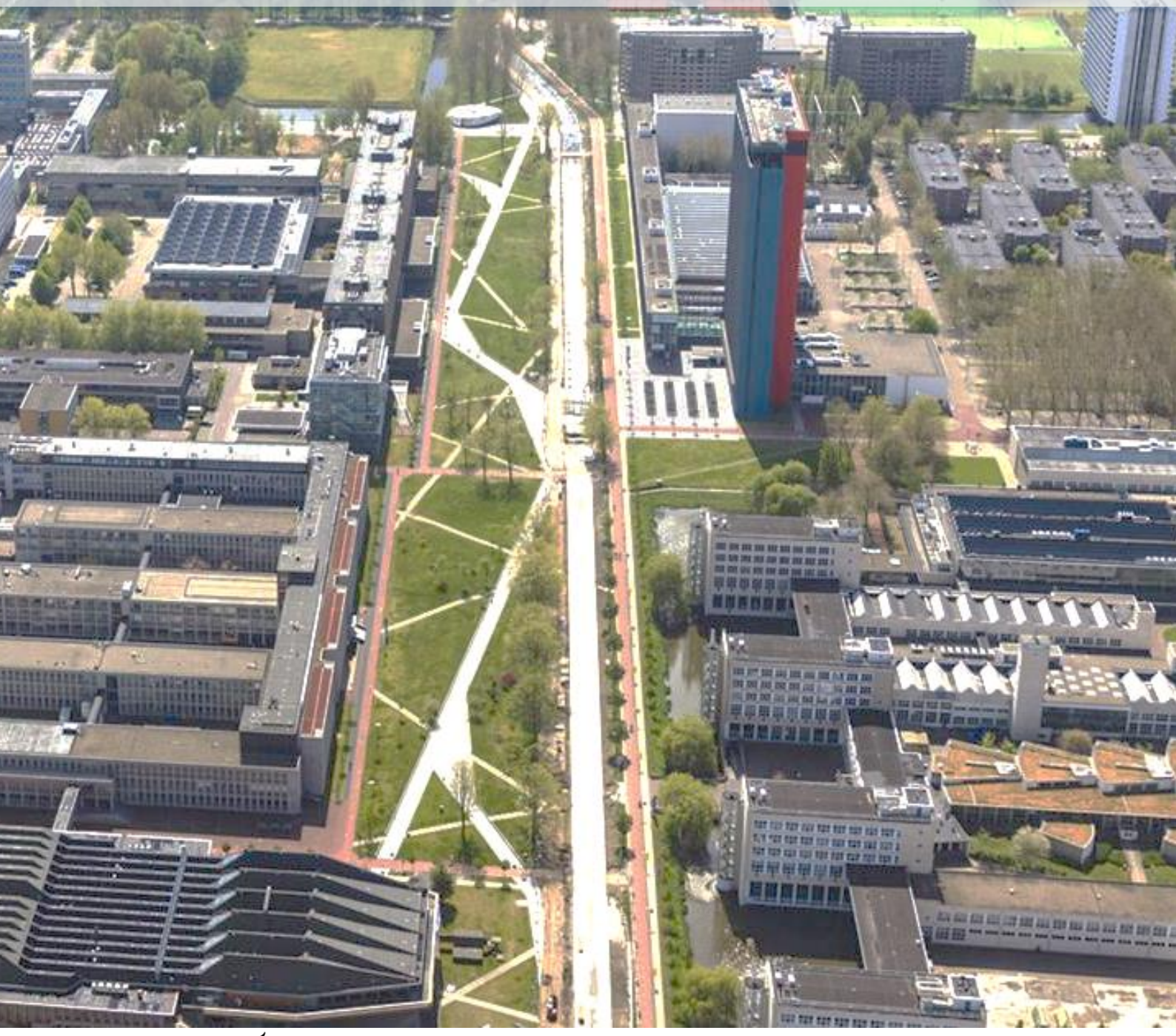
Data-driven Predictive Control for Heating Demand in Buildings

Method Development and Implementation at TU Delft District Heating Grid

Cristina Jurado López



TU Delft

Deerns

Data-driven Predictive Control for Heating Demand in Buildings

Method Development and Implementation at TU Delft District Heating Grid

By

Jurado López, Cristina 4424700

in partial fulfilment of the requirements for the degree of

Master of Science

in Sustainable Energy Technologies

at the Delft University of Technology,

to be defended publicly on Friday March 24, 2017 at 10:00 AM.

Supervisor:	Dr. ir. Laure Itard	
Thesis committee:	Dr. ir. Laure Itard,	TU Delft.
	Prof. dr. Ir. Ad van Wijk,	TU Delft.
	MSc. Paul Stoelinga,	Deerns Nederland B.V.

This thesis is confidential and cannot be made public until October 1, 2017.

TU Delft Deerns

Acknowledgements

I have always had a passion for the optimal use of energy and the implementation of sustainable energies. It makes me proud to have worked on a master thesis that is strongly related to both topics. My drive for energy sustainability has led me to study the master SET at TU Delft and to ultimately deliver this thesis project.

I have really enjoyed the process of this work due to the field of expertise and the high technical and human quality of the team. Therefore, I would like to dedicate some lines to express my gratitude to the people who gave me this opportunity and supported me throughout this process.

I would like to express my thanks to Laure Itard and Chris Hellinga who, independently, help me to discover this project, and to Paul Stoelinga and Prof. Ad van Wijk who made it happen. I am sincerely grateful to my mentors, Laure and Paul, for their guidance and constructive criticism which were of great value to the quality of my work.

I express my warm thanks to the rest of the IPIN team members at Deerns, Roman Aalbers and Renske Kind, for making my daily work at the office enjoyable and making me feel welcome and part of the team. Special thanks to Roman for his support with LEA and being always happy to answer any question.

I would like to dedicate this work to my parents for raising me to believe that anything is possible. And to Jelmer van der Meer, for his unconditional support which makes everything possible.

Cristina Jurado López

Delft, March 2017

Summary

Following the approved Long-Term Agreement on Energy Efficiency (MJA-3), TU Delft has set the target to improve its energy efficiency by 40% by 2020, compared to 2005 levels. In order to reach this goal, 25% of the total energy consumption should be generated by renewable sources by 2020 and the total CO_2 emissions should be reduced by half [1]. Several projects are being launched in TU Delft with the main focus of improving energy management to reduce the electricity and heating consumption in the campus buildings. The optimization of the current heating system has a crucial role to play in ensuring the transition towards a more sustainable campus. This is due to the fact that heating accounts for 47% [2] of the total primary energy demand and it is estimated that between a 10 to 20% [2] of the energy generated in the heat networks is lost.

The optimization of the district heating grid involves lowering the current supply temperature range of the district heating system (DHS) from 130 – 80°C (conventional DHS) to 80 – 40°C (medium temperature DHS). This temperature reduction will lead to the increase of the full loading hours of the two CHP units, the reduction of transport heat losses, the possibility to use waste heat and the implementation of other renewable sources such as geothermal energy. In order to minimize the supply temperature, the Smart Grid Innovation Programme ('Innovatieprogramma Inteligente Netten' - IPIN) was launched.

This thesis is part of IPIN which builds and implements a prototype of a dynamic heating network via a Model Predictive Control system (MPC). This MPC system is formed by two simulation packages: (1) LEA (Low Energy Architecture), developed by Deerns, which estimates the temperature supply at building level (primary supply temperature) by predicting the energy demand; and (2) Wanda, developed by Deltares, which determines the required supply temperature at the central heating plant (secondary supply temperature). This project focuses on the heating prediction at building level and was initiated by Deerns under the IPIN project.

LEA is a physics-based simulation tool (simulator) which gives good estimations of the thermal energy demand of the building. However, as every physics-based model it requires a large number of input parameters that are unknown and need to be estimated. The estimation of parameters for a large scale implementation is a very time-consuming task which requires tedious expert work. Moreover, the estimation of parameters could lead to the introduction of input data errors, decreasing the accuracy of the heating demand prediction and the performance of the thermal grid. This thesis was proposed in order to give an optimal solution to the inconvenience mentioned above.

The goal of this research is to study the possibility of using simple and fast data-driven statistical models to obtain the heating demand of the building with enough accuracy and physical meaning. The scope of the project is limited to the buildings being tested for the current phase of IPIN (phase 2). These are 3mE (Faculty of Mechanical, Maritime and Materials Engineering), IO (Industrial Design) and TPM (Technology, Policy and Management).

Before starting the procedure of building the model, the data set needs to be studied and selected. Two type of data sets were available: (1) simulated data set (calculated by the simulator) and (2) actual data sets (building measurements). By comparing the two data sets, a mismatch between simulated and actual data was observed. The causes of the mismatch between actual and simulated data were studied by comparing the actual buildings and the simulator. Firstly, the accuracy of the simulator was evaluated in order to rule out that the source of this mismatch was due to a possible inaccuracy in thermal balance of the simulator. This analysis concludes that LEA gives a good estimate of the hourly and annual heating demand. Then, the actual cases were compared with the simulator by analysing the most influencing parameters in each case and evaluating the consequences on the data set choice for building the mathematical model. From this

study, it is concluded that the actual data sets are too incomplete to build the mathematical model on and the missing parameters cannot be substituted by the estimated parameters. Moreover, it is proven that the simulated data sets are an accurate representation of the actual data set and their use will contribute to a mathematical model with a better energy and comfort performance. Therefore the mathematical model is built based on the simulated data set.

Once the data set is selected, the multivariate linear models can be built. The major challenge of this part is to find the combination of parameters able to define a correlation which predicts with high accuracy the heating demand for any building. The selection of parameters is based on the thermal energy balance principle of a building. In order to increase the accuracy of the regression model, a new equation is defined in cases where a static variable is influencing the correlation of the linear regression (slope of the least square). In this case study, two equations are distinguished following season changes and three are distinguished for the different operating modes of the building. Each model built is statistically validated by a statistical search procedure. The search procedure analyses the residual of the data set, and quantifies the significance level of both the variables' coefficients and the total model.

In order to build a model that could be applied in the present case study, the selection of parameters is restricted to measurable data at the buildings of the TU Delft campus. With this purpose, it is decided to go an extra mile and obtain a multivariate model independent from all parameters unknown or difficult to obtain. As a result, the final model obtained is independent on building & system characteristics and the influencing parameters that are difficult to measure are replaced. The resulting model is defined by weather data, indoor air temperatures and the internal heat gains of the building. The coefficients of the variables are related to the building and systems characteristics and they are obtained by training the model with historical data.

The model obtained has a high predictive potential and accuracy level. The data collected from the previous season (2.5 months) are able to predict the next month with an accuracy in the range of 73.5 - 99% (weekdays and opening hours). It is expected that the range of the prediction accuracy could be raised up to 90-99% when the equation for weekdays is subdivided in two: (1) including only Mondays and (2) including only Tuesdays-Fridays. Moreover, the model shows a physical relation with the input coefficients, making it possible to get a better insight of the influence of the different parameters in the building response.

In addition to the development of this model (main research question), the calibrations performed previous to this thesis are evaluated and a faster and more accurate calibration procedure is recommended.

This study concludes that the multivariate linear regression model is a more suitable predictive tool than a physics-based model (e.g. LEA) for large scale implementations. Any physics-based simulator (and therefore, LEA) is beyond the complexity limits of this case study because some of the most influencing parameters are unknown and therefore estimated, leading to the introduction of input data errors in the simulated results.

Content

APPENDICES

Table of figures

1 Introduction

1.1 Energy transition

Since the industrial revolution, the energy demand have experienced a growing rise. Fossil fuels have been the base of the worldwide energy production and its use gave place to an exponential increase of the Greenhouse Gases (GHGs). The Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC) estimated that between 1970 and 2004, global greenhouse gas emissions due to human activities rose by 70%, being the main reason of the global warming [3]. Besides the damaging effects that the GHSs have on the climate change, the depletion of the fossil fuels and the expected increase in the energy consumption (related to the world population growth and higher life standards), makes necessary the shift from fossil fuels to renewable sources.

In order to ensure a future energy security and mitigate the climate change, the Intended Nationally Determined Contributions INDC of the European Union (EU) sets out the following targets for 2030 compared with 1990 levels [4].:

- Reducing the GHG emissions at least by 40%;
- increasing the share of renewable energy to at least 27% (of final energy consumption);
- improving the energy efficiency by at least 27%

The 2030 framework builds on the target to reduce EU GHG emissions by 20% by 2020. Figure 1 presents the energy share from renewable energy sources in the EU member states for 2014 and the targets set for 2020. It is observed that the EU is on track of meeting with the targets set for 2020, however The Netherlands is at the bottom of the most European ladder in terms of renewable energy production. In 2014 only a 5.6% of the energy consumed was produced by renewable sources [5].

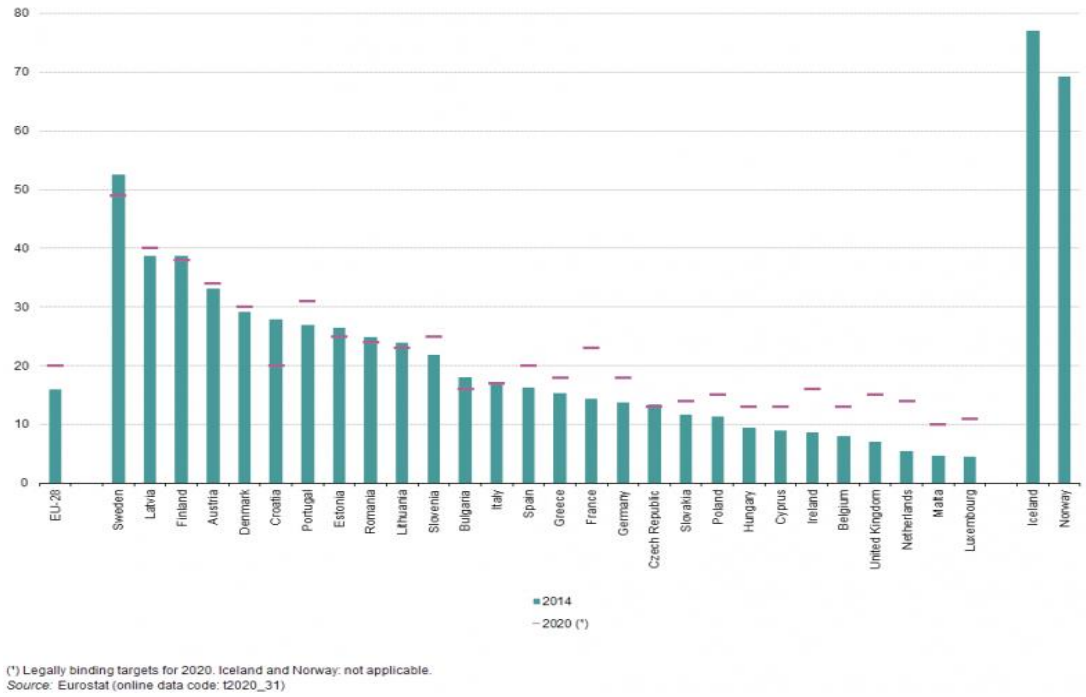
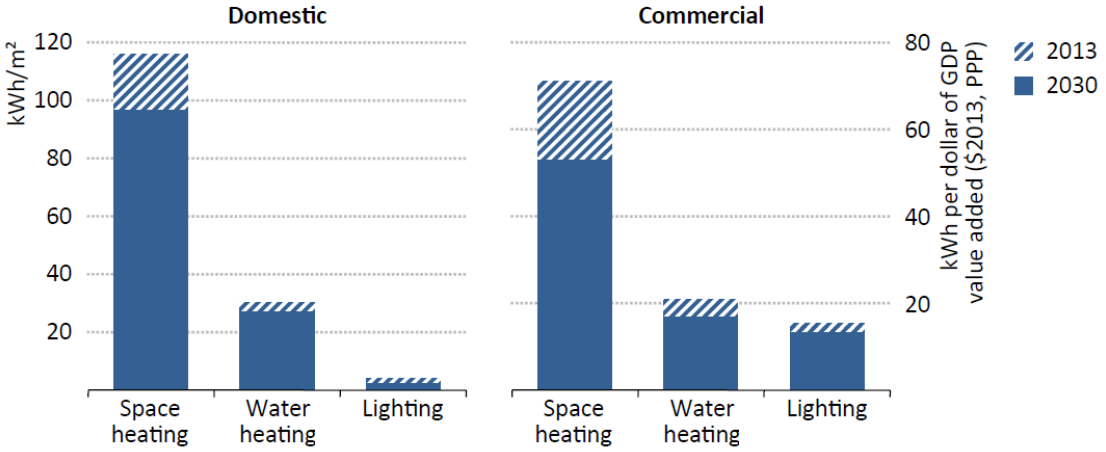


Figure 1 Share of total primary energy from renewable sources in the EU member states, 2014 and 2020 (%) [5]

1.2 Renewable energies in the built environment

The built environment consumes more than the 40% of the global energy used and contributes up to 30% of the global GHGs emissions [6]. Due to the growth in new construction in developing economies and the inefficiencies of the existing building stock worldwide, the GHGs emissions related to the built environment are increasing annually. Between 1971 and 2004 the yearly GHGs emissions related to the built environment have increased at an annual rate of 2.5% for commercial buildings and at 1.7% for residential buildings [7]. Space and water heating in European households accounts for 79% of total final energy use (192.5 Mtoe) and the 84% of heating is still generated from fossil fuels while only 16% is generated from renewable energy [8]. In order to fulfil the European climate and energy goals, the heating sector in building must sharply reduce its energy consumption and cut its use of fossil fuels [8]. Figure 2 shows the energy intensity levels in European domestic and commercial buildings for 2013 and 2030 [4].



Notes: kWh/m² = kilowatt hours per metre squared. GDP is expressed in purchasing power parity (PPP) terms.

Figure 2 European Union energy intensity levels in domestic and commercial buildings, 2013 and 2030

The building sector has a large potential for reducing the energy consumption in both new and existing buildings. By applying proven and commercial technology, the energy consumption can be reduced by a 30 till 80% with potential net profit [6]. Reducing the energy consumed corresponding to heating and cooling in buildings can be achieved by improving the building design and insulations when renovating buildings, providing better information and control of energy use via energy management solutions, and upgrading heating and cooling equipment to the most efficient technologies [6].

The use of fossil fuels can be cut down by using renewable heating and cooling technologies such as reusing waste heat, biomass boilers, geothermal energy or combined heat and power units which produce both heat and electricity. The implementation of renewable energy systems into the existing energy system involves the integration of the different smart grids and the coordination between energy supply and demand in order to obtain a good efficiency of the overall energy system [9], [10]. Therefore, it is expected that smart thermal grids will play a key role to make full use of the potential of distributed renewable technologies. The next section explains the role and challenges of the smart heating grids to reach the corresponding energy reductions and the integration of renewable energy systems.

1.3 From current district heating grids towards smart thermal grids

This section gives an overview of the different district heating generations and describes the role and challenges that may encounter the next district heating generation in the framework of the future energy systems. The definition of the concept of 4th Generation District Heating (4GDH) and smart thermal grid [9] given in this section, is based on the trend followed by the historical development of district heating systems and the existing motivation and need to transform the current energy system.

1.3.1 Tendency of district heating systems

Figure 3 shows the evolutionary tendency of the different district heating systems in history. This presents the main characteristics of the 3 known district heating generations and the 4th predicted district heating network [9].

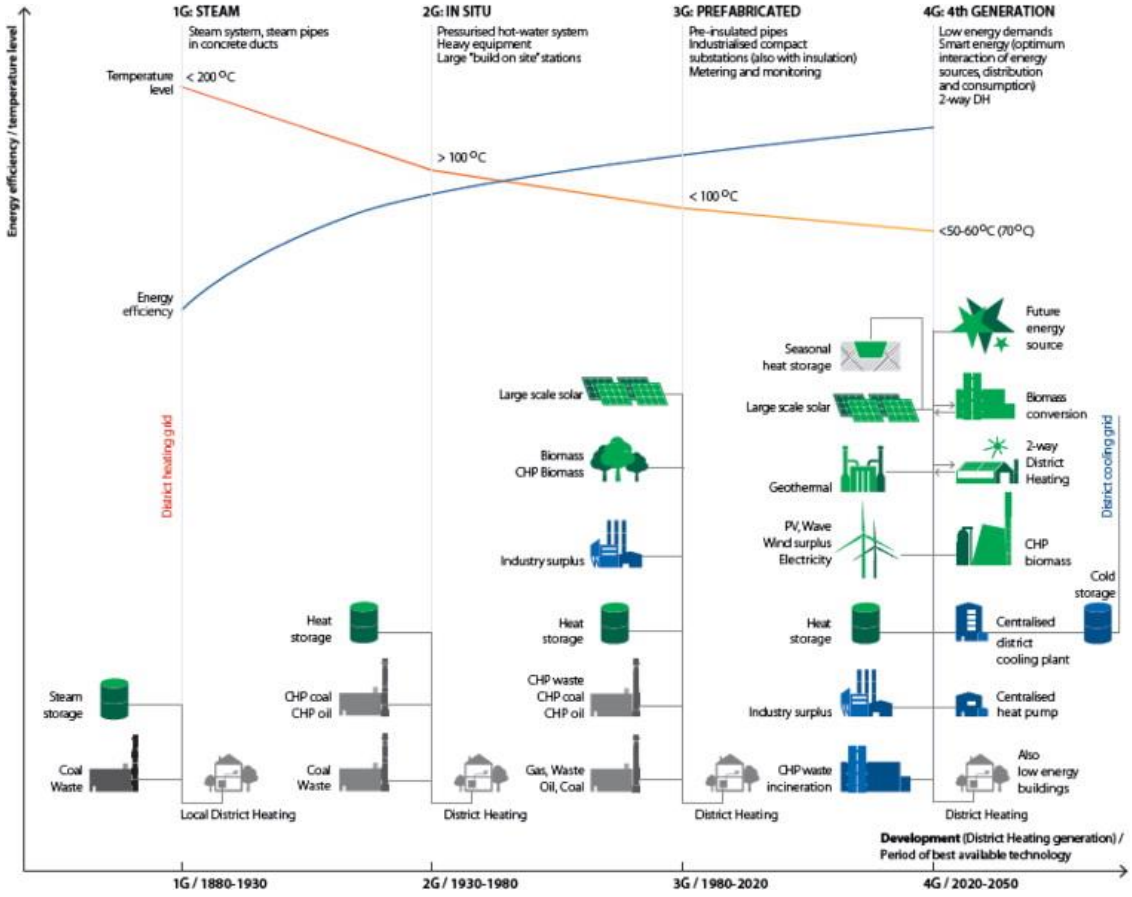


Figure 3 Comparison of the different Generations of District Heating (GDH) [9].

The first generation of district heating systems (1880-1930) was characterised for using steam as heat carrier. The high steam temperatures lead to high heat losses, however the main drive to change these systems was to reduce the risk of accidents due to the explosions of steam boilers. In the second generation (1930-1980), the heat carrier was pressurised hot water at supply temperatures above 100 °C. These systems decreased the heat losses, increased fuel savings and comfort by using combined heating & power (CHP). The third generation of systems (1980-today) also uses pressurized water as the main heat carrier, but at temperatures lower than 100 °C, lowering the heat losses (increase of the energy efficiency). The main motivation for using these systems is to decrease the dependency of oil and increase the security of energy supply (motivated by two previous oil crisis). Therefore, there is a focus on energy efficiency related to CHP by replacing oil by other fuels such as coal, biomass and waste [9].

According to the trend followed by the three previous DHGs, the 4DHG should develop towards lower distribution temperatures and the introduction of the new energy sources.

1.3.2 Role and challenges of the district heating technology for reaching a sustainable energy system

The transition from current energy systems (based on fossil fuel and nuclear energy) towards the future energy systems (based on renewable energy systems) leads to an increase in the fluctuation of the energy supply. As a result, the current energy system needs to be transformed into a smart energy system which focuses on the integration of electricity, heating, cooling and transport sectors, and on using the energy demand flexibility as well as energy storage [9]. This implies the integration of the different smart grids (electricity grids, district heating and cooling grids, gas grids and different fuel infrastructures) and the coordination between them to achieve the optimal solutions for each individual sector and for the overall system [9], [10].

District heating consists in a network of pipes connecting the thermal energy demand and supply sides, allowing any available source of heat to be used (centralised or distributed heat producing units). Therefore, district heating systems plays a crucial role in the implementation of future sustainable energy systems. However, the current district heating system has to be developed in order to decrease grid losses, exploit synergies, and increase the efficiencies of low temperature production units [11], [12], [9]. In order to reach this goal, the future district heating will have to overcome the following challenges [9]:

1. Supplying low temperature district heating to existing and new low energy buildings.
2. Distributing heat in networks with low grid losses.
3. Using low temperature heat sources (waste heat and renewable heat sources such as solar or geothermal heat)
4. Being coordinated with the other energy grids to become an integrated part of the smart energy system.
5. Ensuring suitable planning, cost and motivation structures in relation to the operation and strategic investments related to the transformation into future sustainable energy systems.

In order to achieve the above mentioned challenges, it is important to coordinate the energy performance of the buildings and the district heating systems to improve the energy efficiency of the total system. The requirements that a building may fulfil to reach this synergy between thermal energy demand and supply, can be summarized in the following two requirements [9]: (1) reducing the heating demand in buildings for a better balance between the energy needed during winter and summer, (2) smart control of heating demand for an optimal operation of the buildings and district heating systems.

The reduction of heating demand in buildings must guarantee that comfort is achieved at lower supply temperature, giving rise to lower temperature supply differences between winter and summer [9]. This will lead to facilitating the use of the same systems for cooling and heating, enhancing the implementation of renewable energy sources (low temperature heat sources), decreasing grid losses, and increasing the efficiencies in the production units [9]. The smart control system of the heating demand in buildings optimizes the operation of buildings and district heating systems by:

1. matching demand and supply;
2. shaving heating peak loads in order to maintain the heating supply constant. This leads to an optimal operation of the district heating systems and avoids high investments to cover the high peak loads.

The focus of this master thesis is on the smart control of heating demand by searching for mathematical models able to predict future energy demand during the operation of the buildings.

This master thesis is part of the ongoing Smart Grid Innovation Programme ('Innovatieprogramma Inteligente Netten' - IPIN), which aims at transforming the current TU Delft district heating grid into a smart thermal grid. The data used were obtained from the case study of different buildings at TU Delft. This case study and its context are explained in the next section.

1.4 Transformation of the TU delft District Heating Grid towards a Smart Thermal Grid: IPIN project

Following the approved Long-Term Agreement on Energy Efficiency (MJA-3), TU Delft has set the target improving its energy efficiency by 40% by 2020, compared to 2005 levels. In order to reach this goal, 25% of the total energy consumption should be generated by renewable sources by 2020 and the total CO_2 emissions should be reduced by half [1]. With this purpose, several projects are being launched with the main focus of reducing the electricity and heating consumption in the campus buildings by improving energy management.

The optimization of the current heating system has a crucial role to play in ensuring the transition towards a more sustainable campus. This is due to the fact that heating consumption accounts for 47% [2] of the total energy consumption in TU Delft campus and it is estimated that between a 10 to 20% [2] of the energy generated in the heat networks is lost. This loss is due to a non-optimal use of the heating generation systems and the transport heat losses taking place in the distribution network.

With the purpose of optimizing the current heating generation & distribution system and introducing other energy sources (low temperature heat sources), the Smart Grid Innovation Programme ('Innovatieprogramma Inteligente Netten' - IPIN) was iniciated. IPIN aims at lowering the current supply temperature range of the district heating grid from 130 – 80°C (conventional DHG) to 80 – 65°C (medium temperature DHG) and ensuring a constant heating supply. This will lead to:

- optimising of the use of gas boilers (continuous operating mode due to a constant heating supply);
- reducing the transport heat losses;
- increasing of the full loading hours of the two CHPUs;
- using the waste heat (eg. heat coming from the CHP's flue gas);
- and introducing other renewable sources such as geothermal energy (which supply heat at approximately 70ºC).

0 gives a more detailed information about the current district heating system at TU Delft campus (heating distribution and generation) and how the transformation towards a smart heating grid will lead to the improvements described above.

As it is explained in section 1.3.2 and 0, in order to improve the energy performance of the total heating system, it is important to coordinate the energy demand of the buildings and the district heating system. The synergy between thermal energy demand and supply can be reached by reducing the heating demand in buildings and implementing a smart control of heating demand in buildings. The smart control of the heating demand will ensure a more constant operation leading to an optimal operation of the district heating systems.

Most of the TU Delft buildings have large heating consumptions since they were built under the building regulations of the 1950s and 1960s. In order to decrease the heating demand in TU Delft buildings, large-scale renovations (eg. increasing building insulations) will be the most logical option from the energy point of view, but it requires a very large investment [1]. In order to reach the MJA-3 goal in an affordable manner, it was decided to lower the supply temperature of the heating grid to integrate renewable energy sources.

IPIN develops and implements a prototype of a dynamic heating network via a Model Predictive Control system (MPC) developed by Deerns and Deltares. This MPC adjusts the supply

temperature at both the heating generation (central heating plant) and heating distribution stations (situated at the buildings). In order to lower the supply temperature of the TU Delft District Heating Grid (DHG) without compromising a strong decrease on the comfort level, the MPC predicts and balances the heating demand & supply by taking into account all factors influencing the thermal energy demand in the DHG at the building, systems and pipeline level. Therefore, it includes the physical characteristics of the buildings & systems, the use & operation, weather data and thermal dynamism in the pipeline system. This makes a difference with the current control system which calculates the supply temperature only based on the outdoor temperature, neglecting the other influencing parameters. In the current phase of IPIN project (phase 2), the MPC prototype is being implemented in branch North 2 (see 0 for further details).

Figure 4 shows the schematic overview of the Model Predictive Control system. The MPC system is formed by two simulation packages: (1) LEA (Low Energy Architecture), developed by Deerns, which predicts the thermal energy demand and needed temperature supply at building level (primary supply temperature); (2) Wanda, developed by Deltares, which simulates hydraulic and thermal transients in pipeline systems. Based on the primary supply temperature calculated by LEA, Wanda is used to determine the energy source usage and the required supply temperature at the central heating plant (secondary supply temperature). This project focuses on the heating prediction at building level (green part).

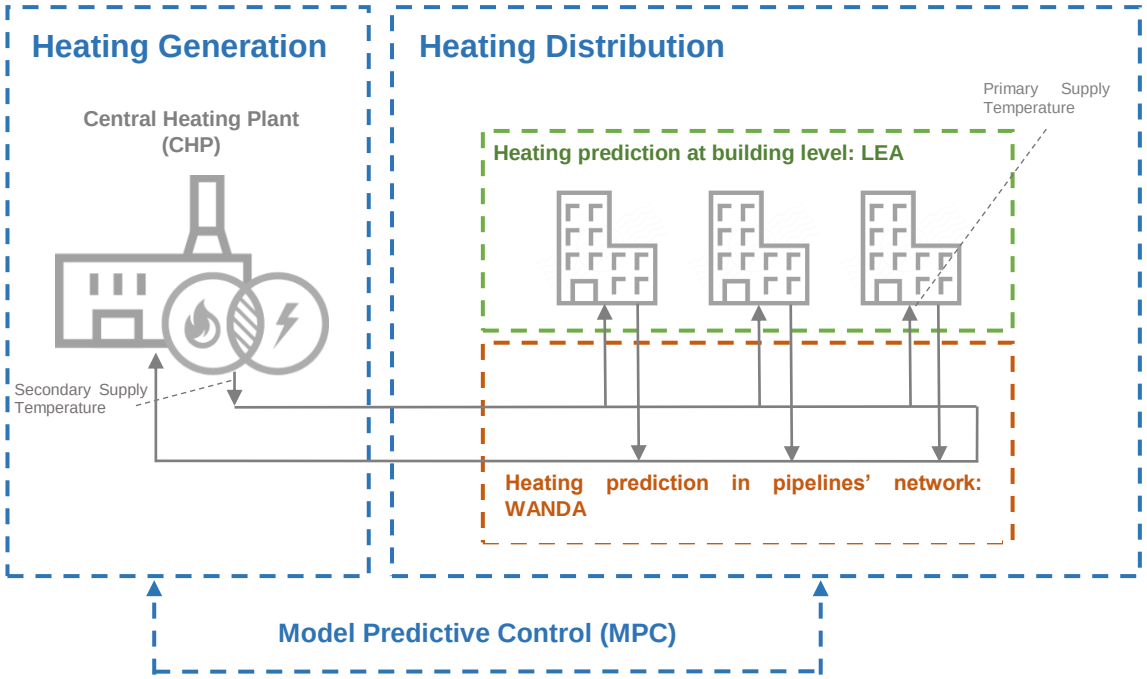


Figure 4 Overview of the Model Predictive Control system [own illustration]

2 Research outline

2.1 Problem definition

As it was explained in the previous section, the prediction of the heating demand in the TU Delft buildings is being calculated by LEA (see section 6.1 for further details). This is a physics-based simulation tool which gives good estimations of the thermal energy demand of the building. However, it requires more than 120 inputs in each case study. These inputs can be classified in weather data, physical characteristics of the building & systems, and operational characteristics. The last two types of parameters were estimated by an inventory of inputs (visiting the place and interviewing the building operators). However, not all parameters were able to be properly estimated. Therefore, a validation with actual historical consumption data was performed for each of the 20 buildings at TU Delft campus (see section 6.2.1).

The use of the mentioned physics-based model and the working procedure involved, lead to the following main challenges for large scale implementation.

1. *Very time-consuming method*

   More than 120 parameters for 20 different buildings had to be estimated. The inventory of these parameters and the validation procedure were very time demanding tasks, which could impact future implementation in other projects.

2. *Expert work required*

   The use of LEA, as every physics-based tool requires high expertise level. A team of 4 experts was formed in order to make the required validations and adjustments in the model for its communication with other softwares (see section 6.2 for further details about the calibrations and validations).

3. *Introduction of errors*

   The estimation of the required parameters could introduce input data errors to the model, which could lead to a reduction in the accuracy of the simulated results (see section 3.1.3).

4. *Mismatch between actual and simulated data.*

   For some of the buildings, the study showed a big mismatch between the actual and simulated data (see chapter 7). This mismatch could affect the efficiency of the Model Predictive Control system leading to a lower performance of the smart thermal grid.

5. *Challenging method to get a good understanding of the most influencing parameters*

   The complexity of the model (containing hundreds of physical equations) challenges the understanding, evaluation and influence of the different parameters on the output.

6. *High computer power demand*

   LEA model is based on hundreds of equations which need to be run every time step, demanding high computer power. This is not a big deal for the current computers, however it limits its implementation to a computer hardware (like PLCs used in control systems). When a model demands less computer power, the prediction of the energy demand can be calculated with simpler hardware such as buildings' controllers or mobile phones (through an application).

2.2 Objective and research questions

The main purpose of this research is to study the possibility of using simple and fast data-driven statistical models to obtain the heating demand of the building. This model should give a solution to the challenges encountered by using physics-based models. Therefore, the following objectives should be reach.

1. *Simplicity* to avoid the introduction of input errors, reduce expertise of the user and time spent for its implementation.
2. *Fast calculation* in order to facilitate its implementation in simpler hardware and reduce the operational cost related to software licenses.
3. *Accuracy* to be able to control the smart thermal grid at a high performance.
4. *Physical relation with the input coefficient parameters* in order to make the interpretation of the results easy and being able to get a better insight of the influence of the different parameters in the building response.

Therefore, this study answers the following main research question:

- ✓ *Is it possible to predict the heating demand of a building through a simple and fast mathematical model with enough accuracy and physical meaning?*

In order to answer this research question, the following sub questions are also approached in the different research phases of this study:

1. Literature review (chapter 3)
    a. What type of prediction model and statistical methodology are the best approach to fulfil the objective mentioned above?
    b. Which are the expected most influencing parameters on actual heating demands in this case study?
    c. Which are the reasons that may cause a mismatch between actual and simulated heating demand in buildings?
2. Analysis real case study (chapter 5)
    a. Which are the main building and operating characteristics in each case study?
    b. How many equations are necessary to build the statistical model?
    c. How should the data set be selected to build each equation?
    d. Is the actual data set provided complete enough to build the statistical model?
3. Analysis physics-based simulator and validation (chapter 6)
    a. Which are the main assumptions of the physics-based simulator used? Are there model assumptions which could affect the simulated results?
    b. Is the physics-based simulator accurate enough to generate a representative data set to build the statistical model?
4. Comparison real case study (buildings) and simulator (chapter 7)
    a. Which are the most influencing key indicators and dynamic parameters in each building for the actual and simulated data? How do they interfere on the actual and simulated heating demand in each case?
    b. Why is there a mismatch between actual and simulated heating demand?
    c. Were the parameters correctly estimated during the calibration procedure performed prior to this thesis? In which cases do the calibrations need to be validated?
    d. Is there a better calibration procedure to estimate the parameters in this case study?

e. Is it acceptable to substitute the missing parameters of the actual data set by estimated parameters? If not, is it acceptable to replace the actual data set by the simulated data set to build the mathematical model?

5. Model design & validation (chapters 8, 9 and 10)
    a. Which is the best combination of variables to obtain an accurate predictive model?
    b. Is the model satisfying the fitting accuracy required?
    c. Are all the variables included in the model available in this case study? If not, is it possible to neglect or replace the variables?
    d. What are the physical interpretations deducted from the coefficient of the parameters?
    e. Which is the best model to apply in each case?
    f. What is the model predictive potential?

3 Literature review

This chapter aims to give direction in the design of a mathematical model which satisfies the research requirements described in chapter 2.

In the first place, the state-of-the-art of the different prediction model techniques is studied and the optimal complexity of building models is analysed. The results of this review lead to the selection of the most appropriate prediction model for the present case study.

In the second place, the mismatch between actual and simulated data found during the calibration period may be signs of inaccurate actual or simulated data. In order to make a good selection of the data set to be used for the construction of the mathematical model, the possible reasons for the gap between actual and simulated data were studied.

Finally, the most influencing parameters on actual heating demand are analyse. The result of this study gives good insights of the physical and operational characteristics of the buildings and systems to be considered in the current case study in order to build the mathematical model.

3.1 Prediction model techniques and optimal complexity

3.1.1 Classification and overview prediction model techniques

There are several categorisation systems to classify the different existing prediction methodologies (see [13]–[17] for further references). Most of the authors classify these methods into white-box models, gray-box models and black-box models [16], [17].

- White-box models (also called law-driven models): are based on building physics equations, therefore they are dependent on a large amount of building's inputs (design requirements). Physic-based simulation tools such as TRNSYS, EnergyPlus, SIMBAD [18] are examples of white-box methods.
- Gray-box models (parameters estimation): are a combination between a white-box and a black-box by using limited building physics in combination with statistical methodologies. Examples of this type of method include some engineering calculations such as the degree-day method [19] or international standards [20], [21].
- Black-box models (also called data-driven models): use statistical models which relates a set of influential inputs parameters to measured outputs, therefore they require training data (historical data) but a very small quantity of inputs. Examples of black-box method include all statistical methods such as regression models [18], [22], [23] or Fourier series model [24]–[26].

3.1.2 Comparison between different prediction model techniques

The selection of the prediction model technique for a certain application will depend on the knowledge of the system modelled (physical equations and availability of inputs) and aim of the application [13], [18]. This section gives an overview of the advantages, limitations and applications of the different model techniques.

The engineering calculations (grey-box models) are easy and fast to use and they normally show a clear relation to physical parameters. Some of these tools do not take into consideration the dynamic processes of the building (eg. dynamic heat transfer) or they make more simplifications than the physics-based simulation methodologies, leading to a lower accuracy level in predicting the energy demand of the building. As a result, they are used in early design stage of buildings design and during energy audits, where energy performance estimations are sufficient for evaluating the performance and predicting the impacts of energy conservation measures [13].

Physics-based simulation tools (white-box models) are very effective for modelling individual buildings (existing or at design stage). However, the accuracy of the modelled results will be highly influenced by the expertise of the modeller and knowledge on the system modelled such as the physical equations and the input parameters available [13], [18]. These tools require a large amount of detailed input data, such as physical characteristics of the building & systems and operational characteristics (eg. internal heat gains profiles, temperature set points, etc), which are uncertain in most of the real life cases increasing the potential error in the predictions (see section 3.1.3 for further details).

Statistical simulation tools (black-box, sometimes grey-box models) provide fast and simple prediction models which are normally developed using a survey database. These models have a much lower number of input data which can be easily known. The main disadvantages of these tools are that they lack a clear relation to physical parameters, resulting in a difficult interpretation of the results and difficulties identifying errors [13].

3.1.3 Model uncertainties related to the complexity

As it was mentioned above, the selection of the approach to be used will depend on the purpose of the model and knowledge of the system modelled [13], [18]. The selection of a more complex model does not mean a better solution. This section explains how the complexity of the model is limited by the number of inputs available.

The type of errors in a verified model are classified as follows:

- Abstraction errors: due to the boundary conditions of the model (eg. using an incomplete model of a physical system)
- Input data errors: uncertainties introduced by the estimation of parameters (unknown parameters)
- Numerical errors: due to the discretization steps chosen in the modelling

The numerical error can be controlled by decreasing the discretization steps. However, the abstraction and input data errors are difficult to quantify, increasing the modelling uncertainty and influencing the gap between actual and modelled outputs [27].

Figure 5 shows that when the model complexity rises, the gap between actual and modelled outputs (potential error) decreases while the predictive uncertainty increases. This is due to the fact that when the modelling complexity increases, the model equations define with more detail the reality, leading to a reduction of the potential error. However, the increase in complexity also leads to the higher amount of parameters to solve the physical equations involved, rising the predictive uncertainty (abstraction errors and input data errors). The predictive uncertainty decreases at a higher level of knowledge on the system modelled (physical equations and input parameters). Therefore, the system knowledge available will determine the complexity of the model for which the model error has its minimum (see sum line in Figure 5 [27]), this corresponds with the optimal

complexity. ”There is no sense in going beyond this complexity, as the overall error in the model uncertainty will not decrease” [27].

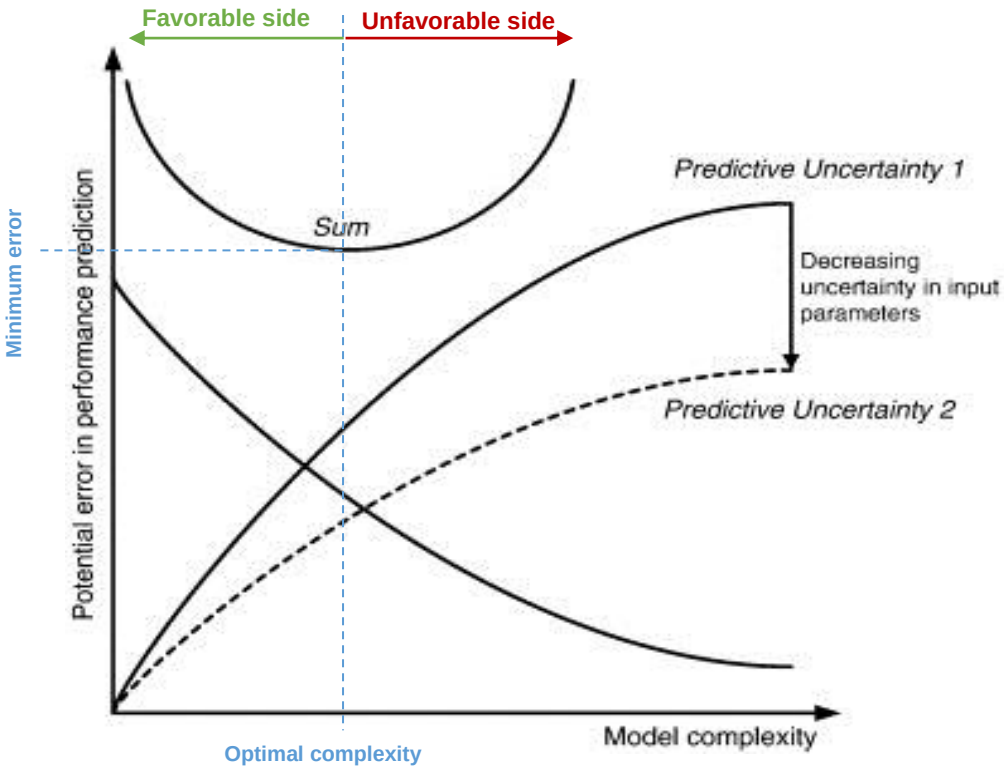


Figure 5 Model uncertainty versus complexity.

3.2 Overview and selection statistical methods

For the last decade, several statistical (mathematical) models have been used for thermal energy demand prediction. The methods most commonly used are regression models, Fourier series models and Artificial Neural Network (ANN) models.

ANN models are a machine learning technique which operates in a black-box principle, which does not require detailed information of the system and is trained based on historical datasets (input and outputs variables) [13]. The application of this method is especially interesting to solve complex problem with many parameters. An example of this type of problems is the automated estimation of the numerous input data parameters used by a physics-based software (physical characteristics of the building & systems and operational characteristics) based on large historical dataset [28]. During the first phase of this project, the possibility of using this method in order to develop a model which could estimate the different input parameters needed by a physics-based model was studied. However, at the time of the definition of this project, the parameters for the 20 different buildings were already estimated and this method did not guarantee a simpler and fast final equation. Therefore, it was decided to proceed with a method which could provide an equation that fulfilled all requirements set (see section 2).

During this study, both regression and Fourier series models were explored. Both models present one equation dependent on multiple independent variables. The multivariable regression model combines the different independent variables according to a linear or polynomial relation between the independent and dependent variable [24]. In the Fourier series approach, the relation between the independent variable and the dependent variable is a combination of sine and cosine functions [24]. This could be useful to model the daily and annual periodicity of the variables in order to obtain the profiles of thermal heating demand. The functional form of these two approaches are represented as follows:

Regression approach:

$$Q_h = constant + \sum_{i=1}^{n} C_i \cdot X_i$$

Where Q_h is the hourly thermal energy demand, C_i are the corresponding coefficient of the dependent variable, 'X' is the dependent parameter and 'I' the index of dependent variables selected for the model.

Fourier series approach:

$$Q_h = constant + \sum_{i=1}^{n} X_i \left[\sum_{x=1}^{182} [\, \alpha_x \cdot \sin\left(\frac{2\pi d}{P_x}\right) + \beta_x \cdot \cos\left(\frac{2\pi d}{P_x}\right)] + \sum_{y=1}^{11} [\, \gamma_y \cdot \sin\left(\frac{2\pi h}{P_y}\right) + \delta_y \cdot \cos\left(\frac{2\pi h}{P_y}\right)] \right]$$

Where the first Fourier series with subscript 'x' represent the annual periodicity and the second Fourier series with subscript y the diurnal periodicity. 'd' is the number of the day with respect to the year (from 1 to 8760) and 'h' is the hour with respect the day (from 1 to 24). P_x is the annual periodicity and P_y the diurnal periodicity for a specific frequency.

The Fourier series approach is more complex than the regression approach since it contains several terms for every parameter, increasing the total number of parameters to be analysed. The advantage of this approach is that it achieves a single equation able to predict the thermal energy demand all year round since it takes into account the annual periodicity of the different variables. However, this advantage was not useful for this project since the heating demand data set available was limited to the seasons when only heating demand was required, neglecting the periods when both heating and cooling are required. Therefore, it was decided to build the mathematical model following the regression approach.

3.3 Influencing parameters on the thermal energy demand

3.3.1 Most influencing parameters on the actual thermal energy demand

Building simulations tools are considered to be unreliable at predicting the energy performance of buildings [29]. The lack of knowledge of the system modelled (as mentioned in section 3.1.3), such as limited information about the buildings´ characteristics, installations [29] and an underestimation of the role of the occupant´s behaviour [30], [31], could lead to a wrong estimation of these parameters and inaccurate thermal energy predictions.

Several researches analyse the most influencing parameters on the actual heating. This section presents the most relevant parameters according to several authors [29-34].

Most influencing building characteristics according to the energy performance of the building and heating systems installed

Ioannou [29] studies the most influential parameters in residential buildings by making a sensitivity analysis for building parameters and occupancy. It was found out that the most influencing parameters vary with the type of energy label (from class A to class F) and the type of heating system installed. Table 1 shows the most influential physical parameters resulting from this study excluding the influence of the occupancy behaviour.

Table 1 Most influencing physical parameters on the actual heating demand according to the performance of the dwelling and heating system implemented [29].

Type heating system → **Label dwellings ↓**	**Radiator heating system**	**Floor heating system**
Class A	Windows U-value, windows g-value, wall conductivity	
Class F	Wall conductivity; windows g-value; orientation of the building	wall conductivity; floor conductivity; windows g-value

The dwellings of class A showed that the most influencing parameters are independent from the heating system installed, being windows U-value, windows g-value and wall conductivity the most influencing parameters for both dwelling types.

For the dwellings of class F, the wall conductivity is the most influencing factor for both heating system types. The floor conductivity has a more influencing effect than the windows g-value in the cases when floor heating systems are installed due to the importance of the heat flux through the floor [29]. For the buildings with radiator heating systems installed, the parameters related with solar gains (windows g-value and orientation of the building) are the most influencing parameters after the wall conductivity.

Behavioural influence (called operational characteristics for office buildings)

When behavioural parameters are introduced by varying the thermostat settings and ventilation flow rate, the influence of the physical parameters on the simulated heating demand is reduced [29]. As a result, the thermostat settings and the ventilation flow rate become the most influencing variables for both type of dwellings.

Although the study described above dealt with household dwellings, it seems logical to assume that the equivalent parameters in office buildings also will have a high influence [33]. In this case, the behavioural parameters in dwellings will correspond to the operational characteristics in office

buildings (temperature set points and ventilation flow rate), which are also related to the occupancy of the building.

Thermal mass influence

The influence of the thermal mass of the building on the heating & cooling energy demands is highlighted by several researches [32], [34]. The internal thermal mass provided in floors and walls absorbs and store internal heat gains. This effect modifies the thermos-physical behaviour of the building, reducing indoor air temperature fluctuations and increasing time lag [32], [34]. As a result, increasing the thermal mass of the building leads a decrease in peak cooling and heating transmission loads and an increase in the time lag. The optimization of the thermal mass can result in maximum savings in yearly thermal energy demand of about 17% for cooling and 35% for heating [32]. There is a minimum critical amount of thermal mass that a building should contain in order to be able to have a potential of energy savings in the range between 90-97% [32].

The insulation in walls does not counteract the effect of the thermal mass, but it has an influence. Figure 6 shows the transmission load variation with time of day in August, January, and November with thermal mass thickness (L_{mas}) of 20 cm and 5 cm for 2 walls with the same insulation value (R-value=2.86 m^2 K/W). (a) presents the results for wall W1 (outside insulation) and (b) for wall W2 (inside insulation). When the thermal mass is increased, load fluctuation is damped and the peaks are reduced in both cases (insulation placed inside and outside). However, when the insulation is placed outside, the increase of the thermal mass provides more damping of load fluctuation and smaller peak load than when the insulation is placed inside, leading to higher energy-saving potentials [32]. This is due to the fact that by placing the insulation outside, the thermal energy is storage in the inside wall, being later dissipated to the inside through the same surface, leading to the corresponding beneficial effect in the indoor climate.

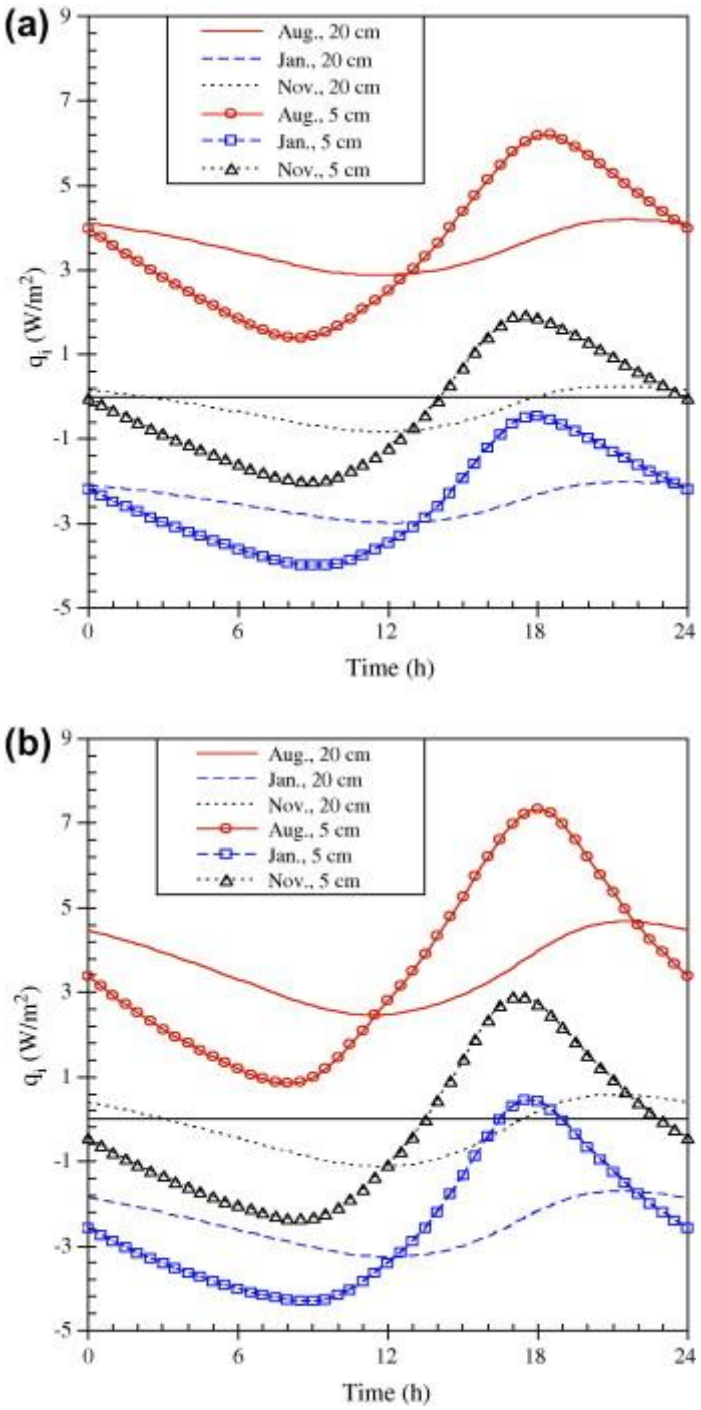


Figure 6 Transmission load variation with time of day in August, January, and November with thermal mass thickness (L_{mas}) of 20 cm and 5 cm; (a) for wall W1 (outside insulation R-value=2.86 m^2 K/W) and (b) for wall W2 (inside insulation, R-value=2.86 m^2 K/W) [32]

3.3.2 Gap between actual and predicted heating demand in buildings

Large differences between actual and predicted energy performance in buildings are observed, ranging from 30% up to 100% [30], [35]–[37]. This difference could be due to different influencing parameters in predicted and actual energy demands.

Majcen studies the influence of several parameters on the actual and predicted heating demands by grouping these parameters in 3 different groups: (1) dwelling characteristics, (2) household characteristics, (3) occupant behaviour. The dwelling characteristic group includes both building and installations (heating & ventilation) characteristics. The household characteristic group is referred to parameters such as incomes or number of occupants. While the occupant behaviour is related to the way the household operates the house, such as ventilation habits, setting thermostat temperature or occupancy profile.

Figure 7 shows the effect expected on the different parameter group on actual and theoretical heating demands, where the thickness of the arrows represent the expected influence magnitude. The study is made in The Netherlands where the space heating is mostly done by gas systems, therefore more than the 95% of gas consumed is for heating proposes. Therefore, the gas use can be translated as heating demand [38].

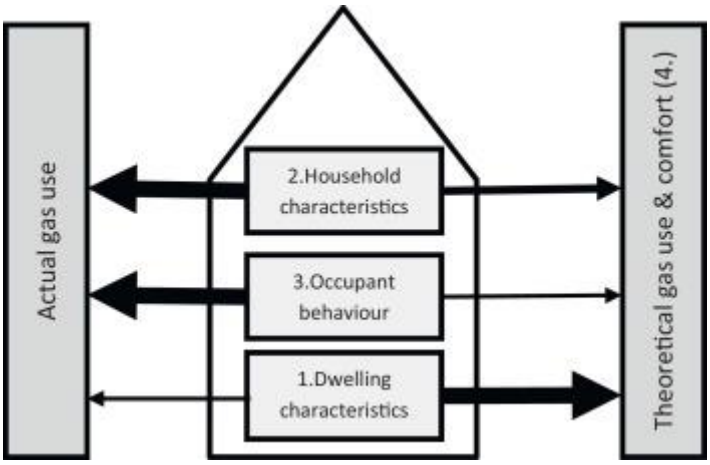


Figure 7 effects of different parameter groups on actual and theoretical heating demands (gas use) [38].

The conclusions of this study shows that the theoretical calculations are dominated by the dwelling characteristics while in the actual heating demand, the household characteristics and the occupant behaviour are the most influencing groups of parameters. This is due to the fact that in the theoretical calculations, the occupant behaviour are included as a normalized parameter, depending mostly on building and system characteristics (dwelling characteristics) [38].

In the present research, the TU Delft buildings are all owned and operated by the university, therefore the household characteristic group is not relevant. However, the occupant behaviour could have an influence in buildings where the thermostat or heaters can be manipulated. The occupant behaviour will be refer to as operating characteristics in the next pages of this report.

3.4 Conclusions

This literature review gave directions to this study with regards to the selection of the prediction model, most influencing parameters on actual thermal energy demand, and the gap between actual and simulated heating demand in buildings. As a result of this review, the following conclusions are drawn:

Concerning the selection of the prediction model technique and statistical method.

- ✓ *The maximum accuracy of a building model is determined by the optimal complexity which is defined by the information available.*

  The selection of the optimal prediction technique for buildings is determined by the optimal complexity level which is dependent on the information available (building & installation characteristics, operating characteristics). The more complex is the model (higher number of physical equations), the higher the amount of parameters needed. Therefore, a more complex model will only be more accurate than a simpler one when all the required parameters are known. This limits the complexity of the building model to the information available.

- ✓ *The accuracy of the thermal energy demand prediction can be improved by using grey-box or black-box models instead of white-models since the first two use a smaller amount of parameters in accordance with the limited information available in buildings.*

  The information available in buildings is limited, especially in older buildings. Since physical models (white-box models) require a large amount of parameters (most of them unknown), it is concluded that the use of physical models for the thermal energy prediction is not the optimal prediction technique. The accuracy of the output can be improved by using techniques depending on a smaller amount of parameters (grey-box or black-box models) which can be determined more easily.

- ✓ *Multivariate regression approach is selected for this case study due to its simplicity, fast calculation and good performance for short term data set.*

  For this case study, statistical methods are selected as the optimal approach considering the required objectives of simplicity and fast calculation. The regression approach is preferred over the Fourier series as the first one is simpler and better for short term data set (seasonal). The main challenge of using this technique to build the mathematical model is to make it physical meaningful and to reach the required accuracy to control and optimize the heat demand of the buildings.

Regarding the study of the most influencing parameters on actual heating demand.

- ✓ *The accuracy and simplicity of the mathematical model will rely on the selection of the most influencing parameters on the thermal energy demand.*
- ✓ *It is important to have a good understanding of the operating characteristics of the building (thermostat settings and ventilation flow rates) since their influence on the heating demand is higher than the building characteristics (physical parameters) for all performance level buildings.*
- ✓ *A good knowledge of the type of heating systems and the related influencing parameters is aim since they may vary the influencing variables (especially in buildings with lower energy performance)*
- ✓ *From all building characteristics parameters, the most influencing one is the U-value of the walls. Therefore, it is important to have the exact U-values of walls, while the U-values and g values of the windows can be estimated.*
- ✓ *In buildings with a lower thermal energy performance, the parameters related to solar gains, such as windows g-value and orientation of the building, may have a higher influence on the heating demand of the building.*

- ✓ *The thermal mass effect modifies the thermos-physical behavior of the building (thermal inertia), influencing the thermal energy savings, hourly heating demand fluctuations and reaction to heat fluxes.*

  The increase in the internal thermal mass of the building reduces indoor air temperature fluctuations and increases time lag, leading to shaving heating demand peaks and a delayed reaction to heat fluxes.

