The impact of adjusted thermostat practices in the residential sector

T.H. Vosmer

The impact of adjusted thermostat practices in the residential sector

Master thesis submitted to Delft University of Technology in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in Complex System Engineering and Management

by

Thijmen Hendrik Vosmer

Student number: 4010310

To be defended in public on July 3 2018

Graduation committee

Chairperson : Prof.dr. K. Blok, Energy and Industry
First Supervisor : Dr.ir. E.J.l. Chappin, Energy and Industry
Second Supervisor : Prof.dr. L.C.M. Itard, Architecture and the Built Environment
External Supervisor : Dr.ir S. Galsworthy, Quby - Data science

Summary

The residential sector is responsible for over 55% of the natural gas consumption in the Netherlands. In the climate accord of Paris, the Netherlands came to an agreement with the rest of the world leaders to limit the overall temperature rise by reducing the consumption of and by switching away from carbon-based fuels. The gas mining induced earthquakes in the northern part of the Netherlands increases the pressure on Dutch society to reduce natural gas consumption. The residential sector can play a role in reducing the consumption of natural gas in accordance with the Paris accord and to mitigate gas mining induced earthquakes. The amount of natural gas consumed per household is dependent on the behavioural aspects of residents, which are the biggest cause of uncertainty in estimating natural gas consumption. Current natural gas consumption based calculations are based upon dwelling characteristics and are not adjusted to individual behavioural aspects. The inside temperature in a dwelling is seen as the primal indicator of residential heat consumption. The behavioural aspects of residents in the form of thermostat interaction are analysed in this thesis. The potential saving in the residential sector is addressed by including thermostat practices of residents in the estimation of potential savings.

The goal of this thesis is to identify thermostat practice in dwellings by the use of disaggregated energy consumption data and estimate the impact of adjusting individual thermostat practices on the natural gas consumption in the residential sector. Disaggregated energy consumption data is seen as detailed individual household consumption data. To reach the goal the following research question is answered:

> *What insights in thermostat practices that influence natural gas consumption of individual households can be identified by a combined analysis of electricity and thermostat use?*

Practice theory is used to understand the underlying mechanisms at play in household interaction with their thermostat. Disaggregated consumption data is used to gain insights in thermostat practices of individual dwellings. The thermostat practices of households are used to group specific practices and indemnify potential savings in the residential sector.

The smart meter/ thermostat Toon is used to gather individual thermostat interactions and gas and electricity consumption data. Grouping of individual households with the use of clustering on the basis of thermostat settings is used to determine similar thermostat practices. Households with similar thermostat practices are grouped together, with the use of unsupervised classification in the form of hierarchal clustering. Similar thermostat practices groups are used to shape potential thermostat adjustments and assess the impact of these adjustments.

Thermostat practices of households are evaluated with the use of occupancy detection to identify potentials savings. A connection between thermostat practices and household occupancy is made with the use of electricity consumption data. Individual electricity consumption data of households is used to determine the occupancy in a dwelling, by detecting moments of relative high consumption. Residential occupancy is detected with model ensemble of a Hidden Markov Model and a Rolling Mean Model to determine an overall occupancy schedule. The combined analysis of occupancy and thermostat practices is used to determine potential savings and determine the extent of these savings.

The generated insight in saving possibilities by a combination of thermostat practices and residential occupancy is used to develop and estimate the impact of 4 different thermostat practice adjustments. The estimations are based upon 3 different calculation methods: relative shift, heat demand and temperature and gas regression model, to gain a comprehensive understanding of the impact of each of the saving options. The potential savings are expressed in condition based lowering of the thermostat settings inside individual dwellings. The relative shift method calculates the potential saving by relating gas consumption to the difference between in and outside temperature. Heat demand calculations are based upon the number of non-active heating hours for each of the saving options. In the regression method, a linear regression model for every individual household is build, to estimate the gas consumption on the basis of the difference between the in and outside temperature.

More than half of the households in the sample group heats their home during daytime while the occupancy of these dwellings is detected at around 50%. The other half of the population has a clear heating pattern of morning and evening heating. For each of the detected thermostat practices, adjusting thermostat settings result in a potential gas saving. There is a factor 2.5 difference in potential saving between the different thermostat practices. Residents are able to save from 2% to 5% depending on their thermostat practice, resulting in an overall average saving of 34 euro per year. The largest saving potential of lowering the thermostat to 15 degrees overnight. Thermostat practices of households have a bigger impact on the gas consumption than currently used household characteristics. Natural gas consumption of households with similar thermostat practices have shown disparate consumption due to dwelling specific characteristics. Household specifics as thermostat practices and thermal inertia have shown to impact natural gas consumption. The detected thermal inertia of dwellings with the use of disaggregated consumption data is not in line with estimated values based upon physical characteristics of dwellings. The potential savings based upon physical characteristics of dwellings over estimates the potential saving in the residential sector. The impact of thermostat practices and thermal inertia of dwellings determined with the use of disaggregated consumption data is substantial. Current energy consumption studies and energy labelling are based upon physical characteristics of dwellings. The mismatch between estimated energy consumption and measured consumption indicates that estimating residential gas consumption for individual households on the basis household characteristics alone is redundant. By including actual residential consumption data in energy labelling and shifting policy towards adapting household behaviour future policy measures can be improved.

The overall potential savings in the residential sector by adjusting thermostat settings are relatively small but can help to reach the climate goals. The results are based upon occupancy detection in a dwelling by analysing dis-aggregated consumption data. By improving the occupancy detection, the confidence in the results and number of included households can be improved. Expanding the size and representativeness of the sample group improves the applicability and confidence in the results. The personal preferences of residents in both the adoption adjustments and the actual impact on their thermal comfort is unknown. Future research is needed in order to close the gap between actual savings and potential savings by thermostat adjustments and to determine the impact on thermal comfort.

Contents

List of Figures

List of Tables

1

Thesis Definition

In this chapter, the broader context of the research conducted in this thesis is discussed by introducing the topic and explaining the research approach. The relevant aspects of natural gas consumption and space heating in the residential sector are discussed in section 1.1. Followed up by the identifying current knowledge gaps in the assessment of heat consumption in the residential sector in section 1.1. In section 1.2 the central problem following from the knowledge gap is discussed. Followed up by the research objective and research questions. This chapter is ended with a description of the research approach used in section 1.2.3.

1.1. Introduction

The world is currently going through an energy transition to limit the worldwide temperature rise. World leaders came to an agreement in Paris to limit the overall temperature rise to 2 degrees above pre-industrial levels [73], by reducing the overall energy consumption and by finding renewable alternatives for the current carbon-based sources. The refusal of the largest carbon polluter in history, the USA [23], to comply with the Paris agreement increases the overall pressure on the remaining countries to reach their combined goal [34]. The residential sector plays a large role in the overall energy consumption and is responsible for 30% of the overall energy consumption [60]. Unlike the industrial sector the demand for energy in the residential sector is increasing [63]. In the European Union, space heating plays the biggest role in residential energy consumption and is responsible for 57% of the overall energy use [36]. In the Netherlands, natural gas is the predominant source of energy, it accounts for over 55% of overall residential energy consumption [16]. Natural gas is used in various ways within households, from generating hot water and cooking to space heating [7]. In line with the Paris agreement, reducing the overall consumption of natural gas is one of ways to reduce the carbon footprint of the residential sector.

The Dutch government is currently under social and political pressure to reduce the consumption of natural gas even further, due to gas mining induced earthquakes. The extraction of natural gas causes continues earthquakes in the province Groningen, the biggest earthquake measured in past five years was last January [54]. This earthquake has reignited the call to reduce the amount of natural gas mined from the Slochteren natural gas field. Gas field Slochteren is the main natural gas deposit of the Netherlands. The extracted natural gas is used domestically and for the export to neighbouring countries. At this moment 7 million Dutch households are dependent on the low caloric natural gas out of the Slochteren natural gas deposit, with an annual consumption of 9 billion cubic meters [38]. Switching to other natural gas sources at this moment is not an option due to loss of income and lack of natural gas mixing facilities that are needed to make the natural gas suitable for Dutch households. Adapting other natural gas sources to current standard in the Netherlands by mixing natural gas will lead to higher natural gas prices and investment in mixing facilities. Investing in a carbon-based energy source and an increased natural gas price does not make it the preferable option [66]. Long term bilateral contract with neighbouring countries make it impossible to reduce the export on the short therm. Instead reducing gas consumption in the residential sector, is a feasible option to reach the climate goals and mitigate gas mining induced earthquakes.

The amount of natural gas consumed varies between individual households' due to multiple underlying factors. These factors can roughly be divided into two groups: building characteristics and household characteristics. Household characteristics consist of an interaction between indoor climate practices of residents and household composition. In this thesis, the choice has been made to group these aspects due to the inherent relationship between these factors. Indoor climate practices describe the behaviour of residents and composition accounts for the factors that can be used to describe individual households. The practices of residents that influence the amount of natural gas used in a dwelling relate to heating, ventilation and air-condition (HVAC) behaviour [52]. This behaviour is shaped by individual preferences and perception of thermal comfort [15] and can be seen as the main source of uncertainty in predicting energy consumption [71]. Household composition relates to the age, marital status, working conditions and number of residents. For example, families with children tend to use more natural gas per households than singles or couples but consume less natural gas per capita [7].

Building characteristics account for the physical aspects of a dwellings like the type of heating system or size of a dwelling and play a large role in the overall natural gas consumption [70]. Other physical characteristics that play a role in the overall energy consumption are: the age, type of dwelling next to the building materials used [21, 59]. Overall around 40% of the variation in natural gas consumption in Dutch dwellings can be explained by dwelling characteristics [28].

The amount of natural gas consumed in dwellings is strongly related to the heating practices of residents which in turn are influenced by the building and heating characteristics [32]. In order to be able to reduce the overall natural gas consumption in the residential sector, the impact of these factors need to be known. The differences in underlying practices need to be determined in order to make any meaningful claim about the natural gas saving potential in the residential sector.

There are many smart metering devices that measure energy consumption with a range of capabilities on the market. These smart metering devices are able to display and program energy use of a dwellings and are becoming more popular and common [8]. Next to the range in devices there is a diversity in definition and capabilities of a smart meter [14]. In this thesis, a smart metering device is seen as a device that can read or measure and display electricity and natural gas use and is able to be used as a thermostat and can interact with other devices. Smarts meters can be used to provide dwellings with insights and advice about their energy practices [42]. With the possibility to influence energy practices in order to reduce residents' overall footprint [33].

There is an overall demand to reduce the amount of natural gas consumed in the Netherlands, to abate the impact of climate change and to subside the natural gas extraction induced earthquakes. The residential sector plays a large role in the overall natural gas consumption and there is a potential to reduce this consumption by adjusting individual space heating practices. But there is ambiguity around the potential and approach due to the uncertainty of embedded factors that influence natural gas consumption.

Knowledge gap

As stated in the previous sections there are multiple underlying factors that influence the natural gas consumption of individual dwellings. Currently scholars have been able to link different underlying factors to the amount of natural gas consumed in dwellings [46, 70]. These findings have been based upon yearly or monthly energy use and through the use of questionnaires. The current gas consumption models are based upon household and dwelling characteristics [39]. The development and emergence of smart metering devices make it possible to analyse detailed residential consumption data. Detailed energy consumption data can be used to assess the underlying factors that influence natural gas consumption of individual dwellings in more detail. There is a gap in the knowledge on how to use the detailed residential consumption data, in the assessment of individual space heating on a large scale. Disaggregated energy consumption data is seen as detailed high frequency individual household consumption data. There are two parts to the knowledge gap surrounding disaggregated energy consumption data in the assessment of residential gas consumption. Firstly, in what way household behavioural aspects can be determined with the use of disaggregated consumption data. Secondly, the potential heat saving in the residential sector based upon individual household energy assessment.

1.2. Problem demarcation

In this section, the central problem of this research is discussed in combination with the research objective. The conducted research is discussed by presenting the research objective and describing the central research questions. Followed up by an explaining of the research approach used to answer the main question and the sub-questions.

Following from section 1.1 there is drive to reduce the natural gas consumption in the Netherlands due to the climate accord and the mining induced earth quakes. The residential sector is a large consumer of natural gas and there are possibilities to reduce the consumption natural gas within this sector. Due to the differences in heating practices of similar household, there currently is no one size fits all approach to reduce natural gas consumption. Even households with similar dwelling and household characteristics have shown to have different heating practices [24]. Individual heating practices shape the overall consumption of natural gas of individual dwellings. The central issue revolves around how to detect heating practices and determine the influence on the overall natural gas saving potential. There are four parts to this central issue, listed below:

1. *The approach to determine thermostat practices with the use of disaggregated residential consumption data*
2. *The impact of individual thermostat practices on residential heat consumption*
3. *The extent of interaction of residents with their programmable smart thermostat*
4. *The potential saving of adjusting individual household's thermostat practices and how to accomplish this*

The overall effect of individual thermostat practices on residential heat consumption is uncertain. Therefore, there is problem in the estimation of heat consumption in the residential sector. Whereby the estimation of potential savings and impact of policy changes is embedded in behavioural uncertainty [53]. The embedded uncertainty reduces the possibility to make any meaningful assessment of the potential heat saving in residential sector.
In current research, natural gas consuming behaviour is analysed at the hand of interviews and questionnaires [60]. Whereby the scope of the research is limited due to the time intensive manner of research. Through the use of smart meter data, a sizable part of the population can be investigated. Smart meters accommodate the availability of detailed consumption data of a large part of the population. How and to what extent this disaggregated consumption data can be used in the assessment of behavioural aspects is unknown.
The majority of the households do not interact with their smart thermostat, 57% of the people do not interact with their thermostat at all [29]. Therefore, the thermostat settings remain on default and are not fitted to the thermal needs of the residents. To what extent default thermostat settings are in line with the thermal needs of residents is unknown.
Current energy assessment studies make use of characteristics, whereby households are described by a set of characteristics instead of individual thermostat use. The use of disaggregated energy consumption data enables the possibility to assess actual energy consumption data and adjust potential savings accordingly on large scale. The potential saving of adjusting savings to individual household consumption and how to establish these savings is unknown. There still is a gap between the identification of savings and progression of potential savings. The identified potential savings need to be carried out by and in accordance with the residents.

1.2.1. Research objective

There are a couple of hurdles to overcome in order to adjust heating practices to individual households. First of all, the actual heating practices of residents need to be known and understood. The actual heating practices of individual dwellings need to be examined in order to produce a picture that is representative of residential heating practices. By analysing actual heating practices of households with the use of disaggregated energy consumption data, the potential of heating practice adjustments for individual dwellings can be assessed. The use of actual disaggregated data enables the possibility to produce a detailed picture of the potential savings for individual dwellings. The individualistic approach enables the possibility to adjust energy interventions to individual needs of households. Individual thermostat adjustments can in turn be used to reduce the overall natural gas consumption of the residential sector. The goal of this research is to firstly identify to what extent thermostat practices of individual households can be detected, with the use of individual disaggregated consumption data. Secondly, to assess the impact of adjusting thermostat practices on overall natural gas consumption. The goal of this thesis can be reach by answering the research questions in section 1.2.2.

1.2.2. Research question

Thermostat practices of individual users impact the gas consumption in the residential sector. Following from the problem demarcation in section 1.2 several problems in the assessment of thermostat practices need to be addressed. The main research question is presented in combination with the sub-questions below.

What insights in thermostat practices that influence natural gas consumption of individual households can be identified by a combined analysis of electricity and thermostat use?

1. What is practice theory and how can it be used to understand heating practices of Dutch households with the use of disaggregated consumption data?
2. To what extent can clustering be used in the identification of specific thermostat practice of Dutch households?
3. To what extent can disaggregated electricity consumption data be used to evaluate thermostat practices through occupancy detection?
4. What is the potential impact of adjusting thermostat practices of individual households on residential natural gas consumption?

1.2.3. Method description

In this section, the research approach is discussed by giving an overview of the steps taken to answer the research question. Figure 1.1 gives an overview of the structure of the thesis, by displaying the dependencies of the different chapter outcomes and the answering of sub-research questions.

A literature study is used to gain insight in practice and feedback theory concerning residential energy consumption. The literature study is used to provide a theoretical basis for the conducted research and gain the state of the art knowledge. Feedback literature study is used in the discussion of the potential heat saving in the residential sector, to provide insights in the gap between identified potential savings and accomplishing these savings. Practice theory is used to understand behavioural aspects of residents concerting thermostat interactions, to determining thermostat practices with the use of disaggregated consumption data. The knowledge gained by reviewing practice theory is used to develop an approach to identify thermostat practices with the use of disaggregated consumption data. The approach to identify thermostat practices is in turn used in the thermostat clustering analysis. Thermostat practices of residents are the main driver of thermal

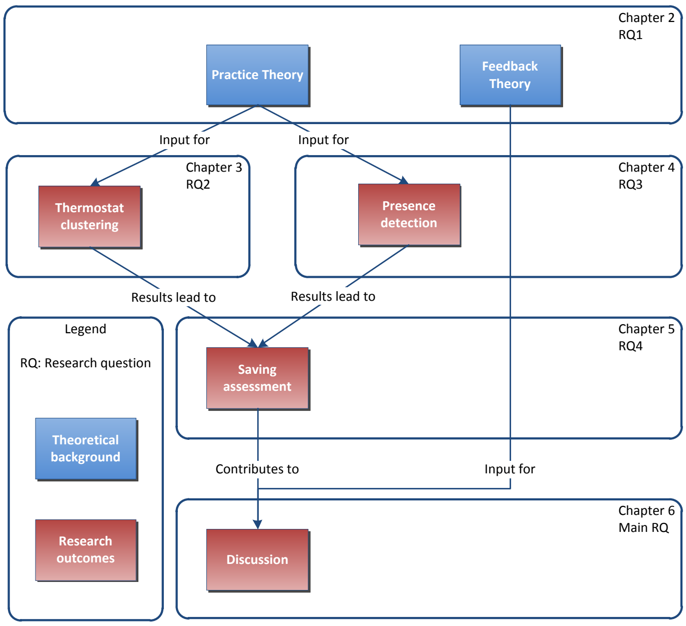


Figure 1.1: Practice theory: elements and example

consumption in the residential sector. But thermostat practice in the sense of thermostat interaction might not be adapted to individual needs due to the lack of interaction. Thermostat interactions are therefore taken into account in combination with occupancy detection.

Disaggregated energy consumption data is used in thermostat clustering to assess the potential of thermostat practice identification. Individual thermostat settings are grouped into profiles and are analysed on the basis of energy consumption and household characteristics. The clustering of thermostat practices is subsequently used to determine the saving potential in the residential sector.

To determine the difference between thermostat practices and actual thermal need of residents, disaggregated electricity data is used for occupancy detection. Individual household electricity consumption data is used to determine the occupancy of residents in a dwelling, whether residents are at home or not. Occupancy detection is used to determine the potential mismatch between thermostat settings and at home presence. The detected residential activity is further used in the saving assessment to adapt thermostat adjustments to individual households.

In the saving assessment, the potential heat saving in the residential sector is analysed by thermostat clustering and presence detection. Whereby potential saving areas are identified and assessed by calculating the impact of thermostat adjustments.

In the discussion, the conducted research is analysed by discussing the limitations of the research and placing the research in context. Furthermore, the research questions are answered and the representability of the identified thermostat practices and saving results is discussed. Feedback theory is used in the discussion to review feasibility of accomplishing the identified residential heat saving.

2

Theoretical background

The biggest cause of uncertainty in the prediction of natural gas consumption in the residential sector is household behaviour Yan et al. [71]. Residential HVAC behavioural actions are the outcome of a complex interplay of multiple underlying factors. To describe the interplay of these underlying HVAC behavioural factors, the current literature concerning practice theory is discussed in section 2.1. Followed up by a literature overview of changing residential energy consumption by feedback mechanisms in section 2.2. In section 2.3 the utilization of practice theory within this thesis is discussed. This chapter is ended by answering the first sub-research question in section 2.4.

2.1. Practice theory

In this section, the emergence of Practice theory is firstly described followed up with the evolution of practice theory concerning energy consumption. This section is ended by the discussion of practice theory with the focus on heating behaviour developed by Gram-Hanssen.

2.1.1. Practice theory emergence

Practice theory describes the workings of how social beings interact with the world in the sense of what mechanisms are at play in changing and interacting with the world. The background of practice theory lies within social theory. Practice theory developed out of philosophical critic by Schatzki on existing social theory by Giddens and Bourdieu [22]. Schatzki developed practice theory to convey his view on the way the social world works, not to create a framework to studying energy consumption. Schatzki described actions of social being as the reaction of what makes sense to them to do. The formulation of practice theory by Schatzki was improved by Reckwitz and further developed by Schatzki himself.

Practice theory developed by Schatzki forms the basis of current energy consumption studies, by identifying main elements in practice theory. General practice theory laid the basis for multiple adaptions of practice theory within different fields, form socio-economic to consumer consumption. Warde was one of the first to relate practice theory to an overall practice based consumption theory. Warde uses practice theory in order to provide new insight into consumption in the form of: how to analyse and structure consumption [69]. Gram-Hanssen combined the different elements that make up a practice described by Schatzki, Warde, Shoeve-Pantzar and Reckwitz into 4 elements that hold practices together [24]. The focus of practice theory developed by Gram-Hanssen is on energy consumption. Based upon the overall practice theory on how the social world works developed by Schatzki. Gram-Hanssen adapted practice theory in order to make practice theory more operational and to understand households' energy consumption, next to developing a way to empirically study consumption by adapting the current practice theories.

Every of the before mentioned adaptions of practice theory identify key elements that together shape a practice. The different authors adapt key elements with similar classifications into one or more new elements. Gram-Hanssen adapted these key elements into a practice theory focused on energy consumption.

2.1.2. Energy consumption based practice theory

There are multiple adaptions of practice theory that have been developed form the original practice theory by Schatzki. The focus in this thesis lies on heat consumption in the residential sector. Therefore, the adaption of practice theory that focuses on energy consumption by Gram-Hanssen is discussed in more detail in this section. To discuss practice theory with the focus on energy consumption a definition of a practice is needed, in the words of Gram-Hanssen [26]:

> "A practice is a collection of sayings and doings performed by individuals but formed and sustained by collectively shared elements."

The primary difference of practice theory with both economic and psychosocial theories of reasoned action is the extent of individual rational within the theory [25]. In practice theory, a practice is seen as a collective that is susceptible to personal differences shaped by attitude and rational knowledge. Instead of actions driven by individual motivation and shaped by individual intention based upon a set of personal believes. Gram-Hanssen identified 4 key elements that form the basis of the collective shared elements that shape a practice. The shared elements that hold a practice together consist of the following 4 elements:

1. *Know-how and embodied habits*, the way a person has taken in things that they have learned and has been socially conditioned and unconsciously continues to do so

2. *Institutionalized knowledge and explicit rules*, the knowledge a person has about something governed by formal and informal institutions in the sense of rules structuring behaviour.

3. *Engagement*, the driver behind what a person wants or means to do and is able to reflect upon. Being able to consciously adapt routines.

4. *Technologies*, shape and guide the direction of a practice. Technology can make a practice more apparent but never determines the action.

The 4 listed elements that are most relevant for the study of energy consumption interact with a practice in multiple ways. First of all, each of the depicted elements can influence multiple practices due to shared elements. Different practices can share on or more elements that hold a practice together. For example, energy saving engagement could on the one hand influence standby practice while on the same time influence heating practice [26].
Vice versa a practice itself can shape one or more of the shared elements defined by Gram-Hanssen as well. Whereby a change in one practice can influence another practice. One practice can influence another practice through the induced change in one of the shared elements between the practices. For example, a change in indoor climate practice might influence the heating know-how of a person. Which in turn could lead to a person advocating at work to change their overall energy consumption.
Reducing heat demand in the residential sector enable the possibility to reach climate goals and mitigate gas minding induced earthquakes. In the assessment of residential heat consumption, indoor climate practice is seen as the biggest source of uncertainty in predicting the gas consumption. How the elements that govern practices influence indoor climate practice is illustrated in figure 2.1.
Each of the elements shown in figure 2.1 that govern indoor climate practice are described in more detail by an example. In the case of heating know-how, someone might heat their home constantly during the day even when they are not at home. The practice of heating a dwelling the entire day might purely be shaped by fact that a person has always done so. Independently of the knowledge a person has about the implications of their actions. On the other hand, someone who has knowledge about the inefficiency of heating a dwelling constantly during the day, might adjust their practice according to their heating system knowledge. Their practice might entail the turning down of the thermostat when no one is actually at home. Concerning heating technology, the possibility to automatically program a thermostat generates the possibility to automatically turn down the thermostat when it is likely that residents are not at home. Whereby technology is able to change the possibility of interaction with a practice but in turn does not shape the practice itself. The last element revolves around the motivation behind a certain indoor climate practice. Someone might have the knowledge and ability to lower the thermostat at night but is not willing to do so. Because they might prefer a warm sleeping room or are not engaged in saving energy.

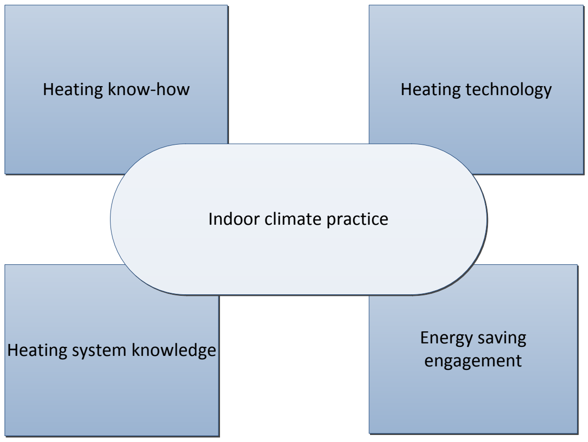


Figure 2.1: Practice theory: elements and example of indoor climate practice [26]

2.2. Feedback devices and energy consumption

There is a gap between the identification of residential heat saving opportunities and actual saving. A transition from identified savings to actual saving has to be made. As discussed in the previous section there are multiple factors involved in the indoor climate practice and in turn adapting this practice. In line with the development of smart metering devices the possibility to provide residents with insights about their energy consumption has grown. Providing energy related feedback to individual residents is seen as way to make this transition [67]. Disaggregated consumption data gathered with a smart metering devices is used to determine residential heating practices and saving potential. Therefore, the focus of feedback theory within this chapter is on the use of in home feedback devices. In section 2.2.1 the notion of in home feedback devices is discussed followed up by a general framework concerting the interaction of feedback devices and heating behaviour. This section is ended with an overview of the overall discussion of the applicability of feedback devices.

2.2.1. Feedback devices and framework

In this section, the impact of feedback devices on energy consumption is discussed. Due to the development and emergence of smart meters it is now possible to monitor real-time energy consumption and provide residents with real time feedback about their energy consumption. Providing residents with information about their energy consumption is an effective tool to promote efficient behaviour [10]. One of the ways to provide residents with feedback about their energy consumption is though In-Home Displays (IHDs). There are a range of possibilities to provide feedback to households through IHDs. From only providing insight in current energy use to tailored energy advice that may be embedded in financial Incentives. There are more factors at play in providing feedback to consumer than the IHDs and the feedback itself. These factors relate to the interaction between a person and the feedback device itself.

General feedback framework

The construct of providing feedback to residents to change their indoor climate practice has been moulded into a framework by Jensen et al. [37]. In this section, the framework by Jenssen is discussed in line with the overlap with practice theory by Gram-Hanssen. The framework by Jensen shows the relationship between feedback devices and heating practice while leaving room for specific individual cases.

The framework developed by Jenssen shown in figure 2.2 describes the interaction of feedback devices and heating behaviour. By identifying the human and technology interaction that is involved in feedback devices and indoor climate practices. The framework makes a clear distinction between the technical feedback loop and human behaviour involved in residential heating. The framework shows how a feedback device interacts between the feedback performance and heating behaviour. Both the interactions of a person with the heating system with and without a feedback device are described by the framework. The framework indicates how a feedback device can influence heating behaviour of residents, by persuasion and situated awareness. Persuasion revolves around convincing resident to adapt their indoor climate practice. Persuasion can take all sorts

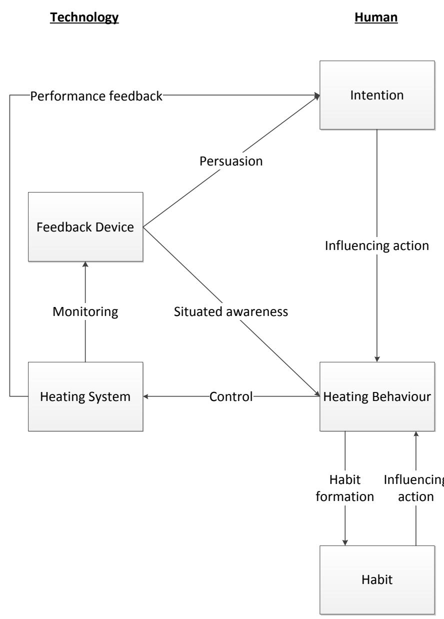


Figure 2.2: Operation of feedback changing heating habit [37]

of shapes and forms from financial incentives to providing energy related information. Situated awareness provides residents with some sort additional information by making them aware of their actions, by providing direct feedback of the implication of their actions. The framework can be used to understand the impact of a feedback device on heating practice while maintaining the human behavioural structure.
Practice theory developed by Gram-Hanssen focusses on the elements that hold a practice together. The framework provided by Jensen shows the role of a feedback device in the change of heating behaviour. Practice theory by Gram-Hanssen and the feedback framework developed by Jenssen share overlapping elements. The way the feedback device interacts with the heating behaviour of occupants, namely by: persuasion and situated awareness. Persuasion focusses actively on the residents' energy saving engagement by providing information on energy use or by financial or competitive incentives. Situated awareness tries to change the behaviour of the resident unconscious and thereby focusses on the heating know-how. Secondly, there is an overlap in the role of technology in both theories/frameworks. Technology is a way to shape the indoor climate practice but does not determine the action.

