Leeghwaterstraat 44
2628 CA Delft
P.O. Box 6012
2600 JA Delft
The Netherlands

www.tno.nl

T +31 88 866 22 00

TNO report

TNO 2020 R10229

TKI Optimaal – Development of data-driven RC-network simulation model for performance guarantees

Date	16 March 2020
Author(s)	Marleen Spiekman Olav Vijlbrief Haseeb Tahir Yasin Bulut
Number of pages	82 (incl. appendices)
Number of appendices	4
Sponsor	
Project name	TKI Optimaal
Project number	060.19794

Summary

Background and purpose of the project
Current monitoring and data analyses provide limited insight in the couses of deviating performances. This is an obstacle to improving and guaranteeing the quality of NoM houses, especially if large-scale roll-out is desired in the future. The aim of the TKI Optimaal project is to develop models and algorithms for data analysis with which, in time, at least 80 to 90% of the deviation between the predicted and actual energy and indoor climate performance of individual NoM houses can be explained.

General approach
The route we are following is to develop a data-driven RC-network simulation model of NOM houses that will allow us to approach the actual performance of those houses on an individual level. We choose the development of a physical model over the use of a model derived via Artificial Intelligents (AI): AI has its limitations, it is a black box that doesn't make use of the physical information in the data, you need training of the model and the model is only valid within the boundaries of the trainingdata. Furthermore, since AI is a black box, it gives at the most limited information about what will happen after changes are made, e.g. to predict what will happen when new technology is used or when users will behave differently. However, AI can be useful on top of a physical model (see next stept)
The development of a data-driven RC-network simulation model of NOM houses involves roughly two steps: 1) the development of the model and 2) the best possible estimation of the parameters in the model. In TKI Optimaal we took a big first step in this by setting up a data-driven RC-network simulation model and filling in the parameters in that model by a combination of expert judgement and fitting of parameters to monitoring data.

In order to achieve this, we have started to set up monitoring in two types of houses, namely social rental houses that have been renovated by BAM to NOM level and new-build houses at NOM level built by van Wijnen in Ermelo. The purpose of the monitoring was to be able to fit the parameters in the model as well as possible. With a good fit, we can ultimately make an assessment of the cause in the event of disappointing energy consumption. Another, somewhat conflicting goal is to make parameter fitting possible with as few sensors as possible, i.e. to be able to estimate parameters well enough on the basis of less data or to be able to generate more insight based on the same amount of sensors.

Conclusions and lessons learned on data quality
Apart from these objectives, good data quality is important in order to be able to make statements based on monitoring. That is why we started monitoring the data and analyzing the data quality. The most important lessons from data quality monitoring and analysis are 1) that to assess the quality of a sensor's data, it is important to know what the sensor measures and where the sensor is placed, which in practice is not always clear and 2) good agreement is required between the data being monitored and the parameters used in the model. In TKI Optimaal first basic algoritms were developed to detect possible incorrect data points. In this project the analysis of the data quality were still project specific. The next step is to develop algorithms to analyze the data quality automatically and how to cope with poor data

and irregularities. The type of sensor and the location where the sensor will be installed will have to be taken into account in these algorithms.

Conclusions and lossons learned on model and parameter fitting based on monitoring data
Parallel to the process of collecting the monitoring data, we've set up a data-driven RC-network simulation model: a model of the individual houses that we will fit on the monitoring data. What the research has shown is that it is feasible to fit a model to monitoring data and to arrive at a good reflection of the actual energy consumption, hourly temperature progression and energy signature. We reached these results despite the fact that the spread in user behavior of residents and neighbors in particular leads to large variations in these factors. On the other hand, we reached these results thanks to the fact that we have been able to map out this behavior through monitoring and surveys. So we can conclude that setting up data-driven RC-network simulation models of the first NOM houses has been a successful first step in order to be able to arrive at a realistic analysis of the performance guarantee of individual houses. To eventually be able to explain at least 80 to 90% of the deviation between the predicted and actual energy and indoor climate performance of individual NoM houses, the big challenge will be to determine the parameters in the model with more certainty and to use fewer sensors than at present. This applies especially, but not exclusively, to the behavioral parameters.

Based on the results of the model, we can learn about renovation concepts that 1) in very well insulated homes, the effect of windows that are opened slightly is significant even when in rooms that are unheated, due to the equalizing effect of the insulation and the heat recovery system, 2) in houses with low temperature systems ventilation losses can result in the situation that the system reaches it maximum capacity, resulting in temperature drops, 3) in very well insulated homes, the effect of the neighbors' heating behavior is also significant, 4) the effect of realistic variations in user behavior in these houses is significant, namely many tens of percentages to sometimes over one hundred percent per aspect.

Next steps
There are a number of methods that will help us to get more certainty about the parameters in the model. Some of these methods are already being concretely developed in follow-up projects, or will be taken up in future proposals:

- Using a probabilistic model to get more certainty about the parameter estimation: to get insight into the effect of the uncertainty of all parameters in the RC-network simulation model, we are working on a probabilistic model. Instead of using estimated values in the model, we use probability curves based on literature sources or directly derived from data.
- Using Artificial Intelligence (AI) to predict parameters from measured data: We have done a first study to see whether we could predict the use of windows and doors based on the measured data. The results were promising: we were able to predict if a window was open or closed with an accuracy of 80% for all hours of the year. This prediction was done for 2 of the houses in Ermelo with an algorithm that was trained by 2 other houses in Ermelo. We plan to expand the study next year and also study where AI (hybrid models) have an added value to models solely based on physics.
- Using fault diagnoses to infer if building components or systems malfunction: In previous projects, we focused on fault diagnoses based on monitoring data, but

mainly in non-residential buildings. These techniques can also be adapted to houses; what are common faults and what are typical patterns in monitoring data due to these faults.

- Using parameter identification by grey box modelling: Fitting parameters using grey box modelling is a technique that combines physical models with statistical models. The technique is proven for models with only a few parameters. A research question is whether the technique will work in more complex models.
- Using physical models to estimate parameters: One of the key parameters in a building model is the building mass. Estimating the building mass from night set back profiles is relatively easy for office buildings and older houses with clear temperature drops at night. However, for Net-Zero houses this proves more difficult since the temperature drop at night is quite small. There might be other ways to do this, e.g. by looking for holiday periods or using free floating temperatures in summer.
- Using a better model to estimate ventilation flows: a parameter with a large influence is the ventilation flow due to open windows. By coupling a more detailed ventilation model (COMIS) to the RC-network, we might be able to estimate this component more realistically. With this module added to the RC-network, it will also be possible to take into account the actual indoor air quality performance in addition to actual energy performance and actual thermal comfort performance.

In addition, we will have to make a step in the automation of all parts of the process: the check of the data quality and the data repair, the fitting of the parameters, the fault diagnosis and the performance test. These are next stept that will be taken in future projects.

And last but not least, more reseach is needed to find out what we can learn from data to improve renovation concepts, and espessially how renovation concepts influence behaviour and how that influences the performance of the concept. In the IEBB project under the MMIP of 2020, research on this topic will start.

Contents

1 Introduction

Background
In order to drive forward the energy transition, stakeholders are looking for ways to accelerate the construction and renovation towards highly energy-efficient buildings that go far beyond the current regulatory bottom-line requirements. To stimulate interest from owners and tenants, construction companies and other suppliers of deep retrofitting solutions have started to give guarantees on the energy performance of these houses and want to do the same for comfort and health related aspects in the future.
With these initiatives, a need has risen for methods that can test these performances in practice. For various reasons, the performance may be disappointing. The person who provides the guarantee, sometimes for 25 years, can be held accountable if the guarantee is not met. This is justified if a disappointing performance is caused by structural or technical installation defects. In practice, however, we see that user behaviour also has a major impact on the performance of a house. In any case, we want insight into the cause of disappointing performances: if an installation is not properly set up or tuned, you want to send an installation engineer to the house as soon as possible, and if a resident's behaviour is the cause of high energy consumption, you may want to send an energy coach. Also insight is needed in user interaction to select and improve concepts that perform better during actual use, using such insights in a proactive way to optimise design, operation and utilization. That's why we want to develop a method with which we can determine at an individual level the actual performance of houses that are in use.
Purpose of the project
The aim of the TKI Optimaal project is to develop models and algorithms for data analysis with which, in time, at least 80 to 90% of the deviation between the predicted and actual energy and indoor climate performance of individual NoM houses can be explained.

General approach
The route we are following is to develop a data-driven RC-network simulation model of NOM houses that will allow us to approach the actual performance of those houses on an individual level. There are many definitions of a data-driven RC-network simulation model, but what we mean by it in this context is a physical model of the performance of a dwelling that approximates actual performance by being tuned using monitoring data. The rationale behind this is that if a model succeeds in approximating the actual performance of a dwelling, it is also clear which aspects determine the performance of that dwelling. After all, these aspects are represented by the parameters in the model. If the actual performance is worse than the predicted performance, it is therefore possible to find an explanation for the deviation with a data-driven RC-network simulation model.

The development of a data-driven RC-network simulation model of NOM houses involves roughly two steps: 1) the development of the model and 2) the best possible estimation of the parameters in the model. On the one hand, the model will have to contain the parameters that will make it possible to approximate the actual performance and, on the other hand, the parameters will have to fit in as well as possible with reality. And of course there is an interaction between the two: a too simple model will not contain the parameters to explain deviations. And a model with too many parameters gives too many solving possibilities to be able to determine the cause of deviations, especially if we presort on a large roll-out and thus keep a limitation in monitoring data in mind.

Concrete steps taken in TKI Optimaal
TKI Optimaal is a "TKI toeslag" project with the aim of building knowledge and laying the foundation of the data-driven RC-network simulation models as described above. The project focuses exclusively on a data-driven RC-network simulation model of the energy performance of NOM houses. The focus is mainly on space heating. To this end, we have taken the following steps in the project:

1. Monitoring, including data quality - to be able to develop a data-driven RC-network simulation model we need good quality monitoring data. Therefore the project started with the development and execution of a monitoring plan and a methodology to check the data quality and repair the data where needed.
2. Model design – The development of the data-driven RC-network simulation model itself started with the setup of a first energy model, that was (and will be) refined in an iterative process.
3. First estimation of parameter values – All parameters in the model were filled based on building information and expert best guesses.
4. Improved fit on the basis of monitoring data – A better fit of some of the parameters were made based on a fit with the monitoring data.
5. Sensitivity analyses - Initial sensitivity analyses were done on some of the parameters to improve certainty and to show the effect of variations in some of the parameters.
6. Lessons learned - Experiences in this project have led to insights where improvements can be made when using the data-driven RC-network simulation model as basis for a performance guarantee. This has led to the implementation of a number of follow-up projects that are currently underway and from which we have positive expectations for the results.

Domestic hot water use
The project also paid attention to a prediction of the use of domestic hot water: the energy use for domestic hot water is relatively high in NOM houses compared to the low energy use for space heating in the highly insulated houses. With a flow meter on the warm water, the prediction of this energy use can be improved. In this project we investigated if a less expensive solution, a temperature sensor, could be an alternative. The findings are described in a separate report.

Reading guide of the report
The first three chapters, namely chapters 2 to 4, describe the monitoring that has taken place in order to be able to set up the basic models for the data-driven RC-network simulation models. Chapter 2 describes the dwellings monitored in the study. Chapter 3 provides an overview of the sensors used in the monitoring and Chapter 4 describes how the data quality was tested, which repairs were carried out and which lessons were generally learned about data quality and data repair for future role out.

Chapter 5 discusses the development of the first data-driven RC-network simulation model. This chapter describes which model was used, how the initial estimation of the parameters took place and how the model was initially fitted to the monitoring data. Some initial sensitivity analyses were also carried out. On the one hand to gain more certainty about the estimation of a few of the parameters in the model, but also to show the effect of possible spread off a few influential parameters on the energy consumption of the NOM houses concerned.

Parallel to the development of the data-driven RC-network simulation model, one of the NOM houses was also modelled in a detailed building simulation model (TRNSYS). During the development of the data-driven RC-network simulation model, we saw that some of the behavioral parameters have a major impact on energy performance. We wanted to investigate this further. The results are described in chapter 6.

Finally, chapter 7 summarizes all the insights from the research and describes what we are working on to get the data-driven RC-network simulation model a step further and which developments we are working on.

2 Dwellings

In this project a total of 10 houses was monitored in two different neighbourhoods.

2.1 Emmen

The first 4 houses are located in Emmen in the north-east of the Netherlands. This is a neighbourhood with social housing. These houses were originally built around 1970. Currently these houses are renovated to NoM (energy neutral) houses. Within the NOM concept; the houses are renovated to become net-zero-energy using an industrialized (pre-manufactured) concept that includes a full refurbishment of the thermal shell, installations and the deployement of local generation. Glassing is replaced, a new façade is placed and additional isolation is applied on the roof. Energy efficient installations are included in an energy module that provides heating (air/water heatpump), fresh air (RCU), storage (Sensible/Water) and connectivity. Solar panels, heat recovery system and a heat pump are installed. Each dwelling has 2 floors. The living room and the kitchen are located on the ground floor. On the first floor, 3 bedrooms and the bathroom can be found. The dwelling has a flat roof.

Figure 2-1 Dwellings in Emmen

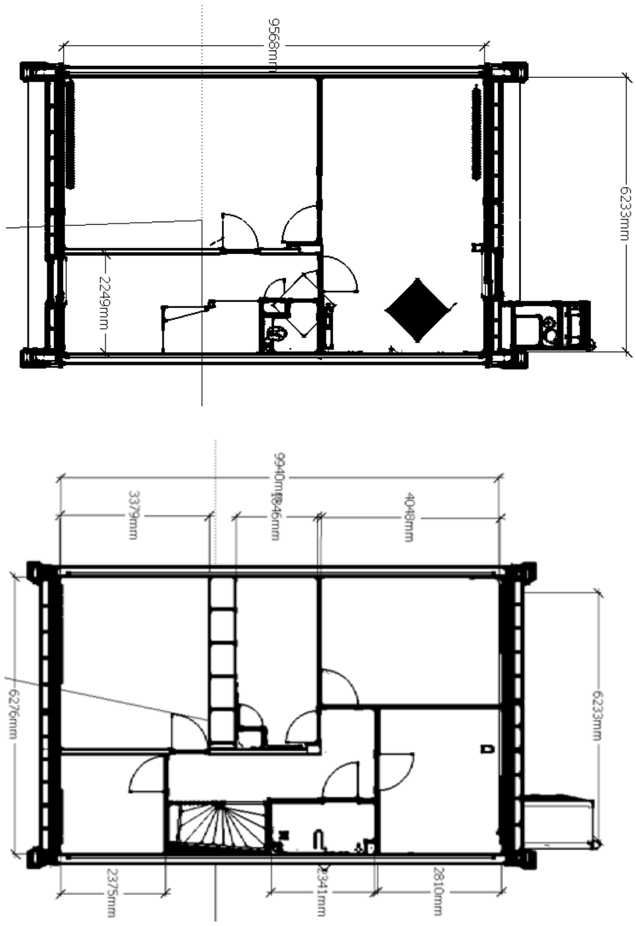


Figure 2-2 Ground plan of the Emmen dwelling

2.2 Ermelo

The other 6 houses are located in Ermelo in the middle of the Netherlands. This is a new neighbourhood with private houses. These houses are newly built and also NoM houses. The houses have a sloped roof which starts at the first floor and ends at the attic. On the roof solar panels are installed. A heat recovery system is installed in these houses, in the ventilation channels between the heat recovery and the blow-in points two heaters are installed to heat up the air. One heater for the ground floor and one heater for the first floor. Hot water comes from an electrical boiler.

Figure 2-3 Dwellings in Ermelo

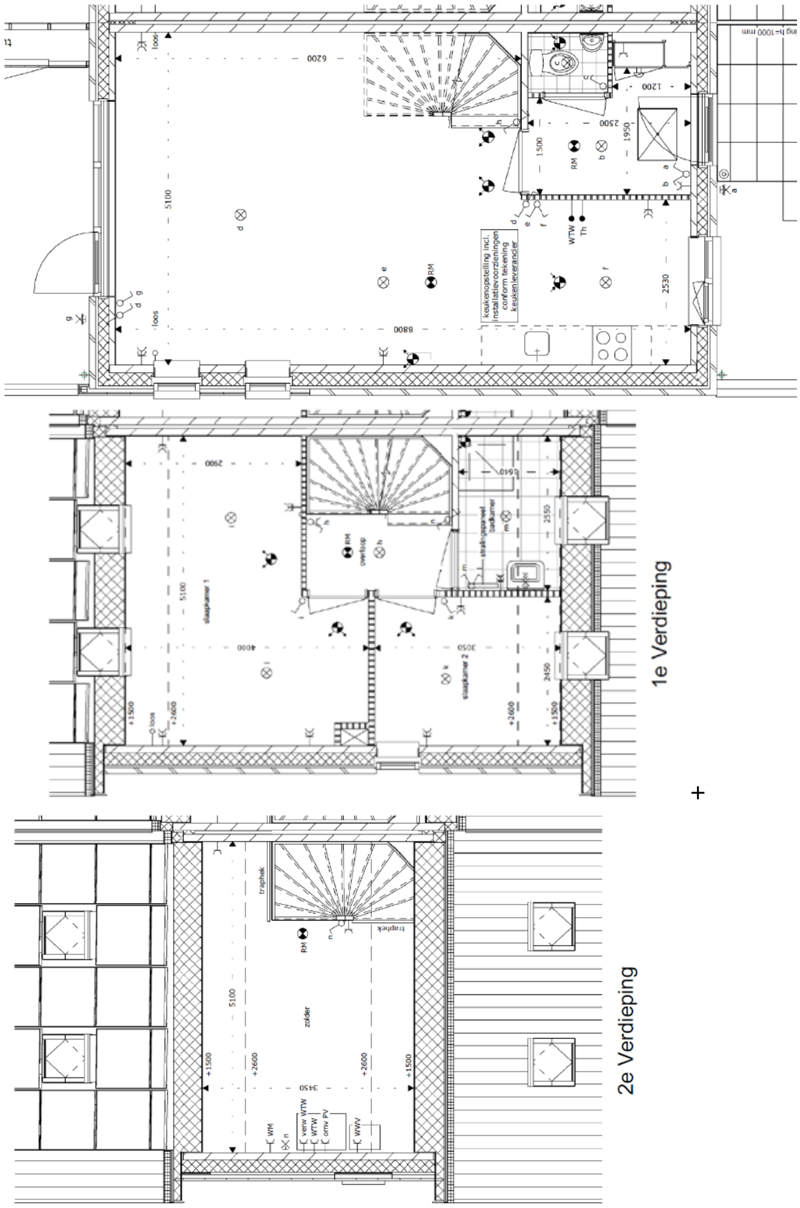


Figure 2-4 Ground plan of the Ermelo dwelling

2.3 Specifications of the dwellings

Table 2-1 Building and installation properties

Building and installation properties	Ermelo	Emmen
Area	98 m^2	107 m^2
Building type	Middle to heavy building mass	Middle to heavy building mass
Heating		
• Source	Electrical heater in ventilation channel	Air-water heat pump
• Afgifte systeem	Air heating	Radiators (traditional)
• Details	Each floor has a thermostat. In the bathroom a radiation panel is installed	Thermostat in the living
Tap water		
• Installation	Electrical Boiler	Air-water heat pump with a 200 liter Boiler
Ventilation		
• Type	Mechanical supply and extract air	Mechanical supply and extract air
• Heat recovery	95%	95%
• CO2 controle	None	None
• Details	Ventilation rate (flow/volume) over the entire house, to be arranged from the kitchen in 3 positions. In addition to manual adjustment, the flow rate is also controls by the heating: if the heating is switched on at the bottom or at the top floor, the ventilation flow rate is set to the maximum.	Ventilation rate (flow/volume) over the entire house, to be arranged from the kitchen in 3 positions. In addition in the bathroom a timer switch is installed.
• Windows/grill	Windows	Windows
Shell		
• Rc floor/facade/roof (m2K/W)	8 / 8 / 11 Head end Rc = 9	5 / 4,7 / 5
• U window/door (W/m2K)	0.8 / 0.8	1,1 / 1,1
• Infiltration qv,10 (ltr/sm2)	0.150	0,39
• Solar shading	None (except the demo house, which has	Shading glass at the ground floor and on the

	shading on two windows on the 1ste floor.)	1st floor at the windows on the south.
• facade with the neighbours	12cm sand-lime brick; 6cm air cavity; 12cm sand-lime brick	Approximately the same as in Ermelo
Solar Power		
• PV	36.4 m^2 (South)	37m^2 up to 39,6m^2
• PVT	None	None

3 Sensors

In each of the dwellings a set of sensors was installed to monitor the energy use, the temperature, but also if doors and windows are open or closed. In the table below all sensors are shown.

Table 3-1 Overview of sensors in the dwellings

Type	Location	Emmen	Ermelo
Temperature	Living	X	X
Temperature	Kitchen	X	[1]
Temperature	Bedroom 1	X	X
Temperature	Bedroom 2	X	X
Temperature	Bedroom 3	X	X
Temperature	Attic		X
Temperature	Tap water	X	X
Setpoint	Living	X	X
Setpoint	1st floor	X	
Energy	Heat pump	X	
Energy	Electrical Heater 1		X
Energy	Electrical Heater 2		X
Energy	Heat recovery	X	X
Energy	Solar panels	X	X
Energy	Smart Meter	X	X
Energy	Electrical Boiler		X
Status	Heat pump	X[2]	
Open/Close	Front door	X	X
Open/Close	Back door	X	X
Open/Close	Kitchen window	X	X
Open/Close	Living window	X	
Open/Close	Bedroom 1 window	X	X
Open/Close	Bedroom 2 window	X	X
Open/Close	Bedroom 3 window	X	X
Open/Close	Bathroom window		X
Open/Close	Attic window		X
Flowrate	Tap water	X	X

[1] In Ermelo the kitchen is part of the living room
[2] The status of the heat pump indicates if it is in use for space heating or for tap water

Most of the sensors are battery powered and communicate via a wireless protocol, the Z-wave protocol. The Z-wave router is connected to a 4g modem, which upload the data to the server.

4 Data quality

While the previous chapter shows all the sensors which are installed in the dwellings, the data cannot be directly used as input for the models. First all data needs to be check on the data quality. In this chapter data quality is discussed. The chapter is build up in three part. First the causes of problems with data quality are discussed. Second the types of incorrect data are shown and as a final part some basic algorithms are shown to detect incorrect datapoints. This chapter is specific for this monitoring project, but many parts are generic and can also be used in other projects. At the end of the chapter there is a session with lessons learned and with conclusions.

4.1 Causes of problems with data quality

4.1.1 Type of sensor

4.1.1.1 Battery powered

Sensor which are battery powered can end up with an empty battery. If the battery is not replaced, no data will be available.

4.1.1.2 Wireless sensor

For most wireless sensors, the distance between the sending and receiving part is given in the free field. However in practice this distance is much smaller and influenced by wall and floor. Beside of these two factors, also metal and furniture plays a role. Therefore the communication can be disturbed for some sensors.

4.1.1.3 Status and open/close

Sensors of this type, only send their data when the value is changed. If this value is not changed for a long time it is hard to define if the data is correct or not. The sensors for open/close exists out of two parts, if one of the parts is aligned incorrect or is fallen down, the sensor will not give the correct reading.

4.1.1.4 Resolution

The resolution of a sensor will influence the results from any calculation or simulation. In this project, the sensors measuring neighbour temperature have a 0.5 degree resolution. But also measurements on kWh and volume (m^3) have a resolution which is sometimes to large. The effect of a relative large resolution is show in the figure below. The blue line has a resolution of 0.1 degree, while the red line has a resolution of 0.5 degree. It can be observed, that a temperature change of 0.1 degree can lead in one case to an overestimation of 0.4 degree, and in another case has no effect at all.

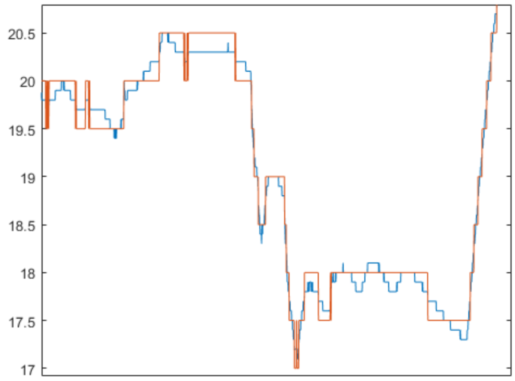


Figure 4-1 Example of the effect of resolution on the measured temperature

4.1.1.5 *Accuracy*

The accuracy of a sensor will influence the results from any calculation or simulation. For this project the best temperature sensors are selected with a measured standard deviation of 0.1 degree.

4.1.2 *Location of sensor*

When a temperature sensor is located on a spot near the window, where the sun can shine on the sensor the measured temperature can be much higher than the actual temperature. This can also happen when the sensor is located near to a heat source. When a temperature sensor is mounted on or very close to the wall, the measured temperature will be influenced by the temperature of the wall. In the figure below the temperature of the air is shown in red. And the measurement of the thermostat, which is wall mounted, is shown in purple.

