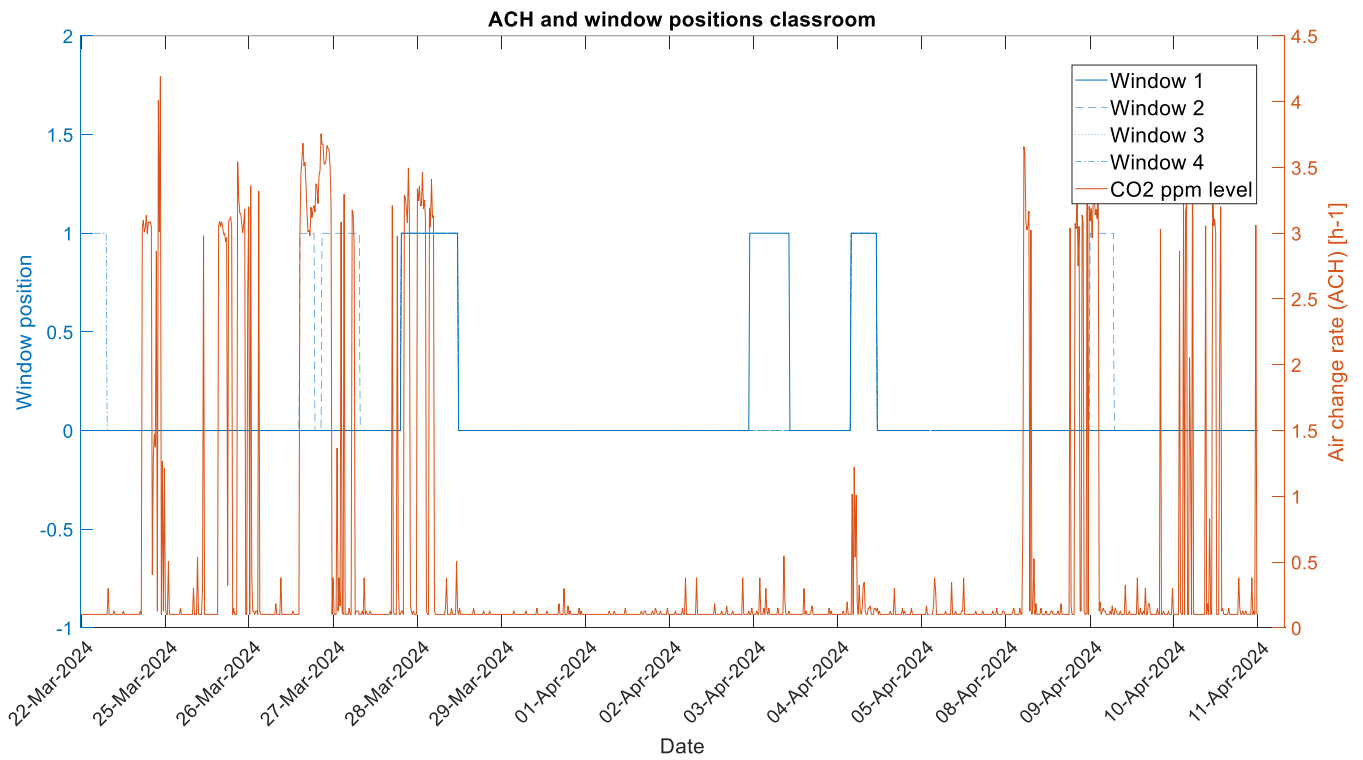
Figure 75: Air change rate of the ventilation in the classroom with the position of the windows
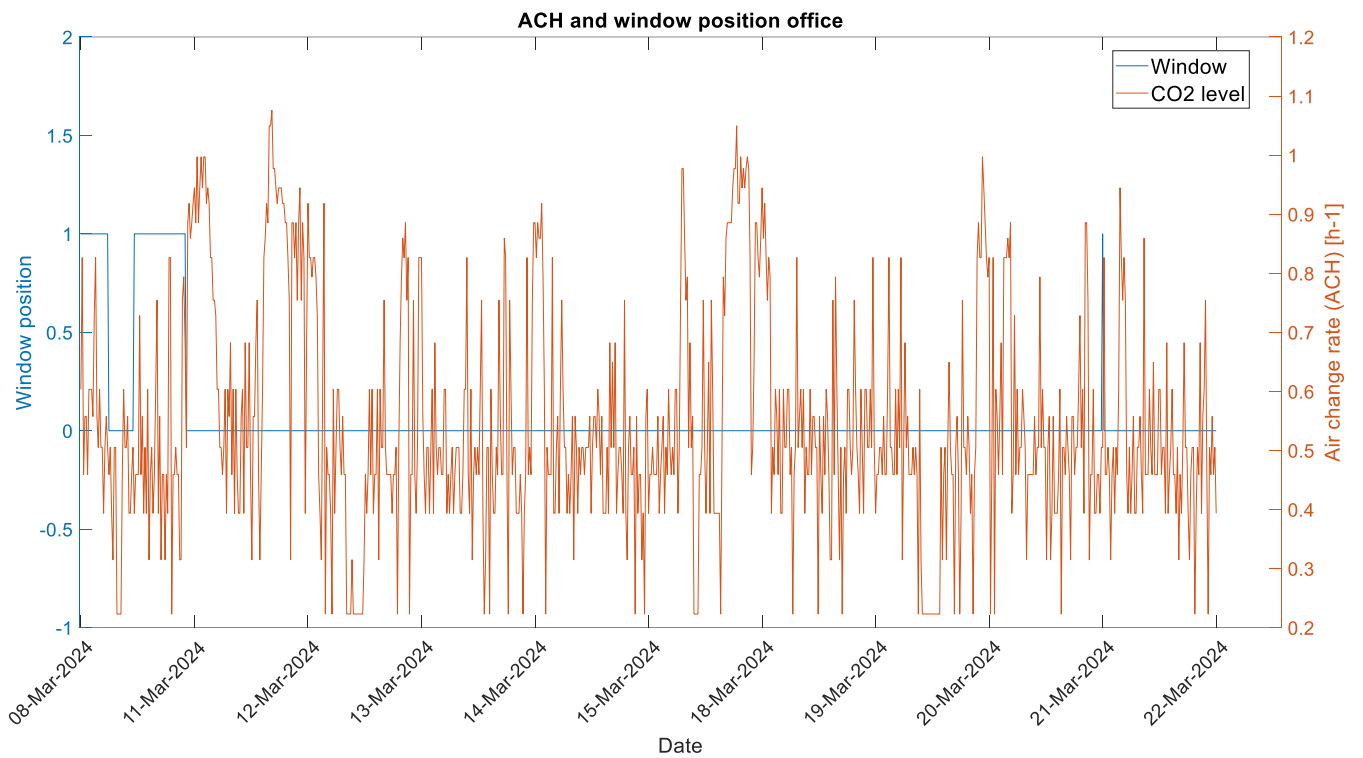

Figure 76: Air change rate of the ventilation in the office with the position of the office

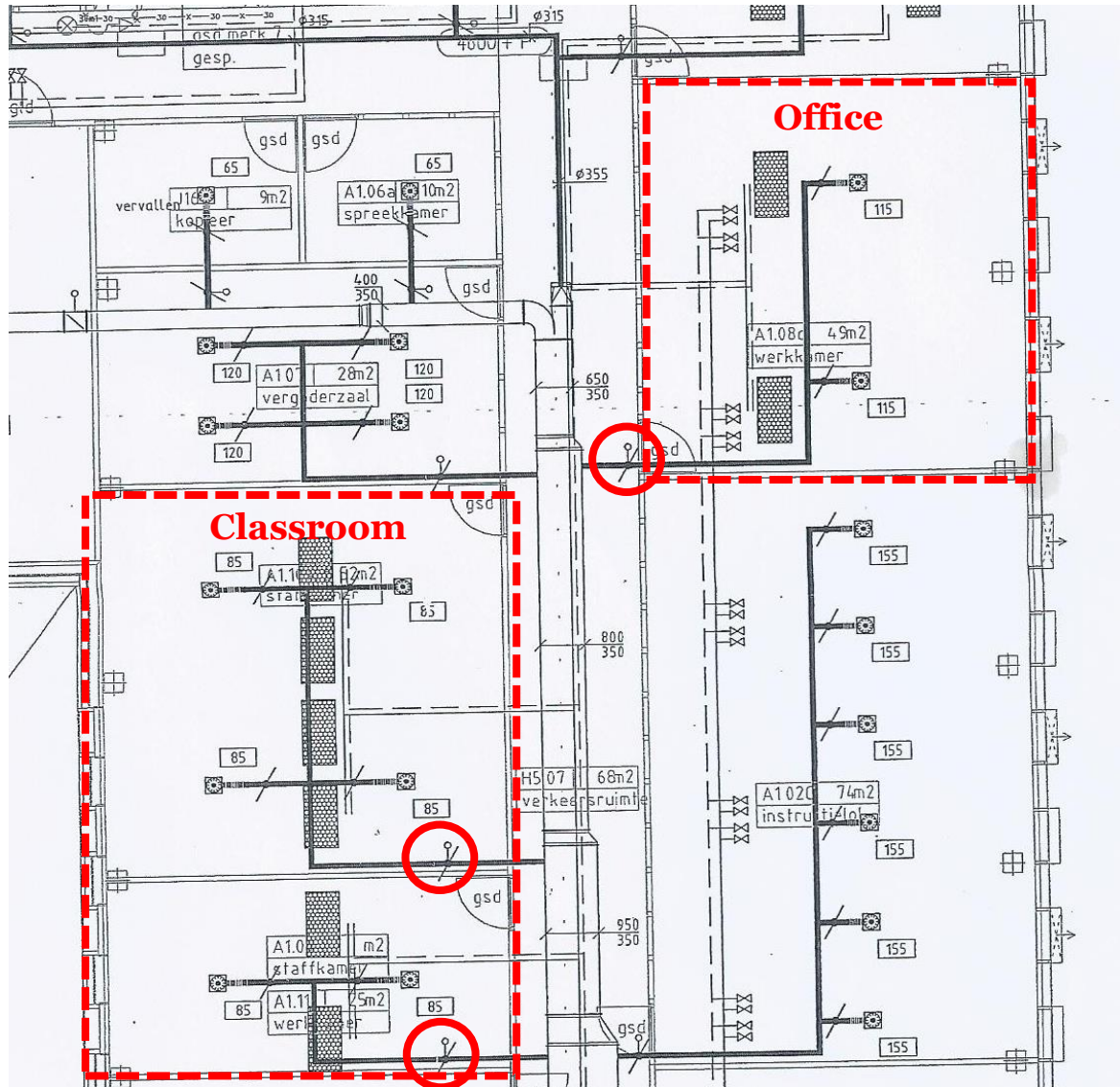# Appendix D: Installation drawing ventilation

Figure 77: Installation drawing of the ventilation system in the test-zone

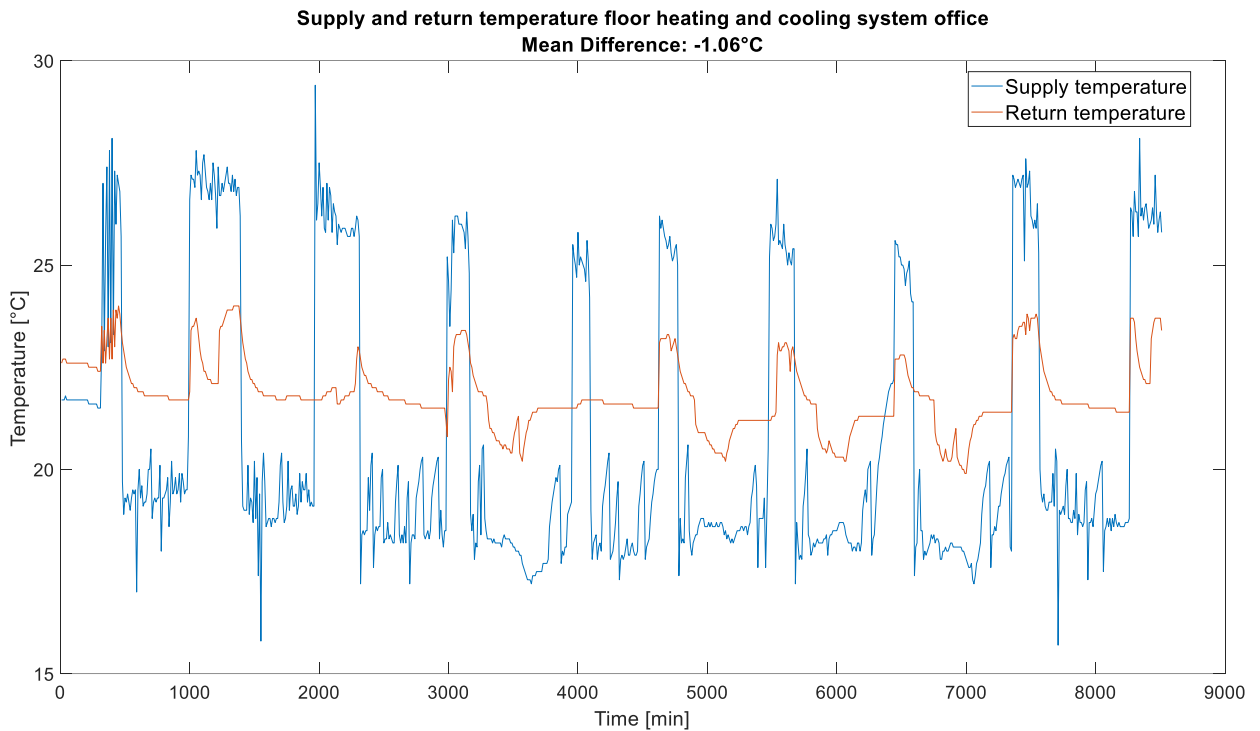# Appendix 5: Supply and return temperatures floor, panel and ventilation system

Figure 78: supply and return temperature floor heating and cooling system office
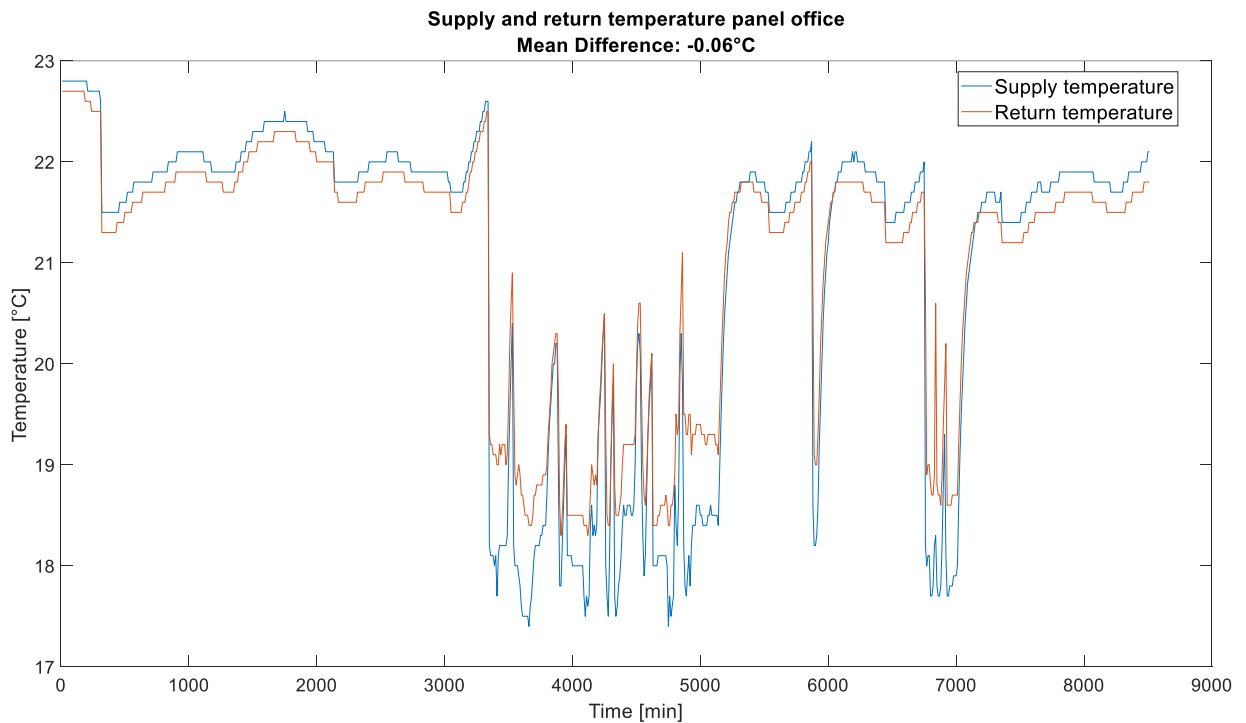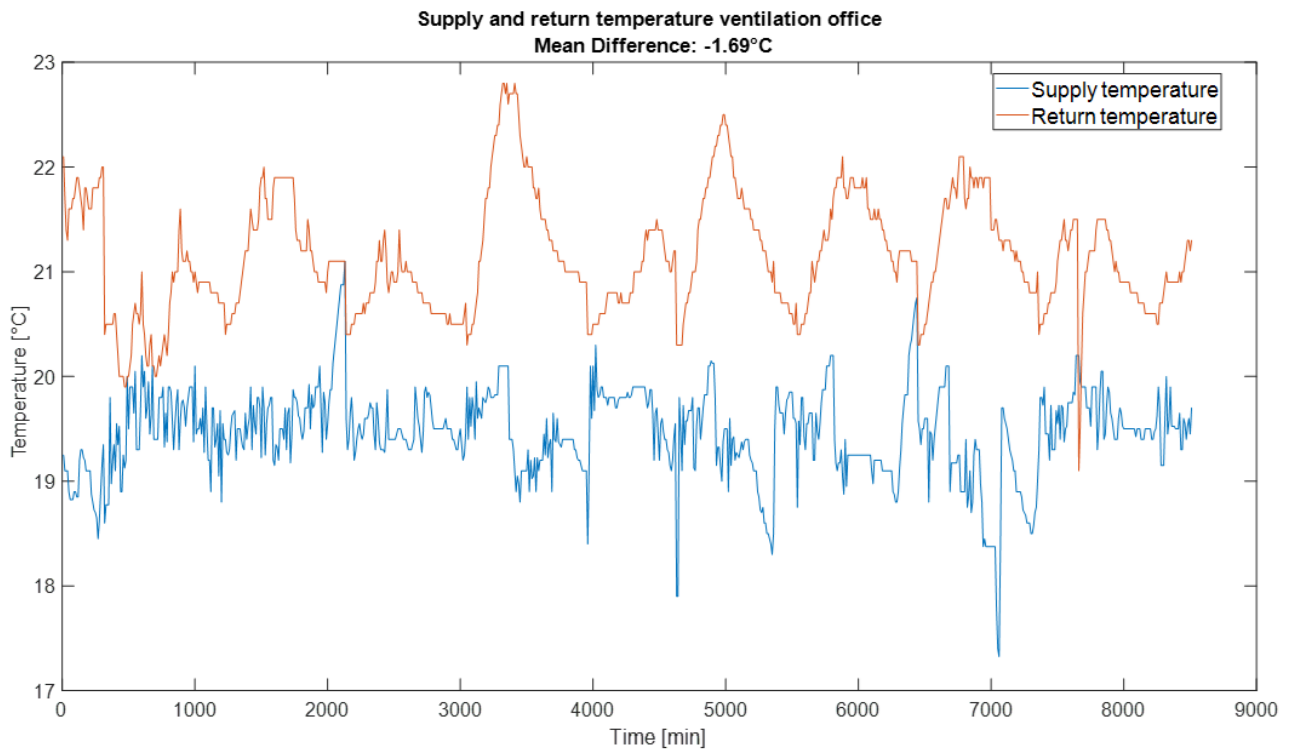