2.2.2. Feedback and contradictory results

Providing feedback to residents to ensure more efficient energy use is reviewed in this section, by debating the impact of feedback devices on energy consumption. Through the discussion of differences in study outcomes in providing feedback to residents and discussing disagreements between scholars. Contradictions in feedback results are discussed to gain a deeper understanding of the complex intertwined factor involved in providing feedback to residents.
There is no consensus between scholars about the effectiveness of feedback devices and energy consumption reduction. There are the studies that have shown a positive effect of feedback devices and reducing energy consumption [1, 18, 58]. But on the other hand, scholars have found several implications that need to be include when looking to feedback devices to reduce energy consumption. The findings of 3 different authors on the effectiveness or elements that can ensure feedback effectiveness are discussed. Firstly, Fischer [19] indicated that the successfulness of the feedback is dependent on the frequency, feedback provision manner and understandability of the provided feedback. Secondly, Šćepanović et al. [57] argue for the role of context in the effectiveness of energy consumption reduction through feedback devices and identifies specific approaches for distinct cases. And thirdly, Buchanan et al. [9] expresses the importance of residential engagement, in order for feedback devices to work and warns for the potential for unintended consequences. Unintended

rebound effects have been detected in the use of feedback through IHDs. The saving in energy consumption might be re-invested in other appliances and in turn reduce the overall effect of behaviour changing feedback [9].

There is general agreement between scholars that providing residents with feedback about their energy consumption could result in a reduced energy consumption. But a successful implementation of providing feedback to residents is dependent on multiple intertwined factors. Residential energy feedback implementation is therefore not as straightforward as providing information to residents about their energy consumption.

The framework by Jensen is used to understand the role of feedback devices within the heating behaviour of residents, by the identification of the 2 ways that a feedback device can influence heating behaviour. The disagreement of scholars about the effectiveness of providing feedback to residents in energy saving is used in the discussion of potential savings. The gained insight in the debate surrounding feedback mechanisms can help to close the gap between potential and actual saving.

2.3. Application of practice theory

In this section, the application practice theory by Gram-Hanssen focused on energy consumption within this thesis is discussed. The behavioural differences of people play an important role in the dissimilarities in residential heat consumption of comparable households. Completely similar dwellings have shown dissimilarity in residential heat consumption of a factor 3 difference [27], due to the difference in behavioural aspects of residents.

Thermostat interactions are one of the residential behavioural aspects that influence the heat consumption of a dwelling. Residents interact with the thermostat to achieve thermal comfort:

> "the condition of mind which expresses satisfaction with the thermal environment."[4]

Thermal comfort is perceived differently between individuals and is furthermore dependent on airflow, clothing and activities performed in the last hour [48]. The impact of household behaviour on heat consumption has to be taken into account in the assessment of potential savings, because of the large differences between energy consumption of similar households and induced uncertainty in predicting residential heat consumption.

Practice theory by Gram-Hanssen on the assessment of energy consumption in the residential sector has been tested by the use of extensive interviews and questionnaires [25]. The time-consuming manner of in the field research is needed to deceiver the key factors that underlie practices. Due to the time-consuming nature of this kind of research, the applicable scale is limited. To get a further understanding of indoor climate practices on a wider a scale a different research method is needed.

To increase the size of the reachable sample group, energy consumption data of the smart meter/ thermostat Toon is used to identify indoor climate practices. To analyse the 4 elements that shape indoor climate practices extensive field research is needed. Therefore, the focus in this thesis lays at the measurable part of the indoor climate practice, the practice itself.

Many studies agree that the indoor temperature can be seen as a good indicator of the impact of residential behaviour on heat consumption [32]. The programmed thermostat settings in a smart thermostat control the inside temperature in a dwelling directly. The indoor temperature settings in the smart meter/ thermostat Toon are used to determine the effect of residential behaviour in indoor climate practices. Indoor climate practices relate to more aspect than thermostat settings alone, for example ventilating. Indoor temperature is a good indicator of the impact of residential behaviour and the indoor temperature is directly controlled by Toon. Therefore, the focus of this research is on the thermostat practices of residents within the context of climate practice.

Thermostat practices of residents are assessed through the detected thermostat settings of individual users, to firstly assess to what extent smart meter/thermostat data can be used in the determination of individual thermostat practices. Secondly to determine the impact of residential thermostat practice on the saving potential in the residential sector. And lastly to determine if thermostat practice with the use of detailed energy consumption data can be used to explain the gap between prediction of residential energy consumption and actual energy consumption.

2.4. Conclusion: sub-research question 1

The first sub-research question of this thesis relates to practice theory on heating practices and use of disaggregated consumption data. In the following way:

What is practice theory and how can it be used to understand heating practices of Dutch households with the use of disaggregated consumption data

Practice theory describes the interaction of people with the world in a way that make sense to them to do. Practice theory has been developed over the years with multiple applications in different fields. Gram-Hansen adopted practice theory to energy consumption by describing the 4 factors that govern energy consumption practices. Indoor heating practices are governed by: heating know-how, heating system knowledge, heating technology and energy saving engagement. The governing elements shape a practice and can in turn be shaped by a practice. The 4 key factors are used to understand a resulting interaction of person with world in the sense of a practice. The governing elements have been analysed for multiple consumption studies by the use of interviews with residents by Gram-Hanssen. The use of disaggregated consumption data imposes an alternative way to assess the indoor climate practice.
The use of disaggregated consumption data in the assessment of indoor practice shift the focus form the 4 key factors to the resulting action. Since the resulting thermostat practice can be detected in the disaggregated consumption data. Detailed energy consumption data can be used to analyse the indoor climate practice aspect of thermostat interaction, to determine similar patterns in thermostat settings and resulting energy consumption. Whereby households can be classified on the basis of their thermostat practice instead of household/ dwelling characteristics, to gain insights in distinctive heat consuming patterns of different thermostat practices.

3 Clustering

The thermostat practices of in this chapter are assessed with the use of thermostat settings of individual households. The thermostat settings are used to determine if thermostat settings of individual dwellings can be grouped together to find similar thermostat practices. A form of automated classification is used to group similar thermostat practices. By grouping thermostat settings, commonalities of household interaction with their thermostat can be explored. The thermostat settings or individual household that make use of the smart thermostat Toon are used to analyse thermostat practices. Grouping household thermal behaviour on the basis of thermostat settings enables the incorporation of household behaviour in energy assessments. Household groups with similar thermostat practices can be used to identify inefficient practices and adapt possible energy interventions accordingly. To group the individual thermostat settings, households with similar thermostat settings are grouped by the use of clustering. Clustering is the unsupervised classification of patterns in data or features into different groups [35]. By the use of various clustering algorithms similarities in different object can be grouped together. Clustering is used in a wide arrange of different scientific fields from pattern recognition to artificial intelligence. Overall clustering is used to group sets of data that are unlabelled and is seen one of the most useful techniques for discovering patterns in underlying data [30]. Combining clustering with the use of a smart thermostat makes it possible to assess thermal household behaviour in a non-intrusive manner. Behaviour aspects can be assessed by remotely determining behavioural aspect in contrast to the commonly used questionnaires.
This chapter begins with a description of the data used in the clustering analysis. In section 3.2 the clustering process is discussed by describing in detail the use and selection of the clustering method. Followed up by presenting the clustering results in section 3.3 and clustering validation in section 3.4. This chapter is ended with the answering of the second sub-research question in section 3.5.

3.1. Clustering data

This section begins with describing the data used for the clustering analysis, by clarifying the data treatment, data source and collection time span. Followed up by comparing the sample data to the general population in section 3.1.1. This section is ended with a preliminary insight in the data used in the clustering analysis in section 3.1.2.
In order to determine thermostat practice of individual dwelling the thermostat stetting of these dwellings need to be gathered. The thermostat settings of individual users are collected through the use of the smart thermostat Toon developed by Quby. This smart thermostat/meter is able on the one hand control the temperature in a dwelling and on the other hand measure gas and electric consumption by interacting with various types of meters. The thermostat settings in Toon consists of two elements: thermostat schedule and manual adjustments. The scheduled settings form the basis of the thermostat settings which can manually be adjusted by the residents. The thermostat settings in Toon, a combination of the programmed and manually adjusted thermostat program, are used in the clustering analysis.
There is a trade-off between the level of detail of the thermostat settings and the computer power needed to analyse these settings. For the clustering analysis, the hourly thermostat settings are gathered by averaging the thermostat settings for every hour. To ensure the functionality of the clustering analysis and because one hour intervals are the default thermostat program interval frequency. Taking the mean thermostat setting for

every hour limits the loss of information within the time interval.
The thermostat settings are collected for a 3-month period in the winter of 2016-2017 for 1319 households. In order to assess the thermostat practices in households there must to be a thermal need. Therefore, the 3 coldest months are chosen of the 2016-2017: December, January and February [41]. The households are selected on their willingness to share their data. Within the entire Toon user population there is a select group of people who have agreed to share their data for research purposes. This group is used in the analysis of thermostat settings.
As described in section 2.1 the information about household heating practices through the measurement of thermostat settings is limited. The individual alterations cannot be explained by measuring the thermostat settings alone. Therefore, the focus of this thesis is on the overall thermostat settings. Meaning that the 3 month individual households' thermostat settings are combined to produce an overall thermostat settings picture. Similar weekdays are combined to determine the overall thermostat settings within the 3-month measuring period. The overall thermostat settings are used to assess the impact of commonly used thermostat practice on the gas consumption in dwellings.
Within the 3-month period only the weekdays are taken into account, the weekends are left out of the scope, due to the different habits of resident during the week and in the weekend. The Dutch statistic agency has shown that residents have different sleeping habits and participate in different outdoor activities during weekends at inconsistent moments [13]. In addition, the data showed that dwellings heat their homes differently during the week compared to weekends. The weekends are separated to mitigate the cluttering of the clusters by irregular weekend habits. Because residential habits in weekends are irregular and inconsistent an overall weekend thermostat practice is not representative. Therefore, weekends are left out of the scope in the determination of thermostat practices of residents.
To account for the residents that have distinct habits for separate weekdays, for example part-time workers, the weekdays are separated. This means that for every user the thermostat settings for similar weekdays are taken into account independently in the clustering. To determine the overall thermostat settings of residents the hourly median temperature settings is taken of similar weekdays, for each of the 1319 households. Meaning that one data case consists of 24 thermostat set points for a single household on a specific weekday. The median hourly data set points for all the 1319 households for every weekday are combined into one data set. To discuss the thermostat settings of an individual household for similar weekdays the following terminology is henceforth used: user weekday.
Households may set their thermostat differently on the same hour for similar weekdays. To ensure that the clustering analysis explores the overall thermostat practices the standard derivation of thermostat settings has been taken into account. For user weekdays, the standard derivation is calculated for similar hours by the use of equation 3.1 for every user.

$$s_x = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}} \tag{3.1}$$

With s_x as the standard derivation, X_i as the thermostat set point for similar hours on similar weekdays, $\overline{X}$ as the mean of X_i and n as the number of days in the 3-month period of a distinct weekday. To illustrate, for a single user all the thermostat set points of a Monday at 8 o'clock are used to calculate the standard derivation in thermostat settings. The average is taken of the hourly standard derivation per user weekday. When the average standard derivation is above 2 degrees the data case is excluded from the clustering analysis. Meaning that for a single user that particular weekday, for example all Mondays, are not used in the clustering analysis. In other words, when the average difference between the mean hourly thermostat set point and the hourly set points for separate days is above 2 degrees. The thermostat settings of that particular weekday for a single user are not taken into account in the clustering analysis.
In table 3.1 the number of households that have thermostat settings within the standard derivation of 2 degrees are shown. Next to the percentage of users that do not have a similar thermostat program for a particular weekday. To give an overall description of the hourly thermostat settings within the data, the standard derivation and variance are shown for each weekday. Table 3.1 shows that on Friday the largest number of households are excluded from analysis due to differences in thermostat settings on Fridays, followed up by Mondays. Both the standard derivation and variance remain similar between weekdays after the household exclusion on the basis of exceeding standard derivation.

Table 3.1: Weekday overview of thermostat settings in sample group

	Number of Households	*Excluded number of Households*	*Standard derivation*	*Variance*
Monday	1231	6.67%	7.83	2.8
Thueseday	1257	4.70%	8.03	2.84
Wednesday	1250	4.23%	7.94	2.82
Thursday	1247	4.45%	7.96	2.82
Friday	1216	7.81%	7.77	2.78

3.1.1. Data population comparison

To give an indication to what extent the sample group is representative of the general household population, the sample group is compared with data from the Dutch statistics bureau (CBS) shown in table 3.2, 3.3, 3.4 and 3.5. The tables compare the household size, build period, dwelling type and dwelling age. The tables show that the data population is not representative for the Dutch population. Some groups are clearly under or overrepresented in the data set. The number of one person households are strongly under-represented in the sample group while the 4 person households are over-represented. There are three build periods within the study that show a mismatch around 10 percentage point with the Dutch population shown in table 3.3. The semi-detached and detached dwellings are under-represented in the clustering data while the terraced dwellings are over-represented, shown in figure 3.4. The dwelling sizes are more closely matched indicated in figure 3.5.

Table 3.2: Representativity check: Household size comparison

Household size	*CBS*	*Clustering Data*
1 Person	37.41%	9.40%
2 Persons	32.89%	37.15%
3 Persons	11.93%	17.44%
4 Persoons	12.58%	27.14%
5 or more persons	4.64%	8.87%

Table 3.3: Representativity check: Build period comparison

Build period	*CBS*	*Clustering Data*
Until 1946	16.32%	11.14%
1946-1964	27.62%	11.37%
1965-1974	14.59%	14.78%
1975-1987	12.84%	26.23%
1988-1999	16.01%	17.13%
2000-2009	8.41%	18.56%
2010-now	4.22%	0.76%

Table 3.4: Representativity check: Dwelling type

Dwelling type	*CBS*	*Clustering data*
Apartment	15%	12%
Terraced	42%	66%
Semi-detached	21%	14%
Detached	23%	7%

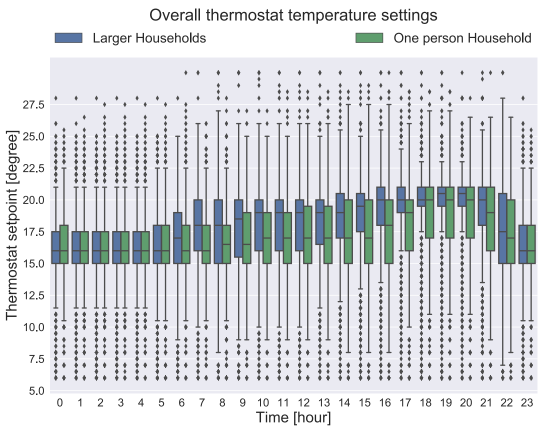
Table 3.5: Representativity check: Dwelling size

Dwelling size [m^2]	*CBS*	*Clustering data*
100-150	68%	75%
150-250	27%	21%
250-500	5%	4%

3.1.2. Preliminarily insights in thermostat data

In this section, an indication of the impact of individual behaviour in residential heat consumption is given, in combination with the discussing similarities in thermostat use of individual dwellings. The data is gathered by the smart thermostat Toon for both the heat consumption and thermostat use.
Figure 3.1 is used to give an indication about the uncertainty in energy consumption predications due to individual behaviour. The number of residents is one of the characteristics used to predict residential heat consumption. In figure 3.1, the thermostat settings of one person households and larger households are shown. The thermostat settings consist of a combination of programmed and manual adjusted settings. The one person and larger households are spitted to portray the biggest difference in household size, the larger households have similar thermostat settings. Figure 3.1 shows the box-plot of the thermostat temperature settings of individual households. The box itself entails the interquartile range accounting for 50% of the thermostat set points, between the first and third quartile. The stripe within the box shows the median thermostat set point for that hour of the corresponding household size. The two stripes of the top and bottom line

show the 95% range of the thermostat set points, with the dots extending these lines account for the outliers. Household size is one of the characteristics that influences heating practices and in turn gas consumption [7]. First of all, figure 3.1 indicates that there is a large overlap between the different household sizes. Overall the different household sizes show a distinguishable heating setting pattern. Household size is one of the indicators used to predict residential heat consumption. Whereby the prediction of heat demand on the basis of more indictors will improve the estimation. Figure 3.1 does indicate that estimations based upon household size will result in sizable uncertainty, due to the large overlap in thermostat settings for different household size groups. There especially is an overlap in thermostat settings for the larger households' sizes.
Both groups show similar thermostat settings around sleeping hours but differ during the day and evenings. The spread in thermostat settings of one person household during the day, is one the one hand caused by households who heat their home during the day and on the other hand household who do not. The overlap in thermostat settings of one person households and larger households in combination with the differences in thermostat settings of one person households, indicates the need for a different separating method to classify individual households.
Figure 3.2 gives an overview of the moments at which households turn their thermostat up a down, with the y-axis showing the percentage of daily thermostat changes at that hour. As in figure 3.1 the thermostats settings consist of both programed and manual changes. More than 90% of the households change their temperatures daily, whereby the figure describes thermostat changes for a sizable part of the population. Figure 3.2 shows clear similar thermostat changes with figure 3.1. In the sense of the moment when residents change their thermostat and the difference in the average thermostat settings. Figure 3.2 illustrates that the majority of the households turn their thermostat on and off in the morning. Although not every household turns the thermostat down again. Around 50% of the daily thermostat changes are caused by turning down the thermostat in the evening, in a time span of a few hours.



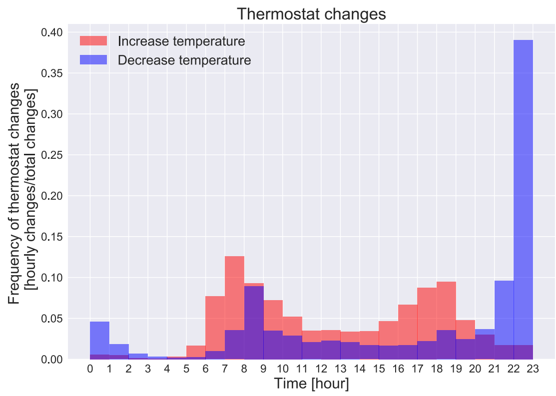


Figure 3.1: Overall thermostat settings for different household sizes Figure 3.2: Thermostat changes during the day

3.2. Clustering: Method and selection

The clustering method to group similar thermostat practices is described in detail in this section. By firstly discussing the different types of clustering and the substantiated clustering type choice. In section 3.2.2 different clustering similarity measures are discussed followed up by the selected similarity measure in section 3.2.3. This section is ended with evaluating the different clustering methods and stopping condition in section 3.2.4.

3.2.1. Clustering types

There are multiple clustering applications and different methods to separate diverse types of data into groups. The main clustering types are discussed to select the best fitting clustering type to grouping thermostats settings of individual dwellings. The main clustering types can be divided into 4 groups: Partitional clustering, Hierarchical clustering, Density-based clustering and Grid based clustering [35]. A description of the different clustering types is listed below.

- *Partitional clustering*, the separation into an integer number of disjoint clusters that optimises a certain criterion function.

- *Hierarchical clustering*, separates clusters into smaller ones (top-down) or combines smaller clusters into larger clusters (bottom up).
- *Density-based clustering*, grouping of neighbouring data sets based upon density conditions.
- *Grid based clustering*, spatial data mining by quantising space into a finite number of cells to perform operations on.

In this research, the choice has been made to use hierarchical clustering, why this clustering type is chosen for the clustering of thermostat practices follows. The clustering is used to find similar thermostat practices, to what extent there are similar thermostat practices between individual households is unknown. Due to the explorative nature of the use of the clustering analysis, a clustering technique that is based upon an ex ante information about clustering groups is not the ideal candidate. Therefore, the choice for partitional clustering which uses an optimised criterion function has not been made.
Heating practices of households may differ at certain moments of the day but have overlapping characteristic on other moments as shown in section 3.1.2. For example, residents with similar overnight thermostat settings may change the thermostat differently during the day. The distinction between similar and dissimilar thermostat settings is harder to make with density based clustering analysis. Therefore, the clustering by density based clustering has been excluded.
Grid based clustering techniques are able to distinguish between arbitrary shapes in the data set and can deal with noise in data [30]. As described in the section 3.1 the noise in the clustering data is limited due to the data processing technique. The hourly thermostat settings of similar weekdays are averaged and household weekdays with a standard derivation above 2 degrees are excluded. The need for shape recognition is limited due to finite shape possibilities in hourly thermostat settings and aim to find overall similar thermostat settings. Grid based clustering is not best fitting candidate due to the limited amount of noise and need for shape separation in the clustering of thermostat practices.
By the use of symmetric similarity measures the hierarchal clustering determines which separate cases to cluster together [61]. The symmetric similarity measure can be seen as a calculation method to determine the distance between two individual cases. In hieratical clustering, every case is seen as a separate cluster in the start of the clustering. The two clusters that have the largest symmetric similarity are clustered together, this process is repeated until a stopping condition is met. Out of the 4 clustering types hierarchical clustering is most suitable to find similar thermostat practices of individual dwellings. Hierarchical clustering is most suitable because the aim of the method is combine data cases that have similar thermostat settings together into groups. The other clustering techniques have capabilities that are not in line with the aim of finding similar thermostat settings. The different similarity measures and the stopping condition are discussed in the following section.

3.2.2. Clustering similarity measures

There are multiple ways to find similarities between data cases and calculate the distances between the different clusters. The similarity measure determines which clusters are grouped together into larger clusters. There are multiple methods to determine the distance between individual clusters, the 7 agglomerative clustering schemes provided by Aerts et al. [2] are discussed below. For all the different agglomerative clustering schemes the Euclidean distance metric is used to determine distance metrics shown in formula 3.2.2. Meaning that the distance between clusters is determined by the use of the Euclidian distance metric. For example, the distance metric between cluster q and p is calculated by combining the distance in thermostat set points of every hour h in the 24-hour user weekday settings.

$$d(p,q) = \sqrt{\sum_{h=1}^{24} (q_h - p_h)^2} \tag{3.2}$$

- *Single*, also known as the Nearest Point Algorithm calculates the minimum distance between all points between different clusters.
- *Complete*, also known as the Farthest Point Algorithm calculates the maximum distance between all points between different clusters.

- *Average*, takes all the points of the clusters into account and calculates the average distance between these points. Also known as the UPGMA (Unweighed Pair Group Method with Arithmetic Mean)
- *Weighted*, also known as the WPGMA (Weighed Pair Group Method with Arithmetic Mean) which makes use of a differential weight to the initial clustering when calculating the distance metric.
- *Centroid*, calculates the distance between the centroid of clusters and recalculates the centroid for every formed cluster. Also known as the UPGMC (Unweighed Pair Group Method with Centroid)
- *Median*, like the centroid method but takes the median distance between centroids. Also known as the WPGMC (Weighted Pair Group Method with Clustering Averaging)
- *Ward*, clusters groups with minimization of variance between similarities of every point in different clusters.

The different clustering methods have been compared to each other to select the clustering method that generates the best results. The methods are compared to determine the method that is most powerful in identifying and separating different thermostat practices. The 7 previously discussed methods have been compared by looking at the output for different clusters sizes for every method. The similarity measure is calculated differently in all the different clustering methods. As discussed in section 3.2.2 the hierarchical clustering methods combines clusters until a stopping condition is met. A stopping condition on the basis of a similar ex ante distance threshold cannot be used to compare similarity measures due to the differences in distance calculation. Therefore, the number of clusters has been used as a stopping condition for comparing the similarity measures.
For every clustering method the cluster sizes 5, 10 and 15 where selected as a stopping condition, the results are shown in table 3.6. The aim of the clustering is to find similar thermostat settings therefore the clusters that represent a significant part of the population are used. The clusters that contain at least 5% of the user population are called the significant clusters, the number of significant clusters for each of the similarity measures are shown in Table 3.6. For example, the similarity measure average with the stopping condition of 5 clusters has only one significant cluster. Meaning that 1 cluster contains almost the entire population and that the other 4 clusters each contain less than 5% of the population. To find similar thermostat practices the clustering that is able to find similar patterns that represent a large part of the population is preferable. Table 3.6 is used to determine the power of the methods to separate the different thermostat practice into common habit groups.

Table 3.6: Clustering separation power for different stopping conditions

Cluster method \ Cluster size	5	10	15
	Significant clusters		
Single	1	1	1
Complete	4	4	6
Average	1	1	1
Weighted	2	2	2
Centroid	1	2	2
Median	1	1	1
Ward	5	6	8

Table 3.7: Weighted standard derivation of hourly thermostat settings

	Clustering method	
Cluster size	*Complete*	*Ward*
5	1.684	1.578
10	1.615	1.478
15	1.562	1.431

3.2.3. Clustering similarity measure selection

The chosen similarity measure used in the hierarchical clustering of thermostat settings is discussed in this section. The 2 most promising candidates are firstly discussed followed up by the choice between these two similarity measures. Table 3.6 show that there are 2 similarity measures that are able to separate thermostat settings in more than 2 similar thermostat practices that contain a significant part of the population. The complete and Ward similarity measure distinguish more than 2 significant clusters for each of the cluster size stopping conditions.
The best candidate is determined by a combination of separating power, visual conformation and a weighted

standard derivation. Table 3.6 indicates that the Ward method generates the most significant clusters under similar stopping conditions. Table 3.6 shows that increasing the number of clusters from 5 to 10 does not generate an additional significant cluster in the Complete method. Indicating that the additional formed clusters do not represent a sizable part of the population. In appendix B for each of the clusters sizes the clustering results of the Complete and Ward method are shown. The figures graphically display the hourly thermostat settings of user weekdays in a box-plot.
There are 2 insights that the clustering with Complete and Ward similarity method provide, for different cluster sizes. First of all, there are visual similarities between the cluster generated in the Complete and the Ward similarity method shown in appendix B. The second insight is the overlap in thermostat practices within the larger cluster sizes for each of the methods. When comparing cluster size 15 for both similarity methods shown in appendix B.4 and appendix B.7. Clusters 2,3 and 4,5 generated by the Complete method shown in appendix B.4 show visual resemblance. For the Ward method shown in appendix B.7 clusters 1,2 and 6,7 resemble one another. Visually comparing the clusters formed in the Complete and Ward similarity method reveals the power of the similarity method to distinguish thermostat practices. The additional significant clusters in Ward method illustrates a distinct thermostat practice. For example, when comparing cluster size 5 of the Complete similarity method in appendix B.1, with the Ward similarity method in appendix B.5. The Ward similarity method is able to distinguish an additional significant cluster that illustrates a distinct thermostat practice of relative high thermostat settings.
Table 3.7 shows the weighted standard derivation of hourly thermostat settings within the different formed clusters. For each of the significant clusters the hourly standard derivation is calculated for the user weekdays. The mean of the standard derivation weighted by the size of the cluster is shown in table 3.7. Table 3.7 indicates that for each of the cluster sizes the standard derivation within the clusters is smaller for the Ward similarity method. The weighted standard derivation of the hourly thermostat settings indicates the similarity in thermostat settings of user weekdays within the clusters.
The Ward similarity method is used in this research in the clustering of thermostat practices of individual users. Because the Ward method is most powerful in separating thermostat practices of individual households, shown in table 3.6. Secondly, the additional clusters that are formed in the Ward method consist of different thermostat practices shown in appendix B. At last the Ward similarity method has a lower standard derivation within the formed clusters compared to the Complete method.

Ward method explanation

To provide insight in how the Ward clustering method works the formula used to calculate the similarity between clusters is discussed with an example. The Ward similarity method combines cluster on the basis of the least variance between the clusters. In the clustering of thermostat practices the variance between each of the 24-hour thermostat settings of individual users/ clusters is calculated. The Ward distance calculation method is shown in equation 3.3. Equation 3.3 shows the distance calculation of two clusters one that consist of user weekdays p and q and the other out of user weekday k.

$$d(p\cup q,k)=\sqrt{\sum_{h=1}^{24}\left(\frac{n_{p,h}+n_{k,h}}{n_{p,h}+n_{q,h}+n_{k,h}}\right)(p_h-k_h)^2+\sum_{h=1}^{24}\left(\frac{n_{q,h}+n_{k,h}}{n_{p,h}+n_{q,h}+n_{k,h}}\right)(q_h-k_h)^2-\sum_{h=1}^{24}\left(\frac{n_{k,h}}{n_{p,h}+n_{q,h}+n_{k,h}}\right)(p_h-q_h)^2} \tag{3.3}$$

$d(p\cup q,k)$ is the Ward similarity measure distance between one cluster consisting of user weekdays p and q and a cluster consisting of user weekday k. p,q,k consist of the overall 24 hourly thermostat setting for a particular user weekday. n accounts for the number of user weekday thermostat settings in the relative cluster at time h.
To give an example of how the clustering methods works 2 clusters are considered that both consist of 2 thermostat set points, one cluster consisting of user weekday p and q and a cluster consisting of user weekday k. The example uses 2 thermostat set points instead of 24 to simplify the example. Cluster $p\cup q$: p[18,20], q [19,20] and k [17,16]. The distance between these clusters is calculated the following way: $d(p\cup q,k)=\frac{1+1}{1+1+1}*(18-17)^2+\frac{2}{3}*(20-16)^2+\frac{2}{3}*(19-17)^2+\frac{2}{3}*(20-16)^2-\frac{1}{3}*(18-19)^2-\frac{1}{3}*(20-20)^2$
$d(p\cup q,k)=4.93$
The example indicates how the variance in hourly thermostat setting is taken into account in the Ward similarity measure. The difference in hourly thermostat settings between each of the user weekday within a cluster is compared with a candidate cluster.