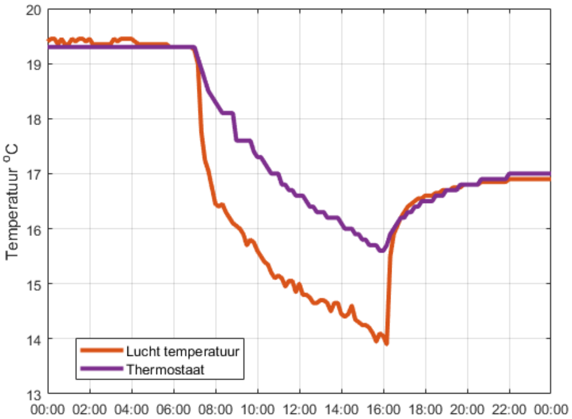


Figure 4-2 Example of the influence of the wall temperature on the temperature reading

4.1.3 *Infrastructure*

The infrastructure between the sensor and the location of data storage plays an important role in the data quality, especially in missing data. In this project the sensors communicate wireless via the Z-wave protocol with the Gateway. The gateway is connected to a 3/4g modem. The gateway communicates with the servers of BeNext. On the servers of BeNext the data is stored. If there is any communication error/failure in the chain no data will be received and stored. During the project there were problems with the 3/4g modem, but also failures in the mobile network.

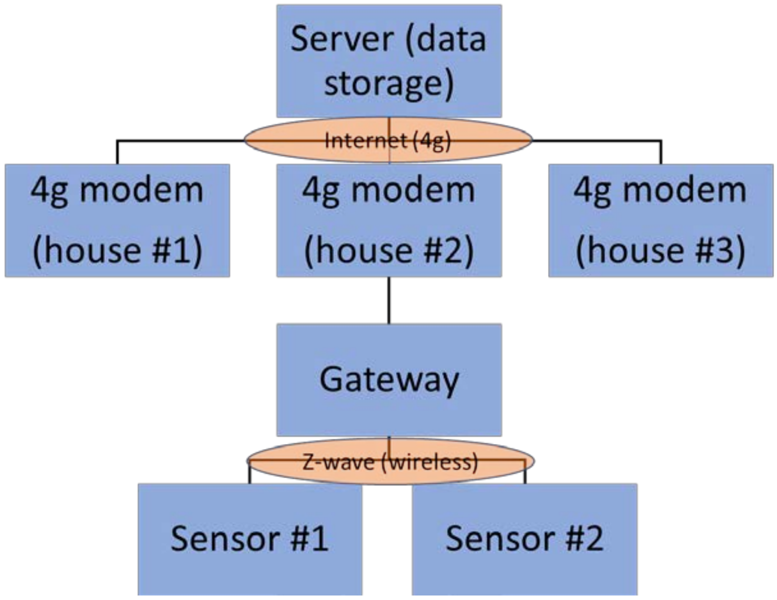


Figure 4-3 Overview of the infrastructure

4.2 Types of incorrect data

4.2.1 Spikes

Spikes are datapoints which are much higher or lower than the previous or next datapoint. These higher or lower values should not be part of a trend. For example a trend is when a sudden temperature rise is followed by a couple of datapoints which also show a rising temperature The datapoints with a spike are random much higher or lower than the measurement point around it.

In the figure below an example of a spike. In this case it shows the energy production from the solar panels in kW.

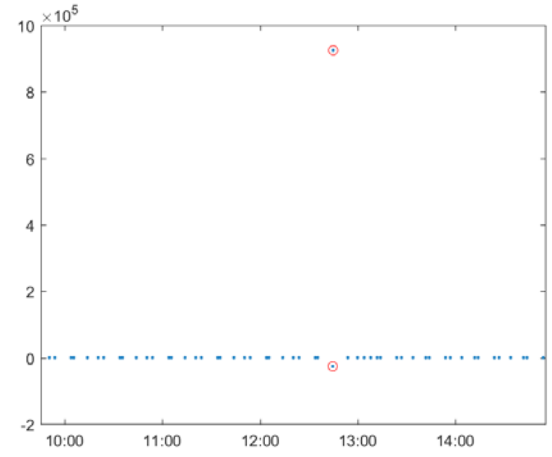


Figure 4-4 Example of a spike

4.2.2 *Gaps*

Gaps in simple words is missing data. It is one or more measurement point which are missing in the dataset. The effect of these gaps depends on the measured quantity, therefore the definition of gaps is different for each sensor. Unfortunate gaps cannot be defined for the on/off type sensors.

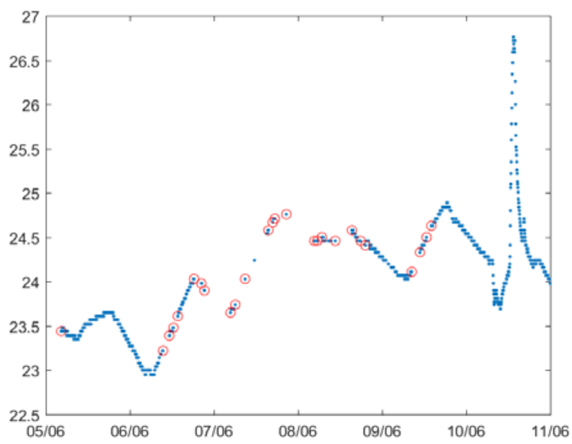


Figure 4-5 Example of data with gaps

4.2.3 *Frozen*

Frozen data means that the value of the measured data stays the same for a period longer than expected. Also this period is different from sensor to sensor. This period depends on the type of sensor, and it location. How quick will a sensor react, or how often will something be used. Beside this also the resolution of the sensor will influence the period. For example if you measure temperature with a resolution of 0.1 or 0.5 degrees. Very small fluctuations will not be "detected" by the sensor with the highest resolution. In the figure below an example. The figure shows the cumulative energy use for domestic hot water. In August the energy use is zero for longer than a day. And for several days in a row. This can indicate frozen data, or as in this case people which are on holidays.

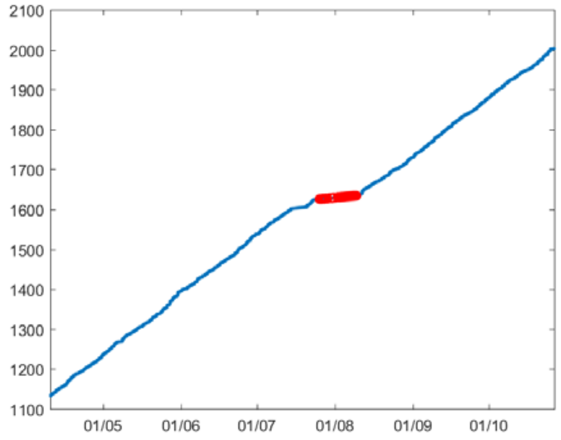


Figure 4-6 Example of data with frozen data

4.2.4 Absolute value

For a measured quantity a minimum and maximum value can be defined. This value should have a logical (physical) background. Also those values are different for each type of sensor and for their purpose. In the figure below an example. The figure shows the temperature in the living room. Temperature above 35 degrees in summer are already odd, but temperature up to 40 degrees where measured in November. In this case the temperature sensor was near an electrical heater. When the heater is switched on the temperature rises very quickly, the opposite happens when the heater is switched off.

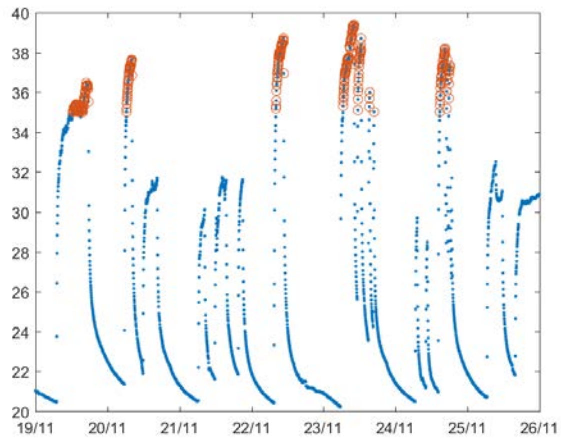


Figure 4-7 Example of temperature sensor which shows unrealistic high values

4.2.5 Criterium

For each sensor location and sensor type criteria has been defined. criteria of when measured data is frozen, has a gap or exceed the absolute values.

For example the temperatures measured in the different rooms of the house. These sensors have a 0.1 degree resolution therefore it is expected that this value will change relative quick. After some investigation in the data we can up with a value of 4 hours (240 minutes) for frozen data. The sensors send their data every 10 minutes and in our simulation we run on a 1 hour time interval, therefore a period of 1 hour is chosen to define a gap. For the absolute boundaries 10 and 35 degrees are chosen.

Also the temperature at the neighbours is measured, but with a sensor with a resolution of 0.5 degree, therefore the criterium for frozen data has been extended to 1 day instead of 4 hours.

For all measurement on the electrical consumption/production data the criterium for frozen data is set to 1 day. Solar panels for example should produce some energy every day.
For the use of tap water a period of 10 hours is used for the frozen data. Almost no use during the night or when the occupants are out for work.

In the tables below the criteria are given for the gap, frozen and absolute values.

Table 4-1 Criteria for gap and frozen data

	dT frozen [minutes]	dT gap [minutes]
Temperature	240	60
Power	1440	60
Flow	600	60
Status	840	60
Volume	1440	60
Neighbour Temperature	1440	60

Table 4-2 Criteria for the absolute value

		Low Criterium	High Criterium
Heat pump	Setpoint Temperature	5	30
Heat pump	Consumed energy kWh	0	3000
Heat pump	Operation Mode	-0.1	6.1
Tap water	Temperature	5	65
Tap water	Volume [l/min]	0	15
Heat recovery	Consumed energy kWh	15	55
Living	Temperature	10	35
Kitchen	Temperature	10	35
Master Bedroom	Temperature	10	35
Bedroom	Temperature	10	35
Bathroom	Temperature	10	35

4.3 Algorithms to detect incorrect data

4.3.1 Spikes

Below a simplified version is shown of how the algorithm works to detect spikes. As a first step the difference between each measurement value and the previous is calculated. From all measurement points with a delta not equal to zero the standard deviation is calculated. The next step is to compare all data points with the 3σ (Sigma) criterium (99% interval). For each datapoint which is out of this range it is checked if it is part of a trend. If the datapoint is part of a trend it will not be marked as spike. A trend means that the previous and next datapoint have the same direction and some order of magnitude (factor 3). As a last check. A spike should also have a “contra” spike If there is no “contra” spike it means that the value will remain much higher of lower for a longer period or will slowly return towards its original value.

```
This function will return a true/false array for each datapoint.
function [true_false_array] = spikes(value)

    calculate the difference between the values
    DeltaValue = diff(value)

    find all differences not equal to 0
    id = find(abs(DeltaValue>0))

    if length(id)>1
    T:
        calculate the 3s boundaries
        MaxDelta = 3*std(DeltaValue(id))

        find values which are outside the bounds
        id1 = 1+find(abs(DeltaValue)>border)

        determine trend of previous datapoint
        Trend1 = DeltaValue(id-1)/DeltaValue(id-2)

        determine trend of next datapoint
        Trend2= DeltaValue(id)/DeltaValue(id-1)

        Check if spike is part of a trend
        TrendCheck = (trend1>0.3)&(trend1<3)|(trend2>0.3)&(trend2<3)

        remove spikes which are part of a trend
        id = id(find(TrendCheck == 0))

        check for each spike if there is a contract spike within the next 5 datapoints
        for nn = 1:1:length(id)

            calculate the time difference between two spikes
            diff_spikes = diff(id(nn-1:nn+1))

            outcome is zero if no contra spikes are in the data
            ContraCheck = sum(diff_spikes<5)

            remove spikes without a contra spike
            id = id(find(ContraCheck==0))
    F:
        no spikes are present
        id_spike = []
```

create the output array, with only good(1) data
```
true_false_array = ones(size(value))
```
add spikes (0) to the array
```
true_false_array(id) = 0
```

4.3.2 Gaps

Below a simplified version is shown of how the algorithm works to detect gaps. As an input it requires, the time axis (time of each measured value), the measured values, but also the given interval in minutes. This last parameter is the criterium. If there is no data for a period that is longer than the criterium it is indicated as a gap. As a first step the time difference is calculated between each measurement point. As a second step all datapoints are marked for which the delta is larger than the criterium. Next also the latest datapoints before a gaps are marked. The total numbers of days (some of all gaps) and the percentage of time are calculated.

this function returns a true/false array for each datapoint. As well as the number of gaps, number of days and the percentage of time
```
function [gaps nr_days perc_time true_false_array] = gap(time_axis,value,interval_min)
```
Calculate DeltaTime in minutes
```
DeltaTime = diff(time_axis*24*60)
```
find when the interval between to datapoints is larger then the given interval
```
id = find(DeltaTime>interval_min)
```
create array with start and end time of gaps
```
gaps =[];
```
create the true/false array with good data(1)
```
true_false_array = ones(size(value))
```
```
for nn=1:1:length(id)
```
add start time of the gap to the array
```
    gaps(nn,1) = time_axis(id(nn)+1)
```
add end time of the gap to the array
```
    gaps(nn,2) = time_axis(id(nn)+2)
```
add the gap to the true/false array
```
    true_false_array(id(nn)+1:id(nn)+2) = 0
```
Calculate the total number of days which are missing
```
nr_days = sum(gaps(:,2)-gaps(:,1))
```
Calculate the percentage of time with missing data
```
perc_time = 100*(1-nr_days/(time_axis(end)-time_axis(1)))
```

4.3.3 Frozen

Below a simplified version is shown of how the algorithm works to detect frozen data. As an input it requires, the time axis (time of each measurement), the measured values, but also the given interval in minutes. This last parameter is the criterium. As a first step the difference between each measurement is calculated. All values which are the same as the previous value are marked. As a next step it looks for the begin and end of periods of frozen data. For each period the duration is calculated and compared with the criterium. All data point within a frozen period are marked as frozen.

```
This function returns a true/false array for each datapoint
function [true_false_array] = frozen_data(time_axis,value,interval_min)

    calculate the difference between each measurement
    DeltaValue = diff(value)

    find all datapoint which are equal to the previous datapoint
    id_nochange = find(DeltaValue==0)

    find the end of a period of data with the same value (frozen data)
    id_endoffrozen = find(diff(id_nochange)>1)

    check if there is (potential) frozen data in de dataset
    if ~isempty(id_enoffrozen)
    T
        determine the end of a frozen period
        frozen_end = id(id2)+1

        determine the begin of a frozen period
        frozen_start = id([1 id2(1:end-1)+1])+1

        calculate the duration of the period
        dT = 24*60*(time_axis(frozen_end)-time_axis(frozen_start))

        check if period is longer than specified
        id_frozen = find(dT > interval_min)

        remove to short periods from this array
        frozen_start = frozen_start(id)

        remove to short periods from this array
        frozen_end = frozen_end(id)

        create true/false array, with all true
        true_false_array = ones(size(value))

        for nn = 1:1:length(frozen_start)
            set true/false array to false for periods with frozen data
            true_false_array(frozen_start(nn):frozen_end(nn)) = 0
    F
        true_false_array = ones(size(value))
```

4.3.4 Absolute value

The basics of this algorithm is very simple. For each datapoint it is check if it is within the specified limits. If the datapoint is outside the limits it is marked.

```
function [true_false_array] = abs_value_check(t,y,limits)

for nn = 1:2:length(limits)
    Check if data is equal or larger than the lower limit
    check_low = (y≥limits(nn))

    Check if data is equal or smaller than the higher limit
    check_high = (y≤limits(nn+1))

    create an array for the lower limit
    check(nn,:) = check_low.*check_high

    create an array for the lower limit
    check(nn+1,:) = check_low.*check_high

create the true/false array
true_false_array  = check(1,:)&check(2,:)
```

4.4 Lessons learned

Before starting a new project an investigation should be made about the type, location and communication of the sensors. This to be sure that the data has the best data quality possible. For example the measured value of the temperature sensor is influenced by the location. When the sensor is mounted on the wall the temperature dynamics are damped by the mass of the wall.

The fact that data is only stored on the server leads to more (smaller) gaps. Or a complete loss of data for a longer period when there is a failure in the network.

Status / open/close sensors only sent a value when changed this makes it really difficult to perform checks on data quality.

Every sensor (type and location) has its own criteria. In this project this values are determined by hand. But this can be improved in the future.

4.5 Conclusion

During this project a start is made for the analysis of data quality. First basic algoritms were developed to detect possible incorrect data points. The work done is project specific. However many parts are generic and can be used as a starting point for new projects. During the project it was investigated what causes there are for problems with data quality. What kind of incorrect data is present in the datasets. As a third step some basic algorithms are developed to detect incorrect datapoint in the measured dataset.

5 Development of the data-driven RC-network simulation models

The energy and indoor climate performance of Zero on the Meter (NoM) houses is sometimes disappointing. At present, the available monitoring tools and data analysis methodologies are unable to identify the causes of deviating performance. We aim at bridging this gap by developing data-driven RC-network simulation models of the individual NoM houses: if a model succeeds in approximating the actual performance of a dwelling, it is also clear which aspects determine the performance of that dwelling. After all, these aspects are represented by the parameters in the model. If the actual performance is worse than the predicted performance, it is therefore possible to find an explanation for the deviation with a data-driven RC-network simulation model. This chapter described the methodology used to develop these data-driven RC-network simulation models. Subquestions were: what model is suitable as a basis for a data-driven RC-network simulation model, how do we determine the parameters in the model and when are we sure enough that this combination of model and parameters is indeed a good representation of reality so it can predict the performance well enough for its purpose.

To answer these questions and to develop the data-driven RC-network simulation models, a basic setup of the model was made and a first estimation of the parameters was done. Paragraph 5.1 gives the general description of the basic RC model of the data-driven RC-network simulation models, the explanation about why we chose this model and about various parameters/inputs required for the model, results obtained by using the RC model and the so-called verification of the model by comparing the model results with the measured data. To improve the certainty of the model, paragraph 5.2 describes a sensitivity analysis, to investigate the influence of some of the parameters in the model on the energy consumption of the dwellings. This chapter concludes with lessons learned in general and with respect to the research questions.

5.1 Modelling Approach

The development of a data-driven RC-network simulation model of NOM houses involves roughly two steps: 1) the development of the model and 2) the best possible estimation of the parameters in the model. In the following paragraphs we'll start with outlining the parameters of the dwellings that form the basis for the model. This is partly a repetition of chapter 3. We'll then go into the model itself and give a first estimation of the parameters. We'll finish the paragraph with a check how good the first parameter fit actually is.

5.1.1 General Description of Dwellings and building parameters

As described in chapter 2 and 3, in TKI Optimal we monitored 4 dwellings located in Emmen and 6 dwellings located in Ermelo. For the development of the data-driven RC-network simulation models we started with 2 of the houses in Emmen of which we had the most and qualitative best data. In this report we will call these houses 'Emmen 224" and "Emmen 228". Figure 5-1 shows the general floor plan of the first floor[1]. The first floor consists of living room, kitchen and the entrance. Figure 2-2

[1] In the Netherlands it is called as "ground floor".

represents the floor plan of the second floor[2]. This floor has 4 bedrooms, one storage room, one bathroom and the hallway.

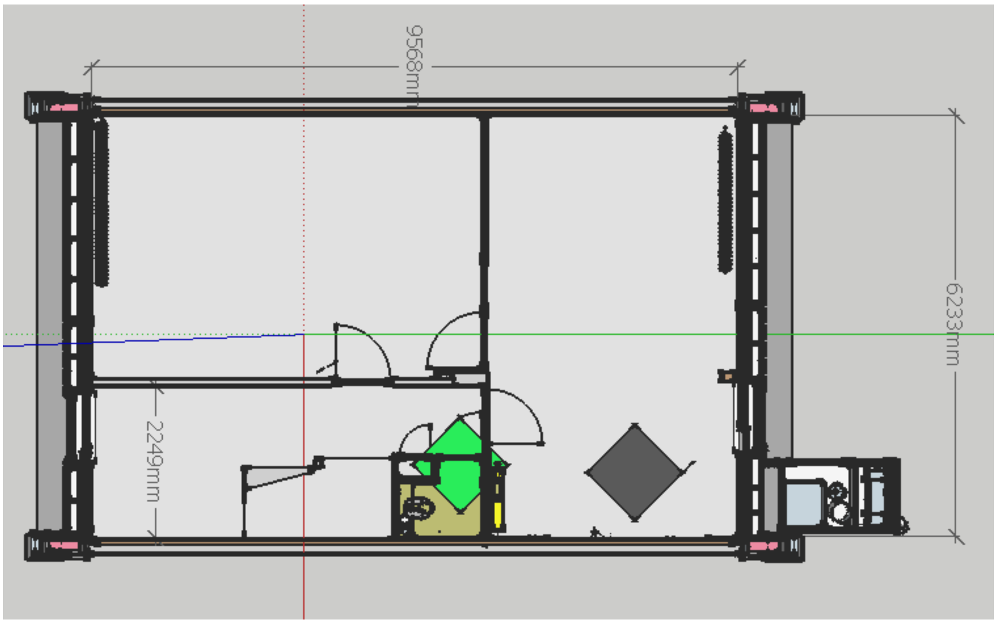


Figure 5-1. Floor Plan for the First Floor

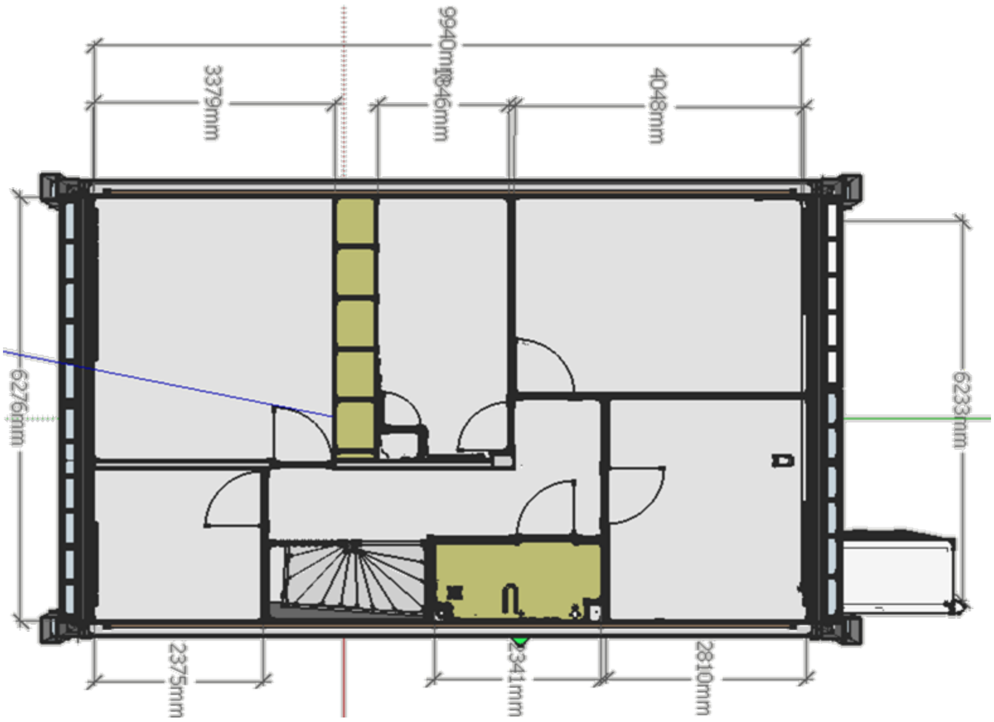


Figure 5-2. Floor Plan for the Second Floor

[2] In the Netherlands it is called as "first floor".

Table 5-1. and Table 5-2 show the parameters used in the RC model for floor 1 and floor 2. Floor areas are derived from Figure 5-1 and Figure 5-2. The Rc and U values are obtained from Table 2-1. There is no exact information about the façade area and window area. Therefore, these values are derived from building pictures.

Table 5-1. The Parameters of Floor 1

Parameters – First Floor	Unit	Emmen 224	Emmen 228
Floor Area	[m^2]	60	60
Floor Height	[m]	2.7	2.7
Facade Area - Total	[m^2]	33.7	33.7
Facade Area (Opaque Part only)	[m^2]	19.6	19.6
Window Area - Total	[m^2]	14.1	14.1
Window Area - South	[m^2]	0	0
Window Area - West	[m^2]	7.3	7.3
Window Area - North	[m^2]	0	0
Window Area - East	[m^2]	6.8	6.8
Rc - Façade (Opaque Part only)	[m^{2}K/W]	4.7	4.7
Rc - Floor	[m^{2}K/W]	5	5
Rc - Roof	[m^{2}K/W]	5	5
U value - Window	[W/m^{2}K]	1.1	1.1

Table 5-2. The Parameters of Floor 2

Parameters – Second Floor	Unit	Emmen 224	Emmen 228
Floor Area	[m^2]	60	60
Floor Height	[m]	2.7	2.7
Roof Area	[m^2]	60	60
Facade Area - Total	[m^2]	33.7	33.7
Facade Area (Opaque Part only)	[m^2]	24.7	24.7
Window Area - Total	[m^2]	9	9
Window Area - South	[m^2]	0	0
Window Area - West	[m^2]	4.8	4.8
Window Area - North	[m^2]	0	0
Window Area - East	[m^2]	4.2	4.2
Rc - Façade (Opaque Part only)	[m^{2}K/W]	4.7	4.7
Rc - Floor	[m^{2}K/W]	5	5
Rc - Roof	[m^{2}K/W]	5	5
U value - Window	[W/m^{2}K]	1.1	1.1

5.1.2 RC Network and Mathematical Model

We chose a hourly 3-zone RC network model as a basis for the data-driven RC-network simulation model of the dwellings in Emmen. This choise was made because the data-driven RC-network simulation model needs to do justice to the dynamics of reality on the one hand, but not contain too many parameters on the other hand. With a model that is too simple, such as a monthly model or a one-zone model, it becomes very difficult to determine the cause of an anomaly at the individual level of a single household: the effect of that anomaly must be clearly reflected in the model. A model with too many parameters, such as a TRNSYS model for example, has too many knobs to turn on and therefore too many uncertain components, all of which have to be tuned to reality.