- ✓ *The thermal inertia of the building is enhanced by placing the insulation layer outside*

  The location of the insulation in walls influences the effect of the thermal mass. In order to optimize its beneficial effect on thermal energy savings and damping fluctuations, the insulation layer should be placed outside.

Concerning the study of the gap between actual and predicted heating demand.

- ✓ *The most influencing parameters may differ from actual and simulated heating demands, leading to large differences between actual and predicted heating demands.*

  The simulated heating demand is dominated by the building and installations characteristics, while the operating characteristics of the building has almost no influence since these parameters are normalized in the theoretical calculations. In contrast, the actual heating demand is mostly influenced by the operating characteristics, the building and installations characteristics may have less influence.

4 Research methodology

This chapter gives an overview of the research methodology followed in this study. The methodology has been changed during the research process due to enhanced insights.

Figure 8 illustrates the flow scheme of the final methodology followed. The continuous arrows indicate the path followed in this study. The complete research procedure can be divided in 3 main parts: (1) Analysis of the actual case study (chapter 5) and physics based model (chapter 6); (2) analysis differences between simulated & actual data and data set selection (chapter 7); and (3) statistical models design & validation.

The initial plan of this master thesis was to build a statistical model based on actual data and simulated data (with a physics-based simulator) which would involve only part (3). However, after analysing the real case study, it is concluded that some of the actual inputs are wrong estimated, and therefore the actual data set is considered incomplete. Therefore, only the data base generated by the physic-based model (simulated data) was used for the design of the mathematical model (multivariate linear models) and further analysis were needed. This aspect increased the scope of this thesis by including parts (1) and (2). These two new parts gave valuable insights which are contributing to improvements in IPIN project at the TU Delft District Heating Grid.

The accuracy of the simulated data was evaluated by studying the accuracy of the physics-simulator. As a result, several improvements are applied in the model and a validation test is performed (chapter 6). From this study, it is concluded that the simulated data are accurate to build the mathematical model.

Based on the knowledge gained on both actual case study and simulator, the actual and simulated heating energy demand patterns are compared and the gap between them is studied (chapter 7). From this analysis, it is concluded that the simulated data can replace the actual data and that this replacement will lead to a mathematical model providing a better energy and comfort performance. As a result, the simulated data were used to build the mathematical model.

The multivariate linear models were built by making correlations of several parameters with the corresponding heating demand. The design of the multivariate linear model is an iterative process which leads to several models with different parameters and accuracies. Firstly, the parameters to be introduced in the model are analysed and selected. The combination of the selected parameters are statistically validated by applying a statistical search procedure. The search procedure validates the model at each step by analysing the residuals of the data set, and quantifying the significance level of both the variables' coefficients and the total model. Therefore, the accuracy and fitting profile of the total model is obtained and analysed. When the accuracy of the model is not enough to be implemented in this case study, further improvements are included and a new model is built. The process finishes when the predictive accuracy of the model is high enough to apply this model in this case study.

The process of the multivariate linear models design has followed both reasoning methods: (1) deductive reasoning (theory→hypothesis→observation→confirmation) and (2) inductive reasoning (observation→pattern→hypothesis→theory). The inductive reasoning is used to analyse the pattern and relationship between the most influencing parameters, leading to a preselection of parameters (see chapter 7). The final selection of parameters is done through deductive reasoning by selecting them according to the building's thermal energy balance principle (see chapters 8 and 0).

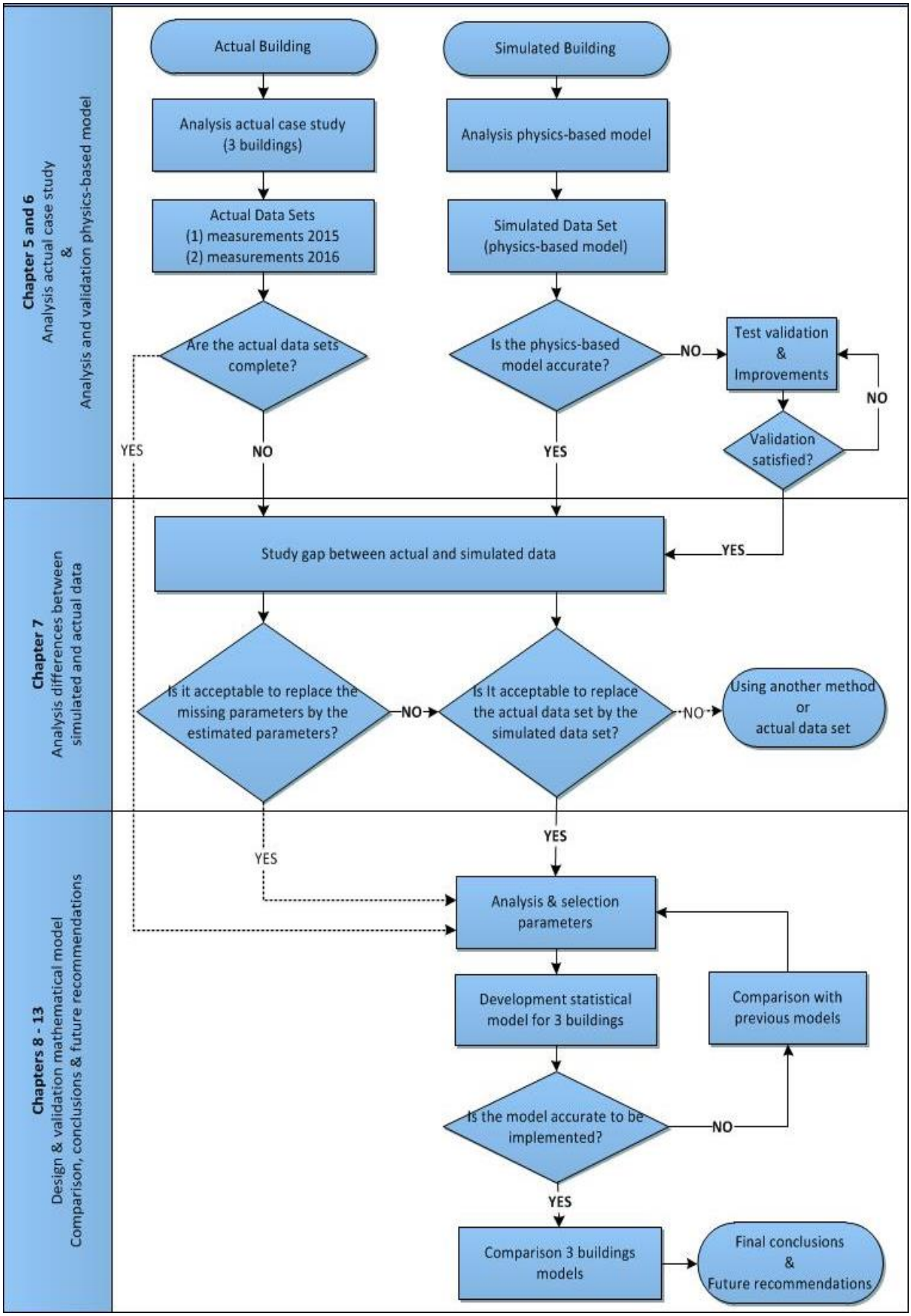


Figure 8 Flow scheme master thesis methodology [own illustration].

5 Description and analysis actual case study

This master thesis analyses the thermal energy behaviour of three buildings at the TU Delft Campus. This chapter gives an overview of the case study and presents a description of the physical and operational characteristics of the buildings & systems for the three buildings studied.

The knowledge gained in this chapter is key to analyse the actual data set and determine the number of equations to build the multivariate linear regression model.

5.1 Overview case study

Most of the buildings at TU Delft campus are connected to the TU Delft District Heating Grid (DHG). The DHG is divided in 4 different branches: North 1 (Noord 1), North 2 (Noord 2), South 1 (Zuid 1) and South 2 (Zuid 2). The buildings studied in this master thesis corresponds to the buildings being tested for the current phase of IPIN (phase 2), which are the ones connected to branch North 2. These are 3mE (Faculty of Mechanical, Maritime and Materials Engineering), IO (Industrial Design) and TPM (Technology, Policy and Management).

Figure 33 shows an overview of the different heating distribution stations of branch North 2 and the situation of the different buildings analysed (3mE in green, IO in yellow and TPM in blue), and the Central Heating Plant (CHP). Branch North 2 contains a total of 11 heat distribution stations of which 1 is situated in TPM, 2 in IO and 3 in 3mE [1]. For further details regarding the current district heating system at TU Delft campus (heating distribution & generation), please refer to 0.

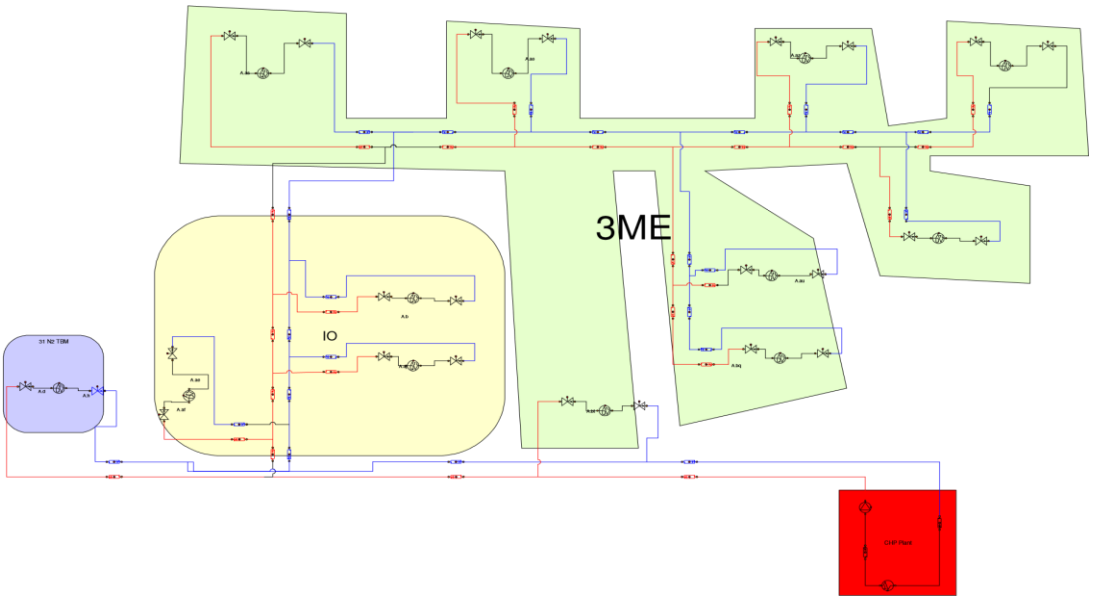


Figure 9 Overview heat distribution stations branch North 2 of TU Delft District Heating Grid [1].

5.2 Building characteristics

This section gives an overview of the building characteristics for IO, TPM and 3mE, and draw some hypothesis regarding the thermal behaviour for each of these buildings. Table 2 presents the most relevant parameters for the three buildings according to the literature review done in section 3.3. Other parameters are presented in Appendix 5. These parameters were estimated during the first phase of IPIN by doing a building inventory or calibrations with real measurements (see section 6.2.1 for further details). For the analysis done in this section, it is assumed that these parameters corresponds to the reality.

Table 2 Main building Characteristics of the buildings analysed

Parameter	Units	TPM	IO	3mE
Gross floor area	m^2	12,000	23,300	36,402
Total indoor air volume	m^3	20,580	46,972.8	64,213.128
Specific air volume	m^3/m^2 floor	1.71	2.02	1.76
Ratio envelop area/gross floor area	-	0.175	0.687	1.34
Specific thermal mass of the building*	Kg/m^2 floor	600	250	300
U-value of windows*	W/m^2 K	3	2.5	2.2
U-value of roof & wall*	W/m^2 K	0.25	0.4	0.4
Ratio windows area/façade area	-	0.098	0.236	0.56
Ratio windows area/floor area	-	N: 0.016 E: 0.0008 S: 0.007 W: 0.002	N: 0.023 E: 0.03 S: 0.02 W: 0.03	NE: 0.05 SE: 0.12 SW: 0.05 NW: 0.12

*parameters estimated by calibrations (see section 6.2.1 for further details).

3mE is the oldest of the three buildings, it was built in 1953 and renovated in 2003 [40]. It has an architectural design were facades and large windows are predominant. IO is dating from 1973 and it was renovated in 2000 [40]. Its shape is more compact and the facades have less windows. TPM is the newest and has the most compact shape of the three buildings, it was built in 2001.

The dates of the buildings indicate the building regulations under which they were built, giving a good indication of the envelop properties. Most of the heat is lost through the building envelop (roof+windows+walls), therefore its insulation level and envelop to floor area ratio are good indicators of the correlation with the outdoor weather conditions. The insulation in walls/roof and windows for IO and 3mE are similar, these are estimated to have a U-value of 0.4 and 2.5 W/m^2 K, respectively. TPM have a higher insulation level in walls but a lower insulation level in windows, 0.25 and 3 W/m^2 K, respectively. However, the percentage of windows in the façade is smaller than IO and 3mE, therefore TPM is better insulated than IO and 3mE. Besides that, TPM has the smallest envelop to floor area ratio, followed by IO (6 times higher) and 3mE (almost 10 times higher). Therefore, according to the envelop insulation and envelop to floor area ratio, it is expected that TPM has the lowest correlation with the outdoor weather conditions, followed by IO and 3mE.

The solar gains of a building are influenced by the quantity, orientation and properties of the windows and thermal energy performance (see section 3.3). The highest windows to façade area ratio is in 3mE, therefore it is expected to be the building with the highest solar radiation influence, followed by IO and TPM. The penetration of the solar radiation inside the building is produced during the morning (East facade) and afternoon (West facade) since the sun has a certain angle with the vertical plane. According to the windows to floor area ratio, during the morning 3mE is the building receiving the highest solar radiation, followed by IO. During the afternoon, both buildings may have similar solar gains. The solar gains in TPM are expected to be negligible. According to the literature review, it is expected a higher influence of the solar gains for the buildings with lower thermal energy performance. In this case, the buildings with lower expected energy performance are also the buildings with the highest windows area ratio. Therefore, the previous classification is in line with this theory.

Internal thermal mass provided in floors and interior walls can absorb and store penetrating solar radiation and internally generated heat. Therefore, for building with higher thermal mass, this effect will reduce the air temperature elevation during daytime and decrease the heating demand peaks. Besides that, the occurrence of peak temperatures and heating demand are expected to be delayed for several hours (see section 3.3 for further details). This effect is expected to be more pronounced in TPM (600Kg/m^2) than in 3mE and IO (300 and 250 Kg/m^2, respectively). However, the solar radiation and internal heat gain will promote the thermal mass effect. Therefore, it is expected a higher thermal mass influence effect in 3mE than in IO.

5.3 Thermal systems and operating characteristics

This section describes the thermal systems, temperature controller and different operating modes for each of the buildings studied. Further details regarding the installations, set points and use of the building can be found in Appendix 5.

5.3.1 Cooling and heating systems

The heat needed to acclimatize IO and TPM is completely supplied by TU Delft heating network and the cooling is through electrical air conditioning systems [40]. TPM has one heating distribution unit, while IO has two units (see Figure 33 in section 5.1).

3mE is partly heated and cooled an underground heat and cold storage system with are also connected to heat pumps (aquifer system) [40]. The remaining heating demand is supplied by the TU Delft heating network and the remaining cooling demand by an air conditioning system. 3mE has 8 heating distribution units (see Figure 33 in section 5.1).

Influence of cooling on heating demand

During moderated weather months in which both heating and cooling are present during the same day, the cooling mode will influence the heating demand. This combination of heating and cooling may have also an influence in the thermos-dynamic behaviour of the wall [32]. Therefore, this phenomena may vary the most influencing parameters on the heating demand for winter and autumn/sprint or their correlation coefficients between them. If this hypothesis is correct, the accuracy of the mathematical model will rise by building one equation for winter (only heating being used) and one equation for autumn/sprint (heating and cooling being used).

5.3.2 Temperature control system

The influence of the heating system on the heating demand will depend on the temperature control system installed in each building. This section analyses the type of operating systems running in the three buildings analysed. It is distinguished two types of operating systems: (1) operating with indoor air temperature controller (TPM) and (2) operating without indoor air temperature controller (IO & 3Me). Independent on the operating system installed in the building, the heaters can be manipulated by the people present in the building by opening and closing the valves in the heaters in the three buildings.

1. Buildings operated with indoor air temperature controller

The operating system of TPM has an indoor air temperature controller. Therefore, the thermal energy demanded by the building will corresponds to the required energy to reach the required indoor air temperature (defined by the indoor air temperature set point). In this way, the heaters turn on ($Q_{demand}>0$) when the indoor air temperature is lower than the indoor air temperature set point. When the indoor air temperature is higher than the indoor air temperature set point, the heaters turn off and the cooling air handling units turn on ($Q_{demand}<0$).

Temperature control system influencing the heating demand

Since the heating demand is regulated by the indoor air temperature, it is expected that the heating demand will have a strong correlation with the indoor air temperature and the dynamic variables influencing the indoor air temperature profile during day time. These are: internal heat loads, solar gains, infiltrations, mechanical ventilation and indoor surface temperature (influenced by the thermal mass in floors and indoor walls).

2. Buildings operated without indoor air temperature controller

The buildings operated without indoor air temperature controller are IO and 3mE. Therefore, the heating demand is dependent on the controllers installed in the supply side (heating distribution units and/or central heating plant) instead of the demand side (building). The building analysed have 2 types of supply systems, TU Delft heating network and the aquifer system. The heat supplied

to IO is only delivered by TU Delft heating network, while the heat at 3mE is supplied by both TU Delft heating network and the aquifer system.

Temperature control system influencing the heating demand

The heat supplied by the TU Delft heating network is limited by the supply temperature in the current central heating plant which is controlled by the heating curve, dependent on the outdoor temperature (see 0 for further details). Therefore, it is expected that the actual heating demand of the building presents a strong dependency on the outdoor temperature.

Since these indoor air temperature control system does not interact with the building (only with the supply systems), one may think that there is no correlation between the heating demand and the indoor air temperature. However, the heaters can be manipulated by the users (users act as operating control system) of the building which may lead to a correlation with the indoor air temperature and the dynamic variables influencing the indoor air temperature. This correlation is expected to be irregular, and thus softer than in the buildings with indoor air temperature controller.

5.3.3 Operating modes

The TU Delft buildings present 3 different operating mode: (1) weekdays during opening hours, (2) weekends during opening hours, (3) closing hours. During the different operating modes, the temperature set point, internal heat gain and ventilation flow rate vary. The temperature set points and ventilation only vary during opening and closing hours (see Table 3), while the internal heat gain also varies from weekdays and weekends (see Figure 3). Since these parameters are unknown for the real data set, they were estimated during the first phase of IPIN by performing calibrations with real measurements (see section 6.2.1 for further details). For the analysis done in this section, it is assumed that these parameters corresponds to the reality.

A description of the different operating modes is given below.

1. Weekdays during opening hours
    a. The heaters are turned on and the people in the buildings are able to turn them off. The temperature set point is only applicable for TPM where the indoor air temperature is controlled in the range between 21 and 24ºC. For IO and 3mE, the heat delivered to the building is not controlled by the indoor air temperature, therefore the temperature set points are not applicable in these cases.
    b. The ventilation during opening hours is higher than during closing hours. The infiltrations are 0 during closing hour, while during opening hours it is assumed 0.3, 0.1 and 0 for TPM, IO and 3mE.
    c. The internal heat gain profile is similar for the three buildings, increasing gradually during the morning, reaching its maximum point at midday, and decreasing gradually in the afternoon. The absolute value of the estimated specific internal heat gain varies per building (W/m^3 air volume), being 3mE the one with the highest specific internal heat gain and IO with the lowest (see Figure 3).
2. Weekends during opening hours. The parameters related to the temperature, ventilation and infiltrations are the same than during weekdays. The internal heat gain during weekends is estimated to be half of the estimated internal heat gain during weekdays. The internal heat gain profile is considered to be the same during weekends and weekdays.
3. During closing hours (weekdays and weekends), the heaters are normally turned off and they only turn on when the indoor air temperature is below 17ºC. This temperature control is installed in the three buildings.

Table 3 Parameters related with the different operating modes during opening and closing hours.

Parameters	Units	TPM		IO		3mE	
		Opening	Closing	Opening	Closing	Opening	Closing
Schedule*	PM	7-22	22-7	7-22	22-7	7-22	22-7
$T^*_{Set\ point}$	ºC	21-24ºC	15-28ºC	20-24ºC	17-28ºC	21-24ºC	15-28ºC
Ventilation*	[air changes/hour]	3	1.5	2	0.5	2	1.5

*parameters estimated through calibrations (see section 6.2.1 for further details).

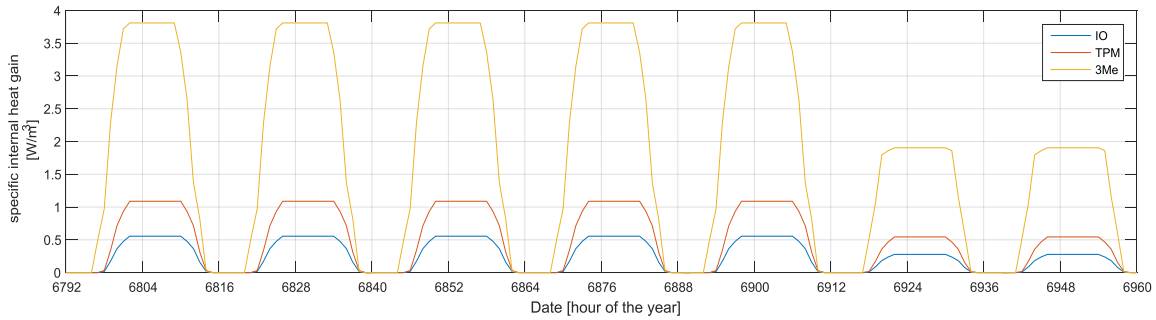


Figure 10 Estimated specific internal heat gain profile for IO, TPM and 3Me during 1 operating week (from Monday till Sunday).

Operating modes interfering in the correlation of variables

The different operating modes interfere in the correlation with the different variables, decreasing the accuracy of the mathematical model. For example, when the set point temperature is 17ºC and the outdoor temperature is 17ºC, the building may not require heating, however if the temperature set point is 25ºC the building is releasing heat to the outdoor environment, therefore heat may be required. This phenomena leads to a different linear correlation of the heating demand with the indoor air temperature (see section 6.3.2). The same can be applied for the internal heat gain and ventilation.

It is concluded that in order to obtain a better linear correlation between the different variables and the heating demand, it will be necessary to make a different equation for each operating mode. This will lead to a better fitting of the mathematical model.

5.4 Conclusions

This chapter described and analysed the physical and operational characteristics of the buildings & systems for the three buildings studied at TU Delft campus. This analysis leads to the following conclusions.

Concerning the number of equations necessary to build the multivariate linear regression model.

- ✓ *It may be necessary to build a different multivariable model for moderated weather months (heating and cooling operating during the same day) and cold months (only heating).*

  The thermo-dynamic behaviour of the wall may be affected when both heating and cooling are operating during the same day. This effect may vary the most influencing variables or their correlation coefficients of the heating demand during winter and moderated seasons (autumn and sprint).

- ✓ *The accuracy of the mathematical model will be increased by defining one equation for each operating mode.*

  The different operating modes affect the correlation between the different variables and the heating demand, decreasing the accuracy of the multivariable model. Therefore, in order to obtain an accurate predictive model, it is necessary to build a different multivariable equation for each operating mode: (1) weekdays during opening hours, (2) weekends during opening hours and (3) closing hours.

Regarding the selection of the data set to build the multivariate linear regression model.

- ✓ *The data set will be selected according to the operating mode and type of months for which the equation is being built.*

- ✓ *In order to optimize the thermal energy and comfort of a building, the mathematical model should be built based on an actual heating demand data set of a building with indoor air temperature controller.*

  The buildings with indoor air temperature controller supply heat only when the building need it, decreasing the heating demand of the building and increasing the indoor comfort. Therefore, only the data survey of TPM could be used in this case study.

- ✓ *The existing data set could be incomplete. If some of the missing parameters are necessary to build the mathematical model, they should be replaced by estimated parameters.*

  In order to use the data survey of TPM, the parameters introduced in the mathematical model should be known. The unknown parameters are estimated by calibrating the simulated heating demand with the actual heating demand (see section 6.2.1). The substitution of missing parameters by estimated parameters in the actual data set could be acceptable when the calibration is accurate enough. Chapter 7 evaluates the accuracy of the calibration process.

6 Description and validation physics-based simulator

The aim of this chapter is to analyse the accuracy of the physics-based simulator used to create the simulated data set. Therefore, this study is key to know whether the simulated data sets are accurate to build the mathematical model.

With the purpose mentioned above, the first section describes the thermal energy principle of the model and analyses the main assumptions. The second and third section evaluate the performance of the thermal energy balance according to the Dutch validation energy test ISSO-54 and a thermal sensitivity analysis.

Previous chapter indicated some of the parameters of the buildings and systems estimated through a calibration procedure before the start of this thesis. Section 6.2.1 explains the calibration process and the main parameters evaluated. The accuracy of the calibrations and the estimated parameters is evaluated in chapter 0.

6.1 Description simulator

6.1.1 Overview and analysis of main assumptions

LEA (Low Energy Architecture) is a simulation program developed by Deerns that calculates the hourly energy needed for heating, cooling, humidification, lighting, ventilation and equipment. The thermal energy is predicted by calculating the energy demand required to reach a corresponding indoor air temperature set point(s) every time step. Therefore, LEA simulates a building operated with an indoor air temperature controller.

LEA simulates the heating demand of the building by modelling a single-zone (instead of multi-zone). Thus, the entire building is represented by a single temperature node, assuming that all the different rooms in the building have the same use function and that in all rooms the same climate prevails. Several studies confirm that modelling the building as multi-zone and single-zone does not produce large differences between the simulated results [29]. Therefore, this assumption is considered to be a good approximation of the real case study.

Figure 11 illustrates the thermal model scheme of LEA with the different physical phenomena considered.

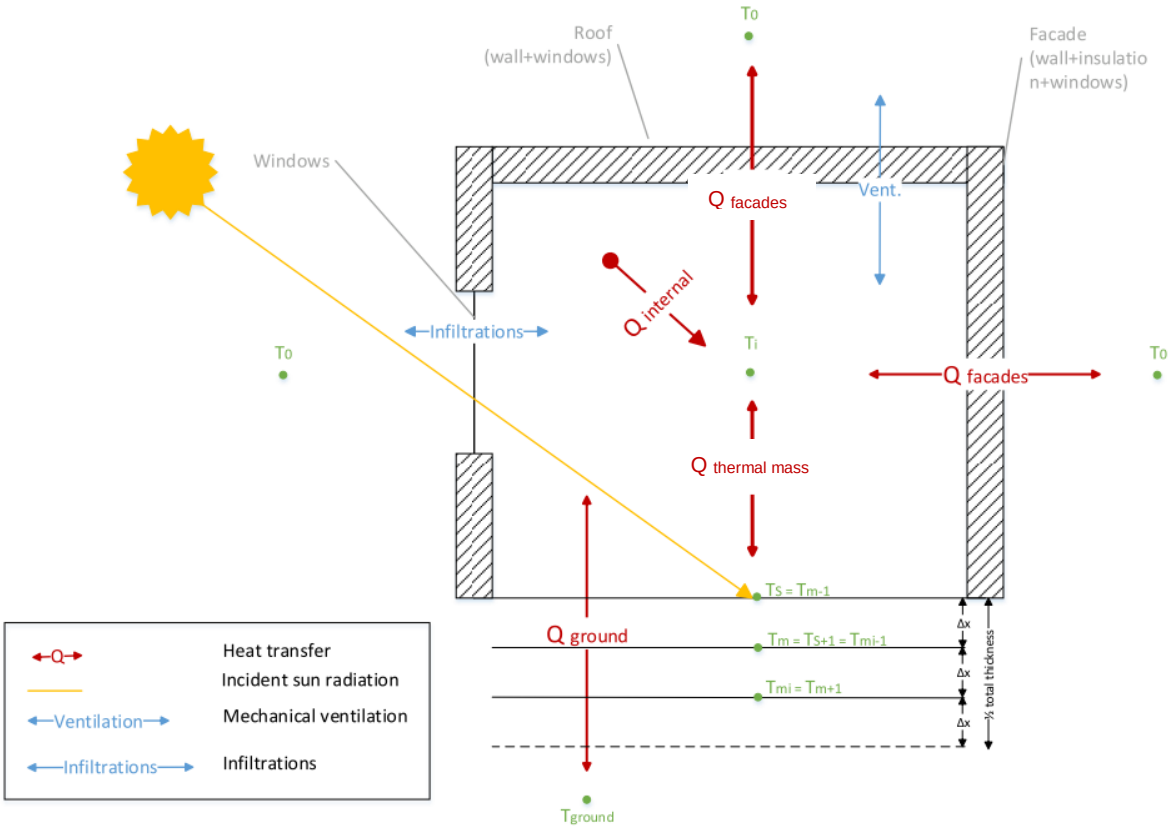


Figure 11 Thermal model scheme of LEA with the different physical phenomena considered [own illustration]

The energy demand for a certain time interval is predicted by calculating the heat transmission through the construction (from floor to ground, facades and roof), mechanical ventilation, infiltrations, solar heat gains, internal heat gains (people, lighting and equipment) and heat accumulation in the thermal mass (indoors). The transmission losses due to thermal bridges are neglected. In more recent buildings, the proportion of heat loss due to thermal bridging is typically 10–15%. This may rise to 30% in better insulated low-energy buildings when insulation and construction details are not properly realized [41]. Therefore, this assumption may have some effects underestimating the simulated heating demand calculations. Section 4.3 of Appendix 4 shows how the thermal bridges can be calculated according to ISO 2007 [42].

The thermal mass of a building defines its thermal inertia which influences the hourly and yearly thermal demand profile. LEA only absorbs heat in the indoor thermal mass. This means that the solar radiation absorbed by the wall outer surface is neglected. In principle this assumption seems to be a good approximation of reality for buildings with wall insulation since a large percentage of the solar radiation stored in the wall outer surface is ultimately dissipated to outside through the same surface [32]. Therefore, the thermal inertia influencing the indoor climate corresponds to the indoor thermal mass which is located in indoor floors and ceilings. LEA accumulates the total mass of the building in the floor. The heat accumulated in the thermal mass is transmitted through convection along the total area of indoor surfaces which are in contact with the indoor air (floors and ceilings). Further details regarding the heat accumulation in LEA can be found in section 2.2 of Appendix 2.

The model is built by combining static and dynamic modelling. All the heat balances corresponding to the room model are calculated considering steady state conditions (static modelling), while the energy balances corresponding to the floor model (heat accumulation) are calculated in an unsteady state (dynamic modelling). The dynamic modelling of the thermal mass represent the

thermos-dynamic behaviour of the wall by using the method finite-difference approximations for one-dimensional unsteady conduction. The accuracy of this method is dependent on the Fourier number defined. When the Fourier number is optimized, this method gives a good approximation of the cyclic temperature variations at the indoor thermal mass layer [25]. Refer to Appendix 2 for further details regarding to the room and floor model.

Figure 12 shows the resistance network corresponding to the thermal model scheme of LEA. The convective heat transfer, radiation heat transfer, conduction heat transfer, accumulation, ventilation and air movements are translated into the symbols indicated in the legend.

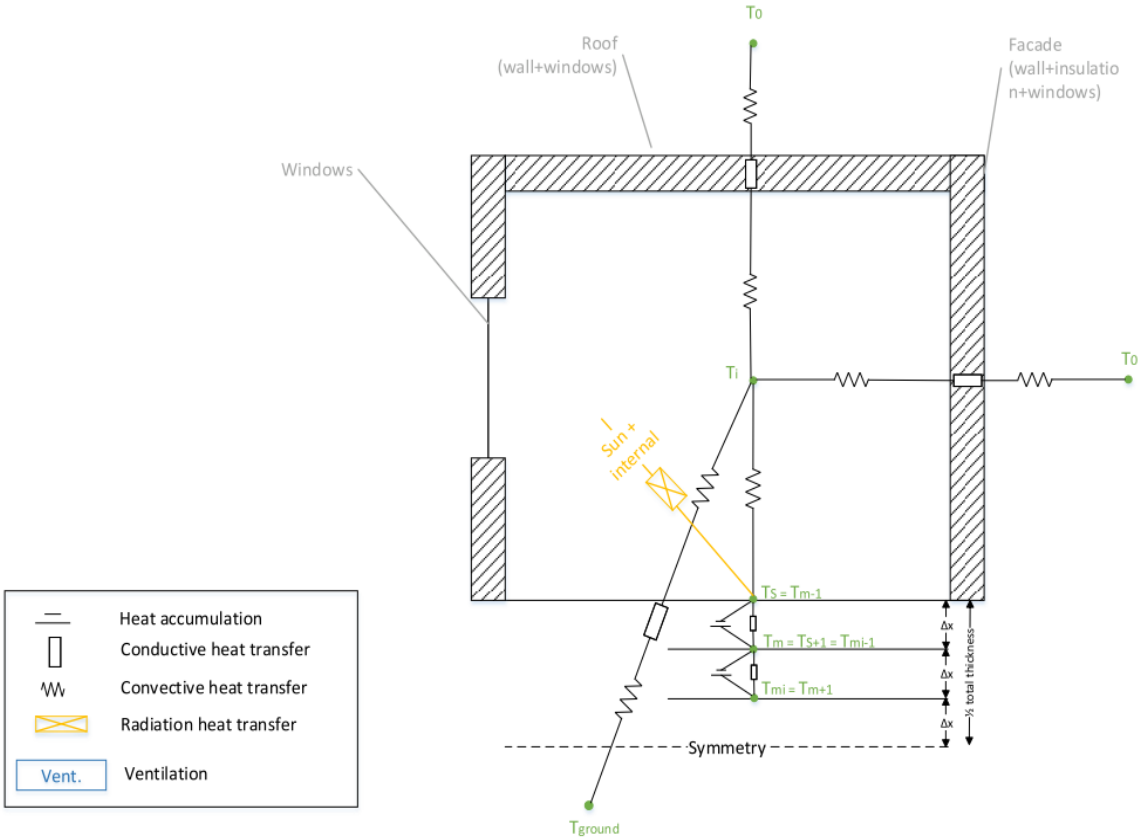


Figure 12 Resistance network of the thermal model scheme of LEA [own illustration].

6.1.2 Thermal energy balance principle

The building heating demand depends on the amount of heat transferred through the building envelop and accumulated in the building. Therefore, these heat fluxes can be divided in 5 main categories: internal heat gains ($Q_{internal}$), solar heat gains (Q_{solar}), envelop losses/gains ($Q_{ground} + Q_{envelop}$), ventilation ($Q_{ventilation}$) and infiltrations ($Q_{infiltrations}$). The indoor thermal mass accumulates the heat absorbed by the solar radiation and internal heat gains. This heat flux ($Q_{thermal\ mass}$) can be assumed to be transferred only with the indoor climate through the total surface of the indoor thermal mass. The thermal energy balance in LEA is described according to the following equation:

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

Where, Q_{ground} and $Q_{envelop}$ are the heat transmission through the floor and the building envelop (walls, windows and roof), respectively. $Q_{thermal\ mass}$ corresponds to the heat transmitted through the total area of indoor surfaces in contact with the indoor air, which simulates the thermos-dynamic behaviour of the indoor surface (thermal capacitance). Q_{solar} and $Q_{internal}$ are the solar heat gains and internal heat gains (people, lighting and equipment), respectively, added to the indoor climate through radiation and convection. $Q_{infiltrations}$ and $Q_{ventilation}$ are the heat transmitted due to infiltrations and ventilations, respectively. Table 4 gives a physical description of the above mentioned heat fluxes. For further details regarding the energy balances and the heat fluxes, please refer to Appendix 2.

Table 4 Description of the heat fluxes present in the thermal energy balance of LEA

Heat flux	Physical description	Parameters definition
Q_{ground}	$Q_{ground} = U_{floor}.\ A_{floor}.(T_{ground} - T_i)$	A_{floor} : area ground floor (building footprint) U_{floor}: heat transfer coefficient of floor T_{ground} : surface temperature of the ground (soil) T_i : indoor air temperature
$Q_{envelop}$	$Q_{envelop} = \sum_i U^j_{envelop}.A^j_{envelop}.(T_o - T_i)$	j: for each façade/roof of orientation $U^j_{envelop}$ heat transfer coefficient of envelop $A^j_{envelop}$ the surface of the facade/roof T_o : outdoor air temperature
$Q_{infiltrations}$	$Q_{infiltrations} = (m_{openings} + m_{cracks}).C_{pair}\ (T_o - T_i)$	$m_{cracks} = V_{building} \cdot 0.15 \cdot \left(\frac{V^2_{wind}}{V^2_{reference}}\right)^{2/3}$ $m_{openings} = \dot{V}_{openings.}.\rho_{air}$ $V_{reference}$: 5m/s $V_{building}$: volume of the building $\dot{V}_{openings}$: volume flow rate (m^3/s), hourly dependent on the weekly and weekend schedule ρ_{air}: air density
$Q_{ventilations}$	$Q_{ventilation} = m_{vent.}.C_{pair}\ (T_{out\ AHU} - T_i)$	Cpair : heating capacity of air (J/kg.K) $m_{vent.}$: mass flow rate of the ventilation air (kg/s). Tout AHU: Temperature of the ventilation air coming out of the air handling unit (AHU).
$Q_{thermal\ mass}$	$Q_{thermal\ mass} = \alpha_i.\ A_{indoor\ surfaces}.(T^t_s - T_i)$	$A_{indoor\ surfaces}$: total area of indoor surfaces in contact with the indoor air. α_i : indoor combined heat transfer coefficients for convection and radiation T_s^t indoor surface temperature

Q_{solar}	$Q_{solar} = Q_{sol\ direct} + Q_{sol\ dif.} + Q_{reflective}$	The solar radiation entering in the building is calculated from the annual and daily solar angle. The diffuse and direct solar radiation are calculated from the global horizontal solar radiation according to monthly average of the reference year.
$Q_{internal}$	$Q_{internal} = n_{people}.Q_{body} + A_{ceilings}.Q_{light} + A_{floor}.Q_{equipment}$	n_{people} is the number of people *A*ceilings : total area of all the ceilings A_{floor} : total area of the floor Q_{body} , Q_{light} and $Q_{equipment}$ corresponds to the heat gain per person, light bulb and equipment unit, respectively.

6.2 Calibration parameters and validation of the simulator

The first part of this section describes the calibration procedure to estimate the building and system parameters before the start of this thesis. The second part explains the validation of LEA and improvements done during this thesis.

6.2.1 Calibration parameters LEA with actual heating demand

During the first phase of IPIN, the physical characteristics of the building & systems, and use & operation of the building were obtained for each of the 20 buildings involved in IPIN project. For each building, around 120 parameters were first estimated by making an inventory, then some of these parameters were calibrated by comparing the simulated heating demand with the actual heating demand (measurements taken during 2015/2016). This calibration was performed before the start of the thesis.

Figure 12 illustrates the flow scheme of the calibration activity which consists in an iterative process. At every time step, the value of one parameter was changed, a new simulation was run and a new comparison between the simulated and actual heating demand was done. This process continued until a reasonable fitting of the simulated data with the actual data was reached. During this task, the following parameters were varied between a logical ranges of values:

- ✓ Physical characteristics of buildings design & systems:
    1. Insulations values (roof, walls, floors and windows);
    2. coefficient of solar radiation through windows (ZTA);
    3. heat recovery efficiency of the HVACs.
- ✓ Use & Operation:
    4. Opening hours;
    5. internal heat gains;
    6. ventilation flow rate during day and night.
    7. temperature set points;
    8. infiltrations;

The main variables changed were the ones related to the use & operation of the building by changing the opening time, internal heat gains and ventilation during the night time. The opening hours were calibrated by matching the start and the end of the heating demand; the ventilation flow rate during night was varied to define the minimum heating ventilation capacity; and the internal heat load to match the heating demand drop during day time.

For some buildings, it was difficult to reach a good fitting during the day, even when the heating load was decreased to 0 W/m^2. In these cases, reasonable assumptions of the previous parameters were made and the following parameters were changed: ventilation flow rate during the day, HVACs heating recovery efficiency and the coefficient of solar radiation entering in the room through the glass (ZTA). If the match between actual and simulated data was still not good enough, then the thermal insulation and the temperature set points were changed.

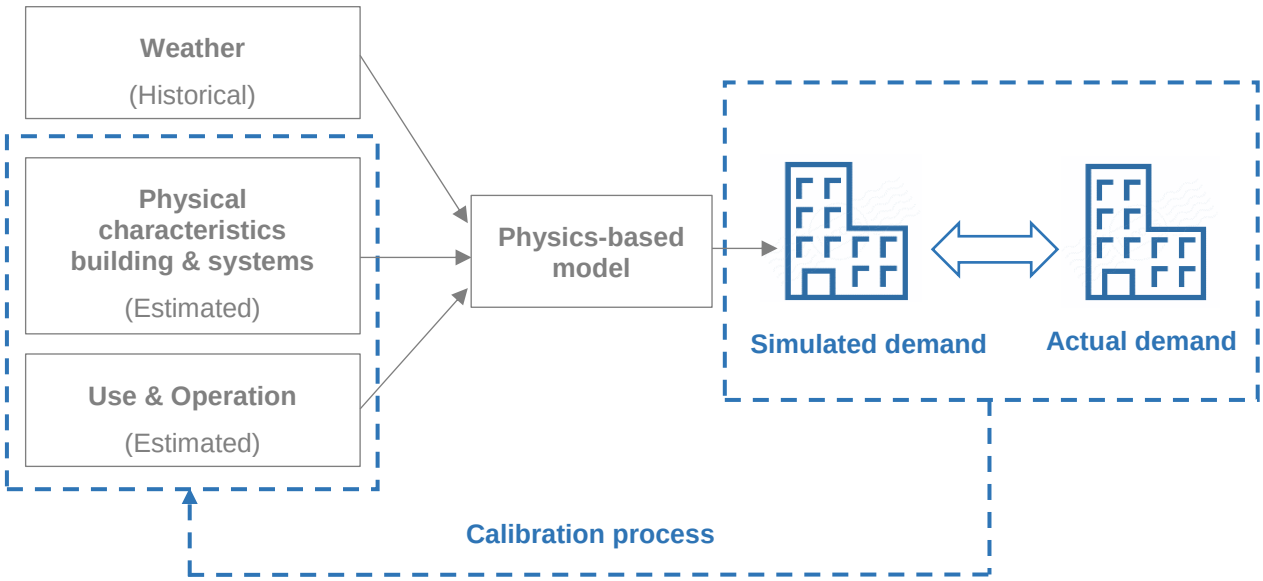


Figure 13 Validation procedure of buildings' inputs in order to use physic-based model as energy predictor [own illustration]

The parameters calibrated has an interaction with several dynamic variables affecting the simulated heating demand and creating a knock-on effect on the other calibrated parameters. This means that a wrong estimation on one of these parameters will have consequences downstream, leading to a wrong estimation of the next parameter. Therefore, in order to complete successfully this task, the actual and simulated data need to be comparable. Thus, the physics-based simulator should be a good representation of the real case study in order to obtain output data comparable with the actual data. Chapter 0 analyses whether the actual and simulated data are comparable, and therefore if the calibration of the parameters was done accurately.

Due to the mismatch between the simulated and actual data in some buildings during the calibration procedure, there were doubts about the accuracy of the thermal energy balance of the physics-based model, therefore a further evaluation of the physics-based simulator was done. Next section explains the improvements done in the physics-based model and the validation performed.

6.2.2 Energy Diagnose and improvements

The performance of LEA was officially validated in the past, however during IPIN some minor changes were done to adapt the software to its new function. It was not expected that these minor changes would have an effect on the thermal energy performance of LEA. Nevertheless, the mismatch observed between the simulated and actual data was an indication that these data set were not comparable. In order to rule out that the source of this mismatch was a low performance of LEA, during this master thesis a new validation to the software was done.

The performance of the thermal energy balance of the version LEA.exe was analysed according to the Dutch validation energy test ISSO-54: 'Energy Diagnose Reference' [43]. The heating demand supplied by the radiators is analysed by applying the tests from A.1.1.01 until A.1.1.16. These tests evaluate the annual heating demand by varying the U-values, thermal mass, opening hours, natural ventilation and internal heat gains (people, lighting, appliances and solar). Appendix 3 gives a detailed explanation about the procedure followed to perform this test in LEA and the results obtained in each of the tests.

The annual heating demand in all the tests performed were within the bandwidth of the reference values. Therefore, it was concluded that the heat balance used in the calculation method of LEA.exe makes a good estimation of the annual heating demand. However, this test does not analyse the hourly thermal energy demand of the building, therefore the hourly fluctuation of the heating demand was not evaluated. This means that a low prediction of heating demand during summer could compensate the high heating demand prediction during winter, leading to a good prediction of the annual heating demand.

A good accuracy for the hourly heating demand is key for the correct performance of the model predictive control, and therefore the smart thermal grid. Thus, a thermal analysis of LEA is done by performing a sensitivity analysis and in depth evaluation of the results of IO. This analysis gives good indications about if the hourly heating demand responds correctly to the input changes. Section 6.3 discusses the main observations during the sensitivity analysis and Appendix 6 presents further results. Besides this validation, some improvements are done in the inputs used for the simulations. Refer to Appendix 4 for further details.

6.3 Thermal analysis simulator

This sensitivity analysis studies the influence of the thermal mass and insulation level on the hourly heating demand profile and the correlation with the outdoor temperature. The thermal mass and the insulation of the facades are the selected parameters to perform this analysis due to their high influence on the hourly heating demand. The effect of these two parameters is studied for IO by varying the thermal mass of the building and the insulation of the facades (only walls part) and roof. This effect is studied during winter period, when only heating is required.

6.3.1 Analysis of the hourly heating demand

Figure 14 presents the simulated hourly heating demand profile for IO during the last week of November 2015 and the first two weeks of December 2015. Both graphs show the results when the R-values of both walls and roof are 1 and 5, respectively. Graph above shows the results for IO with a thermal mass of 600 kg/m^2, while graph below for a thermal mass of 250 kg/m^2.

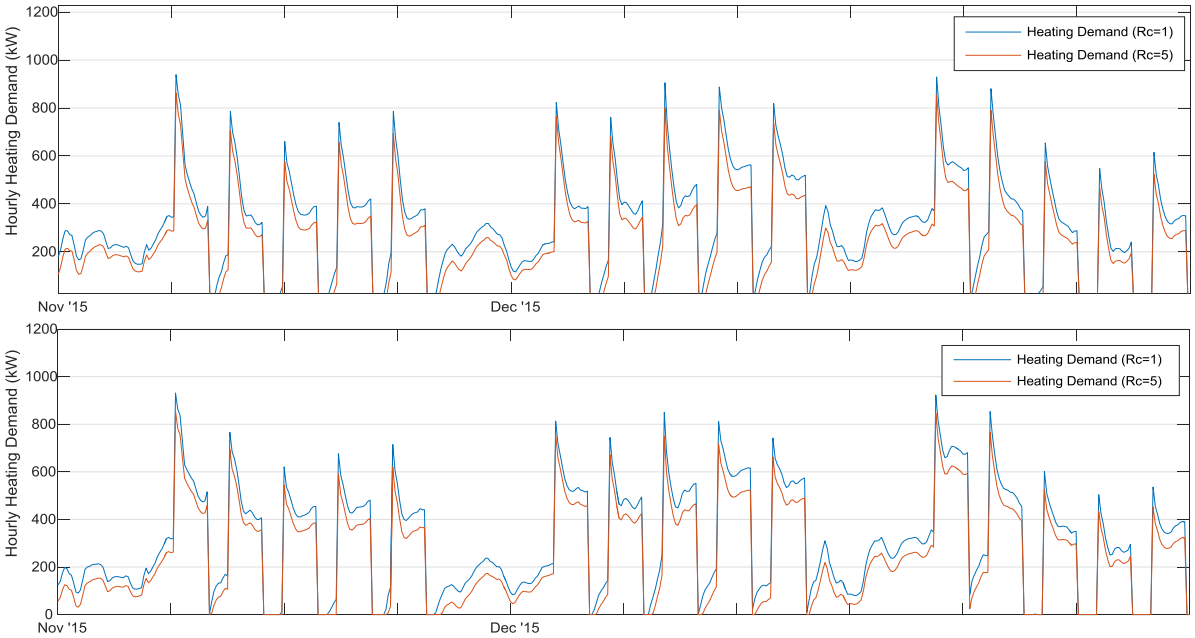


Figure 14 Simulated hourly heating demand profile for IO with façades and roof insulation R-value=1 W/(m^{2}K) and R-value=5 W/(m^{2}K); (above) for a thermal mass of 600 kg/m^2 and (below) for a thermal mass of 250 kg/m^2. Weather data from 21st November – 11th December 2015.

In both cases (graphs above and below), the hourly heating demand presents a higher value for a low insulation level than for a high insulation level. This difference is accentuated after midday, indicating that a higher percentage of the heat added to the building during the morning (through the heaters and the internal heat gains) stays in the building with higher insulation level. As a result, the heating demand after midday drops faster for a building with high insulation level than for a building with low insulation level.

When the thermal mass is higher (graph above), the heating demand peaks increase before midday and decrease after midday due to the thermal mass effect (thermal storage capacity). A building with higher thermal mass has more thermal storage capacity, therefore more energy is required to reach the same surface temperature during heating hours (before midday) than a building with lower thermal mass. Since the heat is exchanged between the indoor surfaces and the indoor air, a building with a higher thermal mass will take more time to reach the indoor air temperature set point than a building with lower thermal mass. After the temperature set point is reached, the heat stored in the thermal mass is released to the indoor environment when the indoor air temperature decreases, reducing the heating demand. Therefore, the higher the thermal storage capacity

(thermal mass), the less heating demand is required after reaching the indoor air temperature set point.

It is important to mention that LEA has two different functions, one for heavy buildings (thermal mass > 100 kg/m^2) and another for light buildings (thermal mass ≤ 100 kg/m^2). Both cases studied in this analysis (600 and 250 kg/m^2) are considered heavy buildings, and therefore modelled in LEA as heavy buildings. The differences observed will be more marked when a building with thermal mass lower than 100 kg/m^2 is simulated.

6.3.2 Linear correlation analysis hourly heating demand versus outdoor temperature

Figure 16.a and Figure 16.b present the scatterplot for the hourly heating demand of IO versus the outdoor temperature when IO is simulated with a thermal mass of 600 kg/m^2 and 250 kg/m^2, respectively. Graphs above shows the relationship for both high and low insulations levels; graphs in the middle for a low insulation level (R-value 1 W/(m^{2}K)); and graphs below for a higher insulation level (R-value=5 W/(m^{2}K)).

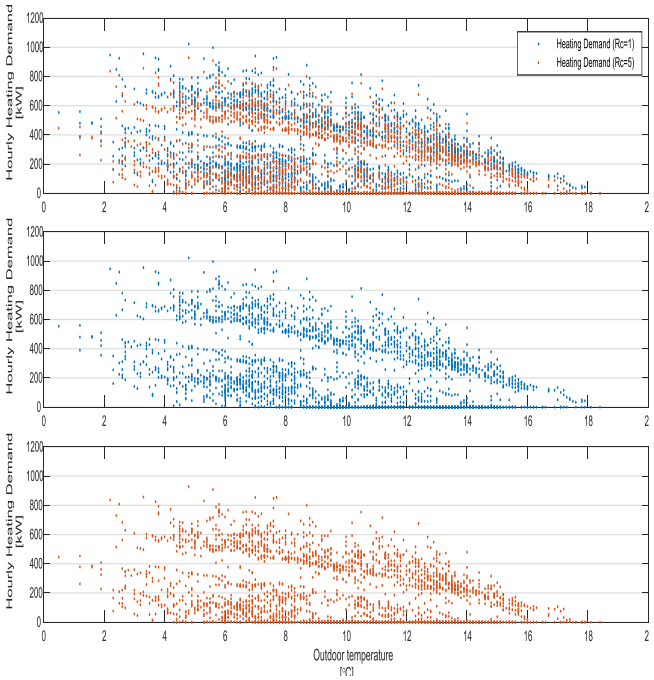


Figure 16.a Influence of the outdoor temperature on the heating demand for IO with a thermal mass of 600 kg/m^2; (above) façades and roof insulation R-value=1 W/(m^{2}K) and R-value=5 W/(m^{2}K), (middle) R-value=1 W/(m^{2}K) and (below) R-value=5 W/(m^{2}K). Weather data from 5th October 2015 - 14th Januarv 2016.

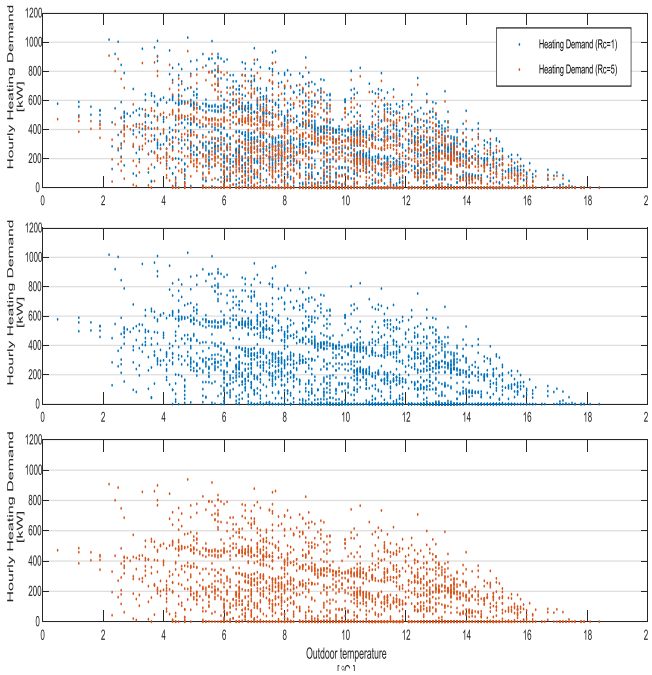


Figure 16.b Influence of the outdoor temperature on the heating demand for IO with a thermal mass of 250 kg/m^2; (above) façade insulation R-value=1 W/(m^{2}K) and R-value=5 W/(m^{2}K), (middle) R-value=1 W/(m^{2}K) and (below) R-value=5 W/(m^{2}K). Weather data from 5th October 2015 - 14th Januarv 2016.

For all cases, it is observed two different well defined cloud of dots which corresponds to the two different operating modes (during day and night). The indoor air temperature set point and the ventilation flow rate are lower during night than during day, leading to less heat required for the same outdoor temperature. This indicates that the different operating modes affects the linear relationship of the heating demand with the influencing variables (eg. outdoor temperature, solar radiation, etc.). Therefore, it is very important to build a different mathematical model for each operating mode.

The dispersion of the cloud indicates the linear correlation of the Y-axis with the parameter in the X-axis. In this case, the higher is the dispersion of dots in the cloud, the weaker is the linear correlation with the outdoor temperature and the higher is the influence with other parameters (eg.

internal heat gains, solar gains, etc). The clouds of dots in Figure 16.a and Figure 16.b show a higher dispersion for a lower thermal mass than for a higher thermal mass building. This indicates that the heating demand dependency with the outdoor temperature is stronger for a heavier building than for a lighter building. This is due to the fact that a heavier building has a higher thermal inertia which has a damping effect on the heating demand fluctuations, leading to a lower influence of the other factors interfering in the heating demand (internal heat gains and solar gains), and vice versa.

The slope of the cloud of points corresponds to the heat losses with the outdoor environment (building envelop and ventilation losses) [44]. The higher the slope, the more heat losses. Figure 16.a and Figure 16.b shows that the slope increases with a lower insulation level, indicating a higher heating demand for the same outdoor temperature. This slope difference is more marked at a lower thermal mass, meaning that the insulation has higher influences in buildings with lower thermal mass. This is because the heat stored in the building acts as a buffer, damping the heating demand fluctuations.

The hourly heating demand fluctuation and its correlation with the outdoor temperature were as expected according to the parameters varied. Therefore, it is concluded that the heat balance of LEA.exe gives accurate simulated data that can be used for building the mathematical predictive model.

6.4 Conclusions

This chapter analysed the accuracy of the simulator and simulated data sets. With this purpose, the main assumptions of the simulator are analysed and the performance of the thermal energy balance is validated. From this study, the following conclusions are drawn.

Regarding the model assumptions.

- ✓ *LEA simulates buildings operated with an indoor air temperature controller (thermostat).*
- ✓ *The building is simulated as a single-zone. The simulated results obtained with a multi-zone and single-zone model are similar* [29].
- ✓ *The transmission losses due to thermal bridges are neglected. This assumption may have some effects underestimating the simulated heating demand calculations [41].*
- ✓ *LEA only absorbs heat in the indoor thermal mass (not in the outdoor wall layer). This assumption approximates good a building with wall insulations where the heat stored in the wall outer surface is ultimately dissipated to outside through the same surface [32].*
- ✓ *The thermo-dynamic behaviour of the wall is simulated with dynamic model by using the method finite-difference approximations for one-dimensional unsteady conduction. This method gives a good approximation of the cyclic temperature variations at the indoor thermal mass layer when the Fourier number is optimized [25].*

Concerning the accuracy of the physics-based model.

- ✓ *According to the validation energy test (ISSO-54 [43]) and the thermal analysis, it is concluded that LEA makes a good estimation of both hourly and annual heating demand.*

Regarding the selection of the data set to build the multivariate linear regression model.

- ✓ *It is assumed that LEA gives accurate simulated data that can be used for building the mathematical predictive model.*
- ✓ *Further analysis needs to be done in order to evaluate whether the simulated data are a good replacement of actual data. This study is presented in chapter 0.*

7 Comparison actual case study and simulator

A gap between the actual and simulated heating demand in the buildings was observed. Previous chapter analysed whether the cause of this mismatch could be due to the lack of accuracy of the simulator. After a detailed evaluation of the thermal balances and the energy performance via validation energy tests, it was concluded that LEA gives a good estimation of the hourly and annual heating demand. Therefore, further research is done to search the source of the mismatch between actual and simulated data. This chapter explains this gap by analysing the differences between the actual case study and the simulator.

The aim of this chapter is to study the causes of the mismatch between actual and simulated data, and analyse their consequences on the data set choice for building the mathematical model. With this purpose, the most influencing parameters on actual and simulated data are analysed and compared via a qualitative analysis. Based on the result of this study, the calibrations performed previous to this thesis are evaluated in order to determine whether the actual data set is suitable for building the multivariate linear regression analysis. In addition, this analysis leads to a guideline for evaluating the calibrations and recommend a faster and more accurate calibration procedure.

7.1 Mismatch between actual and simulated data

This section presents the mismatch between actual and simulated data observed at the start of this thesis. Figure 17 shows the measured and simulated heating demand profile for the three buildings for a representative week. The simulated heating demand is calculated by LEA. The heating demand is expressed in absolute units. Therefore, 3mE presents the highest heating demand while in TPM the lowest as 3mE is the building with the highest volume of air and TPM with the lowest volume of air.