3.2.4. Stopping condition for cluster formation

The extent of combining individual users with similar thermostat practices into clusters is determined by the stopping condition. In this section, the stopping condition for the formation of clusters is discussed, by explaining the selected stopping condition and the resulting clusters. As discussed in section 3.2.3, the Ward method clusters on the basis of least variance between clusters. The variance between hourly thermostat settings of individual users is unknown, due to the exploratory nature of the clustering. Possible typical thermostat practices are determined with the use of thermostat settings of individual dwellings. Therefore, there is no prior knowledge about the distance metric calculated by the Ward method. Because there is no prior knowledge of distance metrics between different thermostat practices, the number of cluster is chosen as the stopping condition. The other reason for selecting the clusters size as a stopping condition, is the direct link that can be created between the number of observed thermostat practices and number of clusters.

The number of clusters is determined by a combination of the Elbow method, number of significant clusters

Table 3.8: Stopping condition: cluster size Ward method

	Cluster size			
	4	*6*	*10*	*14*
Number of significant clusters	4	6	6	7
Excluded users [%]	0	0%	11%	22%
Overlaping thermostat practices	0	0	1	2

and visual conformation. The Elbow method developed by Thorndike [62] is used to determine the cluster sizes where the change in the distance between the formed clusters is the biggest, in this case the change in variance. Figure 3.3 shows the Ward distance between the last formed clusters in the blue line and the change in slope in the green line. Because the Ward method combines clusters with the least variance, the variance of the last formed clusters can be seen as the largest distance between in the overall clustering. The change in slope has peaks at cluster sizes 4, 6, 10, 14. Indicating that at these cluster sizes an additional cluster, results in the biggest reduction in variance within the last formed clusters. In order to group households with similar thermostat practices the variance within a cluster should be minimized. While keeping in mind that individuals are inherently different and in that sense their thermostat practices as well. Therefore, there is a there is a trade-off between the variance within the clusters and the ability to detect typical thermostat practices. To determine the cluster size while keeping this trade-off in mind, cluster sizes 4,6,10,14 are compared to each other shown in table 3.8.

Table 3.8 gives an overview of the number of significant clusters, excluded users and similar practices for the different cluster sizes. In line with section 3.2.2 the power of the clustering to group users into significant clusters is used. Significant cluster are the clusters that represent a significant size of the population. The exclusion of clusters that do not represent a significant part of the population, at least 5%, results in the exclusion of users shown in row excluded users. In appendix C the detected thermostat practice for each of the cluster sizes is shown. Visual conformation is used in the detection of typical thermostat practices, expressed in the number of similar practise in table 3.8. For example, when looking at cluster size 10 in appendix C.3 clusters 4-5 visually show similar thermostat practices. In the sense of similar thermostat settings and thermostat adjustments during the night and day, with minor differences in thermostat set point temperature and the moment when the thermostat is turned on and off. The aim of the clustering is to discover similar thermostat practices, whereby the overlapping practices should be minimized.

Cluster sizes 10 and 14 are excluded due to the high loss of users through insignificant clusters and overlapping thermostat practices between the clusters. For both the cluster sizes 4 and 6 no users are excluded and there are no overlapping thermostat practices. Clusters size 6 has been chosen as the final cluster size due increase of number of significant clusters that portray different practices compared to cluster size 4.

To visualize and provide an overview of the performed hierarchical clustering a denrogram is used. Figure 3.4 provides a schematic overview of the distances at which the different clusters are merged. The horizontal lines indicate at what distance the clusters are merged and each vertical line represent a cluster at that distance. The different colours show the clusters that are grouped for cluster size 6. The y-axis shows the Ward distance of the last formed cluster. By going up and down in the denrogram the merging and splitting of clusters is graphically shown. Figure 3.4 indicates that cluster 6-3 and 4-5 are clustered together for cluster size 4. At Ward distance 300 these clusters have the highest variance and are therefore split, whereby the cluster size increases from 4 to 6. Clusters with the least distance between them are clustered together. The other insight the denrogram provides is why the number of significant clusters drops when the cluster size is increased.

Figure 3.3: Cluster size determination: Elbow method

Increasing the number of clusters results in an increase in the number of clusters that represent less than 5% of the population.

3.3. Clustering results

The cluster outcomes are presented and discussed in this section, by discussing visual representation of the clusters. Followed up by a description of the users that are represented in these clusters. In this section, the 6 clusters that represent a thermostat practice are discussed. A visual representation of the different clusters is given by a box-plot of the hourly thermostat settings of all the user weekdays. As discussed in section 3.1 every user is represented by the daily hourly thermostat settings for each of the weekdays from Monday until Friday. Meaning that the each of the figures shows the box plot of independent daily thermostat settings for each user weekday.

3.3.1. Cluster illustration

In this section, the clustering outcomes are shown graphically in figure 3.6 and discussed by a short description. The user weekdays are grouped in to 6 clusters: Day lowering, Morning-evening program, Adaptable program, Day heating, Gradient day heating and Continues heating in figure 3.6. The figure displays the box plots of hourly the user weekday thermostat settings and the mean thermostat setting for each of the clusters. The box-plot shows the midspread of the 50% of the weekday user thermostat settings within the box, and the 95% spread of the thermostat settings within the lines. The outliers are shown as dots in the graph for every hour.

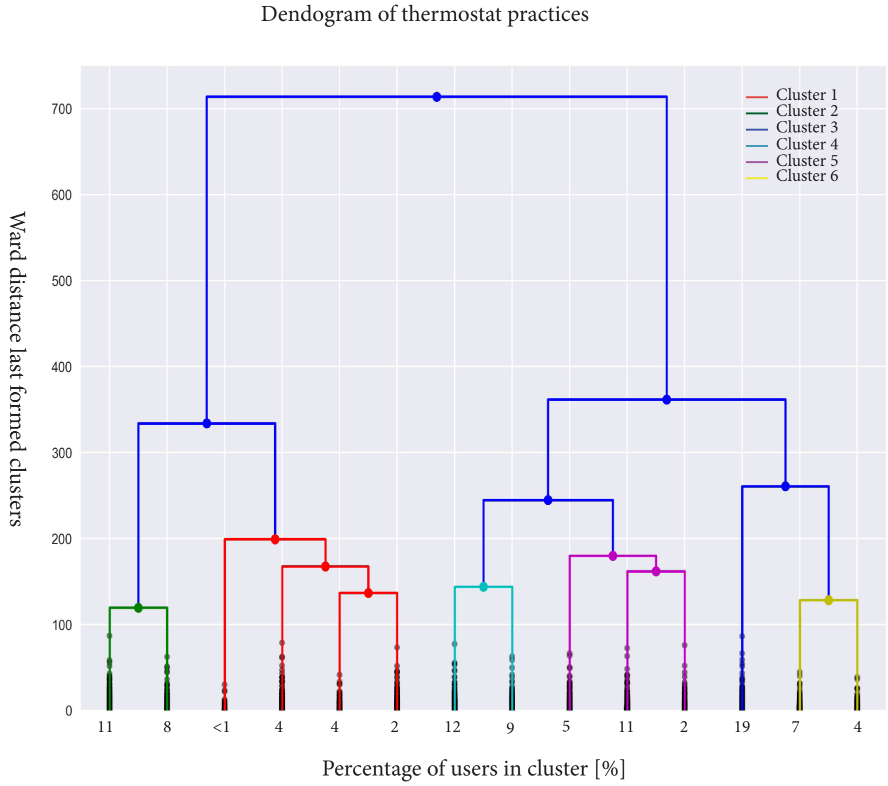


Figure 3.4: Hierarchical cluster formation: denrogram

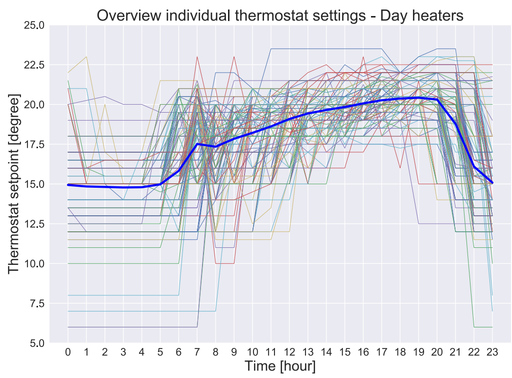


Figure 3.5: Thermostat settings of individual user weekdays

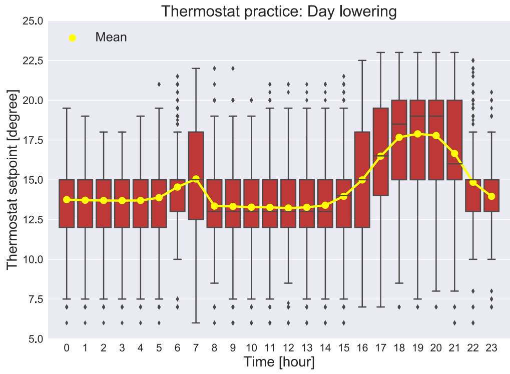
Table 3.9: Spread of users between clusters overview

	Percentage of users in cluster [%]
Day lowering	11
Morning-evening program	18
Adaptable program	19
Day heating	21
Gradient day heating	18
Continues heating	12

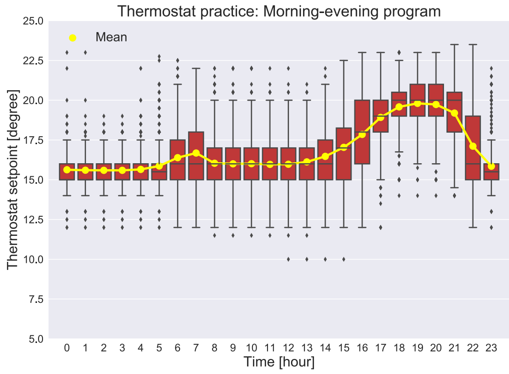
In order to explain the boxplot figures used to describe the different clusters an overview of the individual thermostat settings in given. Figure 3.5 displays all the individual thermostat settings of every user weekday included in the cluster. Figure 3.5 shows the user weekday thermostat settings for the cluster: Day heating. Because the distribution of the thermostat settings is lost in plotting every individual thermostat setting the boxplots are used to describe the different clusters. In the following paragraphs the independent clusters are discussed by giving a short description and discussing the important aspects of each cluster.

Day lowering

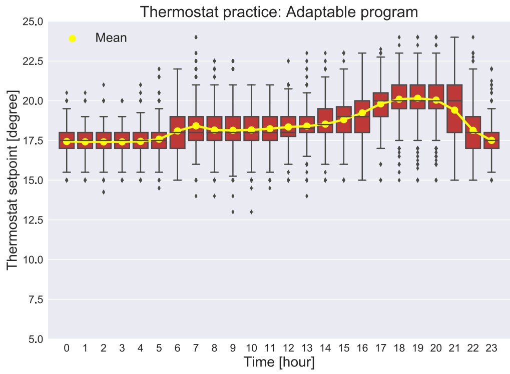
Figure 3.6a gives an overview of the thermostat settings within the Day lowering cluster, accounting for 10% of the population. The spread of thermostat settings within this cluster is large compared to the other clusters. While 75% of the user weekdays set their thermostat below 15 degrees during the night and day. The



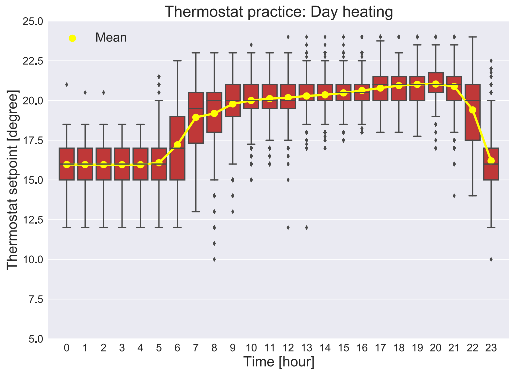
(a) Day lowering



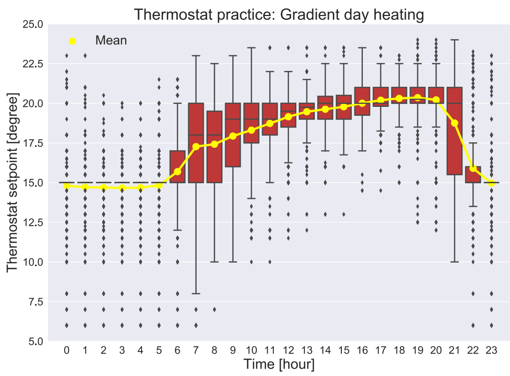
(b) Morning-evening program



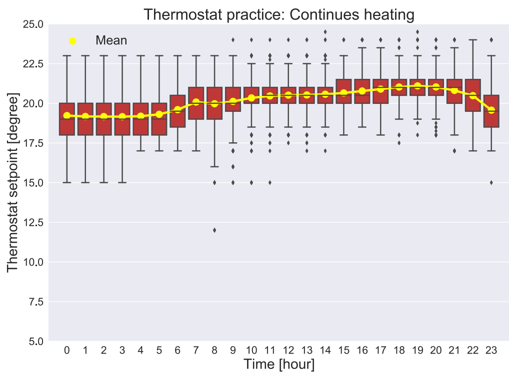
(c) Adaptable program



(d) Day heating



(e) Gradient day heating



(f) Continues heating

Figure 3.6: Clustering overview – cluster size 6

peak in thermostat settings in the morning indicates that some of the users within this cluster increase their thermostat in the morning while the majority of the users keep it below 15 degrees. The spread indicates that this cluster represents on the one had users that keep their thermostat constantly low and on the other hand users that relatively have lower thermostat settings, compared to the other clusters.

Morning-evening program

The Morning-evening thermostat program shown in figure 3.6b illustrates the user weekdays with increased thermostat settings during the morning and evening. With the majority of user weekday thermostat settings below 17 degrees during the night and day. The thermostat settings correspond with a working schedule with increased activity in the morning around 7 and in the evening starting from 5 o'clock. The spread in thermostat settings during the increasing and declining slope indicate a mixture of thermostat settings. The result of increasing and decreasing the thermostat settings hour apart, as shown in figure 3.5. The individual user weekday thermostat in figure 3.5 shows an increase in thermostat settings from 6 o'clock for some users while other users increase the thermostat in the successive hours.

Adaptable program
The spread in thermostat settings of user weekdays with an Adaptable program is limited shown in figure 3.6c. The user weekdays with an Adaptable program overall show a different thermostat setting between night and day. Compared to the Morning-evening program the difference between the overnight and day settings is limited. Indicating that the thermostat is turned on for 75% of the user weekdays the entire day, with a minimum of 17 degrees. The Slight adaptable heating shows a small thermostat peak in the morning, in line with the Day lowering and Morning-evening program.

Day Heating
Figure 3.6d show the overall thermostat settings of the Day heating program. As the name suggest the user weekdays within this cluster heat their homes during the day. During the night the thermostat is lowered on average around 16 degrees and turned on starting from 5 o'clock. As shown in figure 3.5 not every user weekday thermostat setting is turned down to 16 degrees at night. But for 75% of the user weekdays the settings are below 17 degrees during the night. The thermostat settings during the day are around 20 degrees with a small slope from 9 o'clock until 5'oclok. This slope can be explained by comparing figure 3.6d to figure 3.5, some of the users turn their thermostat on later during the day.

Gradient day heating
Gradient day heating shown in figure 3.6e shows overlapping thermostat settings with the Day heating with 2 differences: slope during the day and overnight settings. At 7 o'clock the Gradient day heating shows a spread in thermostat settings accounting for 50% of the population from 15 to 20 degrees. This spread decreases during the day, indicating that the Gradient day heating represent a more diverse before afternoon thermostat program than the Day heating. The second difference is the overnight thermostat settings, more than 75% of the population has an overnight thermostat setting below 15 degrees. For 50% of the population the thermostat settings are at 15 degrees which corresponds to the default Toon overnight thermostat settings.

Continues heating
The thermostat weekdays of a turned-on thermostat setting for the entire day are shown in figure 3.6f. 75% of the population within the Continues heating clusters has set their thermostat above 18 degrees for the entire day, with an even higher set point during the day around 20 degrees. The spread of thermostat settings indicates that some of the user weekday lower their thermostat overnight but an even larger part has turn-on their thermostat even above the 20 degrees.

Overall overlapping findings
In this section, some of the insights that are present in most of the clusters are discussed. The spread in thermostat settings during the morning around 7 and at bedtime is larger compared to different hours. This can be explained by the hourly difference in when people get up and go to sleep. Figure 3.5 clearly shows the differences in first moment people turn down their thermostat around 10 and 11 o'clock in the evening. Figure 3.5 show similar downwards slopes in the evening in successive hours.
There is a clear overlap in moments when residents turn their thermostat up and down shown in figure 3.2 and thermostat changes shown in 3.6. The thermostat practice of adjusting the thermostat settings in the morning and in the evening, is present in all the clusters. The moments when residents' adjuster their thermostat is similar while the thermostat settings itself is clearly different shown in figure 3.6. The spread within the clusters indicate that the overall thermostat practice within a cluster is similar with slightly different thermostat settings.
The Morning-evening program and Adaptable program at first sight display similar looking practices, the difference is the thermostat set point temperature. The temperature set point of the Morning-evening program during the day and night is relatively lower compared to the Adaptable program.
In 3 of the 6 clusters consist of dwellings that are heated during the day, consisting of Day heating, Gradient day heating and Continues heating. Table 3.9 shows that these clusters account for the majority of the population, 51%. Indicating that there is a potential for reducing gas consumption in individual dwellings. Overall clear distinct thermostat practices are detected with the use of hierarchical clustering with similar aspects.

3.3.2. Cluster result exploration

In this section, the clustering of thermostat settings of individual households into 6 different thermostat practices is further analysed. Practices are analysed by looking into 3 different aspects: gas consumption per cluster, household size per cluster and movement of users between clusters.

Daily gas consumption per cluster

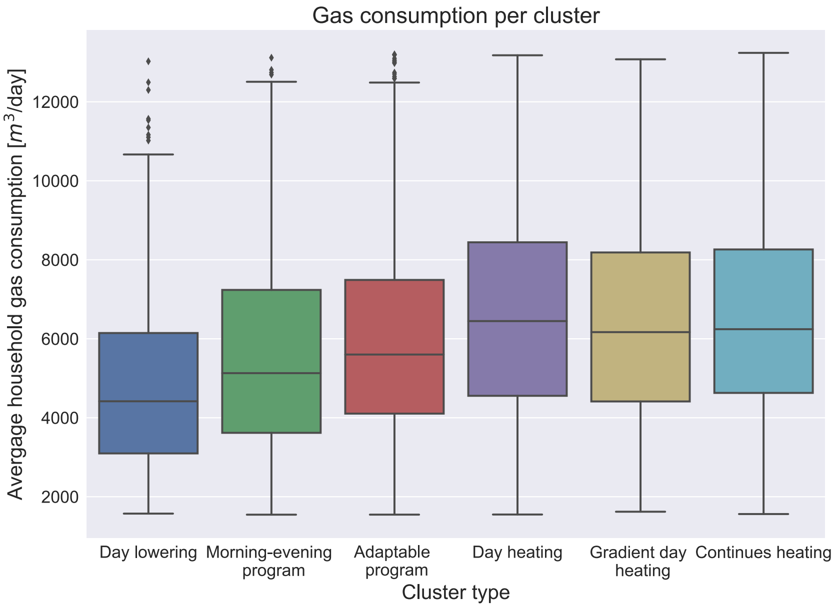


Figure 3.7: Overview of average daily gas consumption per cluster

In figure 3.7 the average daily gas consumption of individual households is plotted in a boxplot for every cluster. The box-plot shows the midspread of the 50% of the average daily gas consumption within the box, and the 95% spread of daily gas consumption within the lines. The average gas consumption is calculated by taking the mean of the daily gas consumption for every household for the same 3-month winter period used in the clustering analysis. All of the clusters display an overlap in average daily gas consumption of individual users. As discussed in chapter 1 there are multiple underlying factors that influence the gas consumption of dwellings that can be divided into building characteristics and household characteristics. The thermostat settings are an element of the HVAC behaviour, namely the heating. Because there are more elements at play than the thermostat settings in the gas consumption of individual dwellings an overlap in gas consumption is expected.

The referencing index letter used in table 3.11 and corresponding thermostat cluster is shown in table 3.10. Figure 3.7 show a clear difference in average daily gas consumption in the clusters in which residents lower their thermostat during the day clusters [A,B,C] and the clusters who do not [D,E,F]. There still is a difference in the midspread and median gas consumption of the cluster who heat their homes the entire day, but minor. The difference in heating patterns and gas consumption indicates that there is an overall gas saving potential for lowering the thermostat during the day. To give an indication of the spread of average gas consumption for different clusters a one–sided ANOVA is used, shown in table 3.11. Table 3.11 clearly shows that there is a significant variance between the average daily gas consumption of all the different clusters. The variance in gas consumption between the clusters who do not heat their home during the day, clusters [a,b,c], is lower than the all the clusters but still significant. This is not the case for the clusters [d,e,f], the variance in average daily gas consumption is low and not significant. Indicating that the influence of other gas influencing factors, like dwelling characteristics becomes more predominant when a dwelling is heated for the larger part of the day.

The declining effect of thermostat practices on the daily gas consumption of dissimilar dwellings is expected. When a dwelling is heated for a majority of the day the impact of thermostat practices is mitigated and dwelling specific characteristics become dominant. The difference in daily gas consumption for different heating practices generate insight the possible saving extent for separate thermostat practices.

Household sizes and clusters

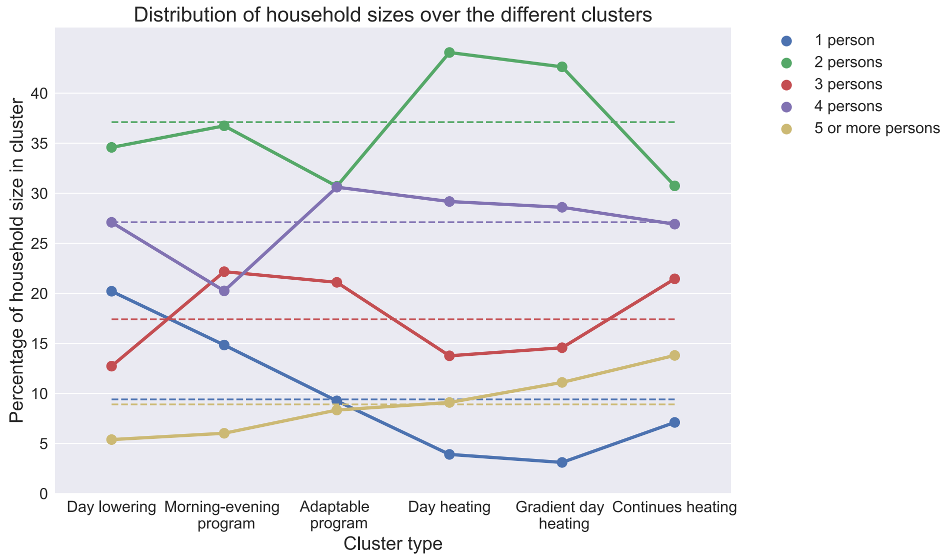


Figure 3.8: Distribution of household sizes per cluster

Different household compositions cause disparate household thermal needs. To give an indication of the composition of household size and thermostat program clustering, figure 3.8 is used. Figure 3.8 gives an overview of the configuration of household size for every cluster and the overall average in the dotted line. In line with the representability of the sample group discussed in section 3.1.1, the household sizes within the data group are not evenly spread.
Figure 3.8 shows that the variance between clusters is largest for the 1 and 2 person households while smallest for the 4 and 5 or more person households. The 1 person households are clearly underrepresented in every clusters except the morning-evening program and day lowering cluster. While the 2 person households are overrepresented in the day heating clusters and underrepresented in the Adaptable program and constant heating cluster. The 3 person households seem to be mixed between the clusters. The largest divergence for the 3 person households is shown in the morning-evening program and day lowering. Although the spread of 5 or more person households between the clusters in minimal, they are underrepresented in the clusters that lower the temperature during the day. The difference of household sizes between the different formed clusters indicate that the thermostat settings in dwellings is influenced by household size.

Movement users between clusters
In this section, the movement of users between different clusters for separate weekdays is discussed. In the clustering analysis, every weekday is evaluated individually to account for the difference between weekdays, discussed in section 3.1. This means that users can move form cluster to cluster on separate weekdays. The movement of the users is shown in figure 3.9. In figure 3.9 the different clusters are on the y-axis and the

Table 3.10: Thermostat practice clusters

Index	*Thermostat practice*
A	Day lowering
B	Morning evening program
C	Adaptable program
D	Day heating
E	Gradient day heating
F	Continues heating

Table 3.11: Variance of daily gas consumption between cluster (ANOVA)

Clusters	*F value*	*p value*
All	50.9	$1.1 * 10^{-51}$
A,B,C	33.6	$4.1 * 10^{-15}$
D,E,F	2.4	0.088

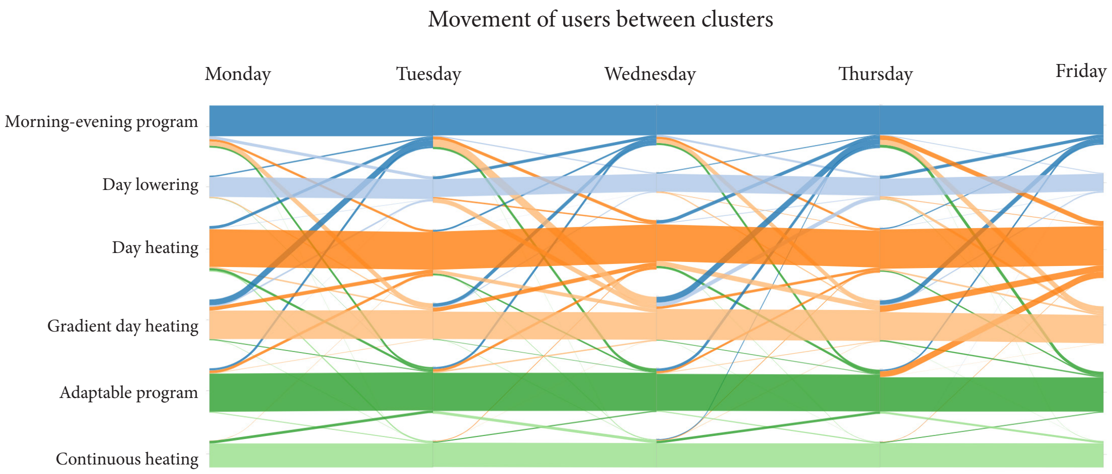


Figure 3.9: Cluster: Cluster movement: Sankey chart

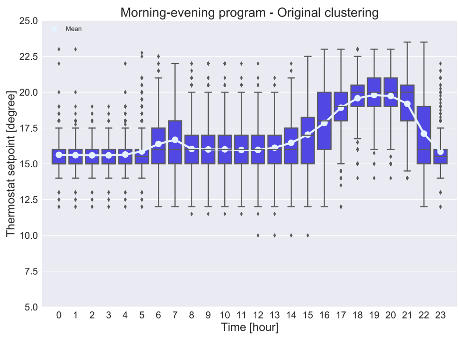
weekdays are shown on the x-axis, the thickness of the line indicates the number of households.
The majority of the users stay within their cluster between different weekdays meaning that the thermostats setting is relatively similar during weekdays. Which either means that households habits and in turn their thermostat practices do not largely change between weekdays or that the same thermostat setting is kept between weekdays independent of habit. The largest movement between clusters can be found between the Morning-evening program and gradient day heating clusters. Indicating that households tend to change their thermostat settings on Tuesdays and Thursday to a program with lowered temperatures during the day. While on Wednesdays and Friday the thermostat settings are changed to Gradient day heating and Day heating, a program with increased temperatures during the day. The movement of users between clusters on Tuesday and Thursday accounts for 5% of the population while on Wednesday and Friday around 6%. The habit of people being at home at Wednesdays and Fridays are in line with expectations of daddy days and an early weekend [68].
55 % of the working population in the Netherlands has the freedom to adapt their working time schedule [13], this effect in not seen in figure 3.9. Minor irregular adjustments to an individual's working schedule are lost in the use of overall thermostat settings. But the freedom to adjust working time schedules and the tendency to work less of from home on Wednesdays and Fridays would suggest a larger movement between clusters. Indicating that the lack of moment of users between weekdays is probably caused by a regular non-adapted thermostat practices.

3.4. Clustering validation

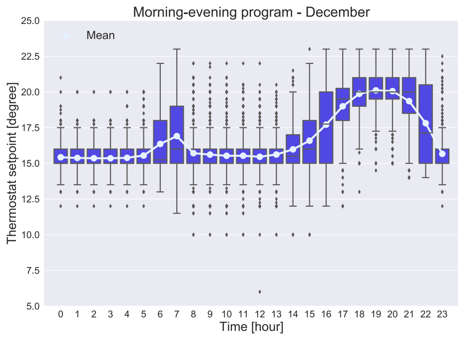
In this section, the validation of the clustering method performed in this thesis is discussed. The clustering method is validated in 2 ways: visual validation and user weekday robustness. The visual validation evaluates the clustering method by comparing the clustering output of spitted data set. The robustness of the clustering is determined by reclassifying the user weekdays post clustering to determine the fitness of the user weekdays within the cluster. The visual validation of the clustering is firstly discussed followed up by the use weekday cluster fitting. This section is ended with discussing the overall validity of the clustering method used.

Visual validation

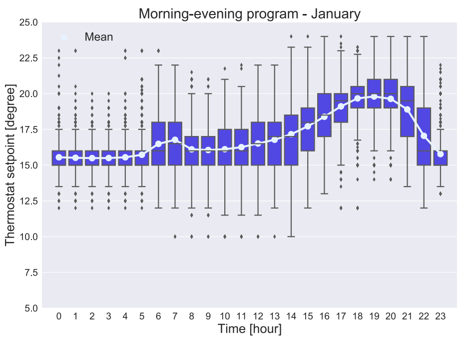
The visual validation is based upon dividing the original data set into 3 separate data sets for comparison. The data is separated by month to evaluate the ability of the clustering technique to distinguish resembling thermostat practices on the basis of a splitted data set. For each of the separate data sets the same clustering method is used as in the original clustering as discussed in section 3.2. In appendix D the resulting cluster outcomes for each of the separate months are displayed. The clusters are grouped by cluster resemblance in line with the original clustering. Overall the majority of the monthly clustering resembles the original clustering, the visual compatibility between resembling clusters is shown in appendix D.