The advantage of modelling 3-zones in the dwelling is that you can make a distinction among various spaces that are heated and used in a diffirent way. The 3 zones were originally devided over the house as followed: zone 1 is the heated living room and kitchen (the whole first floor), zone 2 are all heated spaces on the second floor and zone 3 are all unheated spaces on the second floor. However, since in both houses all rooms on the second floor were either all heated or all unheated, the distinction between zone 2 and 3 was finally made based on orientation and by using a rather arbitrary border (see Figure 5-3). The border was chosen arbitrary to keep the model general. The 60% of the total floor area of the second floor is considered as Zone 2 and the rest is Zone 3.

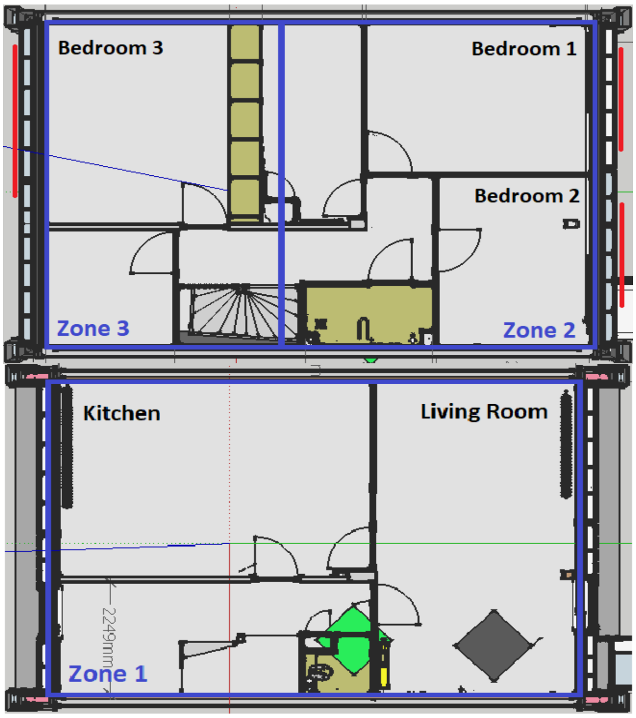


Figure 5-3. The Presentation of 3 Zones in the dwellings in Emmen

The RC network and mathematical model are described for Zone 1 only in this section, otherwise the text becomes unreadable. However, Zones 2 and 3 are almost exact copies, except from the ground floor. All energy flows through Zone 1 are indicated in Figure 5-4. The red dashed rectangle in this figure is used to distinguish describe the inner and outer thermal mass.

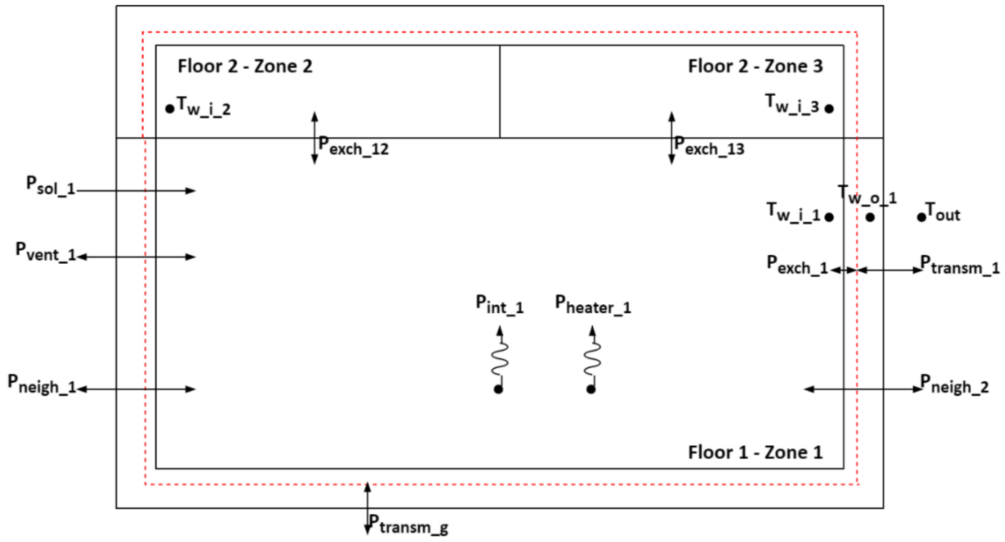


Figure 5-4. Energy Flow Through Zone 1

The RC network in Figure 5-5 is derived based on the energy flow through Zone 1. The full RC network drawing is given for Zone 1 only. The RC network is repetitive for Zone 2 and 3 except for the part of the ground floor.

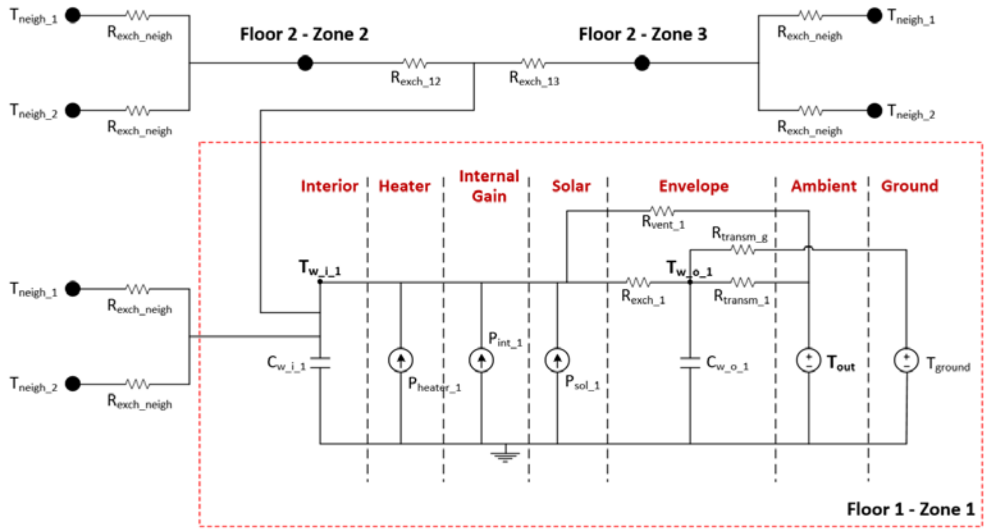


Figure 5-5. RC Network of 3 Zones The dwellings in Emmen

The RC-network has two state variables for each zone, one describing the interior temperature T_{w_i} ($T_{w\_i\_1}$ for Zone 1), which is the lumped temperature of indoor air and the first layer of the wall and one representing the temperature of the building envelope T_{w_o} ($T_{w\_o\_1}$ for Zone 1). The first-order dynamics are represented by the stochastic differential equations below for Zone 1.

$$\frac{\mathbf{dT_{w\_i\_1}}}{\mathbf{dt}} = \frac{\mathbf{1}}{\mathbf{C_{w\_i\_1}}}\begin{pmatrix} -\mathbf{P_{exch_1}} + \mathbf{P_{vent_1}} + \mathbf{P_{sol_1}} + \mathbf{P_{int_1}} \\ +\mathbf{P_{exch\_12}} + \mathbf{P_{exch\_13}} + \mathbf{P_{neigh\_1}} + \mathbf{P_{neigh\_1}} \end{pmatrix}$$

$$\frac{\mathbf{dT_{w\_o\_1}}}{\mathbf{dt}} = \frac{\mathbf{1}}{\mathbf{C_{w\_o\_1}}}\left(\mathbf{P_{transm\_1}} + \mathbf{P_{transm\_g}} + \mathbf{P_{exch\_1}}\right)$$

where

$$\mathbf{P_{exch\_1}} = \mathbf{UA_{exch}}(\mathbf{T_{w\_i\_1}} - \mathbf{T_{w\_o\_1}})$$
$$\mathbf{P_{vent\_1}} = \mathbf{UA_{ventnasis\_1}}(\mathbf{T_{out}} - \mathbf{T_{w\_i\_1}})$$
$$\mathbf{P_{sol\_1}} = (\mathbf{g\ value})(\mathbf{f_{shading}})(\mathbf{Q_{sol}})(\mathbf{A_{glass\_1}})(\mathbf{Solar\ Fraction})$$
$$\mathbf{P_{transm\_1}} = \mathbf{UA_{transm\_1}}(\mathbf{T_{out}} - \mathbf{T_{w\_o\_1}}) + \mathbf{UA_{transm\_g}}\left((\mathbf{T_{ground}} - \mathbf{T_{out}})/2 - \mathbf{T_{w\_o\_1}}\right)$$
$$\mathbf{P_{exch\_12}} = \mathbf{UA_{exch\_12}}(\mathbf{T_{w\_i\_2}} - \mathbf{T_{w\_i\_1}})$$
$$\mathbf{P_{exch\_13}} = \mathbf{UA_{exch\_13}}(\mathbf{T_{w\_i\_3}} - \mathbf{T_{w\_i\_1}})$$
$$\mathbf{P_{neigh\_1}} = \mathbf{A_{neigh\_1}}(\mathbf{UA_{neigh}})(\mathbf{T_{neigh\_1}} - \mathbf{T_{w\_i\_1}})$$
$$\mathbf{P_{neigh\_2}} = \mathbf{A_{neigh\_2}}(\mathbf{UA_{neigh}})(\mathbf{T_{neigh\_2}} - \mathbf{T_{w\_i\_1}})$$

The state equations above do not include $\mathbf{P_{heater\_1}}$ because of the methodology which is followed in the model. As seen in Figure 5-6, first the temporary temperatures are calculated for next time step by state equations. After that, the required heating powers for each zone are calculated based on setpoint temperatures and temporary temperatures. Then the heating capacity is evenly distributed based on the heating needs in each zone. Finally, the temperatures in the next time step are calculated based on adjusted heating powers for each zone.

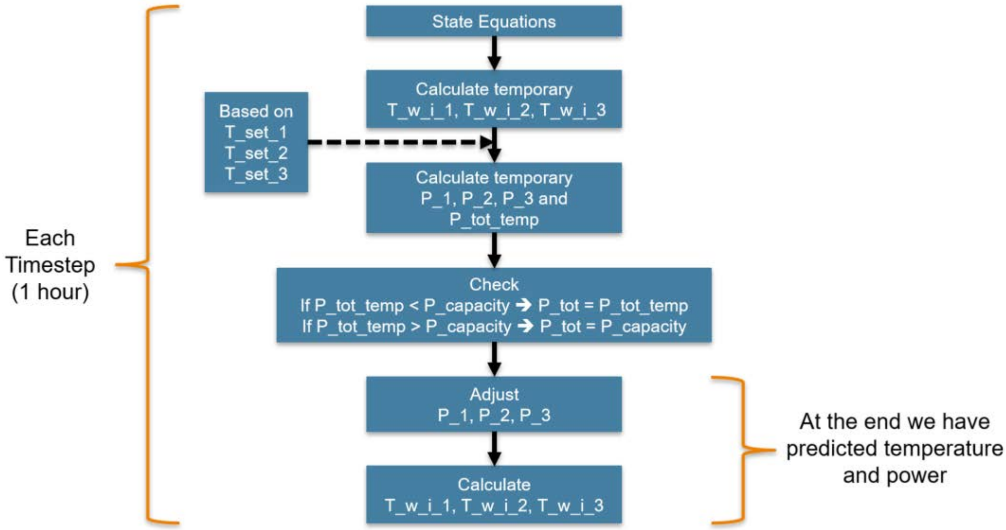


Figure 5-6. Methodology Used in the Model

5.1.3 Parameters and boundary conditions of RC Model

This section describes most parameters and boundary conditions used in the model such as heating setpoint, weather conditions, heating systems, internal heat gains, ventilation system and the thermal mass. The building parameters were already given in paragraph 5.1.1.

5.1.3.1 Temperature Profiles

The following measured temperatures (hourly data) are used in this model:

- **Indoor temperature:** There are 4 temperature sensors used that measure the temperature of the indoor environment (1 in the living room and 3 in the bedrooms). These measured temperatures are compared with the temperatures that are calculated by the model as a mean to check the accuracy of the model.
- **Set point temperature:** The thermostat temperature in the living room is used as the heating setpoint temperature.

- **Neighbor temperature:** Basically this is the indoor temperature in the neighbor's living room. It is used to calculate the heat loss or gain towards or from the neighbor/s.
- **Outdoor temperature:** The outdoor temperature is obtained from a KNMI weather station. The nearest KNMI weather station to the dwellings is in the Hoogeveen (STN:279, LON(east):6.574, LAT(north):52.750, ALT(m):15.80) which is approximately 25-30 km away from the dwellings.
- **Ground floor temperature:** Ground temperature data is not available in the KNMI weather data file. Therefore T_{ground} is assumed as 10 °C and constant during the all year (although it varies from 7 °C to 13 °C during the year). However, the dwellings have a crawlspace which is natural ventilated. Therefore, the average of ground and outside temperature ((T_{ground} +T_{out})/2) is used in the model to calculate the transmission loss through the ground floor.

5.1.3.2 *Thermal Mass*

The basic idea in our model is that out of the total thermal mass of a building, only a relatively small part (typically between 15-35 % for a single-family dwelling)[3] is able to effectively exchange heat with the indoor environment to affect diurnal variations in indoor temperature. The inner part of the total thermal mass typically consists of the indoor walls, floor, roof, etc., with a thickness of up to a few cm of material plus most of the furniture. This part of the thermal mass is called the indoor mass (C_{w_i}). The remaining part of the thermal mass of the building is called the outdoor mass (C_{w_o}).

The total thermal mass of the building is calculated from the volume and density of building materials. For ease of calculation, the simplified generic relations between indoor mass (C_{w_i}) and building volume (V) is: $C_{w\_i} = a * V^b$ where a is 0.18 and b is 0.92[3]. The fraction between the indoor mass and the total building mass is assumed as 0.3 for a medium level insulated the single-family house, as described in the article[3] .

5.1.3.3 *Heating System*

The space heating in the dwellings is provided by an air source heat pump (PUHZ-SW50VHA). Table 5-3. shows the heating capacity and the Coefficient of Performance (COP) of the heat pump which is taken from the datasheet[4] for the nominal operating conditions at 45 °C outlet temperature of the water.

Table 5-3. Heat Pump Specification

Tamb [C]	Capacity [kW]	COP	
-15	3.15	1.46	
-10	4	1.77	
-7	4.4	1.98	
2	5	2.47	
7	6	3.32	
12	7.07	3.63	
15	7.54	3.89	
20	8.04	4.19	
25	**8.54**	**4.49**	**Extrapolated**

[3] Koene, F.G.H. et. al. (2014), Simplified building model of districts, fifth German-Austrian IBPSA conference
[4] For PUHZ-SW50VHA model, which can be accessed by using https://planetaklimata.com.ua/instr/Mitsubishi_Electric/Mitsubishi_Electric_Ecodan_Data_Book.pdf

30	**9.04**	**4.79**	
35	**9.54**	**5.09**	

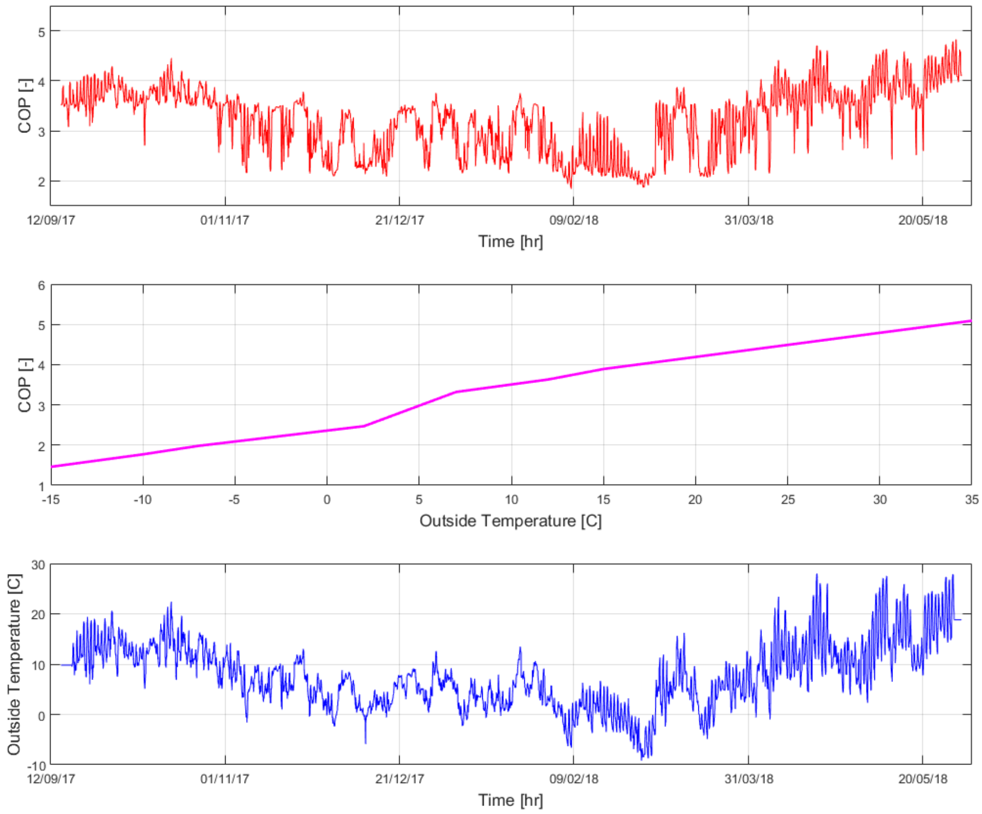


Figure 5-7. COP of Heat Pump

Figure 5-7 represents the variation of COP with the variation in the ambient air temperature in the second plot. Also it represents the time-dependent COP data in the first plot which is used for calculating the heat pump thermal power.
Figure 5-8 shows the heat pump electricity consumption data and the time-dependent COP data in first plot. Then by using the formula ($\text{HP Power}_{\text{thermal}} = \text{COP} * \text{HP Power}_{\text{electric}}$), the thermal power of heat pump is obtained and it is represented in the second plot.

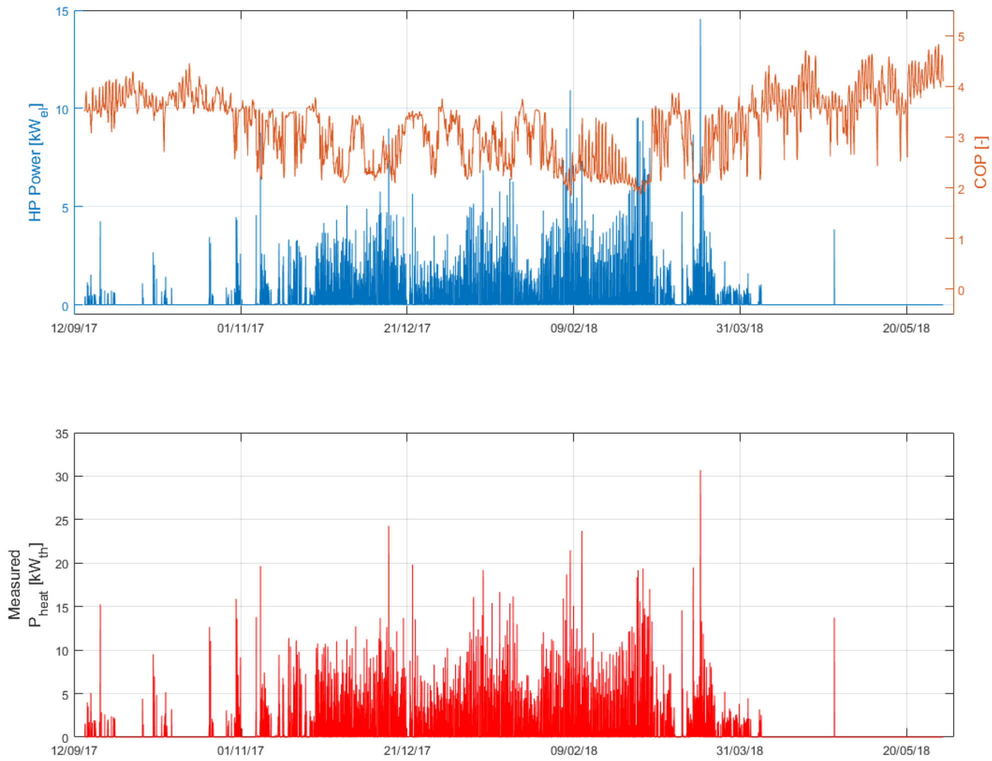


Figure 5-8. Heat Pump Thermal Power

The maximum heating capacity of the heat pump is 5 kW as stated in the data sheet[5] and it is assumed that the capacity is evenly distributed over the zones based on the heating needs of each zone. The explanation of the iterative progress that is used to calculate the heating power per zone is given in Table 5-4..

Table 5-4. Heating power calculation per zone

	Zone 1	Zone 2	Zone 3
Heating power (initial calculation) $P_{tot}= P_1+ P_2+ P_3$	P_1	P_2	P_3
Distribution ratio	$D_1= P_1/ P_{tot}$	$D_2= P_2/ P_{tot}$	$D_3= P_3/P_{tot}$
If $P_{tot}< P_{max}$	$P_{tot\_new}= P_{tot}$		
If $P_{tot}>= P_{max}$	$P_{tot\_new}= P_{max}$		
Distributed heating power	$D_1*P_{tot\_new}$	$D_2*P_{tot\_new}$	$D_3*P_{tot\_new}$

5.1.3.4 *Solar Gain*

As described in section 5.1.2, the solar gain ($\mathbf{P_{sol}}$) is calculated by the formula below.

$$\mathbf{P_{sol}} = (\mathbf{g\ value})(\mathbf{f_{shading}})(\mathbf{Q_{sol}})(\mathbf{A_{glass}})(\mathbf{Solar\ Fraction})$$

Where $\mathrm{g\ value}$ is the glass transmittance, $f_{shading}$ is the shading factor because of external and internal solar blinds and the surrounding objects, (Q_{sol}) is the hourly solar radiation on horizontal surface which is available in the KNMI weather station

[5] For PUHZ-SW50VHA model, which can be accessed by using https://planetaklimata.com.ua/instr/Mitsubishi_Electric/Mitsubishi_Electric_Ecodan_Data_Book.pdf

(the nearest is Hoogeveen, STN:279), A_{glass} is an area of glass and the Solar Fraction is introduced in this equation to convert solar radiation on a horizontal surface to the vertical surface of the corresponding facade. The solar fraction and the shading factor used in the model are described in annex A.

5.1.3.5 *Internal Gain*

Internal heat gain is mainly because of the occupancy, lighting and the appliances, which are discussed as follow:

5.1.3.5.1 *People*

There is no sensor data available in both the dwellings in Emmen related to occupancy. Therefore the number of occupant and occupancy schedules are derived from the questionnaire, which can be found in annex B for both the dwellings. However, the distribution of occupancy to the zones is still unknown. Therefore, a new term, the occupancy fraction[6] per zone is introduced and available in Table 5-5., which is based on the assumption. Finally, the heat generation by occupants is assumed as 100 W per person during the day and evening and 70 W per person during the night.

Table 5-5. Occupancy fraction per zone

	Occupancy Fraction per Zone Zone 1 (Floor 1)	**Occupancy Fraction per Zone Zone 2 and 3 (Floor 2)**
Morning (08:00 - 13:00)	0.75	0.25
Afternoon (13:00 - 18:00)	0.75	0.25
Evening (18:00 - 23:00)	0.75	0.25
Night (23:00 - 08:00)	0.0	1.0

5.1.3.5.2 *Lighting and Appliances*

In order to incorporate the heat gain caused by the appliance, the net electricity consumption for each dwelling is determined by using the measured data of the electricity consumption from the grid, generation from PV panels and the surplus electricity fed to the grid from the PV panels.