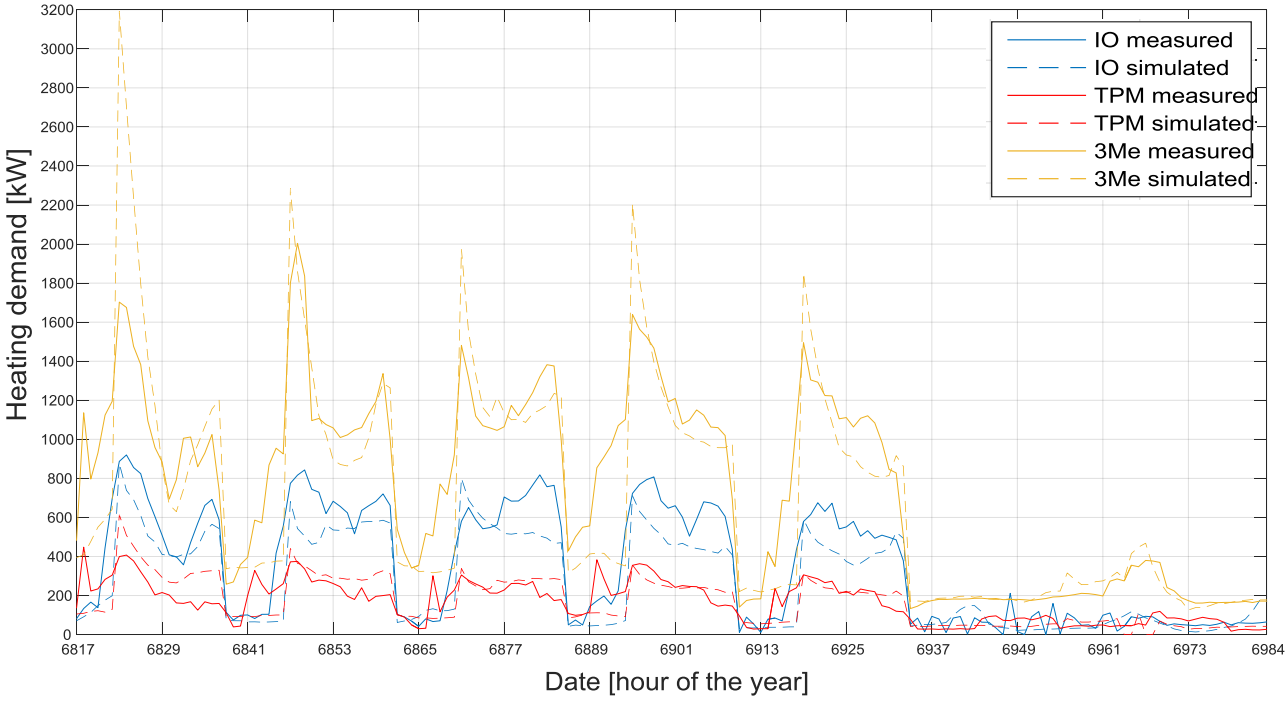


Figure 17 Measured & simulated hourly heating demand profile for the week 12th – 18th October 2015 for IO, TPM and 3Me, respectively.

For IO and 3mE, the measured data do not show a constant decrease of the heating demand during midday and the hourly pattern of both buildings is very similar during weekdays. These are indications that the heat supplied to both buildings is controlled by the same system (supply side: central heating plant by the heating curve), instead of being controlled by the demand side (building). In contrast, the measured data for TPM show a different pattern that the measured data in IO and 3mE, and the profile matches very well with the simulated data. These observations suggest that the heat supplied to TPM is controlled by the building.

Next sections study and discusses with details the causes and consequences of this mismatch between simulated and measured data.

7.2 Analysis gap between actual and simulated data

The gap between actual and simulated data could be due to the fact that the most influencing parameters differ from actual and simulated data as different variables interact in each case. This section presents a qualitative analysis of the most influencing dynamic parameters on simulated and actual data.

This study results in a qualitative matrix which ranks the influence of the dynamic parameter for each building by studying the interaction of the different key indicators. The key indicators are physical characteristics or operational characteristics which differentiate the three buildings analysed and influences the heating demand of each building. Therefore, for the simulated heating demand, the key indicators are based only on the building characteristics, while for the actual heating demand, the key indicators are based on both the building and operational characteristics (different temperature control systems installed in each building and the manipulation of the radiators). Since the key indicators differ for actual and simulated data, the influence on the dynamic parameter could vary for each case. Matrix 1 presents the matrix for the simulated data and Matrix 2 for the actual data. Section 7.2.3 compares these two matrices and gives an insight on the differences between actual and simulated data.

7.2.1 Analysis of the most influencing dynamic parameters on simulated data

This section compares the influence of the dynamic parameters (within one operating mode) on the simulated data for each of the buildings studied by analysing the interaction of the different key indicators. This study results in a qualitative matrix which ranks the influence of each dynamic variable by comparing the three buildings between each other (horizontal comparison). This means that the matrix does not rank the influence of the different variables for one building (vertical comparison).

The key indicators are physical characteristics which differentiate the three buildings analysed and influences the simulated heating demand. In this case, the three buildings have an indoor air temperature controller (demand side control), therefore the buildings characteristics are the predominant factors interfering in the influence of the dynamic variables.

The parameters studied are selected according to the thermal energy balance of the simulator. All parameters are dynamic (time dependent value) within one operating mode. This is because a different mathematical equation will be built for each operating mode (see section 5.3.3). Mechanical ventilation and temperature set points are not included since they have a constant value for the same operating mode. During closing hours (night time), the solar radiation, manipulation of radiators and internal heat gains will not take place since they are null, and therefore static parameter.

Matrix 1 presents the qualitative analysis of the most influencing dynamic parameters on simulated heating demands during opening hours for TPM, IO and 3mE. The matrix should be read horizontally, instead of vertically since the ranking is done comparing the corresponding dynamic variable between the different buildings.

Matrix 1 Qualitative comparison between TPM, IO and 3mE of the most influencing dynamic parameters on simulated heating demands during opening hours (read horizontally instead of vertically).

Dynamic variables	Simulated data			
	TPM	IO	3mE	Key Indicators for simulated data (building characteristics having an effect on the influence of the dynamic variables on the heating demand)
Weather				
Outdoor temperature				↓Insulation, ↑envelop to floor area ratio
wind speed				↓Insulation, ↑envelop to floor area ratio
Solar radiation				↑windows to floor area ratio, ↑façade orientations (W E, S), ↓thermal inertia
Building parameters				
Indoor air temperature				indoor air temperature control system
Indoor surface temperature				↓thermal inertia, (↑solar gains), (↑internal heat gains)
Operating characteristics				
Manipulation of radiators				*No applicable
Internal heat gain				(↑ internal heat gains), ↓thermal inertia

High influence
Moderated influence
Negligible/null influence

() These are dynamic variables, but their heat flux interact with other key indicators

↑↓ The increase or decrease of this key indicator contributes to a higher influence of the dynamic variable.

The manipulation of the radiators do not play a role in the simulated data as it does in the actual data, therefore its influence is null.

The indoor air temperature in LEA is determined by the indoor air temperature set point for the three buildings. This simulates a building with indoor air temperature control system where its heating demand is dependent on the indoor air temperature. As a result, the indoor air temperature will have a high influence on the heating demand.

The wind and the outdoor temperature influence are dependent on the insulation and envelop to floor area ratio. Therefore, these parameters are expected to have its highest influence in 3mE since it is the building with the lowest envelop performance. For IO and TPM, the influence is expected to be moderated.

The solar gains are influenced by the windows characteristics and this effect will be promoted by a low thermal inertia. According to the windows to floor area ratio and windows orientation, the solar radiation is expected to have the highest influence in 3mE, followed by IO, and a negligible effect in TPM. Moreover, the high thermal inertia in TPM damps the indoor air temperature fluctuations, leading to an even lower solar gain influence.

According to the specific internal heat gains, it is expected the highest internal heat gain influence in 3mE, followed by TPM and IO. However, the thermal inertia will decrease the effect of the internal heat gains in TPM. As a result, it is expected a moderated influence in TPM and IO and a high influence in 3mE.

The influence of the indoor surface temperature is increased by higher solar and internal heat gains, and lower thermal inertia which promotes higher indoor surface temperature fluctuations. Therefore, it is expected that the indoor surface temperature has its highest influence on 3mE due to a high amount of solar gains, internal heat gains and a relative small thermal inertia. The lowest influence is expected in TPM due to a high thermal inertia and low solar heat gains.

7.2.2 Analysis of the most influencing dynamic parameters on actual data

This section studies the most influencing parameters on the actual data. The key indicator are physical or operational characteristics which differentiate the three buildings analysed and influences the actual heat supplied to the building. In this case study, the key indicators are dependent on the building characteristics, the different temperature control systems installed in each building (described in chapter 5) and the manipulation of the radiators. In this case, TPM have an indoor air temperature controller (demand side control), while IO and 3mE are controlled by the supply temperature controller in the central heating plant (supply side control). This qualitative analysis is done according to the values estimated for the different parameters during the calibration procedure.

Matrix 2 presents the qualitative analysis of the most influencing dynamic parameters on actual heating demands during opening hours for TPM, IO and 3mE.

Matrix 2 Qualitative comparison between TPM, IO and 3mE of the most influencing dynamic parameters on actual heating demands during opening hours (read horizontally instead of vertically).

Dynamic variables	Actual data			
	TPM	IO	3mE	**Key Indicators for actual data** **(building or operational characteristics having an effect on the influence of the dynamic variables on the heating demand)**
Weather				
Outdoor temperature				↓Insulation, ↑envelop to floor area ratio, temperature control system / manipulation of radiators
wind speed				↓Insulation, ↑envelop to floor area ratio, temperature control system / manipulation of radiators
Solar radiation				↑windows to floor area ratio, ↑façade orientations (W E, S), ↓thermal inertia, temperature control system / manipulation of radiators
Building parameters				
Indoor air temperature				temperature control system / manipulation of radiators
Indoor surface temperature				↓thermal inertia, (↑solar gains), (↑internal heat gains), ↑indoor air temperature fluctuations
Operating characteristics				
Manipulation of radiators				lack of indoor air temperature controller
Internal heat gain				(↑ internal heat gains), ↓thermal inertia, temperature control system / manipulation of radiators

High influence
Moderated influence
Negligible/null influence

() These are dynamic variables, but their heat flux interact with other key indicators

↑↓ The increase or decrease of this key indicator contributes to a higher influence of the dynamic variable.

The influence of the weather parameters is correlated with the façade properties. The wind an outdoor temperature will have its highest influence in the buildings with lower insulation and higher envelop to floor area ratio. Considering these building characteristics, the outdoor temperature and wind will have the highest influence on 3mE. It is expected that IO have a similar influence than 3mE on the outdoor temperature due to a similar insulation level and the same type of temperature control system (outdoor temperature dependent), which may neglect the effect of a lower façade to floor area ratio. However, a lower façade to floor area ratio may lead a moderated influence of the wind in IO due to a lower heat transfer coefficient of the outdoor facade with the outdoor environment and less infiltrations. TPM has the highest insulation and the lowest envelop to floor area ratio, therefore it is expected a moderated influence with both the outdoor temperature and wind.

The manipulation of radiators in the building will have a moderated influence in all buildings since the radiators can be closed and opened by the people. This influence is intensified in IO and 3mE

where the indoor air temperature is not controlled by the building control system and the people are expected to act as the indoor air temperature controllers.

The thermal inertia may have a delaying effect on the solar radiation and internal heat gain, and it influences the indoor surface temperatures. According to the thermal mass, it is expected that TPM has the highest thermal inertia since it is the heaviest of the three buildings.

According to the windows characteristics analysed (windows to floor area ratio and orientation of the windows), the solar radiation is expected to have the highest influence in 3mE, followed by IO, and a negligible effect in TPM. However, the temperature control system at 3mE and IO may moderate the solar gain effect in these buildings since the users will be the only heating demand moderators when the indoor air temperature rises. Therefore, the solar radiation influence at 3mE and IO may be directly related with the manipulation of radiators. As a result, it may be expected a moderated influence of the solar radiation in 3mE and IO. In contrast, in TPM the solar radiation is expected to have a negligible effect. Even though, the temperature control system installed at TPM may make visible the solar gains, the little solar radiation entering in the building and the high thermal inertia of the building are predominating factors, leading to a negligible effect.

Just as the solar heat gains influences, the internal heat gains will be promoted by a low thermal inertia and indoor air temperature control system. In this case, IO and TPM are expected to have the lowest influence of the internal heating gains. IO has the lowest specific heat gains and the temperature control system (independent on indoor air temperature) does not make visual its effect. The high thermal inertia in TPM will decrease the internal heat gain effects. At 3mE it is expected a moderated effect since it has the highest specific heat gains, but the temperature control system decreases this effect.

The influence of the indoor air temperature will be dependent on different parameters due to the different temperature control systems installed in each building. In TPM it is expected that the indoor air temperature control system has a higher influence than the manipulation of radiator, leading to a strong correlation between the indoor air temperature and the heating demand. In contrast, in IO and 3mE, the manipulation of radiators is expected to be the only influencing factor of the indoor air temperature moderation, therefore it is expected a moderated correlation for both buildings.

The main indicators of the influence of the indoor surface temperature on the actual heating demand are the thermal energy sources which increase the surface temperature (internal heat gains and solar radiation), the thermal inertia of the building (internal thermal mass) which indicates the indoor surface fluctuations, and the indoor air temperature fluctuations determined by the indoor air temperature control system. The driving force that releases the heat accumulated in the floor is the temperature difference between the indoor air temperature and the indoor surface temperature, therefore higher indoor air and surface fluctuations will promote this phenomena. The release of this heat leads to the decrease of the heating demand. It is expected that 3mE has the biggest influence of the indoor surface temperature due to higher solar gains and internal heating demand which increases the indoor surface temperature, and higher indoor air temperature fluctuations (due to the temperature control system) which increases the driving force of this phenomena. The influence of the indoor surface temperature in TPM is expected to be negligible due to a low driving force caused by a high thermal inertia (low indoor surface temperature fluctuations) and the indoor air temperature control system (low indoor air temperature fluctuations).

From this study, it is concluded that the interaction of the different key indicators has a large impact on the influence of the dynamic variables. The key indicators may differ for each case study, therefore it is aim to have a good insight of each case study in order to know which key indicators may be predominant and obtain a good analysis. In this analysis, it is considered that the temperature control systems and the manipulation of the radiators have a predominating role, therefore their interference with the other key indicators related with the building characteristics and internal heat gains was marked in some cases, leading to large influences on the dynamic variables.

7.2.3 Comparison of the most influencing parameters for actual and simulated data

Section 7.2.1 and section 7.2.1 analysed the influence of the dynamic parameters on the actual and the simulated data, respectively. This qualitative study ranked the influence of the dynamic parameter for each building by studying the interaction of the different key indicators.Matrix 1 (see section 7.2.1) and Matrix 2 (see section 7.2.1) illustrate the comparison of the dynamic parameters for the three buildings during opening hours on actual and simulated data, respectively. This section compares these two matrices and gives an insight on the differences between actual and simulated data.

The influence of the dynamic parameters differs from Matrix 2 (actual data) and Matrix 1 (simulated data) due to different key indicators that interacts with each other. For both actual and simulated data, the building characteristics and the internal heat gains are assumed to have the same values for each building, and therefore the same influence on the dynamic parameters. However, for the actual data the different temperature control systems and the manipulation of radiators are additional key indicators that dominates the influence of the dynamic parameters, leading to an increase of the differences between actual and simulated data.

The physics-based simulator (LEA) predicts the energy demand required to reach the corresponding indoor air temperature set point(s). Therefore, the simulated data are comparable to a building with indoor air temperature control system (thermostat) where radiators cannot be manipulated. Since TPM has an indoor air temperature controller, the influence of the dynamic parameters are considered to be similar for both actual and simulated data. Therefore, the actual and simulated data may be comparable with each other. Nonetheless, for the real data it is expected to find some irregularities in the heating demand pattern due to the manipulation of the radiators. These irregularities may be easily identified by a discontinuous pattern in the heating demand.

In IO and 3mE the actual heating demand is controlled by the outdoor temperature since the building does not have indoor air temperature controllers. As a result, the outdoor temperature is expected to have a stronger relationship for IO and 3mE on the actual data than on the simulated data, increasing the difference between actual and simulated data. This difference is expected to be more notable in IO than in 3mE as IO has a better envelop performance than 3mE, leading to a lower relationship with the outdoor temperature than 3mE in the simulated data.

The lack of indoor air temperature control system (thermostat) in IO and 3mE leaves the control of the indoor air temperature to the people inside the building which are able to manipulate the radiators, becoming irregular moderators of the indoor air temperature. Therefore, the influence of the indoor air temperature and the parameters affecting the indoor air temperature (especially solar gains and internal heat gains) may decrease for the actual data in IO and 3mE. These three parameters are also expected to have a more irregular hourly profile for the actual data than for the simulated data. The difference between actual and simulated data (influence of the solar gains and internal heat gains) is expected to be more marked in 3mE than in IO. This is due to the fact that 3mE has a higher heat flux of solar gains and internal heat gains than IO.

The indoor surface temperature is expected to be comparable in all cases for actual and simulated data. However in IO and 3mE, it may be expected a higher influence for the actual data than for the simulated data. This is due to the fact that the indoor air temperature fluctuates more in the actual cases than in the simulated ones due to the lack of indoor air temperature controller, promoting the release of the heat accumulated in the surface. For the same reason, in the simulated data a higher influence of the surface temperature is expected in the moments that the indoor air temperature changes, this is during opening and closing hours as the indoor air temperature set points change.

From this study, it is concluded that the effect of the most influencing parameters differ from simulated and actual data due to the interference of different key indicators. The simulated heating demand is dominated by the building characteristics, while the actual heating demand is mostly influenced by the different types of temperature control system and the manipulation of the radiators. As a result, the actual and simulated data are not comparable. Since the calibration of

the parameters (see section 6.2.1) was based on the comparison of these actual and simulated data, it is expected that the parameters are not correctly estimated, especially in IO and 3mE.

Next section analyses the hypothesis drawn in this section and whether the parameters were correctly estimated. The result of this analysis will define if the actual data missing in the data set can be substitute by the estimated parameters.

7.3 Calibration evaluation and new calibration procedure recommended

From the previous section, it was concluded that some calibrations were not done accurately as the actual and simulated data were not comparable. This is applicable to the buildings without indoor temperature control system and in the cases that the cooling effect was not taken into account for the calibrations. Therefore, it is recommended to evaluate the calibrations for the buildings calibrated under the above mentioned conditions. The analytical procedure followed in this section could be applied in the rest of the buildings at TU delft campus that their calibration needs to be evaluated.

This section analyses the accuracy of the calibration procedure by comparing the simulated heating demand with both actual data set (2016 and 2015). The results of this section proves the hypothesis drawn in the previous section and determines whether the actual data set is suitable to be used for building the multivariate linear regression analysis.

The data set of 2015 and 2016 correspond to the measurements taken before and after the implementation of LEA, respectively. Therefore, for the data set of 2015, IO and 3mE do not have indoor air temperature control system, while for the data set of 2016, the indoor air temperature is controlled by LEA. The weather data used for both the simulations and this analysis corresponds to Rotterdam which is the closest meteorological station from TU Delft campus.

7.3.1 Calibrations analysis with data set 2015

For the data set 2015, the measured indoor air temperature was not available or incorrectly measured. The hourly actual heating demand cannot be analysed since the measured and actual data are not comparable. Therefore, it is decided to analyse the correlations between the specific heating demand and some of the most influencing variables.

This analysis is focused on weekdays during opening hours since there are more variables influencing the heating demand. The correlation is analysed for the outdoor temperature, global horizontal solar radiation, wind speed, indoor air temperature and internal heat gains. For weekdays during opening hours the most meaningful correlations are for the outdoor temperature and global horizontal solar radiation, and therefore these are the correlations studied in this section. Appendix 9 and Appendix 10 shows the correlations for the rest of the variables during weekdays (opening and closing hours separately) and weekends for the data set 2015 and 2016, respectively.

In order to compare the heating requirement of the three buildings analysed, the simulated and measured heating demand for the different buildings is expressed in specific heating demand. The specific heating demand is usually represented in [W/m2], however, in this report it is expressed in [W/m^3] in order to take into account the variable height of the building. For this transformation, it is used the total air volume estimated by LEA for each of the buildings (this is 46,972.8 m^3 for IO, 20,580 m^3 for TPM and 64,213.128 m^3 for 3Me).

Insights from the analysis of the outdoor temperature influence

Figure 18 illustrates the scatter plot and the corresponding least square between the specific heating demand and outdoor temperatures during weekdays and opening hours for both actual and simulated data. Table 5 shows the coefficients of the least square for each case, defined as

$$Y = constant + slope \cdot X.$$

Refer to Appendix 9 for further details regarding the influence of the outdoor temperature during weekends.

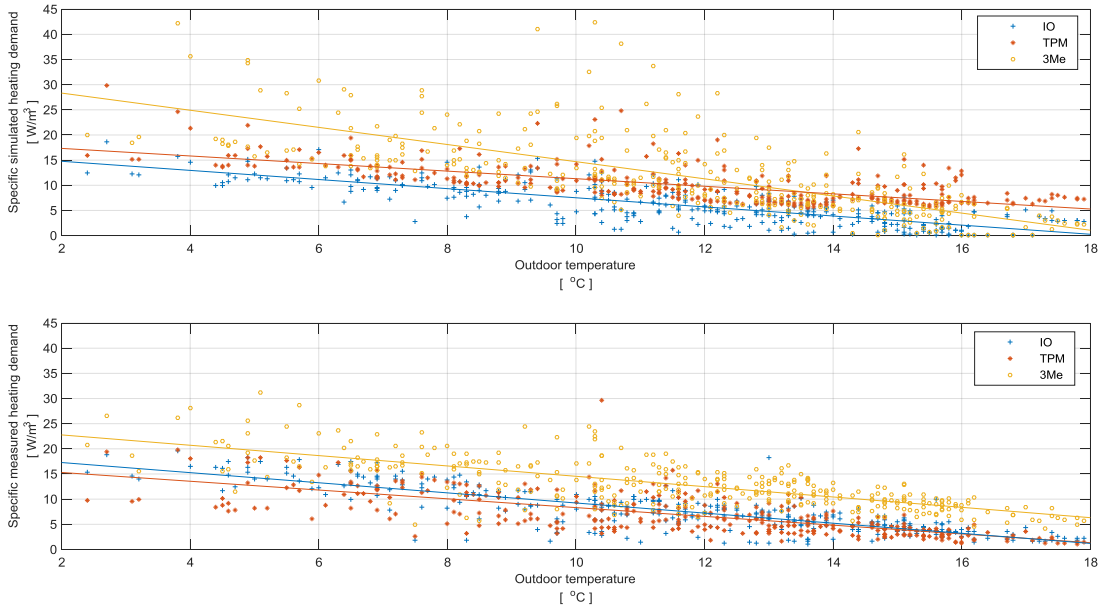


Figure 18 Influence outdoor temperature on the buildings' simulated (above) and actual (below) heating demand during weekdays and opening hours for October 2015.

Table 5 Least square coefficients values for the linear correlation between outdoor temperatures versus simulated heating demand

Data sets →	Simulated data		Measured data	
Buildings ↓	**constant**	**slope**	**constant**	**slope**
IO	17	-0.91	19	-1
3mE	32	-1.7	25	-1
TPM	19	-0.75	17	-0.87

As it was explained in section 6.3.2, the dispersion of the data points in the cloud and the slope of the least square gives information about the linear correlation between the Y-axis and the X-axis. The higher is the dispersion of data points in the cloud, the weaker is the linear correlation of the variable analysed and the stronger is the influence of other parameters. The slope of the least square indicates the influence and or dynamism of the parameter analysed on the heating demand. Therefore, a vertical and horizontal slope indicates that there is not influence of the parameter analysed. A vertical slope shows that the analysed parameter is static (constant value for all heating demand), while the horizontal slope indicates that the analysed parameter is dynamic (changes for the heating demand).

In this case, the slope indicates the energetic quality of the building envelop and ventilation losses. The smaller the slope, the higher the quality [44]. In the graph below (measurements), it is observed that IO and 3mE have the same slope (similar energy performance) while TPM has a lower slope. Therefore TPM presents a higher performance than IO and 3mE. The slope of IO and TPM is slightly smaller for the simulated than for the measured data, indicating that the energy performance of IO and TPM was slightly overestimated. In contrast, the energy performance of 3mE was underestimated. As a result, the insulation values should be increased for 3mE and slightly decreased for IO and TPM, and vice versa for the ventilation losses and infiltrations.

It is interesting to observe that the cloud of data points is higher for 3mE than for IO. This means that 3mE consumes more energy than IO to heat up the same amount of air, being both buildings comparable in their heat losses performance (identical slope). It could be due to an underestimation of the total air of volume in 3mE or and overestimation of the total air of volume in IO due to a wrong approximation of the building dimensions. Considering that the total air volume is correctly estimated, this observation indicates that the indoor temperature set point for 3mE (estimated to be 21^{o}C) is higher than for IO (estimated to be 20^{o}C) and that the heat recovery efficiency of the heating systems are smaller for 3mE than for IO. Since the heat supplied at IO and 3mE is not controlled by the indoor air temperature set point, the second reason is expected to be the main cause for this difference. The estimated heat recovery efficiency for the AHU for 3mE is estimated to be 0.2, while for IO is estimated to be 0.3. In addition, it is expected more heat losses in the heat distribution stations at 3mE than in IO as in 3mE there are 8 stations, while in IO there are 2 stations.

For both actual and simulated data, the scatterplot shows a higher dispersion for 3mE, followed by TPM and IO. This means that the influence of other variables is stronger in 3mE than in TPM and IO. This is mainly due to higher solar gains in 3mE than in IO and TPM (see Figure 19). For IO and 3mE the dispersion for the simulated data is higher than for the measured data. This means that the actual heating supply has a stronger correlation with the outdoor temperature than the simulated heating demand. This was expected since the simulated heating demand is controlled by the indoor air temperature, while in the actual cases, IO and 3mE does not have indoor air temperature controller. In contrast, the dispersion of the cloud for TPM is similar for both the simulated and actual heating demand as the actual building has an indoor air temperature controller, making both data set comparable.

Insights from the analysis of the influence of the global horizontal solar radiation

Figure 19 illustrates the scatter plot and the corresponding least square between the specific heating demand and the global horizontal solar radiation during weekdays and opening hours for both actual and simulated data.

6 shows the coefficients of the least square for each case, defined as $Y = constant + slope \cdot X$.

Refer to Appendix 9 for further details regarding the global horizontal solar radiation during weekends.

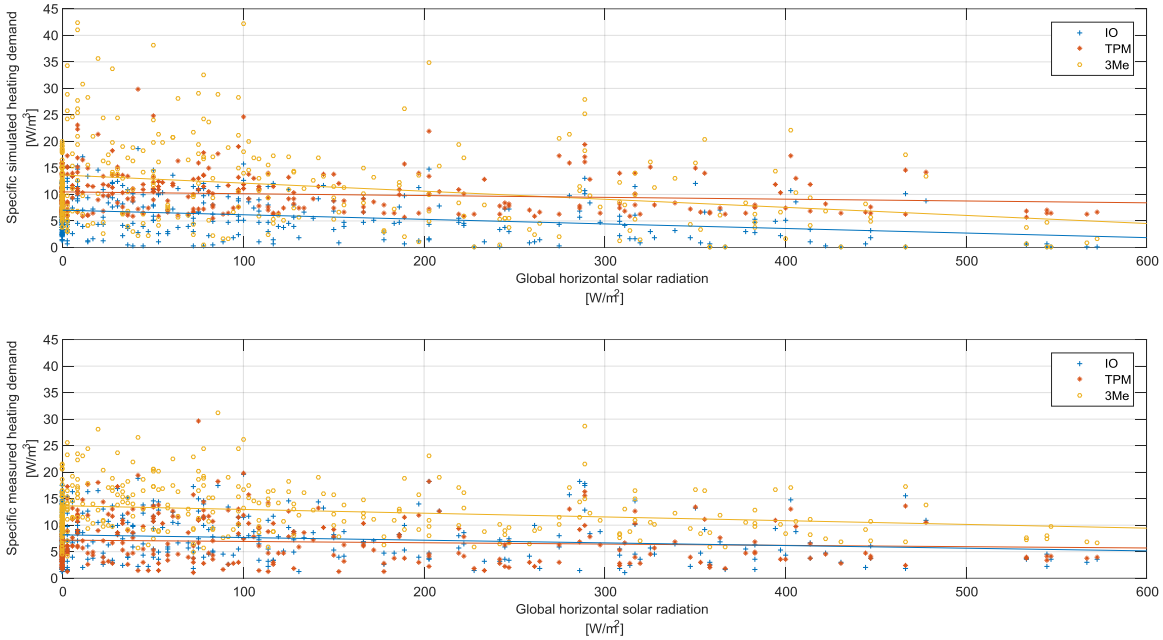


Figure 19 Influence global horizontal solar radiation on the buildings' simulated (above) and actual (below) heating demand during weekdays and opening hours for October 2015.

6 Least square coefficients values for the correlation between global horizontal solar radiation versus simulated heating demand

Data sets →	Simulated data		Measured data	
Buildings ↓	**constant**	**slope**	**constant**	**slope**
IO	7	-0.0086	8.2	-0.005
3mE	14	-0.015	14	-0.0071
TPM	10	-0.0034	7.2	-0.0025

Indicates the accuracy in the calibration of the parameters related with the solar radiation entering in the building (eg. percentage windows, coefficient of solar radiation (ZTA)). However, it is important to note that this solar radiation correspond to the total solar radiation on a horizontal plane. This may be the main reason of obtaining a lower relation between the actual heating demand and the solar radiation. The sun is entering in the building through the E, W and South facades, therefore a better comparison can be made by analysing the solar radiation incident on

the East, West and South vertical planes, respectively. This will give information about the windows characteristics in each façade of the building.

The dispersion of points decrease at a higher global horizontal solar radiation, meaning that the correlation increases at a higher solar radiation and vice versa. This proves that for winter period the solar radiation has a smaller influence on the heating demand than for moderated months (sprint or autumn) during which the solar radiation has a higher intensity. In order to know more about the estimated percentage of windows and ZTA, this analysis should be done with the solar radiation on the corresponding vertical plane.

7.3.2 Calibration analysis with data set 2016

Since 10th of September 2016, LEA was implemented in IO, TPM and 3mE. Therefore, the heating supply in the three buildings is controlled by the heating demand predicted by LEA (based on the simulated indoor air temperature according to the estimated indoor air temperature set points). During the implementation period, actual and simulated data were recorded for both indoor air temperature and heating demand, leading to a new actual data set (actual data 2016). Since LEA is the temperature controller in the three buildings, the actual and simulated data are comparable in this new data set, allowing the calibrations to be validated.

The aim of this section is to evaluate the accuracy of the estimated parameters by analysing the actual data of 2016 and simulated data for both indoor air temperature and heating demand of the building for each building. Figure 20 and Figure 21 present the hourly measured and simulated indoor air temperature and heating demand profile, respectively for TPM, IO and 3mE. The data in both figures correspond to a representative week during the validation period (from 10th until 16th of September 2016).

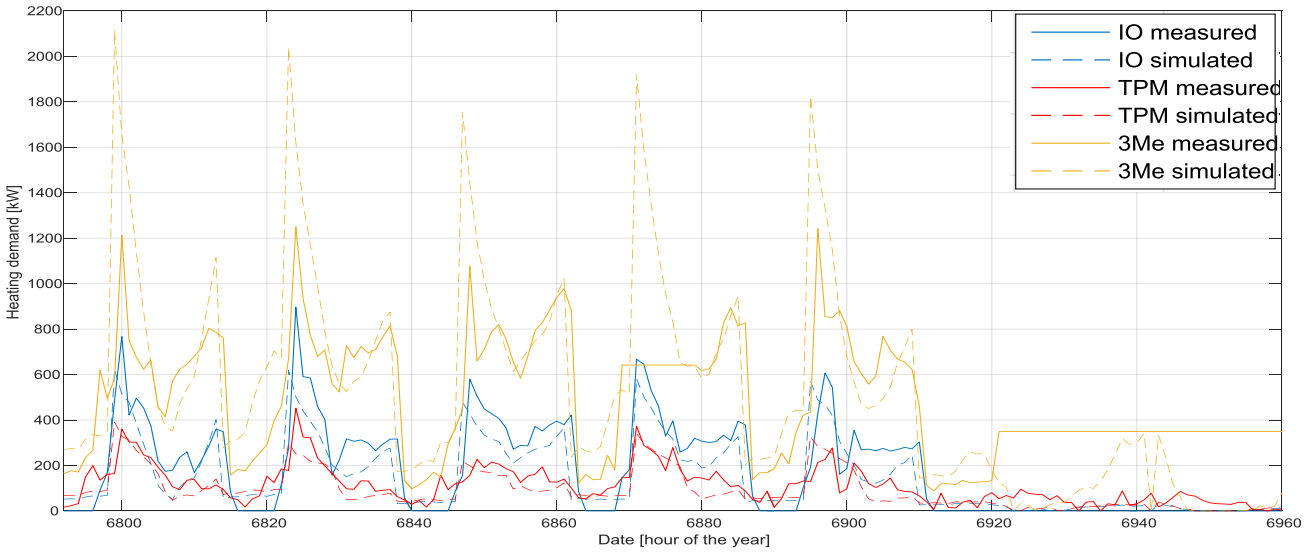


Figure 20 Measured & simulated hourly heating demand profile for IO, TPM and 3Me from for the period 10th – 16th October 2016

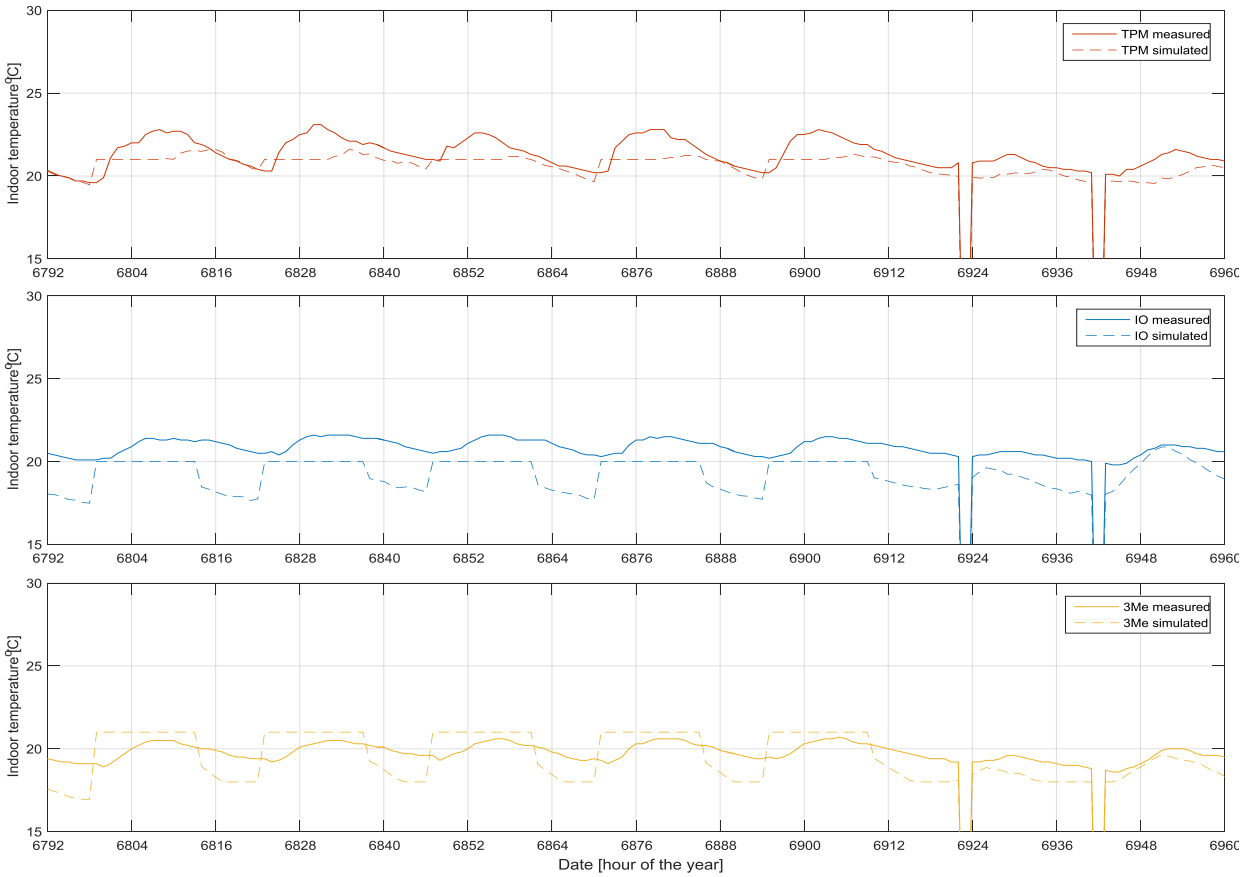


Figure 21 Measured & simulated indoor air temperature profile from 10th – 16th October 2016 for TPM (above), IO (middle) and 3Me (below).

For a correct interpretation of the results, it is very important that the measured indoor air temperature is representative of the average indoor air temperature of the total air volume in the building and the measured heating demand corresponds to reality. For the measured heating demand in 3mE, it is detected a measurement error during the first hours of the morning on Wednesday (between hour 6870 and 6880) and during the weekend (between hour 6920 and 6960). This measurement error is detected by a flat heating demand profile in both cases, while the indoor air temperature profile follows the expected profile. The rest of the measurement data are assumed to be representative of the actual case.

For TPM, it is observed that the measured and simulated indoor air temperature are in phase, therefore the thermal inertia of TPM in the actual case is the same than in the simulated case. This indicates that the thermal mass of the building is well estimated during the calibration procedure. Besides that, the amplitude of the indoor air temperature is 3ºC higher for the measured data than the simulated data, indicating that the building is receiving more heat in the actual situation. Figure 20 shows that the actual heating supplied presents slightly higher peaks than the simulated heating demand during day time. The peaks observed in the actual heating demand are in phase with the temperature peaks. Therefore, it may be that the additional heat increasing the indoor air temperature is caused mainly by the manipulation of radiators. Additionally, the actual heating supply peaks have a constant pattern during all weekdays, which could be an indication that the people inside the building have a constant routine in opening and closing the valves (unlike it was expected).

For both IO and 3mE, the measured indoor air temperature shows a lower fluctuation and a delay with respect the simulated indoor temperature. The delay in IO is around 3 hours while in 3mE around 2 hours. In IO, the actual heating supplied and the heating demand simulated are in phase, indicating that the observed delay is due to an underestimation of the thermal mass. In 3mE, the heat is supplied later than the simulated heating demand, indicating that this delay is due to both an underestimation of the thermal mass and a delay in the heat supply.

A higher thermal mass leads to increasing the heating demand peaks during the opening hours due to a higher heat capacity in the building. This phenomena is observed in IO, where the peaks for the supplied heating are higher than for the simulated heating demand (probably due to manipulation of radiators). In contrast, in 3mE the peaks of the supplied heating demand are lower than in the simulated heating demand, and this difference constantly repeated every day. This could indicate an overestimation of the heating capacity of the heating supply systems (heating distribution stations and/or radiators). Since the heat supplied in IO is higher than predicted and in 3mE is lower, the measured indoor temperature in IO is higher than the simulated one, while in 3mE is lower.

The analysis done in this section proves the hypothesis drawn in the previous section that the estimations of the parameters were incorrect in IO and 3mE as the actual data set of 2015 and simulated data were not comparable. For TPM the parameters are correctly estimated and the manipulation of radiators is mainly the only factor affecting the difference between actual and simulated data. Since the manipulation of radiators have a great impact on the actual indoor air temperature (3^{o}C of difference), this parameter should be measured and included in the mathematical model.

As a result, it is concluded that both actual data set (2015 and 2016) are incomplete to be used and the missing parameters cannot be substituted by the estimated parameters as it would lead to inaccuracies in the mathematical model due to input data errors.

In addition, it was observed that the quality of the measurement data is aim to obtain a good interpretation of the results, leading to a correct validation of the calibration. Therefore, the data collected should be a good representation of the reality and comparable with LEA. In this case, the actual heating demand should corresponds to the total heat supplied by all heating supply systems installed in the building, while the indoor air temperature should represent the average temperature profile of the total air volume contained in the building. This should include the measurements of rooms at each orientation side of the building and the rooms which contains a higher volume of air (eg. canteen, big lecture/study rooms, etc.)

7.3.3 New calibration procedure recommended

This section gives insights of another calibration procedure recommended for the validation of the calibrations done and future estimations of building, systems and operating parameters. With this propose, the correlation between the actual heating demands of IO for the data set of 2015 and 2016 are compared. The data set of 2015 corresponds to the measurements collected before the implementation of LEA, while the data set of 2016 is collected after the implementation of LEA in IO. Figure 22 presents the data set for 2015 above and the data set for 2016 below.

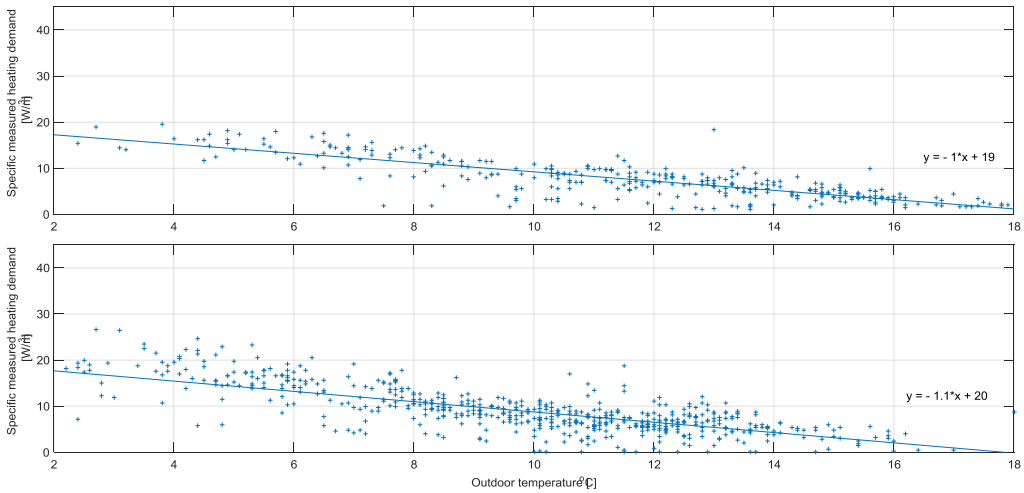


Figure 22 Actual heating demand for IO during weekdays and opening hours before (above) and after (below) the implementation of LEA in the building, respectively. The measurements above corresponds to the data set of 2015 (month of October) and below to the data set of 2016 (from 3rd October 2016 until 25th November 2016)

The data set of 2016 also includes the month of November, increasing the amount of data points for a colder period than for the data set of 2015. As a result, the constant of the least square equation is slightly increased. Since the indoor temperature is controlled for the data set of 2016, the dispersion of the data points decreases with respect the data set 2015. However, the slope does not show appreciable variations. This is due to the fact that the slope is only influenced by the heat losses related characteristics which are constant for each building. This means that both the insulation of the building envelop, and the parameters related to the ventilation losses can be calibrated by comparing the actual and simulated data.

The same calibration procedure can be applied to evaluate the windows characteristics (windows percentage and ZTA) by analysing the slope of the linear correlation between the heating demand and the solar radiation. However, the solar radiation should correspond with the global solar radiation for the vertical plane corresponding to the orientation of the façade analysed.

In order to apply this calibration procedure correctly, the comparison of the slopes have to be done for the same operating mode (temperature set point and heating activated or deactivated). This is due to the fact that a different indoor temperature will lead to a different temperature differential (ΔT = indoor temperature – outdoor temperature), and therefore the heat flux will differ for the same envelop & windows characteristics and ventilation losses. See Appendix 9 and Appendix 10 for further details regarding the slope differences between opening and closing hours. Please note that the slopes during weekends are not meaningful since the opening and closing hours are not distinguished in the graphs.

For temperatures below than 12ºC, the data set of 2016 shows a higher scatter than the data set of 2015. This gives indications that the implementation of LEA in IO has contributed to a lower dependency of the heating demand on the outdoor temperature and higher dependency on other influencing parameters, as it was expected. However, this fact cannot be strongly confirmed since

the data sets analysed can differ due to other parameters influencing the indoor air temperature parameters such as the solar gains, internal heat gains or the manipulation of radiators.

This calibration procedure can be implemented to evaluate other parameters by studying other patterns according to the European prototype for evaluation of building performance [44].

7.4 Conclusions

This chapter analysed the causes of the mismatch between actual and simulated data, and study their consequences on the data set choice to build the mathematical model. The following conclusions are drawn as result of this analysis.

Regarding the mismatch between actual and simulated heating demand:

- ✓ *The gap between actual and simulated data results from the fact that the most influencing parameters differ from actual and simulated data as different key indicators are interfering.*

  The actual heat supplied to the building is influenced by the different types of temperature control system and the manipulation of the radiators. Since these two factors are not playing a role in the simulator, the simulated heating demand is dominated by the building characteristics.

- ✓ *The gap between simulated and actual data is higher for IO and 3mE than for TPM. As a result, the actual and simulated heating demand for IO and 3mE are not comparable, while they can be compared for TPM.*

  In TPM, the manipulation of radiators is the only key indicator which differ from actual and simulated data. However, in IO and 3mE both the manipulation of radiators and the supply temperature control system vary from actual and simulated data. The last one have a great impact on the actual heat supplied, leading to a large mismatch between actual and simulated data.

Regarding the calibration of the estimated parameters (performed before the start of this thesis):

- ✓ *The thermal mass for the buildings without indoor air temperature control system (3mE and IO) is underestimated as a result of a poor calibration performed with incomparable data sets.*
- ✓ *The insulation parameters of the building envelope (windows and wall) should be increased for 3mE and slightly increased for IO and TPM.*
- ✓ *The ventilation losses and infiltrations should be decreased for 3mE and slightly increased for IO and TPM.*
- ✓ *It is observed that IO and 3mE have the same envelop characteristics, however the specific heating demand at 3mE is higher than in IO. This low energy performance of 3mE may be due to a lower energy performance in the heat distribution systems. This conclusion is based on the assumption that the total air volume of the buildings is correctly estimated.*
- ✓ *The calibration procedure based on the analysis of the least square on the scatter plot (section 7.3.3) gives better insights on the accuracy of the parameters estimated than the calibration procedure followed before to this thesis.*

Regarding the selection of the data set to build the multivariate linear regression model.

- ✓ *Both actual data set (2015 and 2016) are incomplete to build the mathematical model and the missing parameters cannot be substitute by the estimated parameters, therefore the mathematical model will be built based on the simulated data set.*

  In all buildings there are evidences that the manipulation of the radiators are causing notable differences of the heat supplied. Therefore, in order to implement the mathematical model with actual data, the manipulations of the radiators should be measured and included in the mathematical model. Besides that, the parameters for IO and 3mE are not correctly estimated, therefore the estimated dynamic variables (internal heating demand) cannot be substituted.

8 Multivariate linear model based on building's thermal energy balance principle

This chapter describes the methodology followed to build the multivariate linear models of this study and presents the results of the first model proposed. The major challenge of building these models was to find the combination of parameters able to define a correlation which predict the heating demand for any building with high accuracy.

The constants and variables defining the model are based on the thermal energy balance principle of a building. The resulting model is defined by weather data and measurable temperatures. The coefficients of the variables correspond to the building and systems characteristics and they are obtained by training the model with historical data. In this way, the model obtained limits the demand of measured data to few measurable parameters, leaving out all parameters related with building & system characteristics (which are unknown in most of the cases).

Based on the findings from previous chapters, the number of equations of the model are defined and the data sets for each equation are selected. The design of several equations and use of the corresponding data set is key to increase the accuracy of the model.

The statistical validation of each model is done by applying a statistical search procedure. The search procedure validates the model at each step by analysing the residual of the data set, and quantifying the significance level of both the variables' coefficients and the total model. The following section explains in details the multivariate regression method and the search procedure followed.

8.1 Description statistical methodology

8.1.1 Multivariable regression

As it was explained in section 3.2, this study uses the regression approach to obtain a mathematical model able of predicting the heating demand. Since there are more than one independent variable related to the dependent variable, a multivariable regression model is created. The analytical form of the general linear model is expressed as follows.

$$Q_h = constant + \sum_{i=1}^{n} C_i \cdot X_i$$

Where Q_h is the hourly heating demand, C_i are the corresponding coefficient of the dependent variable, X is the dependent parameter and 'i' the index of dependent variables selected for the model.

In this case study, the independent parameters change hourly, therefore the general linear model is expressed in matrix term as follows:

$$Q = constant + X \times C$$

Where:

- Q is the matrix of the predictors (hourly predicted heating demand) with dimension nx1 (one column matrix)
- X is the matrix of the variables with dimension nxp. The variables are the hourly values of the inputs selected for the regression analysis such as solar radiation, outdoor temperature, indoor air temperature, heat gains, etc. Each variable is represented in one column of the matrix

- C is the matrix of the coefficients of the variables (unique for each model) with dimension px1 (one column matrix)
- 'n' is the number of hours predicted and 'p' the number of variables introduced in the model

Therefore, for an annual data set (8760 hours in one year), the multiple regression expressed in matrix form is represented as follows:

$$\begin{bmatrix} Q_1 \\ \vdots \\ Q_{8760} \end{bmatrix} = \begin{bmatrix} 1 & X_{1,1} & \cdots & X_{1,p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{8760,1} & \cdots & X_{8760,p} \end{bmatrix} \begin{bmatrix} C_0 \\ \vdots \\ C_p \end{bmatrix}$$

Where C_0 represents the constant of the regression analysis.

The previous matrix form is expressed by the following system of equations:

$$\begin{cases} Q_1 = C_0 + C_1 \cdot X_{1,1} + \cdots + C_p \cdot X_{1,p} \\ \cdots \\ Q_{8760} = C_0 + C_1 \cdot X_{8760,1} + \cdots + C_p \cdot X_{8760,p} \end{cases}$$

In order to solve this system of equations, the coefficients of the variables and the constant (matrix C) are estimated by training the model with historical data set (Q and X are known). The model is trained using regression analysis, which calculates the quantitative relation between the matrix Q (dependent variable) and matrix X (independent variables). This quantitative relation corresponds to the constant and variables' coefficients (matrix C). The values of the variables' coefficients indicates the contribution of the different variables to the predicted heating demand. It is impossible to find a matrix C that perfectly fits all equations. The search procedure (model training) aims at finding the values of C that fits the best all Q/X points.

The values of matrix C are influenced by the type and number of variables introduced in the model and the historical data set selected. Therefore, it is important to make a good selection of the parameters introduced in the model (predictor variables) and data set beforehand (section 8.3).

The most influencing parameters analysed in previous chapters are used as candidates for predictor variables. This means that every possible combination of variables needs to be tried to find the best fitting equation (or model). Therefore, several linear equations are built and compared for each selected data set. In order to build up models with statistical significance and compare them between each other, a search procedure (forward or backward search method) and statistical criteria for the predictor selection is used by applying the Matlab stepwise regression function available in the statistical toolbox in Matlab R2015b [45] , [46] (see Appendix 11). The following section explains the search procedure and statistical criteria used by this function to analyse the statistical significance of the multilinear model. Further details regarding search procedures and statistical criteria can be found in Sá 2007 [46].

8.1.2 Search procedure and statistical criteria used to analyse the statistical significance of the multilinear model

The stepwise regression function fits a regression model of the dependent variable (Q_h) depending on the independent variables (X_i) for performing multilinear models based on the statistical significance in a regression. This function introduces variables consecutively to the model (forward search method) based on their significance until a satisfactory model is found. At the beginning of the procedure, the residuals of the data set are analysed. In each step, the statistical significance of the model and each individual parameters' coefficients are analysed. Therefore, the analyses involved in this function can be classified in 3 groups: (1) analysis of the residuals of the data set, (2) analysis of the individual significant level of the variable coefficients estimated, (3) analysis of the significance level of the model. These analyses are done based on the different statistical criteria described below.

1. Analysis of the residuals of the data set

The statistical criteria used to analyse the residuals of the data set are the mean, variance and distribution. The mean of the residuals should be approximately zero for every value of X, the variance approximate constant for all values of X and the distribution should follow a Normal distribution.

2. Analysis of the individual significant level of the variables' coefficients estimated

The individual significant level of the variables' coefficients, estimated in each regression model, are analysed by using the p-value and t-statistics for the coefficient estimates. The p-value and the t-statistics are statistical criteria used to test the hypothesis and significance level of the independent variable.

The p-value indicates the significance level of the null hypothesis [46], this is the significant level of the coefficient calculated for the dependent variable. The significant level selected for the null hypothesis (or entrance tolerance value) is 5% (p-value<0.05), meaning that there is at least a 95% of confidence level that the coefficient presents the correct correlation between the independent parameter and the dependent parameter. Every time that a new parameter is inserted in the model, the coefficient of the parameters changes and the p-value and t-statistics are recalculated. The minimum p-value for a term to be removed (exit tolerance value) is 0.10. These tolerance values corresponds to the default values given by stepwise regression function [45]. Therefore, only the variables with p-value<0.05 will be introduced in the model while the variables with p-value>0.1 will be removed.

The t-statistics measures the significance of the predictors and is defined as the ratio of the estimated parameter from its notional value and its standard error [46]. Therefore, the variables with higher t-statistics have a higher contribution to the fit of the curve. As a result, the order of priority to insert the variables will be based on the variables with higher t-statistics within the tolerance values described above.

3. Analysis of the significance of the model

The analysis of the significance of the model is aim to avoid the addition of too many parameters which could lead to over-fitting the model. An over-fitted model describes random error instead of the underlying relationship, leading to a poor predictive performance, as it overreacts to minor fluctuations in the training data [47]. The criteria used to analyse the significance of the model are the coefficient of determination (R^2), the coefficient of determination adjusted (R^2 adjusted) and Root Mean Square Errors (RMSE).

R^2 indicates the proportion of the variance that is predictable from the dependent variable, therefore it indicates the goodness of fit of the model and it has to be maximised. Normally, R^2 increases by adding more variables to the model, however a higher fit could lead to an over-fitted model. Therefore, in order to avoid the overfitting, RMSE and adjusted R^2 are also analysed. The RMSE is the square root of the mean variance of residuals, therefore this is minimized [46]. The adjusted

R^2 is a modified version of R-squared that has been adjusted for the number of predictors in the model. It compares the explanatory power of regression models that contain different numbers of predictors, increasing when the new term improves the model more than would be expected by chance [46].

Figure 23 summarizes the search procedure and criterion used for a correct selection of predictor variables. Further details regarding the interactive stepwise function in Matlab can be found in Appendix 11.

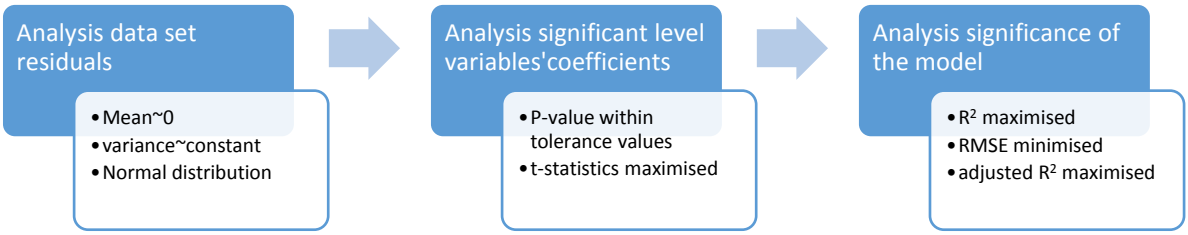


Figure 23 Flow scheme of the search procedure and statistical criteria used for predictor selection [own illustration].

8.2 Definition of the number of equations

Since the scope of IPIN is limited to the heating demand, the mathematical model presented in this study predicts the heating demand during a period of time that only heating is required. A different mathematical model will be needed to calculate the cooling demand since the most influential parameters for the cooling demand differ from the ones for the heating demand. Furthermore, the thermo-dynamics of the walls differ during the moderated months. As a result, in order to calculate the heating and cooling demand for all year round, it is expected that 3 different models will be needed (one for each type of season): (1) heating prediction during winter (or period when only heating is required), (2) cooling prediction during summer, (3) heating and/or cooling prediction during moderated months (when both cooling and heating are required during the same day).

Each operating mode of a building requires a different equation. The TU Delft buildings present 3 different operating mode: (1) weekdays during opening hours, (2) weekends during opening hours, (3) closing hours. During the different operating modes, the temperature set points and ventilation flow rate vary, affecting the correlation between the variables (equation coefficients), and therefore the model accuracy. In order to increase the accuracy of the model, a different equations to predict the heating demand is required for each different operating mode.

This study presents the corresponding equation for weekdays during opening hours (operating mode (1)). This is because the model of this operating mode is the most challenging to build as several parameters influence the heating demand. During closing hours, the prediction of the heating demand is dependent on less parameters (absence of solar and internal heat gains), increasing the correlation with the outdoor and indoor temperature (see Appendix 9, Figure 63 and Figure 68). The linear regression with only these 2 parameters leads to high prediction accuracy.

8.3 Data set selection

As concluded in chapter 7, the existing actual data sets (2015 and 2016) are incomplete and the missing parameters cannot be replaced by the estimated parameters since the calibrations were not performed accurately. Therefore, the simulated data set is selected to build the mathematical models presented in this chapter and chapter 9. It was proven in previous chapters that the simulated data set are accurate and acceptable to replace the actual data set.

The simulated data set selected corresponds to the period of time from 5th of October of 2015 until 14th January 2016. This period of time is chosen since only heating is required, avoiding the influence of cooling. The model presented in this report correspond to operating mode (1) which is built by selecting only the hours corresponding to weekdays during opening hours.

8.4 Model development and selection of parameters

There are numerous factors that influence the energy performance of a building, such as the building's thermal characteristics, architecture, the operation of systems, building use and outdoor weather conditions. The major challenge of this study was to find the combination of the most influencing variables which lead to a high accuracy of the heating demand prediction. Several models with different linear and polynomial combinations were built, but the accuracy reached was not higher than the 60%. Therefore, it was decided to create a combination of parameters based on a simplified thermal energy balance of a building and leaving out all parameters related with building & system characteristics (which are in most of the cases unknown).

Section 6.1.2 described the thermal energy balance of the physics-based simulator (LEA) and defines equation (1). This equation can be rewritten in function of the temperature differences [^{o}C], solar gains [W/m^2], internal heat gains [W/m^3 of indoor air volume], and wind speed [m/s] leading to the specific heating demand [W/m^3 of indoor air volume] described by equation (2).

$$Q_{demand}\left[\frac{w}{m^3}\right] = constant + C_1\left(T_{ground} - T_{indoor}\right) + C_2\left(T_{outdoor} - T_{indoor}\right) + C_3 V_{wind}\left(T_{outdoor} - T_{indoor}\right) + C_4\left(T_{out\ AHU} - T_{indoor}\right) + C_5\left(T^t_{indoor\ surfaces} - T_{indoor}\right) + C_6\, Q_{solar} + C_7\, Q_{internal} \quad \textbf{(2)}$$

Where, C_i are the constant coefficients corresponding to each heat flux. The constant coefficients depend mainly on the building characteristics in each building and the ventilation profiles for each operating mode. Therefore, the values C_1, C_2 and C_3 will be constant for the same building and will vary between buildings. C_3 will depend on the infiltration rate of the building and C_4 on the ventilation profile for each of the operating modes defined. C_6 and C_7 gives a correlation between the solar gains and internal heat gains, respectively. The coefficients of each of the defined parameters are obtained by using the matlab function 'stepwise' (available in statistical toolbox).