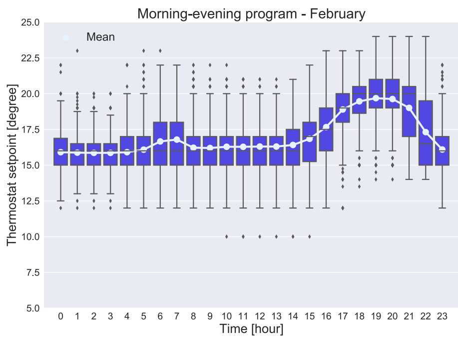
(a) Day lowering - original clustering



(b) Day lowering - December clustering



(c) Day lowering - January clustering



(d) Day lowering - February clustering

Figure 3.10: Visual validation of clustering based upon separated data sets

To give an example, the morning-evening program thermostat practice cluster and the resembling clusters for each of the 3 months are shown in figure 3.10. Figure 3.10 displays the thermostat settings in the same manner as in section 3.3. Figure 3.10 shows that the original clustering and resembling clusters are visually compatible to each other. There are some differences, the midspread of the thermostat settings in December figure 3.10b is lower compared to the other months and original clustering in figure 3.10a. Furthermore, user weekdays with an increased thermostat setting in afternoon are more present in the January clustering in figure 3.10c. Overall the cluster based upon the different months and original clustering show visual matching similarities. The resembling clusters are similar in the moments when the thermostat is turned on and off with a temperature difference of a couple of degrees.
The difference between the clustering outcomes for separated months can be explained by an interaction between hierarchical clustering and alternated starting conditions. As explained in section 3.2.1 hierarchical clustering treats every data case as an independent cluster and clusters user weekdays on the basis of least distance. Therefore, a change in the starting condition, dividing the data set into separate months, changes the formation of the first clusters. After the initial clustering, the hierarchical clustering continues combining user weekday clusters with least distance between them. The change in original starting condition changes the composition of the first formed clusters and in turn the entire clustering process.

User weekday cluster fitting

The robustness of user weekdays in the original clustering is assessed by analysing the difference between user weekdays thermostat settings and the average cluster thermostat setting. The difference in user weekday thermostat settings and the average thermostat settings of the formed clusters as discussed in section 3.3 is used to reclassify the user weekdays. The difference between user weekday thermostat settings and the average cluster thermostat settings is calculated with equation 3.4. For each of the user weekdays the difference between each of the 6 average cluster settings is calculated with equation 3.4. The clusters are reclassified

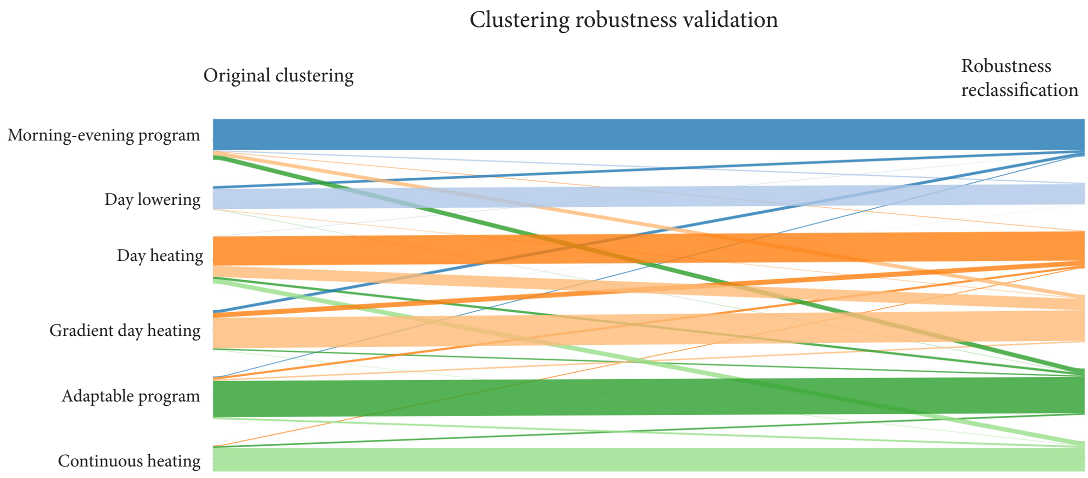


Figure 3.11: Movement of user weekdays after reclassification

into the cluster with the least difference between the average thermostat setting and user weekday settings.

$$D = \sum_{h=1}^{24} (T_{user\,weekday,h} - T_{cluster,h}) \tag{3.4}$$

D as the difference between the user weekday thermostat settings and average cluster thermostat settings, $T_{user\,weekday,h}$ the user weekday thermostat setting at hour h, $T_{cluster,h}$ the average cluster thermostat setting at hour h.

The average thermostat setting of each of the clusters is compared to the individual user weekdays. The user weekdays are reclassified on the basis of least difference with the resulting thermostat practice clusters. Table 3.12 shows the percentage of user weekdays that stay within the same cluster after reclassification on the basis of least difference. On average 77% of the users stay within the same cluster after reclassification. Thermostat practice cluster Day heating and Gradient day heating indicate the biggest difference after cluster reclassification. To determine the movement of user weekdays between original clustering and after reclassification figure 3.11 is used. Figure 3.11 shows the movement between the different clusters after reclassification. In line with table 3.12 the biggest movement after reclassification is in the Day heating and Gradient day heating cluster. Figure 3.11 indicates that the movement of user weekdays of clusters Day heating and Gradient day heating is mainly between each other. There is also movement between the other clusters but relatively low compared to the Day heating and Gradient day heating movement.

The movement between the Day heating and Gradient day heating cluster after reclassification can be explained by looking into the clustering results in section 3.3. Figure 3.6d and figure 3.6e are the most similar clusters in the overall clustering outcome shown in figure 3.6. Whereby a user weekday thermostat settings with settings between the average Day heating and Gradient day heating cluster can be reclassified into the other. The other aspect that explains the movement of user weekdays is the difference in distance calculation in the Ward method compared to the difference calculation used in reclassification. The Ward similarity

Table 3.12: Rigidity of user weekdays after reclassification

Cluster	*Robustness user weekdays within cluster [%]*
Day lowering	83%
Morning evening program	78%
Adaptable program	83%
Day heating	67%
Gradient day heating	73%
Continues heating	88%

method takes more aspect into account than the difference calculation used in reclassification. In Ward hierarchal clustering the entire user weekday thermostat settings within the cluster are taken into account instead of the average hourly thermostat settings.

Overall validation discussion
The overall conclusion that can be drawn about the validity of the clustering form the discussed validation techniques are discussed in this section. The visual validation of separating the clustering data set in to different months indicated the dominance of certain thermostat practices. Similar looking thermostats practices are detected for different months. Whereby a presence of distinct thermostat practices in the winter months of Dutch households is substantiated.
The movement of users between separate days of the week shown in figure 3.9 showed to be relatively modest. The modest movement indicates that households tend to remain in the same cluster for separate weekdays, based upon individual data sets. Whereby can be determined that the clustering repeatedly clusters similar thermostat settings of the equivalent household in the same cluster.
The reclassification of user weekdays on the basis of overall difference between user weekday and cluster thermostat settings generates insights about the rigidity of the clustering method used. Overall 77 % of the user weekdays remain in the same cluster after reclassification. The range in thermostat settings within the clusters shown in figure 3.6 indicates the spread thermostat settings present in the cluster. The largest moment of user between clusters after reclassification is found in clusters that are most similar. Indicating that there are user weekdays thermostat settings that have thermostat settings that can be penned to both clusters.
The visual similarity between the resembling clusters and the original clustering indicate that the detected thermostat practices are dominant within the entire population. The discrepancy in rigidity and similarity on the other hand indicated that the clustering should be seen as classification and not a complete determination of an entire household thermostat practice. Whereby the thermostat practices of individual households cannot completely be fitted within a box.

3.5. Conclusion: sub-research question 2

In this section, the second sub-research question is answered and stated below. The second sub-research question relates to the grouping of thermostat practices and the insights that can be gained from grouping thermostat settings with the use of clustering.

> *To what extent can clustering be used in the identification of specific thermostat practice of Dutch households?*

6 distinct thermostat practices are discovered by the clustering of thermostat settings of individual households for separate weekdays, with the use of hierarchical clustering based upon least variance similarity measure. Clustering is an approach to group similar thermostat practices of individual households. The resulting clusters share similarities in the moments when the thermostat settings are changed. Despite the similarities the thermostat practices in the clustering display clear difference in thermostat practice between clusters.
The grouping of thermostat settings can be used to gain preliminary insights in the natural gas consumption of dwellings. There is an apprehend division in relative high and low natural gas consumption between dwellings related to their thermostat practice. Most of the households do not adapt their specific thermostat practices to separate weekdays. Similar household sizes share commonalities in their thermostat practices, especially one-person household. One-person households are more likely to lower their thermostat during the daytime.
Clustering of overall thermostat practices result in the detection of distinct thermostat practices. The specific thermostat practices generate insight in possibility and potential to reduce energy consumption in the residential sector, in the form of potential inefficient thermostat practice of a sizeable part of the sample group.

4 Habit engine

Individual household thermostat practices are an influential factor in the gas consumption of dwellings. With the clustering of thermostat settings similarities in household thermostat practices can be detected. The majority of residents do not interact with their smart thermostat, 57% do not interact with their thermostat at all [36]. The thermostat settings are not adapted to individual thermal needs and could mismatch individual thermal needs. To determine the fitness of thermostat practices with the thermal needs of resident an additional measure is needed. Potential heat savings in the residential sector can consecutively be analysed by assessing the fitness of the thermostat settings and thermal needs. The exact thermal needs of residents differ from person to person. In order to assess individual thermal needs of residents an extensive study is needed into individual perception of thermal comfort. One element can be detected without studying the thermal needs of residents, whether residents are at home or not. When residents are not at home their thermal needs are not met by the heating of a dwelling. Therefore, residential thermostat practices are assessed in combination with the determination of at home occupancy.
In this chapter, the use of residential electricity consumption to determine at home occupancy is discussed, by firstly discussing how electricity consumption is used to determine at home occupancy with the use of the Habit engine. The Habit engine is a python script developed by Quby to provide insight in at home occupancy. The effectiveness of the Habit engine is discussed by analysing the results of a pilot conducted with the use of the Habit engine in section 4.2. In section 4.3 the use of the Habit engine within this thesis is discussed followed up by the detected at home occupancy results and critically reviewing the habit engine. This chapter is ended with answering the third sub-question in section 4.4

4.1. Habit engine

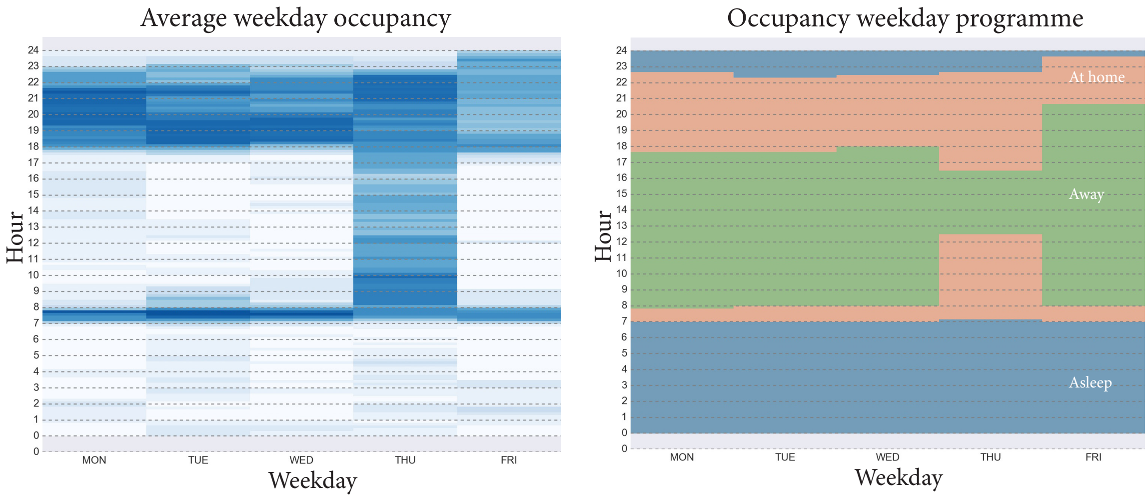
To gain insight in the extent of overlap between thermostat practices and at home occupancy the Habit engine is used. Residential at home occupancy can be detected by analysing household electricity consumption remotely. Through the use of non-intrusive load monitoring (NILM) individual appliances can be detected by looking into specific electricity consumption characteristics of appliances [5]. With the use of NILM, residential occupancy or individual appliance can be detected in dwellings. The use of NILM ensures the determination of at home occupancy without interfering in a dwelling, by looking at the residential electricity consumption at a distance.
As discussed in chapter 1 the capabilities smart metering devices are evolving. Currently it is possible to gather detailed residential energy consumption data remotely, for example with the use of Toon. The relationship between electricity and at home occupancy is quite straight forward. When people are at home they make use of more appliance then when they are away, for example lightning and watching television [55]. Whereby residential electricity is dependent on the occupancy of a dwelling. By the use of smart electricity meters, the occupancy of occupants in the sense of if someone is at home can be achieved to an accuracy around 90% [40].

4.1.1. Habit engine working

The habit engine is used to determine at home occupancy on the basis of individual residential energy consumption data. In this section, the way the Habit engine determines at home occupancy on the basis of elec-

tricity consumption data is discussed. The Habit engine generates an overall weekly occupancy overview. Average electricity consumption data in 10-minute intervals is used to determine at home occupancy. The Habit engine creates an average weekly occupancy overview on the basis of at least 8 weeks of data. The weekly occupancy pattern generated by the habit engine distinguishes between moments when residents are at home, asleep or away.
Residential at home occupancy is detected by the determination of periods of relative high electricity consumption. The habit engine makes use of two methods to detect moments of relative high electricity consumption. These 2 methods need to be able to filter standby energy consumption out of intermittent electricity consumption.
The working of the Habit engine is discussed by giving an overview of the steps taken by the model. Both models determine at home occupancy by following the same steps, listed below.

- 10-minute interval consumption data is classified into active (1) and non-active (0) periods.
- 10-minute interval classification is aggregated for identical weekdays. The data is reclassified by taking the mean occupancy of identical weekdays at 10-minute intervals.
- Periods of consecutive occupancy blocks are identified and grouped together into occupancy periods. To filter out minor disturbance in occupancy and smooth the output.
- The at home classification is assigned to the consecutive occupancy blocks. An inactive period between 2:30 and 3:30 as asleep, an active period as being at home and the other inactive periods as being away.



(a) Average occupancy boolean program

(b) Resulting occupancy program

Figure 4.1: Habit engine program detection

Figure 4.1 shows the average detected weekday occupancy in figure 4.1a and the resulting occupancy program in figure 4.1b. Every 10-minute electricity consumption is classified by the habit engine as active or non-active. In figure 4.1a the occupancy Boolean is averaged out for equivalent 10 interval electricity consumption at the same moment for separate weeks. The darker blue the 10-minute interval block is, the higher the average occupancy for that occupancy block of identical weekdays is. A majority voting system is used to classify the average 10-minute interval occupancy blocks into of active and inactive. When the majority of the 10-minute interval blocks are active for identical weekdays that block is considered as being active. The reclassified 10-minute interval blocks are grouped together into blocks on consecutive occupancy shown in figure 4.1b. A consecutive occupancy block between 2:30 and 3:30 is reclassified as being asleep.
The Habit engine utilizes a Rolling Mean Model (RMM) combined with a Hidden Markov Model (HMM) to detect household occupancy. A HMM can classify a predefined number of states through an assumed probability distribution of hidden states within the system. The model assemble determines the weekly occupancy pattern on the basis of electricity consumption data. The models are used independently of each other, to determine at home occupancy on the basis of the best fitting model for individual households. The selection procedure between the RMM and the HMM is based upon a quality measure, discussed further on in section

4.1.2.
The difference between the use of the RMM and the HMM in the habit engine is the initial classification of the 10-minute interval consumption data. The RMM model takes the mean electricity consumption and defines a period as active when it is above the lower 10% quantile plus the stand derivation of the 10% quantile. The HMM is used to describe a probability distribution over a number of sequences [17]. For this model, the HMM assumes that every 10-minute interval can be in one of 3 different states from a Gaussian Markov process. The top 2 states are classified as active and the other as inactive. The HHM makes use of 3 states instead of the two inactive/ active states to incorporate the ramp up and ramp down. The third state is needed in order to classify the morning and evening as being active.

4.1.2. Quality measure

A quality measurement is used to determine the best fitting program schedule produced by each of the classification methods. The quality measure produces an indication of the overall quality of the generated schedule as well. The quality measure is based upon 7 separate quality checks, to test the Habit engine output on a set of assumptions. The quality measure is used to determine the resemblance of the generated at home occupancy schedule with a general occupancy pattern. The different assumptions used in the quality check of the Habit engine are listed below. The quality measurements are based upon the detection of an occupancy schedule that is in line with a regular schedule. The set of assumptions has been developed in line with the Habit engine creation.

1. The number of occupancy changes between (at home, away, asleep) per day, with a best fitting number changes of 4.
2. The best fitting occupancy rate, which is assumed at 40%.
3. Whether occupants are active before 12 o'clock, is considered best fitting.
4. The number of unique habits per day, 3 is considered correct (sleep, away, at home).
5. The number of hours not at home per day, 8 is considered best fitting.
6. The similarity of habits between different weekdays.
7. The difference between the median occupancy, the higher the contrast the larger the support of the found pattern

The independent quality assumptions are combined in the overall quality assessment. Each of the quality measurements is normalized to an optimum of one and summed up. Therefore, the optimum quality measuring is the sum of its parts equalling 7. In order to assure a minimum quality of the Habit engine output and quality threshold is used. The minimum quality threshold has been detected by visual conformation in line with the Habit engine development and is considered to be 5. To analyse the overall ability of the Habit engine to produce an occupancy schedule that resembles actual residential occupancy a pilot has been conducted. The minimum quality threshold of 5 has been used within the conducted pilot.

4.2. Habit engine pilot

The use of the Habit engine has been tested in a pilot by Eneco, executed by the behavioural research company Totta. In this section, the results of the pilot are discussed to give an indication of the validity of the Habit engine and discuss known issues. The pilot provides insights in the thermostat practice of residents and discusses relevant feedback on the Habit engine. In the beginning of this section the set-up of the pilot and the relevant information provided to the respondents is discussed. The section is ended with discussing the pilot results that are relevant in thermostat practice assessment.

4.2.1. Habit engine pilot overview

The goal of the Habit engine pilot is to assess the potential of giving household personal advice on their thermostat use. With the use of detected at home occupancy and thermostat setting assessment. A group of 550 households where asked to join a pilot concerning giving personal advice on individual thermostat use. In the end 164 of the 550 completed the entire Habit engine pilot.

The advice given to individual households consist of two aspects. Firstly, giving general advice about thermostat settings and suggesting an adopted thermostat program. The general thermostats advice was based upon the temperature settings within Toon for the 4 modes. The 4 modes consist of the temperature settings for being asleep, being at home, being away and comfort modus. The general advice compared these settings to generally assumed optimal settings for these modes provided Milieu Centraal [47].
The second advice given to the respondents consist of a thermostat program based upon the detected occuapncy with the use of the Habit engine and the thermostat modes settings. The thermostat program advice consists the detected occupancy schedule and individual residential thermostat settings. The suggested temperature settings in the program advice are based upon the current modes settings to adjust the advice to individual thermal needs.

4.2.2. Relevant pilot results and representability

In this section, some of the relevant or remarkable outcomes of the pilot are discussed. First of all, the overlap between the general population and the social-demographic factors of the pilot response group. There are two factors known of the pilot group that can give an indication of the representability. The first representability indicator is the gender of the respondents, of the households who responded 97% was male. The majority of the respondents represent more than one person households, based upon general innovators population group shown in table 3.2. Why the predominant gender of the response group is male in relation to interest in in thermostat programs or inner family social construct is a question for a different research.
The other factor in determining the representability of the response group that could be assess is the geographic distribution. The geographical distribution of the respondents is ensured, 67% of the respondent live in different city with an overlap in larger city like Amsterdam and Rotterdam. On the gender of the respondents and the geographical distribution of the dwellings alone, the representability of the respondent group cannot be assessed and is therefore unknown.
The reaction of the respondents to the pilot was overall positive. The respondents indicated that the pilot made them think about their thermostat practice and provided insights in possible improvements. There is an overall overlap between the detected thermostat practice and actual thermostat practice of the residents. Only 24% of the respondent indicated that there was no overlap between their habits and the proposed schedule. The main reason given by the respondents was an irregular living pattern. For example, respondents that work in day and night shifts and respondents that are self-employed did not see overlap. The difference in schedule form week to week and the inconsistency in work load of the self-employed causes mismatch between the overall detected and actual occupancy. Some of the respondents indicated that their program for the weekend does not match due to irregular weekend activities. This feedback is in line with the choice made in the clustering of overall thermostat settings to focus on the weekdays. In line with the findings by Guerreiro et al. [29] respondents indicated that most of them do not interact with their thermostat regularly.

4.3. Habit engine use

In order to assess the potential heat saving in the residential sector the thermostat practices of individual households are analysed. The Habit engine and thermostat settings are combined to determine the overlap between thermostat settings and at home occupancy of residents. By combining thermostat practices with at home occupancy the potential residential saving of thermostat adjustments in the residential sector can be determined.
The Habit engine is used to determine at home occupancy of residents for the 3-month winter period in line with the clustering analysis. The electricity consumption data of all the residents within the clustering analyses is used to determine residential occupancy. The average electricity consumption collected by Toon for 10-minue intervals is used, to determine residential occupancy within the 3 winter months.
Two parts of the Habit engine are used in the assessment of residual occupancy within this thesis. The quality measure that calculates the overlap of the habit engine occupancy schedule with a presumed optimum is used ensure the quality of the detected occupancy. The initial 10-minute interval electricity consumption classification is used in the evaluation of at home occupancy of thermostat practices.
In the assessment of residential occupancy, the 10-minute intervals are combined into one hour intervals in line with the thermostat settings, discussed in section 3.1. The hourly at home occupancy is based upon a majority voting system of the average activity within that hour. When four or more of the six 10-minute interval activity blocks are detected as at home the entire hour is classified as being at home.
The quality of the detected at home occupancy is assured by the use of the quality threshold. If the Habit

engine is able to produce and overall thermostat program that scored above 5 on the quality measurement the detected occupancy is taken into account. When the habit engine is able to generate an overall weekly occupancy pattern on the basis of 3 months of data that meets the quality threshold. The method used to classify at home occupancy on the basis of electricity consumption data is considered valid. Whereby the 10-minute interval classification for the separate days are used in the at home occupancy assessment.

4.3.1. At home occupancy results

In this section, the insights the habit engine provides in combination with the clustering of thermostat settings is discussed. Followed up by critically reviewing the capabilities of the habit engine itself. The Habit engine is used to provide additional insight in the thermostat practices of individual dwellings, through the detection at home occupancy. Figure 4.2 gives an overview of the average occupancy over the day for each of the thermostat practice clusters. The average occupancy within the cluster is composed of taking the mean of the occupancy Boolean, (1) being active and (0) as being inactive for each user weekday within the cluster. Therefore, the y-axis displays the percentage of user weekdays that are detected as being active at that time within the relative cluster.

There are 2 main finding that can be distinguished from figure 4.2, a clear habitual pattern and difference between the clusters. The clustering analysis distinguished a clear habitual pattern of an increased thermostat setting in the morning and decreased setting at night. With an increased thermostat setting during the evening hours. The average at home occupancy shown in figure 4.2 shows a similar pattern, for a substantial part of the day. During the night, the occupancy is around 0 while in the evening the occupancy is around 1. Indicating that the thermostat is turned down when residents are inactive at night and turned on in the evening when they are at home. The slope in the morning and evening is caused by the hourly difference in when residents are active or inactive. Suggesting that the majority of the residents wake up around 7-8 o'clock, with the reverse of this pattern happening in the evening around 11.

The second main finding is the difference between the different clusters. For most of the evening and night their detected occupancy is similar but not during the day. The working thermostat schedule of the day lowering and morning evening program is in line with the detected occupancy. The average occupancy of the

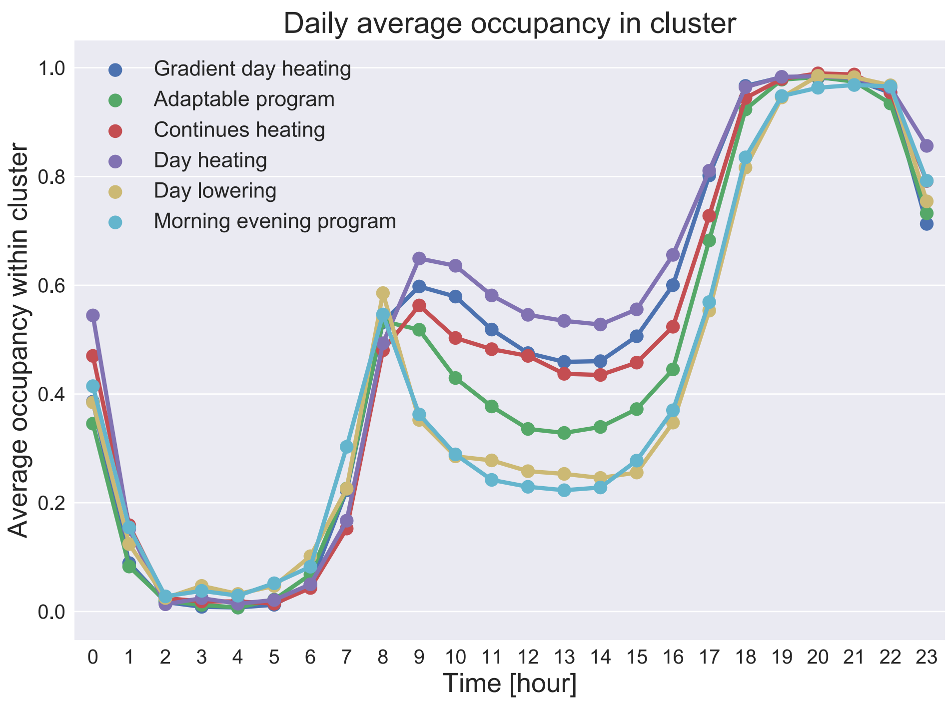


Figure 4.2: Overall Habit engine outcomes for different thermostat practices

day lowering and morning evening clusters during the day is around 20%. Indicating that the majority of the households within this cluster are not at home during the day. The clusters consisting of users who heat their home during the day tend to be active for an extended hour in the morning. The contrast between the users in the clusters who heat their home during the day and who do not is detected the daily occupancy. The average detected occupancy of the users within the heating clusters is around 50%. Indicating that these users heat their home during the day but are not all actively at home.
The relative high inactiveness of the users within the cluster that do heat their homes the entire day indicates a saving potential. When on average only half of the users who heat their home during the day are actually detected as being at home. While viewing these results the limitations of the Habit engine have to be kept in mind. The detected occupancy is not absolute certain therefore these findings have to be seen as an indication. However due to the large number of observations and user's, overall conclusions can be drawn.

4.3.2. Habit engine capability

The potential short comings of the habit engine are discussed in this section. The conducted pilot of the habit engine indicated that for a majority of the households the produced occupancy schedule is in line with their occupancy habits. While for the households with an irregular occupancy schedule habit engine was not able to detect a representative schedule. The conducted pilot provides insights in the overall detected occupancy pattern and occupancy habits of households. The pilot does not provide insight in the correctness of the detected at home activity in 10-minute intervals.
There are 2 mechanisms at play that influence the capability of the habit engine to detect at home activity correctly. Firstly, the electricity consumption of households is not completely representative for their occupancy habits. For example, when households turn on their dishwasher before they go to bed, their electricity consumption is higher while being asleep. Henceforth relative high electricity consumption might be disqualified as detected occupancy while being asleep. The second mechanism is the quality indicator, which check the habit engine output to a presumed occupancy pattern. The presumed occupancy pattern is in line with a working schedule. When a households' occupancy habits are not in line with the presumed occupancy pattern the habit engine generates a lower quality measure. The detected occupancy pattern might be qualified as insufficient while the detected occupancy pattern is correct. The quality measure potentially effects the at home occupancy for the different thermostat practices. Because the quality measure is based upon a working schedule the occupancy rate of the heating clusters might be overestimated. Households that heat their home during the day and are at home might be excluded due to an insufficient quality measure. Overall the habit engine is able to provide insight in the occupancy of households but the exact precision of the detected occupancy is unknown.

4.4. Conclusion: sub-research question 3

The third sub-question relates the detection of at home occupancy with the use of residential electricity consumption. To evaluate overlap between residential thermal needs and thermostat practices with the following question:

> *To what extent can disaggregated electricity consumption data be used to evaluate thermostat practices through occupancy detection?*

Disaggregated electricity consumption data can be used to determine whether residents are at home or not, by detecting periods of relative high electricity consumption out of average 10-minute interval electricity consumption data. The pilot has shown that for 76% of the households the Habit engine is able to generate a weekly occupancy program that is in line with their occupancy habits. The successfulness of an overall detected at home occupancy is dependent on the regularity of residential agendas.
Residential thermostat programs of households are in line with at home occupancy during the night and evenings. Where by the thermostat is lowered when residents are away. During the day, the thermostat programs are not in line with the at home occupancy of residents. The majority of the households who do not heat their home during daytime tend away. This is not the case for households who heat their dwelling at daytime. Around 50% of the residents are actually detected as being at home of the household who heat their home during daytime.
The mismatch between thermostat practises and at home occupancy of residents indicate potential gas saving in the residential sector. Disaggregated electricity consumption data can be used to determine the occupancy in dwellings. The fitness of thermostat practices and thermal need can be analysed, to identify

potential savings and assess the extent of possible thermostat practice adjustments in the residential sector. The precision of the habit engine is currently unknown and can be assess by evaluating the use of the habit engine in line with at home monitoring sensor. Physical movement sensors, sound and geo-fencing could be used to determine the precision of the habit engine.