In the first step, the total electricity consumption is calculated by using the following formula below.

$$\begin{aligned}\textbf{Total electric consumption }[\textbf{kW}] = (&\textbf{Total consumption from grid}\\ &+\textbf{PV generation} - \textbf{Returned to the grid form PV})\end{aligned}$$

Afterward, the electrical consumption of appliances, such as the TV, the stove, the washing machine etc. is derived by using the following calculation:

$$\begin{aligned}\textbf{Net electric consumption }[\textbf{kW}]\\ &= \textbf{Total electric consumption}\\ &- (\textbf{Electric consumption by the heat pump for water and space heating}\\ &+ \textbf{electric consumption by the ventilation and WTW heat recovery system})\end{aligned}$$

For most of the appliances, all electric energy consumed converts into heat but still the conversion factor is unknown . Moreover, most of the heat produced by the

[6] It is the fraction of total occupancy present during different time period of the day in a particular zone.

dishwasher and the washing machine goes to waste. Thereby, It is assumed that 90% of the net electricity consumption will generate internal heat gain. Moreover, it is also assumed that 60% of the net electricity is used on the ground floor, while only 40% is used on the first floor[7]. This results in:

$$\mathbf{P_{light\ and\ applianc\_\ ground\ floor}\ [kW] = Net\ Electric\ consumption * 0.9 * 0.6}$$
$$\mathbf{P_{light\ and\ appliance\_first\ floor}\ [kW] = Net\ Electric\ consumption * 0.9 * 0.4}$$

The $\mathbf{Net\ Electric\ consumption}\ [\mathbf{kW}]$ is negative for some timesteps as it can be seen in Figure 5-9. It is because all the energy sensors are not sending information at the exact same time to the data logger(timestep mismatch for different energy sensors) and because of that the net electric consumption is negative for some timesteps. However, the cumulative electricity consumption increases linearly as expected. It means the data is still reasonably useable. Therefore it has been decided to use the daily average values which are always higher than zero.

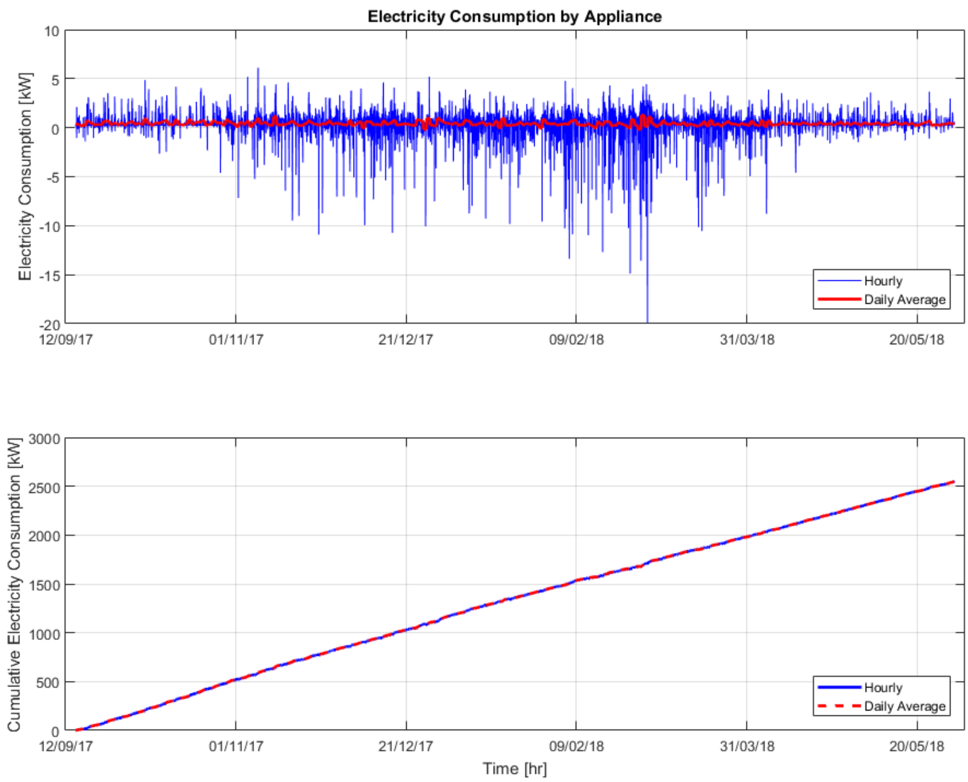


Figure 5-9. Electricity Consumption by Appliance

5.1.3.6 Ventilation System

5.1.3.6.1 Mechanical Ventilation

In the dwellings, the mechanical ventilation rate has been measured by flow finder for each position of ventilation. Table 5-6. shows the measurement results per zone. Resuming, zone 1 consists of the kitchen and the living room. Zone 2 contains 2 bedrooms, while zone 3 has only one bedroom..

[7] Based on expert best guess at time

Table 5-6. Flowrate (L/s) for the different ventilation positions (1/2/3) for all three zones of each building

	Emmen 224			Emmen 228		
	Stand 1	**Stand 2**	**Stand 3**	**Stand 1**	**Stand 2**	**Stand 3**
Zone 1	22	25	27	20	21	26
Zone 2	11.5	14.5	14.5	14	15	16.5
Zone 3	4	5	6	16	20	25

In Table 5-6., it can be easily seen that the mechanical ventilation rate varies depending on the position of the ventilation (positions 1/2/3). However, there is no direct sensor data available to obtain the ventilation position. Therefore measured data of the energy consumption of the heat-recovery/ventilation system (named: WTW_Vermogen_kWh) in Figure 5-10 (blue) is used to estimate the position (1/2/3) of ventilation in each zone. It is obvious in this figure that there are 3 different energy consumption levels which correspond to ventilation position. The ranges which includes these 3 different levels and their corresponding ventilation position are summarized in Table 5-7.. Figure 5-10 also shows how the position of ventilation is varying by changing the ventilation heat recovery (WTW), in kW, for Emmen 224. As it is already mentioned, this is derived from the WTW data by using the conditions as described in the Table 5-7.

Table 5-7. Ventilation position for the various ranges of the WTW_Vermogen_kWh

WTW_Vermogen_kWh	Position of ventilation
0.01-0.028 kW	1
0.029-0.043 kW	2
>0.043 kW	3

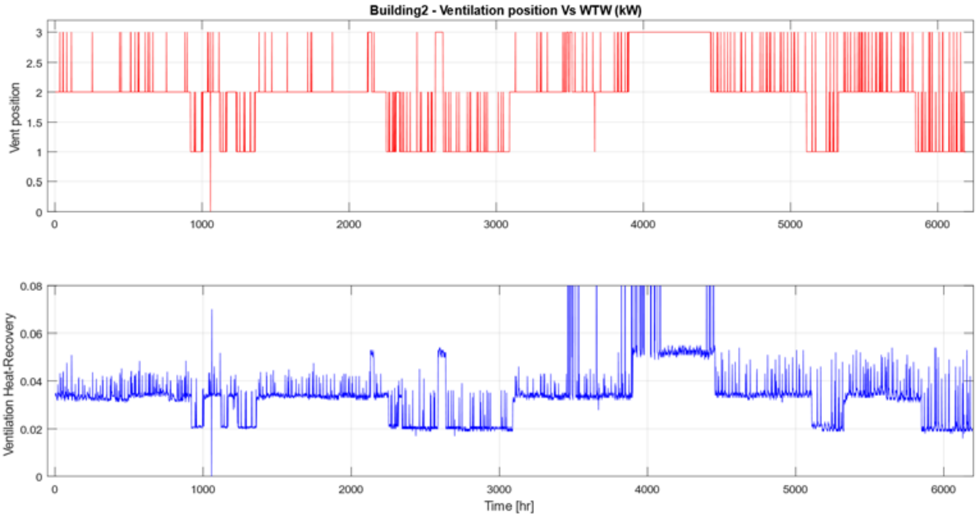


Figure 5-10. Ventilation position Vs WTW for Emmen 224

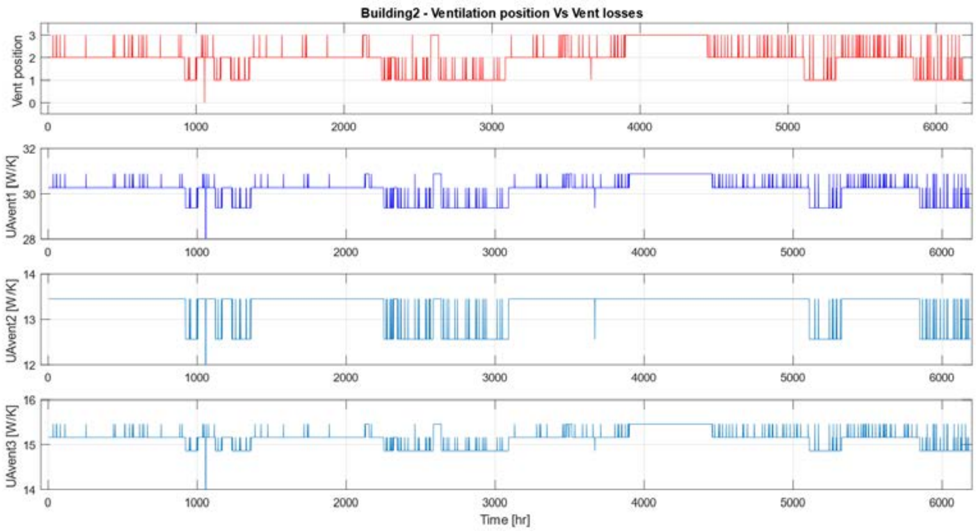


Figure 5-11. Ventilation position and UAvent for all 3 zones of Emmen 224

Figure 5-11 gives the variation in ventilation losses (UAvent) of all three zones by changing the ventilation position. It can be observed that ventilation losses are relatively higher at ventilation position 3 compared with position 1. This is because of the higher ventilation flowrate at position 3, which consequently increases the losses. Finally, the efficiency of the heat recovery unit is assumed to be 80%[8].

5.1.3.6.2 Ventilation via Open Windows and Doors

The methodology to estimate the flow rate through open windows and doors is given in Figure 5-12: firstly the theoretical pressure difference ($\Delta P_{theo}(t)$) between the indoor and the outdoor is calculated based on the wind velocity which is obtained from the KNMI weather station data.

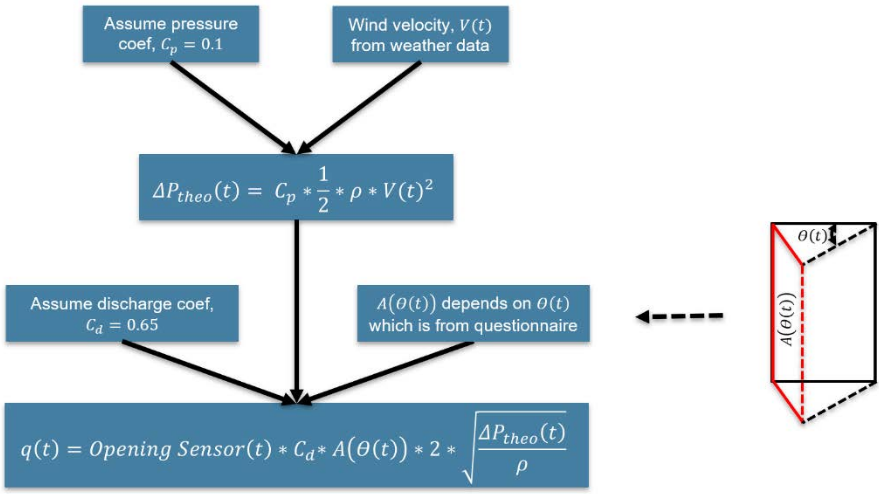


Figure 5-12. Methodology to estimate the flow rate through window and door

The time-dependent opening area ($A(\theta(t))$) is one of the biggest uncertainty in this model. In the dwellings in Emmen there are sensors on all windows and doors. An

[8] According to the building information document the efficiency of heat recovery system is 95% and the correction factor of 85% is used (it is assumed that for 80% of the time this efficieny is achieveable)

example of such a sensor is given in Figure 5-13. Thereby, it is possible to detect when window and door are closed (0) or open (1). The data from these opening sensors are given for Emmen 224 in the Figure 5-14. Sometimes the data is in between 0 and 1. For instance, if the sensor gives 0.5 for a particular timestep, it means the window/door is open only for half an hour.

Figure 5-13. Opening sensor

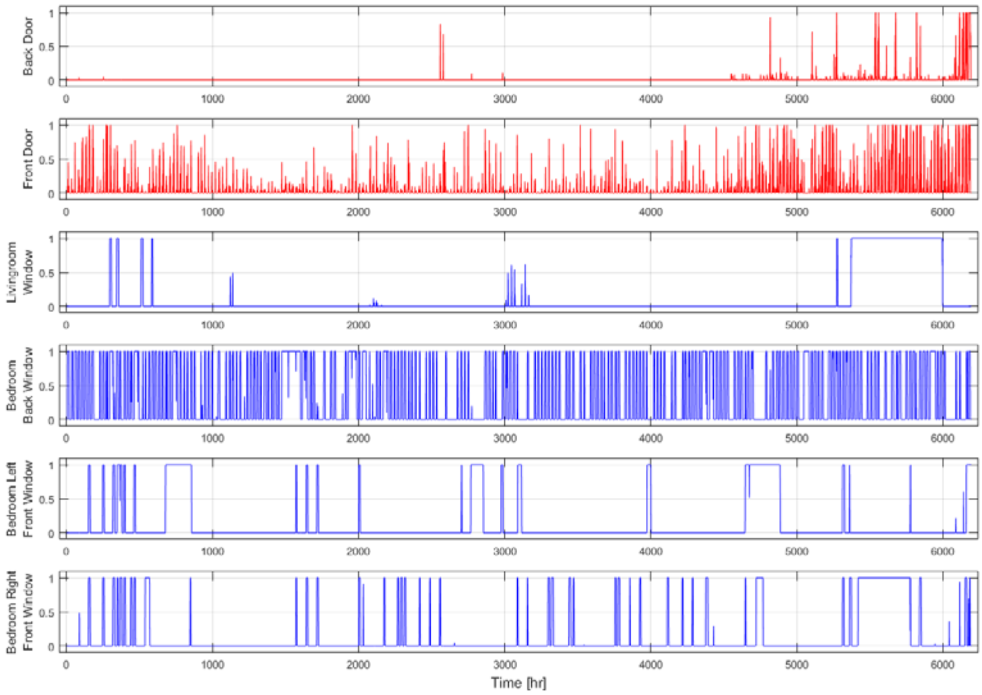


Figure 5-14. Window and door opening sensor for Emmen 224

5.1.3.6.3 *Infiltration*

Infiltration losses can basically be modeled by what we call the “Power Law”, which is the function of an air pressure differential across a building envelope and the flow characteristic of the shell:

$$q_V = C * \Delta P^n$$

where q_V is volumetric airflow, C air permeability coefficient, ΔP pressure difference and n flow exponent. Theoretically, the flow exponent n lies between 0, for fully developed turbulent airflows, and 1 for fully developed laminar airflows.
The measured infiltration rate (q_{V10}) for the dwelling is 1 L/s.m^2 at 10 Pa pressure difference. The q_{V10} is extrapolated back to more typical pressures inside the dwellings by using the following formula. The pressure difference inside the dwelling is assumed to be around 1 Pa and the n is assumed 0.5 in this model[9]. Although the pressure difference inside the dwelling changes over the year due to wind orientation and wind speed, it is assumed constant in this model.

$$\mathbf{q_{V1} = q_{V10} * \left(\frac{1}{10}\right)^{n}}$$

5.1.4 Results of RC Model fit on measured data

The RC model is run with a one-hour simulation time step for the whole measurement period, starting from 22 September 2017 until 28 May 2018. This period covers almost all the space heating period throughout the year. The temperature, energy demand and energy signature results for both dwellings are provided below. A suitable model should have a close agreement between the predicted and the measured data. Thereby, the results obtained from RC models such as indoor temperature, energy demand and the energy signature are compared with the corresponding measured data.
During the initial run of the model the results were not very promising, mainly because of the uncertainty in the ventilation losses. Later on, we tuned the windows opening fraction to match the temperatures, heating demand and energy signature with the corresponding measured data.

To investigate whether the parameter fit was successful we used 3 indicators: 1) the actual energy consumption, 2) the actual hourly temperature changes in the different zones for a whole year and 3) the energy signatures. Fitting the model on only one indicator is easy, but hardly gives any convidence whether the parameters in the model represent reality. If the model is ably to closly follow the 3 indicators, including the hourly pattern of the temperatures in the zones, this gives more confidence that the model actually represents reality. The following paragraphs describe the fit of the models for the 3 indicators: temperature, energy demand and energy signature.

5.1.4.1 Temperature

Figure 5-15 and Figure 5-16 represent the indoor temperature results of the dwellings in Emmen for a simulation with one hour time step, for 22 September 2017 until 28 May 2018. In these plots, the temperature results of RC model are compared to measured temperatures for all the three zones. The corresponding ambient temperature is also provided in the plots to analyse the temperature results. As we can observe, the predicted temperature for all 3 zones are following the measured temperature for both dwellings. In case of Emmen 224 all zones are heated while Zone 2 and 3 are not heated in Emmen 228.
As we can see in the following figures, there is somehow a close agreement between the predicted and measured temperature (for most of the observation period), which implies that the model is capable to predict the actual temperature inside the dwelling.

[9] Based on expert best guess at time

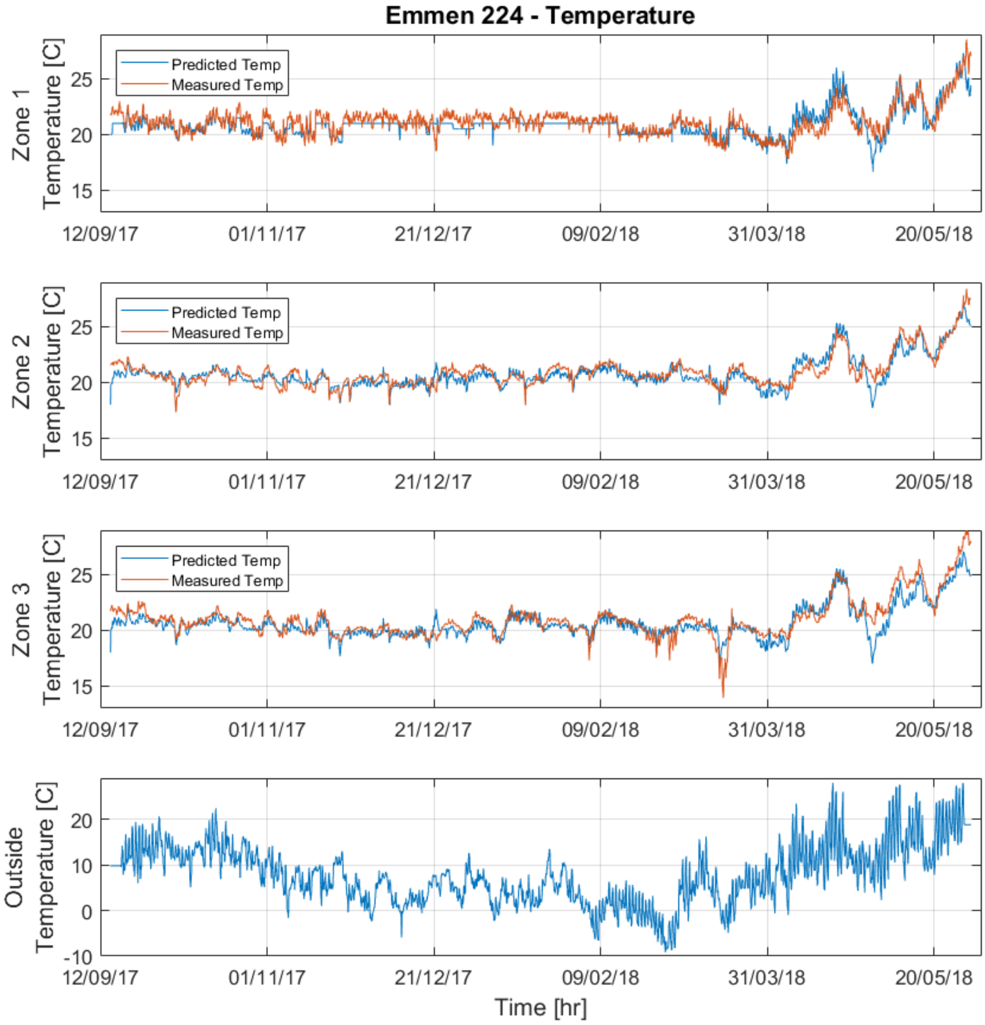


Figure 5-15. Temperature results for Emmen 224

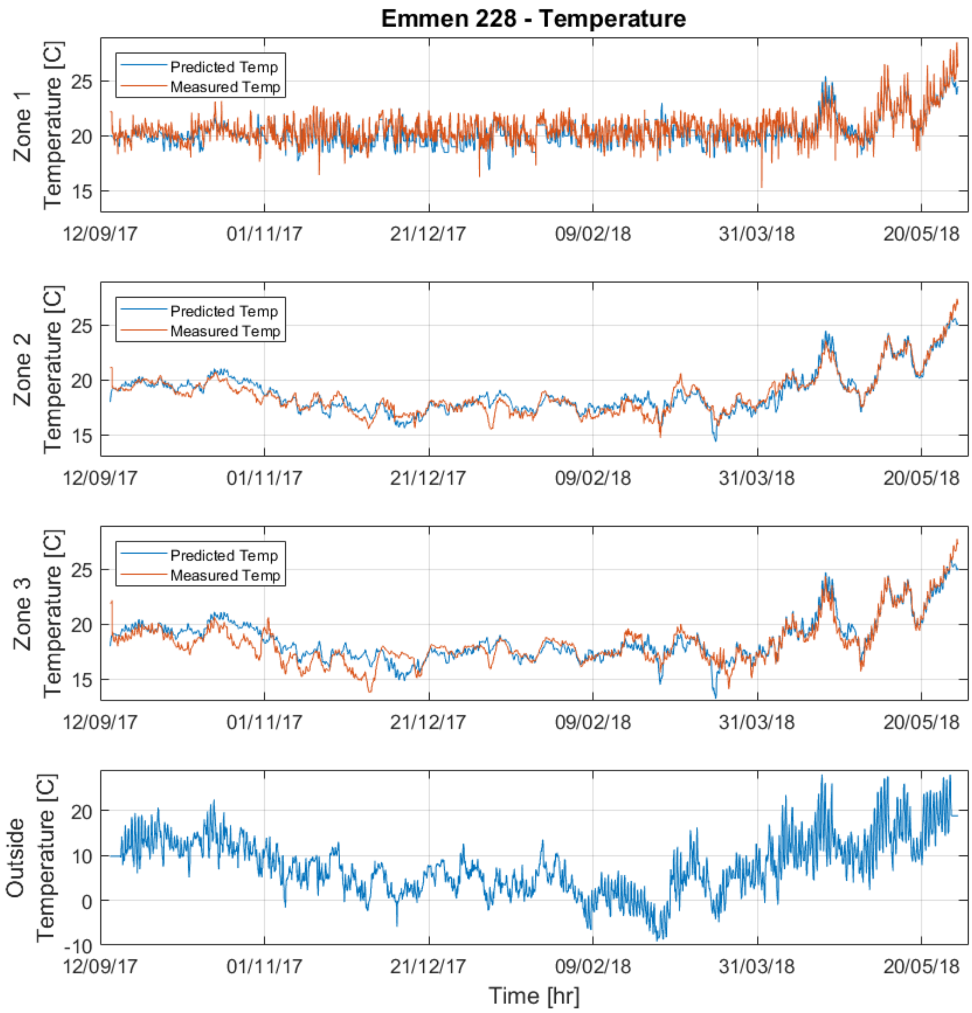


Figure 5-16. Temperature results for Emmen 228

5.1.4.2 *Energy Demand*

Energy demand is one of the key performance indicator, which provides the opportunity to evaluate the RC model. Comparing the predicted energy demand with the measured energy provides the so-called verification of the model.

Figure 5-17 and Figure 5-18 represent the cumulative heating demand results of the dwellings in Emmen for the above-mentioned time period. In these plots, the results of RC model (Predicted P) are compared to the measured heating demand of dwellings.

Since the measurement period starts in September, the heating required at the beginning of the simulation period is low. The RC model for Emmen 224 provides the total heating demand of 4149 kWh, which is comparable with the measured heating demand (4165 kWh).