Table 7 gives the physical equivalence of each of the constant coefficients based on the thermal energy balance of LEA (equation (2). Since the equation above is a regression model (including a constant), these coefficient are just related to these physical parameters, but their values are not equal.

Table 7 Physical description of the constant correlation coefficients considered

Coefficients	Physical equivalence	Parameters definition
C_1	$C_1 \sim U_{floor}.\ A_{floor}$	A_{floor} : area ground floor (building footprint) U_{floor}: heat transfer coefficient of floor
C_2	$C_2 \sim \sum_i U^j_{envelop}.A^j_{envelop}$	j: for each façade/roof of orientation $U^j_{envelop}$ heat transfer coefficient of envelop $A^j_{envelop}$ the surface of the facade/roof
C_3	$C_3 \sim \dot{V}_{openings} \cdot \rho_{air} + V_{building} \cdot \frac{0.15}{5^3} \cdot \rho_{air}$	$V_{building}$: volume of the building $\dot{V}_{openings}$: volume flow rate (m^3/s) hourly dependent on the weekly and weekend schedule ρ_{air}: air density
C_4	$C_4 \sim m_{vent.}.C_{pair}$	C_{pair} : heating capacity of air (J/kg.K) $m_{vent.}$: mass flow rate of the ventilation air (kg/s)
C_5	$C_5 \sim \alpha_i.\ A_{indoor\ surfaces}$	$A_{indoor\ surfaces}$: total area of indoor surfaces in contact with the indoor air. α_i : indoor combined heat transfer coefficients for convection and radiation
C_6	Effect variation total horizontal solar radiation on the heating demand	
C_7	Effect variation total internal heat gains on the heating demand	

The significant value of the independent variable is evaluated by the p-value and the t-statistics (as explained in section 0) but not by the magnitude of the coefficient since the order of magnitude of the variables differ. The temperature differences vary around the range of magnitude between 15 and 30, the solar radiation between 0 and 900 [W/m^2] and the internal heat gains between 0 and 3.8 [W/m^3]. The symbols of the coefficients do not have a physical meaning since they depend on the constant and variables combination.

The only coefficient comparable between buildings is C_6 (solar gains). This is because the heating demand is in specific units [W/m^3 of indoor air volume] and the solar radiation is the same for the three buildings. The variables related to the temperature differences are not comparable between buildings since they depend on the indoor temperature set point which differs in each building. The internal heat gains are expressed in specific units, but they have different magnitude for each building.

8.5 Results and discussions

The model obtained shows an excellent fitting profile for the three buildings studied, 98.6% for IO, 99.6% for 3Me and 97.52% for TPM. Table 8 shows the coefficients and statistical parameters of the model.

Table 8 Coefficients and statistical parameters of the multivariate regression for the specific heating demand prediction (W/m^3) for IO, 3mE and TPM, respectively. Data set: weekdays during opening hours from 5th October 2015 until 14th January 2016.

	IO	3mE	TPM
Constant	-6.09	-52.65	61.14
C_1	-0.90	-10.50	11.65
C_2	**-0.76**	**0.45**	**-2.49**
C_3	-0.01	-0.01	-0.01
C_4	0	7.17	-7.93
C_5	**-2.50**	**-7.37**	**-7.33**
C_6	0.0012	0.0019	0.0009
C_7	-1.92	-0.77	-1.43
Adjusted R^2	98.60%	99.59%	97.50%
RMSE	0.48	0.52	0.66

This table shows the statistical parameters used to analyse the significance of the model. The other statistical parameters (used to analyse the data set residuals and significant level of the variables' coefficients) are within the limit values.

For the three buildings, the variables $(T_{indoor\ surfaces} - T_{indoor})$ and $(T_{outdoors} - T_{indoor})$ are the most influencing parameters on the hourly heating demand profile. Only including these two variables in the model, the three buildings reach a goodness fit between 96 – 99%. This shows that the most influencing parameters on the hourly heating demand of the building are the envelop characteristics and the thermal mass of the building.

Figure 24 illustrates the fitting profile for the multivariate regression model built in this chapter (model 1) for IO, 3mE and TPM, respectively. It is observed that the regression model follows the pattern of the heating demand predicted by the simulator with a high accuracy for the three buildings. It is observed that the regression model does not reach some maximum peaks. The difference is in all cases less than a 10% of the specific heating demand. Since this inaccuracy is localized and follows a constant pattern, it can be mathematically corrected if needed.

The high accuracy of this model make this equation able to be implemented in practice. However, the temperature of the indoor surfaces is a parameter which is not available in the real case study. When the variable $(T_{indoor\ surfaces} - T_{indoor})$ is neglected from the equation, the fitting drops to 79.3% for IO, 53% for 3mE and 63.5% for TPM (See model 1b in Appendix 12, section 12.1 for further details on the results obtained) due to the influence of the thermal mass on the heating demand. As a result, the indoor surface temperature should be replaced by a known variable in order to be applied in this case study. In the next chapter a new model independent from the indoor surface temperature is presented.

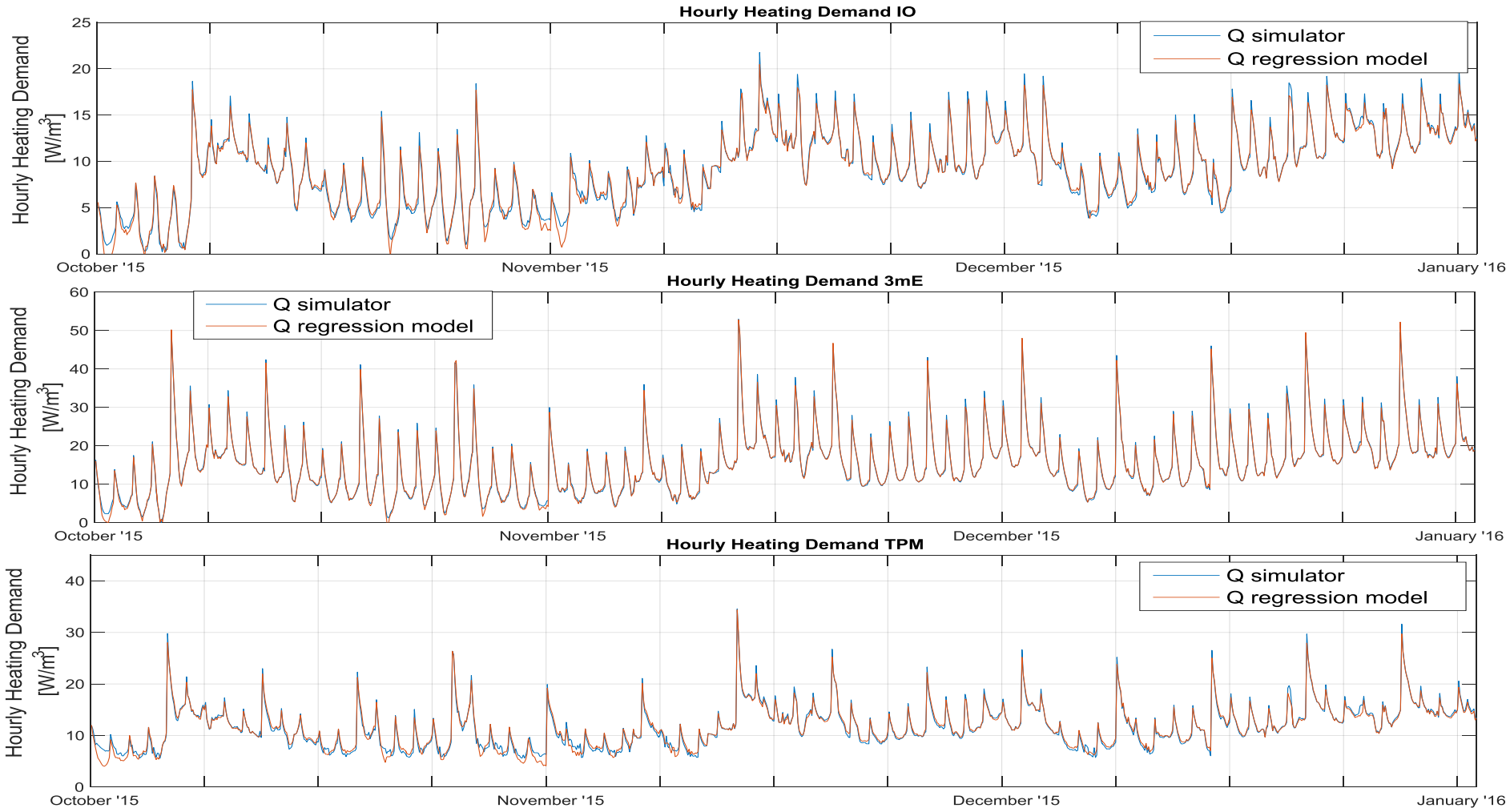


Figure 24 Fitting profile of the multivariate regression model for the specific heating demand prediction defined by equation (2) for IO (above), 3mE (middle) and TPM (below), respectively. Data set: weekdays during opening hours from 5th October 2015 until 14th January 2016.

9 Multivariate linear model improvement towards application into practice

Previous chapter developed a multivariate linear regression model (model 1) very accurate in function of the temperature differences, internal and solar gains. However, one of the most influencing parameters (indoor surface temperature) is an unknown parameter in this case study. Therefore, it is decided to go a step further and develop a model in which the indoor surface temperature could be replaced by existing variables, and therefore could be put into practice for the heating demand prediction of the buildings at TU delft campus.

This chapter builds a model independent on the indoor surface temperatures. With this purpose, a sequence of different models are analyse and developed. Each step gave place to a sub model (called model 2a, 2b and 2c) as follows.

1. Building a multivariate linear regression model with independent temperatures, instead of the temperature differences (model 1). Model 2a includes the indoor surface temperature.
2. The influence of the indoor surface temperature is analysed by neglecting this variable. This step gives place to model 2b.
3. The indoor surface temperature is replaced by known variables (model 2c).

9.1 Regression model with independent parameters (including and excluding the indoor surface temperature)

The model developed in this section includes the same variables than model 1, however the input parameters correspond to the independent temperatures (instead of the temperature differences). This model is described by equation (3).

$$Q_{demand}\left[\frac{w}{m^3}\right] = constant + C_a\,(T_{outdoor}) + C_b\,(T_{indoor}) + C_c(\mathrm{V}_{wind}) + C_d\,(T_{ground}) + C_e\,(T_{out\ AHU}) + C_f\,(T^t_{indoor\ surfaces}) + C_g\,Q_{solar} + C_h\,Q_{internal} \quad \textbf{(3)}$$

Table 9 gives the physical equivalences of the correlation coefficients of model 2. The equivalences of the coefficients of model 2 with respect model 1 are also indicated (see section 8.4 for further details).

Table 9 Physical equivalences of the correlation coefficients considered

Coefficients	Physical equivalence
C_a	$C_a \sim C_2 + C_3$
C_b	$C_b \sim -C_1 - C_2 - C_3 - C_4 - C_5$
C_c	$C_c \sim C_3 \sim \dot{V}_{openings} \cdot \rho_{air} + V_{building} \cdot \frac{0.15}{5^3} \cdot \rho_{air}$
C_d	$C_d \sim C_1 \sim U_{floor} . A_{floor}$
C_e	$C_e \sim C_4 \sim m_{vent.} . C_{pair}$
C_f	$C_f \sim C_5 \sim \alpha_i . A_{indoor\ surfaces}$
C_g	$C_g \sim C_6$
C_h	$C_h \sim C_7$

The outdoor temperature coefficient (C_a) depends on the envelop characteristics and infiltrations in the building. The indoor air temperature coefficient (C_b) depends on the coefficients of the ground temperature, air outflow temperature of the AHU, infiltrations, outdoor temperature and indoor surface temperature. Therefore, a variation in one of the dependent coefficients will influence C_a and C_b, and vice versa.

As it was explained in the previous chapter (section 8.4), the significant value of the independent variable is evaluated by the p-value and the t-statistics, but not by the magnitude of the coefficients since their order of magnitude differs between each other. The symbols of the coefficients are missing physical meaning as they depend of the constant and the combination of variables in the equation.

In this model, the independent coefficients (from C_c until C_h) can be compared between buildings since the heating demand is in specific units [W/m^3 of indoor air volume] and the variables are independent.

The fitting of this model is analysed including and excluding the indoor surface temperature for the three buildings during week days and opening hours (see section 9.3.1 and Appendix 12 for details on the results). Including the indoor surface temperature, the goodness of fit is very similar to model 1 (98.52% for IO, 99.58% for 3mE and 97.47%) for TPM. When the indoor surface temperature is excluded, the accuracy of the model for the three buildings is decreased (down to 80% for IO, 52.9% for 3mE and 63.52% for TPM). Therefore, it is key to replace the indoor surface temperature in order to improve the fitting for the three buildings. Next section gives an overview of the procedure followed to substitute the indoor surface temperature by other known variables.

9.2 Regression model with indoor surface temperature replaced

This section explains the procedure applied to build an accurate multivariate linear regression model independent on the indoor surface temperature. In the first place, the surface temperature profile and its dependency with other variables is analysed in order to replace it by other known parameters. Once the dependent parameters are studied, the accuracy of this relationship is measured by performing a regression analysis for the indoor surface temperature. Finally, these parameters are included in the final equation, leading to model 2c.

9.2.1 Analysis of the indoor surface temperature

Analysis of the dependent parameters

The indoor surface interchanges heat with the indoor air temperature and the thermal mass below (floor layers below) which also accumulates heat from internal and solar gains. Figure 25 shows the daily pattern of the indoor surface temperature, indoor air temperature, internal heat gains and solar gains for IO during a representative week for the data set 2015. The Y-axis is unit less since the magnitudes are normalized.

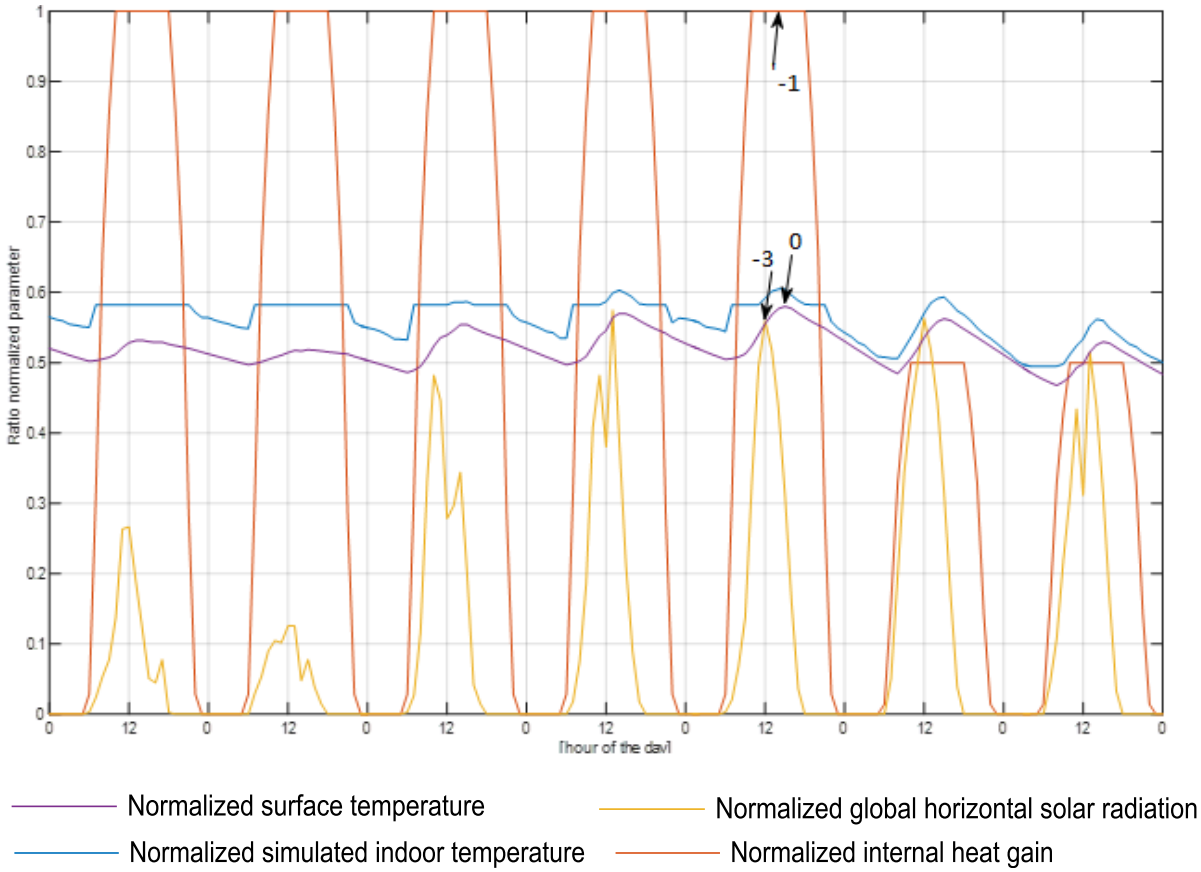


Figure 25 Normalized profiles of the indoor air temperature, internal heat gain, global horizontal solar radiation and surface temperature (modelled in LEA) for IO during a representative week (5th - 11th October '15)

The indoor surface temperature is in phase with the indoor air temperature. However, the internal and solar heat gains are delayed, indicating that the thermal mass takes time to accumulate the internal and solar gains. The indoor surface temperature is delayed 1 hour (1 time step) with respect the internal heat gain profile and 3 hours with respect the maximum global horizontal solar radiation. The delay of the solar radiation is higher than the internal heat gain because the solar radiation enters in the building after midday, when the sun has a certain angle with the vertical to be able to enter through the windows and reach the floor.

Figure 26 shows the daily course of the solar radiation for summer and winter. During winter, the global horizontal solar radiation (h) is delayed 3 hours with respect the global vertical solar radiation on the west façade (w). Since the heat accumulation time is 1 hour (1 time step), the solar radiation around 14:00 is the radiation that has its highest influence on the increase of the indoor surface temperature.

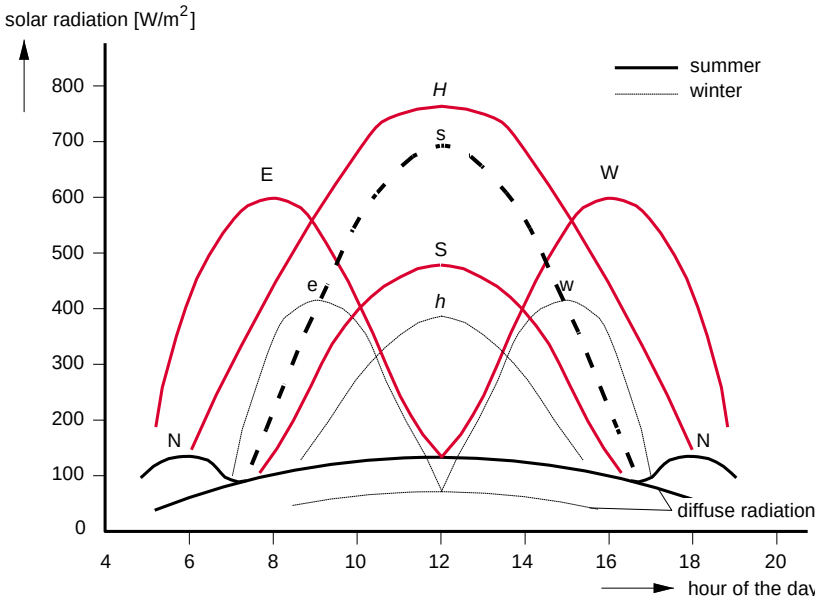


Figure 26 Daily course solar radiation at horizontal and vertical surfaces in summer and winter [48].

According to these observations, the indoor surface temperature is replaced by the internal heat gain 1 hour later ($Q_{internal,1a}$), the solar heat gain 3 hours later ($Q_{solar,3a}$) and the indoor air temperature in phase. In this way, the solar radiation profile on the inclination plane at 14:00 is simulated with the values of the global horizontal solar radiation. This assumption can be done during winter since the solar radiation on the different planes is similar (see solar radiation on the horizontal plane (h) and on the vertical planes (e,w)). However, during spring and summer time, it is advised to make an average between the horizontal plane radiation and the vertical plane radiation profiles (when the solar radiation on the corresponding plane is not available).

Analysis of the hourly profile of the indoor surface temperature

Figure 28 shows the hourly indoor surface temperature profile including all hours (weekdays and weekends) for the data set studied for IO, 3mE and TPM, respectively.

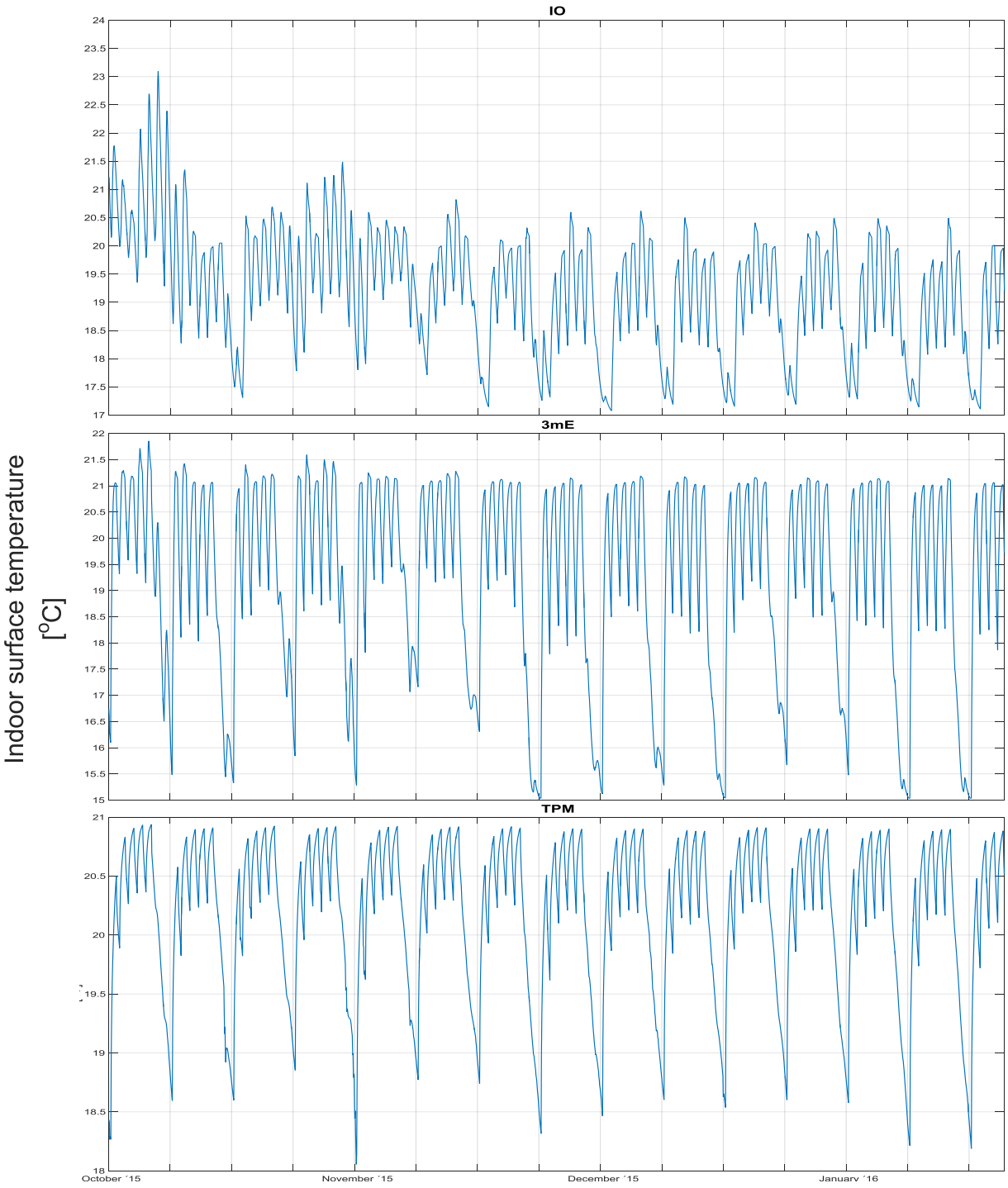


Figure 27 Simulated hourly indoor surface temperature profile. Data set: weekdays and weekends during opening and closing hours from 5th October 2015 until 14th January 2016.

During weekdays, the indoor surface temperature simulated by LEA shows a different pattern for Mondays than from Tuesday until Friday. Every Monday at the first hour of the morning, the indoor surface temperature is lower than the first hour of the rest of the weekdays. This is because during the weekend, the indoor air temperature set point is lower, the internal heat gains are smaller and the range of opening hours decreases. As a result, the indoor surface temperature drops slowly during the weekend, reaching its weekly minimum on Monday at 7:00. From Tuesday until Friday, the building is cooling down only during closing hours (9 hours), and as consequence the decrease of the indoor surface temperature is less marked than during the weekends. This phenomena leads to a higher increase of the indoor surface temperature on Mondays than during the rest of the weekdays. On Mondays, the indoor surface temperature increases by 2.5ºC for IO, 4.5ºC for 3mE and 2ºC for TPM. In contrast, from Tuesday until Friday, the indoor surface temperature is increased by 1.5ºC for IO, 2.5ºC for 3mE and 0.5ºC for TPM. As a consequence, the indoor surface temperature pattern on Mondays differs from the rest of the weekdays. This phenomena leads to an irregular data set (2 different patterns) which could result on a poor linear correlation.9.2.1

The weekly pattern of the indoor surface temperature for TPM is more stable than for IO and 3mE due to a higher thermal mass in TPM. This means that the indoor surface temperature in TPM is less dependent from other parameters which could reduce the success of the replacement by other variables.

For IO and TPM, the maximum temperature reached on Mondays is around 0.5ºC lower than the maximum temperature reached on the rest of the weekdays (which corresponds to the indoor air temperature set point). In contrast, for 3mE the maximum temperature reached is similar every weekday. Moreover, for all buildings the indoor air temperature set point is satisfied, meaning that the indoor air temperature is satisfied even when the indoor floor temperature is 0.5ºC lower. Therefore, the extra heat accumulated in the floor of 3mE (leading to 0.5ºC extra) should come from other heat source (internal or solar gains). If this is the case, in 3mE the indoor surface temperature will present a stronger correlation with the solar radiation and internal heat gains accumulated in the floor than in IO and TPM.

The accuracy of the replacement of the indoor surface by the variables mentioned above is quantified for the three buildings by performing a linear regression analysis for the indoor surface temperature. Next section presents the results of the regression model of the indoor surface temperature.

9.2.2 Results regression model of the indoor surface temperature

The regression model of the indoor surface temperature is performed during opening hours for IO, TPM and 3mE. The best fitting is found for 3mE (63%), followed by IO (52.44%) and TPM (23%). As it was mentioned above, 3mE shows the most accurate correlation due to a stronger influence of the internal heat gains and solar gains than IO and TPM (as the ratio windows to floor area and internal gains are the highest). In contrast, TPM presents the lowest correlation due to a high thermal mass and the low influence of the solar and internal gains. IO has the same thermal mass than 3mE, but lower heat and solar gains, decreasing the explanation power of these two parameters.

Equation (4) corresponds to the multivariate regression model obtained for the indoor surface temperature for IO during weekdays and opening hours. Figure 28 illustrates the surface temperature modelled for IO by the regression model (equation (4)) and LEA, respectively.

$$T_{indoor\ surface} = -26.3 + 0.04\ (Q_{internal,1a}) + 2.25\ (T_{indoor}) + 0.004(Q_{\text{solar,3a}}) \quad \textbf{(4)}$$

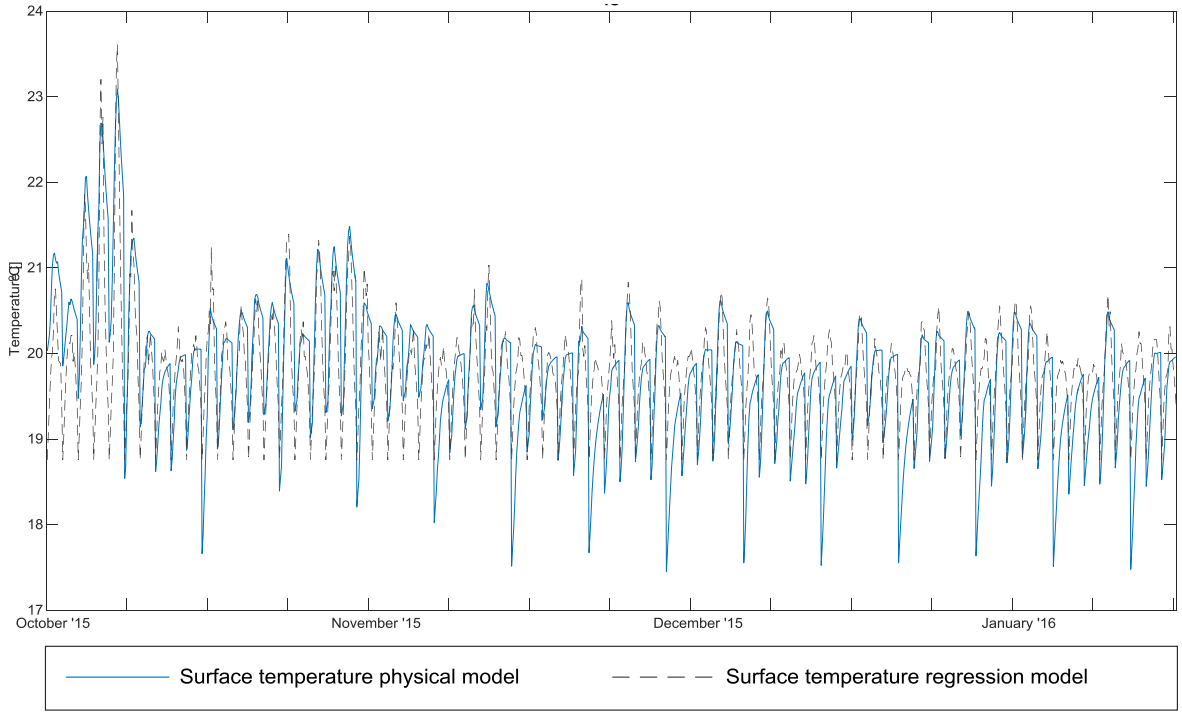


Figure 28 comparison surface temperature regression model versus the surface temperature physical model. Data set: weekdays during opening hours from 5th October 2015 until 14th January 2016.

As explained in previous section, the two different patterns shown by the indoor surface temperature (Mondays and Tuesday-Friday) can lead to inaccuracies in the regression model. This effect is appreciable in the estimation of all minimums and the maximum on Mondays. A lower minimum on the first hour of Monday lowers the least square of the linear correlation, as a result the higher minimums (from October until mid-November from Tuesday until Friday) are underestimated. An overestimation of the maximums occurs especially on Mondays since the regression is higher due to the other weekdays. This effect is visualised in the three buildings. The accuracy of the model can be improved by separating the data sets in 2 and building a regression model for each data set: (1) for Mondays during opening hours and (2) from Tuesdays until Fridays during opening hours.

9.2.3 Regression model with surface temperature replaced

The indoor surface temperature is replaced by the three above mentioned variables and is expressed by equation (5), leading to model 2c.

$$Q_{demand}\left[\frac{w}{m^3}\right] = constant + C_a\,(T_{outdoor}) + C_b\,(T_{indoor}) + C_c(\mathrm{V}_{wind}) + C_d\,(T_{ground}) + C_e\,(T_{out\ AHU}) + C_{f,int1a}\,(Q_{internal,1a}) + C_{f,solar3a}\,(Q_{solar,3a}) + C_g\,(Q_{solar}) + C_h\,(Q_{internal}) \quad \textbf{(5)}$$

This replacement improves substantially the accuracy of the model. The fitting rises by 30% for 3mE, 11.3% for IO and 9.48% for TPM. Section 0 shows and discusses the comparison between the three models built in this chapter for IO (for further details on the fitting profile and results obtained for 3mE and TPM, see comparison model 2b and 2c in Appendix 12).

9.2.4 Physical meaning of the coefficients

The influence of the building parameters on the variables' coefficients is analysed by means of a sensitivity analysis by varying the most influencing building characteristics. The thermal mass, insulation level and windows fraction is increased, leading to an equation with different coefficients. The new coefficients are compared with the coefficients of the baseline scenario. In the next chapter, the results of this analysis are presented and discussed.

9.3 Results and discussions

This section presents the result and discussions of model 2 developed in this chapter described by equation (6). The first subsection compares the results of the three sub models developed (model 2a, 2b and 2c). The second subsection compares the final model (model 2c) for the three buildings. The last subsection analysis the influence of different building parameters on the coefficients of the variables.

$$Q_{demand}\left[\frac{w}{m^3}\right] = constant + C_a\,(T_{outdoor}) + C_b\,(T_{indoor}) + C_c(\mathrm{V}_{wind}) + C_d\,(T_{ground}) + C_e\,(T_{out\ AHU}) + C_f\,(T^t_{indoor\ surfaces}) + C_{f,int1a}\,(Q_{internal,1a}) + C_{f,solar3a}\,(Q_{solar,3a}) + C_g\,(Q_{solar}) + C_h\,(Q_{internal}) \quad \textbf{(6)}$$

9.3.1 Comparison model 2 with indoor surface temperature included, excluded and replaced.

Table 10 presents the corresponding coefficients and statistical parameters for model 2a, 2b and 2c for IO during weekdays and opening hours. Figure 29 illustrates the fitting profile of the hourly heating demand predicted for IO by the regression models 2a, 2b and 2c, respectively. Appendix 12 shows the results of the equations and the fitting profile for the three sub models for 3mE and TPM.

Table 10 Coefficients and statistical parameters of the multivariate regression model 2 for the specific heating demand prediction (W/m^3) for IO with indoor surface temperature included (model 2a), indoor surface temperature excluded (model 2b) and indoor surface temperature replaced (model 2c). Data set: weekdays during opening hours from 5th October 2015 until 14th January 2016.

Coefficients	Ts included (model 2a)	Ts excluded (model 2b)	Ts replaced (model 2c)
Constant	-17.69	147.63	32.34
C_a	-0.82	-0.97	-0.92
C_b	4.23	-6.37	-0.65
C_c	0.14	0.19	0.16
C_d	0	0	0
C_e	0	0	0
C_f	-2.50	-	-
$C_{f,int1a}$	-	-	-11.42
$C_{f,solar3a}$	-	-	-0.009481
C_g	0.0013505	-0.000059	-0.000676
C_h	-1.93	-4.62	8.44
Adjusted R^2	98.52%	79.42%	90.70%
RMSE	0.48	1.82	1.22

T_s: indoor surface temperature

- Coefficient corresponding to a variable excluded/neglected in the multivariate regression model

This table shows the statistical parameters used to analyse the significance of the model. The other statistical parameters (used to analyse the data set residuals and significant level of the variables' coefficients) are within the limit values.

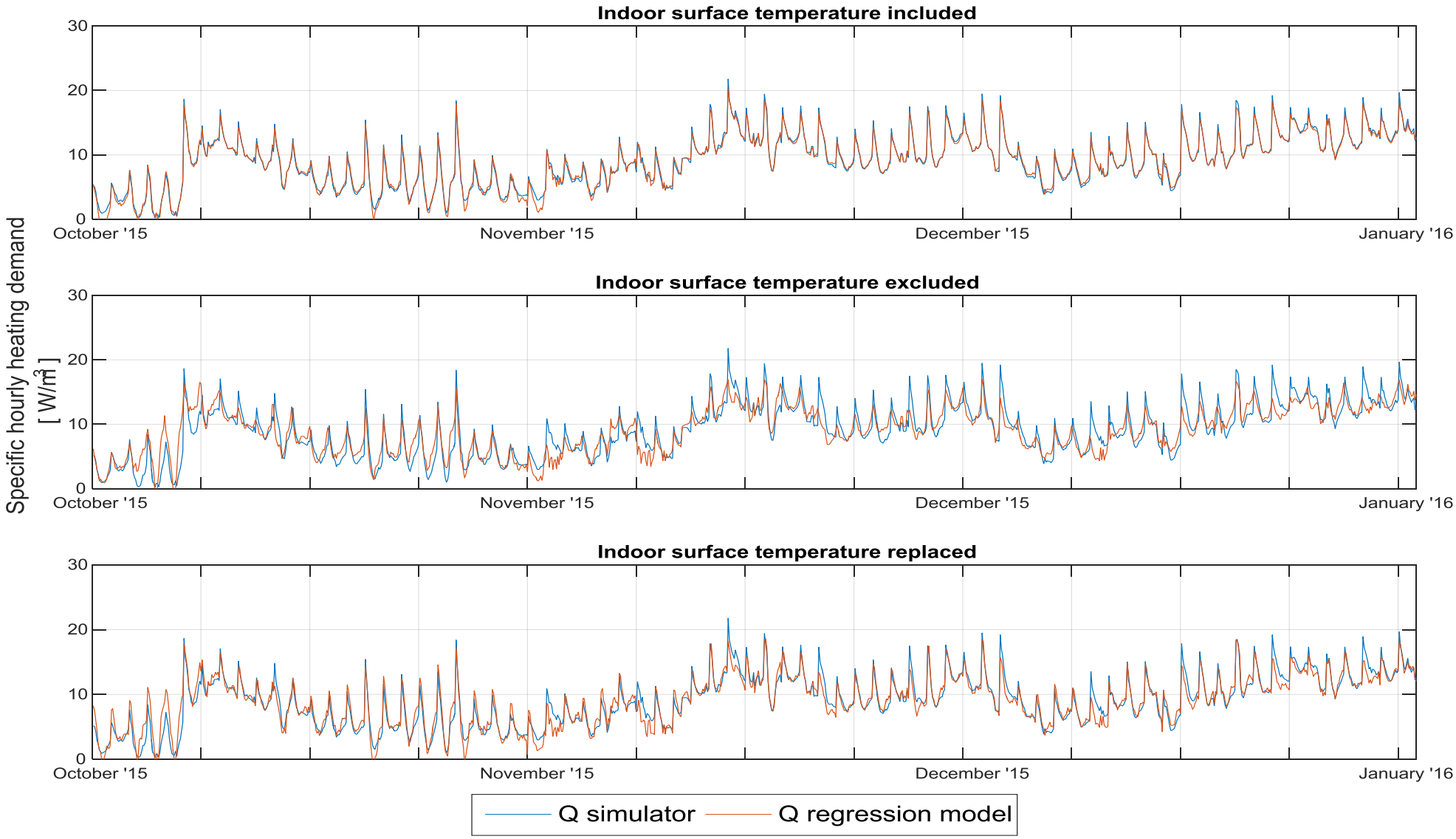


Figure 29 Fitting profile of the multivariate regression model for the specific heating demand prediction defined by model 2 for IO with indoor surface temperature included (above), excluded (middle) and replaced (below), respectively. Data set: weekdays during opening hours from 5th October 2015 until 14th January 2016.

Table 10 shows that the coefficients of the ground temperature and the outflow temperature of the AHU is zero ($C_d = C_e = 0$). This is because the ground temperature has a constant value in the model and the out coming temperature of the AHU has small variations, therefore they are static (or semi-static) parameters and their contribution to the model is zero.

As mentioned earlier, the magnitude and symbol of the coefficients are dependent on the combination of variables. This is appreciable by comparing the coefficients of the three models above (model 2a, 2b and 2c). For example, when the indoor surface temperature is excluded (model 2b), the coefficient of the indoor air temperature (C_b) and the constant changes symbol with respect model 2a. This is because C_b is influenced by the indoor surface temperature coefficient (C_f), interfering in the correlation and changing the symbol of C_b. As the indoor surface coefficient changes, the outdoor temperature coefficient also changes. As a result of this changes, the constant value and the internal heat gain coefficient are also influenced.

Table 10 and Figure 29 show that the model with the indoor surface temperature included (model 2a) presents the best fit for the three buildings (98.52%for IO, 99.58% for 3mE and 97.47% for TPM). The profile of the hourly specific heating demand shows the same accurately pattern than in model 1. As in model 1, model 2a presents some maximum peaks that are not reached by a 10%, leading to a minor underestimation of the heating demand during the first hours in the morning. The pattern of this error is constant and localized, therefore it can be corrected mathematically if needed.

The most influencing parameters in the model for the three buildings are the outdoor temperature and the indoor surface temperature. By introducing only these two parameters, the fitting profile shows an accuracy of 96% for TPM, 97% for IO and 99% for 3mE. Therefore, when the indoor surface temperature is available, the model can be simplified to these two variables.

When the indoor surface temperature is excluded, the fitting profile for the three buildings decreases substantially (79.42% for IO, 52.9% for 3mE and 63.52% for TPM), therefore a higher number of variables are added to rise the fitting up. The mismatch between the patterns followed by the heating demand predicted by the regression model and the simulator are mainly located in the minimum and maximum points for IO. However for 3mE and TPM, the error pattern is irregular which make it difficult to correct mathematically. These differences are between a 30% and 90% which are inacceptable, making the implementation of model 2b impossible.

The replacement of the indoor surface temperature leads to great improvements on the heating demand profile in comparison with model 2b. This is especially notable in 3mE where the fitting rises by more than a 30%. In IO the fitting increases by 11.3% and in TPM by 10%. The replacement is more significant in 3mE than in IO and TPM as the internal heat gains and solar radiation (higher ratio windows to floor area) replacing the indoor surface temperature $(Q_{int,1a}, Q_{solar,3a})$ has a higher influence on the internal surface temperature. The final model with the indoor surface temperature replaced shows a fitting profile of 83.2% for 3mE, 73% for TPM and 90.7% for IO. Next section compares this final equation and the fitting profile for the three buildings.

9.3.2 Comparison model with indoor surface temperature replaced for the three buildings

This section presents the results of the model with the indoor surface temperature replaced (model 2c), and therefore the model that can be implemented to predict and control the heating demand at the buildings of TU Delft. Figure 30 presents the fitting profile of the regression model for the specific heating demand prediction for IO, 3mE and TPM, respectively. Table 11 shows the coefficient of the variables and the statistical parameters of this model for the three buildings.

Table 11 Coefficients and statistical parameters of the multivariate regression model with indoor surface temperature replaced (model 2c) for the specific heating demand prediction (W/m^3) for IO, 3mE and TPM, respectively. Data set: weekdays during opening hours from 5th October 2015 until 14th January 2016.

Coefficients	IO	3mE	TPM
Constant	32.34	827.98	918.42
C_a	-0.92	-9.08	-8.57
C_b	-0.65	-5.25	-9.73
C_c	0.16	-0.02	0.02
C_d	0	0	0
C_e	0	-34.08	-34.55
C_f	-	-	-
$C_{f,int1a}$	-11.42	-7.09	-6.86
$C_{f,solar3a}$	-0.009481	-0.02	0.000086
C_g	-0.000676	0.0062	0.006
C_h	8.44	5.60	3.49
Adjusted R^2	90.70%	83.2%	73%
RMSE	1.22	3.36	2.16

- Coefficient corresponding to a variable excluded/neglected in the multivariate regression model

This table shows the statistical parameters used to analyse the significance of the model. The other statistical parameters (used to analyse the data set residuals and significant level of the variables' coefficients) are within the limit values.

The final model shows the best fit for IO (90.7%), followed by 3mE (83.2%) and TPM (73%). This is because in IO, the fluctuations of the heating demand profile follows a more constant pattern than in 3mE and TPM. However, in the three buildings two different patterns of maximum peaks can be differentiated. The smaller peaks correspond to the first hours of the morning from Tuesday until Friday, while the bigger peaks correspond to the first hours of the Mondays. Therefore, these peaks are related to the different indoor surface temperature patterns explained in section 9.2.1. On Mondays morning, the heating demand increases with respect the other weekdays because the indoor surface temperature has cold down further during the weekends. Therefore, the accuracy of the regression model is decreased due to an interference of two different type of data sets. As a result, the peaks on Mondays morning are underestimated and the peaks from Tuesday until Friday are over estimated.

When a higher accuracy is needed, the fitting can be easily improved by making 2 different regression equations: (1) including only Mondays and (2) including all weekdays from Tuesday until Friday. Therefore, the data set need to be divided accordingly, leading to 2 new data sets with a more uniform pattern. As a result, the accuracy of the regression model for 3mE and TPM will be increased.

Even though both 3mE and TPM are equally influenced by the interference of two different patterns, the fitting profile for 3mE is still higher than for TPM. This is due to a higher influence of the internal heat gains and solar gains in 3mE (as explained in previous section), which is also reflected in the coefficients of the corresponding variables ($C_{f,int1a}$, $C_{f,solar3a}$, C_g and C_h) as they are higher in 3mE than in TPM. As a result, these variables has more explanation power of the heating demand profile in 3mE than in TPM, leading to a better fit.

The coefficient corresponding to the air outflow temperature of the AHU (C_e) is zero for IO, however for 3mE and TPM the weight of the coefficient is increased by 6 when compared with model 2a (including indoor temperature surface). This indicates that the weight of this variable is increased when other variables of the model are neglected. Therefore, if the accuracy of the model for 3mE and TPM is increased by dividing the current data set in two data sets, the model may be simplified (eg. ($\boldsymbol{T_{out\ AHU}}$) may be neglected).

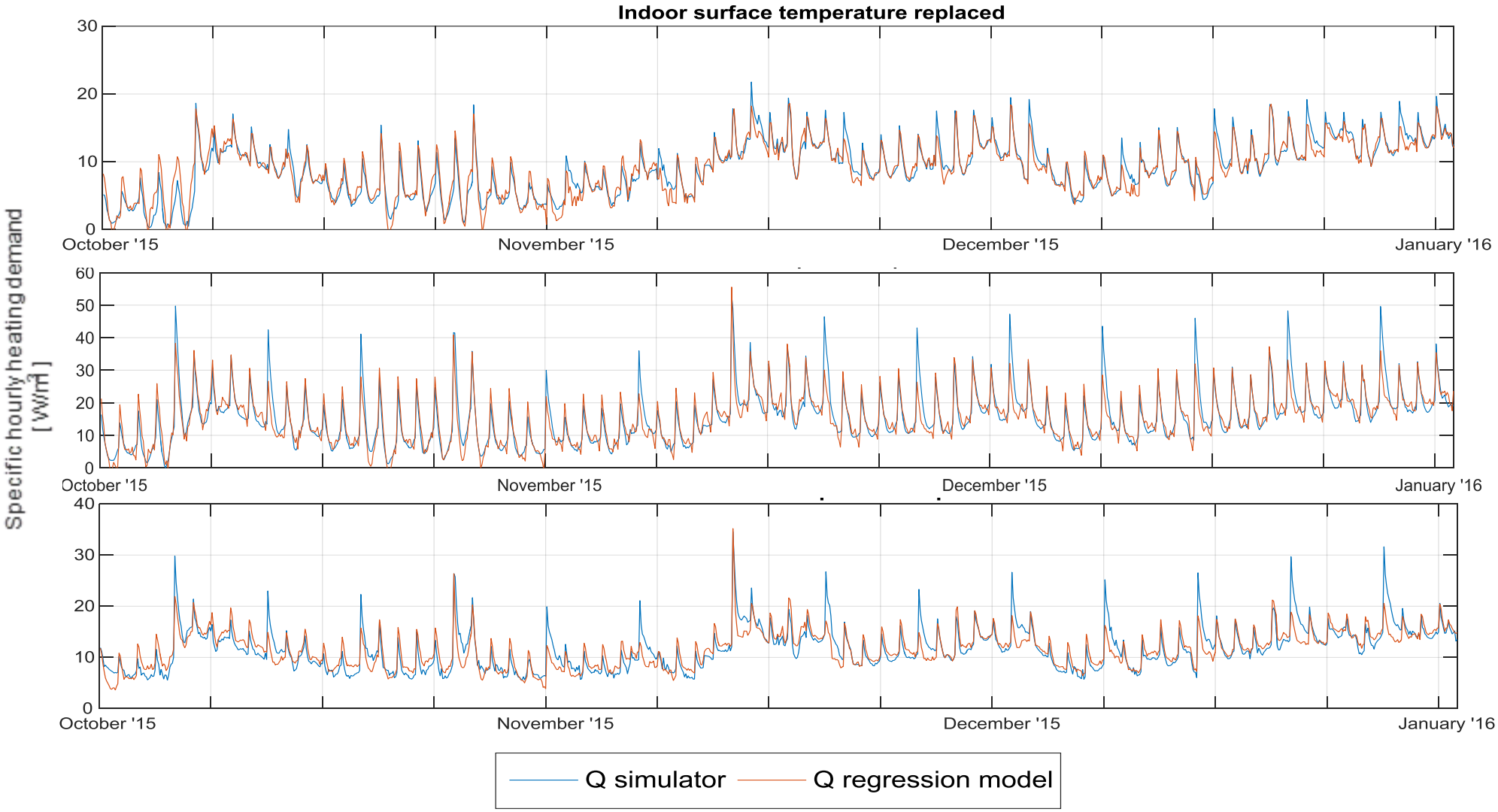


Figure 30 Fitting profile of the multivariate regression model for the specific heating demand prediction defined by model 2c (indoor surface temperature replaced) for IO (above), 3mE (middle) and TPM (below), respectively. Data set: weekdays during opening hours from 5th October 2015 until 14th January 2016. Note: The simulation and regression model of IO is 24 hours forwards with respect TPM and 3mE (see maximum peak on the second week of November '15). This due to a mistake in the starting day of the year in LEA. This mistake is not corrected since it does not affect to the regression model as it is built with the simulated data set and all the variables corresponds to the correct hours of the day

9.3.3 Physical meaning of the coefficients

This section presents the results of the sensitivity analysis performed to analyse the influence of building parameters on the variables'. Table 12 shows the coefficients of the multivariate regression model with the indoor surface temperature replaced (model 2c) for 3 different scenarios. Scenario 1 corresponds to the baseline of IO (original building characteristics values). In scenario 2 and 3 the mass of the building and insulation level in walls and windows is duplicated, respectively.

Table 12 Multivariate regression models for the specific heating demand (W/m^3) for IO for the scenario 1 (baseline), scenario 2 (double specific mass) and scenario 3 (double insulation level in windows and walls); and their influences, respectively. Data set: weekdays during opening hours from 5th October 2015 until 14th January 2016.

	Scenario 1 (baseline A)	Scenario 2 (x2 mass)	Influence scenario 2	Scenario 3 (x2 Rc)	Influence scenario 3
constant	37.14	137.49	270.24%	32.19	-13.33%
C_a	-0.92	-0.96	3.72%	**-0.75**	**-18.62%**
C_b	-0.89	**-5.90**	**566.64%**	-0.87	-1.48%
C_c	0.16	0.19	14.23%	0.16	-0.94%
C_d	0	0	0	0	0
C_e	0	0	0	0	0
C_f	-	-	-	-	-
$C_{f,int1a}$	-11.36	-6.59	-42.03%	-8.41	-25.94%
$C_{f,solar3a}$	-0.012	-0.013	-24.45%	-0.01	-17.81%
C_g	-	-	-	-	-
C_h	8.26	4.52	-45.30%	6.09	-26.29%
R^2	**90.7%**	**88.66%**		**83.65%**	

- Coefficient corresponding to a variable excluded/neglected in the multivariate regression model

This table shows the statistical parameters used to analyse the significance of the model. The other statistical parameters (used to analyse the data set residuals and significant level of the variables' coefficients) are within the limit values.

An increase in the specific mass of the building is detected in the coefficients of the variables which replaces the indoor surface temperature. These are: indoor air temperature, internal heat gain and solar gains accumulated in the floor ($Q_{int,1a}$ and $Q_{solar,3a}$). Since the building has higher mass, the effect of the internal heat gain and solar gains accumulated decreases ($C_{f,int1a}$ and $C_{f,solar3a}$ decrease). Therefore, the indoor air temperature becomes the most significant variable since it acts as the main substitutive of the indoor surface temperature. As a result of a variation in the coefficients of these variables, the constant value increases and the other coefficients are also affected. As a conclusion, an increase in the thermal mass will be notable by an strong increase of C_b and a decrease of $C_{f,int1a}$ and $C_{f,solar3a}$ (and vice versa when the thermal mass decreases).

The increase of the insulation is directly detected in the outdoor temperature coefficient (C_a). The coefficient decreases (slope decreases) as a result of a lower influence of the outdoor temperature. A variation in the coefficient of the outdoor temperature influences the indoor air temperature coefficient, and therefore the coefficients of the variables influencing the indoor surface temperature are also affected (internal heat gain and solar heat gains). As a consequence of all variables variations, the constant changes.

From this analysis, it is concluded that the variation of a coefficient influences several variables coefficients. However, the indoor air temperature coefficient is mainly influenced when the mass is varied, while the outdoor temperature coefficient varies when the insulation value changes. As a result, the indoor and outdoor temperature coefficients become the key indicators of a variation in the mass and insulation level, respectively.

The influence of the solar radiation and the orientation of the windows is evaluated by creating 2 new scenarios. Table 13 illustrates the coefficients of the multivariate regression model with indoor surface temperature replaced for scenarios 4 and 5. In scenario 5, the windows percentage is increased by 3 for the West façade. Since the windows have a lower insulation value than the walls,

a new baseline is created (scenario 4) with the same insulation value for both windows and walls (U-value= 0.4). In this way, the increase of the windows fraction will not interfere the insulation parameters.

Table 13 Multivariate regression models for the absolute heating demand (W/m3) for IO for the scenario 4 (baseline B) and scenario 5 (fraction of windows increased by 3 in West facade); and their influences, respectively. Data set: weekdays during opening hours from 5th October 2015 until 14th January 2016.

	Scenario 4 (baseline B)	Scenario 5 (windows x3 West façade)	Influence Scenario 5
constant	28.99	28.42	-1.99%
C_a	-0.68	-0.68	-0.18%
C_b	-0.78	-0.76	-2.90%
C_c	0.16	0.17	2.63%
C_d	0	0	0
C_e	0	0	0
C_f	-	-	-
$C_{f,int1a}$	-7.93	-7.88	-0.56%
$C_{f,solar3a}$	-0.0101	**-0.0106**	**4.73%**
C_g	0.00427	0.00426	-0.26%
C_h	6.19	6.01	-2.83%
R^2	**81.66%**	**82%**	

- Coefficient corresponding to a variable excluded/neglected in the multivariate regression model

This table shows the statistical parameters used to analyse the significance of the model. The other statistical parameters (used to analyse the data set residuals and significant level of the variables' coefficients) are within the limit values.

The increase of the windows fraction in the West façade is directly reflected by an increase of the coefficient that simulates the solar radiation entering through the West façade (simulated as horizontal solar radiation 2 hours forwards), $C_{f,solar3a}$. Since the influence of the variable on the model increases, the fitting of the total model also increases.

10 Validation model predictive potential

All models built in the present study are based on a data set containing 1096 hours which corresponds to the total opening hours during weekdays from 5th of October 2015 until 14th of January 2016. This chapter validates the predictive potential of model 2c (with indoor surface temperature replaced) on a short period of time.

In order to obtain the minimum size of the population (number of hours) needed, several new regression models are built with different population sizes. The coefficients of the parameters for the new models are compared with the original model (population size = 1095 data).

The results show that for the three buildings and the period of the year studied, a minimum of 825 opening hours (corresponding to 2.5 months) is needed to predict the next 270 opening hours (approx. 1month). As a result, the data collected from the previous season are able to predict with high accuracy level the next month. This can be very convenient in cases where the data from previous years are not available.

Table 14 compares the coefficients of the parameters corresponding to model 2c for the three buildings with a population size of 1095 data and 824, respectively. The coefficients of the parameters for the models with both population sizes are very similar.

Table 14 Comparison coefficients and statistical parameters of the multivariate regression model 2c for the three buildings for a population of 1095 and 824 data, respectively. Data set 1 (1095 data): weekdays during opening hours from 5th October 2015 until 14th January 2016. Data set 2 (824 data): weekdays during opening hours from 5th October 2015 until 19th December 2015.

Coefficients	IO (1095 data)	IO (824 data)	3mE (1095 data)	3mE (824 data)	TPM (1095 data)	TPM (824 data)
Constant	32.34	40.42	827.98	885.94	918.42	991.68
C_a	-0.92	-0.90	-9.08	-9.55	-8.57	-9.17
C_b	-0.65	-1.08	-5.25	-5.97	-9.73	-10.54
C_c	0.16	0.19	-0.02	-0.05	0.02	0.02
C_d	0	0	0	0	0	0
C_e	0	0	-34.08	-36.24	-34.55	-37.39
C_f	-	-	-	-	-	-
$C_{f,int1a}$	-11.42	-10.96	-7.09	-6.73	-6.86	-6.40
$C_{f,solar3a}$	-0.009481	-0.0088	-0.02	-0.0190	0.000086	0.0005
C_g	-0.000676	-0.0004	0.0062	0.0063	0.006	0.0060
C_h	8.44	8.06	5.60	5.30	3.49	3.08
Adjusted R^2	**90.70%**	**89.74%**	**83.2%**	**83.54%**	**73%**	**73.55%**
RMSE	**1.22**	**1.24**	**3.36**	**3.28**	**2.16**	**2.07**

- Coefficient corresponding to a variable excluded/neglected in the multivariate regression model

This table shows the statistical parameters used to analyse the significance of the model. The other statistical parameters (used to analyse the data set residuals and significant level of the variables' coefficients) are within the limit values.

When the population size is smaller than 825 data, the coefficients of the model differ significantly from the original model. In some cases, these models show a better fitting profile for the historical data, but the mismatch with the predicted data increases. The goodness of the fit expressed in Table 14 (Adjusted R^2) corresponds to the historical data used in each case. The fitting profile for both historical and predicted data is illustrated in Figure 31. It is observed that the fitting profile for the predicted data is almost identical to the match observed in section 9.3.2 for the total population size.

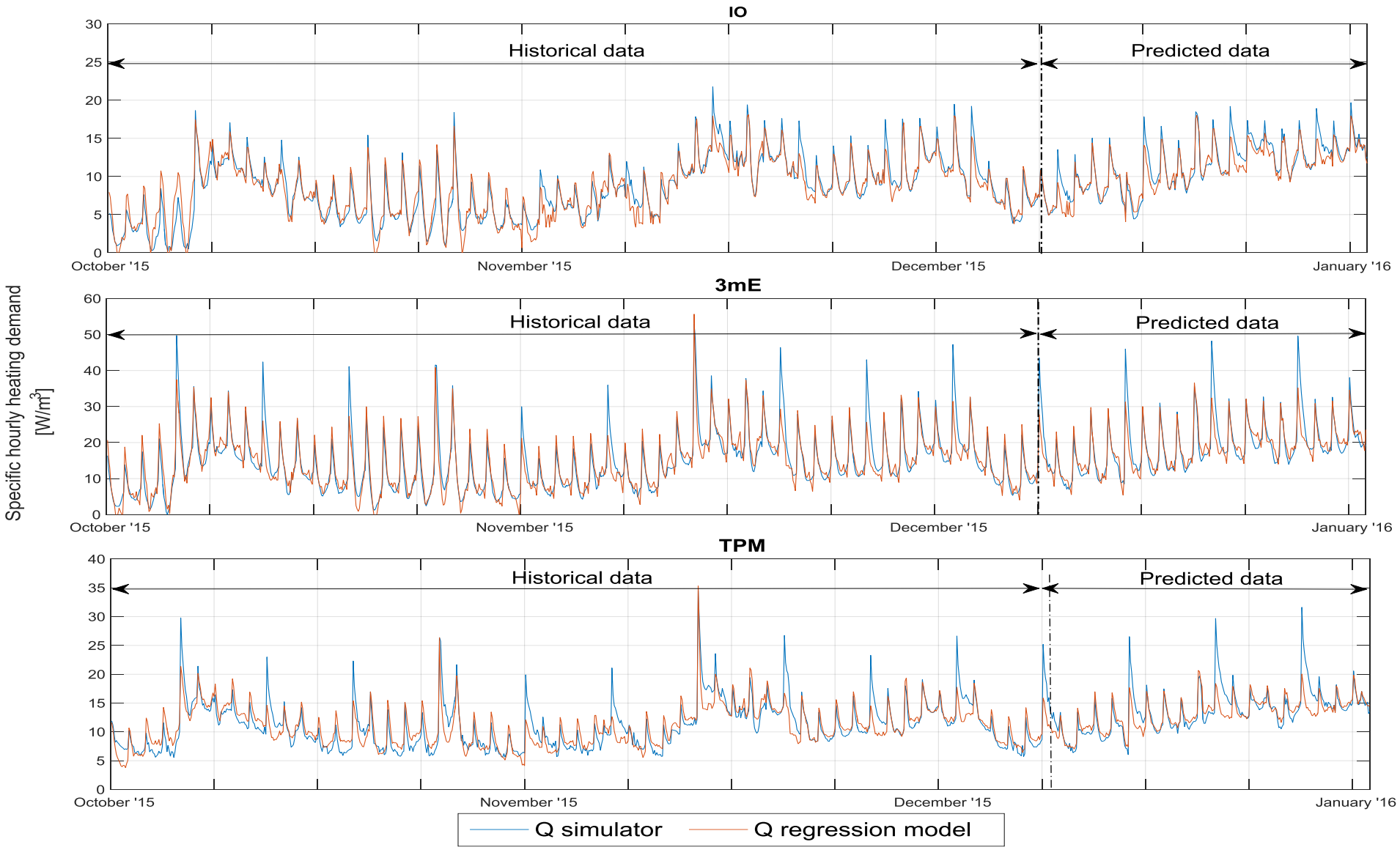


Figure 31 Fitting profile of the multivariate regression model 2c based on the data set October-December 2015 for IO, 3mE and TPM, respectively.