5

Thermostat practice adjustments

The residential sector is responsible for over 55% of the natural gas consumption in the Netherlands with the primary use of space heating. The heating consumption in the residential sector can be reduced by adjusting thermostat settings of individual households. In order to assess the potential impact of individual thermostat adjustments on residential gas consumption, potential savings firstly have to be identified. Followed up by estimating the impact of adapted residential thermostat settings on the potential savings. The savings results can in turn used to compare residential energy saving on the basis of disaggregated consumption data to aggregated data. Disaggregated and aggregated saving studies can be compared to provide insight in the added value of using disaggregated consumption data in residential energy saving assessment. This chapter begins with describing the data used, potential saving options and calculation methods. Followed up with a discussion of potential saving results and disaggregated/ aggregated study comparison. This chapter is ended with the answering of the fourth sub-research question in section 5.4.

5.1. Potential thermostat adjustments

In the beginning of this section the used data is discussed followed up by the identification of potential saving areas and calculation methods. In order to reduce gas consumption of individual dwellings potential savings need to be identified in line with the thermal comfort of residents, as discussed in section 2.3. To change residential thermostat practices, heating adjustments that safeguard thermal comfort of residents need to be identified. Thermal comfort need to be safeguarded in order for residents to accept the adjustments and to ensure engagement with the suggested thermostat adjustments. Therefore, possible thermostat savings need to be identified that do not impact the thermal comfort of residents or do so in a limited way.
The data used to determine potential saving areas and used in the additional analysis is firstly discussed in section 5.1.1. Potential savings that are in line with thermal comfort requirement are discussed in section 5.1.2. Followed up by section 5.2 with a description of the saving calculation methods.

5.1.1. Data structure for adjustment calculations

Multiple data sources are needed to assess the impact of adjusting thermostat settings of individual dwellings. The different data sources can be divided into 3 parts: temperatures, gas and electricity consumption. In line with the clustering analysis, the 3 coldest winter months of 2016-2017 are used in the adjustment calculations, to assess the impact of adjusted thermostat settings when residential thermal consumption is at its peak. The same households used for the clustering analysis are used to determine potential savings.
The temperatures account for the thermostat settings, inside temperature and outside temperature. For the thermostat settings and inside temperature the entire 3-month hourly weekday settings/temperatures data set is used. Unlike the clustering analysis in which the overall thermostat weekday settings are used. The inside temperature is measured by the smart thermostat Toon, for the same time span and frequency as the thermostat settings. The outside temperature is based upon measurements by Dutch weather institute KNMI [41]. The average outside temperature in line with the calculation methods, measured by KNMI in de Bilt is used. De Bilt is chosen as the outside reference point because the location of the individual dwellings is unknown and de Bilt is roughly located in the centre of the Netherlands.
The gas consumption used for heating a dwelling is determined by gas flow metering and boiler communi-

cation. With the use of the open-therm protocol and smart meter /thermostat Toon, gas consumption for heating water can be determined. Consecutively the moments when the boiler is used to heat water for non-heating purposes can be filtered out. The moments when the boiler is operational but not instructed by Toon to heat water for thermal control are filtered out. The possible overlap in gas consumption for space heating and other hot water activities, like showering at the same moment are not taken into account. Subsequently the gas consumption for heating a dwelling per day is used to determine the impact of individual thermostat adjustments.
As discussed in chapter 4, households' electricity consumption can be used to determine at home occupancy. The detection of at home occupancy of residents is used to fit thermostat adjustments to the thermal needs of residents. In the sense that heating a dwelling when its residents are not active or not at home their thermal needs are different. To determine the residential occupancy, the average electricity use of a 10-minute interval is used in combination with the Habit engine. To ensure the quality of the detected at home occupancy of the habit engine, only the households that exceed the quality threshold described in chpater 5.2 are used. Of the 1319 households used in the clustering, 936 meet the quality threshold and are henceforth used in heating adjustment estimation. In line with the thermostat data the hourly median occupancy Boolean is used in determination of at home occupancy. In the sense of, if residents are active or not inside the dwelling in one hour intervals.

5.1.2. Potential saving options

In this section, several adaptions to thermostat setting that limit the overall impact of thermal comfort of residents are discussed. Adapting thermostat practice to individual thermal needs of residents generates the possibility to save energy, by heating a dwelling more efficiently [3]. Through thermostat use there are two ways to lower residential heating consumption, by lowering the temperature or decrease the heating time span. The clustering of thermostat settings and detection of at home occupancy, indicated two areas of potential saving. Firstly, the clustering of residential thermostat settings discussed in section 3.3 indicates that the majority of the residents heat their home during the day and that there is a potential to save energy by lowering the temperature during the night. Secondly the combination of thermostat settings with at home occupancy through analysing electricity consumption, showed that thermostat programs are not all adapted to at home occupancy. Households who heat their home during the day are not detected as being at home, on average around 50% are present. Consequently, there is an opportunity to reduce residential heat consumption by adapting the thermostat settings to at home occupancy. The saving areas have been separated in the following 4 saving options: Active, Absent, Before sleep and Overnight. The 4 saving options are discussed in detail in the following subsection.
The saving options are in line with suggested saving options by van den Ham and van der Vliet [65]. Whereby it is possible to compare residential saving estimations on the basis of disaggregated data to aggregated data use. To ensure the possibility to assess the overall saving potential, the different saving options do not overlap.
In the following sub-sections the 4 different saving options are discussed, by describing what the options entail and what method is used to calculate the potential savings. As described in the previous section there are two ways to reduce residential heating consumption, by lowering the temperature settings and decreasing number of heating hours. The saving options are presented by discussing these to parameters for every option. The minimum thermostat set point of 15 degrees is used as a base line temperature. Lowering the temperature below 12 - 15 degrees could lead to humidity problems within the dwelling [65]. The temperature of 15 degrees is chosen to limit the overall impact on thermal comfort and mitigate possible humidity problems.

Active lowering

The kind and type of activity in which a person is engaged determines the thermal need of a person, for example when people are active their thermal need is lower [50]. Therefore, the temperature in a dwelling could be lowered when residents are more active. Residents tend to be more active during the day than evening, therefore the temperature can be reduced during the day while maintaining thermal comfort. Hanmer et al. [31] indicated that several residents already lower their temperature during the day while being at home. The level of thermal comfort at a certain temperature is strongly dependent on the individual, and differs from person to person [49]. Therefore, the suggested temperature by Milieu Centraal [47] of 19 degrees is used during the day to ensure thermal comfort.
In this saving option, the thermostat is lowered to 19 degrees when residents are at home detected with the

habit engine. When residents are active between 7 o'clock in the morning and 6 in the evening residents the temperature is lowered. The temperature is lowered for every hour residents are at home in this time span to 19 degrees.

Absent lowering

As stated in section 1.2 the majority of the people do not interact with their thermostat at all. In combination with the mismatch between detected at home occupancy and thermostat settings, indicate that dwellings are heated when residents are not at home. The detected occupancy trough the habit engine is used to determine residential occupancy. Where by the thermostat settings are lowered when residents are not detected as being at home.
For every hour that the habit engine does not detect occupancy while residents are not asleep the thermostat is lowered to the baseline temperature of 15 degrees. Absent lowering assesses the impact of lowering thermostat for when people are not at home.

Before sleep lowering

The level of insulation of a dwelling relates to the build period of that dwelling, with developing level of insulation year by year [64]. The decreased heat radiation of dwellings though insulation makes it possible to maintain a certain temperature while not actively heating a dwelling [72]. The insulation of a dwelling enables the possibility to lower the thermostat before people go to sleep, because the temperature decreases gradually in an insulated dwelling. Therefore, it is possible that residents remain thermally comfortably while a dwelling is not actively heated. The drop in temperature of a dwelling is dependent on both the thermal mass and insulation of the dwelling.
In the before sleep lowering saving option the thermostat is lowered one hour before residents go to bed to 15 degrees. Resulting in the turn down of the thermostat one hour before residents go to sleep. The sleeping moment of residents is detected with the use of the habit engine. The hour before the sleep block discussed in the next subsection is used as the before sleep lowering moment.

Overnight lowering

The clustering indicated that 3 of the 6 clusters have a thermostat setting above the baseline temperature of 15 degrees during the night. Residents are able to maintain thermal comfortable with blankets while they are asleep. In addition, the living room of residents does not need to be heated when people are asleep in a different room. Therefore, the thermostat can be lowered when residents are asleep at night.
The occupancy detected by the habit engine is used to determine if residents are asleep or not. The time span between 10 in the evening and 4 o'clock at night is considered as period that residents could be asleep. If residents went to bed in between this time span, the consecutive inactivity period during the night is considered as being asleep. If the consecutive inactivity period extends 3 hours and is less than 12 hours residents are considered asleep. When the consecutive inactivity period is longer than 12 hours the residents are considered as not being at home. The timespan between 10 and 4 is used to detect a sleeping time block. Therefore, inactivity outside this time span is considered as a sleeping time block when the consecutive inactive period overlaps with the sleeping time span. The overnight lowering option adjusts the thermostat to 15 degrees during the detected sleeping time block.

5.2. Saving calculation methods

In this section, the methods used to calculate potential savings by adjusting the thermostat settings are discussed. The saving options are defined as an adjustment to the actually used thermostat settings. Because the impact of the adjusted thermostat settings on residential gas consumption is not measured an estimation has to be made. Multiple methods for estimating the gas reduction are used, to compare methods and mitigate potential over or underestimations by different estimation methods. The following 3 methods will be used to determine the saving potential of thermostat adjustments: relative shift, heat demand reduction and temperature and gas regression.
All of the saving calculation methods translate changes in thermostat settings into potential gas saving. By firstly calculating the effect of changed thermostat settings on the inside temperature of a dwelling, followed up by one of the 3 saving calculation methods. The manner in which thermostat changes effect the indoor temperature is firstly discussed. Followed up by discussing the 3 different saving calculation methods.

5.2.1. Thermal inertia

Building characteristics are one of the underlying factors that influence the gas consumption of individual dwellings discussed in chapter 1. One of the building characteristics that influences the gas consumption is thermal inertia. To account for the difference in heat loss of individual dwellings the specific heat loss is taken into account. The specific heat loss of dwellings is expressed by the time constant, shown in equation 5.2 is used. Equation 5.1 and 5.2 are used to determine the specific heat loss of a dwelling and the change in inside temperature by not actively heating a dwelling [65]. The overall change in inside temperature by adjusting the thermostat settings is used to determine the potential saving of the different options.

$$T_{t_v} = T_e + (T_0 - T_e)e^{-\frac{t_v}{\tau}} \tag{5.1}$$

$$\tau = \frac{-t_v}{ln(1 - \frac{T_0 - T_{t_v}}{T_0 - T_e})} \tag{5.2}$$

With T_{t_v} as the inside temperature after t_v number of hours, T_e as average outside temperature, T_0 inside temperature before the heating interruption and τ as the time constant.
For each of the individual dwellings τ is calculated for each night in which the thermostat has been turned off for more than 4 hours. To calculate the time constant τ the data described in 5.1.1 is used. The outside temperature is based upon KNMI measurements and for the thermostat settings and inside temperature the smart thermostat Toon is used. For every dwelling τ is determined for a 20-day period within the 3-month wither period. The average τ is calculated for each dwelling and based upon the number moments the thermostat has been turned off for more than 4 hours within 20-day period.
The calculated τ can in turn be used to calculate the inside temperature for every hour within the dwelling of adjusted thermostat settings. Equation 5.1 describes the heat loss of an individual dwellings after a number of hours. In the following saving options, the inside temperature based upon the time constant and thermostat adjustments are used. The adjusted inside temperature is calculated by translating the change in thermostat setting into change of inside temperature by the use of equation 5.1.

5.2.2. Relative shift method

The relative shift method is based upon the gas needed to heat a dwelling compared to the outside temperature. Dwellings lose heat through two ways: by conduction through walls, floors, ceilings and windows and through the moving air in and out of the building [12]. When the inside temperature is higher than the outside temperature a dwelling loses heat through airflow and conduction. Dwellings can gain heat by the heating system, internal gains and solar gains. The difference between the in and outside temperature can be used to make an estimation of the gas consumption of a dwelling [56], by estimating the gas consumption per degree difference between the in and outside temperature. This method is a rough estimation because the solar and internal gains are not taken into account in this estimation.
The relative shift method uses the daily gas consumption per degree difference between the in and outside temperature during the entire day, shown in equation 5.3. The relative shift method calculates the saving by multiplying this factor with the relative change inside temperature shown in equation 5.4. The resulting relative shift calculation determines potential saving through the change in the in and outside temperature. Thermal inertia is taken into account in the relative change calculations, the estimated change in inside temperature is calculated with the use of equation 5.1.

$$T_d = \frac{\sum_{h=1}^{24}(T_{in,h} - T_{out,h})}{24} \tag{5.3}$$

With T_d as the daily difference between the average in T_{in} and outside T_{out} temperature.

$$S_h = \frac{\frac{V_g}{T_d} * \left(\frac{\sum_{h=1}^{24}(T_{in,0,h} - T_{out,h})}{24} - \frac{\sum_{h=1}^{24}(T_{in,1,h} - T_{out,h})}{24}\right)}{V_g}$$

Simplified to: (5.4)

$$S_h = \frac{\frac{\sum_{h=1}^{24}(T_{in,0,h} - T_{in,1,h})}{24}}{T_{d,0}}$$

With S_h as the relative heat saving, V_g daily gas consumption, $T_{in,1}$ the changed inside temperature, $T_{in,0}$ temperature before the adjustment.

5.2.3. Heat demand reduction method

The heat demand reduction method calculates the relative residential heat saving by calculating the inside temperature drop induced by not actively heating a dwelling. The potential saving of non-active heating periods in dwellings is described in detail by van den Ham and van der Vliet [65]. The relative saving of non-active heating periods can be calculated by equation 5.5 [65]. Equation 5.5 is based upon the simplification of dividing the reduction in heat demand of a non-heating period by the average heat consumption of that period without non-heating periods. Two correcting factors have been crossed out against each other in setup of this formula, due to the similarity in effect in opposite directions. Firstly the factor that corrects for the heat loss coefficient during the non-active heating periods. Secondly the correcting term for the difference in outside temperature during the non-active heating period.

$$S_h = \frac{t_v}{24} * (1 - \frac{n_0}{n_1}\frac{\tau}{t_v}(1 - e^{-\frac{t_v}{\tau}})) * 100\% \tag{5.5}$$

With S_h as the relative heat saving, t_v non-active heating hours, n_0/n_1 the efficiency of the heating system, τ time constant.

The efficiency of the heating system is the heat output of the system divided by the heat input. van den Ham and van der Vliet [65] have calculated the efficiency of different heating systems by model calculations and multiple field tests performed by AgenschapNL. The efficiency of the heating system ranges from .997 to 1.012 depending on the age and type of heating system. Because the type of heating system in unknown for users within the sample group the efficiency is considered to be 1.

The non-active heating hours result from the saving option calculations. In the case that the temperature drops below the base line temperature the dwelling needs to be heated in order to maintain the baseline temperature. In that case the non-active heating hours t_v are calculated by the use of equation 5.6.

$$T_v = -\tau \ln(1 - \frac{T_0 - T_{base}}{T_0 - T_out}) \tag{5.6}$$

T_v non-active heating hours, T_0 inside temperature at the beginning of the non-active heating period, T_{base} base line temperature, T_{out} the outside temperature.

5.2.4. Temperature and gas regression method

The last calculation method uses a linear regression model for every individual household to calculate the impact of thermostat adjustments. The regression model predicts the gas consumption on the bases of the difference between the in and outside temperature. In this subsection, the use of the regression model is firstly explained followed up by discussing the assumptions underlined in linear regression models. The subsection is ended with describing predictive power of the regression model and the relative gas saving calculation method

Linear regression model

Linear regression is used to describe the relationship between the daily gas consumption and average difference between the in and outside temperature. Through the use of a linear regression model the impact of temperature settings can be related to the gas use of dwellings [20]. The linear regression model fits the observed data to the following equation 5.7.

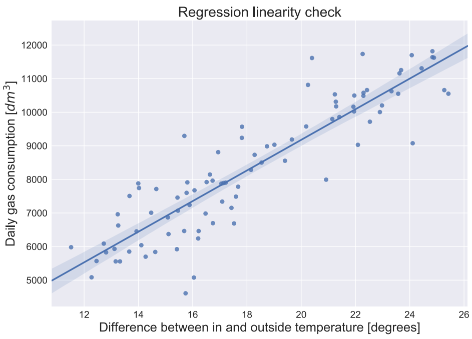
$$Y = B_0 + B_1 X + \epsilon \tag{5.7}$$

With Y as the daily gas consumption of individual dwellings, B_1 as the average difference between the daily in and outside temperature, X the predictor variable that expresses the change in daily gas consumption for one B_1 degree difference, ϵ as the error estimation to accounting for the discrepancy between predicted data and observed data.

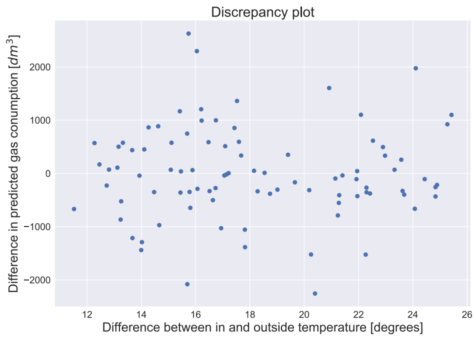
For each of the dwellings the daily gas consumption data and in and outside temperature described in section 5.1.1 are used to build an individual regression model for each dwelling. Every individual regression model is based upon all the 70 weekdays within the 3-month winter period of 2016-2017.

Linear regression models assumptions

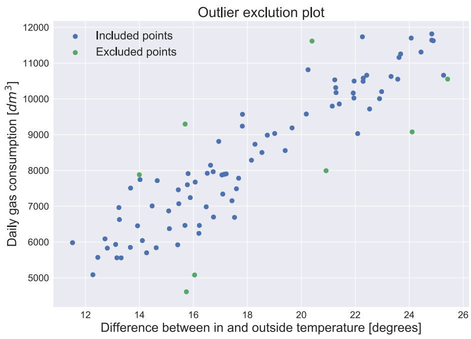
At the basis of linear regression models there are several assumptions, these assumptions need to be checked before linear regression models can be used. The following three assumptions are taken into accounted and have been analysed:



(a) Linarity of gas consumption and inside temperature



(b) Difference between the predicted gas consumption and actual gas consumption



(c) Visual representation of the outlier exclusion

Figure 5.1: Linear regression assumptions analysis

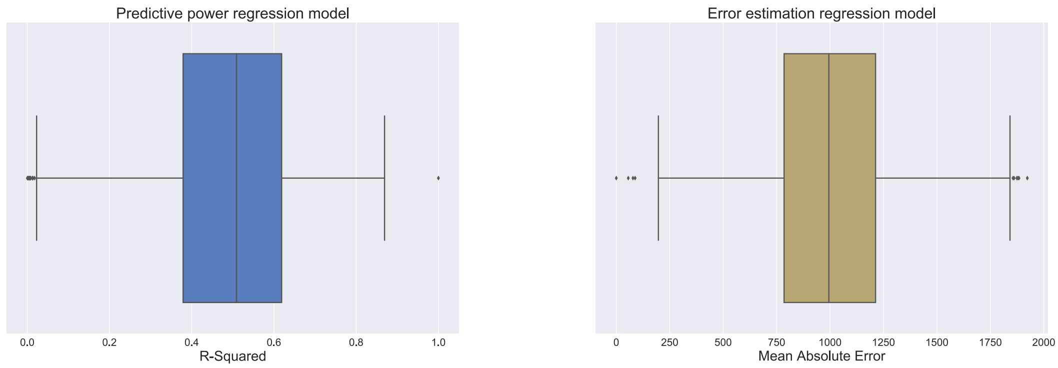
1. *Assumption of linearity*
2. *Assumption of constant variance in residuals, homoscedasticity*
3. *Checking and removing of outliers*

Each of these assumptions is analysed with the help of plotting gas consumption and difference between the inside and outside temperature shown in figure 5.1. The plots in figure 5.1 are based upon the linear regression model of a single users within a 3-month winter period. Figure 5.1a clearly shows a linear relationship between the difference in gas consumption and the daily difference in indoor and outdoor temperature. Whereby the linearity assumption of the regression analysis is met. The difference between the in and outside temperature is henceforth called the daily temperature differences.

The difference between the predicted gas consumption of individual dwellings and the measured gas consumption is shown in figure 5.1b. For each of the weekdays within the winter period the daily measured gas consumption is subtracted from the predicted daily gas consumption. The predicted daily gas consumption is calculated by substituting the daily temperature difference B_1, in the individual linear regression models described by equation 5.7. The measure to what extent the spread of the residual is similar at different points analysed with the homoscedasticity. The spread in residue should be homogeneous at different daily temperature differences. Figure 5.1b indicates that spread within the residual is quite evenly spread for the different daily temperature differences. The amount of observations for each of the daily temperature differences is not. The uneven spread can be explained by the variance in outside temperature. The uneven spread of outside temperature causes the daily difference to be unevenly spread as well. The uneven spread of observations can be neglected because the homoscedasticity is assured. The error in the regression model is homogeneous at different daily temperature differences despite the observation spread.

Outliers can have a dis-appropriate effect on the formation of linear regression models. Therefore, the Bonferroni correction is used to exclude outliers that fall outside the 95% confidence interval in the data set [6]. Figure 5.1c gives an example of the exclusion of data points within a sample set. The Bonferroni correction is used independently for every user within the dataset. Each of the 3 linear regression conditions are met, therefore the linear regression is used to determine the potential gas saving.

Individual linear regression indication



(a) Predictive power regression model

(b) Regression model error: Mean Absolute Error

Figure 5.2: Linear regression quality indicators

In this subsection, the overall quality of individual regression models is discussed with the use of figure 5.2. Figure 5.2 gives an overview of the predictive power of the individual regression models in figure 5.2a and the educed error in figure 5.2b. The boxplots in figure 5.2 give show the midspread of the individual linear regression models within the box and the lines 95% spread within the lines.
To what extent the regression model is able to predict the gas consumption by the difference between the in and outside temperature is shown in figure 5.2a. The R-Squared gives an indication of what percentage of the gas consumption can be predicted by the regression model. For 75% of the users, the linear regression model is able to predict more than 36% of the gas consumption.
Figure 5.2b gives an indication of the error of the predictions made with the regression model. The boxplot shows the Mean Absolute Error (MAE) of the gas prediction by the regression model. In other word the average of the absolute difference between every gas measurement and predicted consumption by the regression model for each household. The figure indicates that for over 50% of the households the regression model results in an MAE between 750 and 1250 dm^3
Overall all the regression model quality checks show a large spread between the linear regression models of different users. For most household, the individual linear regression models are able to explain a large part of the daily gas consumption. The extent of predicted gas consumption indicates that there are more factors that influence the gas consumption of dwelling than the difference between the out and inside temperature alone in line with expectations. The linear regression models able to explain large part of the gas consumption with an error that accounts for 15% of the daily gas consumption on average.

The difference between the individual regression modes is shown in figure 5.3. In figure 5.3 the linear regression model of 20 different households is plotted, each of the lines in figure 5.3 represents a household. For the 20 households in figure 5.3 the linear relationship between the daily natural gas consumption and the difference between the in and outside temperature is shown. The variation in slope of each of the linear regression model indicates the difference between the individual dwellings. The discrepancy in slope indicates the influence of other factors that influence the daily gas consumption for heating in a dwelling. For example, the thermal inertia of a dwelling influences the amount of natural gas needed to keep a dwelling at a certain temperature.

Regression based saving calculation method
To calculate the relative saving of each potential saving option the individual regression models are used. The change in inside temperature based upon the thermostat adjustments is calculated by the use equation 5.1. The daily difference between the in and outside temperature in the old and adjusted situation is entered as B_1 in the regression model. Whereby the daily gas consumption based upon the old and adjusted situation are compared to assess the relative gas saving, shown in equation 5.8. The regression based relative saving

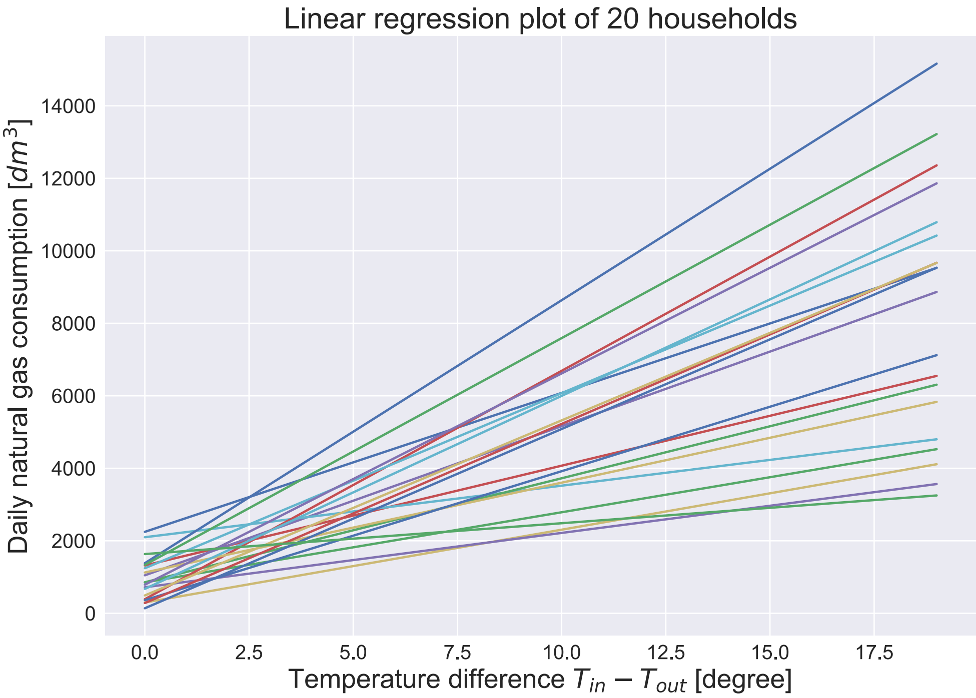


Figure 5.3: Linear regression models of different households

calculation method calculates the relative gas saving by the use of individual regression models. Both the old gas consumption and the adjusted gas consumption are calculated by the use of the regression model. The actual gas consumption is not used in the determination of the relative gas consumption due to the impact of potential mismatch between the regression model and actual gas consumption. The regression model is seen as the description of residential gas consumption based upon the difference between the in and outside temperature.

$$S_h = \frac{B_0 + (T_{d,0})B_1 + \epsilon - (B_0 + (T_{d,1})B_1 + \epsilon)}{B_0 + (T_{d,0})B_1 + \epsilon}$$

Simplified to: (5.8)

$$S_h = \frac{(T_{d,0} - T_{d,1})B_1}{B_0 + (T_{d,0})B_1 + \epsilon}$$

5.3. Potential thermostat adjustments results

In this section, the potential saving for each of the saving adjustments calculated with the use of the different calculation methods is discussed. By presenting the results and discussing the difference between the method outcomes. This section is ended with comparing the saving results of disaggregated data use to use of aggregated consumption data.

5.3.1. Relative saving calculation method results

Following from section 5.1.2, 4 different potential thermostat adjustments have been identified. Each of the potential thermostat adjustments ensures a reduction in the heat consumption of a dwelling while maintaining the thermal comfort of its residents. To recap the identified thermostat adjustment are listed below.

1. *Active lowering*, reducing the temperature during the day in a dwelling to 19 degrees.
2. *Absent lowering*, reducing the temperature in a dwelling when residents are not at home to 15 degrees.
3. *Before sleep lowering*, turning down the thermostat in a dwelling one hour before residents go to sleep

to 15 degrees.

4. *Overnight lowering*, when resident are a sleep the thermostat is turned down to 15 degrees.