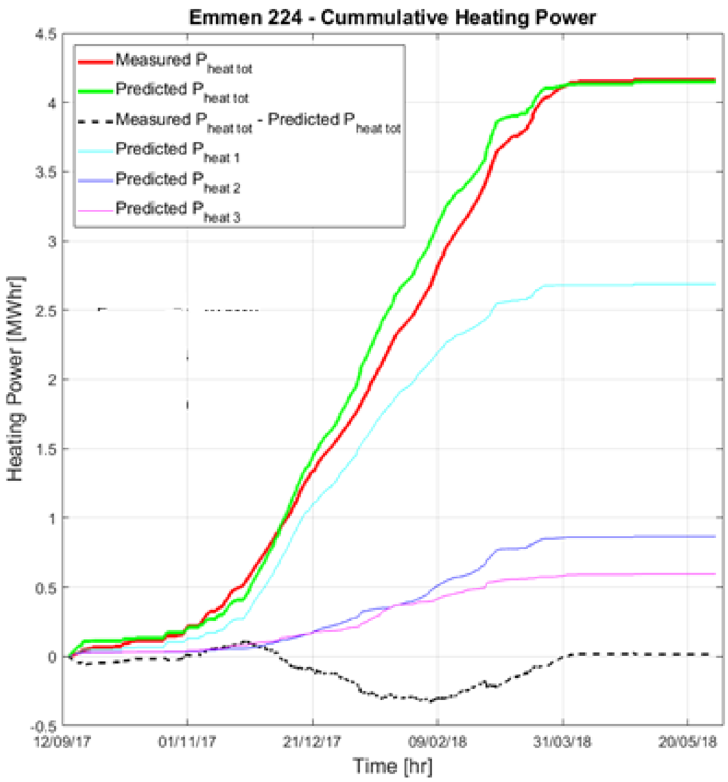


Figure 5-17. Heating demand results for Emmen 224

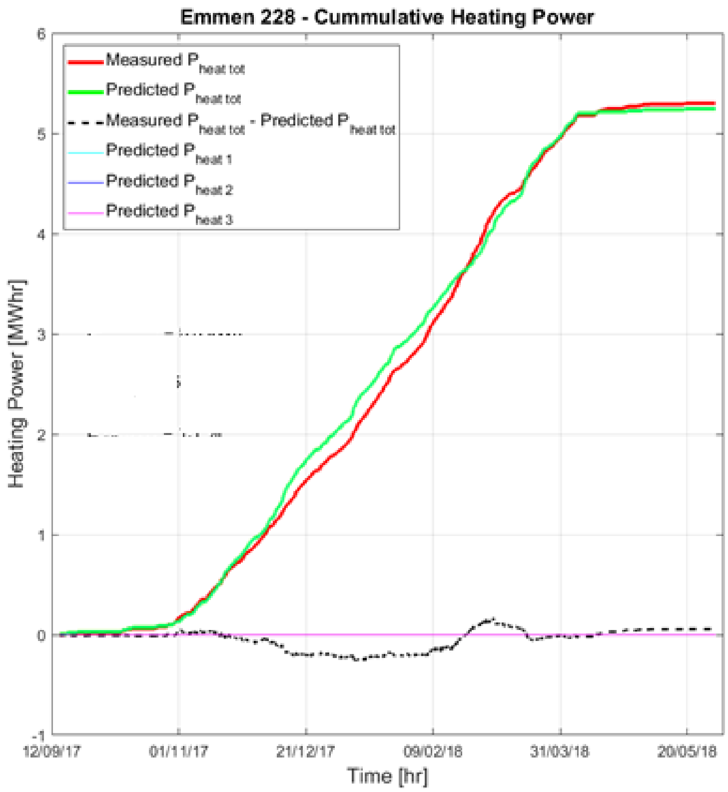


Figure 5-18. Heating demand results for Emmen 228

5.1.4.3 Energy Signature

Energy signature represents the heating demand versus the ambient temperature which are average over a defined period, typically a week. The slope of the lines in energy signature represents the total losses. Higher the slope, the higher will be the losses and consequently the heating demand of the dwelling. Figure 5-19 and Figure 5-20 show the energy signatures of the dwellings in Emmen for the above-mentioned time period. In these plots, the calculated energy use with the RC model (predicted thermal power) are compared to the measured energy use (measured thermal power) of the dwellings. The energy demand and the outside temperature is averaged on weekly-basis. The red and green dots in the figure represent the measured and predicted data points, respectively and the red and green lines represent the linear regression of the corresponding data points. The purple dots represent the electrical power used by HP and it is used to calculate the measured thermal power of HP[10]. Moreover, as expected, the energy demand decreases with the increase in outside temperature.

As we can see in the following figures, the predicted results for the energy signatures have a quite close agreement with the measured energy signatures, which implies that the model is predicting energy performance which is close to the actual performance of these dwellings.

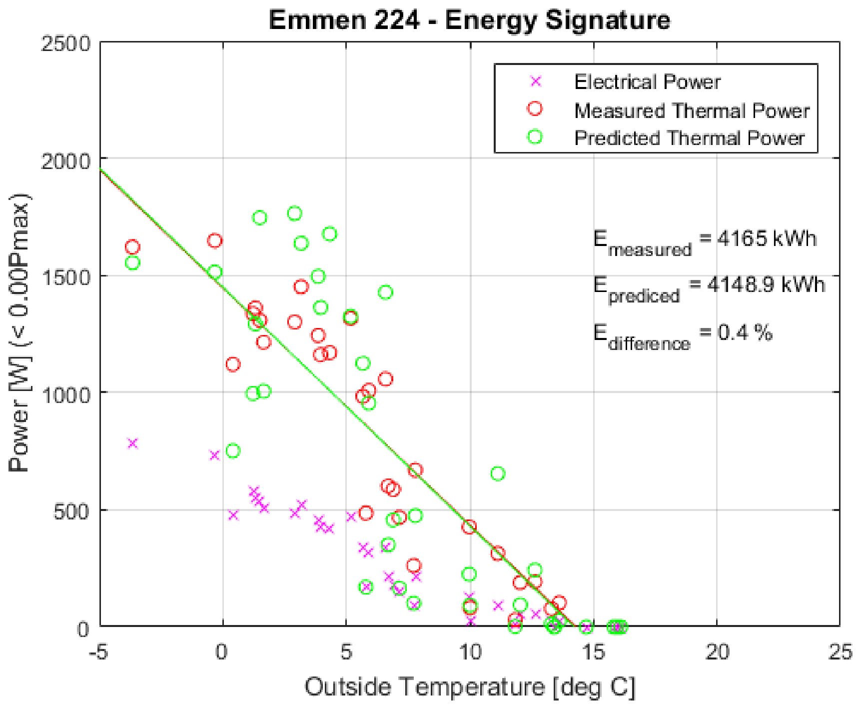


Figure 5-19. Energy signature results for Emmen 224

[10] $\text{HP Power}_{\text{thermal}} = \text{COP} * \text{HP Power}_{\text{electric}}$

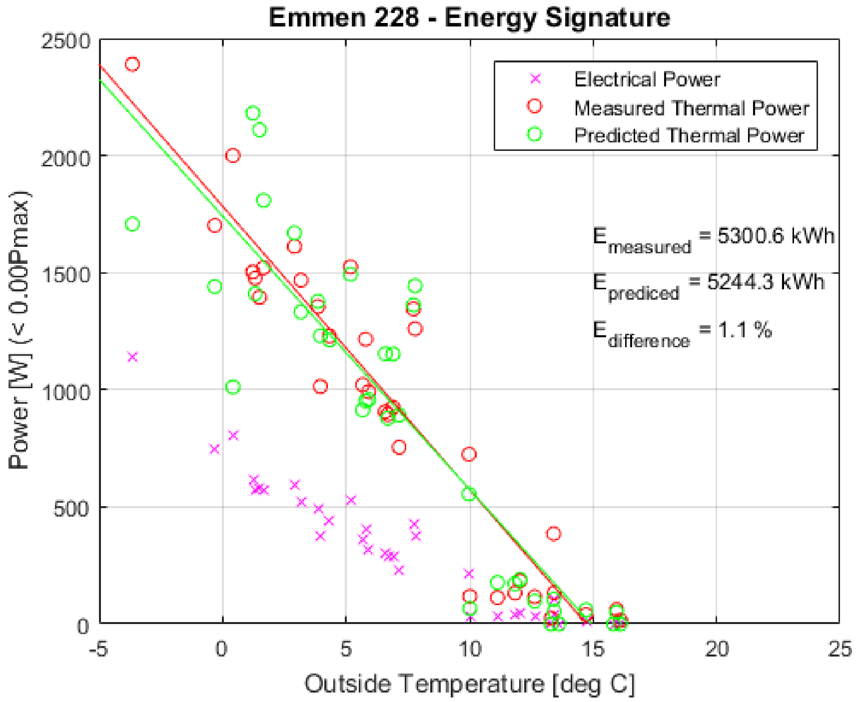


Figure 5-20. Energy signature results for Emmen 228

5.1.4.4 *Conclusions*

In the previous paragraphs we described the first parameter fit for two of the houses in Emmen. These fits are successful: the model succeeds in matching the following indicators with reality: 1) the actual energy consumption, 2) the actual hourly temperature changes in the different zones for a whole year and 3) the energy signatures. This is a promising result. The fact that the model is able to follow these 3 indicators, gives some confidence that the model actually represents reality.

However, the fit techniques are not yet of such a nature that we can be completely sure that the parameters are close to the actual characteristics. On the other hand, the fit has been successful thanks to the monitoring data that was available, while the explicit goal is to reduce the number of sensors in the houses.

We therefore need ways to determine the parameters in the model with more certainty and to use fewer sensors than at present. To gain more confidence in the estimated value of some of the parameters in the model, in the following paragraph a number of sensitivity studies were carried out.

5.2 Sensitivity Analysis

Since there are many parameters in the model and it is very important to have the best possible estimation of the parameters. The aim of the sensitivity analysis is to find out what aspects are important to improve certainty in the model and to show the effect of variations in some of the aspects. The following aspects are considered:

- **Thermal mass**
- **Thermal mass fraction**

- **Temperature setpoints**
- **Neighbor temperatures**
- **Door and window opening fraction**

There are two reasons to choose these aspects. The first one is that many parameters or inputs are used in the model however we are not certain about the accuracy of some of them (thermal mass, thermal mass fraction and door and window opening fraction). The second one is to take the partner's interests into consideration (temperature setpoints and neighbor temperature).
These aspects defined above are investigated based on space heating. This is because the project focuses exclusively on a data-driven RC-network simulation model of the energy performance of NOM houses. Therefore, space heating is the main indicator to evaluate the sensitivity of the model.

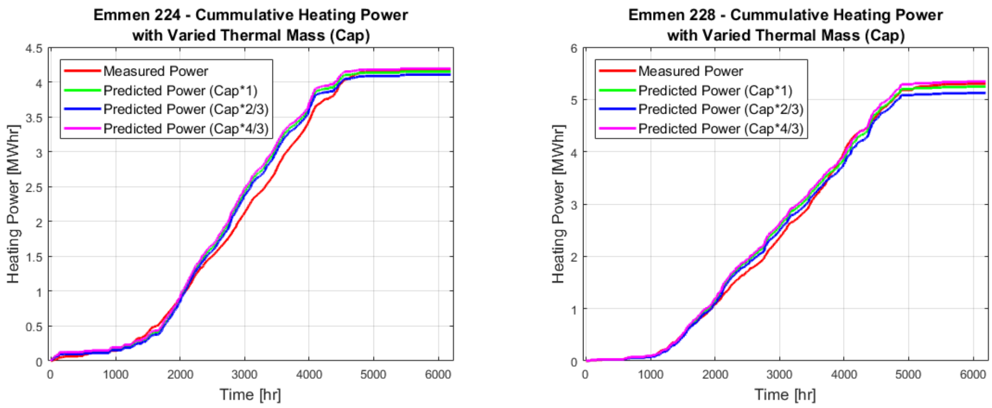
5.2.1 Thermal Mass

As described in Section 5.1.3.2, the thermal mass of the dwelling is calculated from the simplified generic equation[11] based on dwelling volume. However, this equation provides an insight into the thermal mass for a typical dwelling type, not exact results. Therefore this analysis has been performed to quantify the effect of the thermal mass on the heating demand of the dwelling. Here the main focus is heating demand, although it might also influence the temperature result.
The following cases have been studied:

- **Base case:** Thermal mass used in the model (Cap). It is calculated from the simplified generic equation described in Section 5.1.3.2 for the dwellings in Emmen.
- **Case 1:** Lower thermal mass (Cap*2/3)
- **Case 2:** Higher thermal mass (Cap*4/3)

Figure 5-21

Table 6-1 shows the comparison of heating demand for all the above-mentioned cases for both the dwellings in Emmen. As it can be seen in the figure, the total heating demand is only slightly influenced by varying the thermal mass. Therefore we can conclude that the influence of the thermal mass on the energy use for heating is small. We didn't study the effect of the thermal mass on the hourly temperature gradient in the houses. We expect that influence to be larger.



[11] Koene, F.G.H. et. al. (2014), Simplified building model of districts, fifth German-Austrian IBPSA conference

Figure 5-21. Heating demand results for Emmen 224 and 228 for the different cases of thermal mass (Cap)

5.2.2 *Thermal Mass Fraction*

As described in Section 5.1.3.2, the total thermal mass of the dwelling consists of inner and outer parts. The typical thermal mass fraction for a single-family dwelling is between 15-35 %[12] and it represents the ratio between indoor mass and total mass of the dwelling. However, this range provides an insight into the thermal mass for a typical dwelling type, not particularly for this situation. Therefore this analysis has been performed to quantify the effect of the thermal mass on the heating demand of the dwelling. Here the main focus is heating demand, although it might also influence the temperature result.
The following cases have been studied:

- **Base case:** Thermal mass fraction used in the model (f = 0.3)
- **Case 1:** Lower thermal mass fraction (f = 0.2)
- **Case 2:** Higher thermal mass fraction (f = 0.5)

Figure 5-22 shows the comparison of heating demand for all the above-mentioned cases for both the dwellings in Emmen. As can be seen in the figure, the total heating demands in all above-mentioned cases are quite close to each other. Therefore we can conclude that the influence of the thermal mass fraction on the energy use for heating is small. We didn't study the effect of the thermal mass fraction on the hourly temperature gradient in the houses. We expect that influence to be larger.

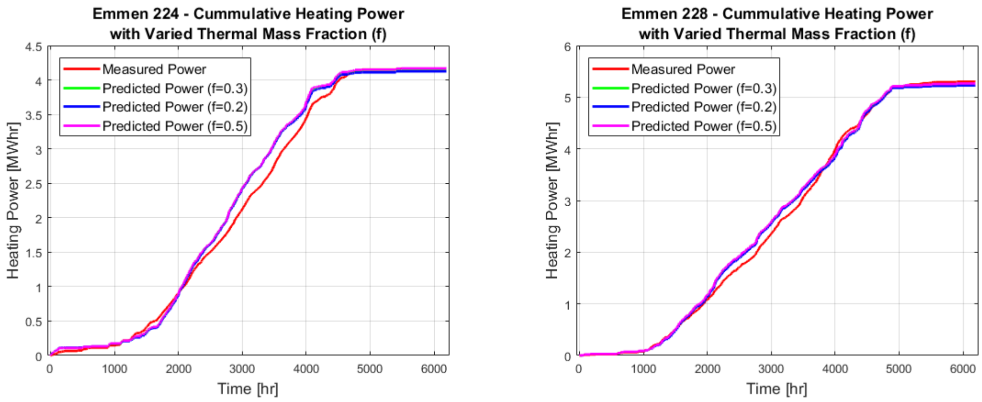


Figure 5-22. Heating demand results for Emmen 224 and 228 for the different cases of the thermal mass fraction (f)

5.2.3 *Temperature Setpoints*

We investigated the influence of the temperature setpoint on the total heating demand of the dwellings. To study this effect, we used the measured temperature setpoint, but increased and decreased this setpoint for the whole period from the measured value by 1 and 2°C.

The cases we studied were the following:

- **Case 1:** Setpoint temperature used in the calculation is 2°C less than the measured setpoint temperature, $\Delta T_{set} = -2$

[12] Koene, F.G.H. et. al. (2014), Simplified building model of districts, fifth German-Austrian IBPSA conference

- **Case 2:** Setpoint temperature used in the calculation is 1°C less than the measured setpoint temperature, $\Delta T_{set} = -1$
- **Base case:** Setpoint temperature used in the calculation is the same as the measured setpoint temperature, $\Delta T_{set} = 0$
- **Case 3:** Setpoint temperature used in the calculation is 1°C higher than the measured setpoint temperature, $\Delta T_{set} = 1$
- **Case 4:** Setpoint temperature used in the calculation is 2°C higher than the measured setpoint temperature, $\Delta T_{set} = 2$

Figure 5-23 shows the comparison of heating demand for all the above-mentioned cases for both dwellings in Emmen.

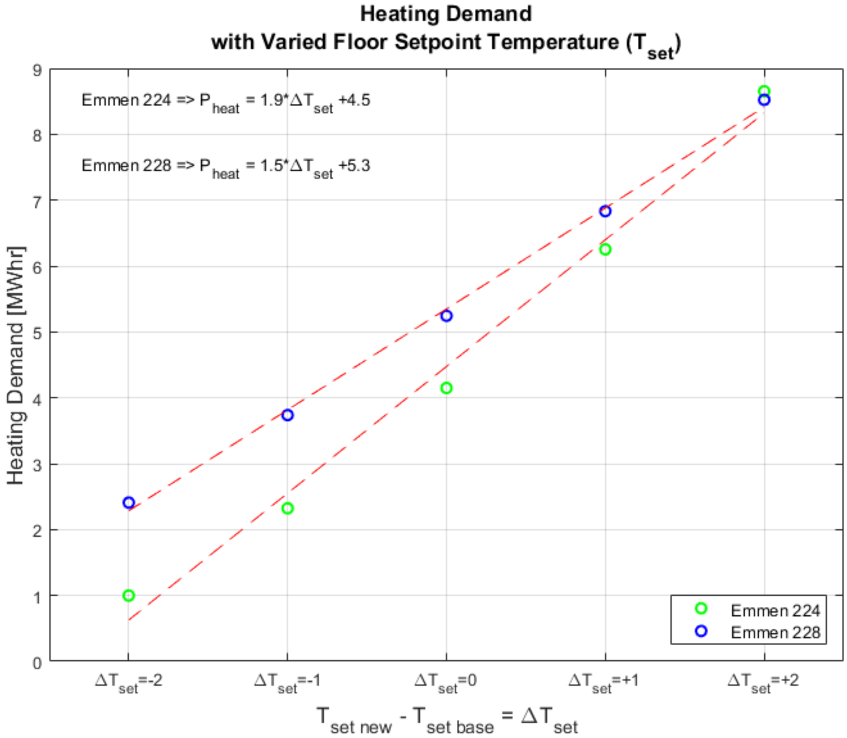


Figure 5-23. Heating demand results for Emmen 224 and 228 for the different cases of setpoint temperature (T_{set})

Table 5-8. represents the variations in the heating demand by varying the heating setpoint from the actual measured setpoint. As it is obvious in the Table 5-8., increase in the setpoint by 1 °C increases the heating demand by 40-50% for Emmen 224. While in case of Emmen 228, the effect of increasing the setpoint of 1 °C on the heating demand is 25-35%. The difference in the increase in heating demand between these dwelling can be related to differences in the user behaviour between the houses: e.g. zone 2 and zone 3 are mostly heated in Emmen 224 and mostly unheated in Emmen 228. In case of heated bedrooms, there are more losses (transmission, ventilation and losses to the neighbors) because of the higher temperature difference.

Table 5-8. Comparison of heating demand for the different cases of setpoint temperature (T_{set})

Setpoint Temperature ($T_{set\_new} - T_{set\_base} = \Delta T_{set}$)	Emmen 224 Heating Demand	Emmen 228 Heating Demand
$\Delta T_{set} = -2$	70 – 80 % decrease	55 – 65 % decrease
$\Delta T_{set} = -1$	40 – 50 % decrease	35 – 45 % decrease
$\Delta T_{set} = 0$	0 % base case	0 % base case
$\Delta T_{set} = +1$	45 – 55 % increase	25 – 35 % increase
$\Delta T_{set} = +2$	100 – 110 % increase	50 – 60 % increase
Note: All period analysis. The change in heating demand is defined based on “base case”.		

In this study we considered the temperature at the neighbors as a fixed pattern. In reality the temperature at the neighbors is influenced by the temperature of the studied house, therefore the changes in heating demand are overestimated.

5.2.4 Neighbor Temperature

The neighbors boundary conditions have a considerable effect on the heating demand of a dwelling. Therefore it is interesting to get insight in the influence of the neighbor’s temperature on the energy performance of the dwelling.
For the base case, the neighbor’s temperature is assumed the same as the measured indoor temperature of the dwelling. This is the so-called ideal case without any heat loss/gain to the neighbors. In addition, other cases are considered by increasing and decreasing the temperature at the neighbors by 1 and 2°C.

The cases we studied were the following:

- **Case 1:** Neighbor’s temperature used in the calculation is 2°C less than the measured indoor temperature, ΔT_{neigh} = -2
- **Case 2:** Neighbor’s temperature used in the calculation is 1°C less than the measured indoor temperature, ΔT_{neigh} = -1
- **Base case:** The neighbor’s temperature used in the calculation is the same as the measured indoor temperature, ΔT_{neigh} = 0
- **Case 3:** Neighbor’s temperature used in the calculation is 1°C higher than the measured indoor temperature, ΔT_{neigh} = 1
- **Case 4:** Neighbor’s temperature used in the calculation is 2°C higher than the measured indoor temperature, ΔT_{neigh} = 2

Figure 5-24 shows the comparison of heating demand for all the above-mentioned cases for both the dwellings in Emmen.

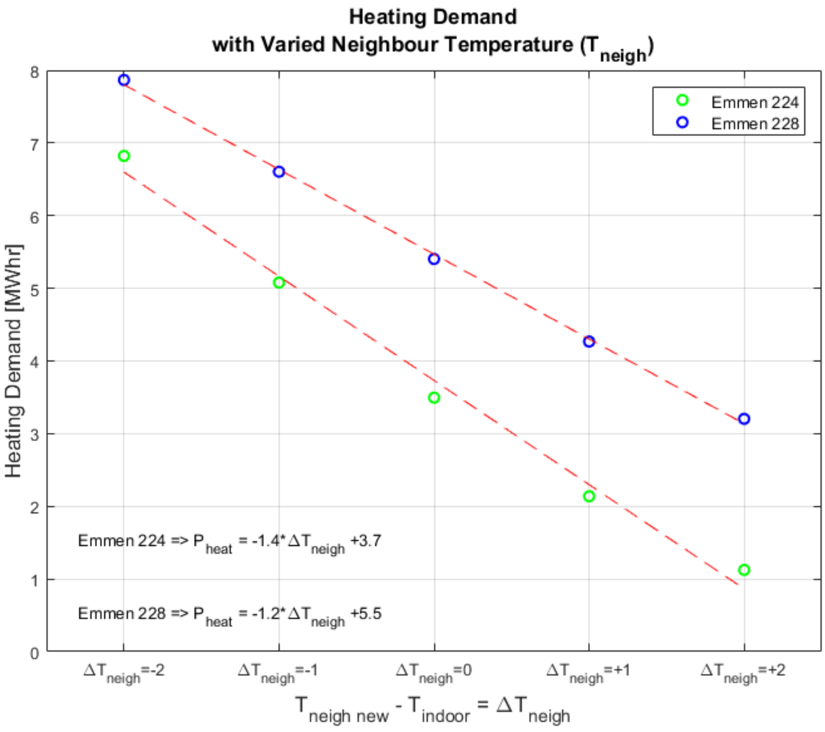


Figure 5-24. Heating demand results for Emmen 224 and 228 for the different cases of neighbor temperature (T_{neigh})

Table 5-9. represents the effect on the heating demand by varying the temperature difference between the neighbor's temperatures and the indoor temperature. The variations in the heating demand from the base case for various above-mentioned cases are represented, for both dwellings in Emmen. The base case in the table represents the scenario when there is no loss to the neighbor's. This is because the neighbor temperatures used in the calculation are assumed the same as the measured indoor temperature of the dwelling. Moreover, the negative temperature difference (ΔT_{neigh}) leads to the increase in heating demand and it increases with the increase in temperature difference (ΔT_{neigh}).

Table 5-9. Heating demand results for Emmen 224 and 228 for the different cases of neighbor temperature (T_{neigh})

Neighbor Temperature ($T_{neigh\_new} - T_{indoor} = \Delta T_{neigh}$)	Emmen 224 Heating Demand	Emmen 228 Heating Demand
$\Delta T_{neigh} = -2$	90 – 100 % increase	40 – 50 % increase
$\Delta T_{neigh} = -1$	40 – 50 % increase	15 – 25 % increase
$\Delta T_{neigh} = 0$	0 % base case	0 % base case
$\Delta T_{neigh} = +1$	35 – 45 % decrease	15 – 25 % decrease
$\Delta T_{neigh} = +2$	65 – 75 % decrease	35 – 45 % decrease
Note: All period analysis. The change in heating demand is defined based on "base case".		

It is interesting to note that the effect of ΔT_{neigh} on the heating demand for both dwellings is different and tihs is because of the difference in user behaviour: e.g. zone 2 and zone 3 are mostly heated in Emmen 224 and mostly unheated in Emmen 228. As it is evident from this analysis, for this specific situation the increase in heating demand for the dwelling with heated bedrooms is almost double compared with the dwelling with an unheated bedroom. This is because, in the case of heated bedrooms, there are more losses or gains because of the higher temperature difference.
In this study we considered the temperature at the neighbors as a fixed patern as given per scenario. In reality this temperature at the neighbors is influenced by the temperature of the studied house, therefore the changes in heating demand are overestimated.

5.2.5 *Window Opening Fraction*

As described in Section 5.1.3.6, both the dwellings in Emmen are equipped with door and window opening sensors. Thereby, it is possible to detect when window and door are closed (0) or open (1) or sometimes in between 0 and 1 (for instance, if the sensor gives 0.5 for a particular timestep, it means the window/door is open only for half an hour). However, the door and window opening area which is one of the main parameters to estimate the flow rate through open windows and doors is still unknown because the opening fraction is unknown and it leads to one of the biggest uncertainties in this model.
From the measured data for the windows opening, it has been observed that the occupants are quite often opening the bedroom windows, even during the heating period. We studies the influence of opening the bedroom windows on the heating demand. We focussed on the bedroom windows only. Thus, this analysis has been performed to quantify the effect of the bedroom windows opening fraction on the heating demand of the dwelling.