11 Final Conclusions

From this research work, it is concluded that the proposed linear regression model shows a promising performance in order to become a fast and simple to use tool for predicting the heating demand in buildings. This model limits the demand of measured data to a few measurable parameters and weather data which make it suitable to be used in older buildings where building characteristics are unknown. Moreover, it shows a physical relation with the input coefficients, making it possible to get a better insight of the influence of the different parameters in the building response.

This chapter presents the main conclusions regarding the multivariate linear model that was developed and the highlights related to the building procedure. Moreover, it gives an insight on the causes of the mismatch between actual and simulated data, and explains the consequences that this mismatch had on the calibrations of the estimated parameters. The evaluation of the calibrations is performed by using a new calibration method which is recommended for further validations.

11.1 Multivariate linear model

In this study, several models are built and compared with each other. The general equation of the final model developed in this study is expressed as follows:

$$Q_{demand}\left[\frac{w}{m^3}\right] = constant + C_a\,(T_{outdoor}) + C_b\,(T_{indoor}) + C_c(\mathrm{V}_{wind}) + C_e\,(T_{out\ AHU}) + C_f\left(T^t_{indoor\ surfaces}\right) + C_g\,(Q_{solar}) + C_h\,(Q_{internal})$$

Depending on the parameters available in the case study, more or less variables will be needed to reach an acceptable correlation. Two main models (for weekdays during opening hours) are differentiated depending on whether the indoor surface temperature is an available parameter or not:

1. Case study when the indoor surface temperature is available

In this case, the model can be expressed in function of two parameters: (1) indoor surface temperature and (2) the outdoor temperature and is expressed according to the equation below. The fitting profile for this model shows an accuracy of 96% for TPM, 97% for IO and 99% for 3mE.

$$Q_{demand}\left[\frac{w}{m^3}\right] = constant + C_a\,(T_{outdoor}) + C_f\left(T^t_{indoor\ surfaces}\right)$$

2. Case study when the indoor surface temperature is not available

In this case the indoor surface temperature should be replaced by additional terms because when the indoor surface temperature is excluded from the general equation, the fitting profile for the three buildings decreases to inacceptable levels (79.42% for IO, 52.9% for 3mE and 63.52% for TPM).

The indoor surface temperature is replaced by the internal heat gains and solar radiation in phase with the indoor air temperature (which is proven to be in phase with the indoor surface temperature). In this case study, the internal heat gain is delayed by 1 hour. When using the global horizontal solar radiation, this parameter is delayed by 3 hours (as it simulates the solar radiation profile on the inclination plane at 14PM). The delay time will vary on every case study and will depend on the thermal mass of the building (lower thermal mass, less delay time). This replacement will be more significant in buildings where the internal heat gains and solar radiation have a higher influence.

The final model in which the indoor surface temperature replaced is expressed by the equation below and shows a fitting profile of 83.2% for 3mE, 73.5% for TPM and 90.7% for IO. For 3mE and TPM, the fitting can be improved by making 2 different regression equations: (1) including only Mondays and (2) including all weekdays from Tuesday until Friday.

$$Q_{demand}\left[\frac{w}{m^3}\right] = constant + C_a\,(T_{outdoor}) + C_b\,(T_{indoor}) + C_c(\mathrm{V}_{wind}) + C_e\,(T_{out\ AHU}) + C_{f,int1a}\,(Q_{internal,1a}) + C_{f,solar3a}\,(Q_{solar,3a}) + C_g\,(Q_{solar}) + C_h\,(Q_{internal})$$

Concerning the optimal predictive methodology for this case study

- ✓ **Multiple linear regression model is a more suitable predictive tool than a physics-based model for this case study.** The complexity level of this tool is in line with the knowledge available of this case study. Any physics-based simulator (and therefore, LEA) is beyond the complexity limits of this case study because main influencing parameters (dynamic parameters and key indicators) are unknown and therefore estimated. This leads to the introduction of input data errors in the simulated results.

Regarding the significant value of the parameters and physical meaning of the coefficients

- ✓ **Significant value of the parameters.** The significant value of the independent variable is evaluated by the p-value and the t-statistics, but not by the magnitude of the coefficients since their order of magnitude differs between each other.
- ✓ **Key indicators coefficients for the detection of the variation of a parameter.** The variation of a coefficient in a multivariate linear regression influences several variables coefficients. However, there are some coefficients that are affected in a higher proportion than the others and act as key indicators of the variation of a parameter.
    - **Mass variation.** The indoor air temperature coefficient is the key indicator for a mass variation. The solar and internal gains coefficients are also affected but in a lower proportion.
    - **Insulation level.** The outdoor temperature coefficient is the key indicator for the variation of the envelop performance.
    - **Fraction windows.** The variation of the fraction windows in a particular façade is indicated by a change in the coefficient of the solar radiation profile on the corresponding inclination plane.
- ✓ **Symbols of coefficients.** The symbols of the coefficients are missing physical meaning as they depend of the constant and the combination of variables in the equation.

Concerning the number of equations necessary to build the multivariate linear regression model.

A new equation will be defined when there is a static variable influencing the correlation of the linear regression (slope of the least square). This can be detected by a visual analysis of the scatter plot of the heating demand versus the outdoor temperature as 2 or more point clouds will be distinguished. For this case study, the following equations are distinguished:

- ✓ **Seasonal differentiation**. It may be necessary to build a different multivariable model for (1) moderated weather months (heating and cooling operating during the same day) and (2) cold months (only heating) as the correlation may be affected when both heating and cooling are operating during the same day (due to the thermos-dynamic effect of the wall).
- ✓ **Operating mode differentiation**. The different operating modes affect the correlation between the variables and the heating demand, therefore it is necessary to build a different multivariable equation for each operating mode: (1) weekdays during opening hours, (2) weekends during opening hours and (3) closing hours.
- ✓ Day of the week differentiation (only needed when indoor surface temperature is not available). When the building cools down further during weekends than during weekdays, the heating demand pattern for Mondays may be differentiate. This phenomena is expected for buildings with low insulation level (lower than 2.5 W/m^2) and/or too heavy buildings (thermal mass > 300kg/m^2) and/or buildings with low influence of the internal and solar gains. In these cases, the fitting can be improved by making 2 different regression equations: (1) including only Mondays and (2) including all weekdays from Tuesday until Friday.

Regarding the selection of the data set to build in the multivariate linear regression model.

- ✓ **The data set will be selected according to the equation built**. Therefore, the data set will be classified according to the season, operating modes and day of the week (when needed).
- ✓ **The simulated data set is selected in this study as actual data sets are incomplete.** Both actual data sets (2015 and 2016) are incomplete to build the mathematical model and the missing parameters cannot be substituted by the estimated parameters as they were not correctly estimated. In contrast, it is proven that the simulated data sets are accurate and representative of the actual data set. Therefore the mathematical model will be built based on the simulated data set.
- ✓ **Variables needed in a complete actual data set for the implementation of this equation**.
    - **Option 1**: outdoor temperature and indoor surface temperature.
    - **Option 2**: outdoor temperature, global horizontal solar radiation, wind speed, internal heat gains, indoor air temperature and outflow temperature of the AHU (less relevant). In the cases where the radiators are manipulated intensively, it is advisable to measure this variable and include it in the equation.
- ✓ **Actual data set should correspond to a building where the heating demand is controlled by the indoor air temperature.** The buildings with indoor air temperature controller supply heat only when the building need it, decreasing the heating demand of the building and increasing the indoor comfort. This will be necessary in order to build an equation that optimizes the thermal energy and comfort of the building. In this case study, only the data survey of TPM fulfilled this requirement.
- ✓ **The optimal size of the data set used to build a linear regression around 2.5 months (seasonal period)**. A smaller size could decrease the accuracy if the period selected is not representative while a higher size will decrease its accuracy due to the annual periodicity presented in the outdoor weather parameters.

11.2 The mismatch between the actual and simulated data

The analysis of the gap between actual and simulated heating demand gives the following findings:

- ✓ **The gap between actual and simulated data results from the fact that the most influencing parameters differ from actual and simulated data as different key indicators are interfering**. The simulated heating demand is dominated by the building characteristics, while the actual heating demand depends on the different types of temperature control system and the manipulation of the radiators.
- ✓ **Actual and simulated data in IO and 3mE are not comparable**. The gap between simulated and actual data is higher for IO and 3mE than for TPM. As a result, the actual and simulated heating demand for IO and 3mE are not comparable, while they can be compared for TPM.

As a consequence of these findings, it is concluded that the calibrations of the parameters (performed prior to this thesis) are not accurate since they were based on the comparison between actual and simulated data. Therefore, the calibrations were evaluated in this study and the main conclusions are presented in the next section.

11.3 Regarding the calibration of the estimated parameters

- ✓ **Thermal mass.** The thermal mass for the buildings without indoor air temperature control system (3mE and IO) is underestimated as a result of a poor calibration performed with incomparable data sets.
- ✓ **Insulations of the building envelop.** The insulation parameters of the building envelope (windows and wall) should be increased for 3mE and slightly increased for IO and TPM.
- ✓ **Ventilation losses and infiltrations.** The ventilation losses and infiltrations should be decreased for 3mE and slightly increased for IO and TPM.
- ✓ **Low energy performance in the heat distribution system at 3mE.** It is observed that IO and 3mE have the same envelop characteristics, however the specific heating demand at 3mE is higher than in IO. This low energy performance of 3mE may be due to a lower energy performance in the heat distribution systems. This conclusion is based on the assumption that the total air volume of the buildings is correctly estimated.
- ✓ **New calibration procedure recommended.** The calibration procedure based on the analysis of the least square of the scatter plot (section 7.3.3) gives a better insights on the parameters estimated than the calibration procedure followed before to this thesis.

12 Recommendations and future developments

For this case study, the multivariate linear regression model is a more suitable predictive tool than a physics-based model (e.g. LEA). This is mainly because the number of parameters in the regression model is smaller and limited to a few measurable parameters. Any physics-based simulator (and therefore, LEA) is beyond the complexity limits of this case study because some of the most influencing parameters are unknown and therefore estimated. The estimation of parameters leads to the introduction of input data errors in the simulated results and is a very time consuming task.

As a result, the implementation of this data-driven predictive tool will save calibration time and will decrease the introduction of input data errors while providing a high accuracy of prediction. Moreover, the significant value of the parameters can be measured, giving a better insight of the influence of the different parameters in the building response than in the physics-based models.

In order to apply the developed linear regression model into practice, the following work is recommended.

- ✓ **Training the model with actual data.** For the application of this model into practice, the coefficients should be trained with actual data of buildings with indoor temperature control system (thermostat). For the buildings where the heating demand is not controlled by a thermostat, this should be installed before taking measurements. For the buildings at TU Delft, LEA can act as a thermostat once the calibrations are validated.
- ✓ **Testing the model for closing hours, weekends and moderated months.** The model presented is built based on data for winter period during weekdays and opening hours because is the most complex and relevant period. This model should be tested during weekends, closing hours and moderated months for its implementation all year round. It is expected that during closing hours, the equation will be mostly dependent on the outdoor temperature. During moderated months, the solar radiation will have higher weight on the equation and the solar radiation profile on the inclination plane should be readjusted to the corresponding season.
- ✓ **Including day of the week differentiation (only needed when indoor surface temperature is not available)**. The heating demand pattern for Mondays may be different for extreme heavy buildings (thermal mass > 300kg/m^2), and/or buildings with low insulation level, and/or buildings with low influence of the internal and solar gains. In these cases, the fitting can be improved by making 2 different regression equations: (1) including only Mondays and (2) including all weekdays from Tuesday until Friday.

Next sections present the guidelines for training this model with actual data and for the validation of the calibrations, respectively.

12.1 Guideline to train the multivariate regression model with actual data for TU Delft buildings.

1. **Data collection of the required parameters.** The data will be collected for around 2.5 months from a similar season than the predicted season. The data can be from a previous year or 2.5 months before the prediction.
    a. **Weather data:** outdoor temperature, global horizontal solar radiation, wind speed.
    b. **Building measurements:** indoor air temperature, internal heat gains, heating demand of the building and outflow temperature of the AHU (less relevant).
        i. The indoor air temperature should represent the average temperature profile of the total air volume contained in the building. The representative sample depends on the physical distribution of the building and it should include the measurements of rooms at each orientation side of the building and the rooms which contains a higher volume of air (eg. canteen, big lecture/study rooms, etc.).
        ii. The actual heating demand should corresponds to the total heat supplied by all heating supply systems installed in the building.
        iii. Outflow temperature of the AHU should correspond to the average temperature proportional to the inflow air of each air handling unit.
2. **Analysis of the solar radiation and internal heat gains alignment with respect the indoor air temperature.** The internal heating demand and solar radiation used should be delayed in order to align them in phase with the indoor air temperature. The alignment done in this thesis can be used for TPM. However, for IO and 3mE the alignment should be recalculated as the thermal mass used for these buildings in this thesis (estimated in the calibrations) is underestimated. As a result, it is expected that the solar radiation should be delayed by 4 or 5 hours (instead of 3 hours) and the internal heat gains around 2 or 3 hours (instead of 1 hour).
3. **Data set selection according to the type of equation**. A good data set selection is key to obtain an accurate prediction of the heating demand.
    a. **Seasonal.** The data set selected should be representative of the season to be predicted (moderated months or winter period). Data from the previous year or the previous 2.5 months can be used.
    b. **Operating modes.** The data set will be divided according to the operating mode to be modelled. For the TU Delft buildings 3 data sets are created: (1) weekdays during opening hours, (2) weekends during opening hours, (3) closing hours.
    c. **Day of the week.** When it is distinguish a different hourly heating demand pattern for the Mondays than for the rest of the weekdays, the data set will be divided as follows: (1) including only Mondays and (2) including all weekdays from Tuesday until Friday.
4. **Search procedure.** The training should be done using a search procedure able to analyse the data set residuals, and the significant level of the variables' coefficients and the model. This study uses the interactive stepwise function available in Matlab 2015b in the statistical toolbox.

12.2 Guideline for the validation of the calibrations (to improve the implementation of LEA).

In case that it is decided to continue the prediction of the heating demand in buildings with LEA, it is recommended to make a validation of the calibrations (performed prior to this thesis) in the following cases:

- ✓ **Buildings without thermostat (indoor air temperature control system).** This is because in these cases the actual and simulated heating demand are not comparable, leading to wrong estimations of the parameters. For these buildings, the new calibration procedure proposed in section 7.3.3 is strongly recommended.
- ✓ **When the existence or absence of cooling mode was not taken into account for the prior calibration procedure.** During the moderated months, the cooling influences the heating demand of the building. Therefore, the cooling mode should be activated to simulate the buildings with cooling mode and deactivated to simulate the buildings without cooling mode.

For performing the validation of the calibrations, the following practices are recommended

- ✓ **New calibration procedure recommended.** This research work recommends a new calibration procedure based on the analysis of the least square line on the scatter data is (see section 7.3.3 for further details). This calibration procedure is expected to be more accurate and less time consuming than the calibration procedure performed prior to this thesis.
- ✓ **Qualitative analysis prior the calibration.** Before performing a calibration, it is recommended to make a qualitative analysis of the most influencing parameters on actual data, simulated data and the comparison between them for the current case study (see section 7.2 for further details). The identification of key indicators and their interaction on the most influencing parameters will give good indications about whether the actual and simulated data set can be compared. This procedure will avoid poor estimation of parameters, leading to time saving and preventing the introduction of input data errors in the heating demand predictions.
- ✓ **Actual data sample representative from reality and in line with the simulated data.** In order to make a correct calibration of the parameters the actual and simulated data have to be comparable. Therefore, the simulator should be adapted to each case study and the actual data sample should be representative from reality and in line with the simulated data. Since LEA simulates the building as a single-zone, the simulated outputs corresponds to the total building. Therefore, the collection of data should be as explained in previous section.

References

[1] P. Stoelinga *et al.*, 'Intelligent Warmtenet Campus TU Delft, Eindrapportage'. Deerns, 23-Dec-2016.

[2] 'Smart heat grid on TU Delft campus-IPIN'. Ministry of Economic Affairs. Netherlands Enterprise Agency, Sep-2015.

[3] B. Metz *et al.*, 'IPCC - Intergovernmental Panel on Climate Change'. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

[4] International Energy Agency., *World energy outlook 2016*. Paris : Organisation for Economic Co-Operation and Development, 2016.

[5] 'File:Share of renewables in gross final energy consumption, 2014 and 2020 (%) YB16.png - Statistics Explained'. [Online]. Available: http://ec.europa.eu/eurostat/statistics-explained/index.php/File:Share_of_renewables_in_gross_final_energy_consumption,_2014_and_2020_(%25)_YB16.png. [Accessed: 31-Jan-2017].

[6] 'UN Environment Programme - Environment for Development : Buildings and Climate Change: Summary for Decision Makers |'. 2009.

[7] M. Levine *et al.*, 'Residential and commercial buildings, Climate Change 2007: Mitigation, Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change'. Cambridge University Press, Cambridge, U.K. & New York, NY, U.S.A.

[8] 'Heating and cooling - Energy - European Commission', *Energy*. [Online]. Available: /energy/en/topics/energy-efficiency/heating-and-cooling. [Accessed: 01-Feb-2017].

[9] H. Lund *et al.*, '4th Generation District Heating (4GDH): Integrating smart thermal grids into future sustainable energy systems', *Energy*, vol. 68, pp. 1–11, Apr. 2014.

[10] H. Lund, *Renewable Energy Systems: A Smart Energy Systems Approach to the Choice and Modeling of 100% Renewable Solutions*. Academic Press, 2014.

[11] D. Connolly *et al.*, 'Heat Roadmap Europe: Combining district heating with heat savings to decarbonise the EU energy system', *Energy Policy*, vol. 65, pp. 475–489, Feb. 2014.

[12] 'Heat Roadmap Europe'. [Online]. Available: http://www.heatroadmap.eu/publications.php. [Accessed: 26-Jan-2017].

[13] E. H. Borgstein, R. Lamberts, and J. L. M. Hensen, 'Evaluating energy performance in non-domestic buildings: A review', *Energy Build.*, vol. 128, pp. 734–755, Sep. 2016.

[14] D. Coakley, P. Raftery, and M. Keane, 'A review of methods to match building energy simulation models to measured data', *Renew. Sustain. Energy Rev.*, vol. 37, pp. 123–141, Sep. 2014.

[15] H. Zhao and F. Magoulès, 'A review on the prediction of building energy consumption', *Renew. Sustain. Energy Rev.*, vol. 16, no. 6, pp. 3586–3592, Aug. 2012.

[16] Z. Li, Y. Han, and P. Xu, 'Methods for benchmarking building energy consumption against its past or intended performance: An overview', *Appl. Energy*, vol. 124, pp. 325–334, Jul. 2014.

[17] T. A. Reddy and K. K. Andersen, 'An Evaluation of Classical Steady-State Off-Line Linear Parameter Estimation Methods Applied to Chiller Performance Data', *HVACR Res.*, vol. 8, no. 1, pp. 101–124, Jan. 2002.

[18] T. Catalina, V. Iordache, and B. Caracaleanu, 'Multiple regression model for fast prediction of the heating energy demand', *Energy Build.*, vol. 57, pp. 302–312, Feb. 2013.

[19] 'ASHRAE Guideline 14-2014 -- Measurement of Energy, Demand, and Water Savings', 2014.

[20] ‘ISO 16346:2013 - Energy performance of buildings -- Assessment of overall energy performance’, 2013.

[21] ‘ISO 13790:2008(en), Energy performance of buildings — Calculation of energy use for space heating and cooling’, 2008.

[22] T. Olofsson and S. Andersson, ‘Long-term energy demand predictions based on short-term measured data’, *Energy Build.*, vol. 33, no. 2, pp. 85–91, Jan. 2001.

[23] J. C. Lam, S. C. M. Hui, and A. L. S. Chan, ‘Regression analysis of high-rise fully air-conditioned office buildings’, *Energy Build.*, vol. 26, no. 2, pp. 189–197, 1997.

[24] A. Dhar, T. A. Reddy, and D. E. Claridge, ‘Using Fourier Series to Model Hourly Energy Use in Commercial Buildings’, 1993.

[25] H. Hens, *Building Physics -- Heat, Air and Moisture.* Wiley, 2007.

[26] H. S. L. C. Hens, *Applied Building Physics: Boundary Conditions, Building Performance and Material Properties*. John Wiley & Sons, 2010.

[27] M. Trčka and J. L. M. Hensen, ‘Overview of HVAC system simulation’, *Autom. Constr.*, vol. 19, no. 2, pp. 93–99, Mar. 2010.

[28] J. M. de Nijs, ‘Inverse modeling of buildings with floor heating and cooling systems for benchmarking operational energy use’, Master Science Thesis, Technische Universiteit Eindhoven, Eindhoven, 2015.

[29] A. Ioannou and L. C. M. Itard, ‘Energy performance and comfort in residential buildings: Sensitivity for building parameters and occupancy’, *Energy Build.*, vol. 92, pp. 216–233, Apr. 2015.

[30] V. I. Soebarto and T. J. Williamson, ‘Multi-criteria assessment of building performance: theory and implementation’, *Build. Environ. Perform. SimulationCurrent State Future Issues*, vol. 36, no. 6, pp. 681–690, Jul. 2001.

[31] O. Guerra Santin, L. Itard, and H. Visscher, ‘The effect of occupancy and building characteristics on energy use for space and water heating in Dutch residential stock’, *Energy Build.*, vol. 41, no. 11, pp. 1223–1232, Nov. 2009.

[32] S. A. Al-Sanea, M. F. Zedan, and S. N. Al-Hussain, ‘Effect of thermal mass on performance of insulated building walls and the concept of energy savings potential’, *Appl. Energy*, vol. 89, no. 1, pp. 430–442, Jan. 2012.

[33] H. Christina Johanna, ‘Uncertainty and sensitivity analysis in building performance simulation for decision support and design optimization’, PhD diss., Eindhoven University, 2009.

[34] V. Cheng, E. Ng, and B. Givoni, ‘Effect of envelope colour and thermal mass on indoor temperatures in hot humid climate’, *Sol. Energy*, vol. 78, no. 4, pp. 528–534, Apr. 2005.

[35] D. Majcen, L. C. M. Itard, and H. Visscher, ‘Theoretical vs. actual energy consumption of labelled dwellings in the Netherlands: Discrepancies and policy implications’, *Decades Diesel*, vol. 54, pp. 125–136, Mar. 2013.

[36] D. Majcen, L. Itard, and H. Visscher, ‘Actual and theoretical gas consumption in Dutch dwellings: What causes the differences?’, *Energy Policy*, vol. 61, pp. 460–471, Oct. 2013.

[37] O. Guerra-Santin and L. Itard, ‘The effect of energy performance regulations on energy consumption’, *Energy Effic.*, vol. 5, no. 3, pp. 269–282, Aug. 2012.

[38] D. Majcen, L. Itard, and H. Visscher, ‘Statistical model of the heating prediction gap in Dutch dwellings: Relative importance of building, household and behavioural characteristics’, *Energy Build.*, vol. 105, pp. 43–59, Oct. 2015.

[39] R. Kramer, J. van Schijndel, and H. Schellen, ‘Simplified thermal and hygric building models: A literature review’, *Front. Archit. Res.*, vol. 1, no. 4, pp. 318–325, Dec. 2012.

[40] ‘TU Delft Energy Monitor’. [Online]. Available: http://www.energymonitor.tudelft.nl/Building. [Accessed: 27-Jan-2017].

[41] *Sustainable Urban Environments - An Ecosystem Approach* | *Ellen M. van Bueren* | *Springer*. .

[42] 'ISO 14683:2007 - Thermal bridges in building construction -- Linear thermal transmittance -- Simplified methods and default values', *ISO*. [Online]. Available: http://www.iso.org/iso/catalogue_detail.htm?csnumber=40964. [Accessed: 13-Feb-2017].

[43] 'ISSO-publicatie 54 Energie Diagnose Referentie'. [Online]. Available: http://kennisbank.isso.nl/docs/publicatie/54/2011. [Accessed: 16-Feb-2017].

[44] C. Neumann, 'Description of European Prototype Tool for Evaluation of Building Performance and the national tools', *Build Up*, 27-Sep-2009. [Online]. Available: http://www.buildup.eu/en/practices/publications/description-european-prototype-tool-evaluation-building-performance-and. [Accessed: 02-Mar-2017].

[45] 'Stepwise regression - MATLAB stepwisefit - MathWorks Benelux'. [Online]. Available: https://nl.mathworks.com/help/stats/stepwisefit.html. [Accessed: 23-Jan-2017].

[46] J. P. M. de Sá, *Applied Statistics Using SPSS, STATISTICA, MATLAB and R*. Springer Science & Business Media, 2007.

[47] 'Overfitting', *Wikipedia*. 02-Jan-2017.

[48] van Paassen, 'Indoor climate fundamentals. TU Delft course (wb4426). Master mechanical engineering, section energy technology in build environment.' .

[49] M. Veldkamp, 'FTO Monitoring IPIN TU Delft, Functioneel en Technisch Ontwerp IPIN Warmtenet'. van Beek, 14-Jun-2016.

Appendix 1 Current district heating system at TU Delft campus

1.1 Heating distribution: Overview TU Delft District Heating Grid

TU Delft District Heating Grid (DHG) supplies the heating needed to most of the buildings at TU Delft campus. Some buildings such as 3mE (Faculty of Mechanical, Maritime and Materials Engineering), EWI(Faculty Electrical Engineering, Mathematics and Computer Science) and the university library are partially heated and cooled by an underground heat and cold storage system combined with heat pumps [40]. The DHG is divided in 4 different branches: North 1 (Noord 1), North 2 (Noord 2), South 1 (Zuid 1) and South 2 (Zuid 2). Figure 32 presents the overview of the TU buildings connected to the different branches.

Figure 32 Overview of the buildings connected to TU Delft District Heating Grid [49].

The buildings studied in this project and in the current phase of IPIN (phase 2) are the ones connected to branch North 2. These are 3mE (Faculty of Mechanical, Maritime and Materials Engineering), IO (Industrial Design) and TPM (Technology, Policy and Management). Figure 33 shows an overview of the different heating distribution stations of branch North 2 for each of the buildings analysed: 3mE (green), IO (yellow) and TPM (blue). Branch 2 contains a total of 11 heat distribution stations, where 1 is situated in TPM, 2 in IO and 8 in 3mE [1].

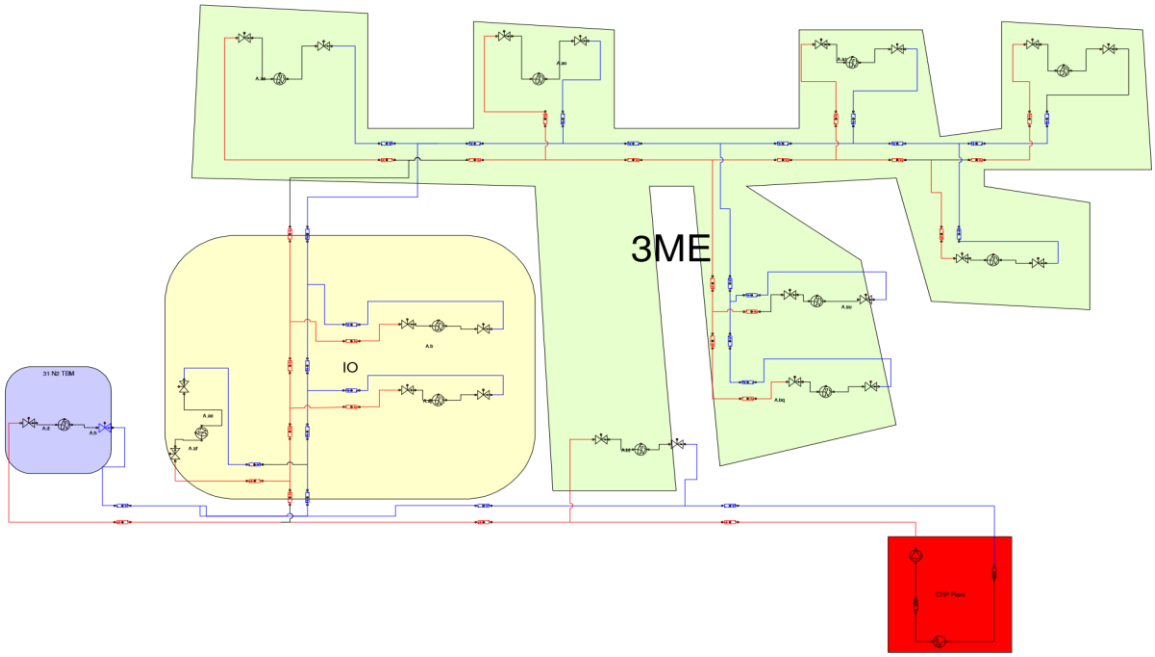


Figure 33 Overview heat distribution stations model branch North 2 of TU Delft District Heating Grid [1].

1.2 Heating Generation: Description Central Heating Plant

The heat distributed at the TU Delft DHG is generated in a Central Heating Plant (CHP) (represented in Figure 33 in red). The CHP has a total thermal capacity of 84 MW_{th} and the heat is generated by combining 3 gas boilers and 2 Combined Heat & Power Units (CHPU). The gas boilers have a total thermal capacity of 15, 30 and 35 MW_{th}, respectively, and the CHPUs 2MW_{th} each one [1].

Figure 34 presents the scheme of the heating generation system at the TU delft Central Heating Plant.

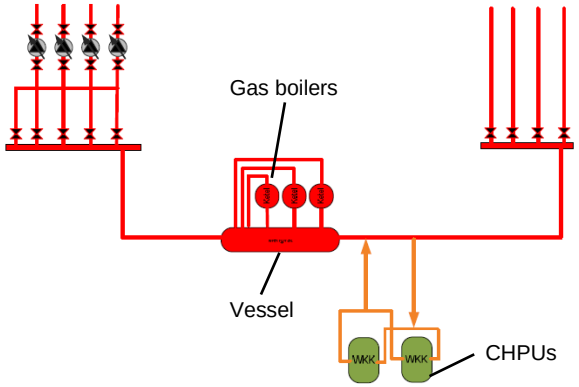


Figure 34 Scheme of the heating generation system at TU Delft Central Heating Plant [1].

The gas boilers heats up the water at high temperatures (between 100 and 130ºC), while the CHPUs at medium temperatures (between 80 and 100ºC) and work in a range temperature between 50 and 83°C [1]. When the return water coming from the thermal grid has a lower temperature than 83ºC, the water is (pre) heated with the CHPUs before being sent to the vessel. On the contrary, the CHPUs are bypassed when the return temperature is higher than 83ºC. In both cases, the gas boilers supply the remaining heat to reach the required supply temperature when needed. Currently, the required supply temperature is just dependent on the outdoor temperature (maximum of 130ºC when the outdoor temperature is -10ºC) and according to the heating curve.

The efficiency of the heating generation system increases at a continuous operating mode due to the start-up time needed for the gas boilers and CHPUs (average start-up time of the gas boilers approximately 1 hour, this is dependent on the capacity). A continuous operating mode for the gas boilers is determined by a constant heat supply, therefore: $\left(T_{supply} - T_{return}\right) \cdot \dot{V} = constant$, where T_{supply} and T_{return} are the water supply and return temperatures, respectively; and $\dot{V}$ is the water flow rate (m^3/s). While a continuous operating mode for the CHPUs is determined by both a continuous heat supply and a $T_{return} \leq 83^oC$.

The heat consumption in the heating generation & distribution system is reduced at a lower supply temperature. In the heating generation system, a lower supply temperature reduces the ΔT (T_{supply} - T_{return}), leading to a decrease in heating supply and contributing to a lower return temperature. A low temperature district heating grid will lead to a reduction in the transport heat losses due to a lower heat transference with the outdoor environment.

According to the heating generation systems implemented in the TU Delft CHP, it is concluded that the following objectives need to be pursue in order to achieve a better energy performance of the current heating generation & distribution system:

1. Continuous operating mode by:
    a. Constant heating supply: $\left(T_{supply} - T_{return}\right) \cdot \dot{V} = constant$. This optimizes the use of the gas boilers and CHPUs.
    b. Lowering return temperature ($T_{return} \leq 83^oC$). The use of the CHPUs is maximized and optimized.
2. Lowering the supply temperature of the TU Delft DHG. This will reduce the heat consumption (lower ΔT) and the transport heat losses.

Appendix 2 Energy balances in a building energy simulation model

Section 6.1.1 explained the overview and the main assumptions of the thermal balances in LEA. This appendix gives a detailed explanation of the methodology used for the thermal balances in the room (static model) and the floor (dynamic model).

2.1 Room model

The thermal model of the room is calculated according to a steady state situation for every time interval. The energy demand of the building corresponds to the following thermal energy balance:

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

Every time step, LEA predicts the energy demand (Q_{demand}) required to reach the indoor air temperature according to the corresponding temperature set point (s).

The calculation method of the different heat losses/gain are described below. The transmission losses due to thermal bridges (thermal losses occurring at the junction of walls, ceiling, floor or wall and windows frame) are neglected.

Heat transmission through the envelop ($Q_{envelop}$):

LEA calculates the heat transmission for each of the four facades (orientation dependent) and the roof. The heat transmissions for each façade and the roof are expressed according to the following equation:

$$Q_{envelop} = \sum_i U^j_{envelop}.A^j_{envelop}.(T_o - T_i)$$

Where, for each façade/roof of orientation j, $U^j_{envelop}$ corresponds to the heat transfer coefficient and $A^j_{envelop}$ the surface of the facade/roof. The heat transfer coefficient of each façade and roof is defined as follows.

$$U^j_{envelop} = (U_{windows}.\%windows + U_{wall/roof}.\%wall)$$

Where the heat transfer coefficient for the walls and roof is defined as $U_{wall/roof} = \frac{1}{\frac{1}{\alpha_i}+R_c+\frac{1}{\alpha_o}}$

Where, α_i corresponds to the indoor combined heat transfer coefficients for convection and radiation. The outdoor combined heat transfer coefficients (α_o) are dependent on the wind speed. In LEA, both α_i and α_o are assumed to be constant values. The heat transfer coefficient for the windows corresponds to a constant value, thus LEA assumes that it is independent from α_i and α_o. The thermal resistance of the composed wall (R_c) is dependent on the wall and roof characteristics. $A_{envelop}$ is the area corresponding to the roof or façade for the respective orientations. T_i corresponds to the indoor air temperature. T_o corresponds to the outdoor temperature.

Heat transmission through the ground

The heat transmission through the ground is calculated according to the following equation:

$$Q_{ground} = U_{floor}.\ A_{floor}.(T_{ground} - T_i)$$

Where, T_{ground} corresponds to the surface temperature of the ground (soil) and A_{floor} the area corresponding to the ground floor surface. The heat transfer coefficient for the floor is defined as $U_{floor} = \frac{1}{\frac{1}{\alpha_i}+R_c}$

Internal heat gains

The internal heat gain refers to the energy flow produced by people, lighting and other electrical devices. Thus, the total internal heat gain is calculated according to the following equation.

$$Q_{internal} = Q_{int.people} + Q_{int.light} + Q_{int.appliances}$$

The internal heat gain comes as heat into the building through convection and radiation, therefore the percentage of the total internal heat transmitted by radiation and convection are differentiated in LEA.

The quantity of heat released by the people will depend on the occupancy and the heat dissipated by a human body and is calculated by the following equation.

$$Q_{int.people} = n_{people} . Q_{body}$$

Where, n_{people} is the number of people and depends on the week and weekend schedule. Q_{body} [wh] corresponds to the heat gain per person which depends mainly on the type of activity deployed, therefore LEA assigns the heat gain depending on the use of the building.

The internal heat gain due to artificial lighting is calculated according to the following equation.

$$Q_{int,lighting} = A_{ceilings} . Q_{light}$$

Where $A_{ceilings}$ is the total area of all the ceilings, therefore it is assumed that the lightings are distributed along the whole ceiling area. $Q_{light\ [\frac{wh}{m^2}]}$ corresponds to the lighting power which is hourly dependent on the week and weekend schedule.

The internal heat gain due to electrical appliances is calculated as follows.

$$Q_{int,appliances} = A_{floor} . Q_{applicances}$$

The appliances are assumed to be equally distributed along all the floor area (A_{floor}). The total power of all appliances ($Q_{appliances}\ [\frac{wh}{m^2}]$) is dependent on the use of the building and the week and weekend schedule.

Solar heat gain

The total solar gain is the amount of solar energy accumulated in the building. LEA calculates the percentage of solar heat gain transmitted by radiation and convection, and uses a different calculation method for direct, diffused and reflective solar radiation.

$$Q_{solar} = Q_{sol\ direct} + Q_{sol\ dif.} + Q_{reflective}$$

The total solar heat gain depends on the percentage of windows in the facade and façade orientation. The direct solar radiation depends on the reflection properties of the window glazing. The reflective solar radiation corresponds to the solar radiation reflected on the ground and is calculated based on the direct, diffused solar radiation and the albedo (ground reflection). Further details about the calculation of the reflective solar radiation can be found in the LEA's function 'bereken zonbelasting'

Mechanical ventilation

The ventilation air is preheated in an air handling unit (AHU) and enters in the room at a defined temperature ($T_{\text{out AHU}}$). The heat transferred through mechanical ventilation is calculated by the following equation.

$$Q_{ventilation} = m_{vent.} . C_{pair}\ (T_{\text{out AHU}} - T_{\text{i}})$$

Where C_{pair} corresponds to the heating capacity of air (J/kg.K) and $m_{vent.}$ to the mass flow rate of the ventilation air (kg/s). The mass flow rate is defined as $m_{vent.} = \dot{V}_{vent.} . \rho_{air}$; where $\dot{V}_{vent.}$ is the volume flow rate of ventilation (m^3/s) which is hourly dependent on the weekly and weekend

schedule. The calculation of the ventilation air ($T_{\text{out AHU}}$) is described in the set of functions AHU (LBK).

Infiltrations

The infiltrations correspond to the outside air which enters into the buildings through cracks in the construction and openings. The infiltration losses are calculated as follows.

$$Q_{infiltrations} = (m_{openings} + m_{cracks}).C_{pair}\ (T_{\text{o}} - T_{\text{i}})$$

Where, $m_{openings}$ is the mass flow rate of air entering through openings which is defined as $m_{openings.} = \dot{V}_{openings.}.\rho_{air}$; where $\dot{V}_{openings}$ is a determine volume flow rate (m^3/s) which is hourly dependent on the weekly and weekend schedule. m_{cracks} corresponds to the mass flow rate of air entering through cracks in the building which is calculated dependent on the wind speed according to the following equation [NEN-EN 12207].

$$m_{cracks} = V_{building} \cdot 0.15 \cdot \left(\frac{V_{wind}^2}{V_{reference}^2}\right)^{2/3}$$

Heat transmission through indoor surfaces ($Q_{\text{thermal mass}}$)

The heat accumulated in the building (and transferred to the indoor air) is determined by the heat transmitted through the indoor surfaces in contact with the indoor climate ($Q_{thermal\ mass}$). Therefore, the heat is transmitted through the total area of indoor surfaces in contact with the indoor air ($A_{indoor\ surfaces}$). For buildings accumulating most of their heat in floors and ceilings, this indoor surface area corresponds to the total area of floors and ceilings in contact with the indoor air.

The heat transmitted to the indoor surfaces is calculated as follows:

$$Q_{thermal\ mass} = \alpha_i.\ A_{indoor\ surfaces}.(T_s^t - T_i)$$

Where T_s^t is the temperature corresponding to the indoor surface and is calculated in the floor model (see section 2.2).

2.2 Floor model

The thermal mass of the building is accumulated in the total mass of the building and the heat is transmitted through the indoor surfaces (floors and ceilings). The thermal model of the floor is calculated using the method finite-difference approximations for one-dimensional unsteady conduction (Millers,1999).

The accuracy of this method will depend on the number of temperature nodes defined, these are determined in function of the number of slabs. LEA defines the number of slabs dependent on the total mass of the building. Table below shows the number of slabs in the building depending on the total mass of the building.

Table 15 Number of slabs corresponding to the mass of the building

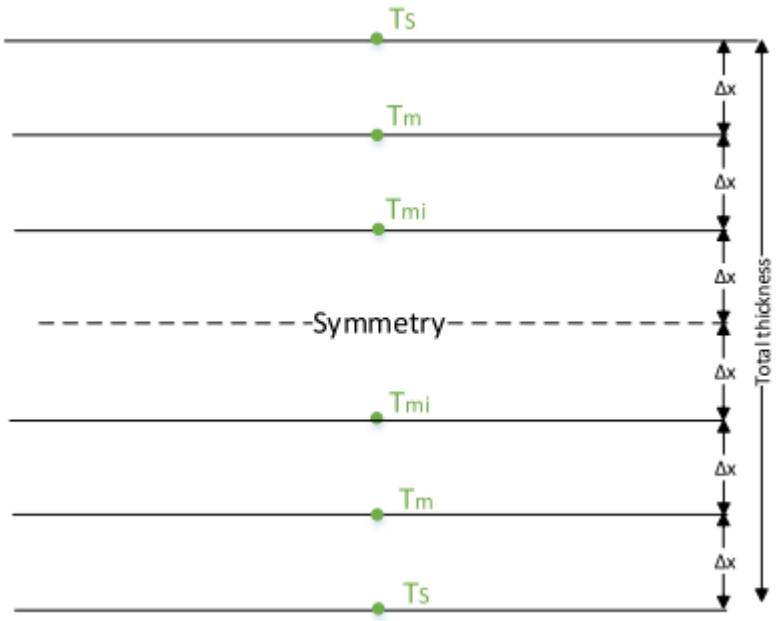
Building mass [kg/ m^2]	N_{slabs}
<100	2
<250	3
>250	4

The thickness of each slab is directly proportional to the half of the total mass of the building and the total number of slabs according to the equation below.

$\Delta x = \frac{X}{2 \cdot N_{slab}}$; where X is the sum of the total floor and ceilings thickness defined as:

$X\ [m] = \frac{mass\ building\ [\frac{kg}{m^2}]}{\rho_{concrete}\ [\frac{kg}{m^3}]}$;

The equation of the nodes can be broken down in 3 characteristic equations: node at the surface of the indoor floor with heat accumulation in one side (node 's'), nodes at the inside of the indoor layer of the floor with heat accumulation in both sides (nodes 'm'), node at the middle of the inside layer (node 'mi'). Figure below shows the representation of the 3 characteristic nodes.



The heat balances according to the 3 mentioned characteristic equations are explained below.

Heat accumulation in floor surface ($T_s^t \rightarrow T_s^{t+\Delta t}$)

The surface temperature of the indoor surfaces is calculated by considering the absorption of the solar radiation. The heat accumulated in the floor surface varies the surface floor temperature from T_s^t to $T_s^{t+\Delta t}$ every time interval (Δt) and it is accumulated in both sides of the slab. The heat accumulated in node 's' is assumed to be equal to the sum of the heat fluxes between the node temperature 's' (T_s) and 'I' (T_i), and the heat conduction between the node temperature 's' (T_s) and 's+1' (T_{s+1}). The heat fluxes connected to the indoor air temperature corresponds to the solar radiation, internal heat radiation, absorbed radiation and heat convection with the indoor air.

Heat accumulation in interior nodes ($T_m^t \rightarrow T_m^{t+\Delta t}$)

For the interior nodes, it is assumed that the accumulation is equal to the sum of the conductions at both sizes. This calculation is based on the Fourier equation, expressed as follows:

$$\rho_i * c_{p_i} * \frac{\partial T}{\partial t} = k_i * \frac{\partial^2 T}{\partial x^2}$$

The method used to solve this differential equation is the finite-difference approximations for one-dimensional unsteady conduction (Millers,1999).

The heat balance in node m is expressed as follows:

$$\frac{\lambda \cdot A}{\Delta x} \cdot (T_{m-1}^t - T_m^t) + \frac{\lambda \cdot A}{\Delta x} \cdot (T_{m+1}^t - T_m^t) = \frac{\rho \cdot \Delta x \cdot c_p}{\Delta t} \cdot A \cdot \left({T_m}^{t+\Delta t} - {T_m}^t\right);$$

Where the heat transference phenomena taking place are the following:

- Conduction transference between m and m-1: $\frac{\lambda \cdot A}{\Delta x} \cdot (T_{m-1}^t - T_m^t)$
- Conduction transference between m and m+1: $\frac{\lambda \cdot A}{\Delta x} \cdot (T_{m+1}^t - T_m^t)$
- Accumulation in node m from t to t+Δt: $\frac{\rho \cdot \Delta x \cdot c_p}{\Delta t} \cdot A \cdot \left({T_m}^{t+\Delta t} - {T_m}^t\right)$

The heat balance above is simplified in the following equation (Millers,1999):

$$T_m^{t+\Delta t} = Fo\ (T_{m-1}^t + T_{m+1}^t) + (1 - 2 \cdot Fo) T_m^t$$

Where, Fo is the mesh Fourier number: $Fo = \frac{\alpha \cdot \Delta t}{\Delta x^2} \leq \frac{1}{2}$, being α the thermal diffusion coefficient of the material (concrete in this case), expressed as $\alpha = \frac{\lambda}{\rho \cdot c_p}$

Heat accumulation in middle node ($T_{mi}^t \rightarrow T_{mi}^{t+\Delta t}$)

For the heat accumulation in the middle node, it is assumed an adiabatic surface condition. This means that the next node (mi+1) has the same temperature as node mi ($T_{mi} = T_{mi+1}$), thus there is no heat transferred neither accumulated between node mi and mi+1. Therefore, the accumulation in node mi is assumed to be equal to the conduction with the previous node (mi-1). As a result, the accumulation is only taking place in half side of the slab, therefore, Δx is divided by 2.

The heat balance corresponding to the node 'mi' is expressed as follows:

$$\frac{\lambda}{\Delta x} \cdot A \cdot (T_{mi-1}^t - T_m^t) = \frac{\rho \cdot (\frac{\Delta x}{2}) \cdot c_p}{\Delta t} \cdot A \cdot \left({T_{mi}}^{t+\Delta t} - {T_{mi}}^t\right);$$

Where the heat transference phenomena taking place are the following:

- Conduction transference between mi-1 and m: $\frac{\lambda \cdot A}{\Delta x} \cdot (T_{mi-1}^t - T_{mi}^t)$
- Conduction transference between mi and mi+1=0

- Accumulation in node m from t to t+Δt: $\frac{\rho\cdot(\frac{\Delta x}{2})\cdot c_p}{\Delta t}\cdot A\cdot\left({T_{mi}}^{t+\Delta t}-{T_{mi}}^{t}\right)$

This heat balance can be expressed by the following equation (Millers,1999):

$$T_{mi}^{t+\Delta t}=2\cdot Fo\cdot T_{mi-1}^{t}+(1-2\cdot Fo)T_{mi}^{t}$$

Where, $Fo=\frac{\alpha\cdot\Delta t}{\Delta x^2}\leq\frac{1}{2}$ and $\alpha=\frac{\lambda}{\rho\cdot C_p}$

Appendix 3 LEA Energy Diagnose

The calculation method in LEA was verified by performing the energy test ISSO-54: 'Energy Diagnose Reference' [43]. This test validates the heating & cooling demand, air handling unit and other installations like the water heating, lighting and photovoltaic energy. In this case, only the annual heating demand is verified.

The goal of this diagnose is to verify the accuracy of the heat balance in LEA's calculation method. With this purpose, the heating demand supplied by the radiators is analysed by applying the tests from A.1.1.01 till A.1.1.16 (except A.1.1.13 since the result is not provided by BRL 9501). These tests are meant to analyse the annual heating demand by varying the U values, thermal mass, opening times, natural ventilation and internal heat gains (people, lighting, appliances and solar).

The hourly heating demand cannot be analysed with ISSO-54, therefore other type of analyses will be needed to study the hourly heating demand fluctuations. In order to verify the calculation method used in the mechanical ventilation, other tests available in ISSO-54 can be performed.

3.1 Methodology: assumptions and changes made

The input data given in ISSO are prepared for commercial software, this means that LEA's inputs vary in format. Therefore, in some cases, the code is changed in order to adapt the inputs of the ISSO study case into LEA. In some cases, test's inputs are not used by LEA due to an approximation in the calculation method used. This section presents the assumptions and code changes done in order to adapt the ISSO inputs into LEA format.

All the tests are based on the study case presented in test A.1.1.01 (baseline) and the inputs introduced in LEA can be found in the following section.

Climate file

The climate file used for the modelling of IO are obtained from KNMI. The horizontal diffused radiation is assumed to be 15% of the horizontal radiation (global radiation on the horizontal plane) and the horizontal direct radiation the other 85%. However, in order to analyse correctly the calculation method used in LEA, the weather file have to be as accurate as possible. Therefore, the weather data used are all measurements. For the horizontal diffused radiation, horizontal direct radiation, temperature and relative humidity, the weather file of Laure is used. The rest of the variables (wind direction, wind speed, relative pressure) are obtained from KNMI.

The reflective coefficient albedo given by ISSO-54 is 0.2 (ISSO 54, section 2.1.1.1). This corresponds to the input: invoer.overig.reflectiecoefficient.

Total building mass

LEA requires the total building mass, however ISSO gives this information in terms of material specifications (material density, thickness and area). The total building mass is calculated considering that the different components of the roof, façade and floor covers all the surface area of the roof, façade and floor, respectively. For this calculation, all the building materials are taken into account (including the windows mass). The density of the windows indicated in ISSO is 2.5 kg/m3. Considering that glass density is in the order of 2500 kg/m3, it is assumed that ISSO is not correct, therefore for the calculations it is assumed that the windows density is 2500 kg/m3. See appendix F for further details on the mass building calculation.

Windows

ZTA value changes with the sun angle of incidence. LEA calculates the angle of incidence of the sun, however, ZTA is a constant (it is not function of the angle of incidence). Therefore, the ZTA considered for the tests is a constant value of 0.831 which corresponds an angle of incident of 45^{o}.

The LTA value is not used in the current version of LEA, therefore a 0 value is given.

Shadow

The specification of the tests do not include shadow, therefore all LEA input values for the shadow are 0.

Internal gains

The internal gains are calculated in LEA by the function 'bereken_interne_warmtelast'. This function gives the output 'int_wl' (int_wl.personen, int_wl.apparatuur, int_wl.verlichting) which is used to calculate the internal heat gain profile. It is distinguished a total of three types of internal heat gain profiles in the tests realized, therefore three different type of functions are introduced in the code (input: functie type), these are the following:

1. Functie.type=0. This option is used in the tests where the internal heat production is constant (all tests performed except A.1.1.07 - A.1.1.10). The constant value is introduced directly in the function.
2. Functie.type=0.1. This option is used when the internal heat production is constant during the day and changes its value during summer and winter (tests A.1.1.07 and A.1.1.09)
3. Functie.type=0.2. This option is used when the internal heat production has a variable daily and seasonal (summer and winter) profile. The inputs are changed from the input file (s.invoer.last). (tests A.1.1.08 and A.1.1.10)

For further details on the code changes, see appendix G.

The input 'functie.type' is also present in other functions such as 'Bereken_bedrijfstijden' and 'bereken_genzen' (office limits calculation). Therefore, new cases for functie.type: 0, 0.1 and 0.2 are added.

LEA assume that a 30% of the internal heat is transmitted by radiation and the other 70% by convection (input: convection factor for people, lighting and equipments). ISSO test assumes that 40% is transmitted by convection and 60% by radiation.

Ventilation

According to the tests specifications, the mechanical ventilation is cancelled for all the tests, therefore the air is renovated only by natural ventilation. The functions affected by this change are: 'Bereken_infiltratievoud', 'Ventilatiedebieten ',Bereken_coefficient', 'Doe_berekening' and 'Simuleer_zwaar_gebouw'. The corresponding changes are made by adding a switch case for functie.type: 0, 0.1, 0.2.

Heat transfer coefficient

LEA uses the combined heat transfer coefficient for convection and radiation (used in balans (2), (3) and (6) in function: simuleer_zwaar_gebouw). Therefore, in the test it is used the combined heat transfer coefficient for convection and radiation (α_i=8, α_o=23). In the case of the indoor surfaces, there is a distintion between horizontal and vertical surfaces in case that there is upward and downward heat. Since LEA does not distinguish between upward and downward heat flows, the combined heat transfer coefficient for convection and radiation used is for vertical surfaces.

3.2 Inputs for the different tests respect to the baseline study case

Table 16 shows the LEA's code used and the inputs modified for each test performed with respect to the baseline case study (A.1.1.01). The baseline case study uses function type (functie.type)=0, has a constant internal heat production, natural ventilation of 1 change of the volume per hour and a daily temperature profile which does not differ from the week days and weekends.

Table 16 Version LEA used and inputs modified for tests from A.1.1.01 till A.1.1.16 with respect to the baseline case study

Test number	Name	LEA's code used	Input modified respect to the baseline
A.1.1.01 (baseline)	energy losses	LEA.exe_for validation - Cristina_V2	(baseline).
A.1.1.02	continuous operation	LEA.exe_for validation - Cristina_V3	Input functie.type=0 Constant temperature set point: the set point is fixed all year and during night and day at 20^{o}C
A.1.1.03	distinction week/weekend	LEA.exe_for validation - Cristina_V3	Input functie.type=0 Variable temperature set point: Monday-Friday: From 7-23h T=20^{o}C and from 23-7h T=15^{o}C Saturday-Sunday: From 0-24h, T=15 ^{o}C
A.1.1.04	modified U/value	LEA.exe_for validation - Cristina_V2	Input functie.type=0 modify U-value (input: dichtgevels.Rc) Roof Rc decreased from 2.95 to 1. External facades from 1.66 to 0.91. Floor keeps the same value: 25.1.
A.1.1.05	modified ventilation quantity	LEA.exe_for validation - Cristina_V2	Input functie.type=0 ventilatie.kantoor.natuurlijk=2

A.1.1.06	further increase fan/old	LEA.exe_for Cristina_V2	validation	-	Input functie.type=0 ventilatie.kantoor.natuurlijk=4
A.1.1.07	reduced internal heat production	LEA.exe_for Cristina_V3.	validation	-	Input functie.type=0.1 Internal heat production is constant during the day and changes its value during summer and winter (this input is changed in fx:bereken_interne_warmtelast) summer: 240 winter: 32
A.1.1.08	variable internal heat production	LEA.exe_for Cristina_V3.	validation	-	Input functie.type=0.2 Internal heat production has a variable daily and seasonal (summer and winter) profile. The inputs are changed from the input file (s.invoer.last).
A.1.1.09	increased internal heat production	LEA.exe_for Cristina_V3.	validation	-	Input functie.type=0.1 Internal heat production is constant during the day and changes its value during summer and winter (this input is changed in fx:bereken_interne_warmtelast). The internal heat is increased respect test 1.1.07 summer: 800 winter: 320
A.1.1.10	variable/increased internal heat production	LEA.exe_for Cristina_V3.	validation	-	Input functie.type=0.2 Internal heat production has a variable daily and seasonal (summer and winter) profile. The inputs are changed from the input sheet (s.invoer.last).
A.1.1.11	impact glass percentage	LEA.exe_for Cristina_V2	validation	-	Input functie.type=0

					modify % glass (input: dichtgevel.raamfractie) from 0.56 to 0.28
A.1.1.12	impact orientation	LEA.exe_for Cristina_V2	validation	-	Input functie.type=0 window orientation changed.
A.1.1.14	influence of ZTA-value	LEA.exe_for Cristina_V2	validation	-	Input functie.type=0 Modify ZTA-value for 45^{o}C from 0.31(baseline) to 0.64 Convection factor changed from 0.012 to 0.039 Note: LEA model does not use most of the parameters given, therefore only the parameters mentioned were introduced.
A.1.1.15	influence U-value glass	LEA.exe_for Cristina_V2	validation	-	Input functie.type=0 changed U value and ZTA (for 45^{o}C).
A.1.1.16	influence thermal mass	LEA.exe_for Cristina_V2	validation	-	Input functie.type=0 Modify building mass: the calculation method of LEA accumulates all the mass of the building in the floor. Therefore, the mass of the building is recalculated for the lighter walls but it was not observed appreciable change in the total mass. The mass density stays at 82.93 kg/m^2 The Rc value of the external façade is changed to 1.65 m^2 k/W

3.3 Results and discussions

This section presents the results obtained for the simulations performed for the tests case studies A.1.1.01-A.1.1.16 (except A.1.1.13). Table 17 presents the reference values (BRL 9501) and results for the annual heating demand for the tests performed in this diagnose.

According to the results obtained, it is observed that the annual heating demand in all the tests performed lay within the bandwidth of the reference values. Therefore, the heat balance used in the calculation method of LEA.exe makes a good estimation of the annual heating demand, however this test does not verify the accuracy in the hourly heating demand. This means that the low prediction of heating demand during summer compensates the high heating demand prediction during winter, leading to a good prediction of the annual heating demand.

A good accuracy for the hourly heating demand is key for the correct performance of the model predictive control, and therefore the Smart thermal grid. For that reason, the next sections studies the hourly heating demand by performing a multivariable and sensitivity analysis.

It is important to remind that for all tests the mechanical ventilation was turned off, and therefore its possible influences on the heat balance are not analysed.

Table 17 Reference values (BRL 9501) and results of the simulations for the annual heating demand for the tests A.1.1.01-A.1.1.16 (based on weather reference year 1964/65).