Table 5.1: Overview potential saving per calculation method

	Calculation methods		
Saving options	*Relative shift*	*Heat demand*	*Regression*
Active lowering	0,89%	0,97%	0,92%
Absent lowering	1,15%	1,00%	1,18%
Before sleep lowering	0,06%	0,04%	0,06%
Overnight lowering	1,92%	2,27%	1,99%

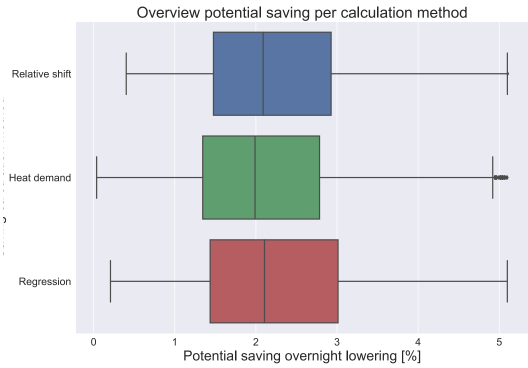


Figure 5.4: Potential saving difference calculation method: Overnight lowering

Table 5.1 shows the potential saving for each of the saving options calculated with the 3 different calculation methods. Table 5.1 displays the average relative gas savings for the 936 households within a 3-month winter period. The difference in the potential saving estimation by the different methods shown in table 5.1 is less than .35% percentage point. When comparing the different calculation methods, the heat demand calculation method underestimates and overestimates the potential saving for 2 of the 4 options.
The largest difference in percentage point potential saving is found for the overnight lowering saving option, when comparing the different calculation methods. Figure 5.4 is used to compare the different calculation methods on the basis of saving options that indicates the biggest difference between the calculation methods. Figure 5.4 gives an overview of the potential saving of the overnight lowering option for each of the calculation methods. The Boxplots in figure 5.4 provides an overview of the calculated potential saving for each day within the 3-mont winter period. The box in figure 5.4 contains the midspread and the lines contain 95% of the potential saving spread. Figure 5.4 indicates that there is a large overlap in the average saving potential of the different calculation methods. Besides the overlap between the different saving calculation methods the range is saving potential for 50% of the population is below 2 percentage points. The difference in potential saving is expected due to the difference between individual households both in physical and behavioural aspects expressed in the potential saving of overnight lowering.
The differences between the different calculation methods are analysed by comparing the calculated saving potential of each method based upon non-heating hours. Each of the saving calculation methods is adapted to determine the potential saving of non-heating hours. Household specifics are used to determine potential saving of each household on the basis of non-heating hours. The potential saving is adjusted to individual households by taking into account the specific thermal inertia and individual regression models. The resulting potential saving of non-heating hours is shown in figure 5.5. Figure 5.5 clearly shows the difference between the relative shift and regression estimations compared to the heat demand calculation method. The difference between the calculation methods advances over time. Figure 5.5 indicates that the heating demand and relative shift calculation methods overlap to a greater extent in the estimation of potential saving than the regression method. The variance in saving potential spread between figure 5.5 and figure 5.4 is caused by the difference in data set use. Figure 5.5 calculates the potential saving on the basis of non-heating hours while figure 5.4 indicates the average potential saving of individual households adjusted to daily household behaviour.
While the differences between the calculation methods is visible over time the estimated potential savings overlap. Each of the boxes in figure 5.5 describing 50% of the households evidently overlap. The boxplots shown in figure 5.5 are based upon the individual characteristics of dwellings. Concerning the specific heat loss of a dwelling and the gas consumption per daily temperature difference. Henceforth the spread in relative saving indicates that the individual dwelling characteristics are an important factor in the estimation of potential savings. The individual difference between the dwellings cause the spread in the potential saving per non-active heating hour. As expected the potential saving of a households is dwelling specific due to the impact of dwelling characteristics on the thermal inertia of a dwelling.
In the following sections the different calculation methods are combined to give an overall estimation of the

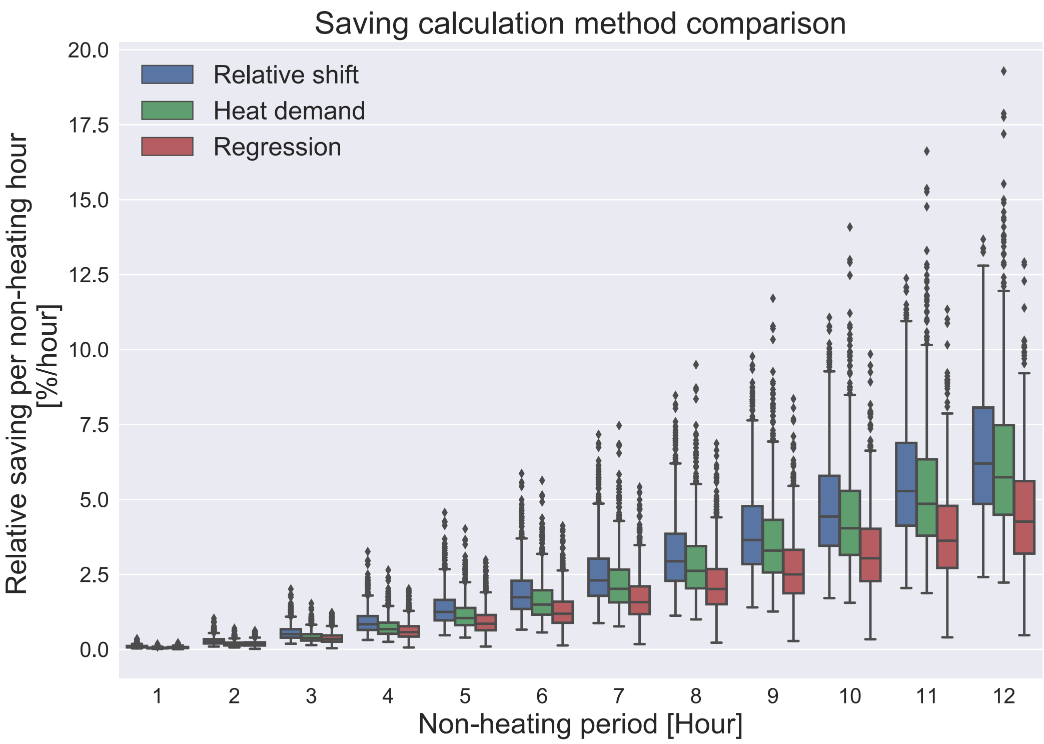


Figure 5.5: Hourly saving per calculation method

potential saving. The different calculation methods are combined by taking the average of the potential saving methods. The overlap in potential saving calculation shown in figures 5.5, 5.4 shows that there is not a clear separation between the calculation methods. The different calculation methods are combined to mitigate the potential over or under estimation by one of the calculation methods and enable potential savings compression in a convenient matter.

5.3.2. Influencing factors in relative saving potential

In this section, different factors that influence potential gas saving in the residential sector are discussed. The potential saving of different thermostat practices are described and compared to the estimated savings in a study performed by van den Ham and van der Vliet [65].

Table 5.2: Potential saving of different thermostat practices

	Potential savings per cluster				
Clusters	*Absent lowering*	*Active lowering*	*Before sleep lowering*	*Overnight lowering*	*Total*
Day lowering	0.49%	0.27%	0.03%	1.20%	1.99%
Morning evening program	1.10%	0.35%	0.05%	1.51%	3.01%
Slight adaptable program	1.63%	0.47%	0.06%	2.10%	4.26%
Day heating	1.20%	1.23%	0.06%	2.19%	4.68%
Gradient day heating	1.06%	0.79%	0.04%	1.60%	3.49%
Continuous heating	1.39%	1.03%	0.07%	2.48%	4.97%

Potential saving per thermostat practice

As discussed in the previous section the characteristics of individual dwellings influences the overall gas saving potential of that dwelling. Thermostat settings play a role in the potential saving of a household, next to specific dwelling characteristics. In figure 5.2 the potential saving per thermostat practice for each of the

different saving options is shown. Figure 5.2 indicates clear differences in saving potential for different thermostat practices. With the clear difference in saving potential between the thermostat settings extremes of Day lowering to Continues heating thermostat practices. The difference in saving potential between these practices is more than a factor 2.
Table 5.2 shows that even with a relatively reduced thermostat setting there is still a potential to save energy, when looking at the Day lowering thermostat practice. In general, the potential saving per thermostat practice is arranged from low too high potential in table 5.2. The gradient increase in saving potential is in line with the overall level of thermostat settings within the clusters show in figure 3.6. With the exception of the Gradient day heating cluster. The lower saving potential in absent, active and overnight lowering can be explained by two factors. Firstly, the overlap between residential at home occupancy and thermostat settings. When the thermostat settings are adapted to the at home occupancy of residents, the potential saving of the active and absent saving option is lower. The adaption of thermostat settings to at home occupancy is in line with the movement of users between clusters discussed in section 3.3.2. The users within the Gradient day heating cluster tend to adjust their thermostat settings to different weekdays. Secondly the overnight thermostat settings of the Gradient day heating cluster are around 15 degrees, demonstrating the lower saving potential of the overnight saving option.

Potential saving based upon build period
To provide insight in the additional value of using disaggregated energy consumption data in the evaluation of residential gas saving potential. The results of this study are compared to calculations performed by van den Ham and van der Vliet [65] in a potential saving study. To assure the compatibility between the two studies the different saving option classification is similar. The study conducted by van den Ham and van der Vliet [65] is based upon the classification of dwellings types on the basis of household characteristics. The potential savings are then based upon assumptions that are calculated for different dwelling type classification. The following assumptions are used in the estimation of saving potential:

1. *Active*, the thermostat is turned down by one degree.

2. *Absent*, the thermostat is turned down for 8 hours per day

3. *Before sleep lowering*, the thermostat is lowered from 20 to 15 degrees for 1 hour

4. *Overnight*, for a sleeping period of 8 hours the thermostat is lowered from 20 to 15 degrees.

Table 5.3: Potential saving: Study comparison

	Absent lowering		**Active lowering**		**Before sleep lowering**		**Overnight lowering**	
Build period	*Thermostat adjustments*	*Milieu Centraal*	*Thermostat adjustments*	*Milieu Centraal*	*Thermostat adjustments*	*Milieu Centraal*	*Thermostat adjustments*	*Milieu Centraal*
Before 1967	2.7%	11.1%	0.1%	6.0%	1.1%	1.4%	1.8%	7.3%
1976-1988	2.0%	10.3%	0.1%	6.1%	0.8%	1.0%	1.2%	4.3%
1989-2000	1.7%	7.3%	0.1%	6.0%	0.7%	0.8%	1.1%	2.3%
2000-until now	1.5%	4.5%	0.0%	5.9%	0.8%	0.2%	0.9%	1.4%

The household characteristics of build period is used to compare the results of the different studies shown in figure 5.3. The results of the calculations performed by van den Ham and van der Vliet [65] is shown in the Milieu Centraal tab and results of this study in the thermostat adjustment tab. The Milieu Centraal potential saving is based upon theoretical gas consumption with the use of a combination of dwelling type and build period. The overall potential saving for the different build periods is shown in table 5.3.
There are two clear differences between the studies: the amount of potential saving and the difference in saving for the different build periods. The difference in potential saving can be explained by the assumptions made in the Milieu Centraal based calculations. The estimations based upon actual thermostat use indicate an overall lower saving potential. Indicating that the assumptions made in the Milieu Centraal based calculations are not in line with the thermostat use of the Toon innovator population group.
The second discrepancy relates to the difference in saving potential for different build periods. Both studies show a decreased saving potential for newer dwellings. In line with the increased level of insulation in newer dwellings. The difference between the different build periods is larger in the Milieu Centraal estimations. This effect could be explained by the difference in actual household heat consumption and theoretical heat

consumption on the basis of household characterises. The gas consumption of poorly insulated dwellings tends to be overestimated compared to an underestimation of well insulated dwellings [44]. The variance in complex household characteristics underpin the discrepancies between theoretical and actual heat consumption [45].
The impact of the assumptions made in the milieu centraal report is further analysed by assessing the impact of these assumptions. The 3 main differences between the assumptions made in the milieu centraal report and the Toon users are listed below.

1. *Estimated* τ, is lower on average than the measured τ for the Toon users, shown in table 5.4
2. *Thermostat settings*, is assumed at 20 degrees instead of the diversity in settings of the Toon users shown in section 3.3.
3. *Absent heating period timespan*, the average absent period detected by the Habit engine for the Toon users is 5.5 hours instead of the 8 hours assumed in the Milieu Centraal based calculations.

Table 5.4 shows the estimated τ used in the milieu central report and the measured average τ, for different build periods. The impact of a different τ is shown in table 5.4 by calculating the potential saving of an 8-hour non-heating period. Table 5.4 expresses the impact of a different τ on the calculated potential saving. The potential saving is calculated by the use of the heat demand reduction method formula 5.5.
There is a clear difference in the potential saving for different τ shown in table 5.4. The lower the τ the faster a dwelling losses heat and in turn the higher the potential saving is. The effect of declining difference between the estimated τ and calculated τ in table 5.4 can be detected in table 5.3 as well. A possible explanation for the difference in measured τ and assumed τ on the basis of dwelling characteristics is residential feedback. Residents in a dwelling with a relative low τ might have improved the insulation of their dwelling or adapted their heating practices.
The difference between the potential saving calculation between figure 5.5 and figure 5.4 is caused by the calculation technique used. In figure 5.5 the potential saving is adjusted the thermal inertia of a dwelling. When a base line temperature is reached due to the cooling down of a dwelling the dwelling is re-heated to ensure a constant temperature. The potential saving shown in figure 5.4 is not adjusted for the cool down rates of dwellings which causes the potential saving to be higher than in figure 5.5. An estimated inside temperature of 20 degrees overestimates the potential saving of absent heating periods in two ways. Firstly, the timespan of an absent heating period is impacted when the inside temperature is lower, because the base line temperature is reached more easily. Secondly, the potential of lowering a thermostat declines when the inside thermostat settings are already lowered. The relationship between the number of absent heating hours and potential saving is quite straight forward. A decreased number of absent heating hours reduces potential natural gas saving.

Table 5.4: Potential residential saving depending on τ and build period

	Time constant [τ]		**Potential saving non-heating period**	
Build period	*Millieu Centraal*	*Thermostat adjustments*	*Millieu Centraal*	*Thermostat adjustments*
Before 1967	14	40	7.8%	3.1%
1976-1988	28	50	4.1%	2.5%
1989-2000	49	57	2.2%	2.2%
2000-until now	80	71	1.2%	1.8%

The potential gas savings for different thermostat practices shown in table 5.2 display a bigger impact on the potential savings than the differences in building period figure 5.3. The larger impact of thermostat practices on the gas saving potential compared to building period indicates two things. Firstly, that different thermostat practices are influential in the potential heat saving in the residential sector. Secondly, the individual differences between household thermostat interaction need to be taken into account. To assess the space heating saving potential of the residential sector thermostat interaction needs to be taken into account, in line with complex underpinnings of dwellings specific heat consumption.
In appendix E the potential saving based upon several household and dwelling characteristics is shown. The potential saving for different household sizes, dwelling types and dwelling sizes is shown in appendix E, some of the remarkable findings are discussed in this section. The potential saving of absent lowering is seen to decrease for larger households in figure E.1. Probably caused by an increased at home occupancy for larger

household sizes and in turn a lower potential to lower the thermostat when not at home. The potential savings of overnight lowering for apartments and terraced dwellings is lower than semi- and detached dwellings shown in figure E.2. The larger a dwelling is the lower the potential saving of overnight lowering is shown in figure E.3. The overall difference between saving potentials for household with the previous mentioned dwelling characteristics is lower than the build period.

5.3.3. Overall potential of thermostat adjustments

An indication of the overall saving potential of the different saving options is shown in figure 5.5. Table 5.5 is based upon a gas price of 63 euro cent per m^3 and a gas consumption of 1300 m^3/year [11]. Lowering the thermostat during the night clearly results in the overall largest potential residential heat saving. Followed up by not heating a dwelling when away and lowering the thermostat during the day. The yearly saving of turning-off the thermostat before residents go to sleep does not generate any substantial savings. Overall the average Toon user could save 33,99 euro on their energy bill by adapting their thermostat practices.
In table 5.6 the yearly potential savings for each of the thermostat practices shown. Table 5.6 indicates that there is a 24 euro difference between the potential saving of residents with an overall lowered thermostat program compared to a heating one. The potential saving difference between households with different thermostat practices accounts for over 70% of the average savings. The thermostat practices of households play a major role in the overall individual saving potential in the residential sector.

Table 5.5: Potential residential heat savings per year: Saving options

Saving option	*Potential annual saving*
Active lowering	€7,59
Inactive lowering	€9,09
Before sleep lowering	€0,44
Bedtime lowering	€16,87
Total saving	€33,99

Table 5.6: Potential residential heat savings per year: Thermostat practice

Thermostat practice	*Potential annual saving*
Day lowering	€16,33
Morning evening program	€24,61
Adaptable program	€34,89
Day heating	€38,36
Gradient day heating	€28,58
Continues heating	€40,70

5.4. Conclusion sub-research question 4

The fourth sub-research question relates to the saving potential of thermostat practice adjustments of individual dwellings. The potential saving for the residential sector is assessed in the following way:

> *What is the potential impact of adjusting thermostat practices of individual households on residential natural gas consumption?*

Adjusting thermostat settings of individual households on average leads to a saving of 4.2% resulting in €34 energy bill reduction per year. The potential saving is dependent on the individual thermostat practices and dwelling characteristics. The difference in thermal inertia of dwelling has a large impact on the overall potential saving for dwelling with similar thermostat practices.
The individual thermostat practices have shown to have a larger effect on the potential residential saving than classic classification characteristics as building period. Whereby the individual thermostat practices need to be taken into account in the overall residential energy assessment, in combination with the dwelling specific heat inertia.
The disparate residential potential heat saving of individual dwellings indicates the importance of adjusting energy assessments accordingly. Both thermostat practices as dwelling specific heat loss has shown to have a substantial influence. The current assessment of dwelling specific characteristic has shown to mismatch actual thermal inertia of a dwelling. Future residential energy predications should include thermostat practices and actual thermal inertia assessment, due to the impact of these factors. Therefore, energy consumption in the residential sector should include individual dis-aggregated consumption data.

6
Discussion, conclusion and context

The performed research is discussed by debating research method and assumptions made in section 6.1. After which the research questions are answered in section 6.2. This chapter is ended with placing the conducted research within this thesis in context.

6.1. Discussion

The performed research method and underlined assumptions are discussed in the beginning of this section by discussing 3 aspects. This section is ended with discussing the relevance of the research in line with the Complex System Engineering and Management master program. The following 3 aspects have been taken into account:

1. *Sample group*, the data is gathered from households with a smart meter/ thermostat device that are willing to share their data.

2. *Thermal comfort*, the extent of thermal comfort of residents with the thermostat adjustments.

3. *Non-intrusive practice detection*, individual practices are based upon the measurable action of residents with the use of thermostat and electricity data.

Sample group

The analysed households within this research are part of the innovators group within the Quby, creator of Toon. Therefore, each of the analysed households have a smart metering device installed in their dwelling. The default thermostat settings follow a standardized program. The smart meter/ thermostat Toon enables the possibility to adjust a thermostat program in multiple ways, through the use of an IHD and a mobile application. Besides the smart metering device within the researched dwellings the analysed users are part of an innovator group. Meaning that the residents have actively chosen to share their consumption data for research purposes.

The presence of a smart metering device inside a dwelling and the affinity of households to share their data for research purposes might underestimate the potential savings. Easier access to thermostat control and affinity with research might induce an increased general knowledge/ interest in of heat saving. As discussed in section 3.1.1 the demographic characteristics of the sample group are not in line with the general population. The influence of different demographic characteristics is uncertain due to the substantial impact of different characteristics as thermostat practice and thermal inertia. Whereby household demographics do not represent a meaningful way to assess the general population in the sense of individual practices and dissimilar dwelling characteristics.

The number of household used in the assessment of thermostat practices of individual dwellings is relative small. The conducted research can be seen as proof on concept of thermostat practice assessment with the use of disaggregated consumption data. The developed approach and data evaluation scripts can be used directly on the entire Toon user group of over 300.000 households.

Thermal comfort
Thermal comfort differs from person to person and is next to the temperature dependent on aspects as ventilation, radiation and metabolism. Whereby the thermal comfort of a person is different in altering position in a room. Within this research, the thermal comfort has been taken into account by temperature alone and without adjusting for separate rooms. The thermals comfort has been taken into account by an average comfortable inside temperature for when residents are at home of 19 degrees.
The individual differences in thermal comfort of people might lead to discomfort with a certain daytime temperature. By taking an average suitable thermal comfort temperature, the adjusted temperature settings to individual needs cancels each other out. In the assessment of general saving potential. The effect of other factors besides the inside temperature in the assessment of thermal comfort are not taken into account and therefore unknown. Current average ventilating behaviour is embedded in the assessment of dwellings by individual assessment of households. Whereby the overall ventilation behaviour is encapsulated in the heat inertia of dwellings.

Non-intrusive practice detection
Residential consumption practices are governed by 4 key elements that together shape a practice. In this research, the outcome of the interplay of the key elements in practice theory is analysed, in the form of thermostat practices of individual households.
The non-intrusive manner of analysing individual practices increases the potential reachable population while limiting the possibility to analysing underlying factors. The key elements are essential in the assimilation of practices. Therefore, the link between the actual adjustments of thermostat practices and potential saving cannot be made, without addressing or analysing the underlying key elements that make up a practice. The currently used at home detection with the use of electricity consumption data with the use of the habit engine is not completely validated. The habit engine has proven to be able to determine adequate week programs but the real time at home detection has not been validated. Whereby the exact at home detection capabilities of the habit engine are uncertain.

Vestigial COSEM link
The COSEM master degree focuses on designing in socio-technical systems. Whereby the role of sociological systems is combined with technical systems. Both elements are studied to be able to design an artefact that can intervene in the complex intertwined socio-technical world. In this research, the heating system is studied by looking into the interaction of social behavioural aspects with the technical system of a smart meter.

6.2. Conclusion

The conclusion and recommendations are discussed in this section, by firstly restating the problem question and its sub-research questions. Followed up by the concluding remarks of answering the research question, ended with the recommendations. The research question is discussed by giving a short recap of the problem and restating the research question. The primal form of energy consumed in the residential sector is natural gas with the main use of heating. Gas mining induced earthquakes in combination with goals set out in the climate accord intensifies the pressure on the Dutch government to reduce the overall natural gas consumption. Adjusting thermostat practices of Dutch households is an approach to reduce gas consumption on the relative short term. Disaggregated energy consumption data is seen as high frequency individual household consumption data. There are several uncertainties in the assessment of natural gas reducing possibilities. To gain insights in the possibilities of thermostat practices adoption the following questions are answered:

1. What is practice theory and how can it be used to understand heating practices of Dutch households with the use of disaggregated consumption data?

2. To what extent can clustering be used in the identification of specific thermostat practice of Dutch households?

3. To what extent can disaggregated electricity consumption data be used to evaluate thermostat practices through occupancy detection?

4. What is the potential impact of adjusting thermostat practices of individual households on residential natural gas consumption?

To answer the following research question:

What insights in thermostat practices that influence natural gas consumption of individual households can be identified by a combined analysis of electricity and thermostat use?

6.2.1. Answering sub-research question

Each of the sub-research question is answered in this section to answer the main research question. The sub-questions are answered in the previous four chapters, a recap of the answers is stated below.

Sub-research question 1: Practice theory
Residential behaviour is the main source of uncertainty in predicting energy consumption. Practice theory describes the interaction of people with the world in a way that make sense to them to do. Practice theory adopted by Gram-Hansen describes 4 factors that govern energy consumption practices. Where by indoor heating practices are governed by: heating know-how, heating system knowledge, heating technology and energy saving engagement. The governing elements shape a practice and can in turn be shaped by a practice. Practice theory is used gain a deeper understanding of the factors involved in residential heat practices. Practice theory is used in this thesis to determine a way to detect residential behavioural aspect with the use of disaggregated consumption data. The 4 key factors are used to understand a resulting interaction of person with world in the sense of a practice. The governing elements have been analysed in multiple consumption studies by the use of interviews with residents by Gram-Hanssen. The use of disaggregated consumption data in practice assessment imposes an alternative way to assess the indoor climate practice.
The use of disaggregated consumption data in the assessment of indoor practice shift the focus form to the resulting practice action. Since the resulting thermostat practice can be detected in the disaggregated consumption data. Detailed energy consumption data is used to analyse an aspect of indoor climate practice, thermostat practices. Thermostat practices are analysed to determine similar patterns in thermostat settings and resulting energy consumption. Households can be classified on the basis of their thermostat practice instead of household/ dwelling characteristics.

Sub-research question 2: Clustering of thermostat practice
Hierarchical clustering is used to determine similar thermostat practices with use of disaggregated energy consumption data. The clustering of thermostat practices results in the detection of 6 distinct practices of individual households. Whereby clustering is an approach to group similar thermostat practices of individual households. The resulting thermostat practices represented by the different clusters display clear difference in thermostat settings.
The clustering of thermostat practice result in preliminary insight in the gas consumption of dwellings, with an apprehend division in relative high and low natural gas consumption between dwellings. Most of the households do not adapt their specific thermostat practices to separate weekdays. Similar household sizes share commonalities in their thermostat practices, especially one-person household. One person households are more likely to lower their thermostat during the daytime.
Clustering of overall thermostat practices result in the formation distinct thermostat practices. The specific thermostat practices generate insight in possible and potential to energy consumption savings in the residential sector, by determining potential inefficient thermostat practices in a sizeable part of the sample group.

Sub-research question 3: Occupancy detection
To what extent thermostat practices are fitted to at home occupancy is determined with the use of electricity consumption data. Disaggregated electricity consumption data can be used to determine whether residents are at home or not, by detecting periods of relative high electricity consumption out of average 10-minute interval electricity consumption data. For 76% of the households in the pilot group, analysing electricity consumption data with the use of the Habit engine results in a weekly activity program that is in line with their habits. The successfulness of an overall detected at home activity is dependent on the regularity of the agendas of the analysed residents.
The thermostat settings of Toon users are in line with their detected overnight and evening occupancy. The thermostat is lowered overnight and turned on again in the evening. During the day, the thermostat programs mismatch the at home activity of residents. The majority of the households who do not heat their home during daytime tend to be away. This is not the case for households who heat their dwelling at daytime.

Around 50% of the residents are actually detected as being at home of the household who heat their home during daytime.
The mismatch between thermostat practises and at home activity of residents indicate potential gas saving in the residential sector. Disaggregated electricity consumption data can be used to determine the occupancy in dwellings. The fitness of thermostat practices and occupancy can be analysed, to identify potential savings and assess the extent of possible thermostat practice adjustments in the residential sector.

Sub-research question 4: Heating adjustments
Combing thermostat practices with at home occupancy detection provides insights in the potential saving in the residential sector. By identifying potential areas of saving combined with calculating the potential saving. Adjusting thermostat settings of individual households on average leads to a saving of 4.2%, resulting in €34 energy bill reduction per year. The potential saving is dependent on the individual thermostat practices and dwelling characteristics. Individual thermostat practices influence the saving potential of individual households by a factor 2. The thermal inertia of a dwelling impacts the natural gas consumption of dwellings as expected but is not in line with current classical energy assessments.
The individual thermostat practices have shown to have a larger effect on the potential residential saving than classic classification characteristics. The individual thermostat practices need to be taken into account in the overall residential energy assessment in combination with the dwelling specific heat inertia.
The disparate residential potential heat saving of individual dwellings indicates the importance of adjusting energy assessments accordingly. Both thermostat practices as dwelling specific heat loss has shown to have a substantial influence. Therefore, energy consumption in the residential sector should include individual disaggregated consumption data to assess these factors.

6.2.2. Answering research question

There are 3 insights that can be gained from analysing thermostat practices with the use of disaggregated consumption data for the analysed households. First of all, half of the household have a thermostat practice of heating their dwelling during the day. Around 50% of the residents who heat during the day are actually detected as being at home. Secondly the impact of individual heating patterns on the gas consumption of dwelling is bigger than classical household characteristics. The specific gas consumption of individual dwellings with similar thermostat practices differs., indicating that household specific characteristics play large role in the overall energy consumption. Thirdly the potential natural gas savings of the residential sector is on average 4.2% with the majority of the saving generated by lowering the thermostat during the night to 15 degrees. Residents can save on average €34 on their natural gas consumption bill with more than a factor 2 difference for separate thermostat practices. The relative low potential gas saving by adjusting thermostat settings of individual dwellings indicate that adjusting thermostat practices in the residential sector can assist in reaching the climate goals., but more measures are needed in order to significantly reduce the natural gas consumption in the residential sector.
There are two insight that are applicable to general population. Firstly, the individual thermostat practices are influential in the saving potential in the residential sector. For households with unmatched dwelling the individual thermostat practice influence the potential saving by more than a factor 2. Secondly, the current assessment of the thermal inertia of dwellings with the use of dwelling characteristics is not in line with the detected thermal inertia of dwelling. The predicted thermal inertia based upon the age and size of a dwelling overestimates the heat loss of a dwelling. The influence of individual thermostat practices and thermal inertia of dwellings can be taken into account with the use of disaggregated consumption data. Disaggregated smart meter/ thermostat data should be taken into account in the assessment of residential heat consumption. By using disaggregated consumption data in energy prediction models the current uncertainty around embedded factors that influence gas consumption can be reduced.
The findings of this research represent the saving potential for residents with a smart thermostat. The potential savings of residents with a manual thermostat to a less user friendly programmable thermostat is disparate, due to the difference in thermostat interaction [51]. The overall estimated saving for the residential sector could both be under and overestimated, depend on thermostat interaction. The possible over or under estimation is independent of the impact that thermostat practices have in the residential sector.

6.3. Placing research in context

In this section the relevance of the conducted research for Quby, scientific community and policy makers is firstly discussed. Followed up by a discussion of the ethical implications of at non-intrusive monitoring and potential savings.

Research context: Quby

There are 2 aspect of this research that are important for the smart metering production company Quby. The first aspect is how to improve and use the habit engine within the company. As discussed in chapter 4 the habit engine is currently only validated on its capability to produce an overall at home week program. To improve the detection of at home activity in dwelling more variables can be taken into account. The detection of at home activity can be improved by including sound, detection sensors and mobile location determination. Sound, detection sensor and mobile location data can be used on the one hand to evaluate at home detection on the basis of electric consumption data. On the other hand, be used to determine at home occupancy. Next to energy assessment, at home occupancy detection can be used in different fields. For example, the health sector for improving assisted living of elderly.

The research showed that the potential saving of residents is disparate and dependent on dwelling and household characteristics of dwellings. One of the factors in successful energy practice adjustment is the engagement of residents. In order to keep residents engaged the provided information should be correct and understandable. Therefore, communicated potential saving information should be based on their energy consumption and adjusted to their thermostat practices, to keep households engaged. Overestimating potential savings will lead to a mistrust information and disengagement of residents. The discussion between scholars about the effectiveness of providing feedback to residents revealed several potential complications. A possibility to mitigate the negative feedback implications is by providing residents with an automated thermostat program. Whereby instead of adjusting thermostat behaviour the thermostat settings are adjusted to the individual needs with the flick of a button. In line with providing customers with information the automated thermostat program should sufficiently correct to ensure user engagement.

Research context: Scientific community

Current bottom-up estimation models are still predominately based upon dwelling specific characteristics [43], while the impact of household behaviour has been addressed by scholars. The impact of household behaviour on energy consumption in the residential sector has been assessed with the use of questionnaires. The use of questioners is a time expensive manner of conducting research. The executed research in this thesis showed that the influence of residential behaviour can be assessed with the use of disaggregated consumption data of individual households. In order to adapt energy consumption models, disaggregated consumption data can be used to generate deeper insight in the saving potential of the residential sector.