We studied the following cases:

- **Case 1:** Whenever bedroom windows are open, the opening duration is decreased to 0. Factor = 0.0. So we assume all wbedroom windows to be closed all the time.

- **Case 2:** Whenever bedroom windows are open, the opening duration is decreased by 0.5. Factor = 0.5. So we assume all windows to be opened only half of the time that they are actualy opened.
- **Base case:** Actuel measured bedroom windows opening schedule, Base Case.
- **Case 3:** Whenever bedroom windows are open, the opening duration is increased by 1.5. Factor = 1.5. So we assume all windows to be opened one and a half times more that they are actualy opened.
- **Case 4:** Whenever bedroom windows are open, the opening duration is increased by 2. Factor = 2.0. So we assume all windows are opend twice as much as they are actualy opened.

In this analysis all bedroom windows are taken into account. In Figure 5-25 only the window opening/closing pattern of the left front bedroom for Emmen 228 is shown. The base case (factor=1) represents the measured opening schedule. While other cases represent opening schedules with more (factor>1) or less (factor <1) opening times. Opening schedules for the 5 cases are shown in the following figure.

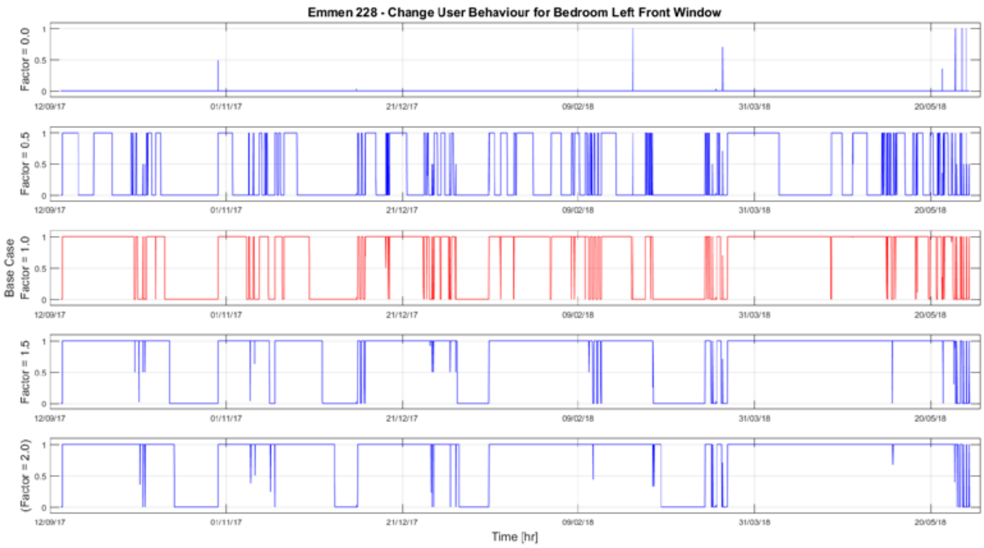


Figure 5-25. Change of windows opening behavior for Emmen 228 (for left bedroom window)

Figure 5-26 shows the comparison of heating demand for all the above-mentioned cases for the bedroom windows opening, for both dwellings in Emmen. The X-axis in the figure represents the window opening hours based on the different opening factors, as explained above. As expected, the heating demand of both the dwellings in Emmen increases with the increase in opening hours of the windows.

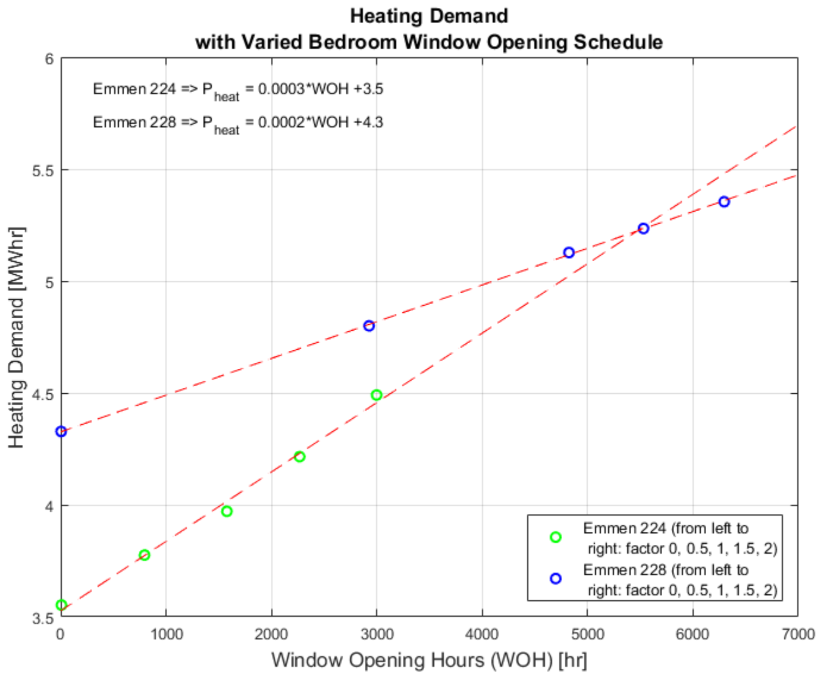


Figure 5-26. Heating demand results for Emmen 224 and 228 for the different cases of window opening hours

Table 5-10. shows the effect of window opening hours on the heating demand of the dwellings in Emmen. As seen in the table, the effect of an increase in the opening factor on the heating demand is different for both dwellings because of the varying user behaviour in the houses.

Table 5-10. Heating demand results for Emmen 224 and 228 for the different cases of the window opening

Emmen 224	Opening Hours (sum of 3 bedroom windows)	Heating Demand
Factor 0	0	7 – 13 % decrease
Factor 0.5	800	2 – 8 % decrease
Factor 1	1580	0 % base case
Factor 1.5	2270	3 – 9 % increase
Factor 2	3000	10 – 16 % increase
Emmen 228	**Opening Hours (sum of 3 bedroom windows)**	**Heating Demand**
Factor 0	0	12 – 18 % increase
Factor 0.5	2930	3 – 9 % increase
Factor 1	4820	0 % base case
Factor 1.5	5530	1 – 5 % decrease
Factor 2	6300	2 – 8 % decrease

Note: Heating period analysis. The period is 3700 hr (from 900 to 4600) for Emmen 224 and 4000 hr (from 900 to 4900) for Emmen 228.
Only change in bedroom window opening schedule (not kitchen and living room windows and doors).
The change in heating demand is defined based on “base case”.

It is interesting to see that the heating demand for Emmen 228 is higher compared to the heating demand of Emmen 224, even the bedrooms in Emmen 228 are unheated for most of the time over the measured period. This finding reflects how occupancy

can affect heating demand. In this particular scenario, this is mainly because of the opening of windows. As we can see in
Table 5-10. the hours of bedroom windows opening for Emmen 228 are almost 3 times that of Emmen 224.

5.3 Conclusions and lessons learned

The aim of TKI Optimaal is to develop models and algorithms for data analysis with which, in time, at least 80 to 90% of the deviation between the predicted and actual energy and indoor climate performance of individual NoM houses can be explained. The approach that we followed in TKI Optimaal is to develop a data-driven RC-network simulation model of NOM houses that will allow us to approach the actual performance of those houses on an individual level as good as close as possible. The conclusions of the development are the following:

- In TKI Optimaal we took a big first step by setting up a data-driven RC-network simulation model and filling in the parameters in that model by a combination of expert judgement.
  On of the parameters that is unknown in the model and that will have a big effect on the nergy use of the houses is the opening fraction of the windows. We succeeded in tuning the model by tuning this factor. This resulted in a close agreement of the predicted hourly indoor temperature, heating demand, and energy signature with the corresponding measured data. The fact that the model is able to follow these 3 indicators, gives some confidence that the model actually represents reality. However, the fit techniques are not yet of such a nature that we can be completely sure that the parameters are close to the actual characteristics.
- It is interesting to observe that the heating demand of Emmen 228 is higher compared to the heating demand of Emmen 224, even though the bedrooms in Emmen 228 are unheated, while Emmen 224 has heated bedrooms for most of the time over the measured period. This is because in Emmen 228 the total hours of window opening are almost 3 times higher compared to that of Emmen 224. This situation shows how occupancy window behavior can influence the heating demand even when the rooms where the windows are opened are unheated.
- Sensitivity analysis for thermal mass and thermal mass fraction (for the indoor and outer mass) show a very slight effect on the total heating demand of the dwellings in Emmen. However, the effect of these parametric analyses on the variation in the indoor temperature might be significant and it can help us to get a better estimation of the thermal mass and the mass fraction for the indoor and outdoor surface.
- For neighbor’s sensitivity case, the variations in the heating demand are overestimated. Because in this analysis we consider a fixed temperature difference between the neighbors and the dwelling. However, in real situations, the dwelling leads to heat up the neighbor’s or the neighbor’s heats up the dwelling. Thus, it will reduce the heat losses/gains compared to the estimated ones in this analysis. The same goes for the sensitivity analysis of heating setpoint .
- The sensitivity analysis for the heating setpoint reveals other interesting information. In Emmen 224 the increase in the setpoint by 1 °C increases the heating demand by 40-50%. While in the case of Emmen 228, the effect of

increasing the setpoint by 1 °C only increases the heating demand by 25-35%. The difference in the increase in heating demand between these dwelling can be related to differences in user behaviour: e.g. zone 2 and zone 3 are mostly heated in Emmen 224 and mostly unheated in Emmen 228. This is because, in case of heated bedrooms, there are more losses (transmission, ventilation, and losses to the neighbors) because of the higher temperature difference between the indoor and ambient.

- Measurement of the window opening fraction is possible, but probably not feasible. In addition, we are looking for ways to reduce the sensor set, beginning with the window and door sesnors. Therefore we need sophisticated techniques to estimate the parameters of the model. A more sophisticated ventilation model might help to incorporate the airflow through the windows and doors based on the dynamic weather conditions (wind speed and wind direction). In addition, it is also important to consider the air coupling among the different zones. By adding this module, it will also become possible to take into account actual indoor air quality perfomences. In addition, artificial intelligent techniques might help to predict window opening patterns.

6 Influence of user behavior

Parallel to the development of the data-driven RC-network simulation model (described in chaper 5), one of the NOM houses was also modelled in a detailed building simulation model (TRNSYS). During the development of the data-driven RC-network simulation model, we saw that some of the behavioral parameters have a major impact on energy performance. We wanted to investigate this further. The results are described in this chapter.

The purpose of this chapter is to address the following questions:

1) What is the energy consumption of one of the dwellings in Emmen if modeled in TRNSYS by using the measured user behavior? And is this anywhere near the measured consumption?
2) What is the influence of 1) ventilation behavior for heated and unheated bedrooms and 2) the temperature of the neighbor’s house on the heating demand for a dwelling in Emmen.

To investigate the above questions, this chapter includes the description of the TRNSYS model and the parameters used in the model. Furthermore, the sensitivity analysis is performed for the windows opening (natural ventilation) and the neighbor's boundary temperature by considering the different scenarios.

6.1 Model Approach

Emmen 224 is modeled in TRNSYS to evaluate the space heating profile. This is a middle-terrace house with a flat roof, which is oriented West-East. Table 5-1 and Table 5-2 in chapter 5 show the parameters used in the TRNSYS model. More information about the dwelling is available in section 5.1.1. The other parameters used such as heating system, internal gains, ventilation network, etc. are available in annex C. What is different in this TRNSYS model compared to other TRNSYS simulations is that all used behavior that was monitored, this moitoring data is used instead of fixed, standardised user patters.

6.2 Results of TRNSYS Model

The simulation is run for the whole measurement period, starting from 22 September 2017 until 28 May 2018. This period covers almost all the space heating periods throughout the year. Figure 6-1 represents the heating profile for a one-hour simulation time step, for the above-mentioned time period. Since the measurement period starts in September, less heating is required at the beginning of the simulation period. It can be observed that the maximum heating capacity is 6 kW. The TRNSYS model for Emmen 224 provides the total heating demand of 13.6 GJ, which is comparable with the measured heating demand (14.8 GJ).

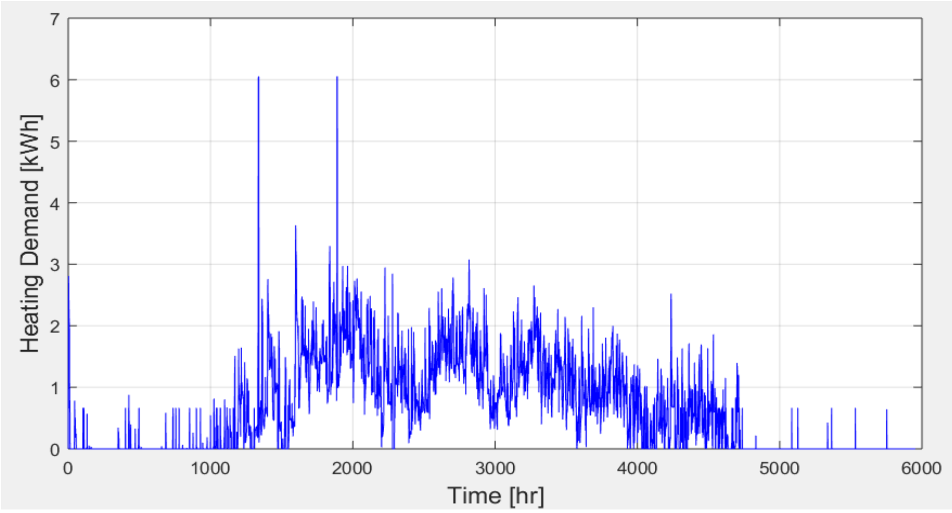


Figure 6-1 Space heating profile

6.3 Sensitivity Analysis

Based on the outcome of other projects, we know the ventilation and the neighbor's boundary conditions have a considerable effect on the heating demand of a dwelling. Parallel to the development of the data-driven RC-network simulation models, we used TRNSYS to perform a parametric analysis.
The aim of this sensitivity analysis is to quantify the effect of ventilation with windows and temperature differences with the neighbor's on various chosen scenarios: it is interesting to observe how different scenarios effect the heating demand of a dwelling where all bedrooms are either unheated or heated. The temperature pattern, the energy signature and the heating demand for a cold and a typical winter week are presented for each case.

6.3.1 Ventilation with Windows

As the ventilation through windows is a considerable factor for determining the heating demand of the dwelling. The purpose of this sensitivity analysis is to compare the various cases with the so-called ideal or base-case when there is no loss through the windows. The following 4 cases have been chosen for this aspect:

- **Base case:** All windows are closed
- **Case 1:** Windows in ''Kiepstand'' position and measured opening/closing profile
- **Case 2:** Same as case 1, but in addition, all the bedroom windows are assumed to be fully-open for an hour every day
- **Case 3:** Same as case 1, but in addition, the bedroom windows on the East-side of the house are fully-open for an hour every day

The aforementioned cases are considered for a house where all bedrooms are either unheated or heated. Moreover, during these calculations, the thermal equilibrium is assumed towards the neighbors and the same thermostat profile as measured in the living room is used as the heating setpoints for the heated bedrooms.
Table 6-1 shows the comparison of heating demand for all the above-mentioned cases for the house where all bedrooms are either unheated or heated; namely the ''Unheated Bedroom'' and the ''Heated Bedroom'' respectively.

Table 6-1 Comparison of heating demand for the different cases of ventilation with Window

		Unheated Bedroom		Heated Bedroom	
Cases	**Description**	Heating demand [GJ]	Difference [%]	Heating demand [GJ]	Difference [%]
Base Case	**Window closed**	**7.0**	**-**	**8.1**	**-**
Case 1	**Kiepstand**	10.0	43	13.6	68
Case 2	**Kiepstand + 1hr**	11.0	57	16.5	104
Case 3	**Kiepstand + 1hr (east -side)**	10.7	53	15.6	93

It is interesting to note that in the base-case scenario, the difference in heating demand of the house with the unheated and heated bedrooms is not very significant. This is because the R_c value between the floors is very low (0.57 m^2 K/W). Moreover, the ventilation system recovers the heat from the ground floor and adds this heat to the bedrooms. In this case, the energy savings can be improved by increasing the internal floor insulation (R_c) and making the separation for the ventilation heat recovery system.
In case 1, as expected, we can observe the increase in heating demand which is 43% for the unheated bedroom and 68% for the heated bedroom scenario, compared with the base-case. For the unheated bedroom scenario in case 2, it is interesting to observe that the heating demand increases by 1 GJ compared to the unheated bedroom scenario for case 1. It implies that the heat losses from the ground floor increase due to the full opening of bedroom windows for an hour every day. For the heated bedroom scenario, there is an even bigger increase in the heating demand as compared to that of case 1.
However, the increase in heating demand is not as significant as we expected. This is because of the limited capacity of the heating system. The abrupt decrease in the indoor temperature has been observed as a consequence of fully opening the bedroom windows for 1-hour. Since the heating system has a limited heated capacity, it limits the heating power during the particular hours at the expense of the decrease in indoor temperature.
For case 3, the heating demand for both the unheated and heated bedroom case is slightly lower than that of case 2. This is due to the reduced heat losses through the windows because only the bedroom windows which are on the East-side are fully-opened for an additional hour of the day, so no cross ventilation took place.

6.3.1.1 Energy Signature

Energy signature represents the heating demand versus the ambient temperature which is averaged over a defined period, typically a week. The slope of the lines in the energy signature represents the total losses. The higher the slope, the higher will be the losses and consequently the heating demand of the dwelling. Figure 6-2 shows the energy signature for all the above-mentioned cases for the dwelling with both the unheated and heated bedroom scenario. The energy demand and the outside temperature is averaged on weekly-basis. The dots in the figure represent the data points and the lines are obtained from these data points by fitting. The dotted-lines represents the cases for the dwelling with unheated bedrooms while the solid-lines for the dwelling with heated bedroom. As expected, the energy demand decreases with the increase in outside temperature.

It is obvious from the following figure, that the heated bedroom requires more energy as compared to the unheated bedroom scenario for the same case. Therefore, the energy signatures are higher for the dwelling with heated bedrooms compared to the same dwelling with unheated bedrooms. We can also see from the following figure that the energy signature for the case when all bedroom windows are fully open for an additional hour every day is higher compared to the case when all bedroom windows are open in 'kiepstand' position. This is because of the higher losses when the windows of all bedrooms are fully open for an additional hour.

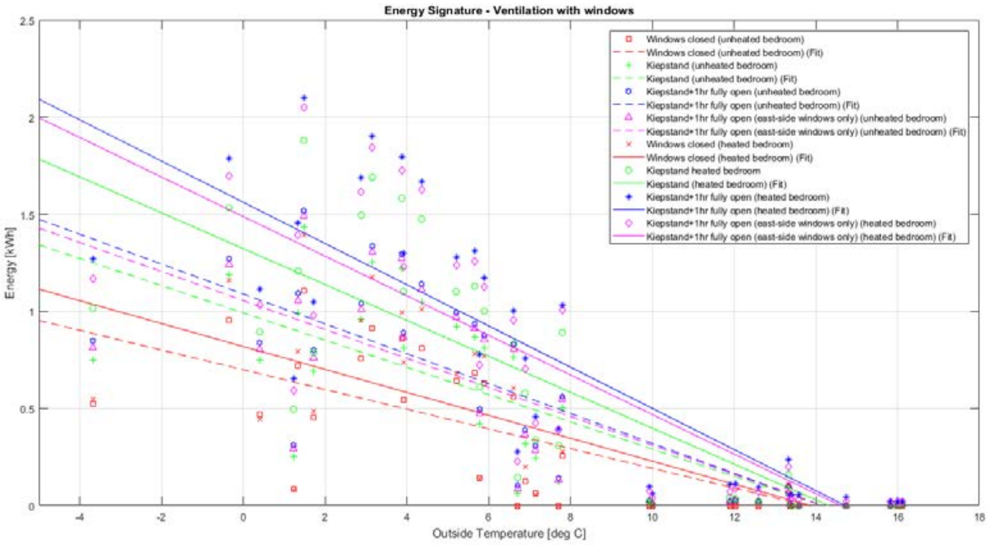


Figure 6-2 Energy signature - Ventilation with windows

6.3.1.2 Typical and Coldest Winter Week Analysis

We analysed the above mentioned scenarios for a typical winter week and the coldest winter week.

Typical winter week

The typical winter week that is looked at starts from the 28th of December until the 3rd of January. Figure 6-3 and Figure 6-4 show the variations in indoor temperature for both the first floor and the second floor along with the heating power required for the different scenarios, for the dwelling with unheated and heated bedroom cases. The temperature fluctuations on the first floor are because of the opening of the front door which results in an abrupt decrease in indoor temperature (because the entrance temperature decreases and the floor temperature is the simple average of all the zones temperatures on the floor). For the second floor, we can observe the decrease in temperature for ''Kiepstand+1hr" and ''Kiepstand+1hr (for east-side windows)" because during these scenarios, the windows are fully open for one hour every day (from 9 pm-10 pm), which increases the ventilation losses through the windows and consequently increases the heating power to maintain the heating setpoint. Because the heating system has a limited heating capacity and thereby unable to add enough heat immediately to maintain the indoor setpoint. The variation in heating power for different cases for the dwelling with unheated and heated bedroom cases can be observed in the following figures.

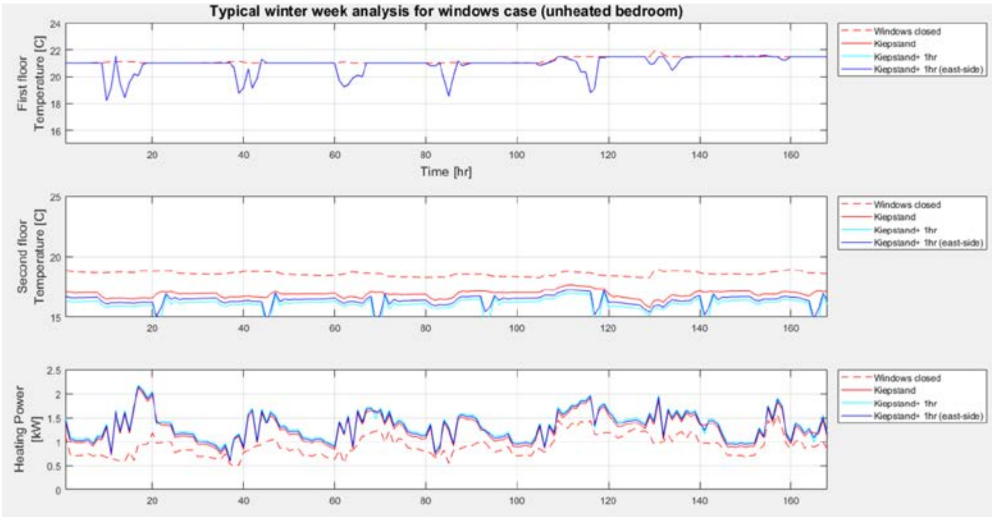


Figure 6-3: Typical winter week analysis for windows case (unheated bedrooms)

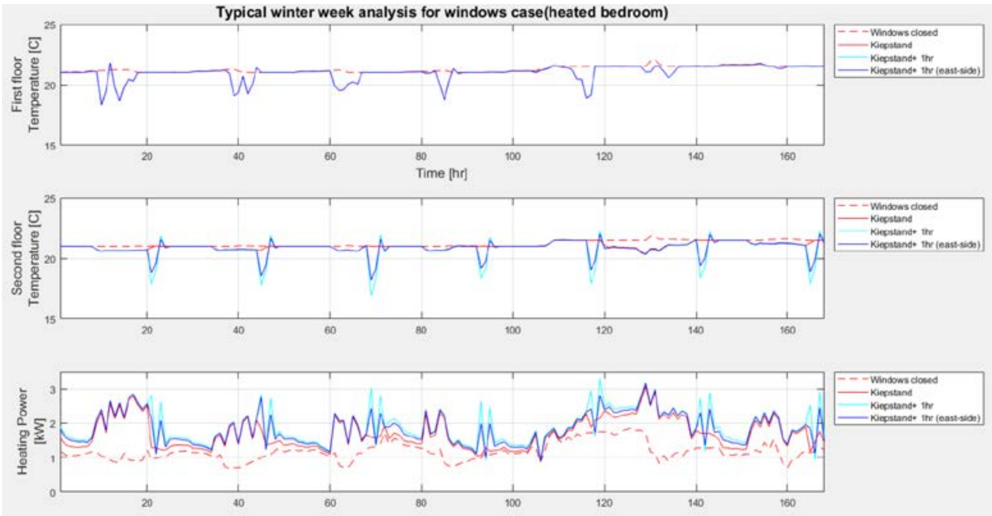


Figure 6-4: Typical winter week analysis for windows case (heated bedrooms)

Coldest winter week

The coldest winter week occurs in mid of February and the average temperature is around -2 °C. This analysis reflects similar results as we discussed in the previous section for the typical winter week analysis. The details for this analysis are described in annex D.