Test	Description	Reference values		simulation results	
		Max(GJ/year)	Min (GJ/year)	Energy demand (kW.h/year)	Energy demand (GJ/year)
A.1.1.01	energy losses	30.2	20.2	7.97E+03	28.70
A.1.1.02	continuous operation	35.7	23.8	9.17E+03	33.00
A.1.1.03	distinction week/weekend	27.3	18.2	7.22E+03	26.01
A.1.1.04	modified U/value	42.9	28.6	1.12E+04	40.40
A.1.1.05	modified ventilation quantity	42.7	28.5	1.03E+04	37.19
A.1.1.06	further increase fan/old	68.2	45.5	1.51E+04	54.34
A.1.1.07	reduced internal heat production	33.9	22.6	8.88E+03	31.95
A.1.1.08	variable internal heat production	32.8	21.8	8.53E+03	30.70
A.1.1.09	increased internal heat production	18.6	12.4	4.99E+03	17.97
A.1.1.10	variable/increased internal heat production	25.6	17.1	6.84E+03	24.63
A.1.1.11	impact glass percentage	26.6	17.8	6.48E+03	23.33
A.1.1.12	impact orientation	34.5	23	7.45E+03	26.83
A.1.1.14	influence of ZTA-value	32.9	21.9	8.28E+03	29.80
A.1.1.15	influence U-value glass	17.6	11.7	4.61E+03	16.59
A.1.1.16	influence thermal mass	30.7	20.5	7.99E+03	28.75

Appendix 4 Corrections in LEA´s inputs

4.1 Weather file

For the validations of the measurements 2015/2016 done for IO, TPM and 3ME, two inaccuracies are observed: (1) the weather data used for the simulations are corresponding to De Bilt and (2) the horizontal direct and diffused solar radiation is not estimated properly.

The outdoor weather parameters between the De Bilt and Rotterdam (considered the closest location to TU Delft) are compared in order to analyse the differences between the temperature, wind, global solar radiation and relative humidity. It was found that the outdoor weather parameters from these two locations are very different, especially for the outside temperature (differences up to 10^{o}C). Therefore, it was decided to continue the multivariable analysis with the weather data obtained from Rotterdam (closest location to TU delft).

4.2 Diffused and direct solar radiation

The horizontal direct and diffused solar radiation is not available at KNMI, therefore a fraction of the global radiation is applied. The horizontal diffused radiation was assumed to be 15% of the horizontal global radiation and the horizontal direct radiation the other 85%. In order to check this assumption, the monthly average of the global, diffused and direct solar radiation on a horizontal plane for the collection of data ‘year 1964/1965’ and the ‘average from the years 1986 till 2005’ [1] was evaluated (see table below).

year	1964/1965					average 1986/2005				
	global	diffuse	direct	%diffuse	%direct	global	diffuse	direct	%diffuse	%direct
January	66	49	17	0.742	0.258	70	49	21	0.700	0.300
February	136	91	45	0.669	0.331	131	83	48	0.634	0.366
March	262	151	111	0.576	0.424	250	149	101	0.596	0.404
April	339	204	135	0.602	0.398	393	218	175	0.555	0.445
May	570	284	286	0.498	0.502	539	273	266	0.506	0.494
June	539	296	243	0.549	0.451	526	291	235	0.553	0.447
July	510	303	207	0.594	0.406	530	297	233	0.560	0.440
August	434	243	191	0.560	0.440	464	262	202	0.565	0.435
September	346	188	158	0.543	0.457	302	183	119	0.606	0.394
October	184	113	71	0.614	0.386	188	119	69	0.633	0.367
November	66	47	19	0.712	0.288	84	60	24	0.714	0.286
December	50	39	11	0.780	0.220	51	38	13	0.745	0.255

Table 18 Total global, diffuse and direct solar radiation on a horizontal plane [W/m^2]

It is observed that the fraction of diffuse and direct solar radiation is different than the initially assumed in LEA. A wrong estimation in the direct and diffused solar radiation causes high differences in the heating demand simulations, especially during midday. Figure 35 shows the hourly heating demand (ventilation no applied) at IO for: (1) the measurements, (2) simulations with weather data from De Bilt assuming solar diffused radiation a 15% of global solar radiation, (3) simulations with weather data from De Bilt assuming solar diffused radiation a variable % of global solar radiation (according to average 1986/2005) and (4) simulations with weather data from Rotterdam assuming solar diffused radiation a variable % of global solar radiation (according to average 1986/2005).

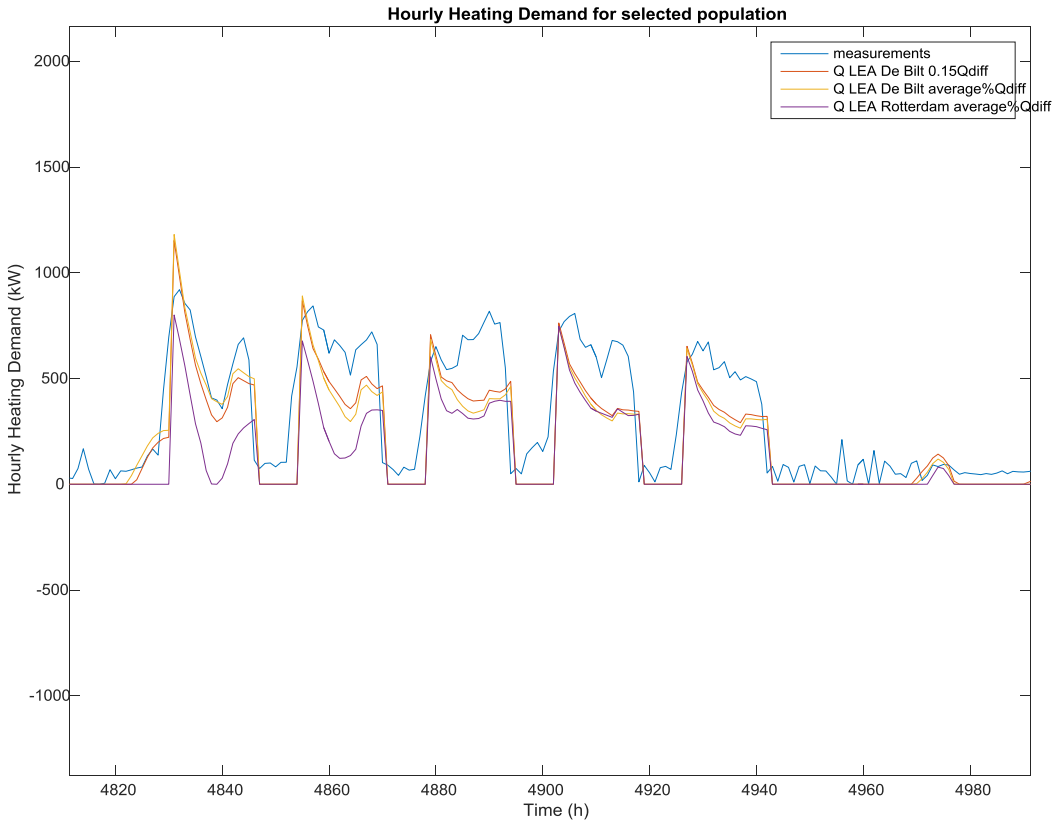


Figure 35 hourly heating demand (ventilation no applied) at IO for: (1) the measurements, (2) simulations with weather data from De Bilt assuming solar diffused radiation a 15% of global solar radiation, (3) simulations with weather data from De Bilt assuming solar diffused radiation a variable % of global solar radiation (according to average 1986/2005) and (4) simulations with weather data from Rotterdam assuming solar diffused radiation a variable % of global solar radiation (according to average 1986/2005).

According to this analysis, it was decided to correct the initial estimation of the monthly fraction of direct and diffused solar radiation by the estimation corresponding to the monthly average 1986/2005.

4.3 Future improvements

In more recent buildings, the proportion of heat loss due to thermal bridging is typically 10–15%. This can rise to 30% in better insulated low-energy buildings when insulation and construction details are not properly realized [41]. According to ISO 2007 [42], the heat losses through thermal bridges are calculated using the following equation.

$$P_{trans.bridge} = \varphi\, L(T_o - T_i)\ [WK]$$

Where *L* is the length of the thermal bridge (m) and φ the linear thermal transmittance of the thermal bridge W/(mK). The values of φ are between 0.02 and 1.

Appendix 5 Inputs of the studied buildings

This appendix collects the main inputs of the studied buildings (IO, TPM and 3mE) used by the simulator (physics-based software or LEA) to obtain the simulated heating demand. These inputs corresponds with the ones used in IPIN phase 2 during the second implementation of LEA (from 10th September 2016 until 25th November 2016).

The inputs are divided in the following 4 categories: (1) building characteristics, (2) use of the building, (3) installations and set points, and (4) coefficients (heat transfer coefficients and convective factors). Section 0 gives a description of some of the parameters. Sections 5.2, 5.3 and 5.4 give the corresponding input values for IO, TPM and 3mE, respectively.

5.1 Inputs´ description

Table 19 Description building characteristics used for the simulations (physics-based software)

Building characteristics		
Building dimensions		
gross surface	m²	Total area of the floor
fraction floor	-	This indicates the total surface of floor which is in contact with the internal air
gross height	m	high from floor (0) to floor (1) (it takes into account the floor thickness. Therefore, it is air+concrete)
ratio height	-	high from floor (0) to roof (0) (percentage of the real high. Therefore, space where there is air)
specific mass of the building	kg/m² bvo	calculated in function: bereken_infiltratievoud
surface facade	m²	total surface facade each orientation
fraction windows	-	(surface glass +frame)/wall area
Walls		
Rc-value	m².K/W	
Glass		
U-value	W/(m²·K)	U-value glass + frame
ZTA	-	Absolute zontoetredingfactor (absolute solar factor). coefficient of incident solar energy entering in the room through the glass.
LTA	-	absolute lichttoetredingfactor (absolute light factor).light transmission coefficient. Percentage of light entering in the room through the galss. LTA=1-light absorbed by glass-light reflected by glass.
Convective factor	-	convective factor of the windows. Indicates the percentage of solar energy that is absorbed by the air due to the convection.
Shading effects	-	This is an input parameter to calculate the LTA, ZTA and CF in case there is shading effects. If there are no shading effects, then the ZTA, LTA and CF is given above. 0= no shadow, 1=indoor blinds, 2=outdoor blinds.
indoor shading		
ZTA	-	coefficient of incident solar energy combined blinds+glass entering in the room.
LTA	-	light transmission coefficient combined blinds+glass. Percentage of light entering in the room. LTA=1-light absorbed by glass-light reflected by glass.
Convective factor	-	convective factor combined blinds+glass. Indicates the percentage of solar energy that is absorbed by the air due to the convection.
solar radiation/facade surface	W/m² facade surface	radiation values direct+diffused radiation
outdoor shading		
ZTA	-	coefficient of incident solar energy combined blinds+glass entering in the room.
LTA	-	light transmission coefficient combined blinds+glass. Percentage of light entering in the room. LTA=1-light absorbed by glass-light reflected by glass.
Convective factor	-	convective factor combined blinds+glass. Indicates the percentage of solar energy that is absorbed by the air due to the convection.
solar radiation/facade surface	W/m² facade surface	radiation values direct+diffused radiation

Table 20 Description installations & set points used for the simulations (physics-based software)

Installations and set points

set points indoor climate		
minimum temperature during opening hours	°C	
maximum temperature during opening hours	°C	
minimum temperature during closing hours	°C	
maximum temperature during closing hours	°C	
minimal RH during opening hours	%	
maximal RH during opening hours	%	
minimal RH during closing hours	%	
maximal RH during closing hours	%	
mechanical ventilation		number of times that the volume of the air of the building needs to be changed
ventilation during opening hours	-	
ventilation during closing hours	-	
ventilation during holidays period	-	
Infiltrations		infiltration causes due to the use of the building (eg. Opening windows, doors, etc). The infiltrations through the facade of the building is calculated in function of the wind speed.
inflitrations during opening hours	-	
infiltrations during closing hours	-	
infiltrations during holidays period	-	
Air Handling Unit (AHU)		
heat recovery site	-	0 = no present; 1 = present
type heat recovery	-	0 = no heat recovery; 1 = heat recovery with on-off control; 2 = heat recovery model
efficiency heat recovery	-	
presence reheating	-	0 = no present; 1 = present
fan site	-	0 = no present; 1 = present
pressure drop across the fan	Pa	
efficiency fan	-	
efficiency engine	-	
presence humidificator	-	0 = no present; 1 = present
efficiency humidificator	-	
outlet water temperature at outdoor temperature 0°C or lower	°C	
outlet water temperature at outdoor temperature 20°C or higher	°C	
Heating instalations		
heat exchanger	W	Maximum heating capacity of the heat exchanger of the building between the heat coming from the grid and heating systems of the building
AHU	W	maximum heating capacity, specified at a heating temperature range of 90-70C
radiators	W	maximum heating capacity, specified at a heating temperature range of 90-70C
floor heating	W	maximum heating capacity, specified at a heating temperature range of 90-70C
radiation panels	W	maximum heating capacity, specified at a heating temperature range of 90-70C
Working temperature limits building's heating systems		inlet temperature limitations for frost protection
Minimal inlet temperature for a outdoor temperature of 20C or higher	°C	
Minimal inlet temperature for a outdoor temperature of 0C or lower	°C	

Table 21 Description coefficients used for the simulations (physics-based software)

coefficients		
heat transfer coefficient between outdoor facade and outdoor air	W/m^2/K	αo: heat transfer coefficient of the floor/ceiling outdoors surface
heat transfer coefficient between indoor facade and outdoor air	W/m^2/K	αi: heat transfer coefficient of the floor/ceiling indoors surface
convective factor people	-	% of people's heat which is convection
convective factor lighting	-	% of lighting's heat which is convection
convective factor equipment	-	% of equipments's heat which is convection
reflective coefficient albedo	-	this is the coefficient of soil reflectance which is the amount of sun radiation reflected divided by the total amount of dominating radiation. it is used to calculate the reflection of the ground due to the sun radiation (function: grondreflectie).

5.2 Industrial Design (IO)

Table 22 Building characteristics used for the simulations (physics-based software) of IO

Building characteristics													
Facade orientations			**N**	**NE**	**E**	**SE**	**S**	**SW**	**W**	**NW**	**roof**	**floor to ground**	**floor to air**
Building dimensions													
gross surface	m^2	23300											
fraction floor	-	0.8											
gross height	m	3.6											
ratio height	-	0.7											
specific mass of the building	kg/m^2 bvo	250											
surface facade	m^2		1800	0	2460	0	1800	0	2460	0	7485	7485	0
fraction windows	-		0.3	0	0.3	0	0.3	0	0.3	0	0.4	0	0
Walls													
Rc-value	m^2.K/W		2.5	0	2.5	0	2.5	0	2.5	0	2.5	2.5	0
Glass													
U-value	W/(m^2·K)		2.5	0	2.5	0	2.5	0	2.5	0	2.1	2.1	0
ZTA	-		0.2	0	0.2	0	0.2	0	0.2	0.2	0.6	0.6	0
LTA	-		0.6	0	0.6	0	0.6	0	0.6	0	0.6	0.6	0
Convective factor	-		0.024	0	0.024	0	0.024	0	0.024	0	0.024	0.024	0
Shading effects	-		0	0	0	0	0	0	0	0	0	0	0
indoor shading													
ZTA	-	0.42											
LTA	-	0.21											
Convective factor	-	0.45											
solar radiation/facade surface	W/m^2 facade surface	150											
outdoor shading													
ZTA	-	0.19											
LTA	-	0.43											
Convective factor	-	0.05											
solar radiation/facade surface	W/m^2 facade surface	150											

Table 23 Coefficients used for the simulations (physics-based software) of IO

coefficients		
heat transfer coefficient between outdoor facade and outdoor air	W/m²/K	28.49
heat transfer coefficient between indoor facade and outdoor air	W/m²/K	3
convective factor people	-	0.7
convective factor lighting	-	0.7
convective factor equipment	-	0.7
reflective coefficient albedo	-	0.5

Table 24 Inputs related to the use of Industrial Design used for the simulations (physics-based software)

Use		
Hourly intervals	**PM**	
schedule		
opening time	PM	7
closing time	PM	22
days per week opened	-	5
internal heat gain (people)	m² bvo per person	50
Equipment		
internal heat gain (equipment)	W/m² bvo	0
lighting		
internal heat gain (lighting)	W/m² bvo	0

Table 25 Installations & set points used for the simulations (physics-based software) of Industrial Design

Installations and set points

set points indoor climate		
minimum temperature during opening hours	°C	20
maximum temperature during opening hours	°C	24
minimum temperature during closing hours	°C	17
maximum temperature during closing hours	°C	28
minimal RH during opening hours	%	0
maximal RH during opening hours	%	100
minimal RH during closing hours	%	0
maximal RH during closing hours	%	100
mechanical ventilation		
ventilation during opening hours	-	2
ventilation during closing hours	-	0.5
ventilation during holidays period	-	0.5
Infiltrations		
inflitrations during opening hours	-	0.1
infiltrations during closing hours	-	0
infiltrations during holidays period	-	0
Air Handling Unit (AHU)		
heat recovery site	-	1
type heat recovery	-	1
efficiency heat recovery	-	0.3
presence reheating	-	0
fan site	-	1
pressure drop across the fan	Pa	1200
efficiency fan	-	0.7
efficiency engine	-	0.85
presence humidificator	-	0
efficiency humidificator	-	0.65
outlet water temperature at outdoor temperature 0°C or lower	°C	19
outlet water temperature at outdoor temperature 20°C or higher	°C	16
Heating instalations		
heat exchanger	W	3514080
AHU	W	787467
radiators	W	1510750
floor heating	W	0
radiation panels	W	0
Working temperature limits building's heating systems		
Minimal inlet temperature for a outdoor temperature of 20C or higher	°C	40
Minimal inlet temperature for a outdoor temperature of 0C or lower	°C	20

5.3 Technology, Policy and Management (TPM)

Table 26 Building characteristics used for the simulations (physics-based software) of TPM

Building characteristics												
Facade orientations			**N**	**NE**	**E**	**SE**	**S**	**SW**	**W**	**NW**	**roof**	**floor to ground**	**floor to air**
Building dimensions													
gross surface	m²	12000											
fraction floor	-	0.7											
gross height	m	3.5											
ratio height	-	0.7											
specific mass of the building	kg/m² bvo	600											
surface facade	m²		756	0	42	0	323	0	77	0	897.12	656.3	288
fraction windows	-		0.25	0	0.25	0	0.25	0	0.25	0	0	0	0
Walls													
Rc-value	m².K/W		4	0	4	0	4	0	4	0	4	4	4
Glass													
U-value	W/(m²·K)		3	0	3	0	3	0	3	0	2	2	2
ZTA	-		0.2	0	0.2	0	0.2	0	0.2	0	0.4	0.4	0.4
LTA	-		0.6	0	0.6	0	0.6	0	0.6	0	0.6	0.6	0.6
Convective factor	-		0.1	0	0.1	0	0.1	0	0.1	0	0.1	0.1	0.1
Shading effects	-		0	0	0	0	0	0	0	0	0	0	0
indoor shading													
ZTA	-	0.4											
LTA	-	0.2											
Convective factor	-	0.4											
solar radiation/facade surface	W/m² facade surface	300											
outdoor shading													
ZTA	-	0.2											
LTA	-	0.2											
Convective factor	-	0.2											
solar radiation/facade surface	W/m² facade surface	300											

Table 27 Coefficients used for the simulations (physics-based software) of TPM

coefficients		
heat transfer coefficient between outdoor facade and outdoor air	W/m²/K	28.49
heat transfer coefficient between indoor facade and outdoor air	W/m²/K	8.29
convective factor people	-	0.7
convective factor lighting	-	0.7
convective factor equipment	-	0.7
reflective coefficient albedo	-	0.5

Table 28 Inputs related to the use of TPM for the simulations (physics-based software)

Use		
Hourly intervals	**PM**	
schedule		
opening time	PM	7
closing time	PM	22
days per week opened	-	5
internal heat gain (people)	m² floor surface per person	30
Equipment		
internal heat gain (equipment)	W/m² floor surface	0
lighting		
internal heat gain (lighting)	W/m² floor surface	0

Table 29 Installations & set points used for the simulations (physics-based software) of TPM

Installations and set points

set points indoor climate		
minimum temperature during opening hours	°C	21
maximum temperature during opening hours	°C	24
minimum temperature during closing hours	°C	15
maximum temperature during closing hours	°C	28
minimal RH during opening hours	%	0
maximal RH during opening hours	%	100
minimal RH during closing hours	%	0
maximal RH during closing hours	%	100
mechanical ventilation		
ventilation during opening hours	-	3
ventilation during closing hours	-	1.5
ventilation during holidays period	-	0
Infiltrations		
inflitrations during opening hours	-	0.3
infiltrations during closing hours	-	0
infiltrations during holidays period	-	0
Air Handling Unit (AHU)		
heat recovery site	-	1
type heat recovery	-	1
efficiency heat recovery	-	0.3
presence reheating	-	0
fan site	-	1
pressure drop across the fan	Pa	1200
efficiency fan	-	0.7
efficiency engine	-	0.85
presence humidificator	-	0
efficiency humidificator	-	0.65
outlet water temperature at outdoor temperature 0°C or lower	°C	18
outlet water temperature at outdoor temperature 20°C or higher	°C	16
Heating instalations		
heat exchanger	W	1460000
AHU	W	375687
radiators	W	542502
floor heating	W	0
radiation panels	W	0
Working temperature limits building's heating systems		
Minimal inlet temperature for a outdoor temperature of 20C or higher	°C	40
Minimal inlet temperature for a outdoor temperature of 0C or lower	°C	20

5.4 Mechanical, Maritime and Materials Engineering (3mE)

Table 30 Building characteristics used for the simulations (physics-based software) of 3mE

Building characteristics														
Facade orientations			**N**	**NE**	**E**	**SE**	**S**	**SW**	**W**	**NW**	**roof**	**floor to ground**	**floor to air**
Building dimensions														
gross surface	m²	36402												
fraction floor	-	0.7												
gross height	m	3.6												
ratio height	-	0.7												
specific mass of the building	kg/m² bvo	300												
surface facade	m²		0	4799	0	10969	0	4799	0	10969	17529	14529	0	
fraction windows	-		0	0.4	0	0.4	0	0.4	0	0.4	0.1	0	0	
Walls														
Rc-value	m².K/W		0	2.5	0	2.5	0	2.5	0	2.5	2.5	2.5	0	
Glass														
U-value	W/(m²·K)		0	2.2	0	2.2	0	2.2	0	2.2	2.2	2.2	0	
ZTA	-		0	0.1	0	0.1	0	0.1	0	0.1	0.6	0.6	0	
LTA	-		0	0.6	0	0.6	0	0.6	0	0.6	0.6	0.6	0	
Convective factor	-		0	0.024	0	0.024	0	0.024	0	0.024	0.024	0.024	0	
Shading effects	-		0	0	0	0	0	0	0	0	0	0	0	
indoor shading														
ZTA	-	0.42												
LTA	-	0.21												
Convective factor	-	0.45												
solar radiation/facade surface	W/m² facade surface	150												
outdoor shading														
ZTA	-	0.19												
LTA	-	0.43												
Convective factor	-	0.05												
solar radiation/facade surface	W/m² facade surface	150												

Table 31 Coefficients used for the simulations (physics-based software) of 3mE

coefficients		
heat transfer coefficient between outdoor facade and outdoor air	W/m²/K	28.49
heat transfer coefficient between indoor facade and outdoor air	W/m²/K	8.29
convective factor people	-	0.7
convective factor lighting	-	0.7
convective factor equipment	-	0.7
reflective coefficient albedo	-	0.5

Table 32 Inputs related to the use of 3mE for the simulations (physics-based software)

Use		
Hourly intervals	**PM**	
schedule		
opening time	PM	7
closing time	PM	22
days per week opened	-	5
internal heat gain (people)	m² floor surface per person	50
Equipment		
internal heat gain (equipment)	W/m² floor surface	0
lighting		
internal heat gain (lighting)	W/m² floor surface	8

Table 33 Installations & set points used for the simulations (physics-based software) of 3mE

Installations and set points

set points indoor climate		
minimum temperature during opening hours	°C	21
maximum temperature during opening hours	°C	24
minimum temperature during closing hours	°C	15
maximum temperature during closing hours	°C	28
minimal RH during opening hours	%	0
maximal RH during opening hours	%	100
minimal RH during closing hours	%	0
maximal RH during closing hours	%	100
mechanical ventilation		
ventilation during opening hours	-	2
ventilation during closing hours	-	1.5
ventilation during holidays period	-	2
Infiltrations		
inflitrations during opening hours	-	0
infiltrations during closing hours	-	0
infiltrations during holidays period	-	0
Air Handling Unit (AHU)		
heat recovery site	-	1
type heat recovery	-	1
efficiency heat recovery	-	0.2
presence reheating	-	0
fan site	-	1
pressure drop across the fan	Pa	1200
efficiency fan	-	0.7
efficiency engine	-	0.85
presence humidificator	-	0
efficiency humidificator	-	0.65
outlet water temperature at outdoor temperature 0°C or lower	°C	18
outlet water temperature at outdoor temperature 20°C or higher	°C	16
Heating instalations		
heat exchanger	W	4044000
AHU	W	1076540
radiators	W	2661900
floor heating	W	37000
radiation panels	W	0
Working temperature limits building's heating systems		
Minimal inlet temperature for a outdoor temperature of 20C or higher	°C	40
Minimal inlet temperature for a outdoor temperature of 0C or lower	°C	20

Appendix 6 Sensitivity Analysis of Simulator

6.1 Hourly heating demand

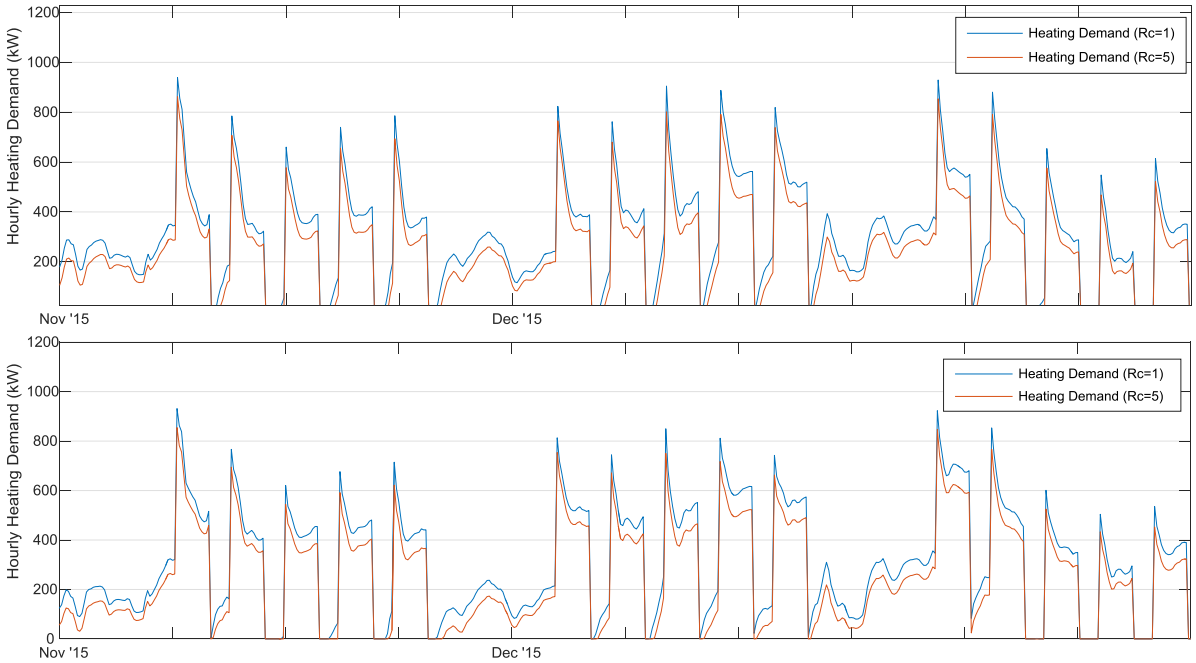


Figure 36 Simulated hourly heating demand profile for IO with façade insulation R-value=1 W/(m^{2}K) and R-value=5 W/(m^{2}K); (above) for a thermal mass of 600 kg/m^2 and (below) for a thermal mass of 250 kg/m^2. Weather data from 21st November – 11th December 2015.

6.2 Indoor versus outdoor temperature

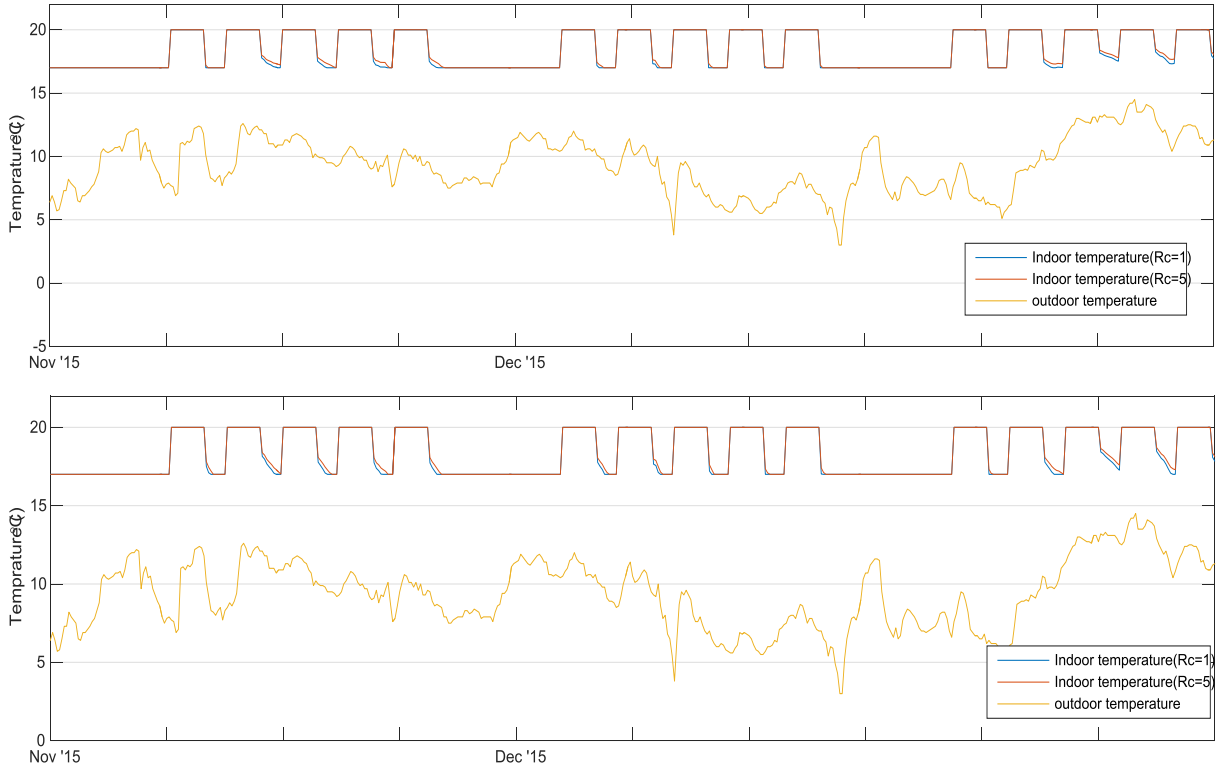


Figure 37 Outdoor temperature and simulated indoor temperature profile for IO with façade insulation R-value=1 W/(m^{2}K) and R-value=5 W/(m^{2}K); (above) for a thermal mass of 600 kg/m^2 and (below) for a thermal mass of 250 kg/m^2. Weather data from 21st November – 11th December 2015.

6.3 Outdoor temperature

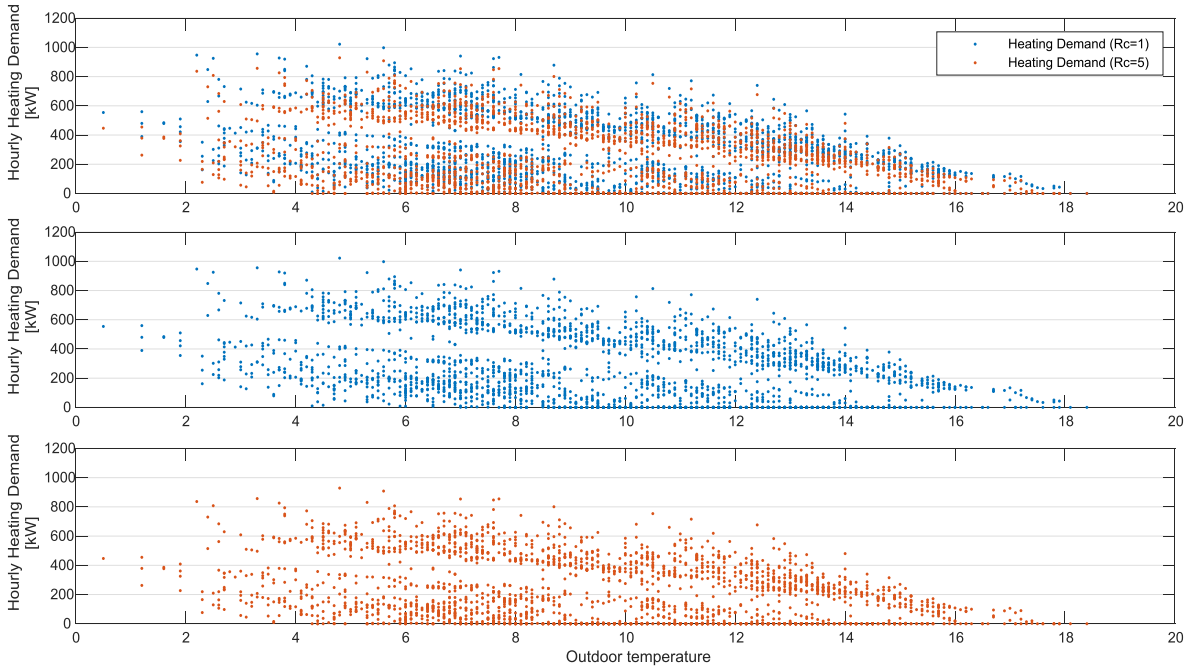


Figure 39 Influence of the outdoor temperature on the heating demand for IO with a thermal mass of 600 kg/m^2; (above) façade insulation R-value=1 W/(m^{2}K) and R-value=5 W/(m^{2}K), (middle) R-value=1 W/(m^{2}K) and (below) R-value=5 W/(m^{2}K). Weather data from 5th October 2015 - 14th January 2016.

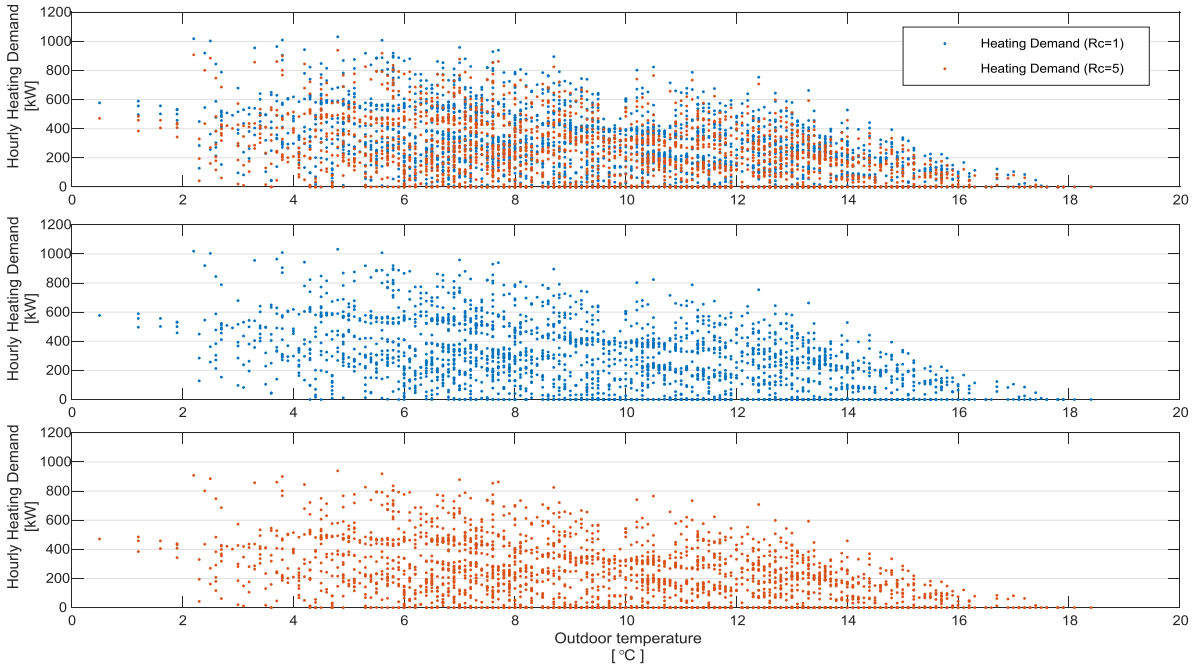


Figure 38 Influence of the outdoor temperature on the heating demand for IO with a thermal mass of 250 kg/m^2; (above) façade insulation R-value=1 W/(m^{2}K) and R-value=5 W/(m^{2}K), (middle) R-value=1 W/(m^{2}K) and (below) R-value=5 W/(m^{2}K). Weather data from 5th October 2015 - 14th January 2016.

6.4 Direct solar radiation

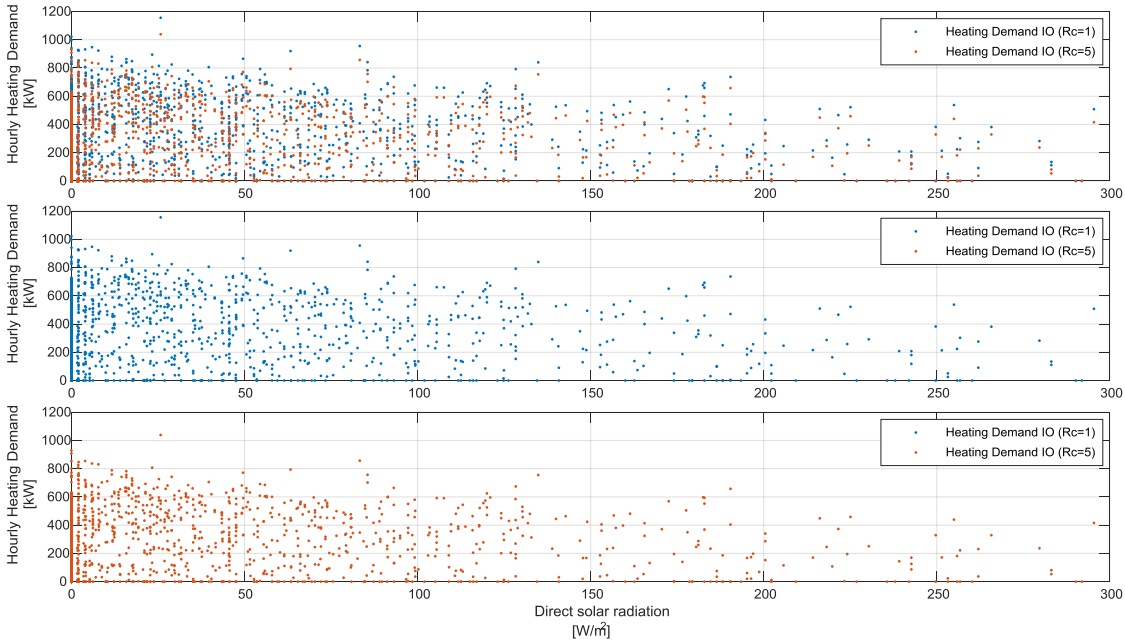


Figure 41 Influence of the direct solar radiation on the heating demand for IO with a thermal mass of 600 kg/m²; (above) façade insulation R-value=1 W/(m²K) and R-value=5 W/(m²K), (middle) R-value=1 W/(m²K) and (below) R-value=5 W/(m²K). Weather data from 5th October 2015 - 14th January 2016.

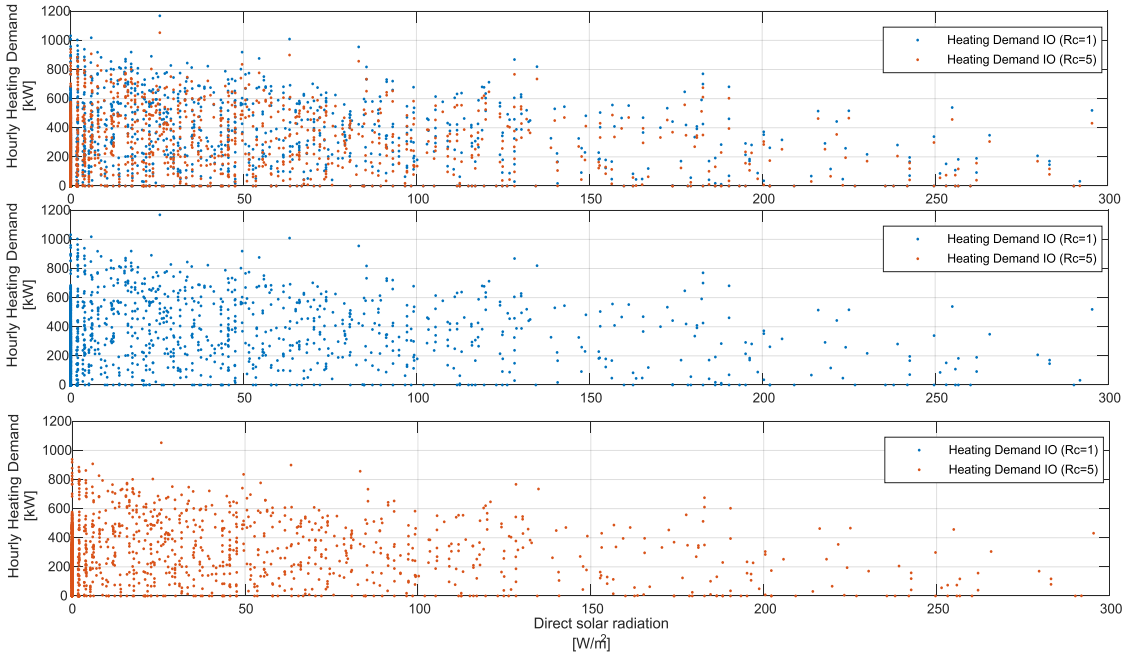


Figure 40 Influence of the direct solar radiation on the heating demand for IO with a thermal mass of 250 kg/m²; (above) façade insulation R-value=1 W/(m²K) and R-value=5 W/(m²K), (middle) R-value=1 W/(m²K) and (below) R-value=5 W/(m²K). Weather data from 5th October 2015 - 14th January 2016.

6.5 Diffused solar radiation

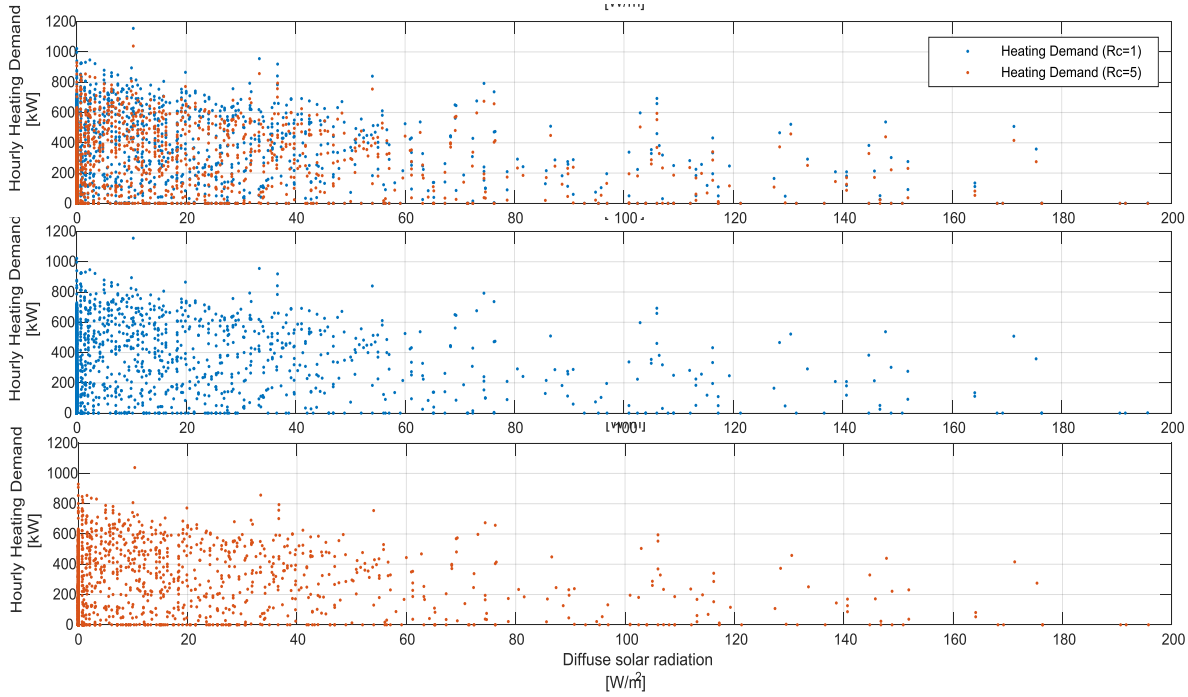


Figure 43 Influence of the diffused solar radiation on the heating demand for IO with a thermal mass of 600 kg/m²; (above) façade insulation R-value=1 W/(m²K) and R-value=5 W/(m²K), (middle) R-value=1 W/(m²K) and (below) R-value=5 W/(m²K). Weather data from 5th October 2015 - 14th January 2016.

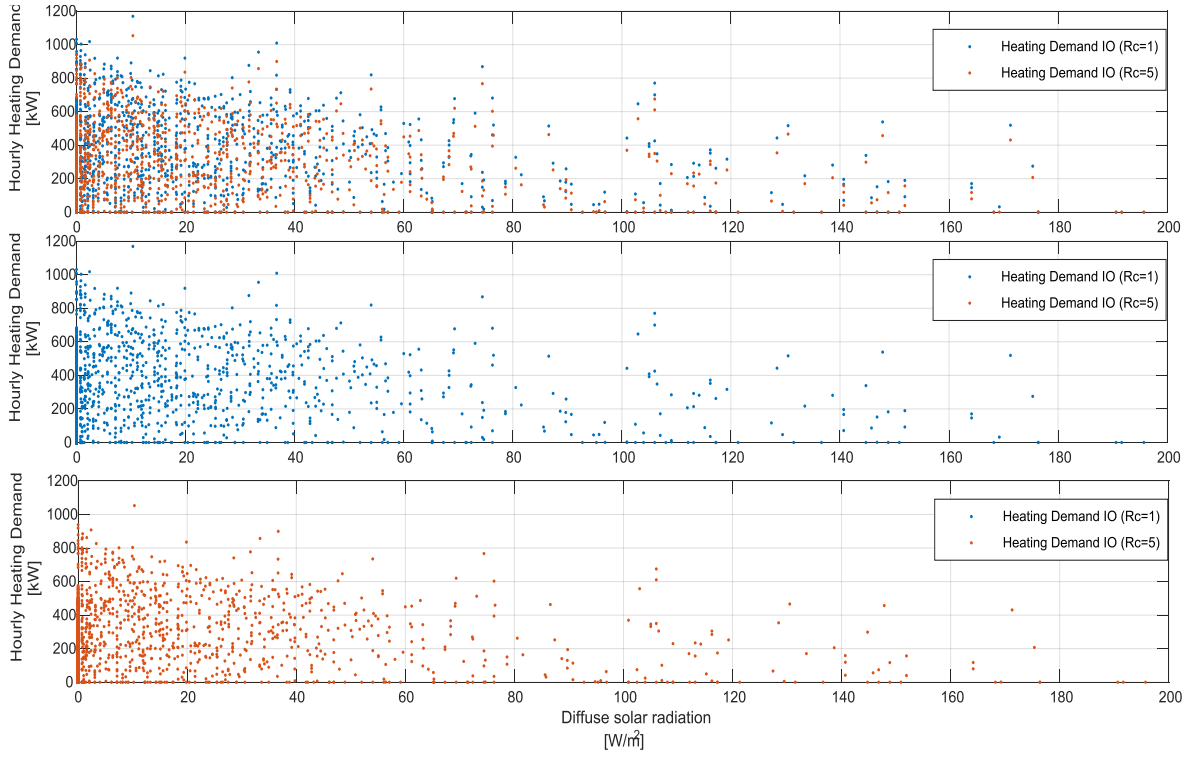


Figure 42 Influence of the diffused solar radiation on the heating demand for IO with a thermal mass of 250 kg/m²; (above) façade insulation R-value=1 W/(m²K) and R-value=5 W/(m²K), (middle) R-value=1 W/(m²K) and (below) R-value=5 W/(m²K). Weather data from 5th October 2015 - 14th January 2016.

Appendix 7 Hourly profile heating demand and other parameters for data set 2015

This appendix presents the hourly profile of the most influencing parameters during the measurement period of 2015 (data set 2015). The selected period of time is from 1th October (hour: 6553) until 31th December 2015 (hour: 8760). During this period, only heating is required. The representative week selected is 12th – 18th October 2015.

7.1 Hourly heating demand

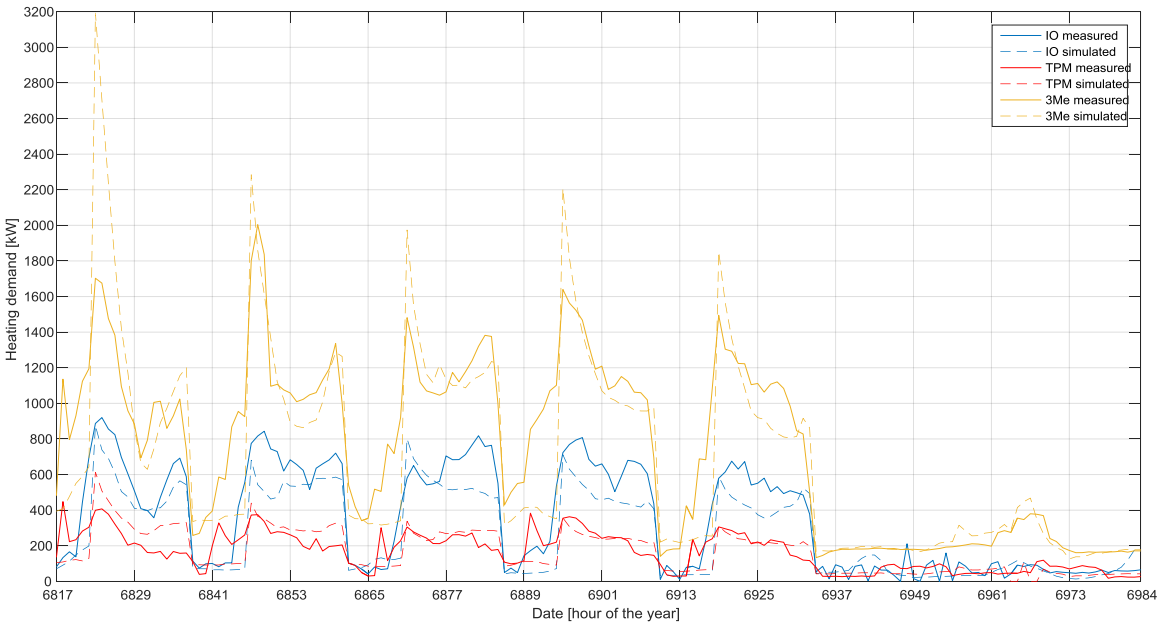


Figure 44 Measured & simulated hourly heating demand profile for IO, TPM and 3Me for the week 12th – 18th October 2015

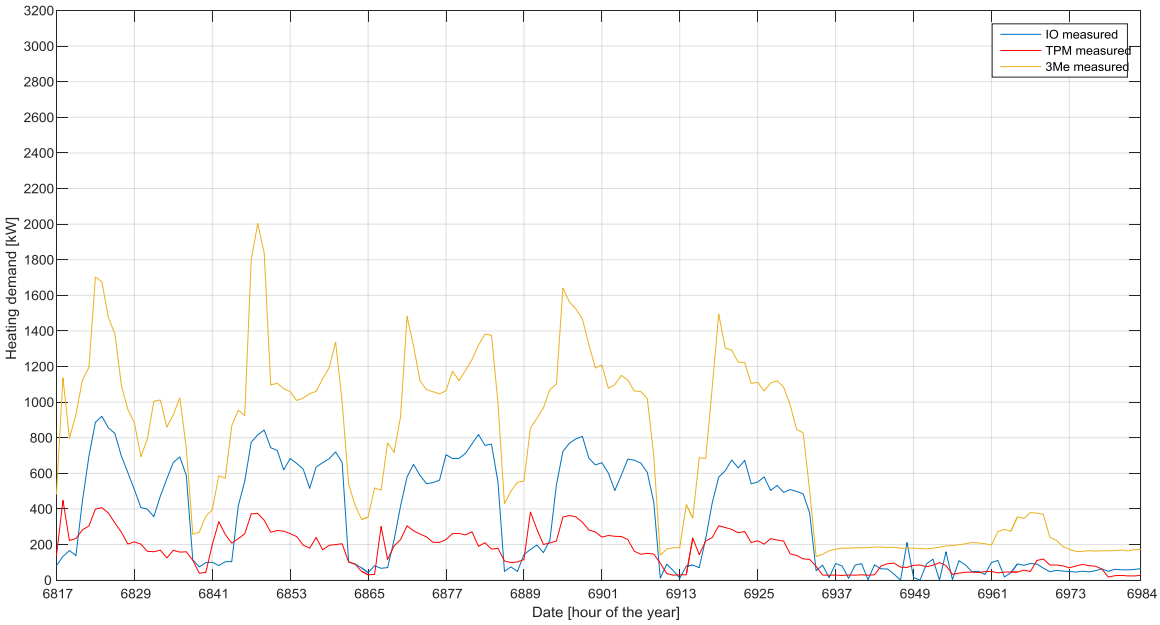


Figure 45 Measured hourly heating demand profile for IO, TPM and 3Me for the week 12th – 18th October 2015

7.2 Outdoor temperature

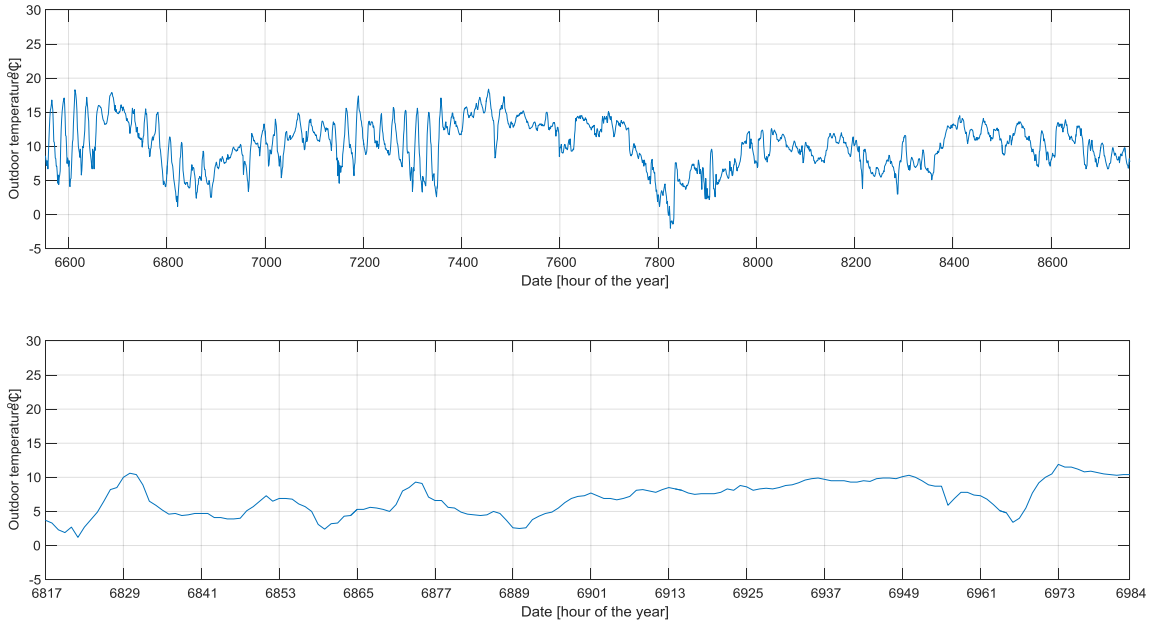


Figure 46 Outdoor temperature profile for 1th October – 31th December 2015 (above) and 12th – 18th October 2015 (below)

7.3 Global horizontal solar radiation

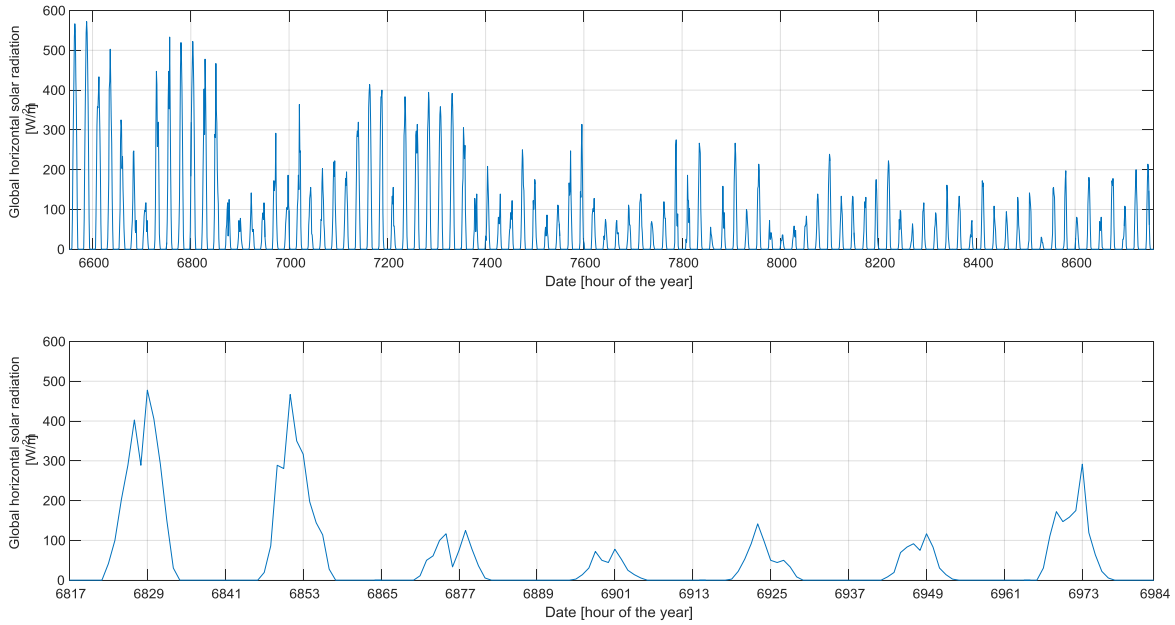


Figure 47 Hourly global horizontal solar radiation profile for 1th October – 31th December 2015 (above) and 12th – 18th October 2015 (below).