Analysing behavioural aspects with the use of disaggregated energy consumption data can be used in energy transition the world is currently going through. Behavioural aspects play an important role in the development of smart grids and move from carbon based fuels. Next to the behavioural aspect the thermal capabilities of dwellings can be assessed with the use of disaggregated consumption data. For example, the thermal inertia of dwelling must be sufficient in order for a heat pump to work, an alternative for gas heating. Preliminary insight in the potential roll out of heat pumps can be assessed by the use of disaggregated consumption data.

Placing research in context: Policy makers

The impact of thermostat practices on the potential gas saving of unmatched dwellings indicates the role of thermostat behaviour in residential heat demand. The mismatch between estimated savings in the residential sector governed by assumptions and actual saving indicate the need to adjust policy measures. Current policy measures are based upon energy labels that classify energy consumption on the basis of physical dwelling characteristics. Chapter 5 showed that the assumptions based energy assessment used in energy labels overestimate the potential savings in the residential sector. In order for policy to be successfully implemented the measures need to be shaped to actual energy consumption. Therefore, the assessment of energy labels of dwellings should firstly include different thermostat profiles and can be improved further by including behavioural aspects of residents.

The difference between thermostat practices of residents showed another policy measure approach. Instead of steering on energy labels and physical characteristics of dwellings policy measures should be aimed at residential behaviour. Because the impact of thermostat practices of residents has more than a factor 2 impact on the potential saving for unmatched dwellings.

Ethical implications of non-intrusive monitoring and climate change

Climate change impacts the entire world and can lead to severe weather induced disasters to agricultural implications through draught. Climate change can be seen as the complex result of an increase of global warming gasses in the atmosphere. Whereby the direct impact of burning a cubic meter of natural gas is uncertain. In the view of Utilitarianism an action can be assessed by determining the impact of the action in the sense of wellbeing. The overall utility of an action is the result of summing all positive effects of an action minus the negative ones. The implication of estimating the overall utility of an action is that the sense of wellbeing is inherently personal and cannot be compared from person to person. In the case of assessing potential saving of adjusting thermostat of individual households there is contrast between residential privacy and climate change impact. The average saving potential of adjusting thermostat practices of toon users results in a 4.2% percent saving. The European Union personal privacy rights have been ensured and reinforced with a new law in May. The question is if it is ethical sound to infringe a person's right of privacy to reduce global warming pollution. On the grand scale the potential loss of life due to weather and drought disasters outweigh uncertain effect of reducing energy consumption embedded in potential privacy infringement, in my option. In the case of smart meter/ thermostat use to determine potential energy savings the implication on privacy can be mitigated. By separating data variables and anonymising data the privacy of residents can be guaranteed to an extent of theft and hacking of a device. In my opinion when privacy is guaranteeing in this way the benefits outweigh the costs even for relative small overall saving potential of 4.2%. In the end, the impact of world-wide climate change and impact on future generations outweigh potential privacy infringement, when data is used for the purpose of mitigating the effect of climate.

Bibliography

[1] Wokje Abrahamse, Linda Steg, Charles Vlek, and Talib Rothengatter. The effect of tailored information, goal setting, and tailored feedback on household energy use, energy-related behaviors, and behavioral antecedents. *J. Environ. Psychol.*, 27(4):265–276, 2007. ISSN 02724944. doi: 10.1016/j.jenvp.2007.08.002.

[2] Dorien Aerts, Joeri Minnen, Ignace Glorieux, Ine Wouters, and Filip Descamps. A probabilistic activity model to include realistic occupant behaviour in building simulations. *eSIM*, page 15, 2014. URL `http://www.ibpsa.org/proceedings/eSimPapers/2014/3A.4.pdf{%}5Cnhttp://www.ibpsa.org/?page{_}id=455`.

[3] E. E. Alders. Adaptive heating, ventilation and solar shading for dwellings. *Archit. Sci. Rev.*, 60(3):150–166, 2017. ISSN 17589622. doi: 10.1080/00038628.2017.1300132. URL `http://www.tandfonline.com/action/journalInformation?journalCode=tasr20`.

[4] Noortje Alders. *Adaptive thermal comfort opportunities for dwellings Providing thermal comfort only when and where needed in dwellings in the Netherlands*. PhD thesis, Delft University of Technology, 2016. URL `http://journals.library.tudelft.nl/index.php/faculty-architecture/article/view/alders/pdf{_}alders`.

[5] Kaustav Basu, Vincent Debusschere, Seddik Bacha, Ujjwal Maulik, and Sanghamitra Bondyopadhyay. Nonintrusive load monitoring: A temporal multilabel classification approach. *IEEE Trans. Ind. Informatics*, 11(1):262–270, feb 2015. ISSN 15513203. doi: 10.1109/TII.2014.2361288. URL `http://ieeexplore.ieee.org/document/6915905/`.

[6] C. E. Bonferroni. *l calcolo delle assicurazioni su gruppi di teste*. Rome, 1935.

[7] Dirk Brounen, Nils Kok, and John M. Quigley. Residential energy use and conservation: Economics and demographics. *Eur. Econ. Rev.*, 56(5):931–945, 2012. ISSN 00142921. doi: 10.1016/j.euroecorev.2012.02.007. URL `https://ac.els-cdn.com/S0014292112000256/1-s2.0-S0014292112000256-main.pdf?{_}tid=19cfb992-be4b-11e7-bdf1-00000aab0f01{&}acdnat=1509461722{_}41d68a2bcc868cf15de3518aa3a5d147http://ac.els-cdn.com/S0014292112000256/1-s2.0-S0014292112000256-main.pdf?{_}tid=4f3650ee`.

[8] Kathryn Buchanan, Riccardo Russo, and Ben Anderson. Feeding back about eco-feedback: How do consumers use and respond to energy monitors? *Energy Policy*, 73:138–146, 2014. ISSN 03014215. doi: 10.1016/j.enpol.2014.05.008.

[9] Kathryn Buchanan, Riccardo Russo, and Ben Anderson. The question of energy reduction: The problem(s) with feedback. *Energy Policy*, 77:89–96, 2015. ISSN 03014215. doi: 10.1016/j.enpol.2014.12.008. URL `http://ac.els-cdn.com/S0301421514006739/1-s2.0-S0301421514006739-main.pdf?{_}tid=c050af0e-38e7-11e7-b59b-00000aab0f27{&}acdnat=1494795547{_}038ad7f95705c74aa2e44894d5701bfc`.

[10] Sandro Casal, Nives DellaValle, Luigi Mittone, and Ivan Soraperra. Feedback and efficient behavior. *PLoS One*, 12(4):e0175738, 2017. ISSN 19326203. doi: 10.1371/journal.pone.0175738. URL `http://dx.plos.org/10.1371/journal.pone.0175738`.

[11] CBS. Energieverbruik particuliere woningen; woningtype en regio's. URL `http://statline.cbs.nl/Statweb/publication/`.

[12] CIBSE. *Degree Days*. Glasgow, United Kingdom, 1993. URL `http://www.cibse.org/getmedia/5a988d9d-7dd4-45e7-985c-eebb0564f6cf/FEB07-Degree-Days-1993.pdf.aspx`.

[13] Mariëlle Cloïn. Met het oog op de tijd: Een blik op de tijdsbesteding van Nederlanders. *Soc. en Cult. Planbur.*, pages 13–14, 2013. URL http://www.sportdeelname.nl/sportdeelname.nl/rapportages/leeswijzers/leeswijzer-met-het-oog-op-de-tijd-scp.pdf.

[14] Sarah Darby. Smart metering: What potential for householder engagement? *Build. Res. Inf.*, 38(5):442–457, 2010. ISSN 09613218. doi: 10.1080/09613218.2010.492660.

[15] Richard J. De Dear and Gail S. Brager. Thermal comfort in naturally ventilated buildings: Revisions to ASHRAE Standard 55. *Energy Build.*, 34(6):549–561, jul 2002. ISSN 03787788. doi: 10.1016/S0378-7788(02)00005-1. URL https://ac.els-cdn.com/S0378778802000051/1-s2.0-S0378778802000051-main.pdf?{_}tid=dabae466-b34b-11e7-b568-00000aacb35e{&}acdnat=1508252583{_}815e55844b10618d5ac1b12b248bacd0ht//linkinghub.elsevier.com/retrieve/pii/S0378778802000051.

[16] ECN, Energie-Nederland, and Netbeheer Nederland. Energietrends 2016. Technical report, ECN, Energie-Nederland en Netbeheer Nederland, 2016.

[17] Sr Eddy. Profile hidden Markov models. *Bioinformatics*, 14(9):755–763, 1998. ISSN 13674803. doi: btb114[pii]. URL http://bioinformatics.oxfordjournals.org/content/14/9/755.short.

[18] Ahmad Faruqui, Sanem Sergici, and Ahmed Sharif. The impact of informational feedback on energy consumption-A survey of the experimental evidence. *Energy*, 35(4):1598–1608, 2010. ISSN 03605442. doi: 10.1016/j.energy.2009.07.042.

[19] Corinna Fischer. Feedback on household electricity consumption: A tool for saving energy? *Energy Effic.*, 1(1):79–104, feb 2008. ISSN 1570646X. doi: 10.1007/s12053-008-9009-7. URL http://link.springer.com/10.1007/s12053-008-9009-7.

[20] Nelson Fumo. A review on the basics of building energy estimation. *Renew. Sustain. Energy Rev.*, 31:53–60, 2014. ISSN 13640321. doi: 10.1016/j.rser.2013.11.040. URL http://ac.els-cdn.com/S1364032113007892/1-s2.0-S1364032113007892-main.pdf?{_}tid=3ec170ca-5437-11e7-a8c0-00000aacb35d{&}acdnat=1497798371{_}e5788f596b14231f005559bade155896.

[21] Nelson Fumo and M. A. Rafe Biswas. Regression analysis for prediction of residential energy consumption. *Renew. Sustain. Energy Rev.*, 47:332–343, 2015. ISSN 18790690. doi: 10.1016/j.rser.2015.03.035. URL http://ac.els-cdn.com/S1364032115001884/1-s2.0-S1364032115001884-main.pdf?{_}tid=3a2f072a-2bf6-11e7-9ebf-00000aacb35e{&}acdnat=1493372399{_}5d975f406d4dc603afcfda0815e48020.

[22] Ray Galvin and Minna Sunikka-Blank. Schatzkian practice theory and energy consumption research: Time for some philosophical spring cleaning? *Energy Res. Soc. Sci.*, 22:63–68, 2016. ISSN 22146296. doi: 10.1016/j.erss.2016.08.021. URL https://ac.els-cdn.com/S2214629616302043/1-s2.0-S2214629616302043-main.pdf?{_}tid=3badc742-beec-11e7-83d0-00000aacb362{&}acdnat=1509530928{_}58ee278c2892ef60f694672074617b87.

[23] J. Gillis and N. Popovich. The U.S. Is the Biggest Carbon Polluter in History. It Just Walked Away From the Paris Climate Deal., jun 2017. URL https://www.nytimes.com/interactive/2017/06/01/climate/us-biggest-carbon-polluter-in-history-will-it-walk-away-from-the-paris-climate-deal.html.

[24] Kirsten Gram-Hanssen. Residential heat comfort practices: Understanding users. *Build. Res. Inf.*, 38(2):175–186, 2010. ISSN 09613218. doi: 10.1080/09613210903541527.

[25] Kirsten Gram-Hanssen. Standby consumption in households analyzed with a practice theory approach. *J. Ind. Ecol.*, 14(1):150–165, 2010. ISSN 10881980. doi: 10.1111/j.1530-9290.2009.00194.x.

[26] Kirsten Gram-Hanssen. Understanding change and continuity in residential energy consumption. *J. Consum. Cult.*, 11(1):61–78, mar 2011. ISSN 14695405. doi: 10.1177/1469540510391725. URL http://journals.sagepub.com/doi/pdf/10.1177/1469540510391725http://journals.sagepub.com/doi/10.1177/1469540510391725.

[27] Kirsten Gram-Hanssen. Efficient technologies or user behaviour, which is the more important when reducing households' energy consumption? *Energy Effic.*, 6(3):447–457, 2013. ISSN 1570646X. doi: 10.1007/s12053-012-9184-4.

[28] Olivia Guerra Santin, Laure Itard, and Henk Visscher. The effect of occupancy and building characteristics on energy use for space and water heating in Dutch residential stock. *Energy Build.*, 41(11): 1223–1232, 2009. ISSN 03787788. doi: 10.1016/j.enbuild.2009.07.002.

[29] Susana Guerreiro, Susana Batel, Maria Luísa Lima, and Sérgio Moreira. Making energy visible: sociopsychological aspects associated with the use of smart meters. *Energy Effic.*, 8(6):1149–1167, dec 2015. ISSN 15706478. doi: 10.1007/s12053-015-9344-4. URL https://link.springer.com/content/pdf/10.1007{%}2Fs12053-015-9344-4.pdfhttp://link.springer.com/10.1007/s12053-015-9344-4.

[30] Maria Halkidi. On Clustering Validation Techniques. *J. Intell. Inf. Syst.*, 173(2):107–145, 2001. URL http://web.itu.edu.tr/sgunduz/courses/verimaden/paper/validity{_}survey.pdf.

[31] Clare Hanmer, Michelle Shipworth, David Shipworth, and Edwin Carter. Household thermal routines and their impact on space heating demand patterns. In *ECEEE SUMMER STUDY Proc.*, pages 1133–1141, 2012.

[32] Anders Rhiger Hansen, Kirsten Gram-Hanssen, and Henrik N. Knudsen. How building design and technologies influence heat-related habits. *Build. Res. Inf.*, 46(1):83–98, 2018. ISSN 14664321. doi: 10.1080/09613218.2017.1335477.

[33] Sander Hermsen, Jeana Frost, Reint Jan Renes, and Peter Kerkhof. Using feedback through digital technology to disrupt and change habitual behavior: A critical review of current literature. *Comput. Human Behav.*, 57:61–74, 2016. ISSN 07475632. doi: 10.1016/j.chb.2015.12.023. URL http://ac.els-cdn.com/S0747563215302946/1-s2.0-S0747563215302946-main.pdf?{_}tid=5a20e7a4-38e6-11e7-9dbe-00000aab0f01{&}acdnat=1494794946{_}9e3112f62f6ea791bea746a1980ad2ef.

[34] J.A. The flaws in Donald Trump's decision to pull out of the Paris accord. *Econ. Gr. Ltd.*, jun 2017. URL http://www.economist.com/blogs/democracyinamerica/2017/06/america-and-climate-change.

[35] A. K. Jain, M. N. Murty, and P. J. Flynn. Data clustering: a review. *ACM Comput. Surv.*, 31(3):264–323, 1999. ISSN 03600300. doi: 10.1145/331499.331504. URL http://portal.acm.org/citation.cfm?doid=331499.331504.

[36] R Janssen. Towards energy efficient buildings in Europe. *Eur. Alliance Co. Energy Effic. Build.*, page 74, 2004. URL http://www.voltimum.es/files/gb/others/R/20070917934295232007091445502007083097032007083090406750EU.pdf.

[37] Thorben Jensen, Georg Holtz, and Émile J L Chappin. Agent-based assessment framework for behavior-changing feedback devices: Spreading of devices and heating behavior. *Technol. Forecast. Soc. Change*, 98:105–119, 2015. ISSN 00401625. doi: 10.1016/j.techfore.2015.06.006.

[38] Jurre Van den Berg. Bijtende gasadviezen stellen minister voor haast onmogelijke opgave - wat kan Wiebes nog?, feb 1. URL https://www.volkskrant.nl/binnenland/bijtende-gasadviezen-stellen-minister-voor-haast-onmogelijke-opgave-wat-kan-wiebes-nog{~}a4564911/

[39] M. Kavgic, A. Mavrogianni, D. Mumovic, A. Summerfield, Z. Stevanovic, and M. Djurovic-Petrovic. A review of bottom-up building stock models for energy consumption in the residential sector. *Build. Environ.*, 45(7):1683–1697, jul 2010. ISSN 03601323. doi: 10.1016/j.buildenv.2010.01.021. URL http://ac.els-cdn.com/S0360132310000338/1-s2.0-S0360132310000338-main.pdf?{_}tid=f89faa46-542a-11e7-b176-00000aab0f27{&}acdnat=1497793099{_}cb46b1fc3245779da2e7bdd9d97a89e8http://linkinghub.elsevier.com/retrieve/pii/S0360132310000338.

[40] Wilhelm Kleiminger, Christian Beckel, and Silvia Santini. Household occupancy monitoring using electricity meters. In *Proc. 2015 ACM Int. Jt. Conf. Pervasive Ubiquitous Comput. - UbiComp '15*, pages 975–986, New York, New York, USA, 2015. ACM Press. ISBN 9781450335744. doi: 10.1145/2750858.2807538. URL http://dl.acm.org/citation.cfm?doid=2750858.2807538.

[41] KNMI. Daggegevens van het weer in Nederland, 2016. URL `http://www.knmi.nl/nederland-nu/klimatologie/daggegevens`.

[42] Tamar Krishnamurti, Alexander L. Davis, Gabrielle Wong-Parodi, Jack Wang, and Casey Canfield. Creating an in-home display: Experimental evidence and guidelines for design. *Appl. Energy*, 108:448–458, 2013. ISSN 03062619. doi: 10.1016/j.apenergy.2013.03.048.

[43] Hyunwoo Lim and Zhiqiang John Zhai. Review on stochastic modeling methods for building stock energy prediction. *Build. Simul.*, 10(5):607–624, oct 2017. ISSN 19968744. doi: 10.1007/s12273-017-0383-y. URL `https://link.springer.com/content/pdf/10.1007{%}2Fs12273-017-0383-y.pdfhttp://link.springer.com/10.1007/s12273-017-0383-y`.

[44] D. Majcen, L. C.M. Itard, and H. Visscher. Theoretical vs. actual energy consumption of labelled dwellings in the Netherlands: Discrepancies and policy implications. *Energy Policy*, 54:125–136, mar 2013. ISSN 03014215. doi: 10.1016/j.enpol.2012.11.008. URL `http://dx.doi.org/10.1016/j.enpol.2012.11.008http://linkinghub.elsevier.com/retrieve/pii/S0301421512009731`.

[45] Daša Majcen, Laure Itard, and Henk Visscher. Actual and theoretical gas consumption in Dutch dwellings: What causes the differences? *Energy Policy*, 61:460–471, oct 2013. ISSN 03014215. doi: 10.1016/j.enpol.2013.06.018. URL `http://linkinghub.elsevier.com/retrieve/pii/S030142151300503X`.

[46] Daša Majcen, Laure Itard, and Henk Visscher. Statistical model of the heating prediction gap in Dutch dwellings: Relative importance of building, household and behavioural characteristics. *Energy Build.*, 105(2015):43–59, oct 2015. ISSN 03787788. doi: 10.1016/j.enbuild.2015.07.009. URL `http://linkinghub.elsevier.com/retrieve/pii/S0378778815301262`.

[47] Milieu Centraal. Bespaartips verwarming, 2017. URL `https://www.milieucentraal.nl/energie-besparen/snel-besparen/bespaartips-verwarming/`.

[48] Asit Kumar Mishra and Maddali Ramgopal. Field studies on human thermal comfort - An overview. *Build. Environ.*, 64:94–106, jun 2013. ISSN 03601323. doi: 10.1016/j.buildenv.2013.02.015. URL `https://ac.els-cdn.com/S036013231300070X/1-s2.0-S036013231300070X-main.pdf?{_}tid=464b41c8-b49e-11e7-a13b-00000aacb362{&}acdnat=1508397933{_}7e9d570340bb74c363ef6e4c1ac92c58ht//linkinghub.elsevier.com/retrieve/pii/S036013231300070X`.

[49] J. Fergus Nicol and Susan Roaf. Rethinking thermal comfort. *Build. Res. Inf.*, 45(7):711–716, oct 2017. ISSN 14664321. doi: 10.1080/09613218.2017.1301698. URL `https://www.tandfonline.com/doi/full/10.1080/09613218.2017.1301698`.

[50] Leen Peeters, Richard de Dear, Jan Hensen, and William D'haeseleer. Thermal comfort in residential buildings: Comfort values and scales for building energy simulation. *Appl. Energy*, 86(5):772–780, 2009. ISSN 03062619. doi: 10.1016/j.apenergy.2008.07.011. URL `https://ac.els-cdn.com/S0306261908001785/1-s2.0-S0306261908001785-main.pdf?{_}tid=5f8827be-b49e-11e7-99ea-00000aab0f27{&}acdnat=1508397976{_}c4a503ad86b57dfee829f5f2ceb3ece7`.

[51] Therese Peffer, Marco Pritoni, Alan Meier, Cecilia Aragon, and Daniel Perry. How people use thermostats in homes: A review, 2011. ISSN 03601323.

[52] Luis Pérez-Lombard, José Ortiz, and Christine Pout. A review on buildings energy consumption information. *Energy Build.*, 40(3):394–398, jan 2008. ISSN 03787788. doi: 10.1016/j.enbuild.2007.03.007. URL `http://ac.els-cdn.com/S0378778807001016/1-s2.0-S0378778807001016-main.pdf?{_}tid=ac101772-57f1-11e7-a1c5-00000aab0f01{&}acdnat=1498208294{_}4a995ab223f1885837bf5b6a82bdaaa6ht//linkinghub.elsevier.com/retrieve/pii/S0378778807001016`.

[53] Marco Pritoni, Jonathan M. Woolley, and Mark P. Modera. Do occupancy-responsive learning thermostats save energy? A field study in university residence halls. *Energy Build.*, 127:469–478, 2016. ISSN 03787788. doi: 10.1016/j.enbuild.2016.05.024. URL `http://ac.els-cdn.com/S0378778816303851/1-s2.0-S0378778816303851-main.pdf?{_}tid=4240bfc8-55c5-11e7-a186-00000aacb35d{&}acdnat=1497969318{_}0f2d91b43f5524ca63bc8a66f992ce49`.

[54] Raoul du Pre. Wiebes: grote bedrijven moeten binnen 5 jaar stoppen met Gronings gas, jan 2018. URL https://www.volkskrant.nl/economie/wiebes-grote-bedrijven-moeten-binnen-5-jaar-stoppen-met-gronings-gas{~}a4561358/.

[55] Ian Richardson, Murray Thomson, David Infield, and Conor Clifford. Domestic electricity use: A high-resolution energy demand model. *Energy Build.*, 42(10):1878–1887, oct 2010. ISSN 03787788. doi: 10.1016/j.enbuild.2010.05.023. URL https://www.sciencedirect.com/science/article/pii/S0378778810001854.

[56] H. Sarak and A. Satman. The degree-day method to estimate the residential heating natural gas consumption in Turkey: A case study. *Energy*, 28(9):929–939, 2003. ISSN 03605442. doi: 10.1016/S0360-5442(03)00035-5.

[57] Sanja Šćepanović, Martijn Warnier, and Jukka K. Nurminen. The role of context in residential energy interventions: A meta review. *Renew. Sustain. Energy Rev.*, 77:1146–1168, sep 2017. ISSN 18790690. doi: 10.1016/j.rser.2016.11.044. URL www.elsevier.com/locate/rserhttp://linkinghub.elsevier.com/retrieve/pii/S1364032116308310.

[58] K. Shimada, Y. Ochi, T. Matsumoto, H. Matsugi, and T. Awata. An empirical study of real-time feedback and dynamic pricing effects on electric power consumption: Field experiment on a remote island in Japan. *SMARTGREENS 2015 - 4th Int. Conf. Smart Cities Green ICT Syst. Proc.*, pages 201–208, 2015. doi: 10.5220/0005434402010208.

[59] Gustavo Sousa, Benjamin M. Jones, Parham A. Mirzaei, and Darren Robinson. A review and critique of UK housing stock energy models, modelling approaches and data sources. *Energy Build.*, 151:66–80, 2017. ISSN 03787788. doi: 10.1016/j.enbuild.2017.06.043. URL https://ac.els-cdn.com/S0378778817301159/1-s2.0-S0378778817301159-main.pdf?{_}tid=499f7e08-a359-11e7-81a9-00000aab0f02{&}acdnat=1506499134{_}64af2db93c43c0e7810b503ffaf87a10.

[60] Lukas G. Swan and V. Ismet Ugursal. Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. *Renew. Sustain. Energy Rev.*, 13(8):1819–1835, oct 2009. ISSN 13640321. doi: 10.1016/j.rser.2008.09.033. URL http://ac.els-cdn.com/S1364032108001949/1-s2.0-S1364032108001949-main.pdf?{_}tid=3aa635ae-542b-11e7-a5d2-00000aab0f6c{&}acdnat=1497793210{_}4242da640b2f282ce1dfab415ab5b788http://linkinghub.elsevier.com/retrieve/pii/S1364032108001949.

[61] Satoshi Takumi and Sadaaki Miyamoto. Top-down vs bottom-up methods of linkage for asymmetric agglomerative hierarchical clustering. In *2012 IEEE Int. Conf. Granul. Comput.*, pages 459–464. IEEE, aug 2012. ISBN 978-1-4673-2311-6. doi: 10.1109/GrC.2012.6468689. URL http://ieeexplore.ieee.org/document/6468689/.

[62] Robert L. Thorndike. Who belongs in the family? *Psychometrika*, 18(4):267–276, 1953. ISSN 00333123. doi: 10.1007/BF02289263. URL https://link.springer.com/content/pdf/10.1007/BF02289263.pdf.

[63] Tsuyoshi Ueno, Fuminori Sano, Osamu Saeki, and Kiichiro Tsuji. Effectiveness of an energy-consumption information system on energy savings in residential houses based on monitored data. *Appl. Energy*, 83(2):166–183, 2006. ISSN 03062619. doi: 10.1016/j.apenergy.2005.02.002. URL www.elsevier.com/locate/apenergy.

[64] Paula van den Brom, Arjen Meijer, and Henk Visscher. Performance gaps in energy consumption: household groups and building characteristics. *Build. Res. Inf.*, 46(1):54–70, may 2018. ISSN 14664321. doi: 10.1080/09613218.2017.1312897. URL https://doi.org/10.1080/09613218.2017.1312897https://www.tandfonline.com/doi/full/10.1080/09613218.2017.1312897.

[65] E.R. van den Ham and J. van der Vliet. Energiebesparing door selectief stoken bij woningen Inhoud. Technical report, Milieu Centraal, TU Delft, Utrecht, 2013.

[66] Erik van der Walle. Grootverbruikers Gronings gas: brief Wiebes intimideert, jan 2018. URL https://www.nrc.nl/nieuws/2018/01/22/grootverbruikers-gronings-gas-brief-wiebes-intimideert-a1589311.

[67] Iana Vassileva, Erik Dahlquist, Fredrik Wallin, and Javier Campillo. Energy consumption feedback devices' impact evaluation on domestic energy use. *Appl. Energy*, 106:314–320, 2013. ISSN 03062619. doi: 10.1016/j.apenergy.2013.01.059.

[68] Anne Vegterlo. Wie werkt er nog op vrijdag?, nov 2015. URL https://www.nrc.nl/nieuws/2015/11/04/woensdag-papadag-vrijdag-vrij-1554532-a540025.

[69] Alan Warde. Consumption and theories of practice. *J. Consum. Cult.*, 5(2):131–153, jul 2005. ISSN 14695405. doi: 10.1177/1469540505053090. URL http://journals.sagepub.com/doi/10.1177/1469540505053090http://www.ncbi.nlm.nih.gov/pubmed/16337333.

[70] Peter Wyatt. A dwelling-level investigation into the physical and socio-economic drivers of domestic energy consumption in England. *Energy Policy*, 60:540–549, 2013. ISSN 03014215. doi: 10.1016/j.enpol.2013.05.037. URL http://ac.els-cdn.com/S0301421513003728/1-s2.0-S0301421513003728-main.pdf?{_}tid=405d0de2-543a-11e7-9a04-00000aab0f6c{&}acdnat=1497799662{_}da78ea14ed0dd0840be16c058259c291.

[71] Da Yan, William O'Brien, Tianzhen Hong, Xiaohang Feng, H. Burak Gunay, Farhang Tahmasebi, and Ardeshir Mahdavi. Occupant behavior modeling for building performance simulation: Current state and future challenges. *Energy Build.*, 107:264–278, 2015. ISSN 03787788. doi: 10.1016/j.enbuild.2015.08.032.

[72] Lina Yang and Yuguo Li. Cooling load reduction by using thermal mass and night ventilation. *Energy Build.*, 40(11):2052–2058, jan 2008. ISSN 03787788. doi: 10.1016/j.enbuild.2008.05.014. URL https://www.sciencedirect.com/science/article/pii/S0378778808001308{#}aep-section-id23.

[73] Xie Zhenhua. The Paris agreement marks an unprecedented political recognition of the risks of climate change. *Econ. Gr. Ltd.*, pages 12–14, dec 2015. URL http://www.economist.com/news/international/21683990-paris-agreement-climate-change-talks.

A

Scientific paper

Heat load profile detection with the use of individual thermostat settings

T.H. Vosmer

Faculty of Technology and Management, Delft University of Technology

Abstract—Natural gas is the predominant source of energy in the residential sector, primarily used for space heating. Reducing the consumption of carbon based sources is one of the pathways to reduce the worldwide temperature rise. Current European policy measures hinge on the energy assessment of dwellings with the use of labels. Energy labels estimate theoretical energy consumption on the basis of physical characteristics of dwellings. There are discrepancies between the estimated energy consumption in the determination of energy labels and actual residential energy consumption. In this paper, residential heat demand is assessed with the use of hierarchical clustering of thermostat programs of individual households, to analyse the influence of individual heat demand on residential energy consumption. The clustering of thermostat settings result in 6 distinct heat load profiles and the identification of two potential residential saving areas. The influence of individual heat load profiles is considerable and should be included in residential energy assessments, to adequately shape policy measures to move away from carbon based sources.