Figuur 79: supply and return temperature panel office

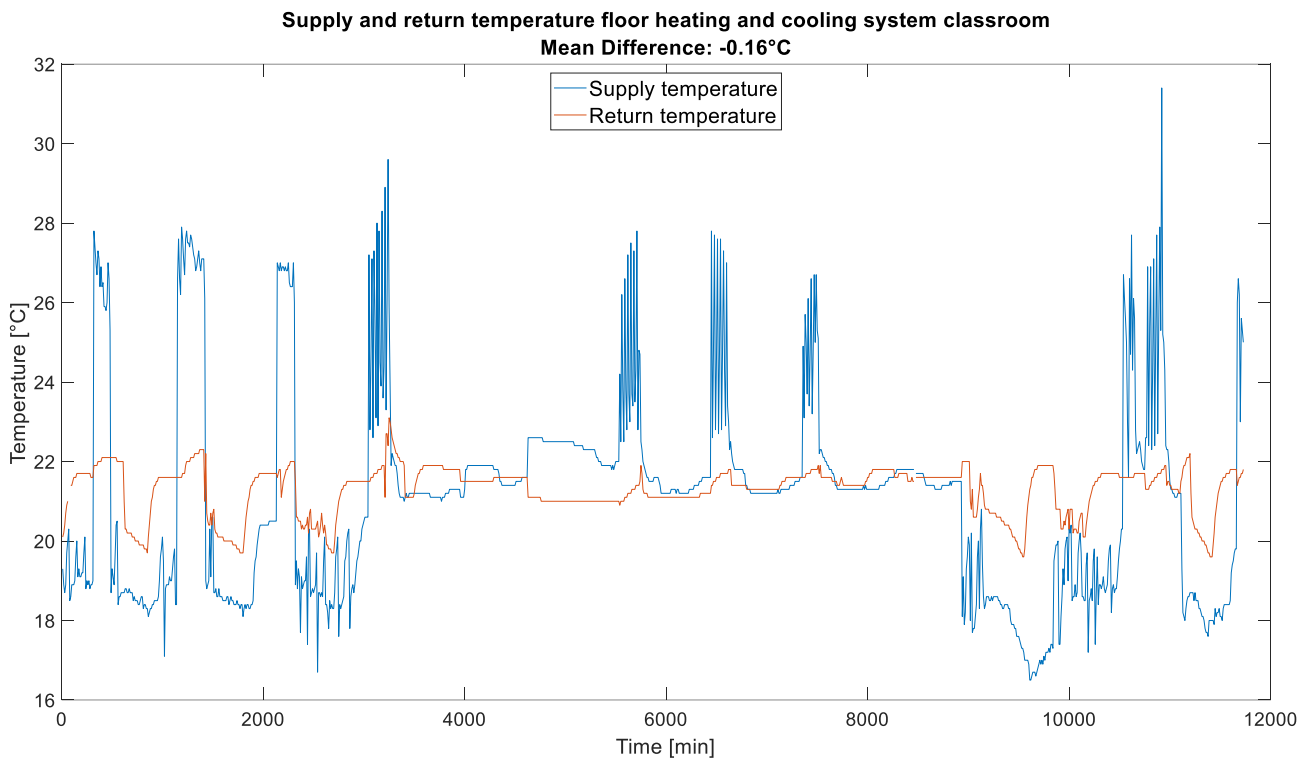*Figure 80: supply and return of the ventilation office*

Figure 81: supply and return temperature floor heating and cooling system classroom
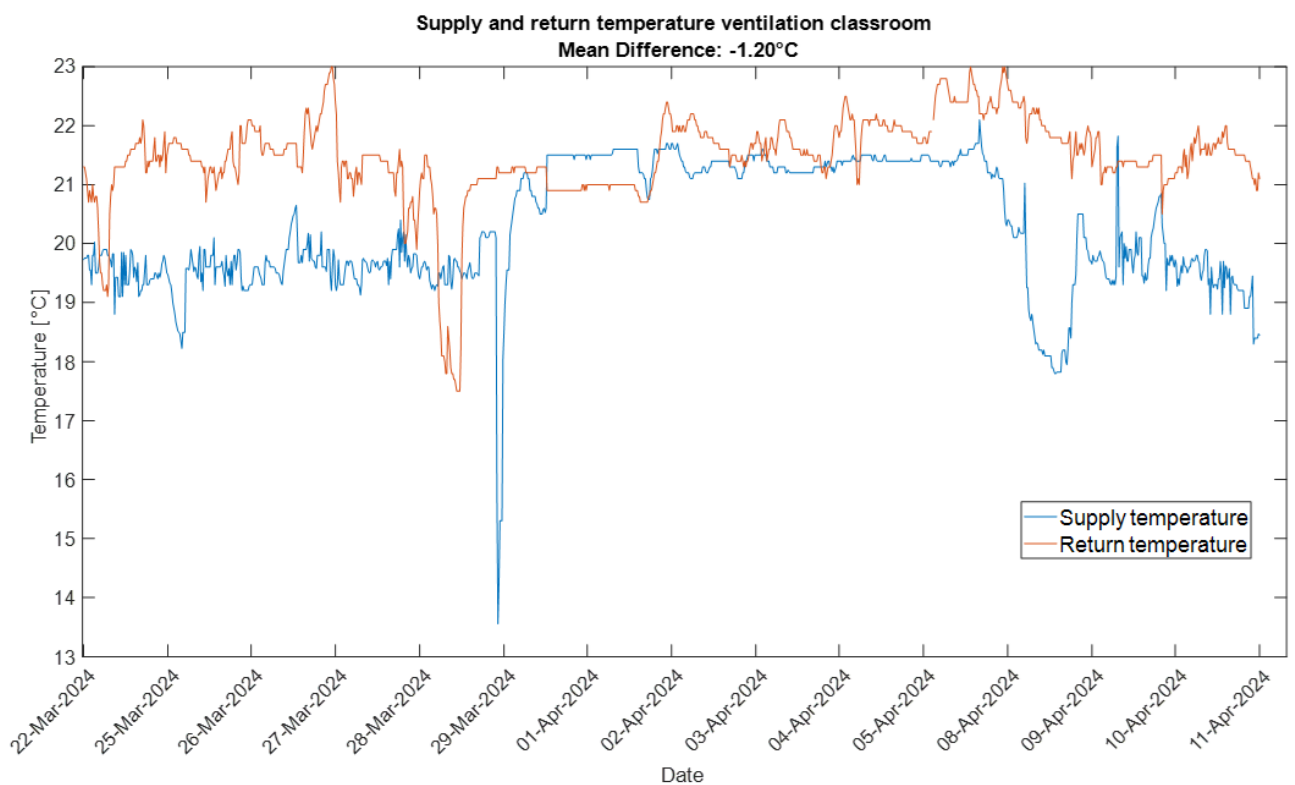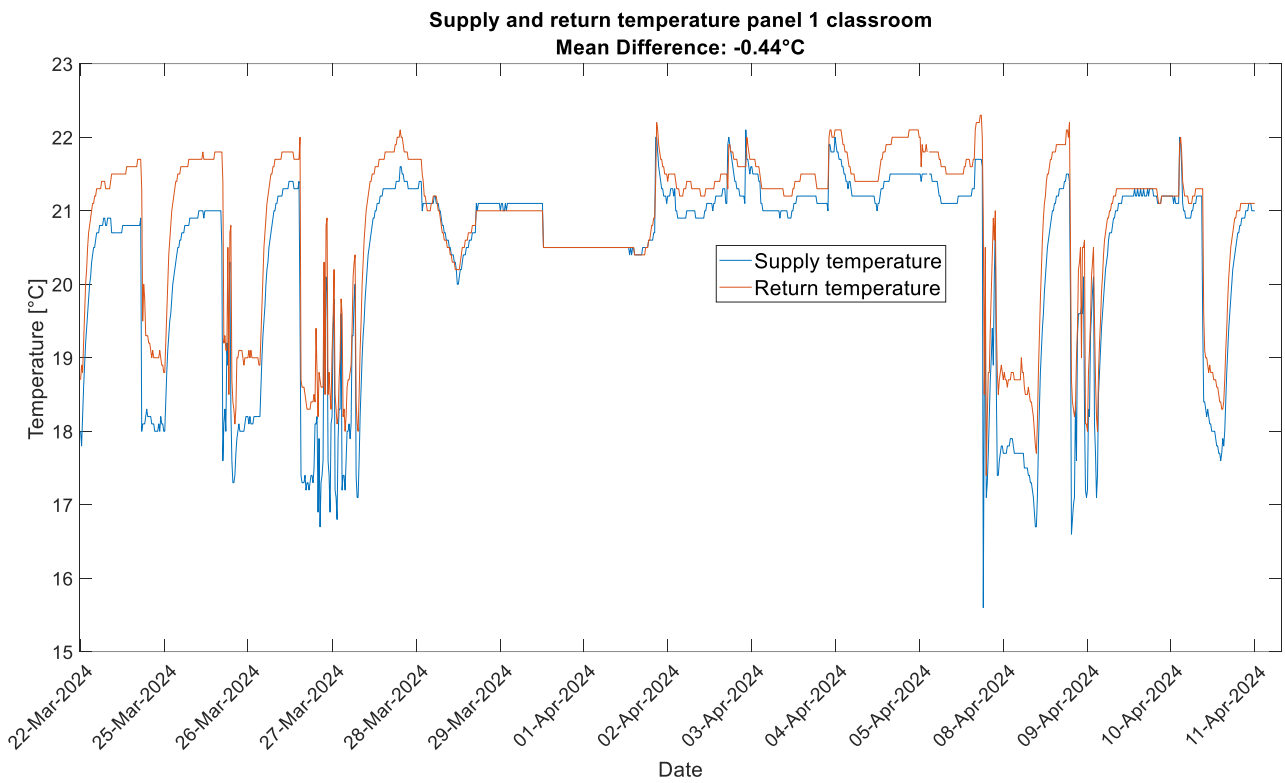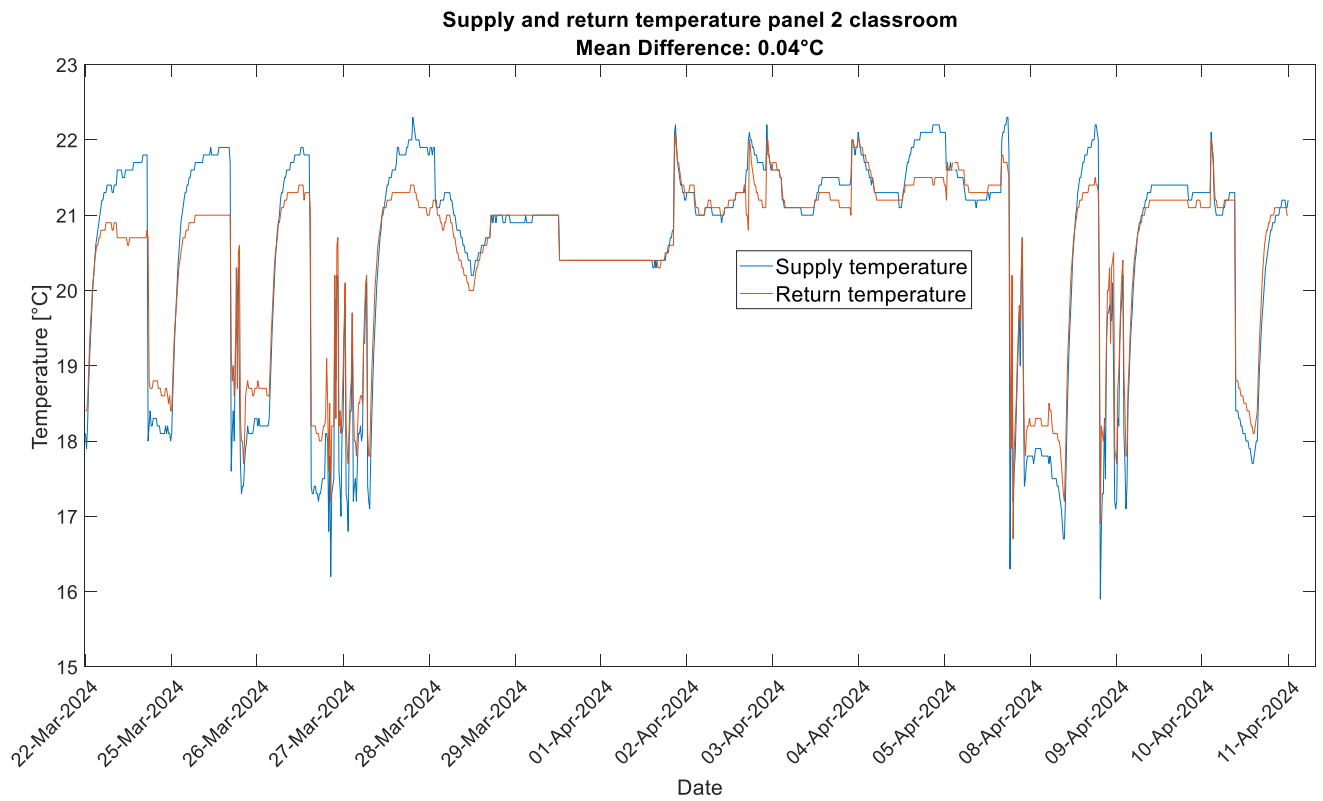

Figure 82: supply and return of the ventilation classroom

Figure 83: supply and return temperature panel 1 classroom

Figure 84: supply and return temperature panel 2 classroom

Appendix E: Graph interpolation supply temperature ventilation
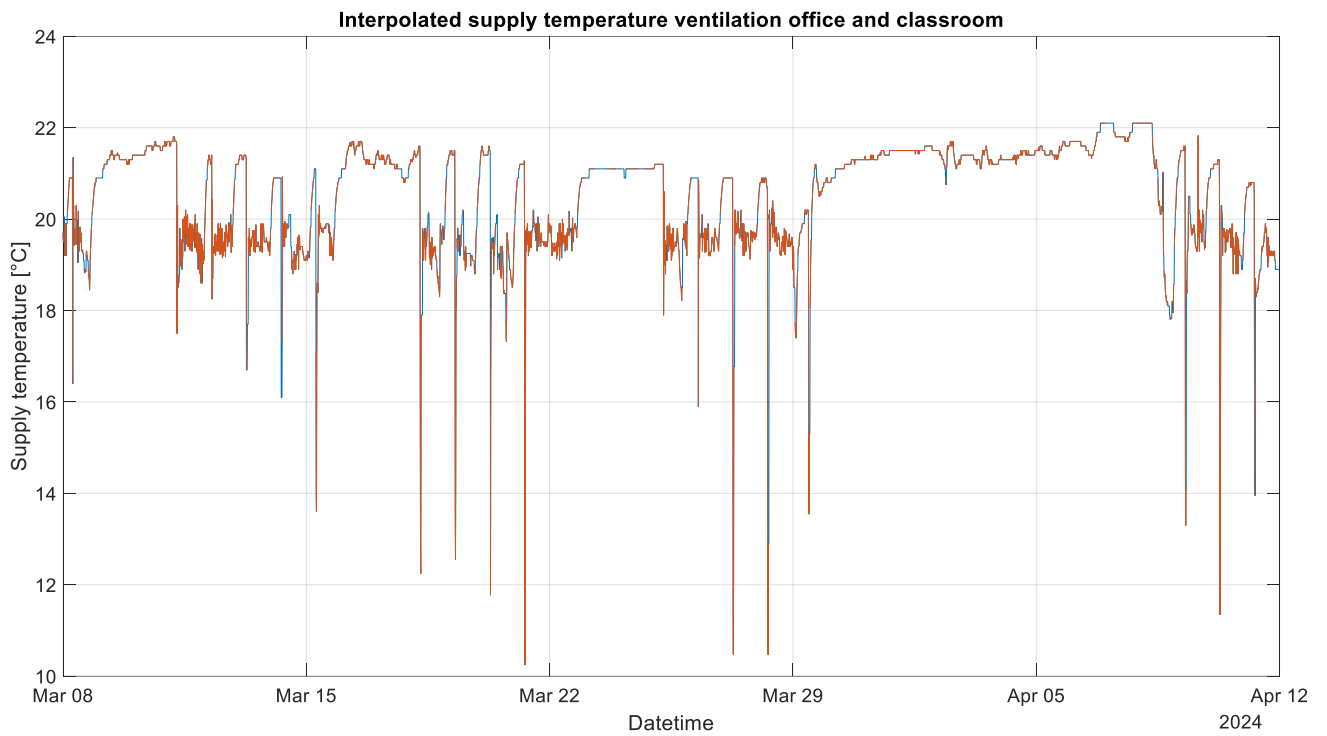
Figure 872: interpolated supply temperature ventilation system, same for office and classroom