7.4 Wind speed

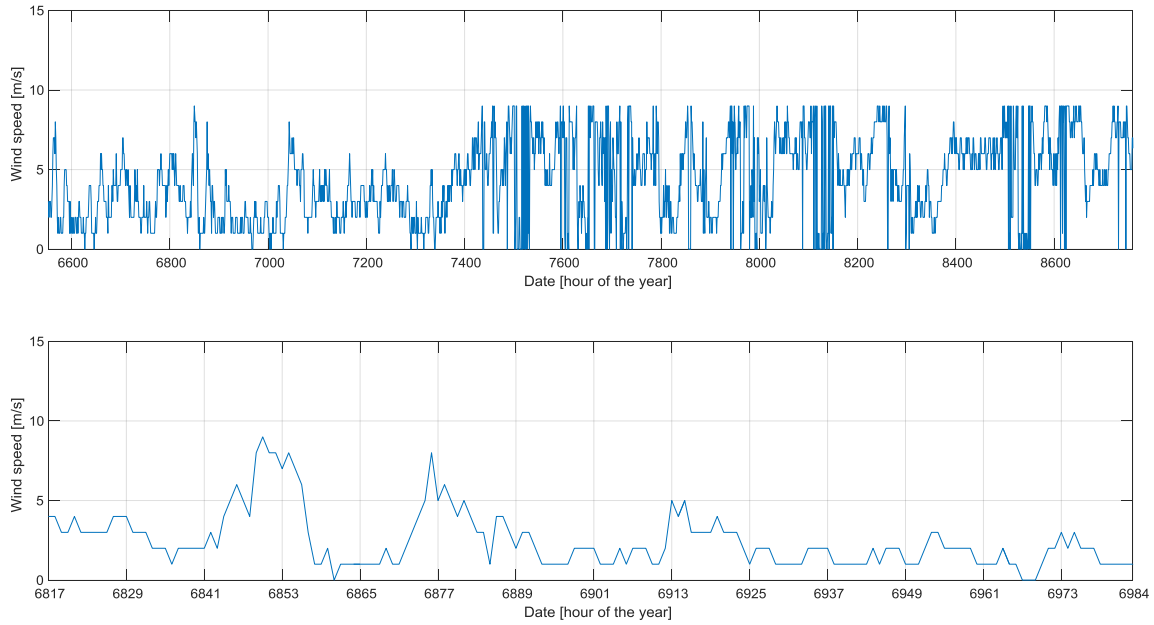


Figure 48 Hourly wind speed profile for 1th October – 31th December 2015 (above) and 12th – 18th October 2015 (below)

7.5 Indoor air temperature

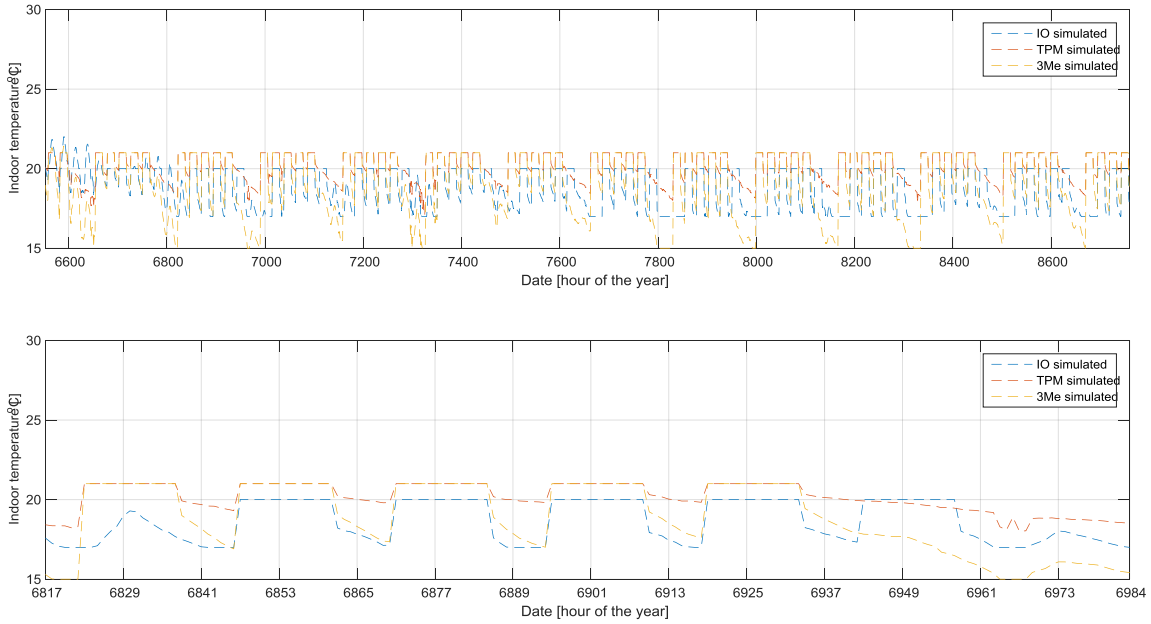


Figure 49 Hourly indoor air temperature profile for 1th October – 31th December 2015 (above) and 12th – 18th October 2015 (below)

7.6 Specific internal heat gain

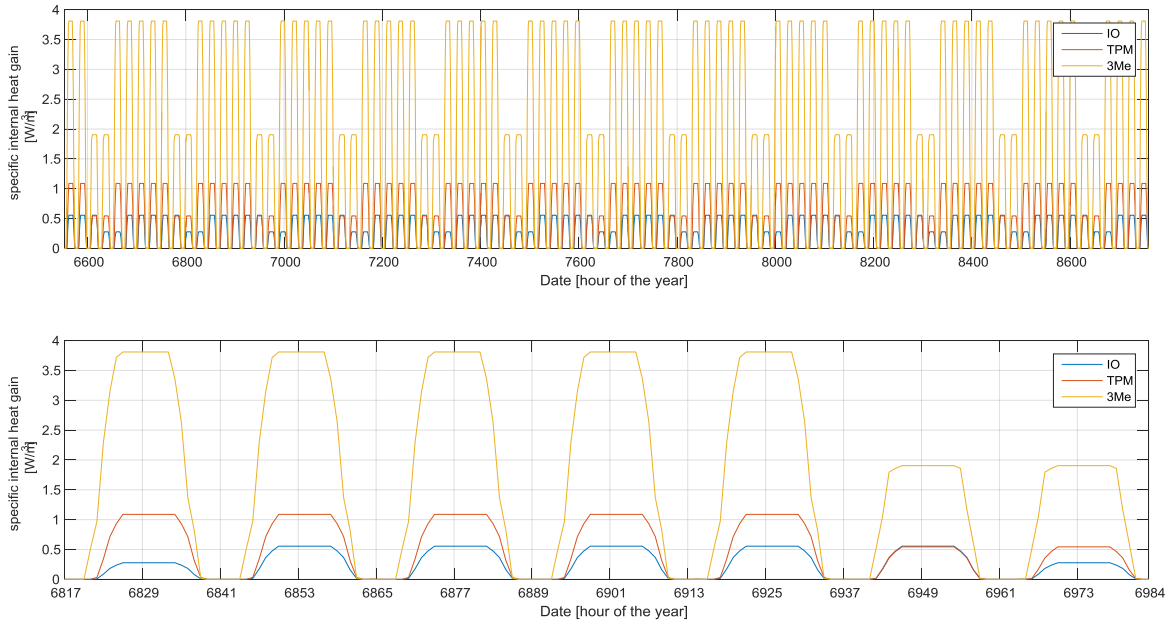


Figure 50 Hourly specific internal heat gain profile for 1th October – 31th December 2015 (above) and 12th – 18th October 2015 (below).

Appendix 8 Hourly profile heating demand and other parameters for data set 2016

This appendix presents the hourly profile of the most influencing parameters during the last implementation period (from 10th September 2016 until 25th November 2016).

8.1 Hourly heating demand

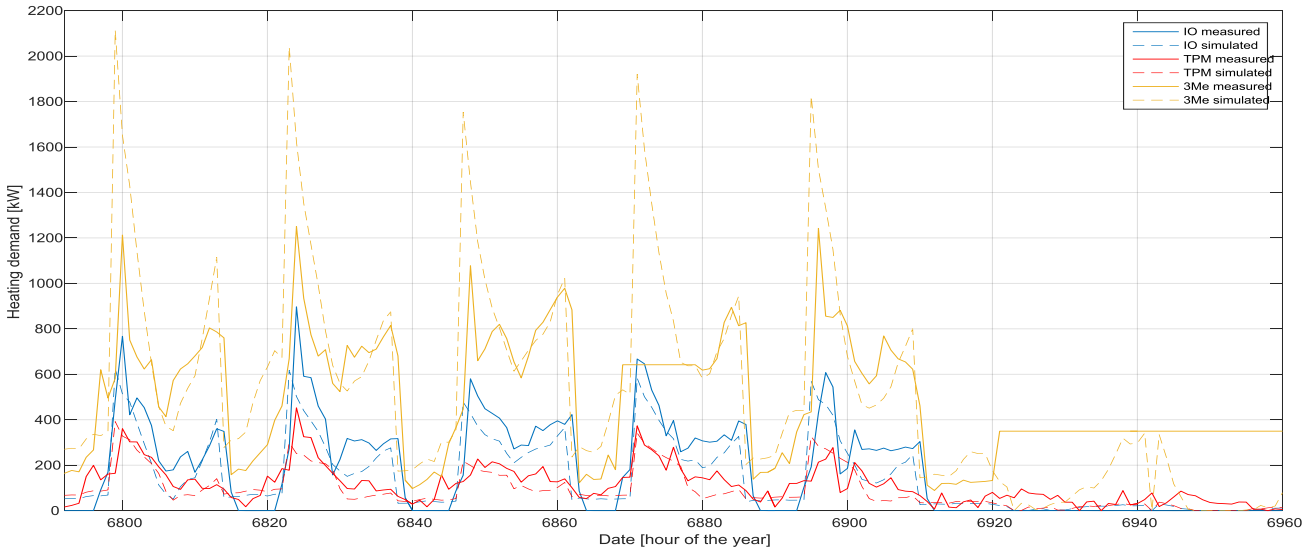


Figure 51 Measured & simulated hourly heating demand profile for IO, TPM and 3Me from for the period 10th – 16th October 2016

Figure 52 Measured & simulated heating demand TPM (above), IO (middle), and 3mE (below). Week period: 10th – 16th October 2016

8.2 Outdoor temperature

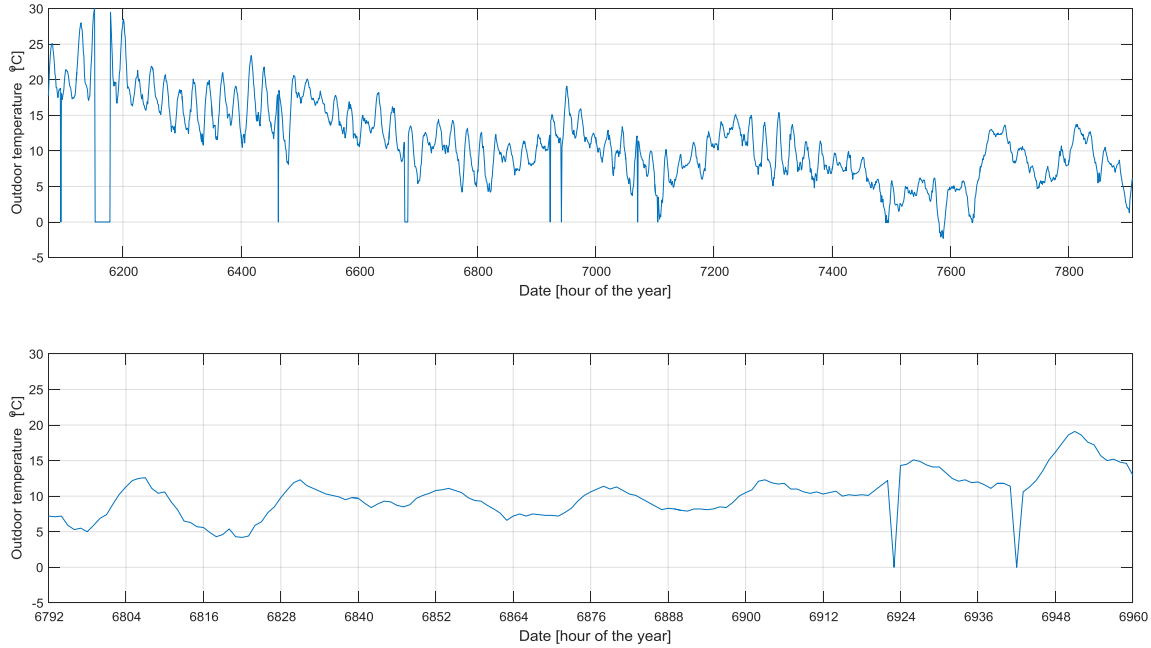


Figure 53 Outdoor temperature profile from 10th September – 25th November ’16 (above) and 10th – 16th October ’16 (below)

8.3 Global horizontal solar radiation

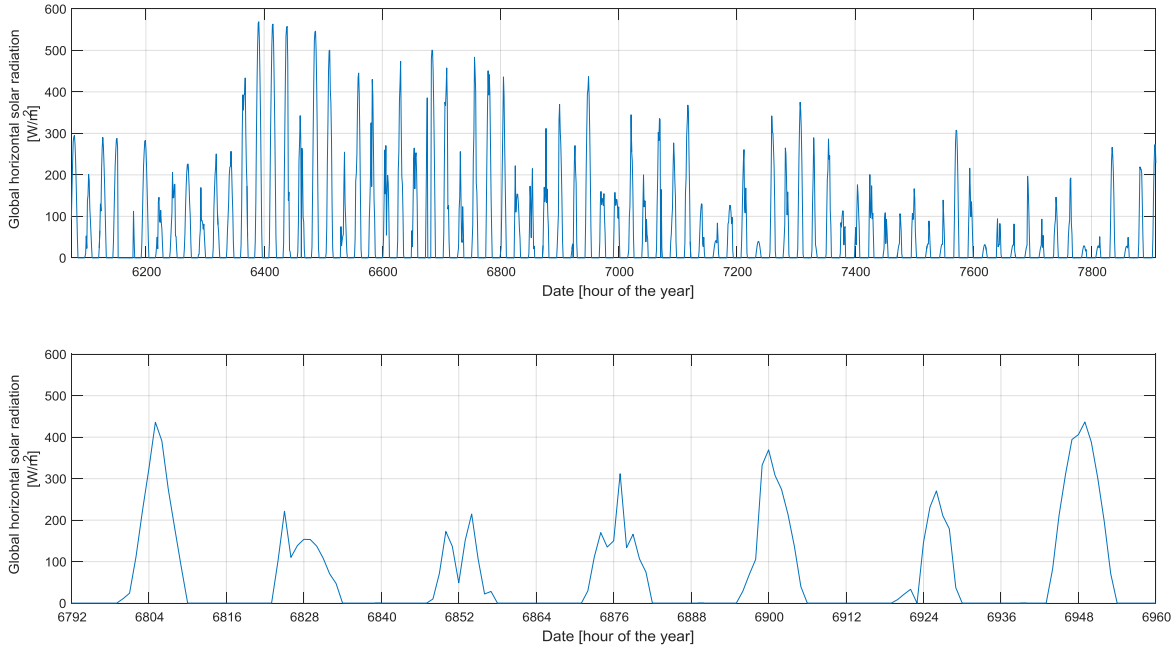


Figure 54 Global horizontal solar radiation profile from 10th September – 25th November ’16 (above) and 10th – 16th October ’16 (below)

8.4 Wind speed

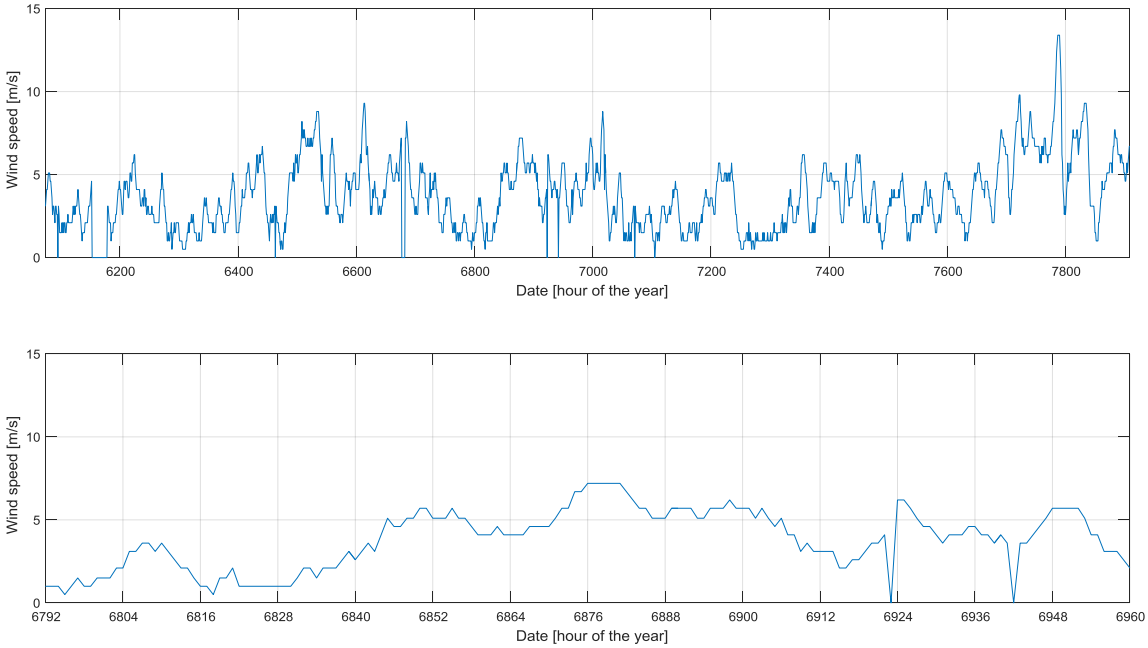


Figure 55 Wind speed profile from 10th September – 25th November '16 (above) and 10th – 16th October '16 (below)

8.5 Indoor air temperature

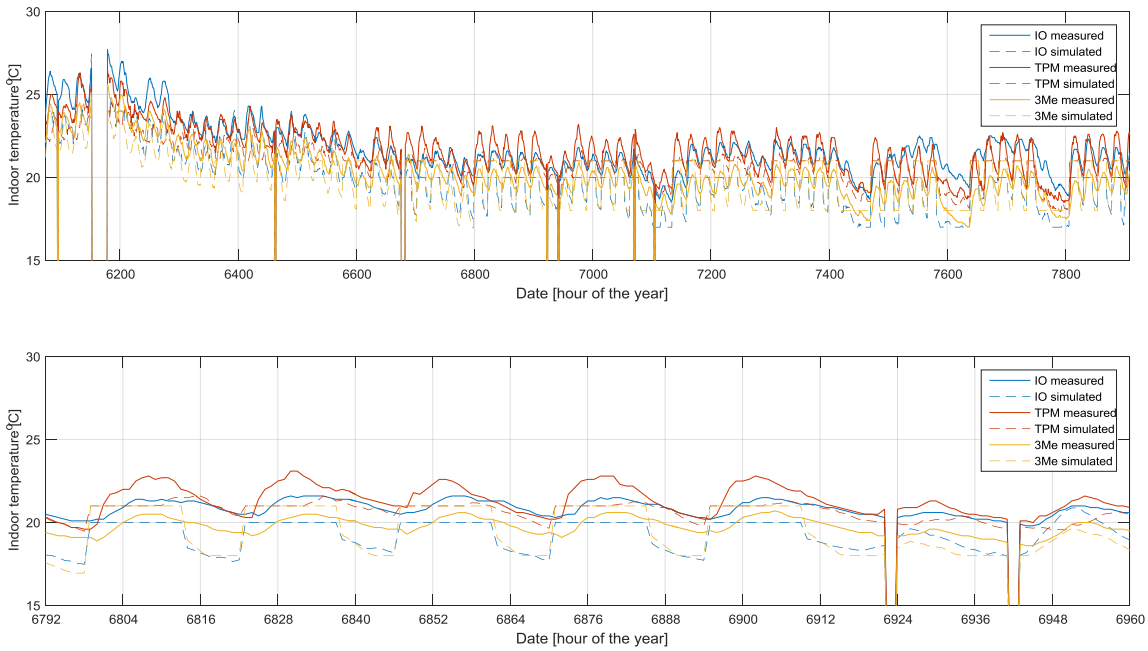


Figure 56 Measured & simulated indoor air temperature profile for IO, TPM and 3Me from 10th September – 25th November '16 (above) and 10th – 16th October '16 (below)

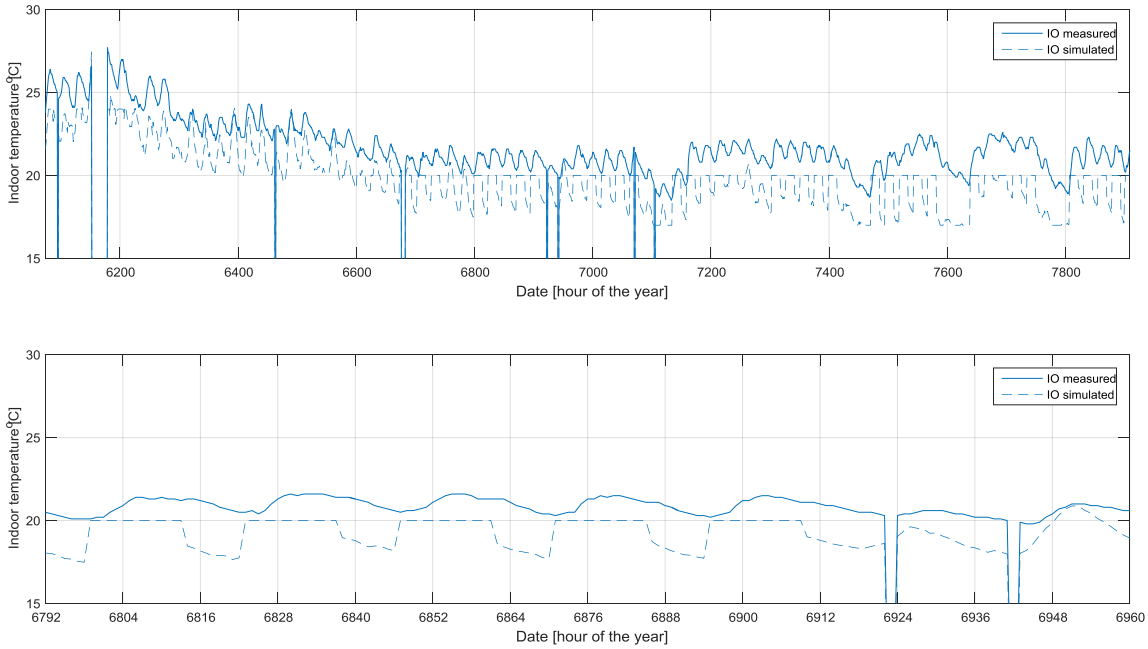


Figure 57 Measured & simulated indoor air temperature profile for IO from 10th September – 25th November '16 (above) and 10th – 16th October '16 (below)

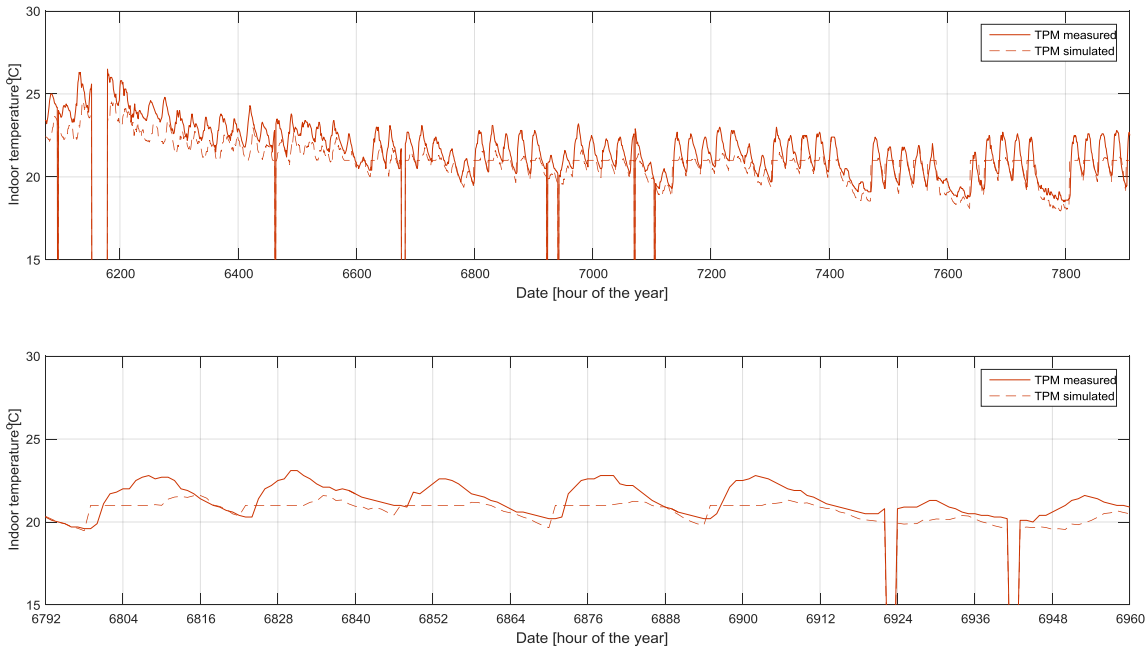


Figure 58 Measured & simulated indoor air temperature profile for TPM from 10th September – 25th November '16 (above) and 10th – 16th October '16 (below)

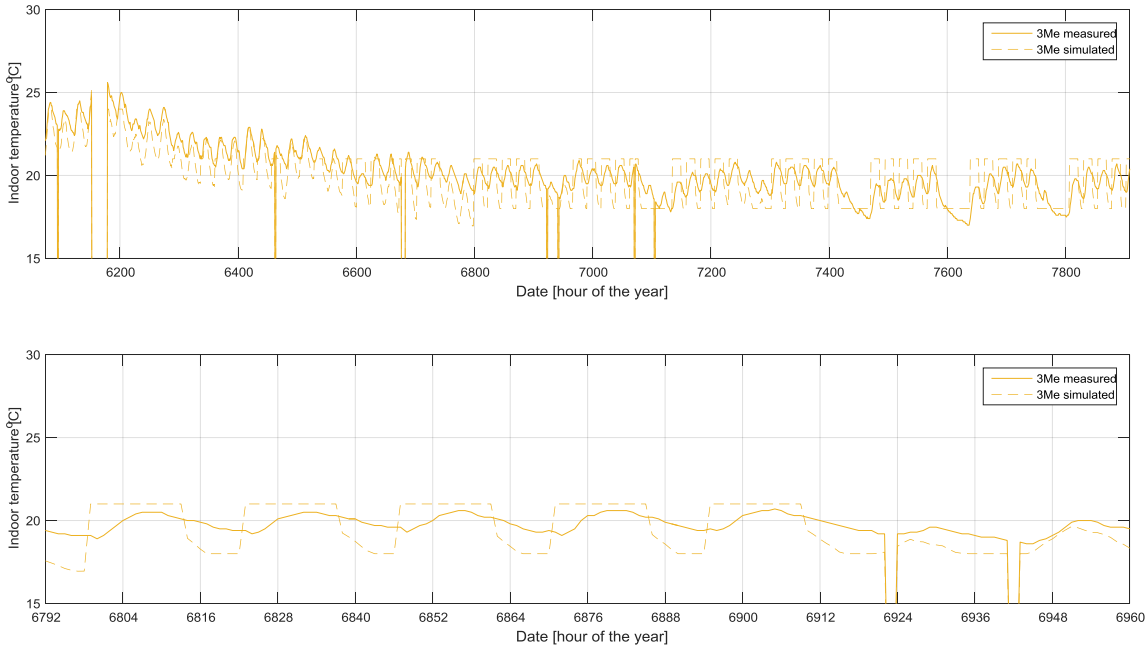


Figure 59 Measured & simulated indoor air temperature profile for 3Me from 10th September – 25th November '16 (above) and 10th – 16th October '16 (below)

8.6 Specific internal heat gain

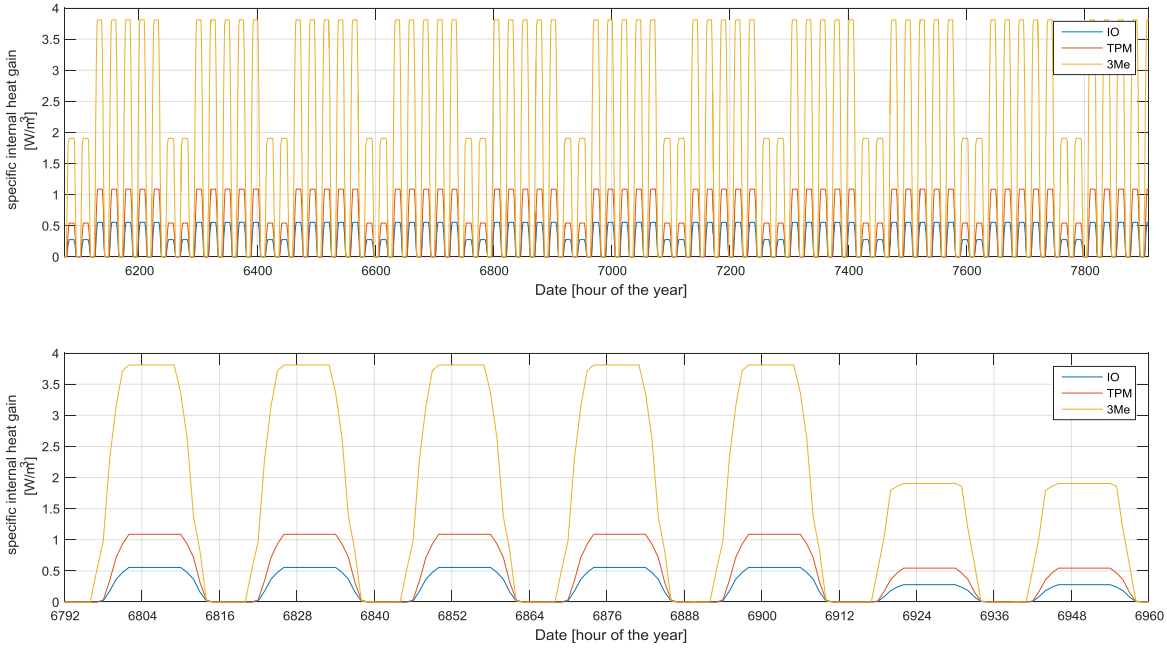


Figure 60 Estimated specific internal heat gain profile for IO, TPM and 3Me from 10th September – 25th November '16 (above) and 10th – 16th October '16 (below)

Appendix 9 Impact of different parameters on the building heating demand of data set 2015

This appendix shows the impact of the most influencing parameters on the simulated and actual heating demand of the buildings for October 2015 (part of data set 2015). The graphs include the data points and least square for each of the buildings.

The period selected is after 1st of October 2015 which corresponds to the period in which only heating was required. The actual heating demand for IO and TPM is available from 1st October 2015 until 13th January 2016, while for 3mE only the month of October 2015. The least square for the month of October is representative of the period from October until January, therefore only the month of October is represented.

9.1 Outdoor temperature

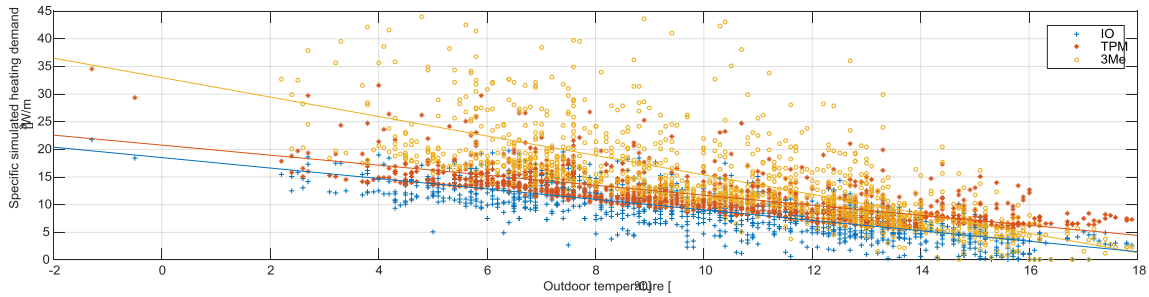


Figure 61 Influence outdoor temperature on the buildings' simulated heating demand during weekdays and opening hours for the period from 1st October until 13th January 2016.

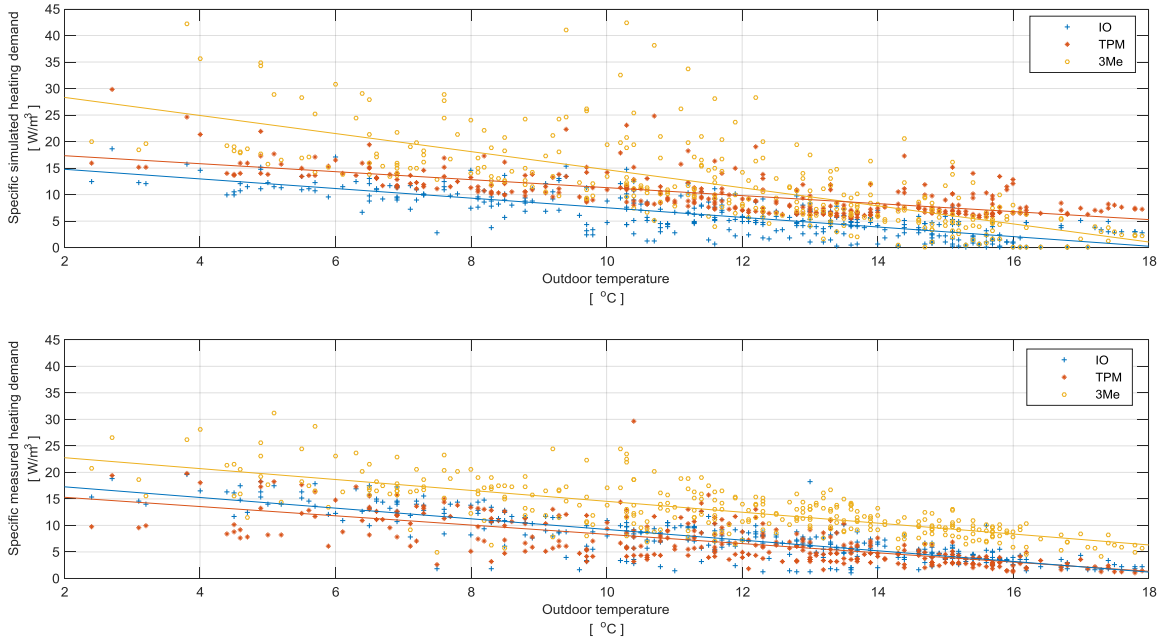


Figure 62 Influence outdoor temperature on the buildings' simulated (above) and actual (below) heating demand during weekdays and opening hours for October 2015.

Table 34 Linear correlation outdoor temperature versus simulated and actual heating demand during weekdays and opening hours for October 2015.

	simulated		measured	
	constant	slope	constant	slope
IO	17	-0.91	19	-1
3mE	32	-1.7	25	-1
TPM	19	-0.75	17	-0.87

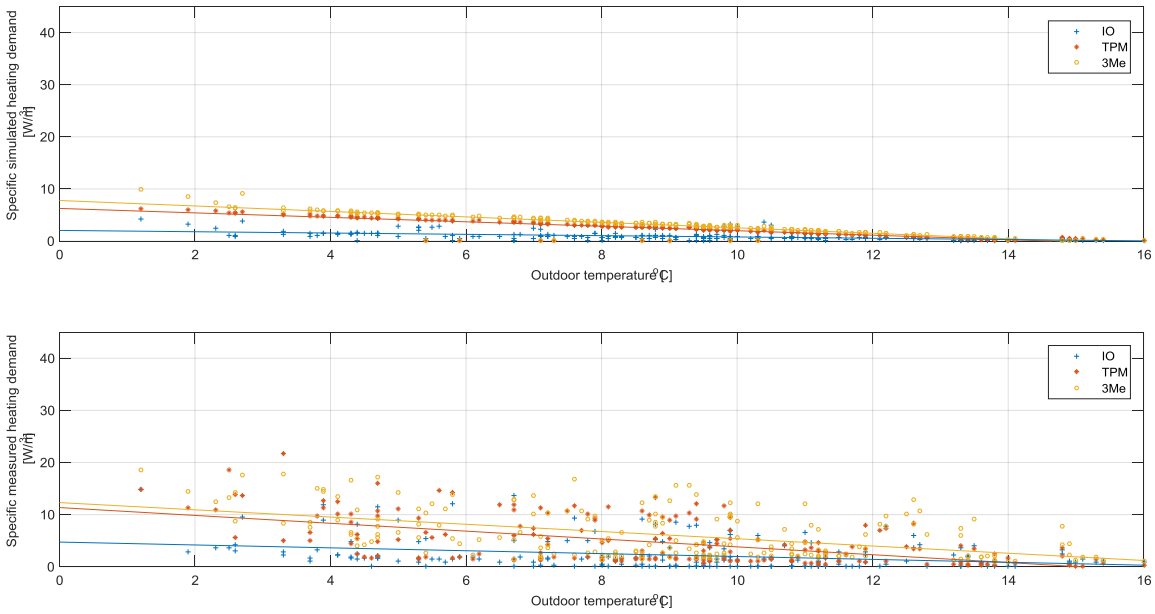


Figure 63 Influence outdoor temperature on the buildings' simulated (above) and actual (below) heating demand during weekdays and closing hours for October 2015.

Table 35 Linear correlation outdoor temperature versus simulated and actual heating demand during weekdays and closing hours for October 2015.

	simulated		measured	
	constant	slope	constant	slope
IO	2	-0.13	4.7	-0.28
3mE	7.8	-0.52	12	-0.69
TPM	6.3	-0.43	11	-0.75

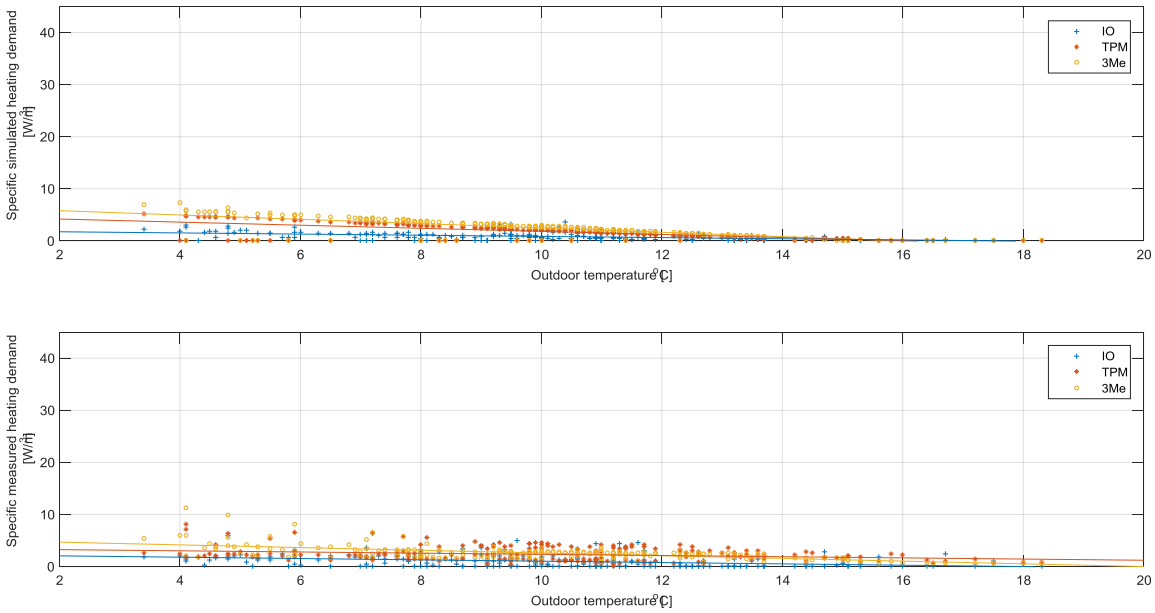


Figure 64 Influence outdoor temperature on the buildings' simulated (above) and actual (below) heating demand during weekends (opening & closing hours) for October 2015.

Table 36 Linear correlation outdoor temperature versus simulated and actual heating demand during weekends (opening & closing hours) for October 2015.

	simulated		**measured**	
	constant	slope	constant	slope
IO	2	-0.11	2.3	-0.13
3mE	6.7	-0.41	5.2	-0.26
TPM	4.8	-0.3	3.5	-0.11

9.2 Global horizontal solar radiation

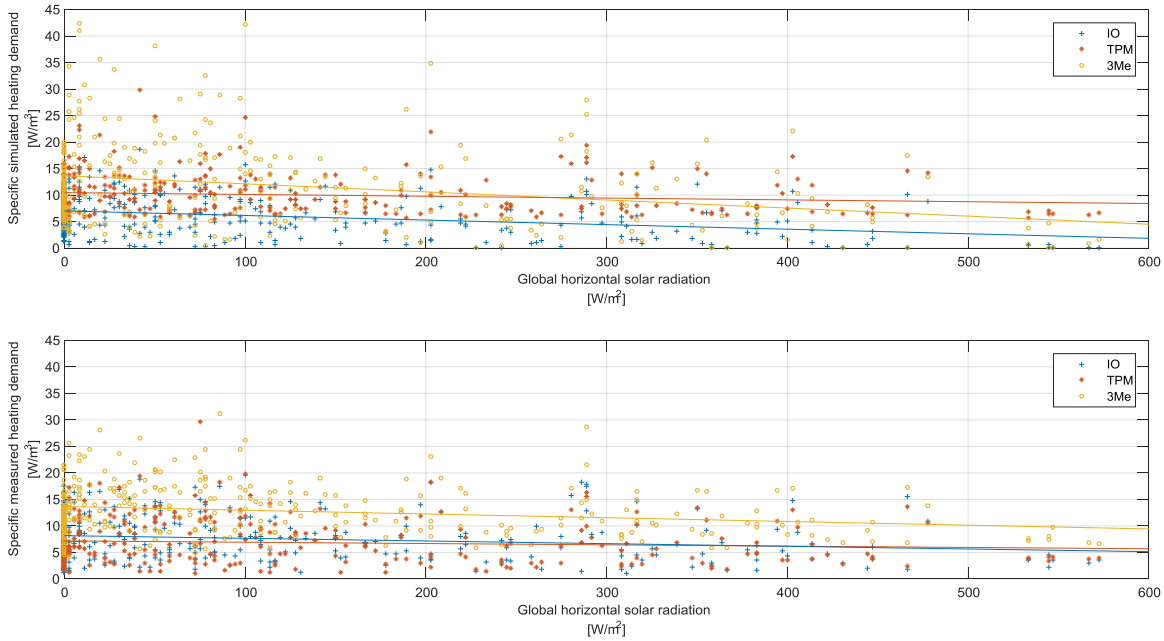


Figure 65 Influence global horizontal solar radiation on the buildings' simulated (above) and actual (below) heating demand during weekdays and opening hours for October 2015.

Table 37 Linear correlation global horizontal solar radiation versus simulated and actual heating demand during weekdays and opening hours for October 2015.

	simulated		measured	
	constant	slope	constant	slope
IO	7	-0.0086	8.2	-0.005
3mE	14	-0.015	14	-0.0071
TPM	10	-0.0034	7.2	-0.0025

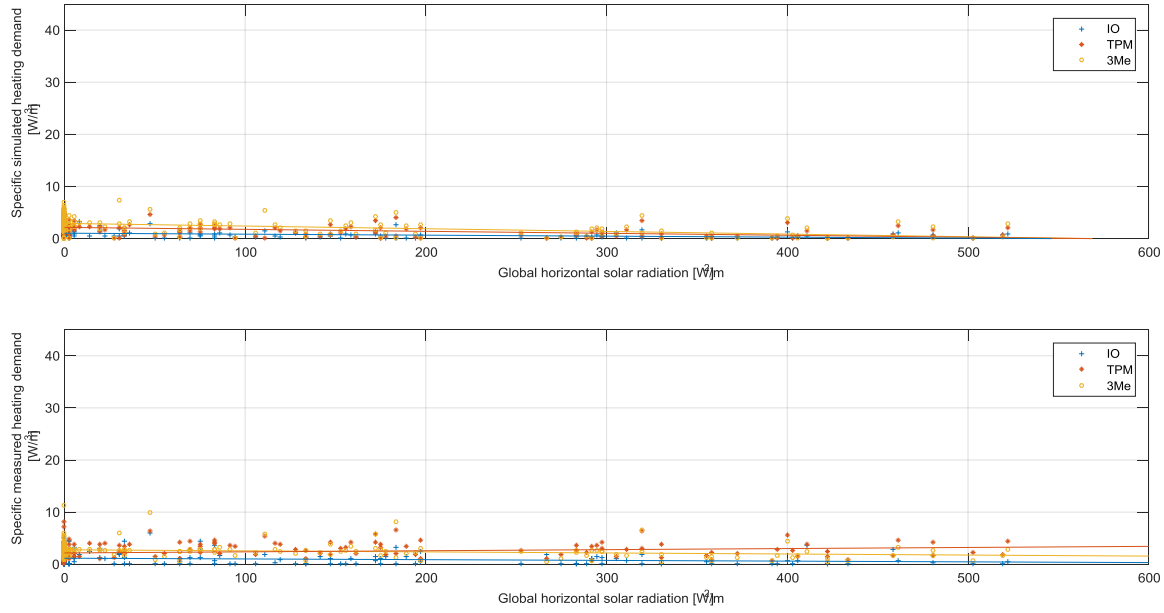


Figure 66 Influence global horizontal solar radiation on the buildings' simulated (above) and actual (below) heating demand during weekends (closing and opening hours) for October 2015.

9.3 Indoor air temperature

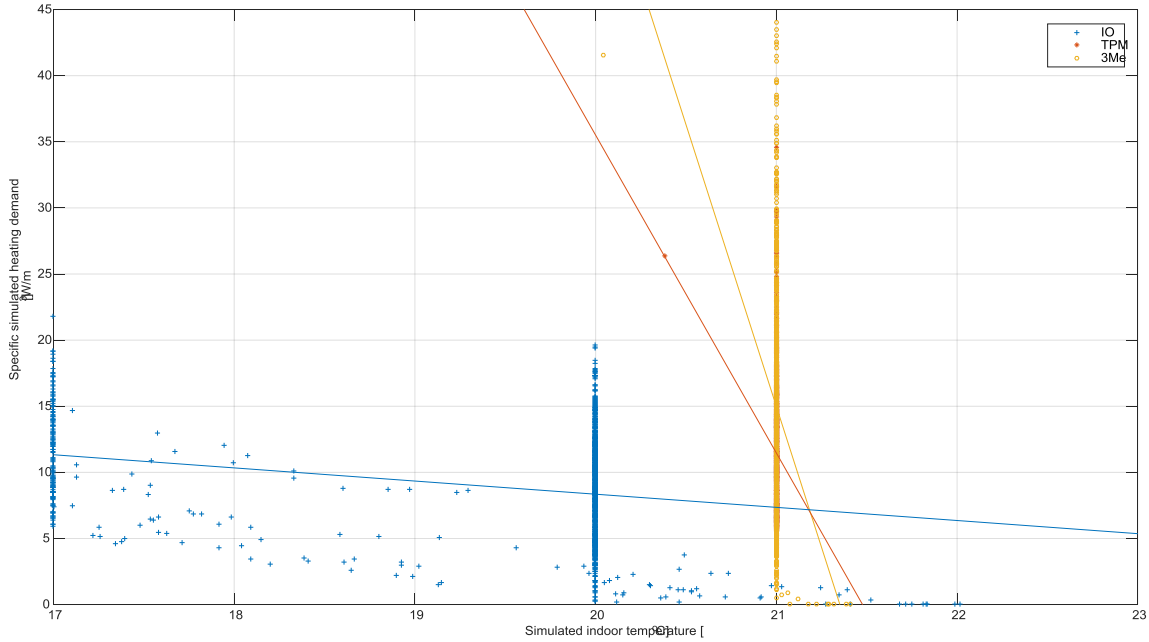


Figure 67 Influence indoor air temperature on the buildings' simulated heating demand during weekdays and opening hours for October 2015.

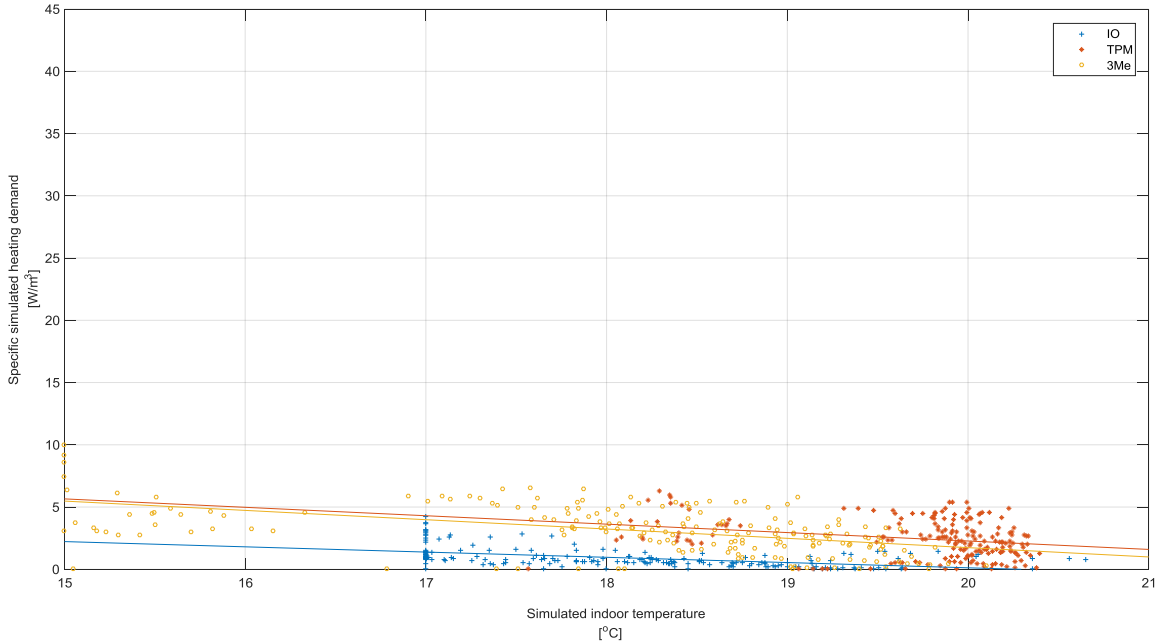


Figure 68 Influence indoor air temperature on the buildings' simulated heating demand during weekdays and closing hours for October 2015.

Table 38 Linear correlation indoor air temperature versus simulated and actual heating demand during weekdays and closing hours for October 2015.

	simulated		**measured**	
	constant	slope	constant	slope
IO	8.5	-0.42	8.5	-0.42
3mE	17	-0.75	17	-0.75
TPM	16	-0.68	16	-0.68

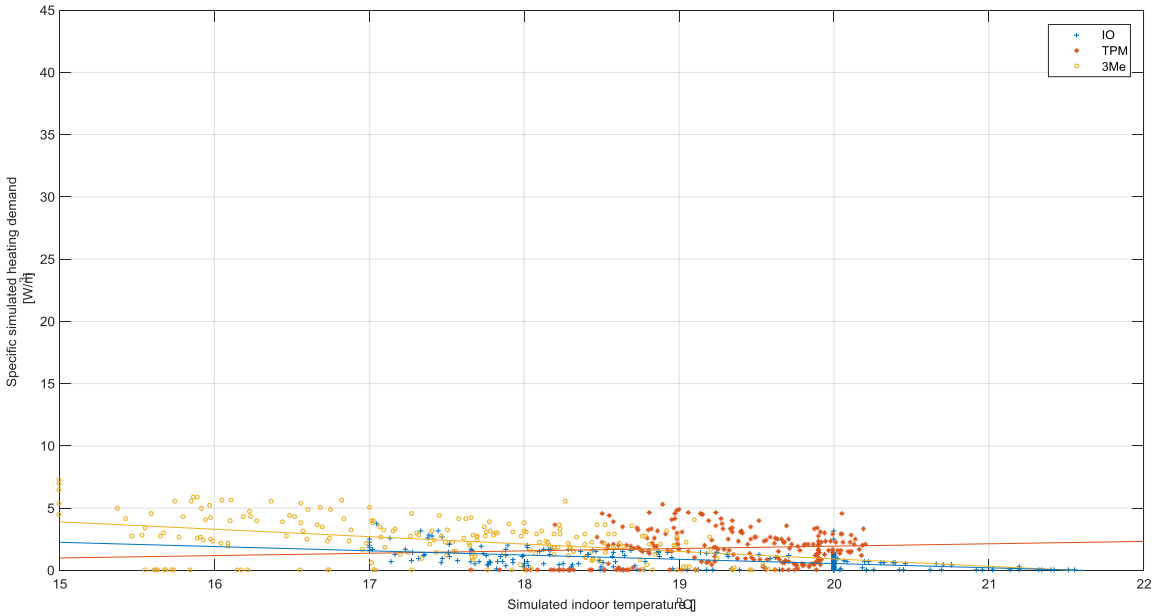


Figure 69 Influence indoor air temperature on the buildings' simulated heating demand during weekends (opening and closing hours) for October 2015.

Table 39 Linear correlation indoor air temperature versus simulated and actual heating demand during weekends (opening and closing hours) for October 2015.

	simulated		**measured**	
	constant	slope	constant	slope
IO	7.4	-0.34	7.4	-0.34
3mE	13	-0.6	13	-0.6
TPM	-1.9	0.19	-1.9	0.19

Appendix 10 Impact of different parameters on the building heating demand of data set 2016

This appendix shows the impact of the most influencing parameters on the simulated and actual heating demand of the buildings. The graphs include the data points and least square for each of the buildings. The actual heating demand of 3Me was wrong, therefore the 3Me data set is excluded in the actual eating demand representations.

The time set in LEA before the 3rd of October was not correct, leading to wrong heating demand predictions. Therefore the data set represented in this appendix corresponds to the period of time from the 3rd of October 2016 at 6:00 (hour of the year: 6630) until 25th November of 2016 at 13:00 (hour of the year: 7909).