6.3.2 Neighbors Cases

This analysis is performed to observe the effect of neighbors on the heating demand of the dwelling. The R_c value of the wall with the neighbors is typically very low, which implies that the neighbors can be a source of heat loss or gain to the dwelling. The following cases have been considered under this aspect:

- **Base-case:** Thermal equilibrium towards the neighbors
- **Case 1:** Heated first floor, unheated second floor at the neighbors
- **Case 2:** Both floors heated at the neighbors

The aforementioned cases are considered for the dwelling with both heated and unheated bedrooms. Moreover, during these calculations, the windows are considered to be on the ’kiepstand’ position (using the open/close pattern as measured). Since it is an attached house, it is assumed that the situation is similar at both neighbors. Moreover, it is assumed that all other boundary conditions at the neighbors are identical to the dwelling under observation except the heating profile. The heating at the neighbors is only present in the early morning (6 AM – 8 AM) and in the evening (6 PM – 11 PM) with the heating setpoint of 20 °C.
This analysis implies that the conditions at the neighbors have a significant effect on the energy demand of the dwelling, about a 38% increase in the heating demand (in this specific case) when both floors are heated and only the first floor is heated at the neighbors. Therefore, it is very important to know about the boundary conditions at the neighbors in order to predict the performance of a dwelling.

Table 6-2 Comparison of heating demand for the different cases of neighbors

		Unheated Bedroom		**Heated Bedroom**	
Cases	**Description**	Heating demand [GJ]	Difference [%]	Heating demand [GJ]	Difference [%]
Base Case	**Thermal equilibrium towards the neighbors**	**10.0**	**-**	**13.6**	**-**
Case 1	**Heated ground floor at the neighbors**	12.5	25	18.8	38
Case 2	**Both floors heated at the neighbors**	10.5	5	16.2	19

Table 6-2 shows the comparison of the heating demand for all the above-mentioned cases for the dwelling with both unheated and heated bedrooms. From Table 6-2, It can be seen that the heating demand is the lowest for the base-case because there are no losses to the neighbors, followed by case 2 and case 1. The heating demand is the highest for case 1 among all the cases for the dwelling with both unheated and heated bedroom scenarios. In case 1, there is a significant increase in heating demand for the dwelling with a heated bedroom case, which is due to the increase in the heat losses to the neighbors from the bedroom floor. This is because the bedroom floor at the neighbors is completely unheated and there is a considerable temperature difference between the indoor temperature and the neighbor’s bedroom.
This analysis implies that the conditions at the neighbors have a significant effect on the energy demand of the dwelling, about a 38% increase in the heating demand (in this specific case) when both floors are heated and only the first floor is heated at the neighbors. Therefore, it is very important to know about the boundary conditions at the neighbors in order to predict the performance of a dwelling.

Table 6-2 Comparison of heating demand for the different cases of neighbors

		Unheated Bedroom		**Heated Bedroom**	
Cases	**Description**	Heating demand [GJ]	Difference [%]	Heating demand [GJ]	Difference [%]
Base Case	**Thermal equilibrium towards the neighbors**	**10.0**	**-**	**13.6**	**-**
Case 1	**Heated ground floor at the neighbors**	12.5	25	18.8	38

Case 2	**Both floors heated at the neighbors**	10.5	5	16.2	19

6.3.2.1 *Energy Signature*

Energy signature represents the heating demand versus the ambient temperature which is averaged over a defined period, typically a week. The slope of the lines in the energy signature represents the total losses. Higher the slope, higher will be the losses and consequently the heating demand of the dwelling. Figure 6-5 shows the energy signature for the neighbors heating cases. The energy demand and the outside temperature is averaged on a weekly basis. The dotted-lines represent the cases for the unheated bedrooms while the solid-lines for the heated bedroom cases. As it can be seen in the figure, The energy signature is lowest for base-case followed by case 2 (both floors heated at the neighbors) and then case 1 (only ground floor heated at the neighbor). Moreover, as expected, the energy demand decreases with the increase in outside temperature.

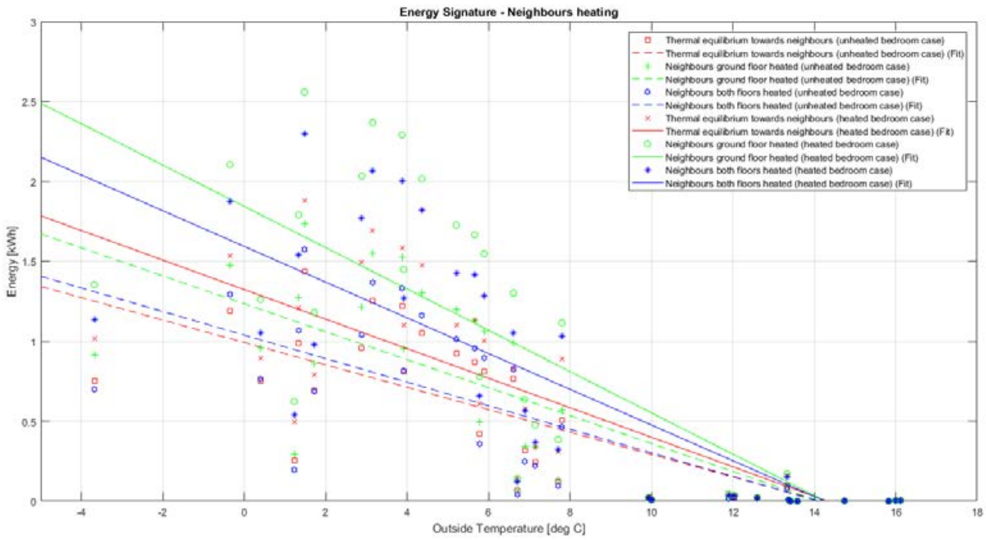


Figure 6-5 Energy signature - Neighbors heating

6.3.2.2 *Typical and Coldest Winter Week Analysis*

As we discussed in Section 6.3.1.2, typical and coldest winter week analysis has been also carried out to analyze the variations in the indoor temperature and the heating power required for all the above-mentioned cases for the neighbor's boundary conditions.

Typical Winter Week

Figure 6-6 and Figure 6-7 shows the variations in indoor temperature for both first and second floor along with the heating power required for the different cases, for the dwelling with both unheated and heated bedroom scenario, respectively. The temperature fluctuations on the first floor are because of the opening of the front door which results in an abrupt decrease in indoor temperature. For the second floor, the temperature is lowest in case 1 (neighbor first-floor heating), because of the higher thermal losses through the second floor to the neighbors (in Figure 6-6).
Moreover, in both cases (unheated and heated bedrooms) the heating demand is highest for case 1 (only ground floor heated at the neighbor) followed by the case 2 (both floors heated at the neighbors) and the base-case (thermal equilibrium towards the neighbors).

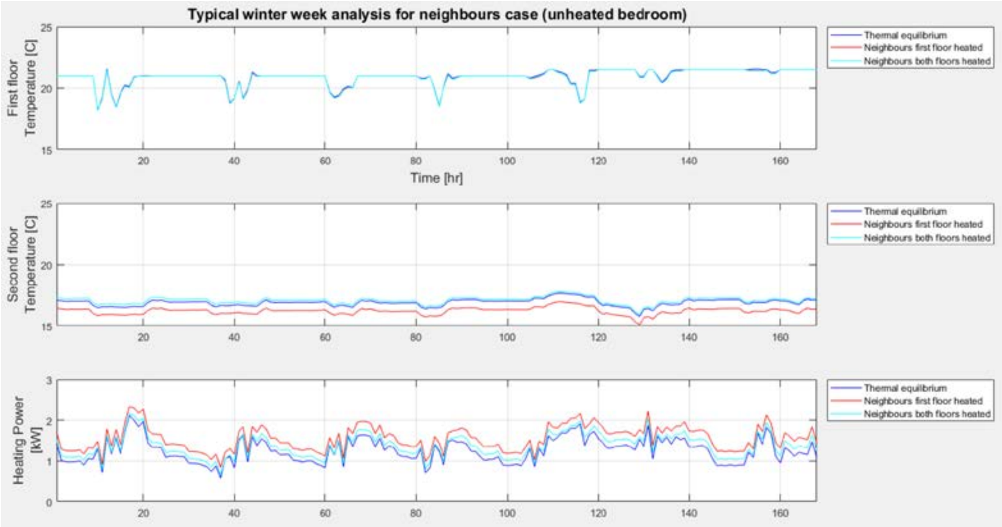


Figure 6-6: Typical winter week analysis for neighbors case (unheated bedroom)

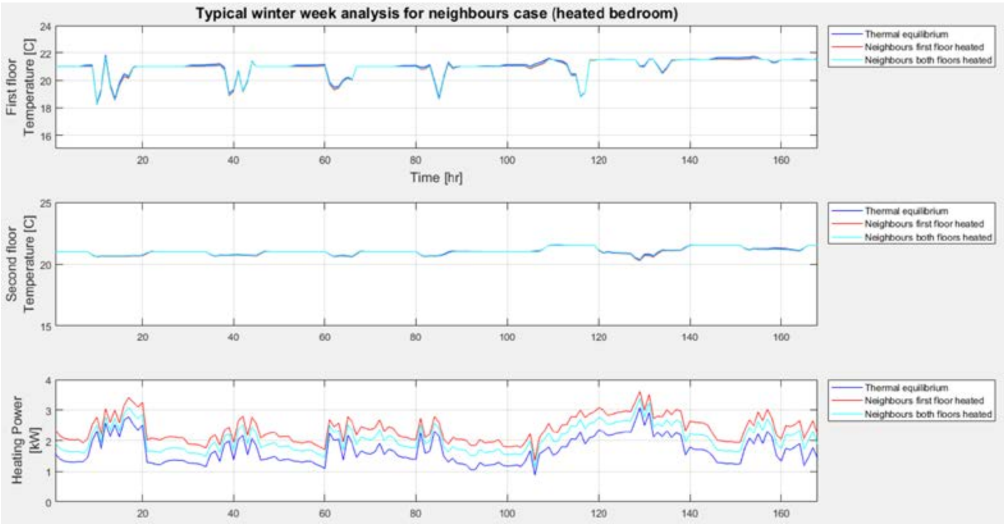


Figure 6-7: Typical winter week analysis for neighbors case (heated bedroom)

Coldest Winter Week

As already mentioned, the coldest winter week occurs in mid of February and the average temperature is around -2 °C. This analysis shows similar results as we discussed in the previous section for the typical winter week analysis. The details for this analysis are described in annex D.

6.4 Conclusions and lessons learned

From the TRNSYS simultions described in this chapter, we derived the following conclusions:

- The dwelling in Emmen is modeled by using the same measured data and using the same assumptions as we used in the RC model (except the difference in the ventilation model which is described in detail in annex C). The heating demand predicted for the dwelling is 13.6 GJ, which is comparable with the measured heating demand (14.8 GJ).

- The difference in space heating demand for the base-case (bedroom windows are closed) for the dwelling with an unheated and heated bedroom scenario is not very significant. This finding is due to this specific situation because the ventilation system recovers the heat from the first floor and add this heat to the bedrooms (even when the intention is not to heat up the bedrooms). In this case, the energy savings can be improved by increasing the internal floor insulation (R_c) and making the separation for the ventilation heat recovery system.
- One of the most important findings is the increase in heating demand for the case when bedroom windows are at kiepstand position, even when the bedrooms are unheated. The increase in heating demand is about 43% in this specific case compared to the base-case (when all bedroom windows are closed). This implies that the opening of the bedroom windows have a considerable effect on the heating demand even if the bedrooms are unheated.
- The typical and coldest winter week analysis reveals another very important information about occupancy behavior. The typical and coldest winter week is picked only based on the variations in the ambient temperature, with the daily-average temperature of 5°C and -2°C, respectively. We normally expect that the coldest week should give us the peak heating power required. However, in a real situation, it also depends on the occupancy behavior. In this specific situation, the heating power required for the typical winter week is higher compared to the coldest winter week for the dwelling with both unheated and heated bedroom scenarios. This is because of opening the external doors on the ground floor. It is observed that occupants tend to open the doors on the ground floor more often during the normal winter period (typical winter week), which requires higher heating demand to maintain the heating setpoint. However, during the coldest week, occupants reflect energy-conscious behavior.
- In case of limited heating capacity, the opening of the bedroom windows for an additional hour (to ventilate the bedrooms) don't have a significant effect on the total space heating demand of the dwelling. However, the indoor temperature can drop below the thermal comfort temperature. Neighbor's boundary conditions do have a considerable influence on the total heating demand. This study implies that the conditions at the neighbors have a significant effect on the energy demand of the dwelling, about a 38% increase in the heating demand (in this specific case) when both floors are heated and only the first floor is heated at the neighbors. Therefore, it is very important to know about the boundary conditions at the neighbors in order to predict the true performance of a dwelling.

7 Conclusions and next steps

7.1 Introduction

The aim of TKI Optimaal is to develop models and algorithms for data analysis with which, in time, at least 80 to 90% of the deviation between the predicted and actual energy and indoor climate performance of individual NoM houses can be explained. The approach that we follow is to develop a data-driven RC-network simulation model of NOM houses that will allow us to approach the actual performance of those houses on an individual level: If a model succeeds in approximating the actual performance of a dwelling, it is also clear which aspects determine the performance of that dwelling. To reach this we need to develop a data-driven RC-network simulation model of NOM houses and fit all parameters in the model as good as possible so that the model gives outcomes that match with monitored data as good as close as possible.

In TKI Optimaal we took a big first step in this by setting up a data-driven RC-network simulation model and filling in the parameters in that model by a combination of expert judgement and fitting of parameters to monitoring data.

7.2 Conclusions on monitoring and data quality

In order to achieve this, we have started to set up monitoring in two types of houses, namely social rental houses that have been renovated by BAM to NOM level and new-build houses at NOM level built by van Wijnen in Ermelo (see Ch2). The purpose of the monitoring was to be able to fit the parameters in the model as well as possible. With a good fit, we can ultimately make an assessment of the cause in the event of disappointing energy consumption. Another, somewhat conflicting goal is to make parameter fitting possible with as few sensors as possible, i.e. to be able to estimate parameters well enough on the basis of less data.

Lessons learned on data quality
Apart from these objectives, good data quality is important in order to be able to make statements. That is why we started monitoring the data and analysing the data quality. In TKI Optimaal first basic algoritms were developed to detect possible incorrect data points. The most important lessons from data quality monitoring and analysis are as follows (for more details, see Ch4):

- To assess the quality of a sensor's data, it is important to know what the sensor measures and where the sensor is placed. Sometimes, for example, a sudden short peak can be an actual measurement, for example with a PM2,5 meter next to a gas stove, but sometimes a sudden short peak is physically impossible and it is clearly a spike (for example, a temperature sensor will never heat up and cool down significantly in a very short period of time, due to inertia).
- In addition, good agreement is required between the data being monitored and the parameters used in the model. For example, if we apply a temperature sensor to a wall and try to fit an air temperature in the model, this can lead to less good results because the mass of the wall will dampen the dynamics of the temperature. A good understanding of what is being monitored, how the sensor works, where the sensor is mounted and the consequences this has on the

measuring signal is important for the use of data in models. Standardising and/or automating monitoring data for performance analysis can help in this respect, whereby the link to the parameter in the model used must also be explicitly included here or determined at a later stage on the basis of this input.
- In order to be able to check data quality, it is important that a sensor regularly gives a signal. If a sensor only gives a signal when the measuring signal changes, it is not possible to determine whether the sensor is functioning properly if no signal is received: in that case it is not clear whether there is no signal because nothing has changed, or because the sensor has temporarily or permanently failed.

Next steps
In this project the analysis of the data quality were still project specific. The next step is to develop algorithms to analyse the data quality automatically. The type of sensor and the location where the sensor will be installed will have to be taken into account in these algorithms.

7.3 Conclusions on model and parameter fitting based on monitoring data

Parallel to the process of collecting the monitoring data, we've set up a data-driven RC-network simulation model: a model of the individual houses that we will fit on the monitoring data. As described earlier, the ultimate goal is to be able to determine the parameters in the model with such certainty that the model provides a good description of reality. If this succeeds, it will be possible to see what the cause is in the event of disappointing energy consumption.

The choice for the 3-zone RC network model
A model developed for this purpose will have to do justice to the dynamics of reality on the one hand, but not contain too many parameters on the other hand. With a model that is too simple, such as a monthly model or a one-zone model, it becomes very difficult to determine the cause of an anomaly at the individual level of a single household: the effect of that anomaly must be clearly reflected in the model. A model with too many parameters, such as a TRNSYS model for example, has too many knobs to turn on and therefore too many uncertain components, all of which have to be tuned to reality. That is why we opted for an hourly 3-zone RC network model with a limited number of parameters.

Ch5 describes how this model is constructed and how the parameters are estimated. The initial estimation of the parameters was based on a mix of methods, namely partly on the basis of measurements, partly on the basis of the builder's specifications and partly on the basis of the expert judgement of the researchers.

For two of the houses in Emmen, we succeeded in making a first parameter fit. The model succeeds in matching the following indicators with reality: 1) the actual energy consumption, 2) the actual hourly temperature changes in the different zones for a whole year and 3) the energy signatures. This is a promising result: it is easy to fit a model on one monitoring result, for example the energy consumption, but the fact that the model is able to follow these 3 indicators, including the hourly pattern of the temperatures in the zones, gives some confidence that the model actually represents reality.

There are two comments to be made:

- The fit techniques are not yet of such a nature that we can be completely sure that the parameters are close to the actual characteristics.
- The fit has been successful thanks to the monitoring data that was available, while the explicit goal is to reduce the number of sensors in the houses.

We therefore need ways to determine the parameters in the model with more certainty and to use fewer sensors than at present. To gain more confidence in the estimated value of some of the parameters in the model, a number of sensitivity studies were carried out (see Ch5 and Ch6).

The most important lessons from these analyses are the following (for more details, see Ch5 and Ch6):

- It was already known that ventilation has a major effect on energy consumption. However, it was not yet clear that in very well insulated homes, the effect of windows that are opened slightly is also significant in unheated rooms. This is due to both the temperature equalizing effect due to the good insulation, and the equalizing effect of the heat recovery system when not zoned, which is usually the case.
- The energy loss due to ventilation can be so large that the heating system has too low a capacity to compensate for the losses. In that case, the energy loss appears to be lower, because the energy consumption does not increase significantly, but the temperature in the room decreases significantly. Especially with low temperature systems, this effect will occur more often. This phenomenon is not reflected if you only fit the energy consumption. However, it does occur if you also take into account the temperature development in the different zones in a building. In homes it is known that especially bedrooms are ventilated with windows and grilles. For that reason also it is important to monitor the temperature in bedrooms.
- A comparison between the energy consumption in an average winter week and the coldest winter week showed that ventilation behavior also has an unexpected influence: the energy consumption in the coldest winter week was lower than in the average winter week because the residents closed more windows and doors in the coldest week.
- In very well insulated homes, the effect of the neighbors’ heating behavior is also significant. However, modelling what happens at the neighbors is not so easy, as the temperature in unheated bedrooms at the neighbors is also influenced by the house we consider. The temperature in the bedrooms of the neighbors was unknown. And since the temperature in unheated neighboring bedrooms is strongly influenced by the temperature in 'our bedrooms', especially if 'our' bedroom is heated and the neighbors are not, this is not just negligible. It becomes considerably easier if there are temperature measurements available for both the adjoining living room/kitchen and the adjoining bedrooms.
- Of course, the effects of the various behavioral components considered (setpoint temperature, ventilation through windows and temperature behavior of the neighbors) depend on the frequency and extent to which they occur. For each of the behavioral effects studied, the effect of realistic variations is many tens of percents to sometimes over one hundred percent.
- Finally, it was not easy to obtain more certainty about the actual values of the parameters in the model on the basis of the available monitoring data, let alone

to be able to reduce the number of sensors. The study has resulted in a number of research areas that will enable us to take this a step further. These are described under next steps.

In sum:
Setting up data-driven RC-network simulation models of the first NOW houses has been a successful first step in order to be able to arrive at a realistic analysis of the performance guarantee of individual houses. What the research has shown is that it is feasible to fit a model to monitoring data and to arrive at a good reflection of the actual energy consumption, hourly temperature progression and energy signature. This has been achieved despite the fact that the spread in user behavior of residents and neighbors in particular leads to large variations in these factors. This has been achieved thanks to the fact that we have been able to map out this behavior through monitoring and surveys. To eventually be able to explain at least 80 to 90% of the deviation between the predicted and actual energy and indoor climate performance of individual NoM houses, the big challenge will be to determine the parameters in the model with more certainty and to use fewer sensors than at present. This applies especially, but not exclusively, to the behavioral parameters.

Next steps
There are a number of methods that will help us to get more certainty about the parameters in the model. Some of these methods are already being concretely developed in follow-up projects, have been worked out in research proposals, such as the MMIP3, or will be taken up in future proposals:

- Using a probabilistic model to get more certainty about the parameter estimation: to get insight in the effect of the uncertainty of all parameters in the data-driven RC-network simulation model, we are working on a probabilistic model. Instead of using estimated values in the model, we use probability curves. The curves are based on literature sources where possible. With these curves it is possible to estimate the spread in outcome of the energy use and the temperature profile in the houses. We can do this for all parameters separately, for instance: what is the effect on the energy use for a realistic spread in air tightness or in deviation from the expected insulation level after construction. Or we can investigate this effect for all parameters together. By using brute computer calculation force, we can see which combinations of parameter values will result in the measured energy use, measured hourly temperature profiles and measured energy signatures. With this it will be possible to see if the parameter estimations we made in TKI Optimaal are among the parameter sets that predict the measured data quite closely. A PhD student is developing this module at the moment. The outcome will help our development of the data-driven RC-network simulation models a lot, but is of course no solution for usage on a large scale.
- Using Artificial Intelligence to predict parameters from measured data: We have done a first study to see whether we could predict the use of windows and doors based on the measured data. The results were promising: we were able to predict if a window was open or closed with an accuracy of 80% for all hours of the year. This prediction was done for 2 of the houses in Ermelo with an algorithm that was trained by 2 other houses in Ermelo. We plan to expand the study next year.
- Using fault diagnoses to infer if building components or systems malfunction: In previous TKI and other projects we focused on fault diagnoses based on

monitoring data, but mainly in non-residential buildings. These techniques can also be adapted to houses; what are common faults and what are typical patterns in monitoring data due to these faults. For example: if the mechanical ventilation of a house is out of balance, this has an effect on the mechanical ventilation losses and heat recovery efficiency. I fit is possible to recognize such imbalance on monitoring data, the imbalance can be resolved.
- Using parameter identification by data-driven RC-network modelling: Fitting parameters using data-driven RC-network modelling is a technique that combines physical models with statistical models. The technique is proven for models with only a few parameters: the amount of parameters in the model is limited by the amount of parameters that are monitored. Using this technique in more complex models is new, but might be valuable especially in combination with other techniques. A lesson from data-driven RC-network modelling is that we probably need monitoring data of time series with time steps of 1 to 5 minutes instead of a step of 1 hour which we now often have. This is something to keep in mind in new monitoring project.
- Using physical models to estimate parameters: One of the key parameters in a building model is the building mass. Estimating the building mass from night set back profiles is relatively easy for office buildings and older houses with clear temperature drops at night. However, for NOM houses this proves difficult since the temperature drop at night is quite small. There might be other ways to do this, e.g. by looking for holiday periods or using free floating temperatures in summer.

Another action we are already taken is couling a more detailed ventilation model (COMIS) to the RC network. With this module added to the RC network, it will be possible to take into account the actual indoor air quality performance in addition to actual energy performance and actual thermal comfort performance.

In addition, we will have to make a step in the automation of all parts of the process: the check of the data quality and the data repair, the fitting of the parameters, the fault diagnosis and the performance test.

And last but not least, more reseach is needed to find out what we can learn from data to improve renovation concepts, and espessially how renovation concepts influence behaviour and how that influences the performance of the concept.

8 Signature

Delft, 16 March 2020

Ir. A.C. Westerlaken
Research Manager

TNO

ir. M.E. Spiekman
Author

A. Solar data

All data in Table A-1, Table A-2Table A-3 which is obtained from NEN5060 - 2008 are used to derive the $\mathrm{Solar\ Fraction}$. Table A-1 includes the average solar radiation on the horizontal surface per month. Table A-2 includes the average solar radiation on the vertical surface per month for each orientation. The ratio between the average solar radiation on the vertical surface and the average solar radiation on the horizontal surface gives the $\mathrm{Solar\ Fraction}$ which is shown in Table A-3 per month for each orientation. Finally, to consider the shading by surrounding objects and the dwelling itself, the shading factor ($f_{shading}$) is introduced. The values of the monthly shading factor that are assumed in RC model are given in Table A-4.