Appendix F: Ranges thermal demand and indoor air temperature
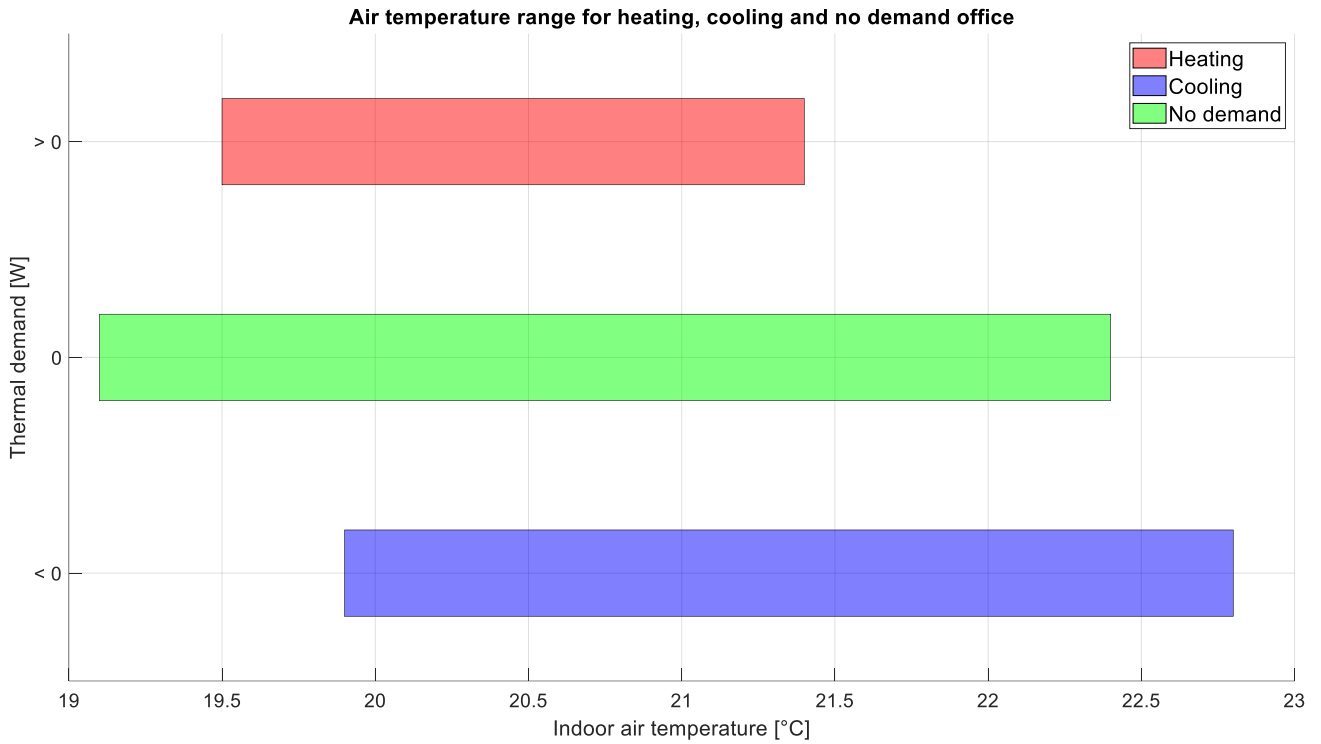
Figure 873: Air temperature ranges for heating, cooling and no thermal demand office
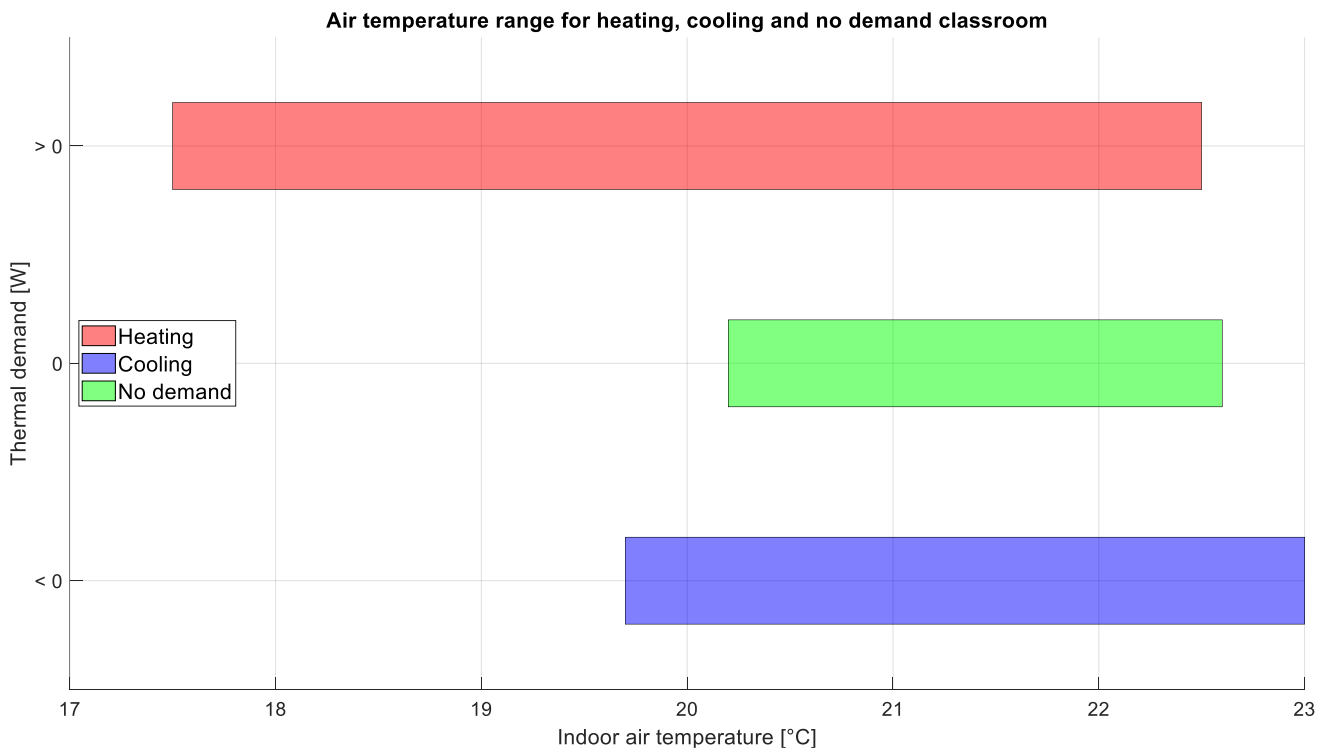
Figure 874: Air temperature ranges for heating, cooling and no thermal demand classroom

Appendix G: Ranges thermal demand outdoor temperature
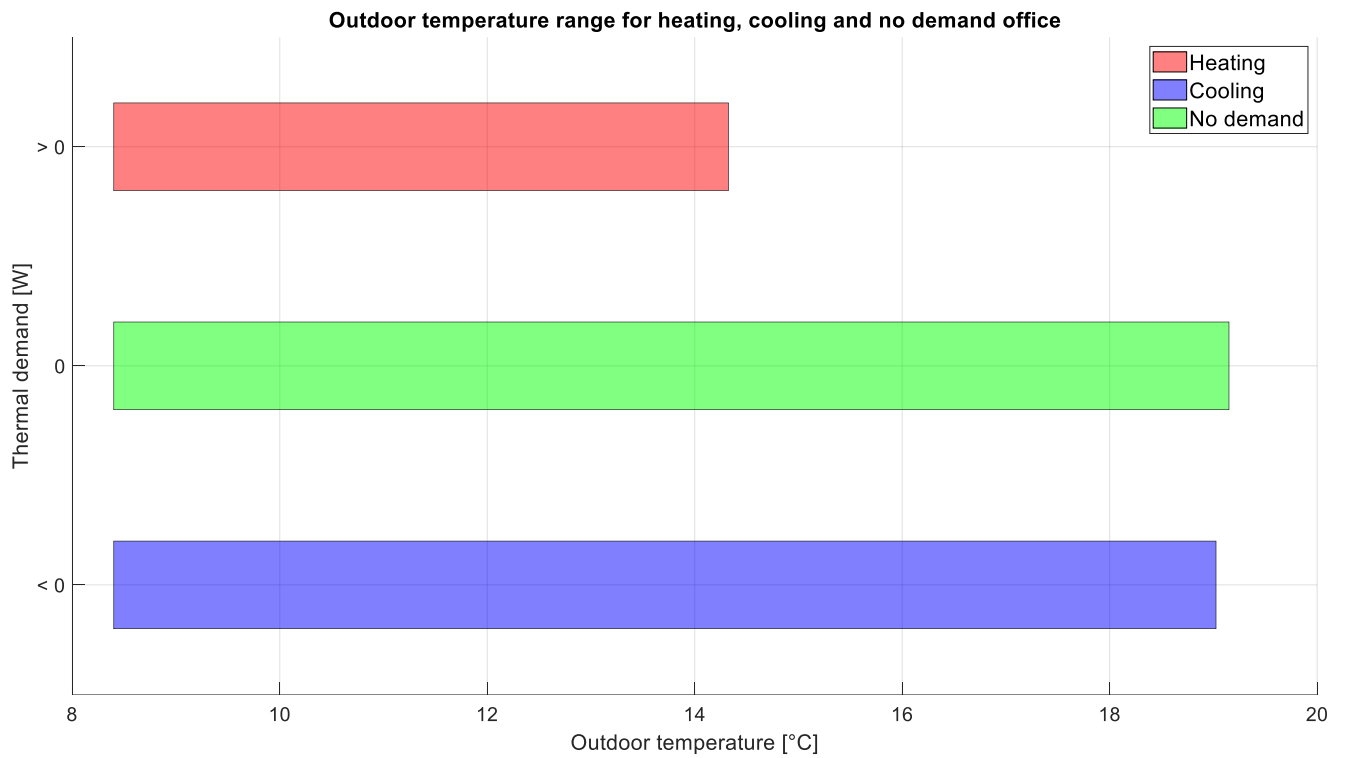
Figure 875: outdoor temperatures ranges for heating, cooling, and no thermal demand in the office
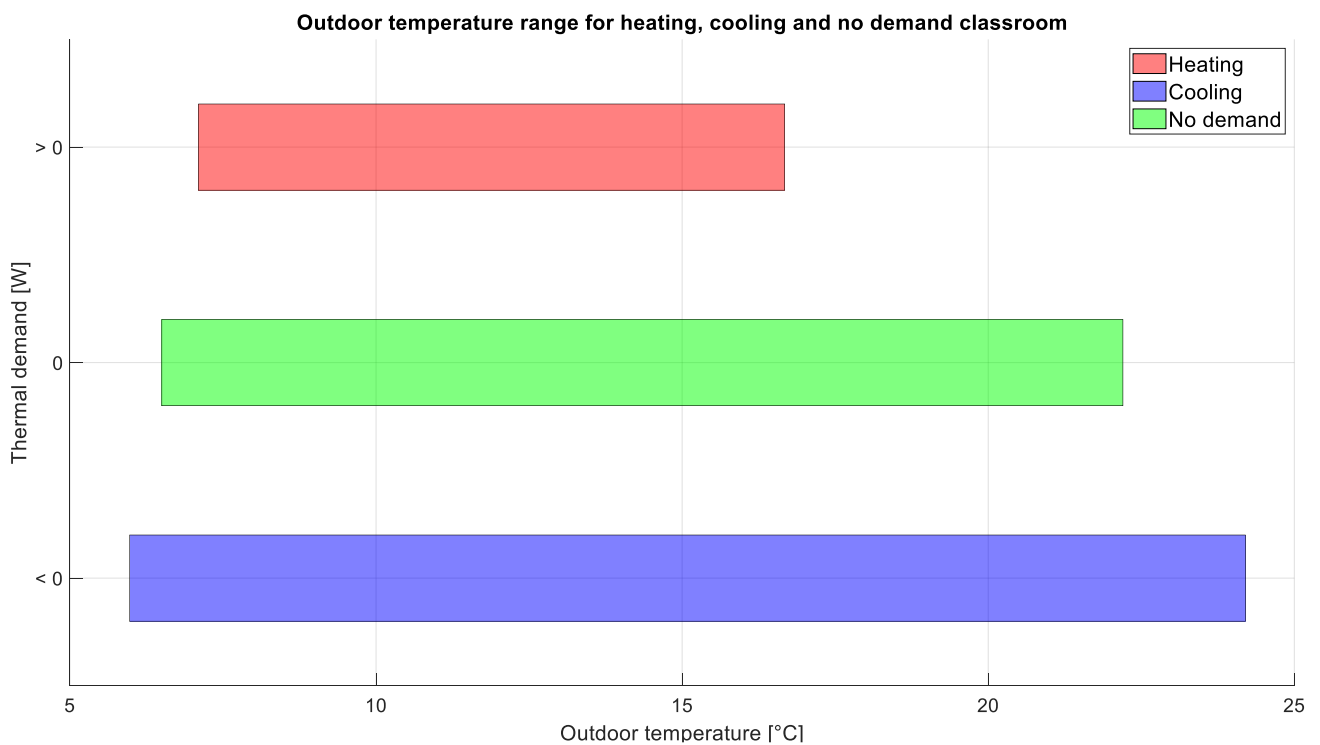
Figure 876: outdoor temperatures ranges for heating, cooling, and no thermal demand in the classroom

Appendix H: Scatterplots independent variables and thermal demand

Indoor air temperature

Figure 77: Scatterplot of the thermal demand against the indoor air temperature in the office
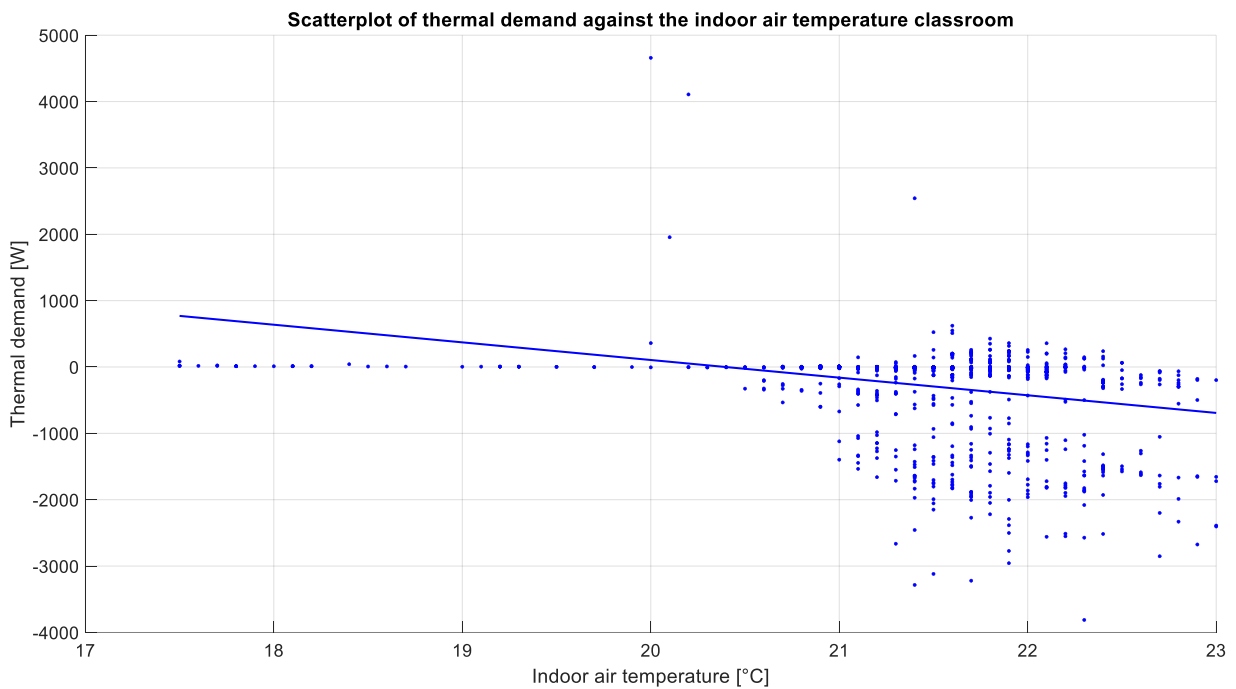
Figure 78: Scatterplot of the thermal demand against the indoor air temperature in the classroom