10.1 Outdoor temperature

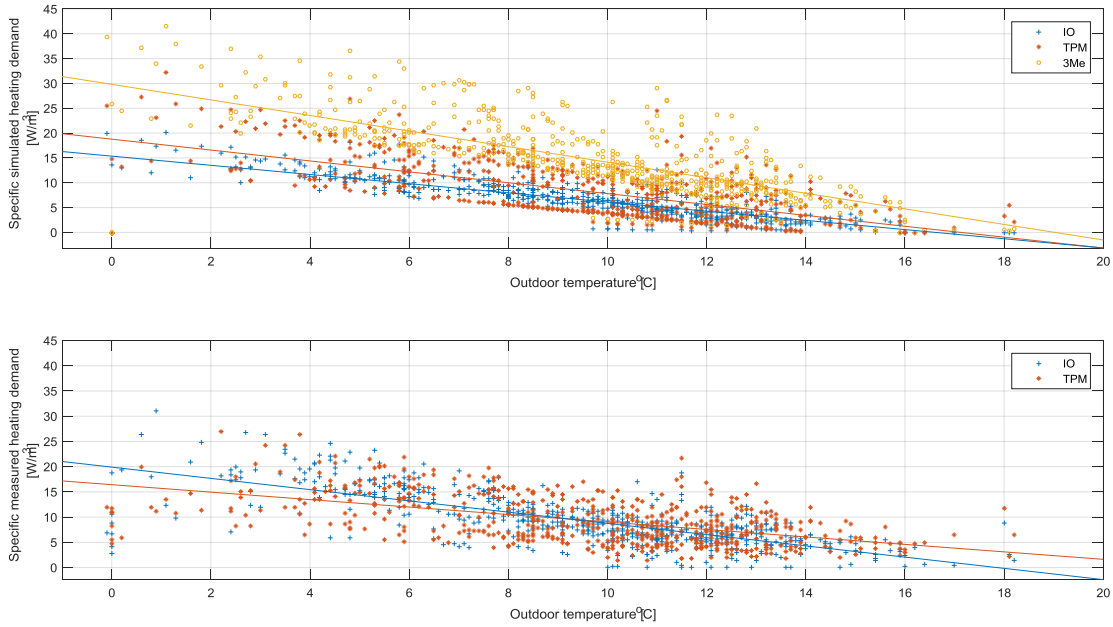


Figure 70 Influence outdoor temperature on the buildings' simulated (above) / measured (below) heating demand during weekdays and opening hours

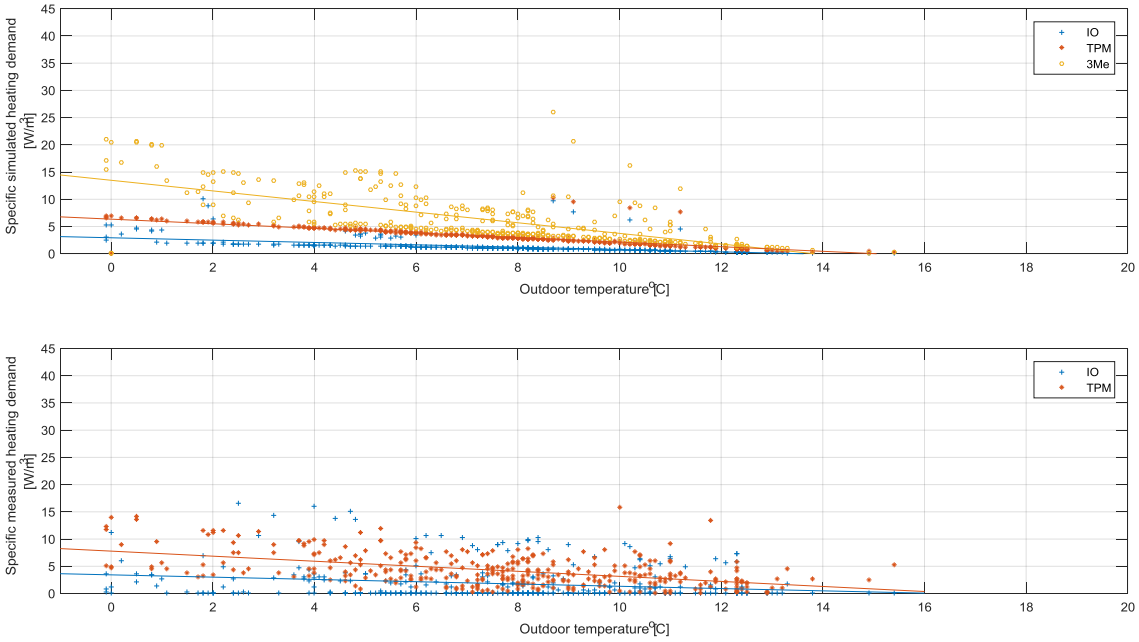


Figure 71 Influence outdoor temperature on the buildings' simulated (above) / measured (below) heating demand during weekdays and closing hours

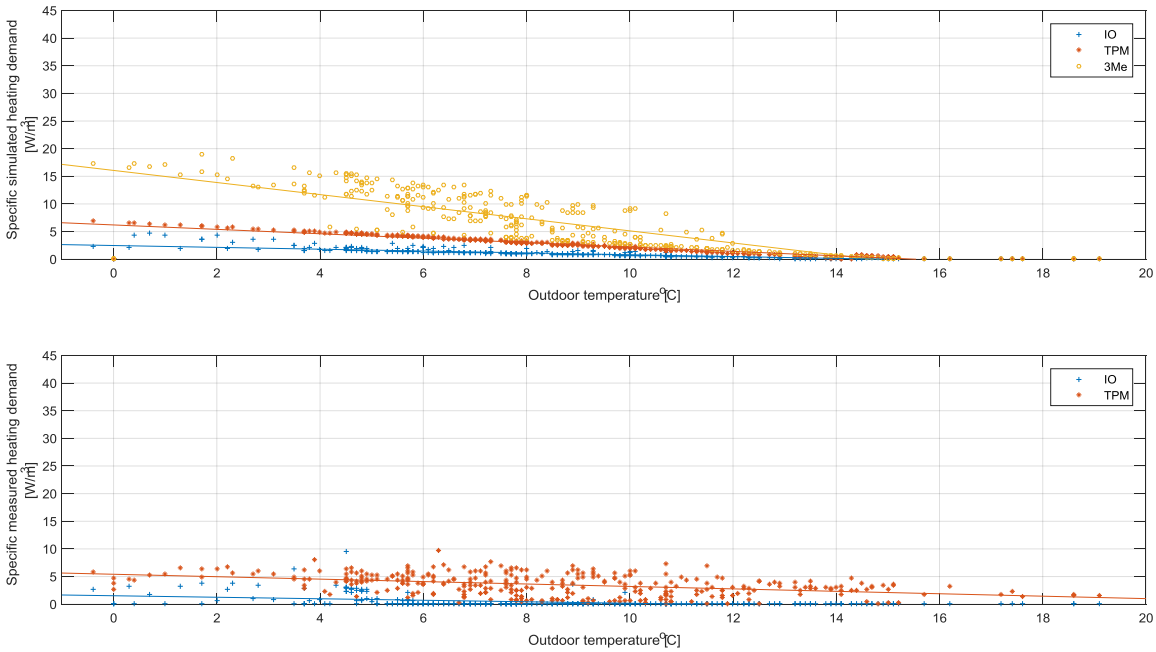


Figure 72 Influence outdoor temperature on the buildings' simulated (above) / measured (below) heating demand during weekends (openings + closing hours)

10.2 Global horizontal solar radiation

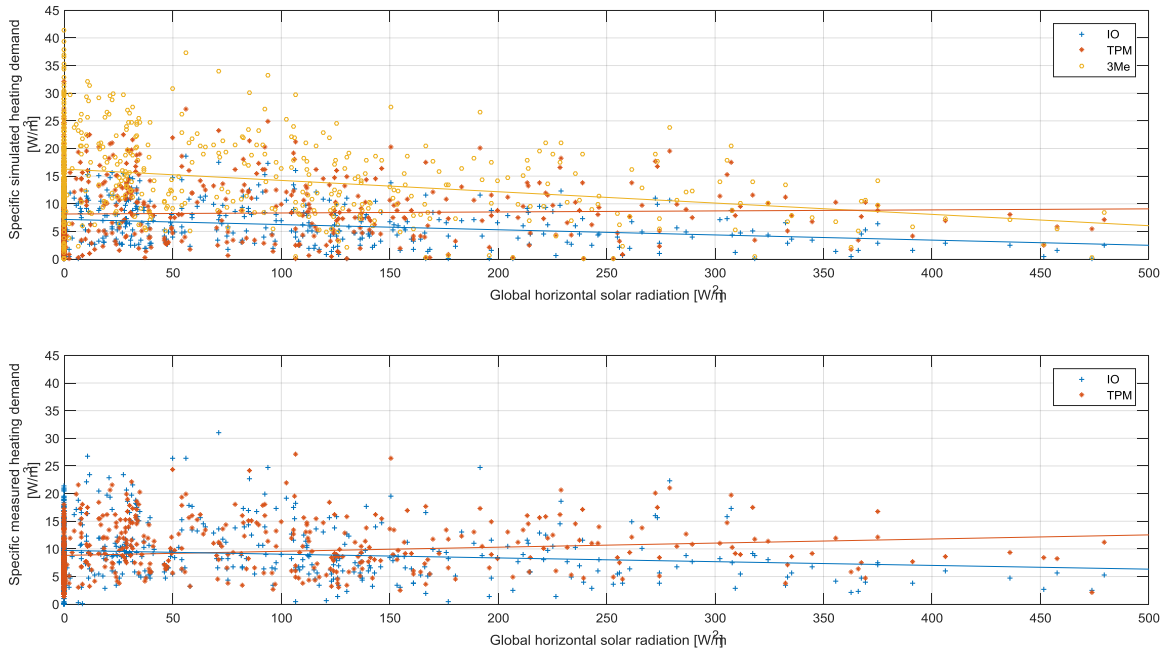


Figure 73 Influence global solar radiation on the buildings' simulated (above) / measured (below) heating demand during weekdays and opening hours

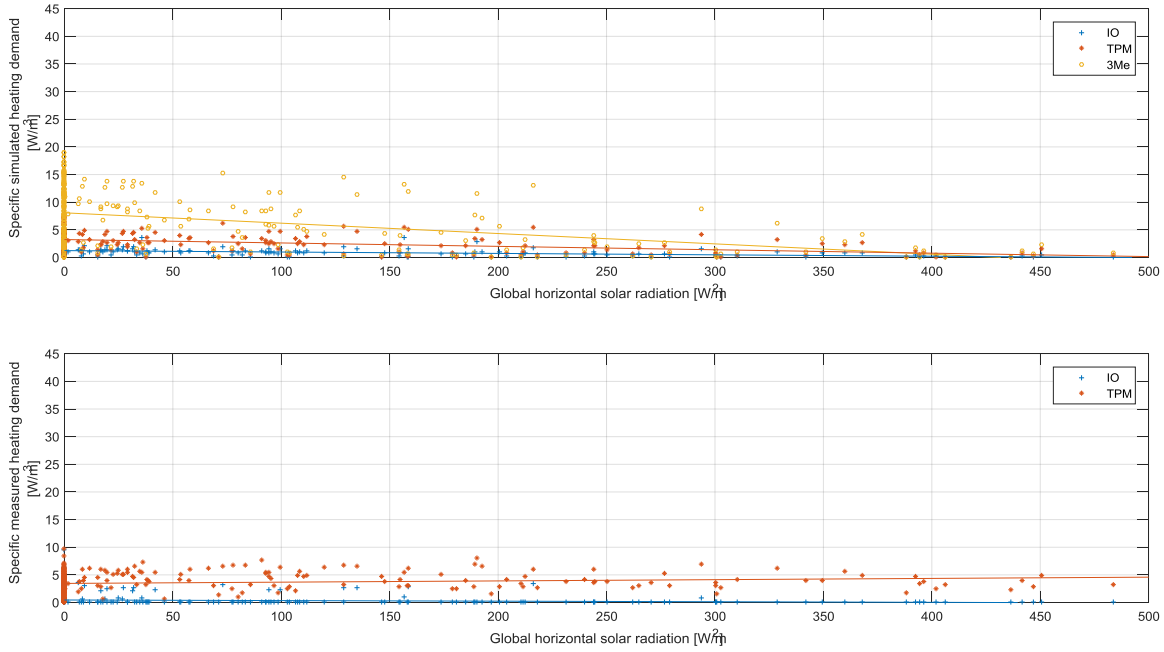


Figure 74 Influence global solar radiation on the buildings' simulated (above) / measured (below) heating demand during weekends (openings + closing hours)

10.3 Wind speed

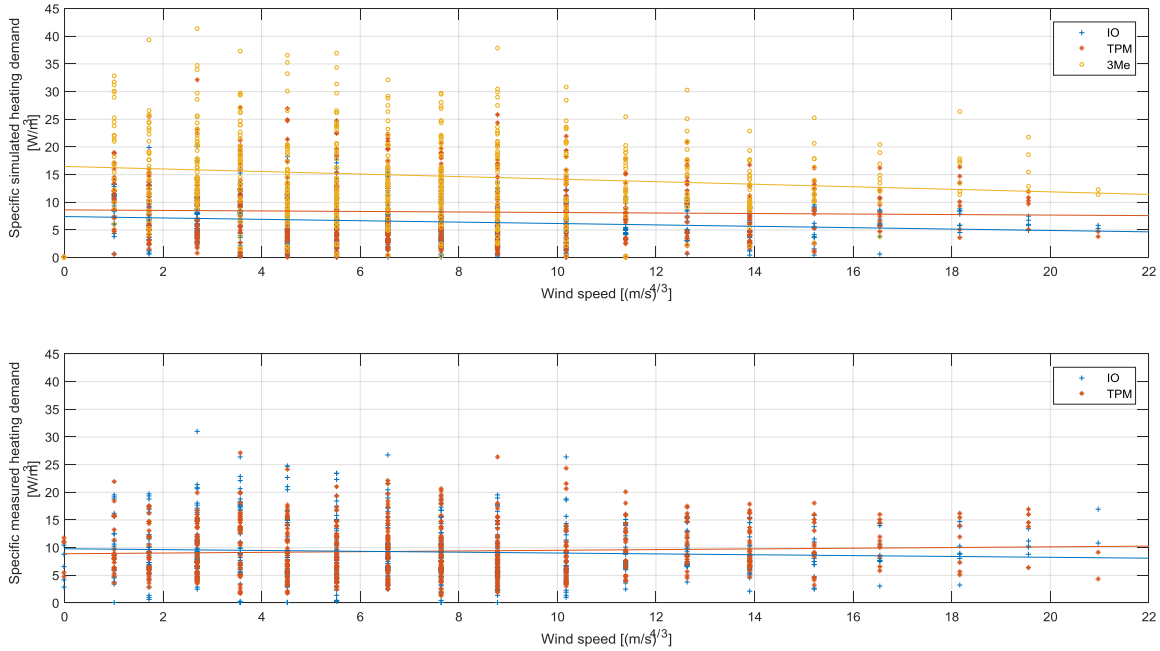


Figure 75 Influence wind speed on the buildings' simulated (above) / measured (below) heating demand during weekdays and opening hours

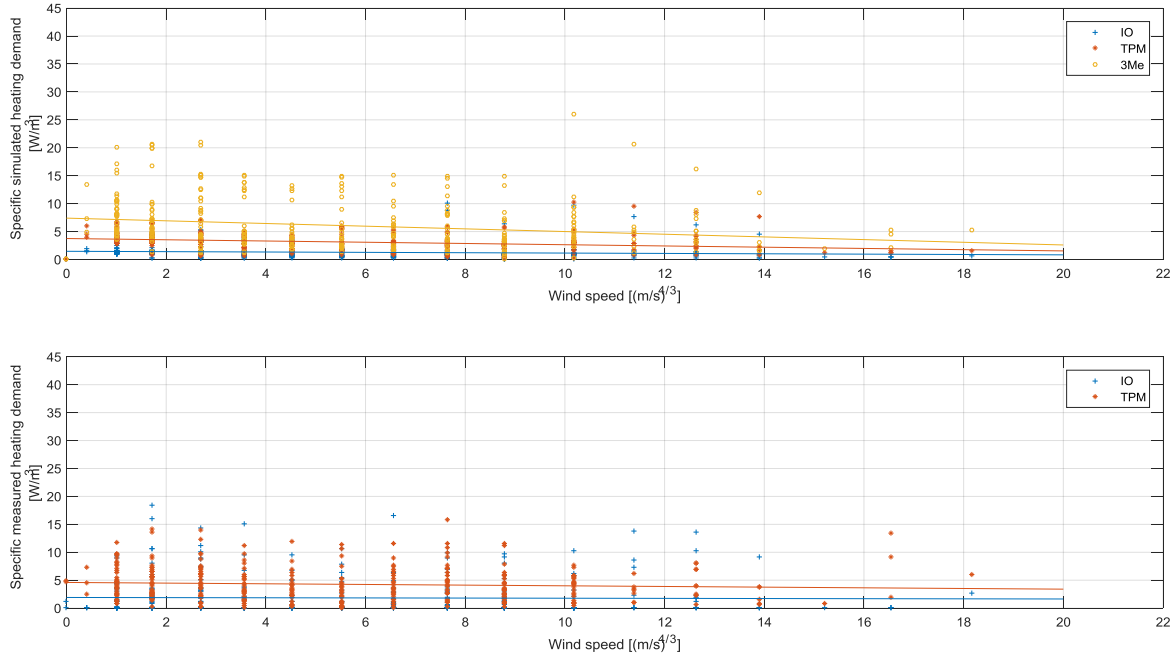


Figure 76 Influence wind speed on the buildings' simulated (above) / measured (below) heating demand during weekdays and closing hours

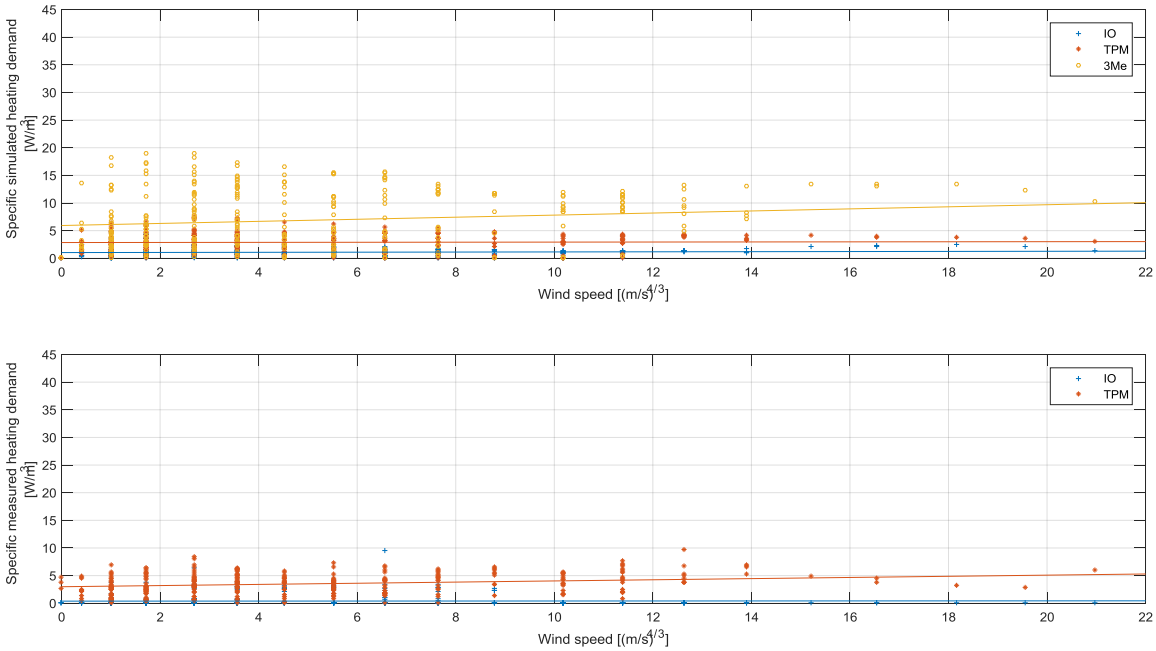


Figure 77 Influence wind speed on the buildings' simulated (above) / measured (below) heating demand during weekends (openings + closing hours)

10.4 Indoor air temperature

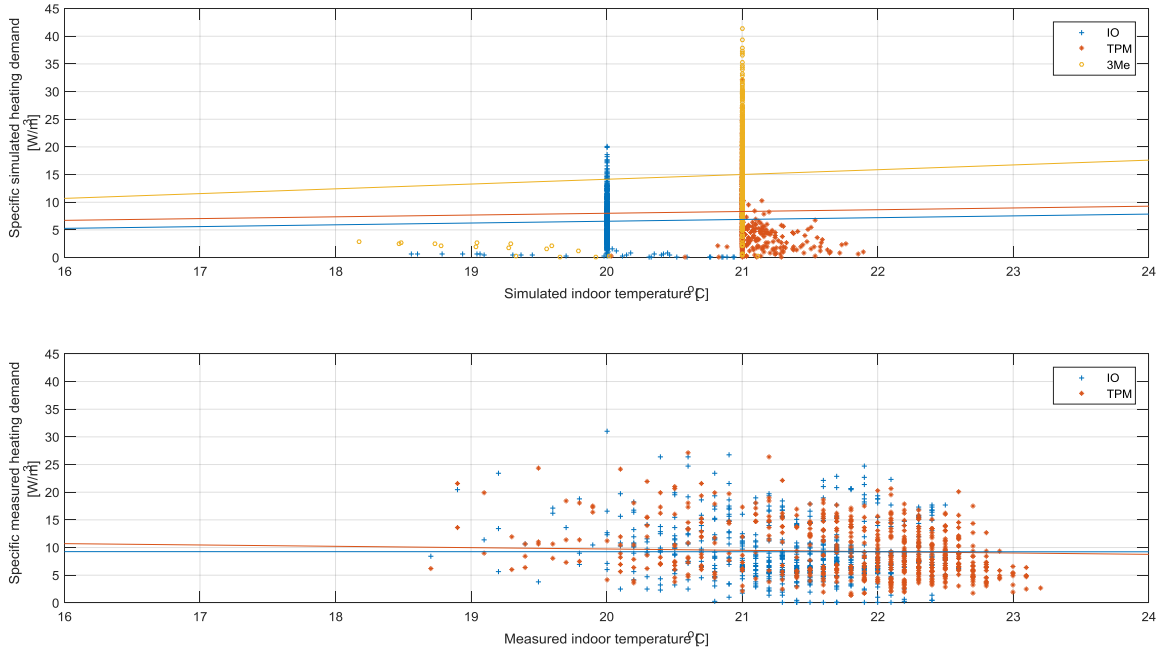


Figure 78 Influence indoor air temperature on the buildings' simulated (above) / measured (below) heating demand during weekdays and opening hours

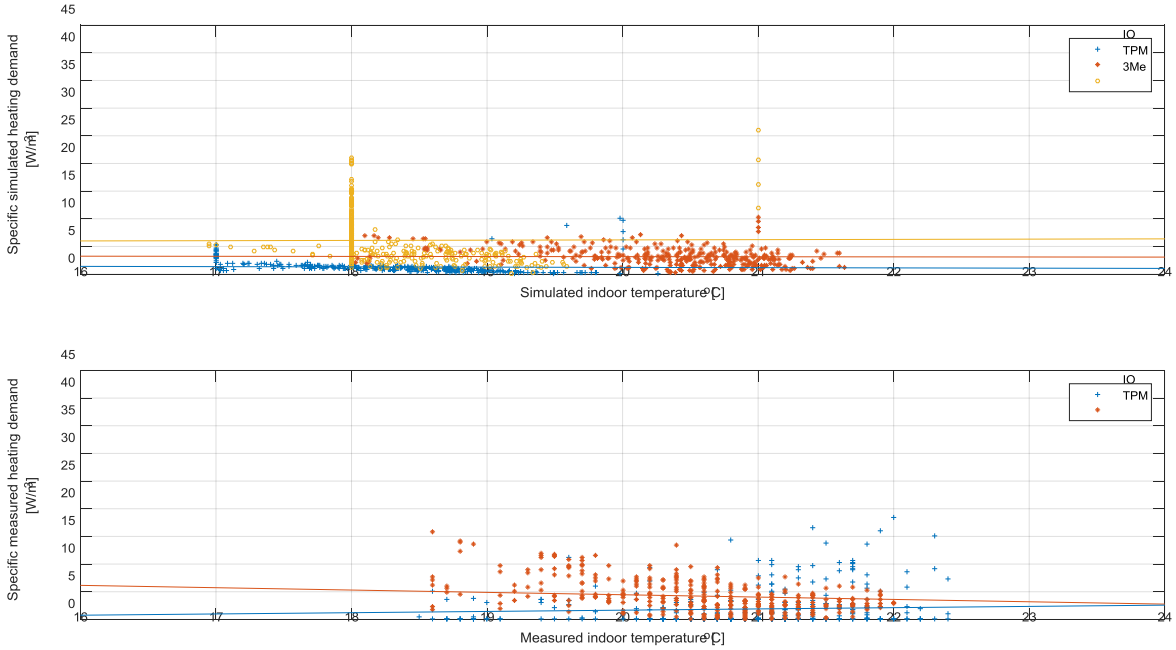


Figure 79 Influence indoor air temperature on the buildings' simulated (above) / measured (below) heating demand during weekdays and closing hours

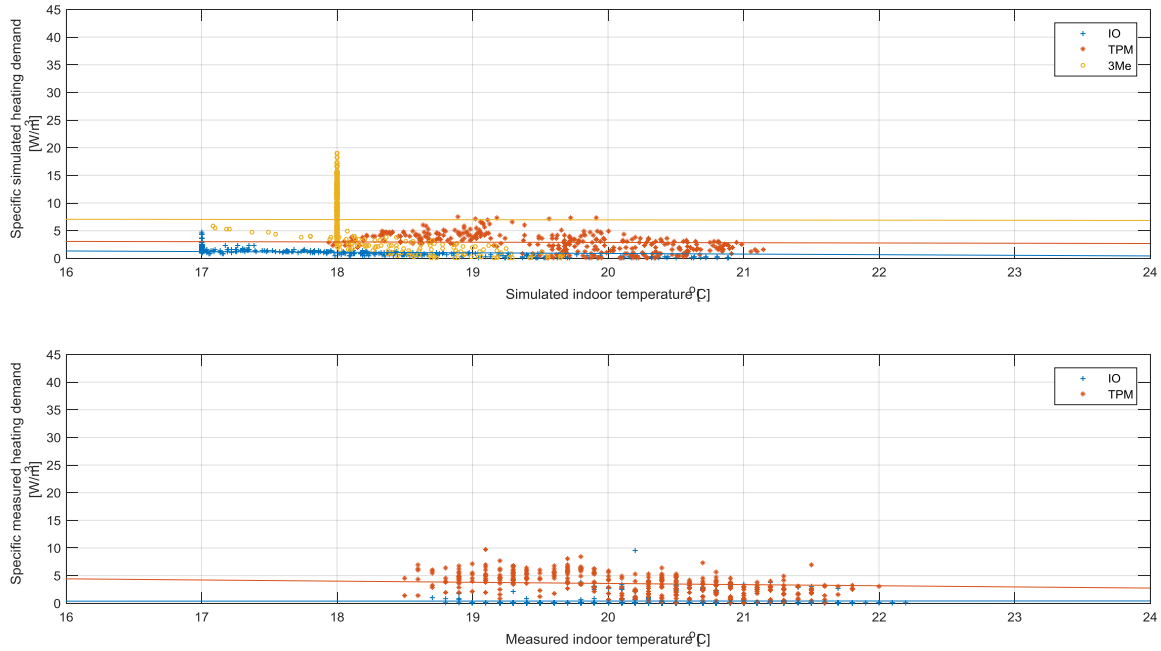


Figure 80 Influence indoor air temperature on the buildings' simulated (above) / measured (below) heating demand during weekends (openings + closing hours)

10.5 Internal heat gain

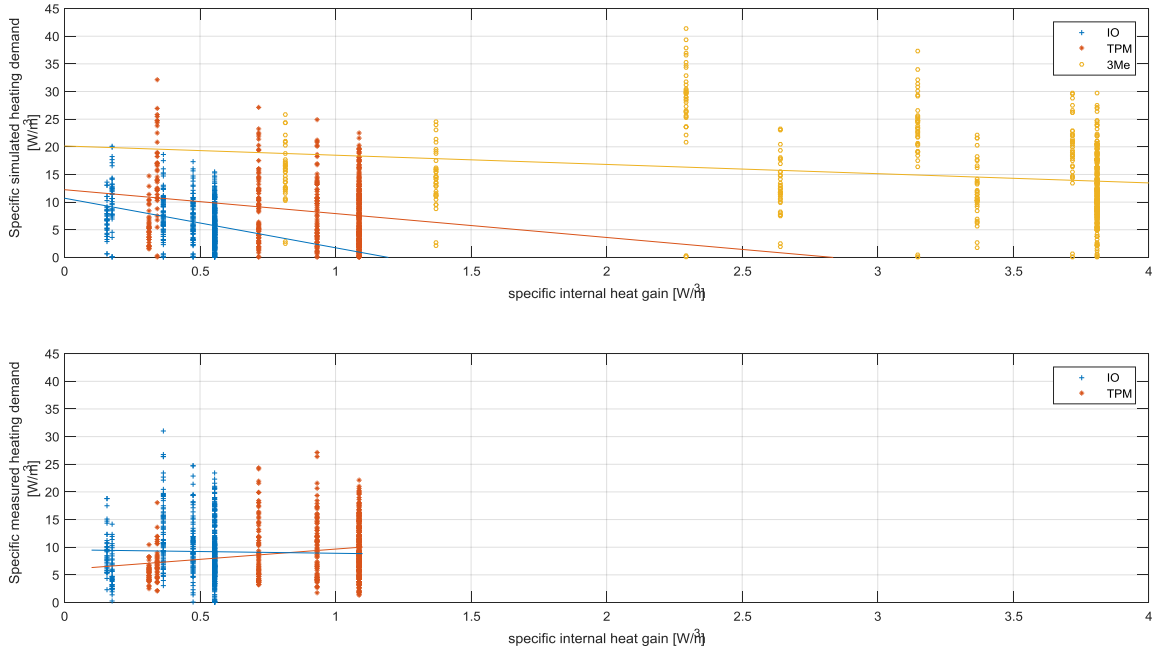


Figure 81 Influence internal heat gain on the buildings' simulated (above) / measured (below) heating demand during weekdays and opening hours

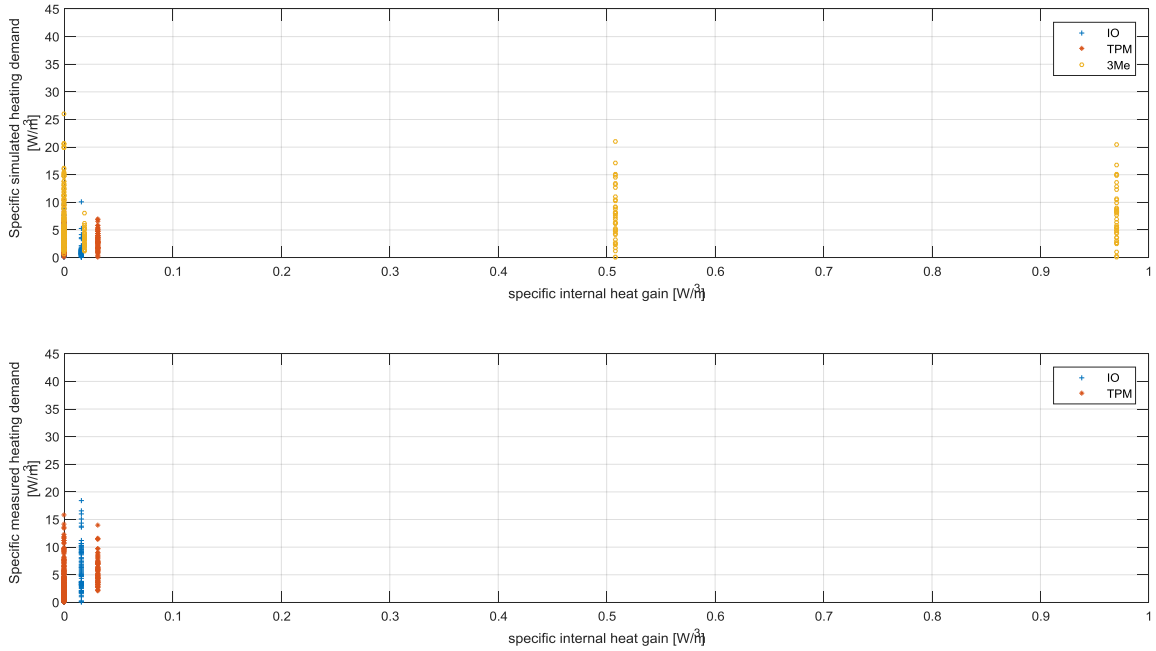


Figure 82 Influence internal heat gain on the buildings' simulated (above) / measured (below) heating demand during weekdays and closing hours

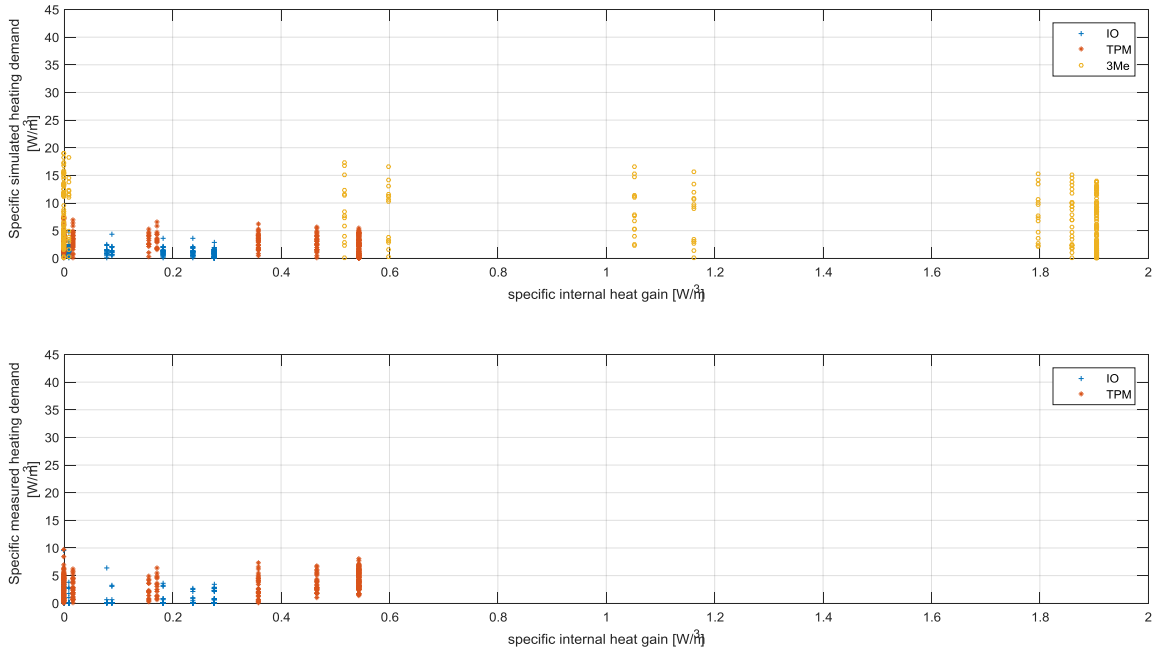


Figure 83 Influence internal heat gain on the buildings' simulated (above) / measured (below) heating demand during weekends (openings + closing hours)

Appendix 11 Appearance interactive stepwise regression function in Matlab

Matlab stepwise regression function is available in the statistical toolbox of Matlab [45]. This function displays an interface for interactively controlling the stepwise addition and removal of the model parameters (see Figure 84). The first square shows the values corresponding to the data set residuals and variables' coefficients. The second square presents a summary of the statistics parameters for the entire model and the third square shows the model history RMSE, which tracks the RMSE from step to step in order to compare the optimality of the different models.

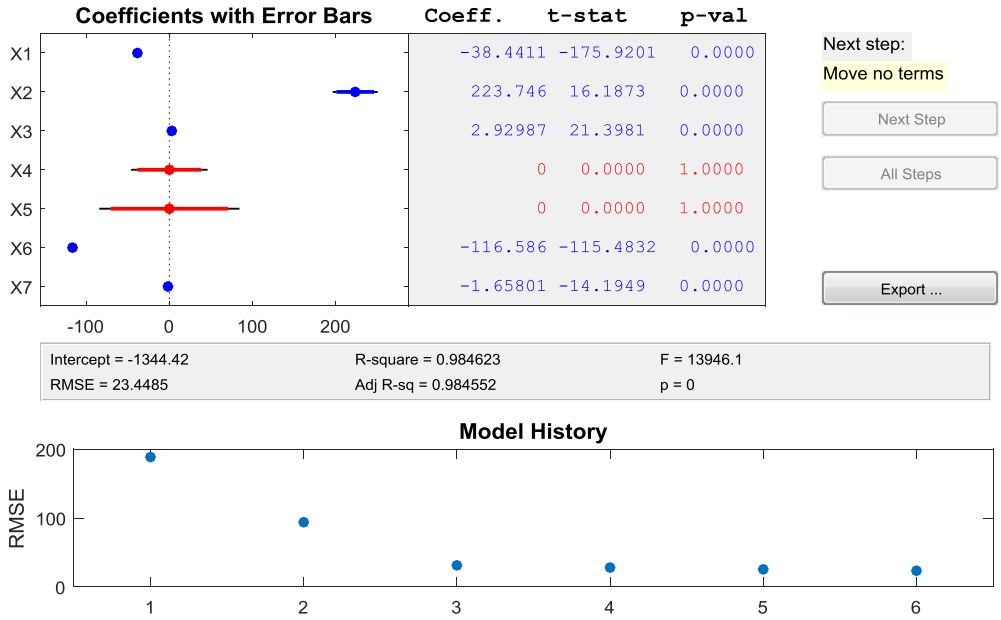


Figure 84 Interface stepwise regression function in Matlab R2015b

The interface indicates the recommended next step. The recommended next step either adds the most significant parameter or removes the least significant parameter until the regression reaches a local minimum of RMSE. At each step, the coefficients of each parameter are recalculated for the new model and the new values for the statistical criteria are given. The method finishes when there is no step that can improve the model [45].

Appendix 12 Results regression models

12.1 Multivariate linear regression model 1 (models 1a and 1b)

The regression model 1a and model 1b correspond to the models developed in chapter 8 (Multivariate models based on the building's thermal energy balance principle) described by the equation below. In model 1a all variables where considered, while in model 1b the variable $\left(T^{t}_{indoor\ surfaces} - T_{indoor}\right)$ was neglected

$$Q_{demand}\left[\frac{w}{m^3}\right] = constant + C_1\left(T_{ground} - T_{indoor}\right) + C_2\left(T_{outdoor} - T_{indoor}\right) + C_3 V_{wind}\left(T_{outdoor} - T_{indoor}\right) + C_4\left(T_{out\ AHU} - T_{indoor}\right) + C_5\left(T^{t}_{indoor\ surfaces} - T_{indoor}\right) + C_6\, Q_{solar} + C_7\, Q_{internal}$$

Table 40 Coefficients and statistical parameters of the multivariate regression model 1 for the specific heating demand prediction (W/m³) for IO, 3mE and TPM, respectively. Data set: weekdays during opening hours from 5th October 2015 until 14th January 2016.

	IO	3mE	TPM
Constant	-6.09	-52.65	61.14
C1	-0.90	-10.50	11.65
C2	-0.76	0.45	-2.49
C3	-0.01	-0.01	-0.01
C4	0	7.17	-7.93
C5	-2.50	-7.37	-7.33
C6	0.0012	0.0019	0.0009
C7	-1.92	-0.77	-1.43
R^2	98.60%	99.6%	97.52%
Adjusted R^2	98.60%	99.59%	97.50%
RMSE	0.48	0.52	0.66

This table shows the statistical parameters used to analyse the significance of the model. The other statistical parameters (used to analyse the data set residuals and significant level of the variables'coefficients) are within the limit values.

Table 41 Coefficients and statistical parameters of the multivariate regression model 2 for the specific heating demand prediction (W/m³) for IO, 3mE and TPM, respectively. Data set: weekdays during opening hours from 5th October 2015 until 14th January 2016.

	IO	3mE	TPM
Constant	60.82	469.95	363.05
C1	7.40	89.13	67.27
C2	-0.89	-14.62	-10.56
C3	-0.02	-0.0013	-0.0036
C4	0	-58.17	-43.43
C5	-	-	-
C6	-0.0002	0.0004	0.0074
C7	-4.59	-0.26	-3.04
R^2	79.43%	53.19%	63.71%
Adjusted R^2	79.34%	52.93%	63.51%
RMSE	1.83	5.63	2.51

- Coefficient corresponding to a variable neglected in the multivariate regression model

This table shows the statistical parameters used to analyse the significance of the model. The other statistical parameters (used to analyse the data set residuals and significant level of the variables'coefficients) are within the limit values.

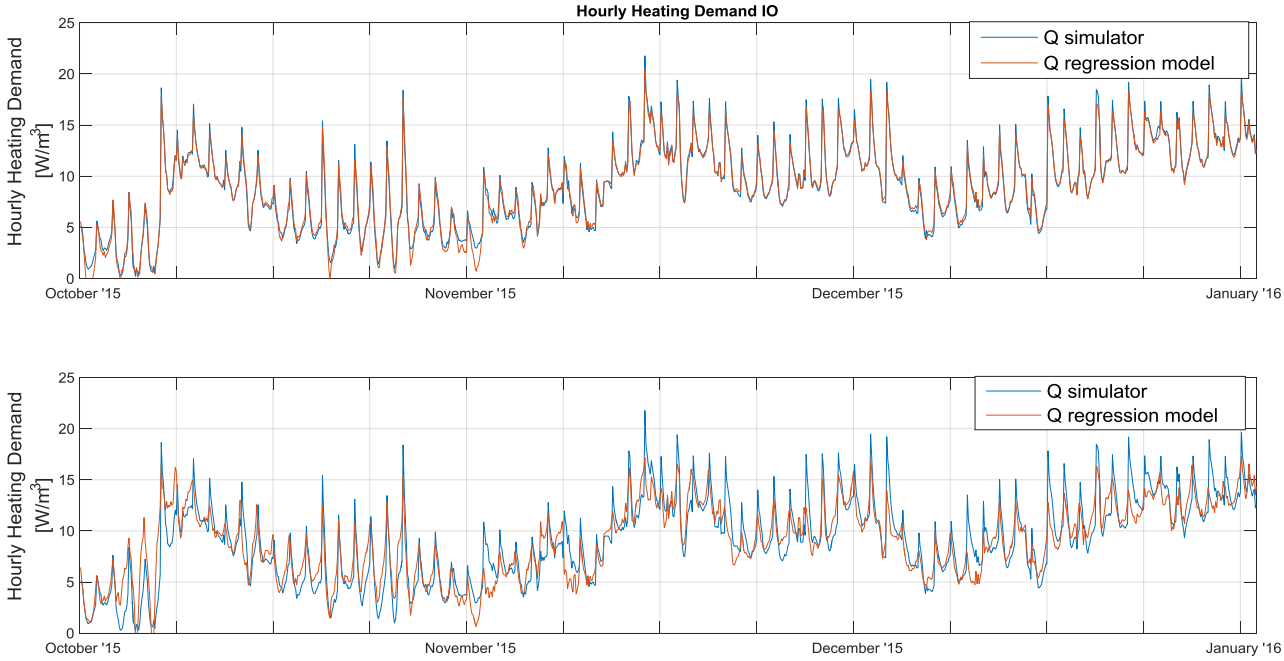


Figure 85 Comparison fitting profile of the regression model 1a (above) and 1b (below) for IO. Data set: weekdays during opening hours from 5th October 2015 until 14th January 2016.

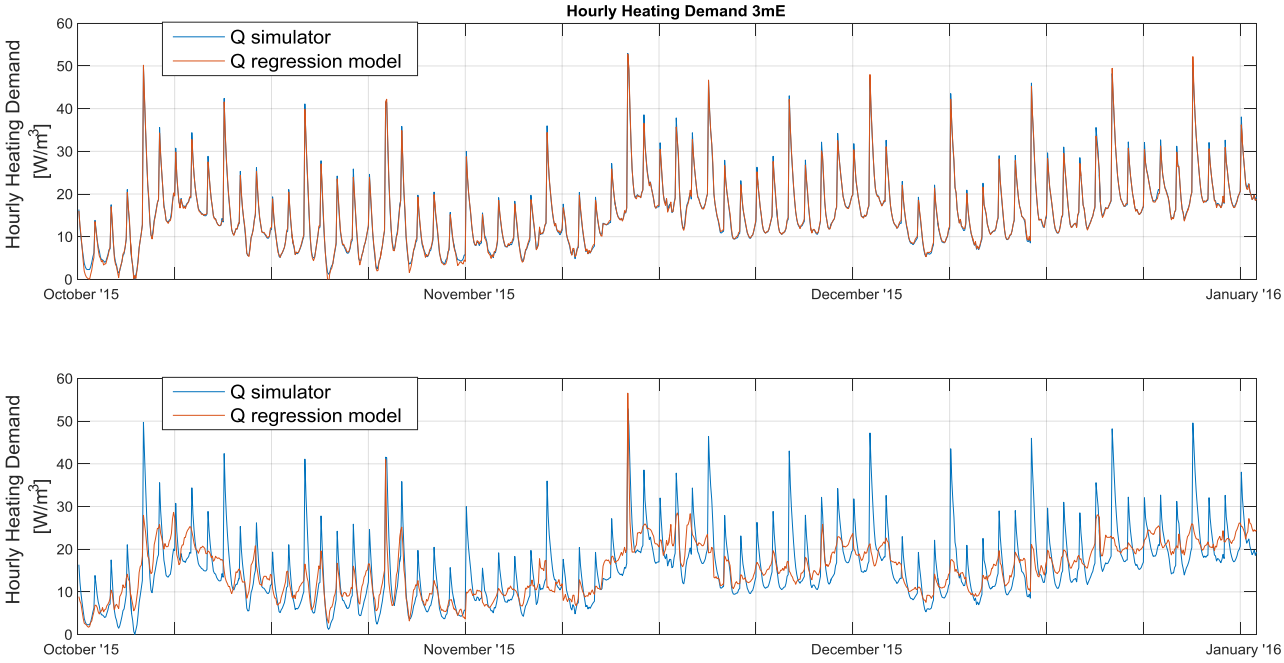


Figure 86 Comparison fitting profile of the regression model 1a (above) and 1b (below) for 3mE. Data set: weekdays during opening hours from 5th October 2015 until 14th January 2016.

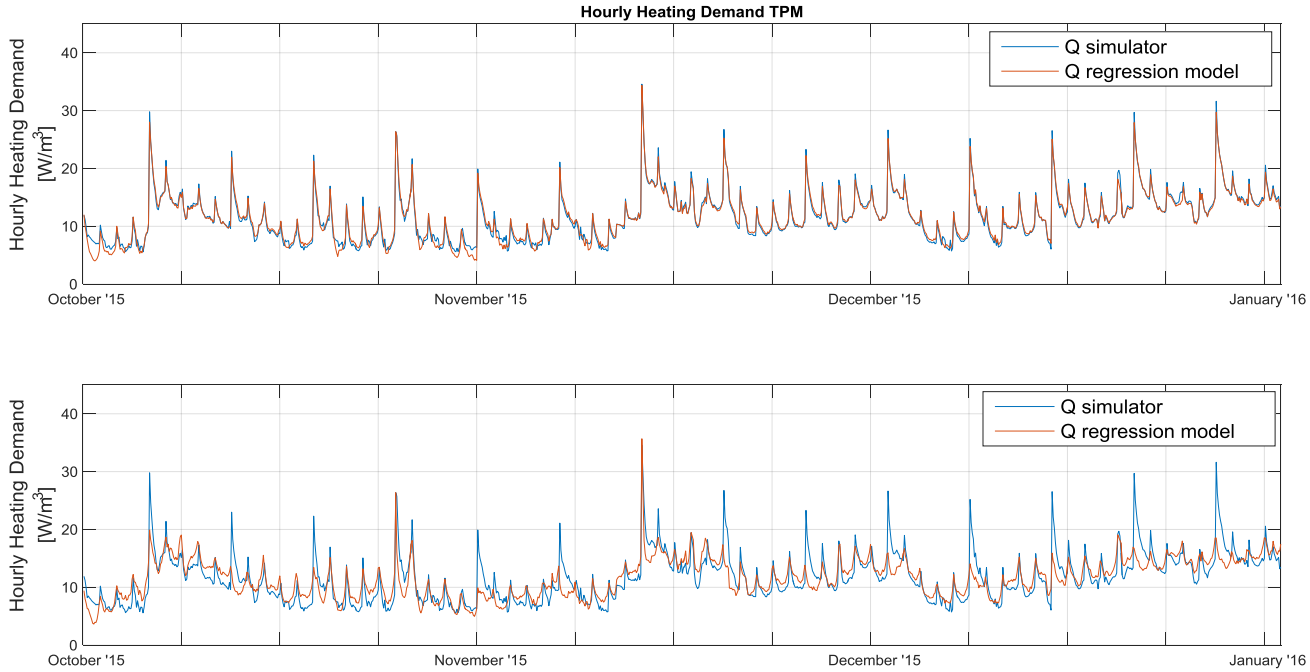


Figure 87 Comparison fitting profile of the regression model 1a (above) and 1b (below) for TPM. Data set: weekdays during opening hours from 5th October 2015 until 14th January 2016.

12.2 Multivariate linear regression model 2 (model 2a, 2b and 2c)

The regression models 2a, 2b and 2c correspond to the models developed in chapter 9 (Multivariate linear model improvement towards application into practice) described by the equation below. In model 2a the indoor surface temperature is included, in model 2b the indoor surface temperature is excluded and in model 2c the indoor surface temperature is replaced.

$$Q_{demand}\left[\frac{w}{m^3}\right] = const + C_a\,(T_{outdoor}) + C_b\,(\boldsymbol{T_{indoor}}) + C_c(\mathrm{V}_{wind}) + C_d\,(T_{ground}) + C_e\,(\boldsymbol{T_{out\ AHU}}) + C_f\,(\boldsymbol{T^t_{indoor\ surfaces}}) + C_{f,int1a}\,Q_{internal,1a} + C_{f,solar3a}\,Q_{solar,3a} + C_g\,Q_{solar} + C_h\,Q_{internal}$$

Results for IO

Table 42 Coefficients and statistical parameters of the multivariate regression model 2 for the specific heating demand prediction (W/m³) for IO with indoor surface temperature included (model 2a), indoor surface temperature excluded (model 2b) and indoor surface temperature replaced (model 2c). Data set: weekdays during opening hours from 5th October 2015 until 14th January 2016.

Coefficients	Ts included (model 2a)	Ts excluded (model 2b)	Ts replaced (model 2c)
Constant	-17.69	147.63	32.34
Ca	-0.82	-0.97	-0.92
Cb	4.23	-6.37	-0.65
Cc	0.14	0.19	0.16
Cd	0	0	0
Ce	0	0	0
Cf	-2.50	-	-
Cf,int1a	-	-	-11.42
Cf,solar3a	-	-	-0.009481
Cg	0.0013505	-5.89E-05	-0.000676
Ch	-1.93	-4.62	8.44
Adjusted R^2	98.52%	79.42%	90.70%
RMSE	0.48	1.82	1.22

- Coefficient corresponding to a variable excluded/neglected in the multivariate regression model

This table shows the statistical parameters used to analyse the significance of the model. The other statistical parameters (used to analyse the data set residuals and significant level of the variables' coefficients) are within the limit values.

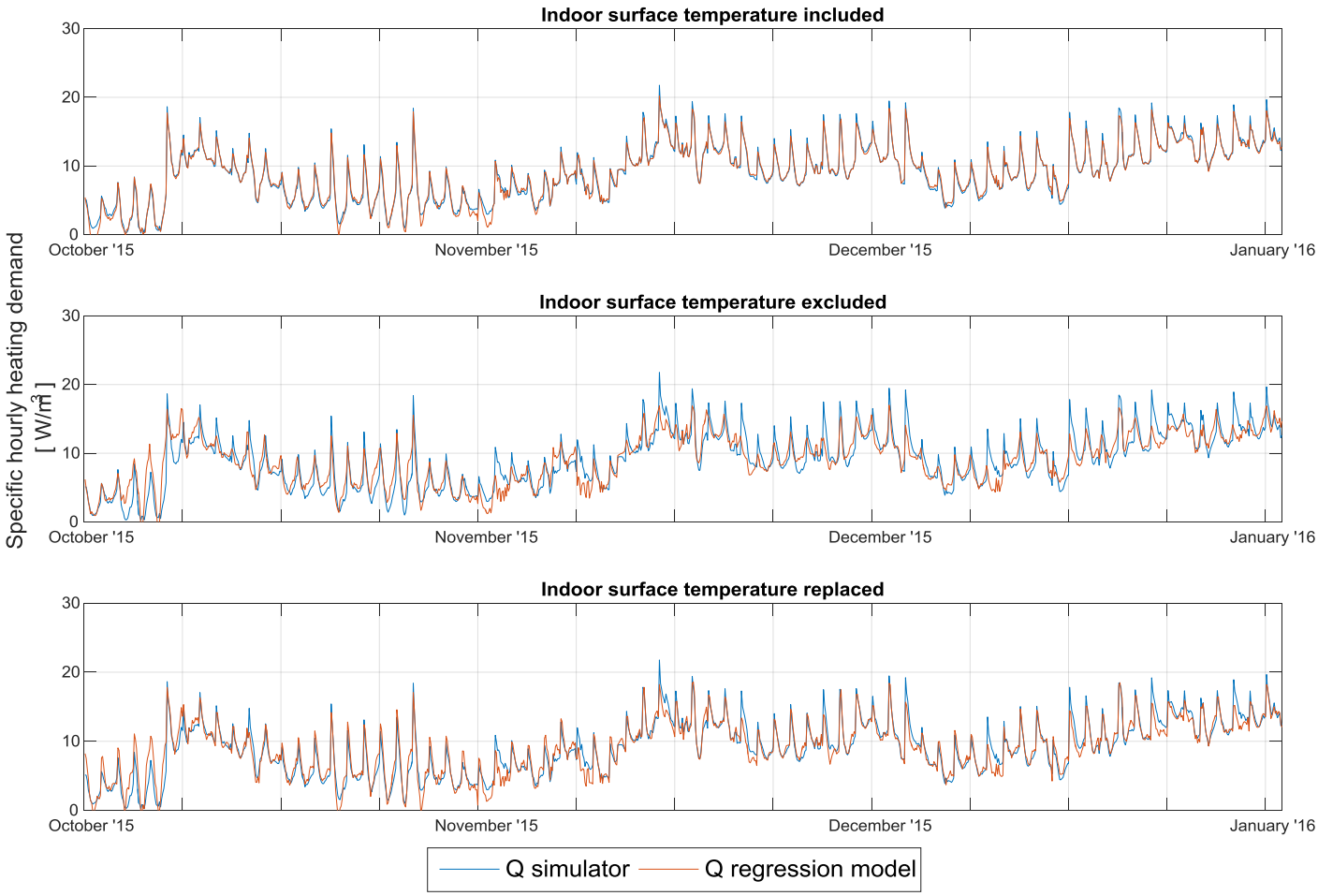


Figure 88 Fitting profile of the multivariate regression model for the specific heating demand prediction defined by model 2 for IO with indoor surface temperature included (above), excluded (middle) and replaced (below), respectively. Data set: weekdays during opening hours from 5th October 2015 until 14th January 2016.

Results for 3mE

Table 43 Coefficients and statistical parameters of the multivariate regression model 2 for the specific heating demand prediction (W/m^3) for 3mE with indoor surface temperature included (model 2a), indoor surface temperature excluded (model 2b) and indoor surface temperature replaced (model 2c). Data set: weekdays during opening hours from 5th October 2015 until 14th January 2016.

Coefficients	Ts included (model 2a)	Ts excluded (model 2b)	Ts replaced (model 2c)
Constant	-219.06	1542.52	827.98
Ca	0.78	-14.65	-9.08
Cb	10.52	-16.39	-5.25
Cc	0.15	0.04	-0.02
Cd	0	0	0
Ce	8.93	-58.27	-34.08
Cf	-7.36	-	-
Cf,int1a	-	-	-7.09
Cf,solar3a	-	-	-0.02
Cg	0.002	0.0005	0.0062
Ch	-0.78	-0.27	5.60
Adjusted R^2	99.58%	52.90%	83.2%
RMSE	0.53	5.63	3.36

- Coefficient corresponding to a variable excluded/neglected in the multivariate regression model

This table shows the statistical parameters used to analyse the significance of the model. The other statistical parameters (used to analyse the data set residuals and significant level of the variables'coefficients) are within the limit values.

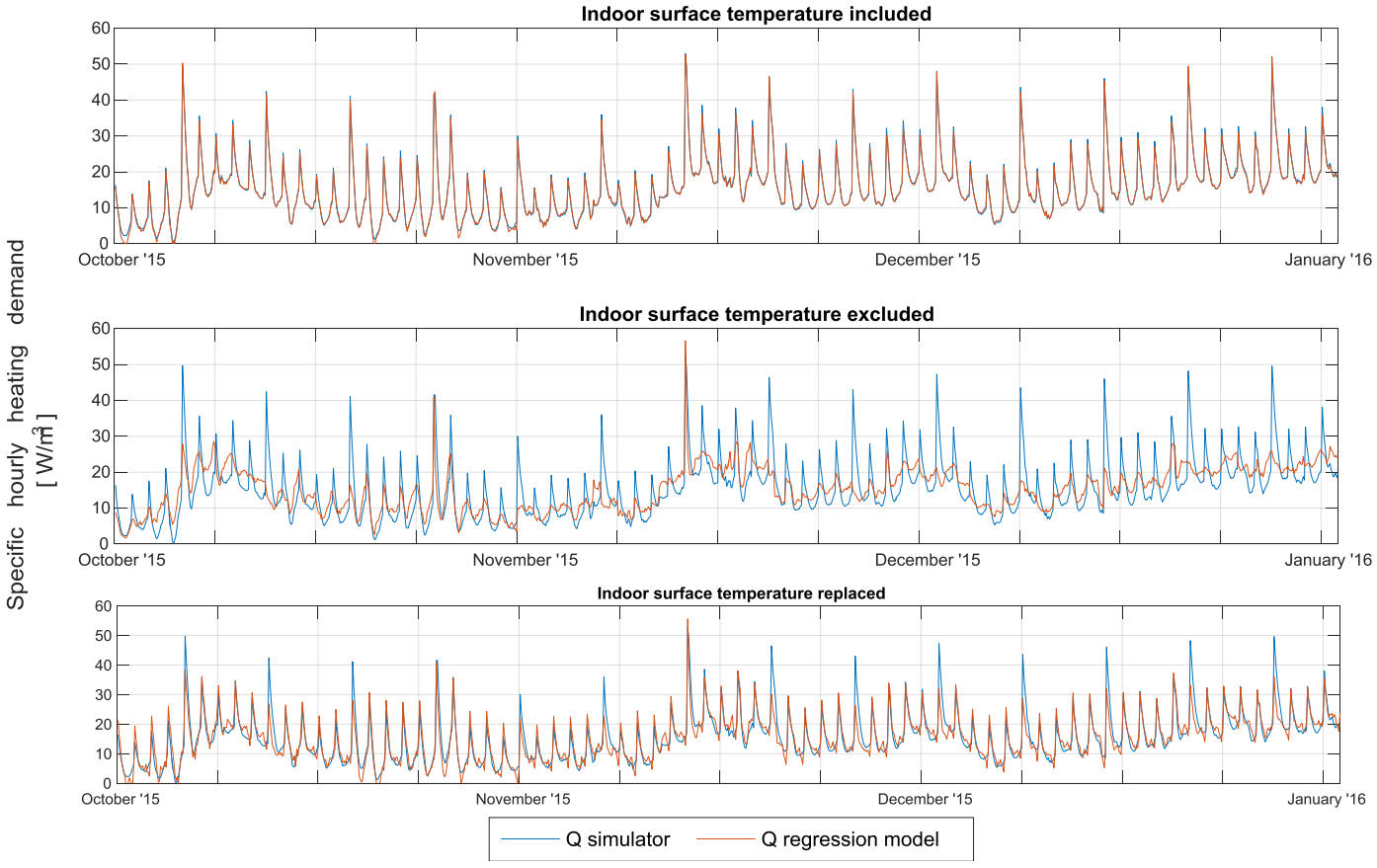


Figure 89 Fitting profile of the multivariate regression model for the specific heating demand prediction defined by model 2 for 3mE with indoor surface temperature included (above), excluded (middle) and replaced (below), respectively. Data set: weekdays during opening hours from 5th October 2015 until 14th January 2016.

Results for TPM

Table 44 Coefficients and statistical parameters of the multivariate regression model 2 for the specific heating demand prediction (W/m^3) for TPM with indoor surface temperature included (model 2a), indoor surface temperature excluded (model 2b) and indoor surface temperature replaced (model 2c). Data set: weekdays during opening hours from 5th October 2015 until 14th January 2016.

Coefficients	Ts included (model 2a)	Ts excluded (model 2b)	Ts replaced (model 2c)
Constant	156.13	1159.36	918.42
Ca	-2.14	-10.48	-8.57
Cb	6.50	-13.18	-9.73
Cc	0.15	0.05	0.02
Cd	0	0	0
Ce	-6.08	-42.98	-34.55
Cf	-7.33	-	-
Cf,int1a	-	-	-6.86
Cf,solar3a	-	-	0.000086
Cg	0.001	0.007	0.006
Ch	-1.43	-3.05	3.49
Adjusted R^2	97.47%	63.52%	73%
RMSE	0.66	2.5	2.16

- Coefficient corresponding to a variable excluded/neglected in the multivariate regression model

This table shows the statistical parameters used to analyse the significance of the model. The other statistical parameters (used to analyse the data set residuals and significant level of the variables' coefficients) are within the limit values.

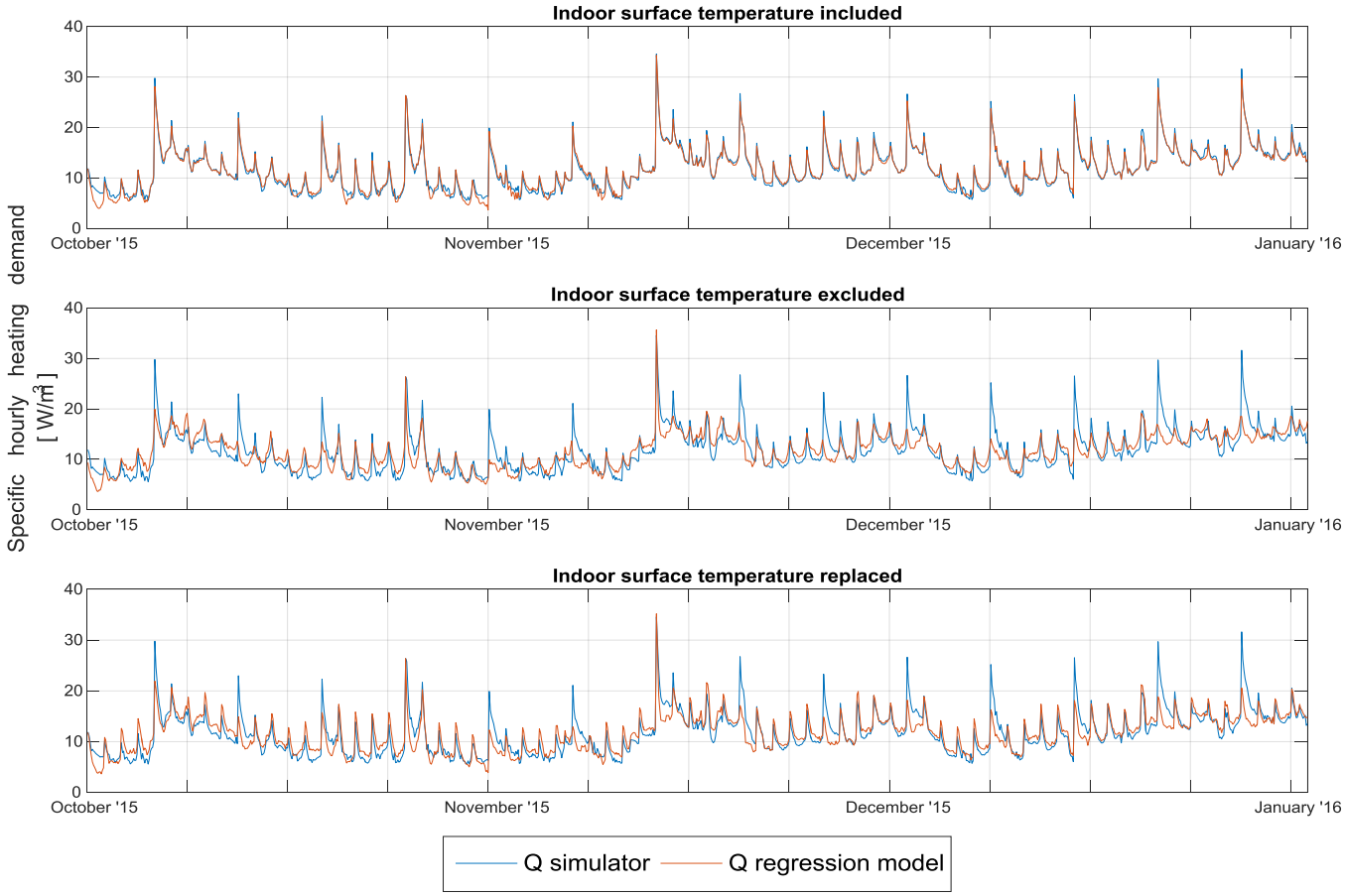


Figure 90 Fitting profile of the multivariate regression model for the specific heating demand prediction defined by model 2 for TPM with indoor surface temperature included (above), excluded (middle) and replaced (below), respectively. Data set: weekdays during opening hours from 5th October 2015 until 14th January 2016.