Table A-1 Average Solar Radiation on Horizontal Surface per Month

Month	Reference Year	Average Solar Radiation on horizontal surface (0 deg) [W/m2] per month
jan	2003	26.8
feb	2004	49.4
mrt	1992	79.6
apr	2002	164.1
mei	1986	212.3
jun	2000	225.2
jul	2002	199.1
aug	2000	185.9
sep	1992	117.5
okt	2004	72.7
nov	2001	32.6
dec	2003	20.9

Table A-2 Average Solar Radiation on Vertical Surface per Month

Average Solar Radiation [W/m2] on vertical surface (90 deg) per month per orientation								
Month	N	NE	E	SE	S	SW	W	NW
jan	10.5	10.8	19.6	41.3	56.1	44.9	22.3	10.9
feb	18.8	21.2	37.4	59	68.5	53	32.6	20.4
mrt	30.1	34.2	50.5	71.3	82.9	72.3	51.6	34.7
apr	52.6	73.7	112.1	136.9	140.2	133.9	109.5	73.1
mei	68.6	88.9	122	136.9	134.6	143.2	132.8	97.9
jun	81.6	104	130.6	131.9	123.4	142.9	143.6	111.5
jul	70.4	95	121.8	127.6	119.2	122.9	112.5	86.5
aug	60.8	84.1	121.5	140.6	135.5	128.3	109.2	79.1
sep	40.3	49.7	79.3	107	115.1	98.6	73.5	50
okt	25.4	28.7	51	83.6	101.6	81.1	49	28.3
nov	12.9	13.5	23.4	44.3	56.4	44.3	23.5	13.6
dec	8.3	8.5	16.1	36.4	47.2	35.5	15.4	8.5

Table A-3 Solar Fraction on Vertical Surface per Month

Solar Fraction [-] on vertical surface (90 deg) per month per orientation									
Month	**N**	**NE**	**E**	**SE**	**S**	**SW**	**W**	**NW**	**Average**
jan	0.4	0.4	0.7	1.5	2.1	1.7	0.8	0.4	1.0
feb	0.4	0.4	0.8	1.2	1.4	1.1	0.7	0.4	0.8
mrt	0.4	0.4	0.6	0.9	1.0	0.9	0.6	0.4	0.7
apr	0.3	0.4	0.7	0.8	0.9	0.8	0.7	0.4	0.6
mei	0.3	0.4	0.6	0.6	0.6	0.7	0.6	0.5	0.5
jun	0.4	0.5	0.6	0.6	0.5	0.6	0.6	0.5	0.5
jul	0.4	0.5	0.6	0.6	0.6	0.6	0.6	0.4	0.5
aug	0.3	0.5	0.7	0.8	0.7	0.7	0.6	0.4	0.6
sep	0.3	0.4	0.7	0.9	1.0	0.8	0.6	0.4	0.7
okt	0.3	0.4	0.7	1.1	1.4	1.1	0.7	0.4	0.8
nov	0.4	0.4	0.7	1.4	1.7	1.4	0.7	0.4	0.9
dec	0.4	0.4	0.8	1.7	2.3	1.7	0.7	0.4	1.1

Table A-4 Monthly Shading Factor

Month	jan	feb	mrt	apr	mei	jun	jul	aug	sep	okt	nov	dec
Shading Factor [-]	0.5	0.5	0.5	0.5	0.3	0.3	0.3	0.3	0.3	0.5	0.5	0.5

B. Questionnaire

Geachte bewoner en deelnemer aan het TKI Optimaal onderzoek,

1. Belang en doel van het onderzoek

In uw woning vinden al enige tijd metingen plaats en worden binnenkort enkele aanvullende metingen gedaan. Ook hebben we enkele vragen voor u. De metingen en de vragen zijn onderdeel van een onderzoeksproject waarin TNO samenwerkt met bouwbedrijven om inzicht te krijgen in het energiegebruik van energiezuinige woningen. Met de resultaten proberen we in de toekomst nog energiezuinigere woningen te kunnen bouwen, waarin het ook heel plezierig wonen is. De antwoorden op deze vragen zijn daarom heel belangrijk voor ons.

2. Wat wordt er van u verwacht?

We vragen u de onderstaande vragenlijst in te vullen. Uw deelname is vrijwillig. Mocht u één of meerdere vragen niet in willen vullen, dan kunt u het antwoord open laten. U hoeft geen toelichting te geven waarom u een vraag niet wilt beantwoorden.

3. Wat gebeurt er met uw gegevens?

We hechten groot belang aan uw privacy en nemen de daarvoor geldende regels in acht. Uw naam- en adresgegevens worden direct na het onderzoek vernietigd. Uw gegevens zijn slechts toegankelijk voor daartoe bevoegde leden van het onderzoeksteam. Derden hebben geen toegang tot de verzamelde gegevens. In publicaties over het onderzoek zijn de antwoorden van individuele deelnemers op geen enkele wijze herkenbaar. Na afloop van het onderzoek kunnen de geanonimiseerde onderzoeksgegevens nog gedurende 15 jaar worden bewaard.

4. Wilt u verder nog iets weten?

Mocht u vragen hebben over één van de vragen in onderstaande vragenlijst, dan kunt u terecht bij één van de onderzoekers die vandaag/binnenkort bij u in huis aanwezig is.
Bedankt voor u medewerking.

Vragenlijst

Hieronder volgen 16 vragen. Vragen 1 & 2 gaan over de aanwezigheid van personen in uw huis. Vragen 3 t/m 10 gaan over het ventileren van uw huis. Vragen 11 & 12 gaan over het gebruik van de zonwering. Vragen 13 & 14 gaan over uw gebruik van de thermostaat en de verwarming en vragen 15 & 16 tenslotte gaat over uw tevredenheid met het comfort in uw huis. Door inzicht te krijgen in hoe u uw huis gebruikt, krijgen wij een beter inzicht in hoe gedrag samenhangt met een lager of hoger energiegebruik.

1. Met hoeveel personen woont u in uw woning (inclusief uzelf): ☐ personen
2. Kunt u aangeven hoeveel personen er op een gemiddelde dag in uw woning aanwezig (graag in ieder leeg vakje het aantal personen invullen):

	Maandag	Dinsdag	Woensdag	Donderdag	Vrijdag	Zaterdag	Zondag
‘s ochtends							
‘s middags							
‘s avonds							
‘s nachts							

3. Als u een raam open zet, hoe ver staat dat raam dan meestal open? Zet voor ieder seizoen een kruisje voor de meest voorkomende stand.

In de winter:	In het voorjaar/najaar:	In de zomer:
○ Op een kiertje	○ Op een kiertje	○ Op een kiertje
○ In de kiepstand	○ In de kiepstand	○ In de kiepstand
○ Flink ver open	○ Flink ver open	○ Flink ver open
○ Anders, nl: …	○ Anders, nl: …	○ Anders, nl: …

4. Heeft u wel eens (minimaal 1x per week) een raam flink ver open staan, bijvoorbeeld om te luchten? Graag beantwoorden voor ieder seizoen.

In de winter:	In het voorjaar/najaar:	In de zomer:
○ Nee	○ Nee	○ Nee
○ Ja	○ Ja	○ Ja

5. Zo ja, Hoe lang staat dat raam dan zover open? Graag beantwoorden voor ieder seizoen waar u hierboven een ja heeft ingevuld.

In de winter:

- ○ Maximaal een kwartier
- ○ Een kwartier tot een uur
- ○ Langer dan een uur

In het voorjaar/najaar:

- ○ Maximaal een kwartier
- ○ Een kwartier tot een uur
- ○ Langer dan een uur

In de zomer:

- ○ Maximaal een kwartier
- ○ Een kwartier tot een uur
- ○ Langer dan een uur

6. In welke kamers staat er wel eens (minimaal 1x per week) een raam flink ver open? Graag voor ieder seizoen de betreffende kamers aankruisen. (Met betrekking tot de slaapkamers: U kunt zelf bepalen welke nummering u voor welke slaapkamer gebruikt. Als u minder dan 4 slaapkamers heeft, kunt u de niet gebruikte nummers leeg laten.)

In de winter:

- ○ In de woonkamer
- ○ In de keuken
- ○ In de badkamer
- ○ In slaapkamer 1
- ○ In slaapkamer 2
- ○ In slaapkamer 3
- ○ In slaapkamer 4
- ○ Op zolder

In het voorjaar/najaar:

- ○ In de woonkamer
- ○ In de keuken
- ○ In de badkamer
- ○ In slaapkamer 1
- ○ In slaapkamer 2
- ○ In slaapkamer 3
- ○ In slaapkamer 4
- ○ Op zolder

In de zomer:

- ○ In de woonkamer
- ○ In de keuken
- ○ In de badkamer
- ○ In slaapkamer 1
- ○ In slaapkamer 2
- ○ In slaapkamer 3
- ○ In slaapkamer 4
- ○ Op zolder

7. Als u een deur naar buiten open zet (bijvoorbeeld de tuindeur), hoe ver staat die deur dan doorgaans open? Zet voor ieder seizoen een kruisje voor de meest voorkomende stand.

In de winter:

- ○ Op een kiertje
- ○ Flink ver open
- ○ Anders, nl:

…

In het voorjaar/najaar:

- ○ Op een kiertje
- ○ Flink ver open
- ○ Anders, nl:

…

In de zomer:

- ○ Op een kiertje
- ○ Flink ver open
- ○ Anders, nl:

…

8. Gebruikt u tijdens het douchen de timer van de afzuiging?

- ○ Nee, nooit
- ○ Ja, soms
- ○ Ja, altijd

9. Zet u de ventilatie in de keuken hoger tijdens het koken?

- ○ Nee, nooit
- ○ Ja, soms
- ○ Ja, altijd

10. Heeft u een kattenluik?

○ Nee
○ Ja

11. In welke vertrekken maakt u regelmatig gebruik van de zonwering. Graag voor ieder seizoen de betreffende kamers aankruisen. (Met betrekking tot de slaapkamers: U kunt zelf bepalen welke nummering u voor welke slaapkamer gebruikt. Als u minder dan 4 slaapkamers heeft, kunt u de niet gebruikte nummers leeg laten.)

In de winter:	In het voorjaar/najaar:	In de zomer:
○ In de woonkamer	○ In de woonkamer	○ In de woonkamer
○ In de keuken	○ In de keuken	○ In de keuken
○ In de badkamer	○ In de badkamer	○ In de badkamer
○ In slaapkamer 1	○ In slaapkamer 1	○ In slaapkamer 1
○ In slaapkamer 2	○ In slaapkamer 2	○ In slaapkamer 2
○ In slaapkamer 3	○ In slaapkamer 3	○ In slaapkamer 3
○ In slaapkamer 4	○ In slaapkamer 4	○ In slaapkamer 4
○ Op zolder	○ Op zolder	○ Op zolder

12. Als u regelmatig zonwering gebruikt, wanneer gebruikt u deze dan? U mag meerdere opties aankruizen.

☐ De hele dag of een groot deel van de dag
☐ Altijd als de zon schijnt
☐ Als de zon schijnt en iemand thuis is
☐ Als de zon schijnt en het wordt binnen te warm
☐ Als de zon schijnt, het binnen warm wordt en er iemand thuis is
☐ Als de zon mij of iemand anders hindert (bv in de ogen schijnt)
☐ Anders, namelijk: ...

13. Veranderen u of uw huisgenoten in het stookseizoen (oktober tot april) de thermostaat instelling in de woonkamer wel eens? U mag meerdere opties aankruizen.

☐ Nee, hij staat meestal op dezelfde stand
☐ Ja, bij langere afwezigheid (bv een weekend weg of een vakantie)
☐ Ja, bij weg gaan en thuis komen (bv naar werk of de supermarkt)
☐ Ja, bij het opstaan en naar bed gaan
☐ Anders, namelijk: ...

14. Hoe tevreden bent u in het algemeen over het comfort in u woning? Graag uw antwoord aankruisen.

- ○ Zeer tevreden
- ○ Tevreden
- ○ Neutraal
- ○ Ontevreden
- ○ Zeer ontevreden

Kunt u uw antwoord toelichten?

15. Hoe tevreden bent u in over het comfort tijdens het douchen? Graag uw antwoord aankruisen.

- ○ Zeer tevreden
- ○ Tevreden
- ○ Neutraal
- ○ Ontevreden
- ○ Zeer ontevreden

Kunt u uw antwoord toelichten?

Dit is het einde van de vragenlijst.
We danken u hartelijk voor uw medewerking.

C. Parameters of the TRNSYS model

Temperature Profiles
Same temperature profiles are used, as described in section 5.1.3.1.

Heating System
Ideal TRNSYS heating with 6 KW heating power is used.

Solar Gain
As described in section 5.1.3.4, the hourly solar radiation on horizontal surface (Q_{sol}) is available in the KNMI weather station (the nearest is Hoogeveen, STN:279). The solar processor is used in TRNSYS to convert solar radiation on a horizontal surface to the vertical surfaces.

Internal Gain
For the occupancy, same heat gains are used as described in the section 5.1.3.5.1. In order to incorporate the heat gain caused by the electrical appliance, the electricity consumption by the appliance such as TV, stove, washing machine is determined, as discussed in section 5.1.3.5.2.
It is assumed that 90% of the total electricity consumption by the appliances will generate internal heat gain. Moreover, it is also assumed that 70% of the electricity is used on the ground floor, while only 30% is used on the first floor. The radiative and convective part is evenly distributed e.g. 50% for each. Moreover, the electric internal gains are evenly distributed based on per m^2 of thermal zones.

Ventilation System
The TRNSFlow package is used to model the infiltration and the ventilation. TRNFLOW is the integration of the multizone airflow model COMIS (Conjunction of Multizone Infiltration Specialists) into the thermal building module of TRNSYS (Type 56). Ventilation model includes infiltration, natural ventilation (windows and doors opening) and mechanical ventilation. In the RC model, it is assumed that the pressure difference between the indoor and outdoor is 1 Pa and infiltration is determine based on this assumption. Whereas in TRNSYS model, infiltration is based on the dynamic pressure difference between the indoor and the outdoor which depends on the weather conditions. Another difference in the ventilation approach used in the RC and TRNSYS model is how we model the natural ventilation (airflow through the windows and doors), as described in section 5.1.3.6.2 for RC model. In the TRNSYS model, it is based on the wind speed and wind direction along with the internal airlinks among the different zones which might influence the air flow through a windows. The approach used to model these phenomena are discussed below.

Mechanical Ventilation
Balanced mechanical ventilation system with the heat recovery efficiency of 75% is considered. The air is extracted from the Kitchen, the bathroom and the entrance (WC), while the fresh air is supplied to the bedrooms, the living room, and the kitchen. The C_s value is evaluated for the flow rate of $1 m^3/h$ by using the following equation:

$$\dot{\mathbf{m}} = \mathbf{C_s}\,(\Delta \mathbf{P}^{0.66})$$

Here ΔP of 200 Pa is assumed in order to consider the flow resistance for ducting and heat recovery unit. Thereby, the C_s of 0.0000235 kg/s is used to consider the flow resistance through the duct and heat recovery unit. The actual measured ventilation rates for the various ventilation positions (1/2/3), as discussed in the previous chapter, are used as an input for the airflow factor in the model.

Window and Door

The flow through the windows and the doors are modeled by using the large opening in the TRNSFlow. For all the windows, the dimensions of 1 m of height and 0.9 m of width is used. For the front and back door, 2 m height and 1 m width are assumed. Moreover, it is assumed that all windows are bottom-hinged and there is no airflow (leakage) when the windows are fully closed. The measured opening and closing schedules for windows and doors are used as an input for the opening factor of windows and doors in the TRNSFlow. The opening fraction of 10% is assumed in case of windows, however it is 60% in case of the doors opening[13].

Table C-1 shows the average number of hours per week the windows are open (based on the measured data), for the whole measurement periods and also for the winter period only. Winter period begins from November until March.

Table C-1: Average number of hours per week the windows are opened

	Living room window	Back door	Front door	Master Bed	Rear Bed 1	Rear Bed 2	Kitchen window
Average no of hour/week	19.3	3.9	10.7	56.2	19.8	22.0	15.4
Average no of hour/week during winter	0.3	4.6	5.9	53.7	13.7	11.7	12.5

To incorporate the internal circulation of airflow, all internal doors are modeled as a crack by using crack area (area available for the flow) of 120 cm^2. The following formula is used to find out C_s

$$\mathbf{C_s = A\sqrt{\frac{2\,(\Delta P)}{\rho}}}$$

The C_s value of 0.0154 kg/s. m is used. The staircase is modeled as a crack by assuming the crack area of 1 m^2.

Infiltration

Measured infiltration rate (q_{v10}) for the dwelling is 1 L/s.m^2. For both floors, the infiltration rate of 432 m^3/h (q_{v10}) is found which is equal to 1.2 (ACH). This value changes over the year based on wind orientation and wind speed.

Infiltration is modeled by using the cracks in the TRNSFlow with the flow coefficient C_s value which represents the above-mentioned infiltration rate. The following equation is used to find C_s

$$\mathbf{\rho\, q_{V10} = C_s\,(\Delta P^{0.66})}$$

The C_s value is distributed for all zones to get the cumulative C_s value (which is obtained from the above equation).

[13] For the windows, ''Kiepstand'' position is assumed and we use the same opening fraction (10%) in the RC model. In case of doors, it is assumed that if someone opens the door to enter or leave the house, it would be on average 60% open.

D. Cold winter weeks

Coldest Winter Week Analysis for Ventilation Cases

Figure D-1 and Figure D-2 represent the variations in indoor temperature for both the ground floor and the first floor along with the heating power required for the different cases, for the dwelling with both unheated and heated bedroom scenario. The same variations in the first-floor temperature and heating power as we observed in the typical winter week analysis are found, as seen in the following figures.

As we can see in Figure D-1 and Figure D-2, the dwelling with the heated bedrooms need higher heating power to maintain the setpoint compared to the dwelling with the unheated bedrooms. We can also observe the drop in temperature for the second floor when the bedroom windows are fully open. Because heating system has a limited heating capacity and thereby unable to add enough heat immediately to maintain the indoor setpoint.

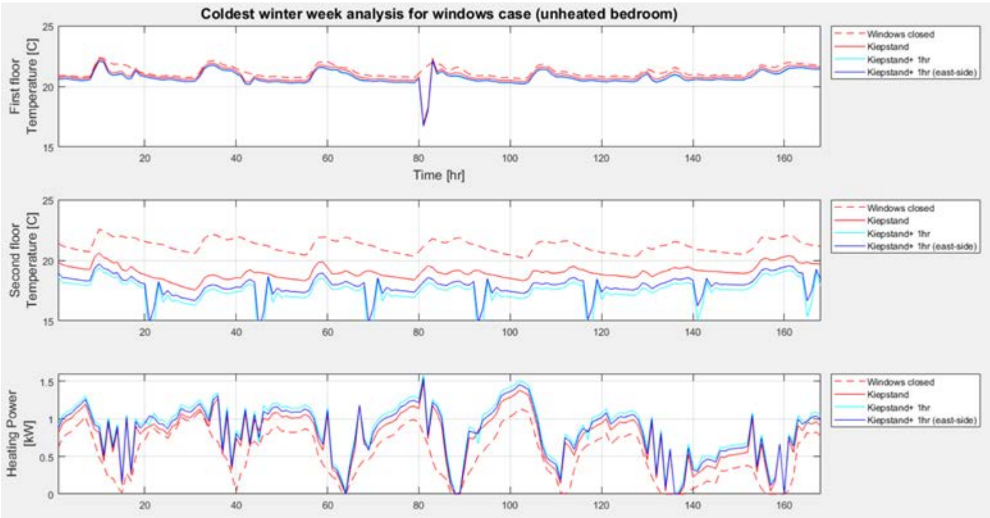


Figure D-1: Coldest winter week analysis for windows case (unheated bedrooms)

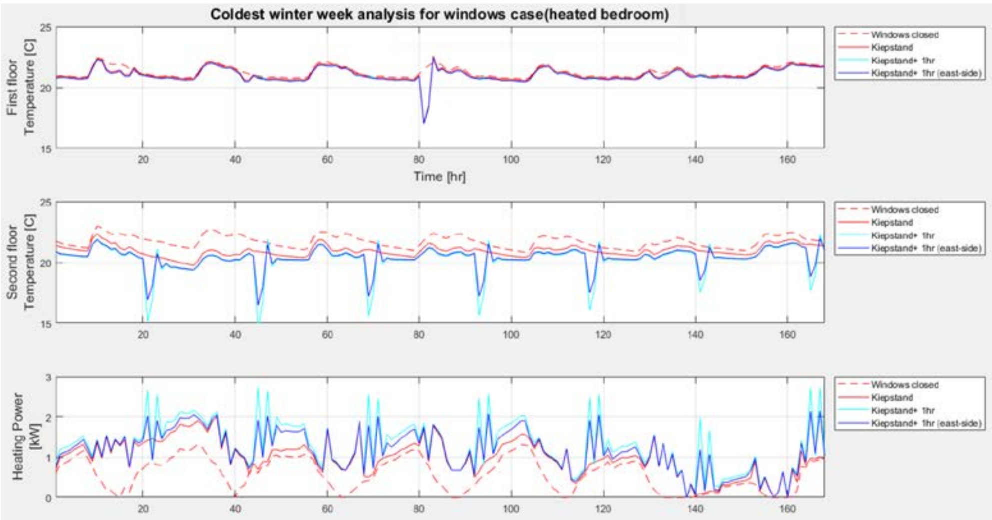


Figure D-2: Coldest winter week analysis for windows case (heated bedrooms

Coldest Winter Week Analysis for Neighbors Cases

Figure D-3 and Figure D-4 represent the variations in indoor temperature for both the ground floor and the first floor along with the heating power required for the different neighour cases, for the dwelling with both unheated and heated bedroom scenario. The same variations in the first-floor temperature and heating power as we observed in the typical winter week analysis are found, as seen in the following figures.

As we can see in Figure D-3 and Figure D-4, the dwelling with the heated bedrooms need higher heating power to maintain the setpoint compared to the dwelling with the unheated bedrooms. We can also observe that the temperature in the second floor is lowest for the dwelling with unheated bedroom scenario when the bedroom floor at the neighbors are unheated.

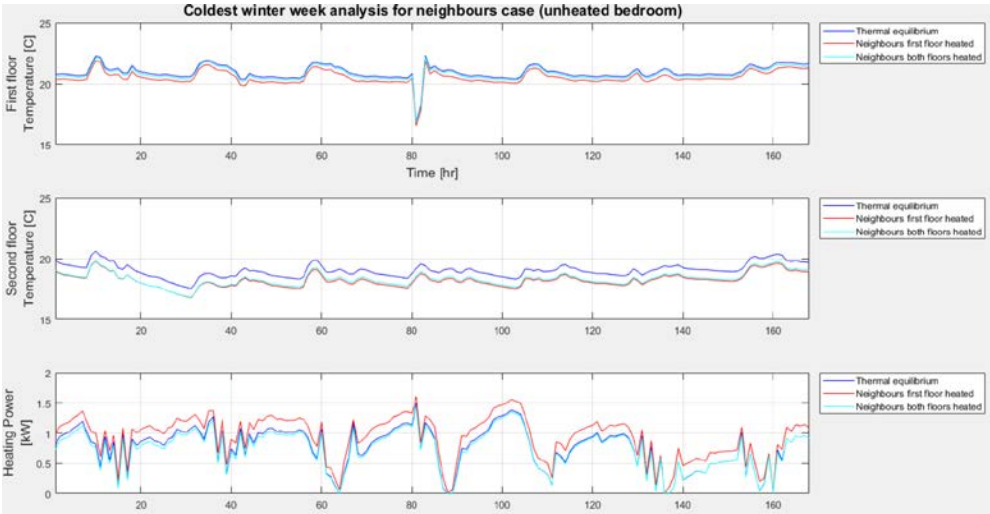


Figure D-3: Coldest winter week analysis for neighbors case (unheated bedroom)

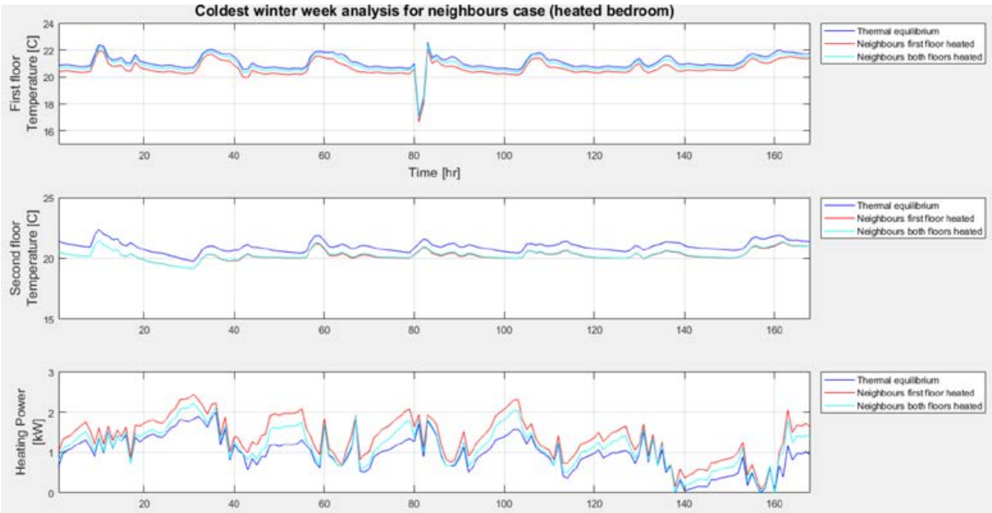


Figure D-4: Coldest winter week analysis for neighbors case (heated bedroom)