Outdoor temperature
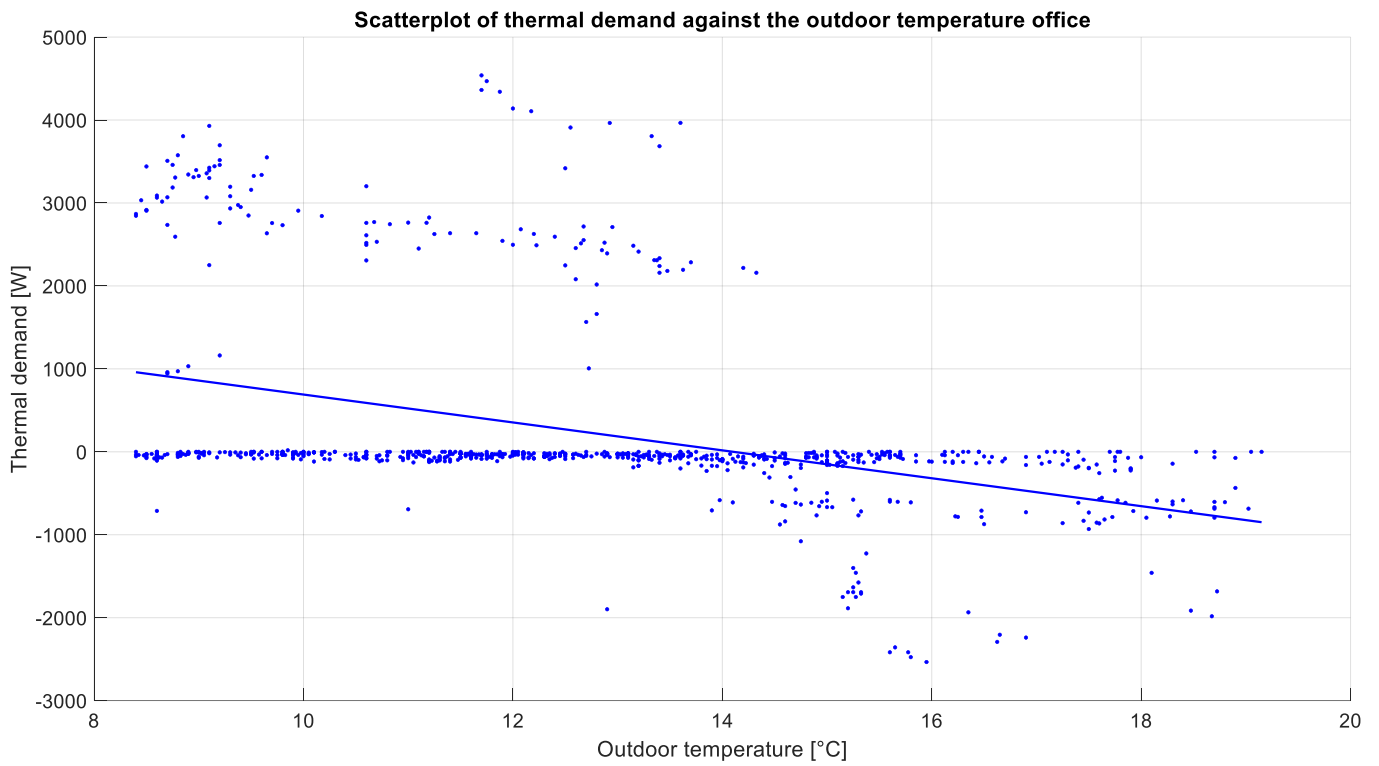
Figure 79: Scatterplot of the thermal demand against the outdoor temperature in the office
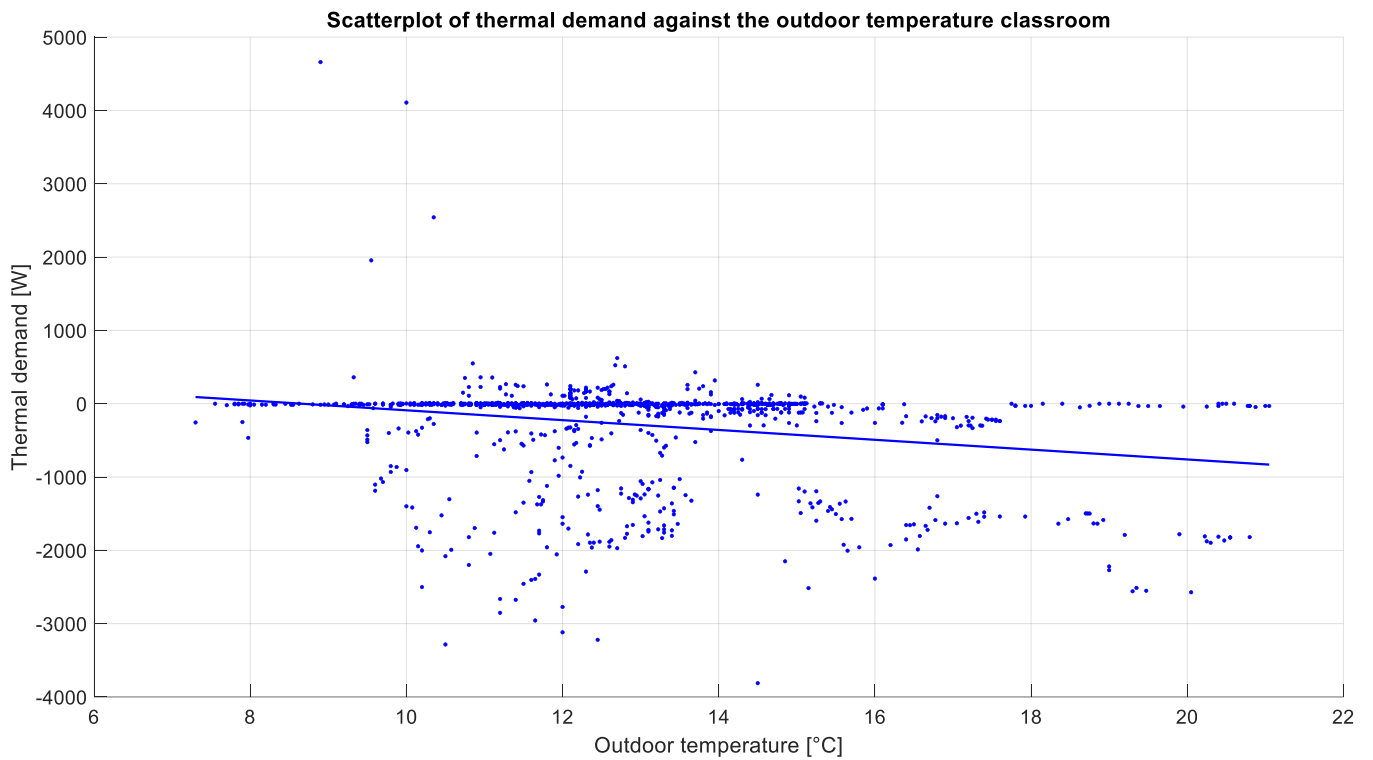
Figure 80: Scatterplot of the thermal demand against the outdoor temperature in the classroom

Indoor surface temperatures
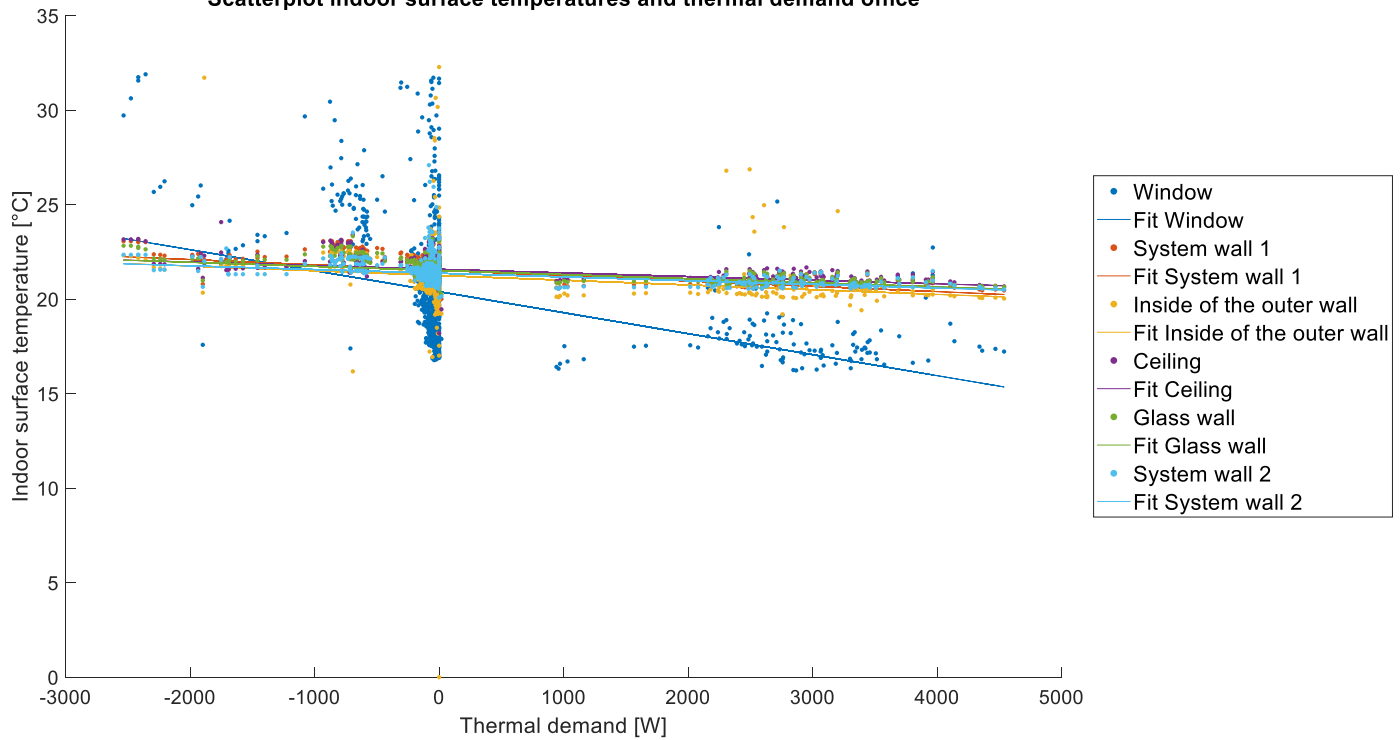
Figure 81: Scatterplot of the thermal demand against the indoor surface temperatures in the office

Figure 82: Scatterplot of the thermal demand against the indoor surface temperatures in the classroom

Internal heat gains
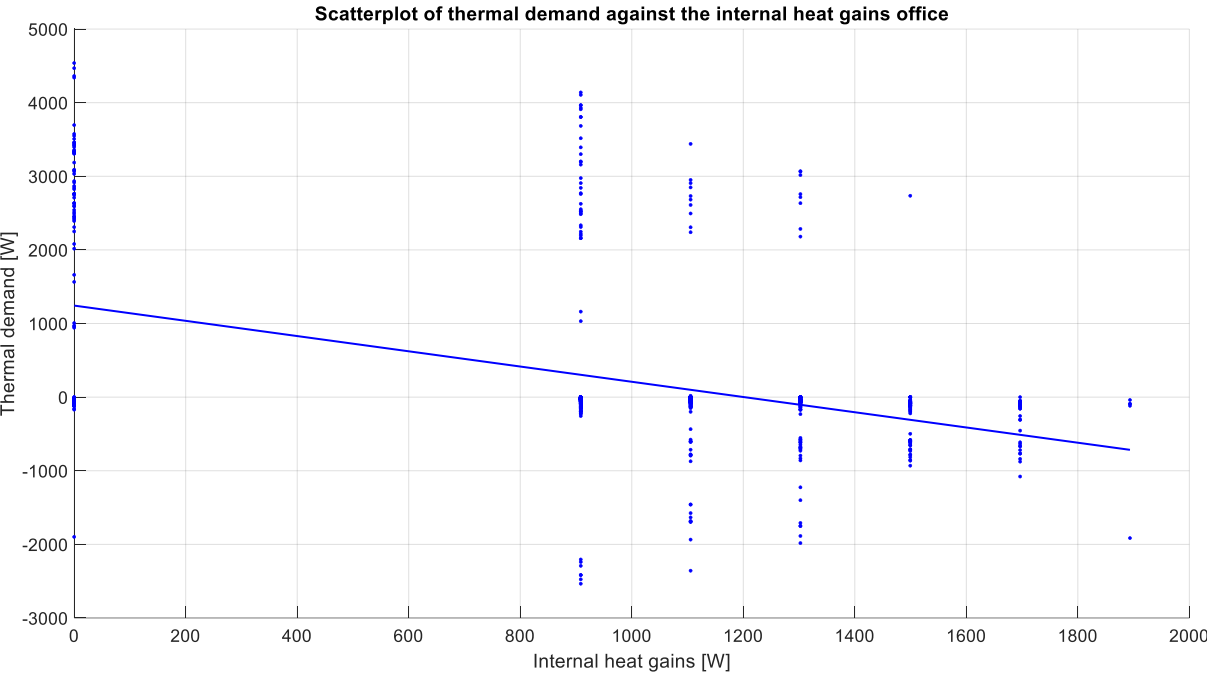
Figure 83: Scatterplot of the thermal demand against the internal heat gains in the office
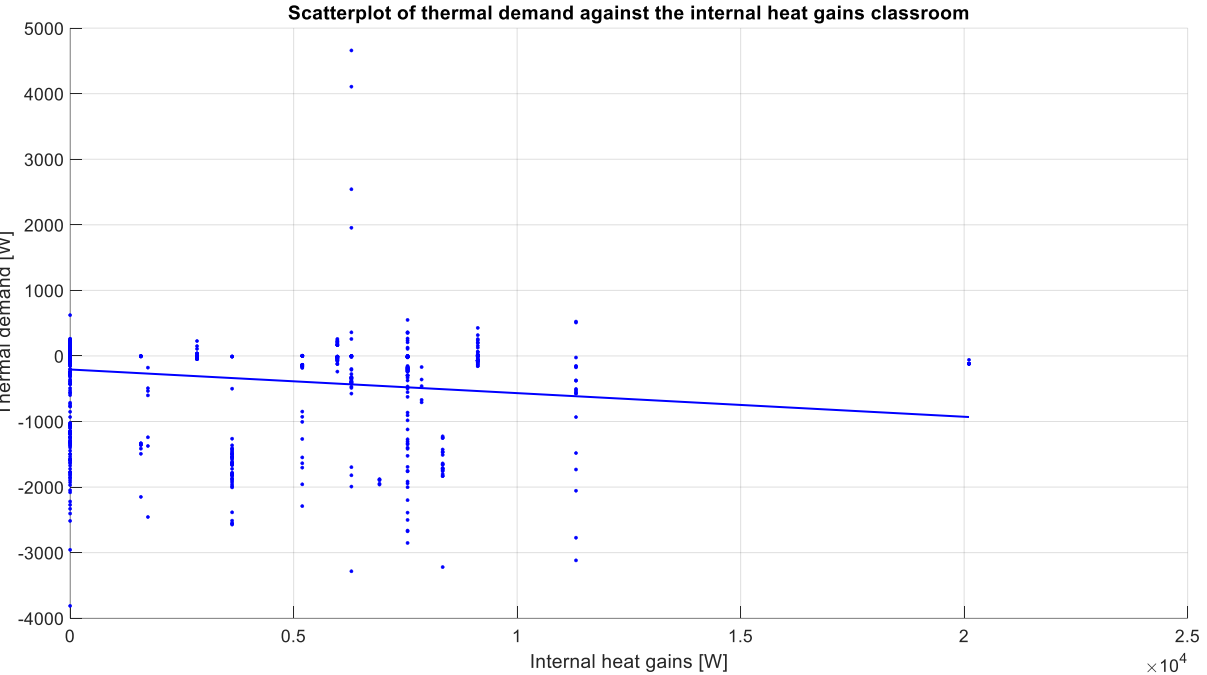
Figure 84: Scatterplot of the thermal demand against the internal heat gains in the classroom

Solar light intensity
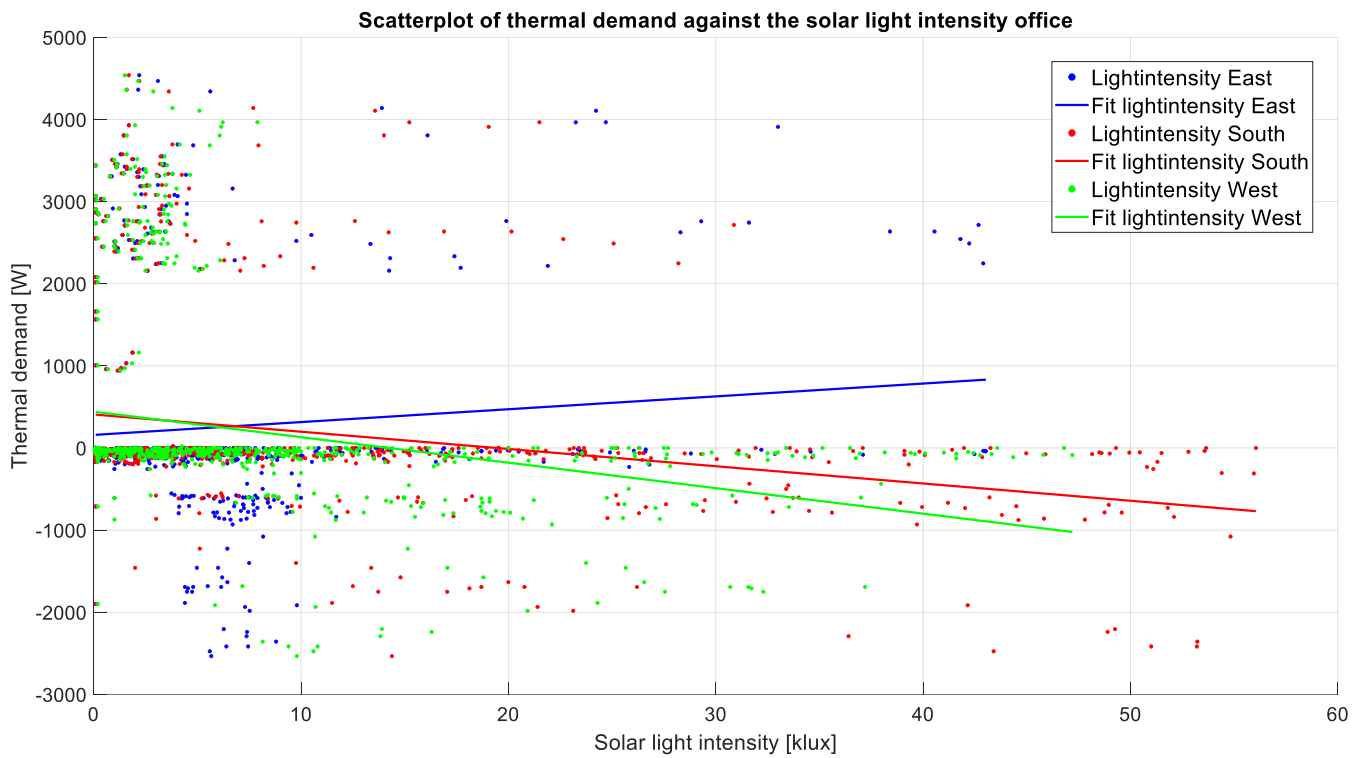
Figure 85: Scatterplot of the thermal demand against the solar light intensities in the office
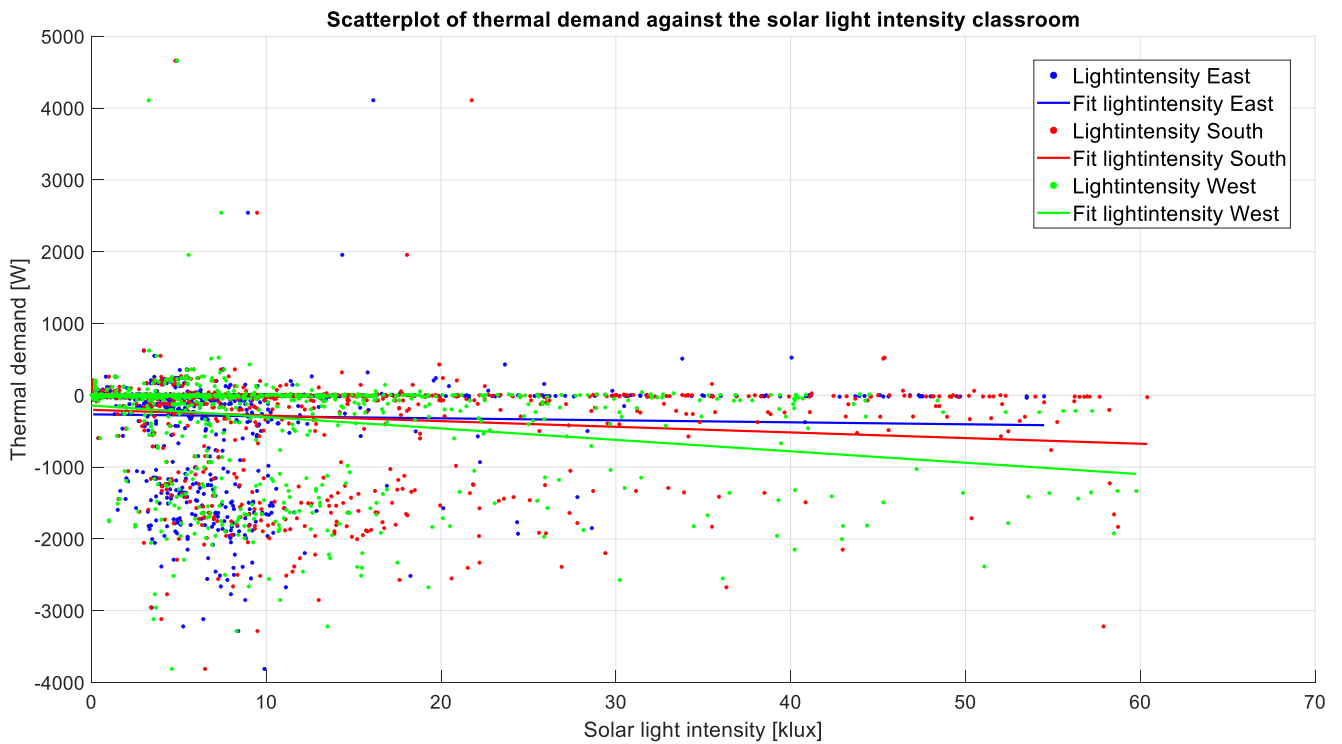
Figure 86: Scatterplot of the thermal demand against the solar light intensities in the classroom
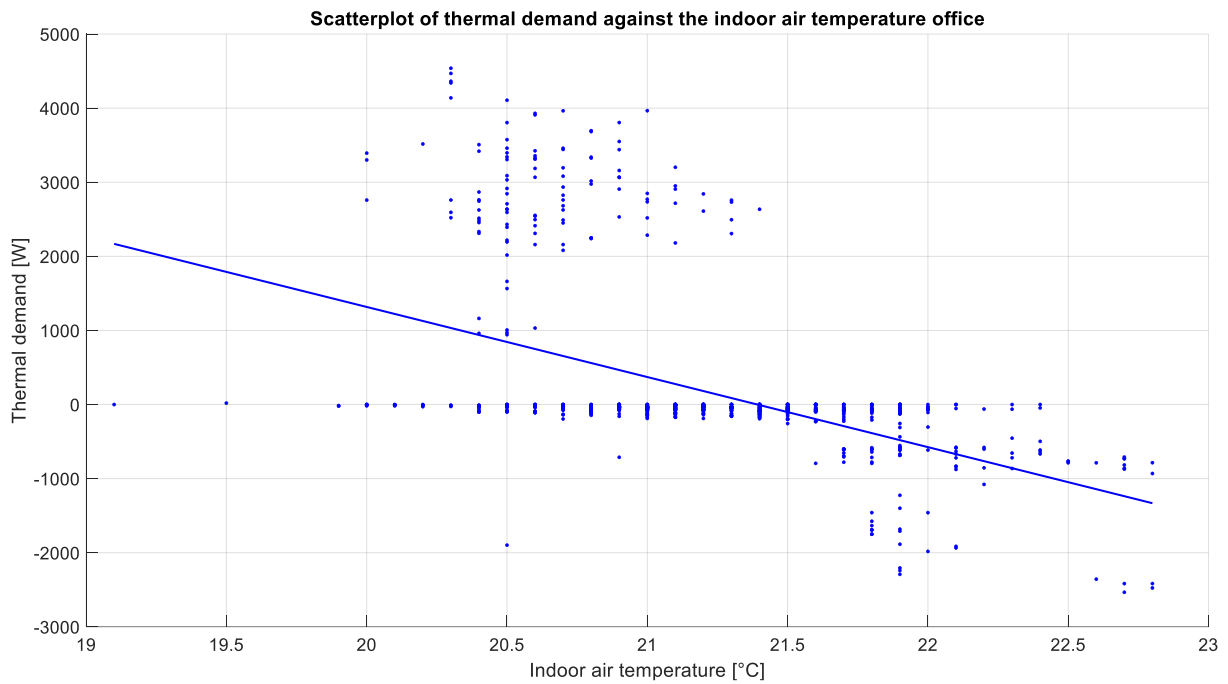
Windspeed
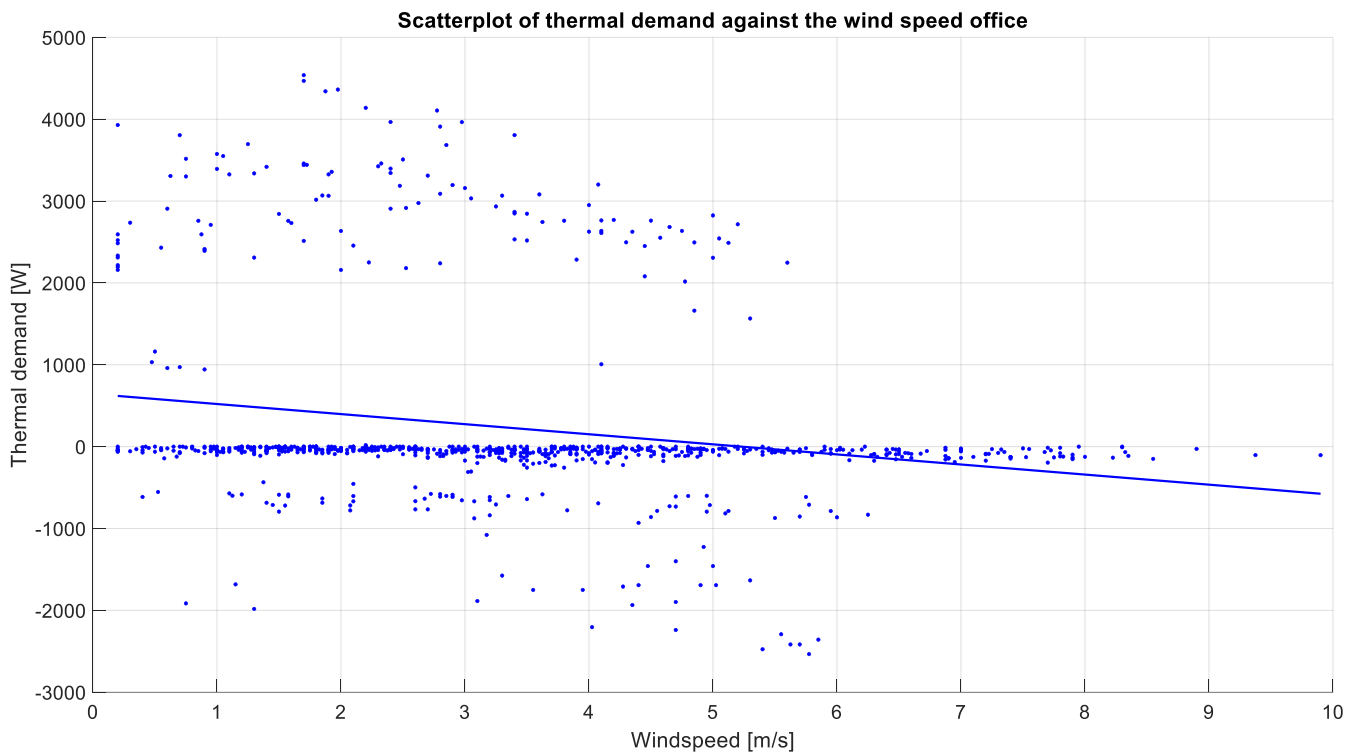
Figure 87: Scatterplot of the thermal demand against the windspeed in the office
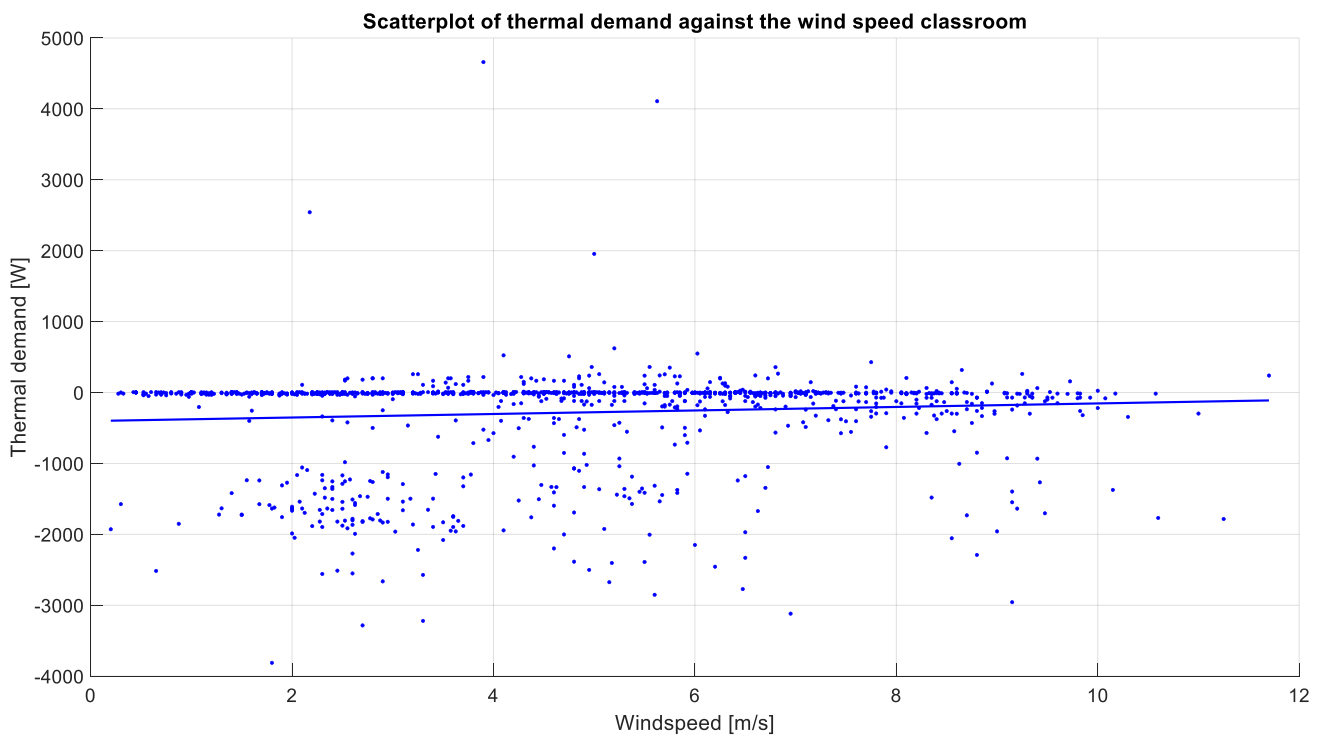
Figure 88: Scatterplot of the thermal demand against the windspeed in the classroom

Appendix I: Scatterplots correlation independent variables
Indoor surface temperatures and air temperatures
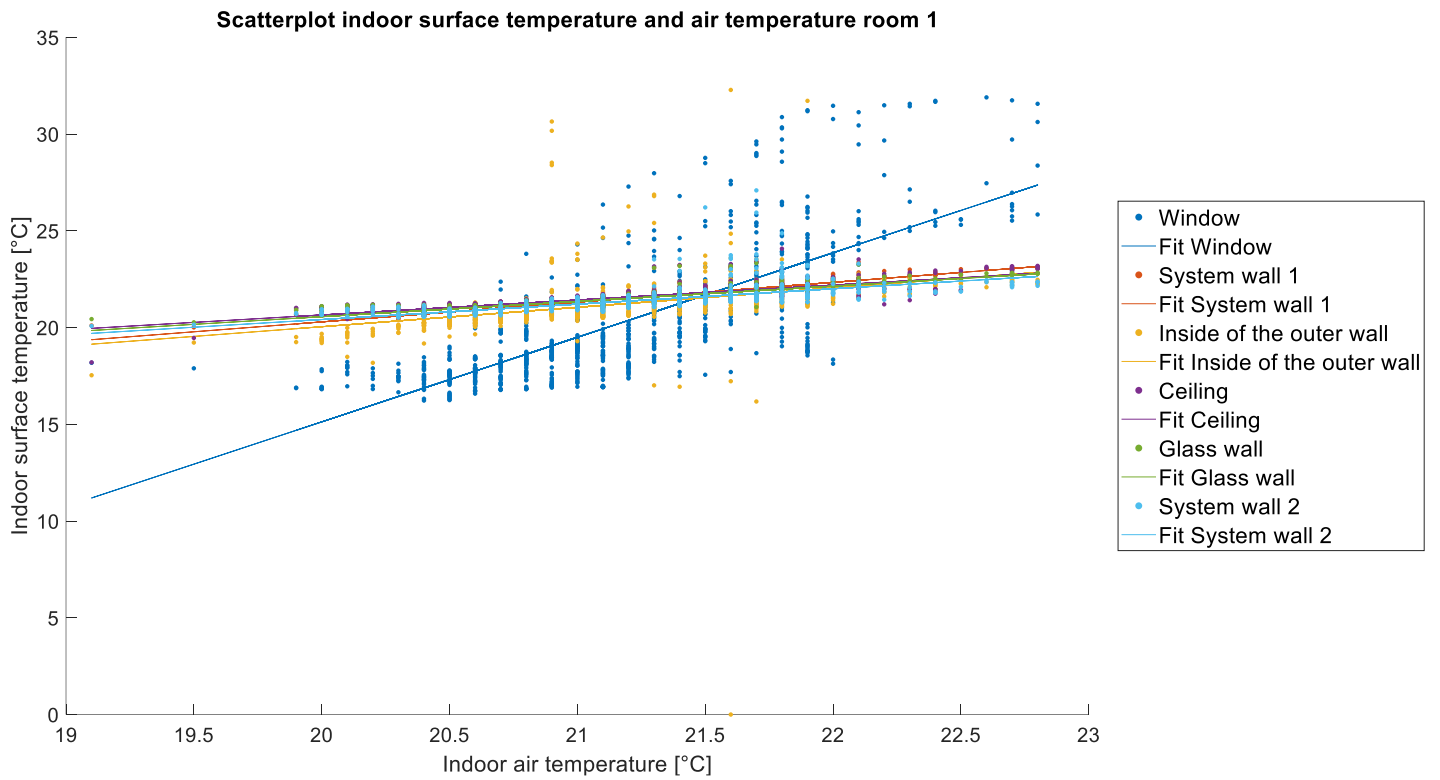
Figure 89: Scatterplot of the indoor surface temperatures against the indoor air temperature in the office
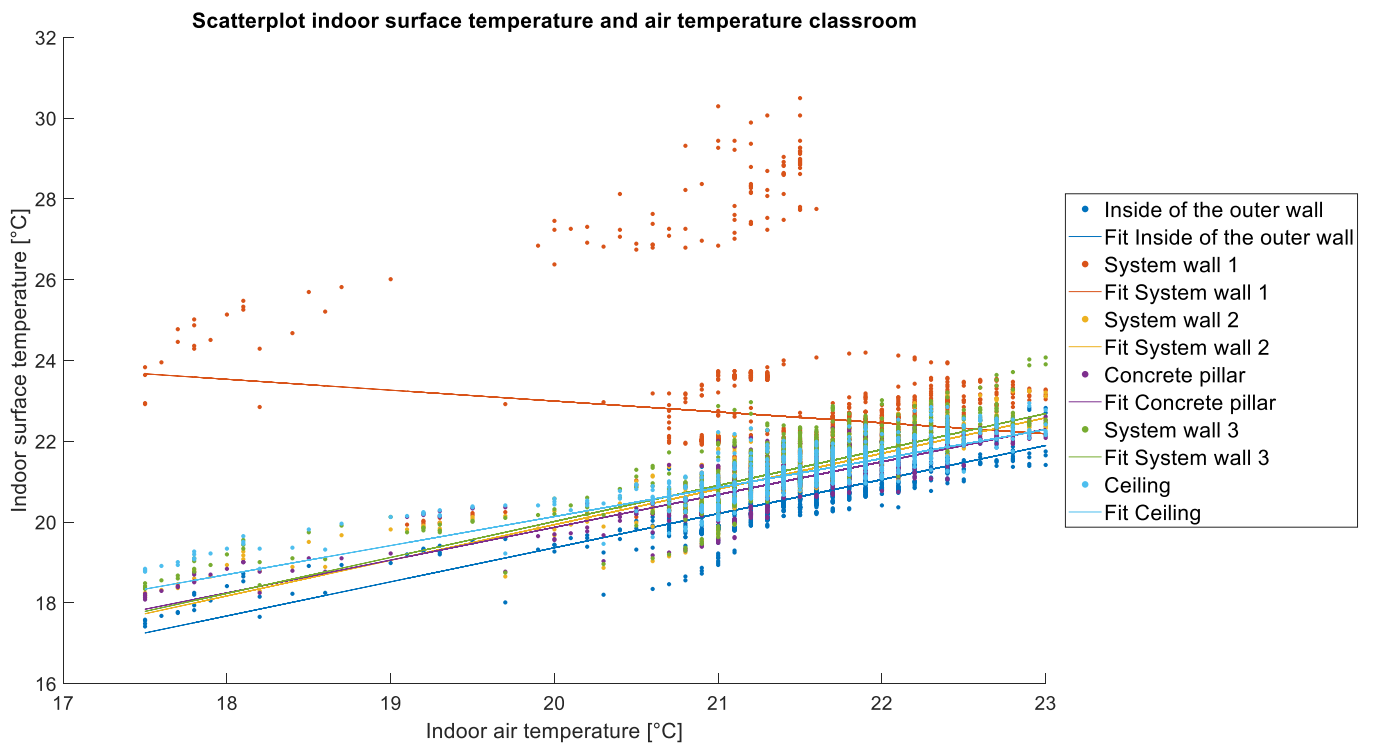
Figure 90: Scatterplot of the indoor surface temperatures against the indoor air temperature in the classroom

Indoor surface temperatures and outdoor temperature
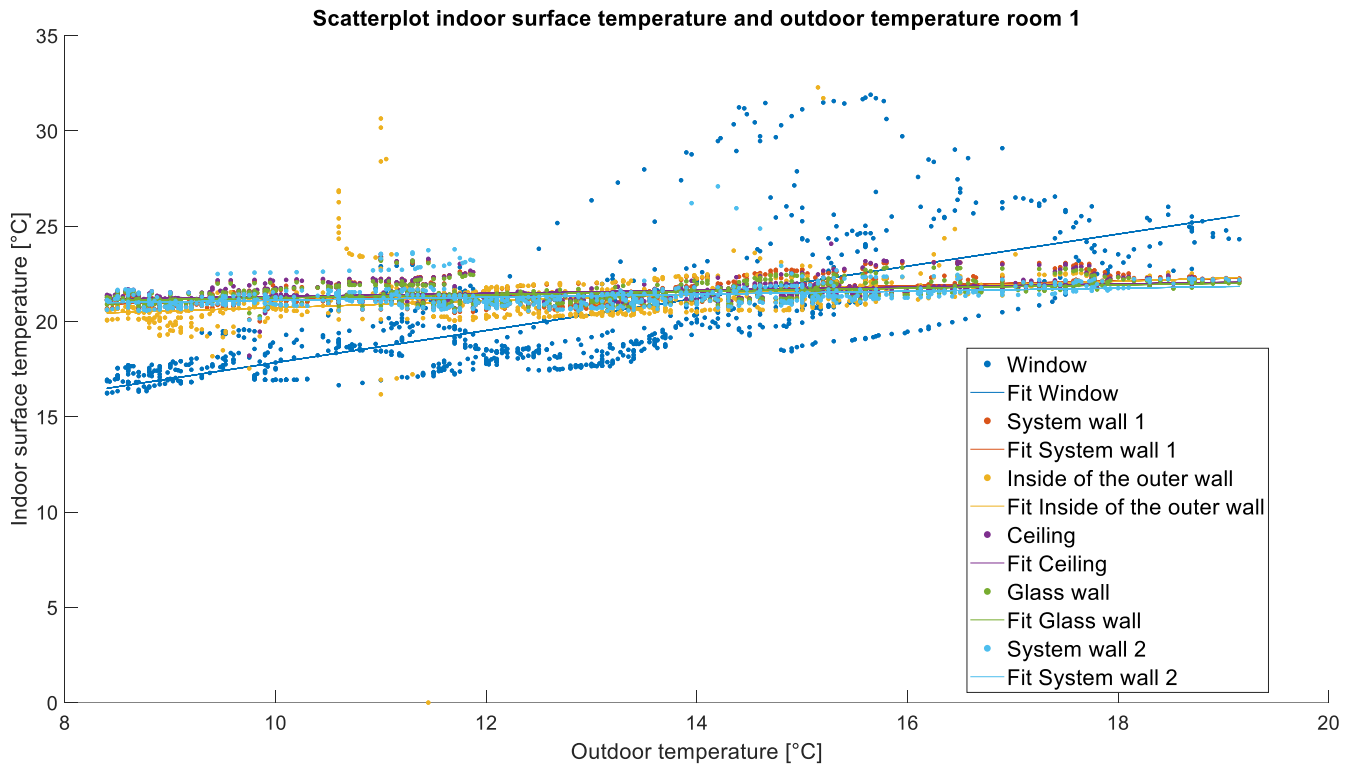
Figure 91: Scatterplot of the indoor surface temperatures against the outdoor temperature in the office
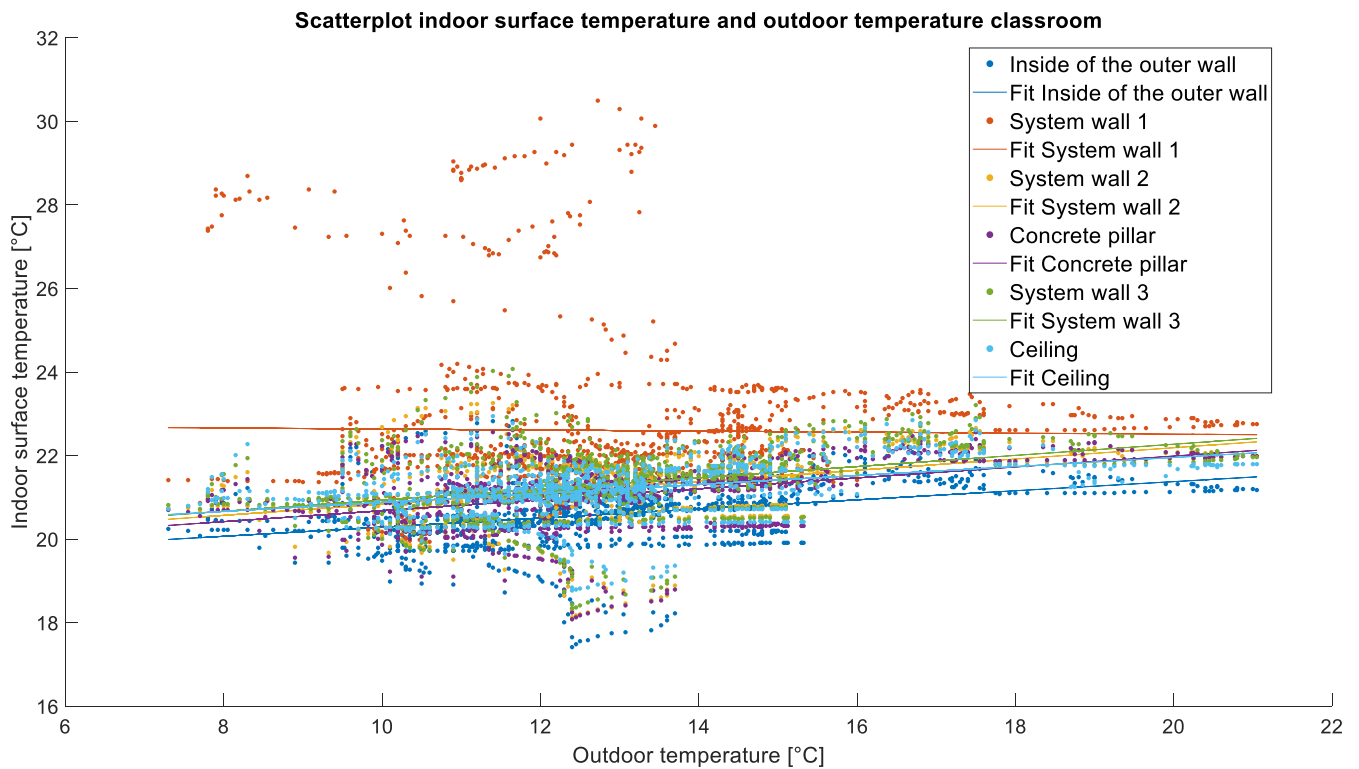
Figure 92: Scatterplot of the indoor surface temperatures against the outdoor temperature in the classroom

Appendix J: Normalized thermal demand graphs to determine delay
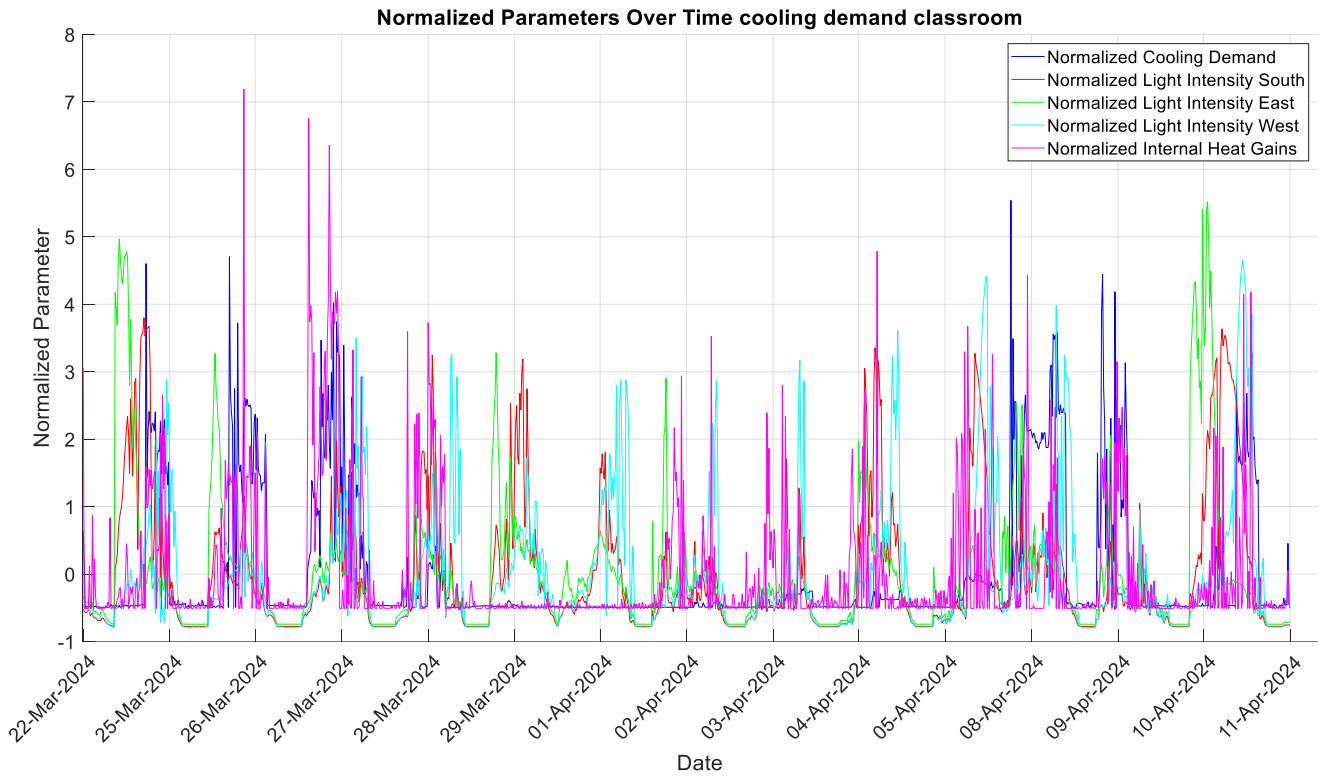
Figure 93: normalized parameters: solar light intensity and the internal heat gains, and the cooling demand in the classroom
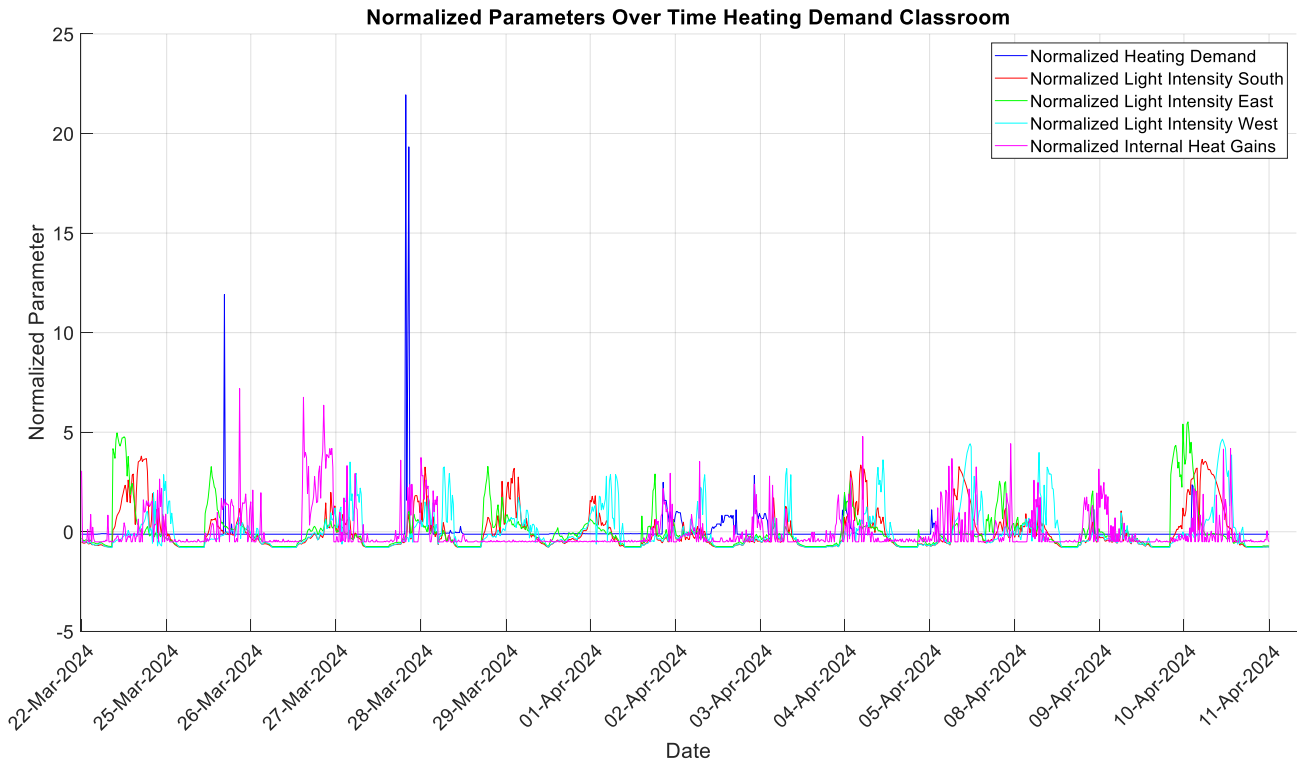
Figure 107: normalized parameters: solar light intensity and the internal heat gains, and the cooling demand in the classroom
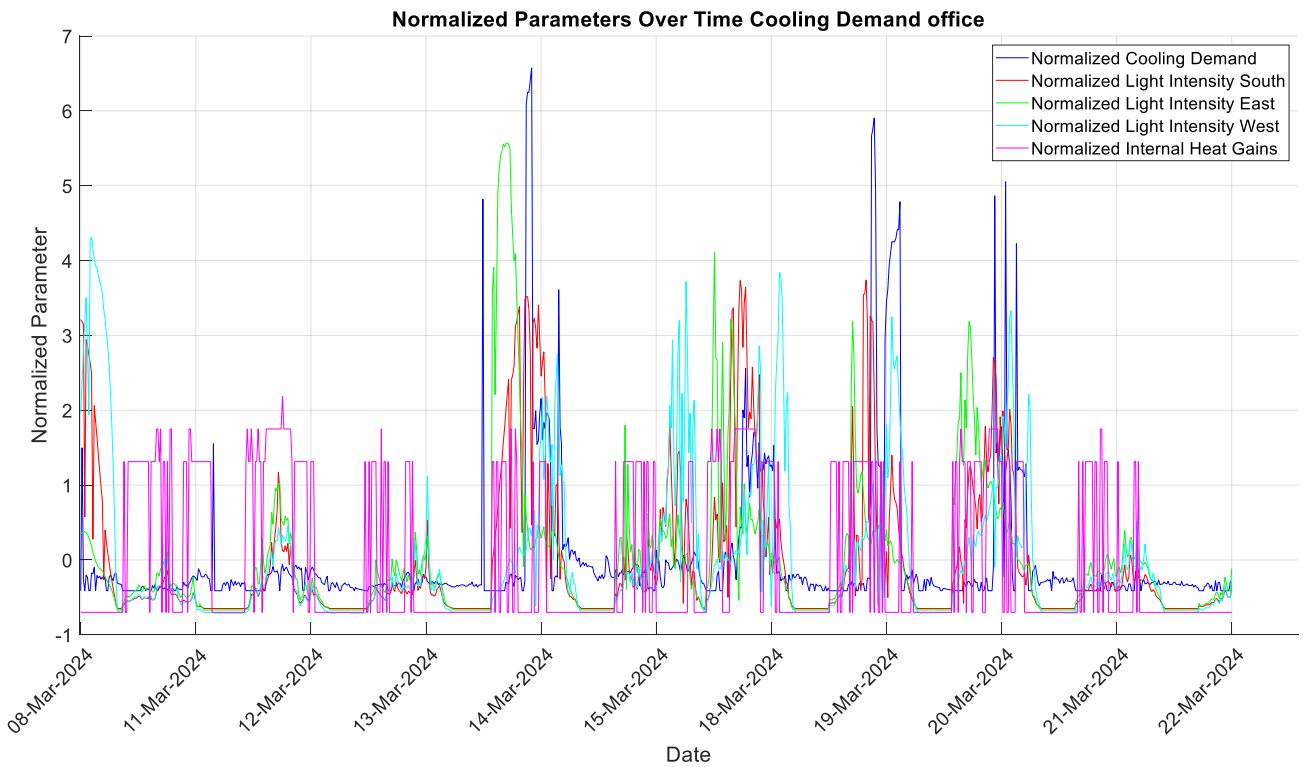
Figure 94: normalized parameters: solar light intensity and the internal heat gains, and the cooling demand in the office
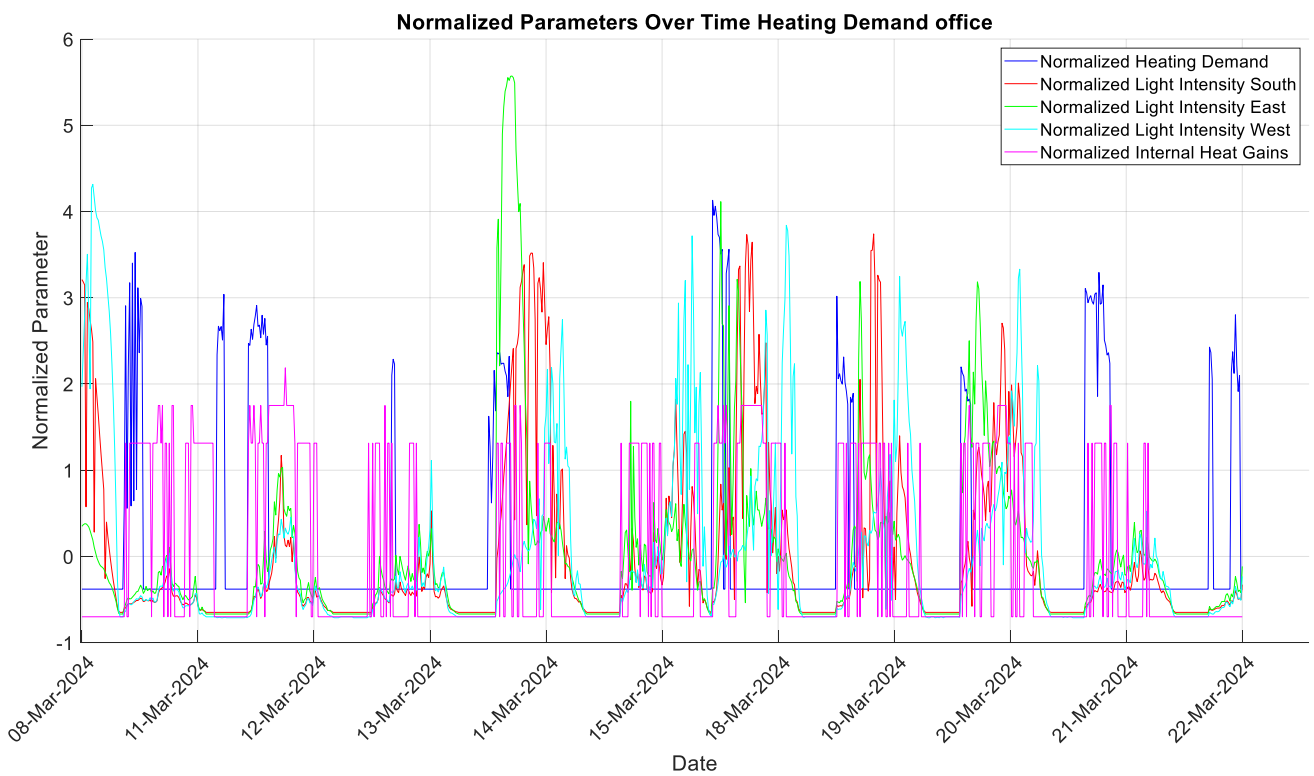
Figure 95: normalized parameters: solar light intensity and the internal heat gains, and the cooling demand in the classroom

Appendix K: Coefficient values model C
Model C_1

Table 40: coefficient values of the independent variables, with the corresponding T-statistics and P-value of the thermal demand Equations derived from model C_1 for the office

Coefficient	Independent variable	Thermal demand		
		Value	T-statistics	P-value
C_0		24826.00	13.53	2.29e-37
C_1	$T_{outdoor}(t)$	-79.36	-4.37	1.41e-5
C_2	$T_{s,window}(t)$	39.90	2.40	0.02
C_3	$T_{s,systemwall1}(t)$			
C_4	$T_{s,insideoftheouterwall}(t)$	-82.25	-2.06	0.04
C_5	$T_{s,ceiling}(t)$			
C_6	$T_{s,glasswall}(t)$	-1056.80	-10.74	4.46e-25
C_7	$T_{s,systemwall2}(t)$			

Table 41: coefficient values of the independent variables, with the corresponding T-statistics and P-value of the thermal demand Equations derived from model C_1 for the classroom

Coefficient	Independent variable	Thermal demand		
		Value	T-statistics	P-value
C_0		676.93	1.26	0.21
C_1	$T_{outdoor}(t)$	-26.00	-3.22	1.32e-3
C_2	$T_{s,insideoftheouterwall}(t)$	-308.01	-2.56	0.01
C_3	$T_{s,systemwall1}(t)$			
C_4	$T_{s,systemwall2}(t)$	794.71	5.67	1.92e-8
C_5	$T_{s,concretepillar}(t)$	-648.05	-6.61	6.44e-11
C_6	$T_{s,systemwall3}(t)$	-1400.70	-14.09	3.58e-41
C_7	$T_{s,ceiling}(t)$	1526.10	17.14	9.30e-58

Model C_2

Table 382: coefficient values of the independent variables, with the corresponding T-statistics and P-value of the thermal demand Equations derived from model C_2 for the office

Coefficient	Independent variable	Thermal demand		
		Value	T-statistics	P-value
C_0		18932.00	14.68	6.99e-43
C_1	$T_{outdoor}(t)$	-59.56	-4.09	4.84e-5

C_2	$T_{indoorair}(t)$	-851.85	-13.08	3.29e-35
C_3	$Windspeed(t)$	-66.09	-3.97	7.97e-5
C_4	$Solar\ light\ intensity\ south(t)$			
C_5	$Solar\ light\ intensity\ east(t)$	34.16	7.14	2.26e-12
C_6	$Solar\ light\ intensity\ west(t)$			
C_7	$Q_{internal}(t)$	0.23	3.13	1.8e-3

Table 43: coefficient values of the independent variables, with the corresponding T-statistics and P-value of the thermal demand Equations derived from model C₂ for the classroom

		Thermal demand		
Coefficient	Independent variable	Value	T-statistics	P-value
C_0		3560.90	7.09	2.60e-12
C_1	$T_{outdoor}(t)$	-51.464	-5.88	5.61e-9
C_2	$T_{indoorair}(t)$	-151.20	-6.05	2.04e-9
C_3	$Windspeed(t)$	78.07	9.74	1.94e-21
C_4	$Solar\ light\ intensity\ south(t)$	-4.39	-2.54	0.01
C_5	$Solar\ light\ intensity\ east(t)$			
C_6	$Solar\ light\ intensity\ west(t)$	-6.51	-2.93	3.46e-3
C_7	$Q_{internal}(t)$	-0.27	-14.84	5.19e-45

Appendix L: Coefficient values model D

Table 44: coefficient values of the independent variables, with the corresponding T-statistics and P-value of the heating and cooling Equations derived from model D for the office

Coefficient	Independent variable	Heating Delay (n)	Value	T-statistics	P-value	Cooling Delay (n)	Value	T-statistics	P-value
C_0			35.99	0.43	0.66		- 7.82	-0.48	0.63
C_1	$T_{outdoor} - T_{indoorair}(t)$								
C_2	$V_{wind} \cdot (T_{outdoor} - T_{indoorair})(t)$		12.43	5.66	2.19e-08				
C_3	Solar light intensity south (t)		11.96	2.56	0.01		10.73	8.29	5.89e-16
C_4	Solar light intensity east (t)		43.95	7.76	3.10e-14		-7.86	-3.86	1.25e-4
C_5	Solar light intensity west (t)								
C_6	$Q_{internal}(t)$		0.32	4.34	1.60e-05				
C_7	$T_{indoorair} - T_{s,window}(t)$		167.29	7.97	6.32e-15				
C_8	$T_{indoorair} - T_{s,systemwall1}(t)$		39.45	3.38	7.64-4		10.48	8.27	6.88e-16
C_9	$T_{indoorair} - T_{s,insideoftheoutewall}(t)$				8.82		3.73	2.09e-4	
C_{10}	$T_{indoorair} - T_{s,ceiling}(t)$								
C_{11}	$T_{indoorair} - T_{s,glasswall}(t)$								
C_{12}	$T_{indoorair} - T_{s,systemwall2}(t)$								
C_{13}	Solar light intensity south $(t-n)$					17			
C_{14}	Solar light intesnity east $(t-n)$					25			

Coefficient	Independent variable	Delay (n)	Value	T-statistics	P-value	Delay (n)	Value	T-statistics	P-value
C_{15}	Solar light intensity west $(t-n)$					1			
C_{16}	$Q_{internal}(t-n)$					20			
C_{17}	$T_{indoorair} - T_{s,window}(t-n)$	21				71			

Table 45: coefficient values of the independent variables, with the corresponding T-statistics and P-value of the heating and cooling Equations derived from model D for the classroom

		Heating				Cooling			
Coefficient	Independent variable	Delay (n)	Value	T-statistics	P-value	Delay (n)	Value	T-statistics	P-value
C_0			-11.56	-0.85	0.39		596.86	5.19	2.53e-7
C_1	$T_{outdoor} - T_{indoorair}(t)$						30.95	3.568	3.78e-4
C_2	$V_{wind} \cdot (T_{outdoor} - T_{indoorair})(t)$						2.47	3.0254	2.55e-3
C_3	Solar light intensity south (t)								
C_4	Solar light intensity east (t)						5.80	3.06	2.33-3
C_5	Solar light intensity west (t)						-3.04	-2.07	0.04
C_6	$Q_{internal}(t)$		0.02	2.07	0.04		0.11	6.95	6.86e-12
C_7	$T_{indoorair} - T_{s,systemwall1}(t)$								
C_8	$T_{indoorair} - T_{s,insideoftheouterwall}(t)$						-311.31	-3.05	2.33e-3
C_9	$T_{indoorair} - T_{s,ceiling}(t)$		-168.27	-4.89	1.19e-06		1197.70	17.33	1.66e-58
C_{10}	$T_{indoorair} - T_{s,systemwall2}(t)$		225.80	3.28	1.07e-3		836.48	6.35	3.34e-10
C_{11}	$T_{indoorair} - T_{s,concretepillar}(t)$		83.92	2.23	0.03		-455.97	-4.89	1.17e-06
C_{12}	$T_{indoorair} - T_{s,systemwall3}(t)$		-136.00	-2.96	3.30e-3		-1218.30	-11.88	2.17e-30
C_{13}	Solar light intensity south $(t-n)$					144	7.36	6.40	2.51e-10

Coefficient	Independent variable	Delay (n)	Value	T-statistics	P-value	Delay (n)	Value	T-statistics	P-value
C_{14}	Solar light intesnity east $(t-n)$					33	5.47	2.83	4.71-e3
C_{15}	$Q_{internal}(t-n)$					1			

Appendix M: Coefficient values model E

Table 46: coefficient values of the independent variables, with the corresponding T-statistics and P-value of the heating and cooling Equations derived from model E for the office

		Heating				Cooling			
Coefficient	Independent variable	Delay (n)	Value	T-statistics	P-value	Delay (n)	Value	T-statistics	P-value
C_0			13370.00	2.69	7.27e-3		-3581.00	-3.21	1.39e-3
C_1	$T_{outdoor}(t)$		77.51	3.92	9.64e-05				
C_2	$T_{indoorair}(t)$						-31.24	-3.32	9.33e-4
C_3	$T_{s,window}(t)$		-86.92	-3.48	5.31e-4		-497.92	-4.68	3.51e-6
C_4	$T_{s,systemwall1}(t)$		-974.32	-3.14	1.76e-3				
C_5	$T_{s,insideoftheouterwall}(t)$								
C_6	$T_{s,ceiling}(t)$		-436.95	-2.63	8.81e-3				
C_7	$T_{s,glasswall}(t)$		942.19	2.30	0.02		2424.10	5.05	5.71e-7
C_8	$T_{s,systemwall2}(t)$								
C_9	$V_{wind}(t)$								
C_{10}	Solar light intensity south (t)		11.45	2.58	0.01		6.49	3.68	2.47e-4
C_{11}	Solar light intensity east (t)		19.21	3.28	1.08e-3				
C_{12}	Solar light intensity west (t)						5.18	2.76	5.88e-3
C_{13}	$Q_{internal}(t)$		0.228	3.01	2.75e-3				
C_{14}	$T_{outdoor}(t-n)$	100				1			

175

Coefficient	Independent variable	Delay (n)	Value	T-statistics	P-value	Delay (n)	Value	T-statistics	P-value
C_{15}	$T_{indoorair}(t-n)$	100	-1179.40	-5.82	9.09e-09	1			
C_{16}	$T_{s,window}(t-n)$	117	173.42	5.80	1.01e-08	8	52.40	5.68	1.97e-8
C_{17}	$T_{s,systemwall1}(t-n)$	104	-2467.10	-7.22	1.41e-12	1			
C_{18}	$T_{s,ceiling}(t-n)$	117	-819.90	-5.62	2.84e-08	1			
C_{19}	$T_{s,glasswall}(t-n)$	101	3461.90	6.90	1.21e-11	1	-1997.00	-4.19	3.14e-5
C_{20}	$T_{s,systemwall2}(t-n)$	94	734.58	3.88	1.13e-4	12	221.00	7.69	5.04e-14
C_{21}	$V_{wind}(t-n)$					142			
C_{22}	Solar light intensity south $(t-n)$					17	8.58	5.85	7.44e-9
C_{23}	Solar light intensity east $(t-n)$					25			
C_{24}	Solar light intensity west $(t-n)$					1			
C_{25}	$Q_{internal}(t-n)$					20			

Table 47: coefficient values of the independent variables, with the corresponding T-statistics and P-value of the heating and cooling Equations derived from model E for the classroom

Coefficient	Independent variable	Heating				Cooling			
		Delay (n)	Value	T-statistics	P-value	Delay (n)	Value	T-statistics	P-value
C_0			25.57	0.10	0.92		-1413.30	-3.02	2.63e-3
C_1	$T_{outdoor}(t)$								
C_2	$T_{indoorair}(t)$								
C_3	$T_{s,insideoftheouterwall}(t)$						-24.75	-3.82	1.44e-4
C_4	$T_{s,systemwall1}(t)$						315.12	3.35	8.42e-4
C_5	$T_{s,systemwall2}(t)$								

C_6	$T_{s,concretepillar}(t)$		-216.16	-3.51	4.64e-4		-771.70	-6.81	1.73e-11
C_7	$T_{s,systemwall3}(t)$		-84.90	-2.22	0.03		458.23	5.26	1.81e-7
C_8	$T_{s,ceiling}(t)$		134.04	2.94	3.34e-3		1030.00	10.18	3.91e-23
C_9	$V_{wind}(t)$		165.35	3.91	9.94e-05		-972.48	-11.67	1.81e-29
C_{10}	Solar light intensity south (t)								
C_{11}	Solar light intensity east (t)						5.98	3.32	9.34e-4
C_{12}	Solar light intensity west (t)								
C_{13}	$Q_{internal}(t)$		0.02	2.01	0.04		0.10	7.05	3.47e-12
C_{14}	$T_{outdoor}(t-n)$					130	41.42	5.81	8.60e-9
C_{15}	$T_{s,insideoftheouterwall}(t-n)$					1			
C_{16}	$T_{s,systemwall2}(t-n)$					1			
C_{17}	$T_{s,concretepillar}(t-n)$					1			
C_{18}	$T_{s,systemwall3}(t-n)$					1			
C_{19}	Solar light intensity south $(t-n)$					144	4.14	3.28	1.09e-3
C_{20}	Solar light intensity east $(t-n)$					33	11.86	6.52	1.16e-10
C_{21}	$Q_{internal}(t-n)$					1			

Appendix N: Coefficient values model F_1

Table 48: coefficient values of the independent variables, with the corresponding T-statistics and P-value of the heating and cooling Equations derived from model F_1 for the office

Coefficient	Independent variable	Delay (n)	Value	T-statistics	P-value	Delay (n)	Value	T-statistics	P-value
		Heating				**Cooling**			
C_0			-4085.10	-1.01	0.31		-7940.30	-8.46	1.55e-16
C_1	$T_{S,systemwall1}\,(t)$								
C_2	$T_{s,ceiling}\,(t)$		-626.09	-7.97	6.45e-15				
C_3	$T_{s,glasswall}(t)$						507.53	6.50	1.52e-10
C_4	$T_{s,window}(t)$						-162.98	-4.31	1.85e-05
C_5	$T_{indoorair}(t)$								
C_6	$T_{outdoor}\,(t)$								
C_7	Solar light intensity west (t)								
C_8	Solar light intensity south (t)						8.39	4.55	6.44e-06
C_9	$T_{S,systemwall1}\,(t-n)$	104	-1492.90	-5.91	5.28e-09	1			
C_{10}	$T_{s,ceiling}\,(t-n)$	117	-1039.30	-7.14	2.41e-12	1	-128.39	-2.03	0.04
C_{11}	$T_{s,glasswall}(t-n)$	101	4281.80	11.637	1.05e-28	1			
C_{12}	$T_{s,window}(t-n)$					1	142.21	3.64	2.91e-4
C_{13}	$T_{indoorair}(t-n)$	100	-906.26	-4.49	8.42e-06	1			
C_{14}	$T_{outdoor}\,(t-n)$					1	28.96	4.44	1.03e-05
C_{15}	Solar light intensity west $(t-n)$					17	-6.21	-3.90	1.04e-4
C_{16}	Solar light intensity south $(t-n)$					8			

Table 49: coefficient values of the independent variables, with the corresponding T-statistics and P-value of the cooling Equations derived from model F₁ for the classroom

Coefficient	Independent variable	Delay (n)	Value	T-statistics	P-value
			Cooling		
C_0			-5966.60	-14.62	7.25e-44
C_1	$T_{S,insideoftheouterwall}(t)$				
C_2	$T_{s,systemwall2}(t)$		-1612.70	-14.21	8.67e-42
C_3	$T_{s,concretepillar}(t)$		1427.30	6.84	1.45e-11
C_4	$T_{s,systemwall3}(t)$				
C_5	$Q_{internal}(t)$		0.19	12.31	1.86e-32
C_6	$T_{S,insideoftheouterwall}(t-n)$	1	563.95	5.99	2.89e-09
C_7	$T_{s,systemwall2}(t-n)$	1			
C_8	$T_{s,concretepillar}(t-n)$	1	-945.23	-4.56	5.70e-06
C_9	$T_{s,systemwall3}(t-n)$	1	872.96	9.54	1.11e-20
C_{10}	$Q_{internal}(t-n)$	1			

Appendix O: Coefficient values model F_2

Table 50: coefficient values of the independent variables, with the corresponding T-statistics and P-value of the heating Equations derived from model F_2 for the office

Coefficient	Independent variable	Delay (n)	Heating Value	T-statistics	P-value
C_0			-3.01e+6	-10.75	6.04e-25
C_1	*Solar light intensity west* (t)		751.31	3.91	1.01e-4
C_2	$T_{s,ceiling}(t-n)$	117	-32062.00	-8.84	8.61e-18
C_3	$T_{s,insideoftheouterwall}(t) \cdot T_{s,glasswall}(t)$		1718.90	9.13	8.10e-19
C_4	$T_{s,window}(t)/T_{indoorair}(t)$		198.69	7.05	4.40e-12
C_5	$T_{s,systemwall1}(t)/T_{indoorair}(t)$		-1419.20	-8.36	3.72e-16
C_6	$T_{s,insideoftheouterwall}(t)/T_{outdoor}(t)$		-291.21	-11.14	1.55e-26
C_7	$T_{s,ceiling}(t)/T_{outdoor}(t)$		543.21	11.48	6.13e-28
C_8	$T_{s,systemwall2}(t)/$ *Solar light intensity west* (t)		-35.04	-3.93	9.22e-5
C_9	$T_{s,window}(t) \cdot T_{s,window}(t-n)$	117	-79.72	-5.10	4.50e-7
C_{10}	$T_{s,insideoftheouterwall}(t) \cdot T_{s,window}(t-n)$	117	229.03	8.56	8.12e-17
C_{11}	$T_{s,window}(t) \cdot T_{s,systemwall1}(t-n)$	$T_{s,systemwall1}$:104	426.31	5.54	4.43e-8
C_{12}	$T_{s,systemwall2}(t)/T_{s,systemwall1}(t-n)$	$T_{s,systemwall1}$:104	390.80	7.93	9.56e-15
C_{13}	$T_{s,systemwall1}(t)/T_{s,ceiling}(t-n)$	117	1472.60	8.72	2.30e-17
C_{14}	$T_{s,glasswall}(t) \cdot T_{s,glasswall}(t-n)$	$T_{s,glasswall}$:101	-5335.00	-10.66	1.41e-24
C_{15}	*Solar light intensity east* $(t)/T_{s,glasswall}(t-n)$	$T_{s,glasswall}$:101	-136.08	-5.15	3.37e-7
C_{16}	$T_{outdoor}(t)/T_{s,systemwall2}(t-n)$	$T_{s,systemwall2}$:94	-247.80	-5.08	5.04e-7
C_{17}	$T_{s,window}(t-n) \cdot T_{s,systemwall2}(t-n)$	$T_{s,window}$: 177	379.53	3.63	3.02e-4

Coefficient	Independent variable	Delay	Value	T-statistics	P-value
C_{18}	$T_{s,systemwall1}(t-n) \cdot T_{s,systemwall2}(t-n)$	$T_{s,systemwall1}$:104 $T_{s,systemwall2}$:94	-4027.70	-9.39	9.85e-20
C_{19}	$T_{outdoor}(t)/T_{indoorair}(t-n)$	T_{indoor}:100	-53.04	-2.79	5.43e-3
C_{20}	$Solar\ light\ intensity\ east\ (t)/T_{indoor}(t-n)$	T_{indoor}:100	138.82	5.18	2.99e-7
C_{21}	$T_{s,window}(t) \cdot T_{outdoor}(t-n)$	$T_{outdoor}$:100	65.91	3.94	9.04e-5
C_{22}	$T_{s,ceiling}(t) \cdot T_{outdoor}(t-n)$	$T_{outdoor}$:100	-402.81	-5.66	2.29e-8
C_{23}	$T_{s,systemwall2}(t)/T_{outdoor}(t-n)$	$T_{outdoor}$:100	-604.19	-7.63	8.25e-14
C_{24}	$T_{outdoor}(t)/T_{outdoor}(t-n)$	$T_{outdoor}$:100	63.97	5.28	1.74e-7
C_{25}	$T_{s,systemwall2}(t-n) \cdot T_{outdoor}(t-n)$	$T_{s,systemwall2}$:94	-301.61	-3.22	1.36e-3

Table 52: coefficient values of the independent variables, with the corresponding T-statistics and P-value of the cooling Equations derived from model F_2 for the office

Coefficient	Independent variable	Delay (n)	Cooling		
			Value	T-statistics	P-value
C_0			1.28e+05	10.54	4.33e-24
C_1	$T_{s,glasswall}(t)$		-665.76	-3.69	2.43e-4
C_2	$Solar\ light\ intensity\ east(t)$		-35.26	-3.66	2.70e-4
C_3	$Solar\ light\ intensity\ west(t-n)$	1	-296.50	-2.42	0.02
C_4	$T_{s,window}(t)/Solar\ light\ intensity\ west(t)$		7.60	-5.00	7.26e-7
C_5	$T_{s,systemwall1}(t) \cdot Solar\ light\ intensity\ west(t)$		-34.30	-3.43	6.40e-4
C_6	$T_{s,insideoftheouterwall}(t) \cdot Solar\ light\ intensity\ west(t)$		-16.44	-6.80	2.38e-11
C_7	$T_{s,ceiling}(t)/Solar\ light\ intensity\ west(t)$		40.17	5.04	6.12e-7

C_8	$T_{s,window}(t)/Solar\ light\ intensity\ east(t)$		2.50	5.72	1.67e-8
C_9	$T_{s,window}(t)/Solar\ light\ intensity\ south(t)$		-3.46	-6.60	8.36e-11
C_{10}	$T_{s,systemwall1}(t) \cdot Solar\ light\ intensity\ south(t)$		37.33	6.73	3.79e-11
C_{11}	$Solar\ light\ intensity\ west(t) \cdot Solar\ light\ intensity\ south(t)$		-0.69	-4.05	5.80e-5
C_{12}	$T_{s,systemwall2}(t)/Q_{internal}(t)$		0.43	3.18	1.55e-3
C_{13}	$T_{s,systemwall2}(t)/T_{s,window}(t-n)$	8	-142.44	-8.62	5.03e-17
C_{14}	$Solar\ light\ intensity\ south(t) \cdot T_{s,window}(t-n)$	8	2.86	4.80	1.98e-6
C_{15}	$T_{s,insideoftheouterwall}(t) \cdot T_{s,ceiling}(t-n)$	1	316.74	11.83	2.40e-29
C_{16}	$Solar\ light\ intensity\ west(t) \cdot T_{s,ceiling}(t-n)$	1	-81.93	-9.01	2.31e-18
C_{17}	$T_{s,systemwall2}(t)/T_{s,systemwall2}(t-n)$	1	86.29	7.21	1.61e-12
C_{18}	$Q_{internal}(t)/T_{s,systemwall2}(t-n)$	1	-0.44	-3.20	1.42e-3
C_{19}	$Solar\ light\ intensity\ west(t)/T_{indoorair}(t-n)$	1	54.61	7.15	2.39e-12
C_{20}	$Solar\ light\ intensity\ west(t) \cdot T_{outdoor}(t-n)$	1	7.09	6.41	2.81e-10
C_{21}	$Solar\ light\ intensity\ south(t) \cdot T_{outdoor}(t-n)$	1	2.58	3.59	3.62e-4
C_{22}	$T_{s,ceiling}(t)/Solar\ light\ intensity\ west(t-n)$	1	13.29	2.37	0.02
C_{23}	$Q_{internal}(t)/Solar\ light\ intensity\ west(t-n)$	1	0.01	4.00	7.14e-5
C_{24}	$T_{s,window}(t)/Solar\ light\ intensity\ south(t-n)$	17	3.03	3.45	5.96e-4
C_{25}	$T_{s,systemwall2}(t)/Solar\ light\ intensity\ south(t-n)$	17	167.44	8.46	1.82e-16
C_{26}	$T_{indoorair}(t)/Solar\ light\ intensity\ south(t-n)$	17	48.18	4.99	7.79e-7
C_{27}	$Solar\ light\ intensity\ west(t) \cdot Solar\ light\ intensity\ south(t-n)$	17	1.53	5.66	2.29e-8
C_{28}	$Solar\ light\ intensity\ south(t) \cdot Solar\ light\ intensity\ south(t-n)$	17	-2.41	-13.30	7.47e-36
C_{29}	$T_{s,systemwall2}(t-n) \cdot Solar\ light\ intensity\ south(t-n)$	$T_{s,systemwall2} : 1$	-114.01	-7.20	1.68e-12

		Solar light intensity south:17			
C_{30}	$T_{indoorair}(t-n)/Solar\ light\ intensity\ south(t-n)$	$T_{indoorair}$:1 Solar light intensity south:17	-70.28	-7.23	1.35e-12
C_{31}	$T_{s,window}(t)/Solar\ light\ intensity\ south(t-n)$	17	8.54	11.13	1.97e-26
C_{32}	$T_{s,systemwall1}(t)/Solar\ light\ intensity\ south(t-n)$	17	-57.94	-5.18	2.95e-7
C_{33}	$T_{s,insideoftheouterwall}(t)/Solar\ light\ intensity\ south(t-n)$	17	-10.16	-4.82	1.77e-6
C_{34}	$T_{s,glasswall}(t)/Solar\ light\ intensity\ south(t-n)$	17	68.99	6.10	1.84e-9
C_{35}	$T_{s,systemwall2}(t)/Solar\ light\ intensity\ south(t-n)$	17	-8.24	-2.37	0.02
C_{36}	$Solar\ light\ intensity\ west(t)/Solar\ light\ intensity\ south(t-n)$	17	0.35	2.87	4.18e-3
C_{37}	$Solar\ light\ intensity\ east(t)/Solar\ light\ intensity\ south(t-n)$	17	-5.98	-12.25	3.57e-31
C_{38}	$T_{s,window}(t)/Q_{internal}(t-n)$	20	-0.09	-5.59	3.34e-8
C_{39}	$T_{s,ceiling}(t)/Q_{internal}(t-n)$	20	-0.52	-4.61	4.85e-6
C_{40}	$Solar\ light\ intensity\ east(t)/Q_{internal}(t-n)$	20	0.04	5.12	4.01e-7
C_{41}	$Q_{internal}(t)/Q_{internal}(t-n)$	20	9.03e-5	2.21	0.03
C_{42}	$T_{s,window}(t-n)/Q_{internal}(t-n)$	$T_{s,window}$:1 $Q_{internal}$:20	0.04	2.92	3.58e-3
C_{43}	$T_{s,glasswall}(t-n)/Q_{internal}(t-n)$	$T_{s,glasswall}$:1 $Q_{internal}$:20	0.55	4.81	1.85e-6
C_{44}	$Solar\ light\ intensity\ south(t)/Q_{internal}(t-n)$	$Q_{internal}$:20	7.83e-3	3.62	3.21e-4

Table 53: coefficient values of the independent variables, with the corresponding T-statistics and P-value of the heating Equations derived from model F_2 for the classroom

		Heating

Coefficient	Independent variable	Value	T-statistics	P-value
C_0		-5951.50	-0.76	0.45
C_1	$T_{s,concretepillar}(t)$	-183.94	-3.55	4.10e-4
C_2	$Solar\ light\ intensity\ south\ (t)$	-2.39	-2.86	4.31e-3
C_3	$T_{s,insideoftheouterwall}(t) \cdot T_{s,systemwall3}(t)$	-260.30	-6.32	4.03e-10
C_4	$T_{s,systemwall2}(t) \cdot T_{s,ceiling}(t)$	388.16	6.20	8.61e-10
C_5	$T_{s,insideoftheouterwall}(t) \cdot T_{indoorair}(t)$	-318.22	-3.77	1.72e-4
C_6	$T_{s,systemwall2}(t) \cdot T_{indoorair}(t)$	834.59	8.99	1.32e-18
C_7	$T_{s,ceiling}(t) \cdot T_{indoorair}(t)$	-669.56	-8.21	6.92e-16
C_8	$T_{s,systemwall2}(t)/T_{outdoor}(t)$	76.55	3.35	8.34e-4
C_9	$T_{s,systemwall3}(t)/T_{outdoor}(t)$	-77.35	-3.39	7.37e-4
C_{10}	$T_{s,systemwall2}(t)/\ Solar\ light\ intensity\ west\ (t)$	22.99	4.28	2.09e-5
C_{11}	$T_{s,systemwall3}(t)/\ Solar\ light\ intensity\ west\ (t)$	-11.41	-2.30	0.02
C_{12}	$T_{s,ceiling}(t)/\ Solar\ light\ intensity\ west\ (t)$	-11.63	-3.22	1.31e-3
C_{13}	$T_{s,systemwall2}(t) \cdot\ Solar\ light\ intensity\ south\ (t)$	-83.00	-6.31	4.34e-10
C_{14}	$T_{s,systemwall3}(t) \cdot\ T_{s,insideoftheouterwall}(t)$	71.91	6.02	2.42e-9
C_{15}	$T_{outdoor}(t)/\ Solar\ light\ intensity\ west\ (t)$	2.22	3.28	1.06e-3

Table 54: coefficient values of the independent variables, with the corresponding T-statistics and P-value of the heating and cooling Equations derived from model F_2 for the classroom

Coefficient	Independent variable	Delay (n)	Cooling		
			Value	T-statistics	P-value
C_0					
C_1	$T_{s,systemwall3}(t)$		-13530.00	-8.78	8.24e-18
C_2	$T_{s,systemwall2}(t)/T_{s,systemwall3}(t)$		692.76	9.87	6.85e-22
C_3	$T_{s,systemwall2}(t)/T_{indoorair}(t)$		-490.48	-6.83	1.59e-11
C_4	$T_{s,systemwall2}(t)/T_{outdoor}(t)$		-408.50	-10.25	2.18e-23
C_5	$T_{indoorair}(t)/T_{outdoor}(t)$		148.68	8.24	6.04e-16
C_6	$T_{indoorair}(t)/Solar\ light\ intensity\ west\ (t)$		-1.38	-3.44	6.15e-4
C_7	$T_{s,insideoftheouterwall}(t)/Solar\ light\ intensity\ south\ (t)$		22.60	4.63	4.12e-06
C_8	$T_{s,ceiling}(t) \cdot T_{s,insideoftheouterwall}(t-n)$	1	-1328.50	-8.63	2.64e-17
C_9	$T_{indoorair}(t)/T_{s,insideoftheouterwall}(t-n)$	1	892.08	7.34	4.79e-13
C_{10}	$T_{s,ceiling}(t) \cdot T_{s,concretepillar}(t-n)$	1	713.34	5.58	3.22e-8
C_{11}	$T_{outdoor}(t) \cdot T_{s,concretepillar}(t-n)$	1	105.91	3.37	7.76e-4
C_{12}	$T_{indoorair}(t)/T_{s,systemwall3}(t-n)$	1	-487.37	-5.23	2.13e-7
C_{13}	$T_{outdoor}(t) \cdot T_{s,systemwall3}(t-n)$	1	110.40	3.85	1.28e-4
C_{14}	$Solar\ light\ intensity\ south\ (t) \cdot T_{s,systemwall3}(t-n)$	1	-15.33	-3.82	1.43e-4
C_{15}	$T_{indoorair}(t)/T_{outdoor}(t-n)$	1	48.88	3.21	1.35e-3
C_{16}	$T_{outdoor}(t) \cdot T_{outdoor}(t-n)$	1	11.86	3.21	1.38e-3
C_{17}	$Solar\ light\ intensity\ west\ (t)/T_{outdoor}(t-n)$	1	2.16	3.31	9.55e-5
C_{18}	$T_{s,ceiling}(t) \cdot Solar\ light\ intensity\ east(t-n)$	1	16.33	3.95	8.47e-5

C_{19}	$T_{outdoor}(t) \cdot Solar\ light\ intensity\ east(t-n)$	1	2.65	2.64	8.41e-3
C_{20}	$Solar\ light\ intensity\ south\ (t) \cdot Solar\ light\ intensity\ east(t-n)$	1	0.21	2.30	0.02
C_{21}	$T_{outdoor}(t-n) \cdot Solar\ light\ intensity\ east(t-n)$	$T_{outdoor}$:1 $Solar\ light\ intensity\ east$:1	-5.81	-8.30	3.68e-16
C_{22}	$Solar\ light\ intensity\ south\ (t) \cdot Solar\ light\ intensity\ south(t-n)$	1	-0.33	-4.04	5.88e-5