Index Terms—Heat load profile, Hierarchical clustering, Natural gas consumption, Energy label, Residential sector

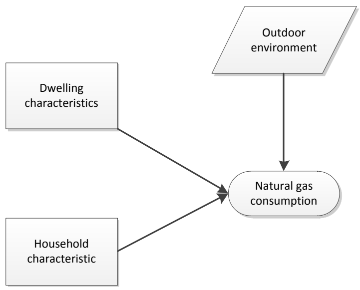


Fig. 1. Influencing factors residential gas consumption

1 Introduction

Reducing the consumption of carbon based energy sources is one of the goals set in the climate accord in Paris, to limit the world-wide temperature rise below 2 degrees [1]. The residential sector is responsible for over 55% of the natural gas consumption in the Netherlands [2], predominantly used for the heating of dwellings [3]. Residential energy policy objectives in the European Union to improve energy efficiency are based upon energy labels of dwellings [4].

The amount of natural gas consumed by individual households is dissimilar and dependent on multiple factors. The compulsory energy label of dwellings in the European Union, described in the Energy Performance of Building Directive (EPBD), is based upon theoretical energy consumption. The estimated heat consumption of a dwelling used for the classification of energy labels is based upon the physical characteristics of that dwelling. There are discrepancies between the estimated heat consumption based upon physical characteristics and actual heat consumption of dwellings [5]. Even households with similar dwelling and similar household characteristics have shown dissimilarities in their heat consumption [6]. Whereby residential behaviour is seen as the largest source of uncertainty in the prediction of energy consumption [7].

Individual residential natural gas consumption is influenced by 3 factors [8], graphically shown in figure 1. Dwelling characteristics and household characteristics can be adjusted by residents while the outdoor environment is dependent on the location of the dwelling. Dwelling characteristics account for the physical characteristics of a dwelling related to the thermal inertia of a dwelling. The influence of residential behaviour and household composition is embedded in household characteristics.

The influence of household characteristics in the form of thermostat settings is analysed by determining heat load profiles. Heat load profiles can be seen as the overall thermostat temperature settings of an individual dwellings. A heat load profile is the resulting thermostat setting dependent on the interaction of household composition and household behaviour. Similarities in heat load profiles are detected with the use of individual thermostat settings of dwellings. Individual household thermostat settings are assessed with the use of smart meter/ thermostat Toon. Individual heat load profiles in the form of thermostat settings are grouped with the use of unsupervised clustering. The impact of heat load profiles is assessed by determining the individual amount of natural gas consumed for heating a dwelling. Residential thermostat settings are clustered on the one hand to analyse

the degree of similarities in heat load profiles. And on the other hand, to analyse the impact of distinctive heat load profiles on residential gas consumption.
In section 2 the data gatherer to detect overall heat load profiles of individual dwellings is discussed. Followed up by the clustering method and type used to group similar heat load profiles. In section 3 the resulting heat load profiles are presented by discussing differences and similarities between heat demand profiles. The difference in natural gas consumption for each of the heat load profiles is discussed in section 4. The concluding remarks and recommendations are given in section 5.

2 METHOD

2.1 Clustering data

Clustering is used to group individual households with similar thermostat settings in the form of heat load profiles. The similar heat load profiles of household are assessed by analysing the thermostat settings and gas consumption of 1319 households. The data is gathered with the smart meter/ thermostat Toon. The thermostat settings are collected in the 3 coldest months of the winter of 2016-2017, December-February. The thermostat settings are collected in one hour intervals for all the weekdays in the winter period. The hourly thermostat settings consist of the average temperature settings within a dwelling for that hour. The thermostat settings for similar weekdays are grouped to produce a general thermostat setting overview of individual households for separate weekdays. To illustrate, the median is taken of the thermostat settings of every Wednesday at eleven within the three-month period for individual households. The heat load profile of households for similar weekdays is based upon the median hourly thermostat setting within the winter collection period. The weekends are left out of the scope due to disparate household habits between weekends. Because the median thermostat setting for similar hours in the weekend does not generate a representative overall thermostat setting program.
The amount of natural gas consumed for heating a dwelling is used to analyse the impact of separate heat load profiles. The natural gas consumption for heating is filtered out of the natural gas consumption by adjusting for non-heating events, with the use of smart thermostat and boiler communication.

2.2 Clustering method

The unsupervised clustering type of hierarchical clustering is used to determine similar heat load profiles. Unsupervised clustering is used to group sets of data that are unlabelled and is seen one of the most useful techniques for discovering patterns in underlying data [9]. Hierarchical clustering treats individual data cases as independent clusters and combines the clusters with the least distance between them. The distance between the clusters is based upon the similarity measure. The Ward similarity measure is used in the detection of overall heat load profiles. Out of the 7 clustering schemes provide by [10], the Ward similarity measure proofed most suitable in detection overall heat load profiles. The Ward similarity measure scored best on

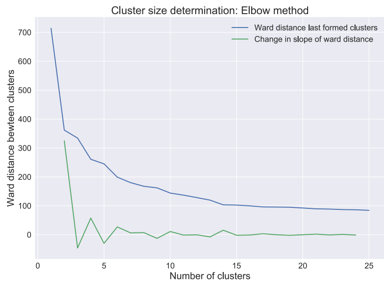


Fig. 2. Cluster size determination: elbow method

separating power and the resulting clusters demonstrate the lowest variance in thermostat settings within the separate heat load profiles.
The Ward similarly measure combines clusters on the basis of least variance between different clusters. The hourly variance in thermostat setting of the individual heat profiles is used in the Ward similarity method to determine the distance between clusters. The individual heat load profiles consist of the median thermostat settings of individual users. The heat load profile of each weekday is taken into account separately in the clustering analysis and henceforth called the user weekdays. User weekdays with the least variance in overall hourly thermostat settings are grouped together until a stopping condition is met.
The stopping condition for the ward similarity measure is determined with the use of the Elbow Method [11]. The Elbow method is used to distinguish the cluster sizes at which the drop variance within the clusters is largest. In figure 2 the Ward distance between the last formed clusters and the change in distance is shown. The distance between the last formed clusters demonstrates the largest distance between 2 clusters in the clustering analysis. Figure 2 indicates that the change in distance between the last formed clusters is largest at cluster sizes [4,6,10,14]. Each of the 4 cluster size stopping conditions are compared on the basis of similarities in resulting heat load profiles and representativeness of the resulting heat load profiles. A stopping condition of cluster size 6 scored best in the determination of similar heat load profiles. At cluster size 6, the similarity in heat load profiles is minimalised and the clusters that represent a substantial part of the population is maximised.

3 DETECTED HEAT LOAD PROFILES

The hierarchical clustering of overall thermostat settings of individual user weekdays is shown in figures 3, 4, 5, 6, 7 and 8. The heat load profiles are illustrated by the box-plot of the hourly thermostat settings for individual user weekdays. The box contains the midspread of 50% of the temperature settings and 95% of the settings are contained within the lines.

The heat load profiles of individual user weekdays share overlapping elements while being clearly distinguishable.

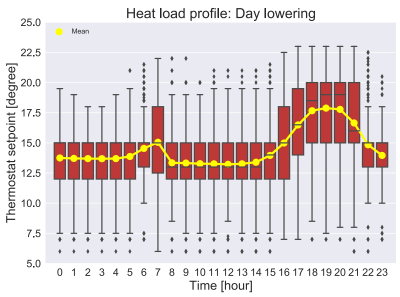


Fig. 3. Heat load profile: Day lowering

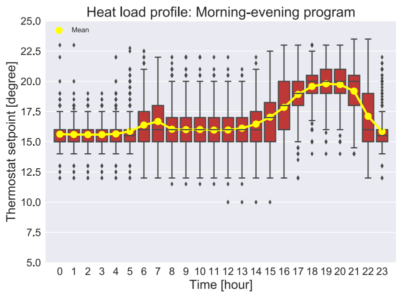


Fig. 4. Heat load profile: Morning-evening program

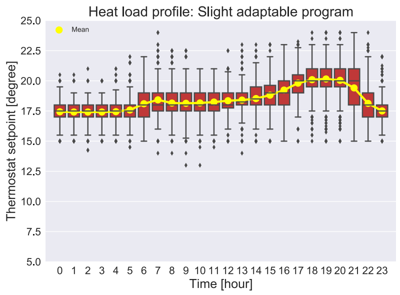


Fig. 5. Heat load profile: Slight adaptable program

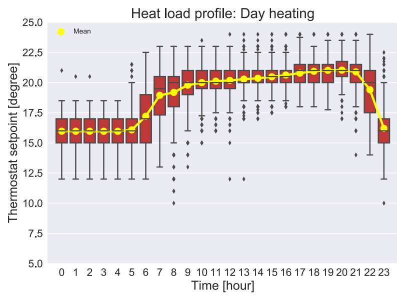


Fig. 6. Heat load profile: Day heating

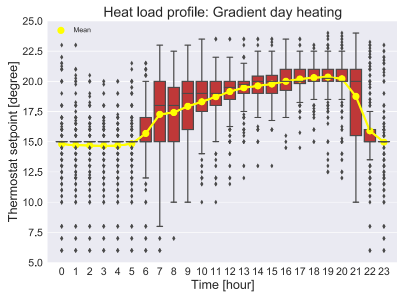


Fig. 7. Heat load profile: Gradient day heating

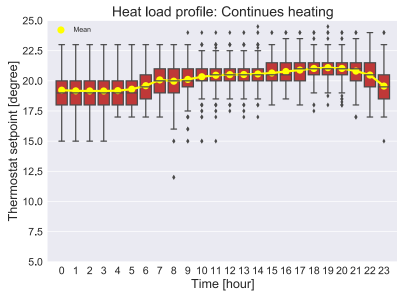


Fig. 8. Heat load profile: Continues heating

The different heat load profiles have 2 similarities in common between all the clusters. First of all, in each of the load profiles the thermostat settings are lowered overnight, compared to the day time settings. Secondly the moments when the thermostat is turned up and down. In each of the heat load profiles the thermostat is turned on in the morning and off again in the evening.

There is a clear separation between the heat load profiles during the day time in temperature settings. Figures 3, 4 and 5 represent the user weekdays with a lowered day time thermostat setting compared to figures 6, 7 and 8 with a heating daytime setting. Within the heat load profiles with a lowered thermostat setting a similar pattern can be detected. The thermostat is turned on and off again in the morning within these heat load profiles.

The exemplification of each of the heat load profiles is

shown in figure 1, in the form of percentage of households in the corresponding cluster. Two potential saving areas can be identified by analysing the different heat load profiles. Firstly in 3 of the 6 heat load profiles accounting for 51% of the user weekday population [D,E,F] have a day time heating pattern. Indicating that there is a potential to save natural gas consumption in the residential sector by lowering the thermostat during the day. The temperature can be lowered when residents are not actually at home or to a lower comfortable temperature. Secondly the majority of the user weekdays in heat load profiles [C,D,F] have an thermostat setting above 15 degrees during the night.
The clustering of heat load profiles indicates the existence of distinctive general heat load profiles. The clear difference between overall heat load profiles indicate the importance of including thermal profiles in the assessment of natural gas consumption in the residential sector. The second insight the clustering of heat load profiles generates is the possibility to target residential energy saving policy to high saving potential households. The consumption of natural gas for heating a dwelling is used to link heat load profiles to energy consumption in section 4.

4 Heat load profiles and gas consumption

The daily natural gas consumption of each of the heat load profiles is shown in a box-plot in figure 9. The box contains the midspread of the daily natural gas consumption in the collection period for 1319 households and 95% of the daily gas consumption is contained within the lines. The referencing index letter and corresponding heat load profile is shown in table 1. The heat load profiles consisting of the user weekdays without day time heating [A,B,C] clearly have a lower natural gas consumption that the heat load profiles who do [D,E,F]. Figure 9 displays the range in daily gas consumption, the natural gas consumption for each of the profiles clearly overlaps. The similarity in daily gas consumption is substantially larger in the heat load profile consisting of day time heating user weekdays.

To statistically determine the variation in daily natural gas consumption between the heat load profiles an one sided ANOVA is used. In the ANOVA test the average daily gas consumption per household in the winter period is used, shown in table 2. The daily average is taken to compare the overall heat load profile to the average daily natural gas consumption. The ANOVA results in table 2 support the findings in figure 9. There is a significant variance between all the different heat load profiles and the heat load profiles representing non day time heating. There is no significant

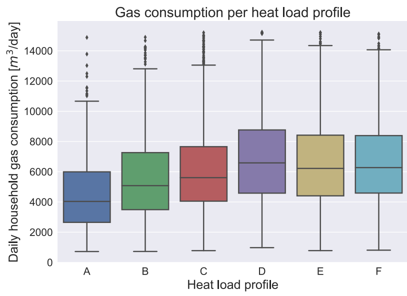


Fig. 9. Daily gas consumption for heat load profile

difference in the natural gas consumption of the heat load profiles of day time heaters.

The overlap in daily natural gas consumption is in line with the multiple factors that influence the natural gas consumption of dwellings shown in figure 1. The role of other influential factors than the thermostat settings is shown in the overlap between daily gas consumption. The considerable overlap in daily gas consumption of the day heating profiles [D,E,F], indicates that the influence of heat load profiles becomes less dominant when a dwelling is heated for longer periods during the day. The limited variance in natural gas consumption of day heating heat load profiles is expected. When a dwelling is heated for most of the day, the influence of alternative thermostat setting programs is limited.
The heat load profiles consist of dwellings with disparate physical characteristics of dwellings. Even for disparate dwellings a difference in daily natural gas consumption can be detected for separate heat profiles, indicating the impact of thermostat settings on natural gas consumption. The variance in natural gas consumption of separate heat load profiles indicate the potential to aim residential energy saving policy measures to inefficient thermostat use with relative high natural gas consumption.

5 Conclusion

With the use of hierarchal clustering of individual thermostat settings overall heat demand profiles can be detected. The clustering of individual thermostat settings results in the detection of 6 distinction heat load profiles. The heat demand profiles show clear differences in the thermostat settings in households with similarities in some aspects. The heat load profiles indicate two potential areas for reducing the residential natural gas consumption for around 50% of

TABLE 1
Heat load profiles and exemplification

Heat load profile		Exemplification heat load profile - Households in cluster [%]
A	Day lowering	11
B	Morning-evening program	19
C	Adaptable program	19
D	Day heating	21
E	Gradient day heating	18
F	Continues heating	12

TABLE 2
Variance daily gas consumption (ANOVA)

Clusters	*F value*	*p value*
All	50.9	$1.1 * 10^{-51}$
a,b,c	33.6	$4.1 * 10^{-15}$
d,e,f	2.4	0.088

the population, by lowering the thermostat overnight and during day time. To assess the extent of saving by lowering the thermostat during the day the thermal needs of residents need to be included. For example, by detecting whether residents are actually at home during the day.

The average daily gas consumption of separate thermal load profiles is dissimilar, especially in thermal load profiles without day time heating. The difference in daily gas consumption is visible for a mixture of dwellings with disparate dwelling characteristics. The difference in average daily gas consumption of separate heat load profiles displays the role of thermostat settings in the natural gas consumption of dwellings.

The clear disparate heat load profiles of individual households indicate the importance of including household characteristics residential energy policy. The heat load profiles conflict with the assumed constant inside temperature of 18 degrees in energy label determination [12]. The variation in natural gas consumption between dwellings with separate heat load profiles indicates the opportunity to aim policy measures at households with inefficient thermostat settings and relative high natural gas consumption.

5.1 Discussion

The thermostat settings of Toon users are gathered in the assessment of residential thermal load profiles. The household composition of the selected Toon users is not representative for the general population. Toon provides a preprogramed thermostat and enables users to easily adapt their thermostat programs to their needs with the use of an in-home display and mobile application. This is not the case for the general population whereby the thermostat interaction differ for users without or a different smart meter/ thermostat [13]. The distribution of users over the different heat load profiles and the profiles itself might differ from the general population due to non-representativeness of the sample group. Whereby the extent of inefficient thermostat setting might not resemble the general population. However the importance of including heat load profiles in residential energy policy is present because of the distinct heat load profiles and corresponding natural gas consumption.

References

[1] X. Zhenhua, "The Paris agreement marks an unprecedented political recognition of the risks of climate change," *Econ. Gr. Ltd.*, pp. 12–14, dec 2015.

[2] ECN, Energie-Nederland, and Netbeheer Nederland, "Energietrends 2016," tech. rep., ECN, Energie-Nederland en Netbeheer Nederland, 2016.

[3] R. Janssen, "Towards energy efficient buildings in Europe," *Eur. Alliance Co. Energy Effic. Build.*, p. 74, 2004.

[4] European Commission, "Energy efficiency: Buildings."

[5] D. Majcen, L. Itard, and H. Visscher, "Energy labels in Dutch dwellings their actual energy consumption and implications for reduction targets," *Eceee*, pp. 1947–1952, 2013.

[6] K. Gram-Hanssen, "Residential heat comfort practices: Understanding users," *Build. Res. Inf.*, vol. 38, no. 2, pp. 175–186, 2010.

[7] L. Yang and Y. Li, "Cooling load reduction by using thermal mass and night ventilation," *Energy Build.*, vol. 40, pp. 2052–2058, jan 2008.

[8] L. Diao, Y. Sun, Z. Chen, and J. Chen, "Modeling energy consumption in residential buildings: A bottom-up analysis based on occupant behavior pattern clustering and stochastic simulation," *Energy Build.*, vol. 147, pp. 47–66, 2017.

[9] M. Halkidi, "On Clustering Validation Techniques," *J. Intell. Inf. Syst.*, vol. 173, no. 2, pp. 107–145, 2001.

[10] D. Aerts, J. Minnen, I. Glorieux, I. Wouters, and F. Descamps, "A probabilistic activity model to include realistic occupant behaviour in building simulations," *eSIM*, p. 15, 2014.

[11] R. L. Thorndike, "Who belongs in the family?," *Psychometrika*, vol. 18, no. 4, pp. 267–276, 1953.

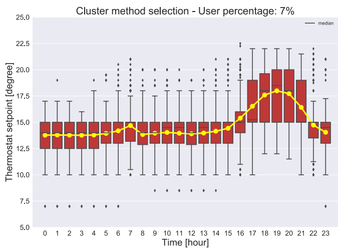
[12] D. Majcen, L. Itard, and H. Visscher, "Actual and theoretical gas consumption in Dutch dwellings: What causes the differences?," *Energy Policy*, vol. 61, pp. 460–471, oct 2013.

[13] T. Peffer, M. Pritoni, A. Meier, C. Aragon, and D. Perry, "How people use thermostats in homes: A review," 2011.

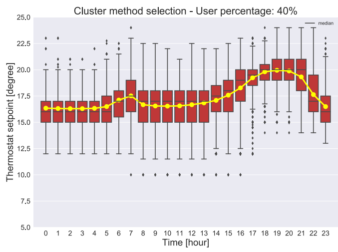
B
Clustering graphs

In this appendix, the clustering results are shown, for the different clusters size comparison, for the method selection and cluster validation.

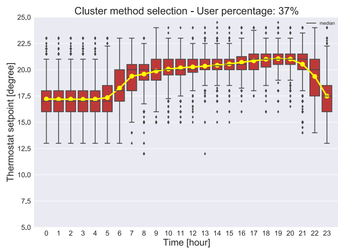
In this section, the 2 clustering method comparison outcomes are shown for the following methods: Ward and Complete. For each of the methods the clustering is in line with the described approach in chapter 3. The cluster outcomes for the stopping condition of clusters size 5,10 and 15 are shown. For the 2 different clustering methods both the significant clusters and the non-significant clusters are shown. The figures shown in this appendix consist of clustering representations by showing the box plot of the hourly median thermostat setting for individual users within that cluster. In figures B.1 B.3 B.4 the cluster comparison of the complete clustering method is shown for each of the different cluster sizes. The ward method clustering comparison is shown in figures B.5 B.6 B.7 for each of the cluster sizes.



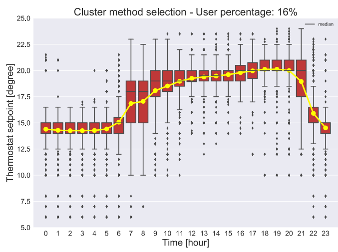
(a) Cluster 1



(b) Cluster 2

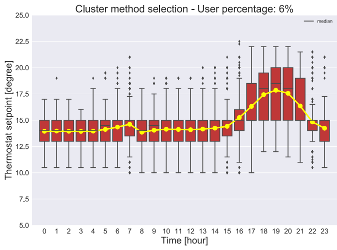


(c) Cluster 3

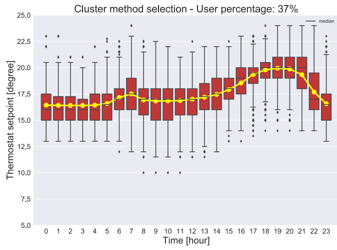


(d) Cluster 4

Figure B.1: Cluster method: Complete, Clustersize 5

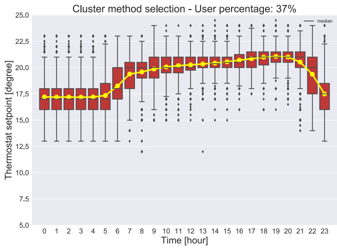


(a) Cluster 1

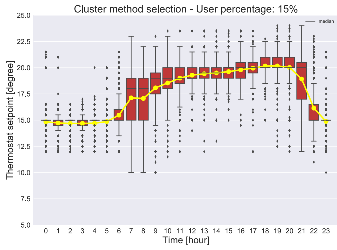


(b) Cluster 2

Figure B.2: Cluster method: Complete, Clustersize 10

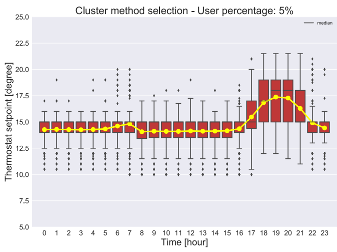


(a) Cluster 3

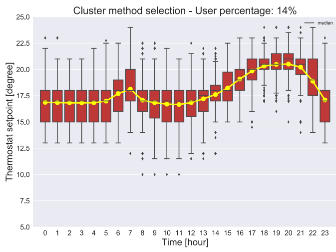


(b) Cluster 4

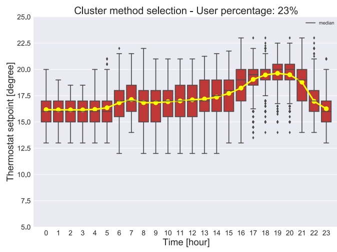
Figure B.3: Cluster method: Complete, Clustersize 10



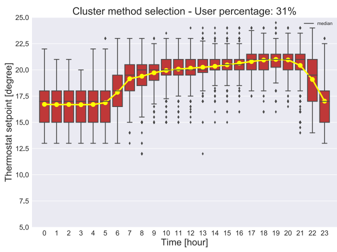
(a) Cluster 1



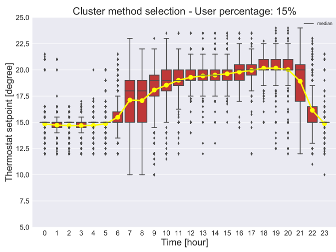
(b) Cluster 2



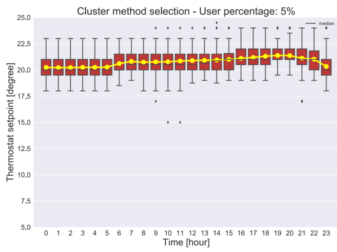
(c) Cluster 3



(d) Cluster 4

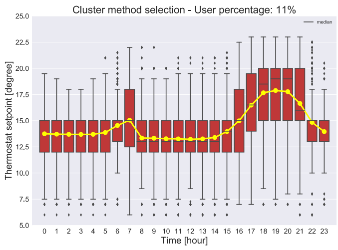


(e) Cluster 5

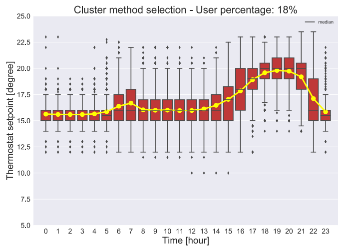


(f) Cluster 6

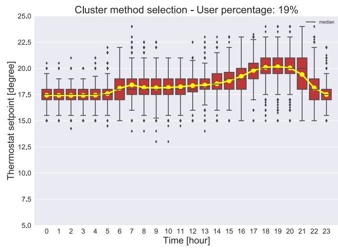
Figure B.4: Cluster method: Complete, Clustersize 15



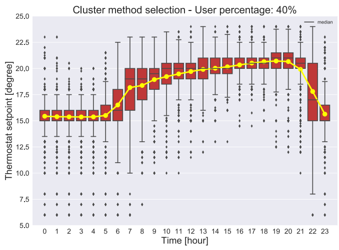
(a) Cluster 1



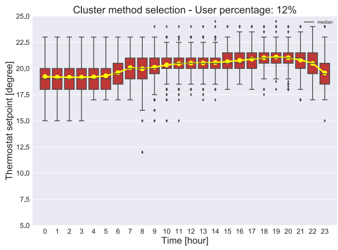
(b) Cluster 2



(c) Cluster 3

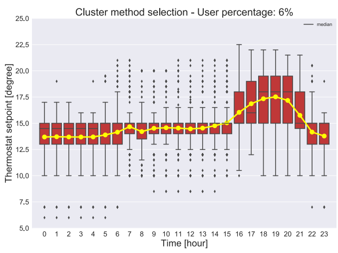


(d) Cluster 4

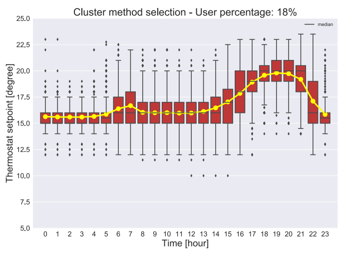


(e) Cluster 5

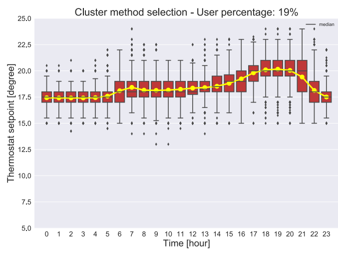
Figure B.5: Cluster method: Ward, Cluster size 5



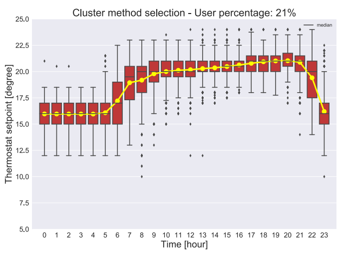
(a) Cluster 1



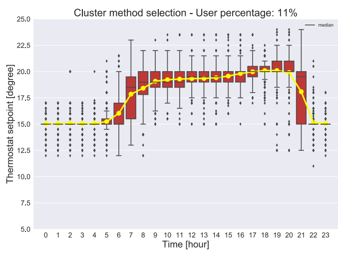
(b) Cluster 2



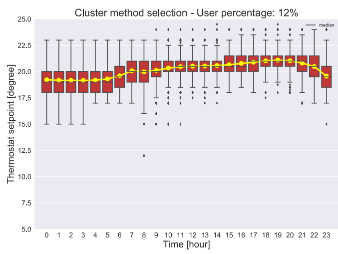
(c) Cluster 3



(d) Cluster 4

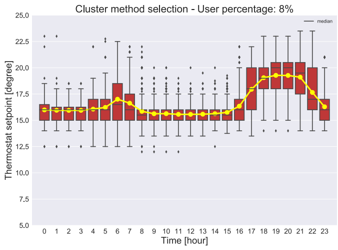


(e) Cluster 5

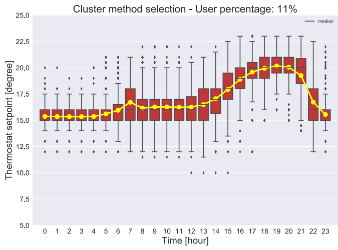


(f) Cluster 6

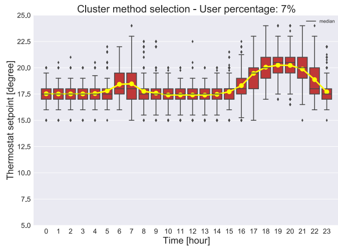
Figure B.6: Cluster method: Ward, Cluster size 10



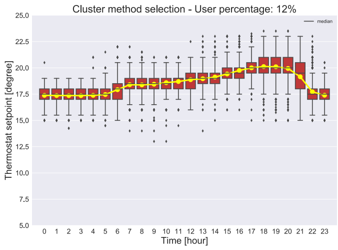
(a) Cluster 1



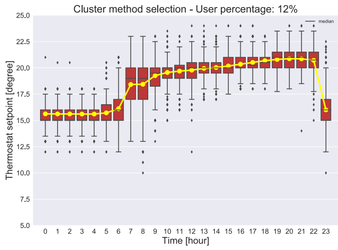
(b) Cluster 2



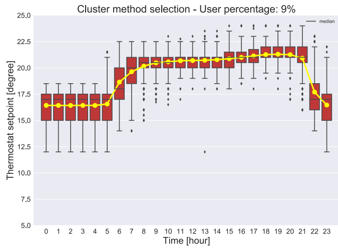
(c) Cluster 3



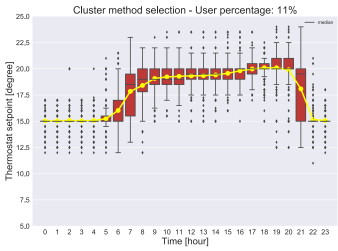
(d) Cluster 4



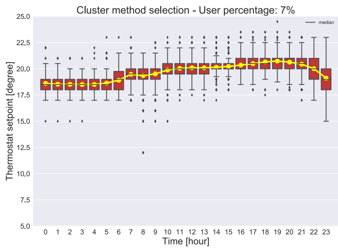
(e) Cluster 5



(f) Cluster 6



(g) Cluster 7

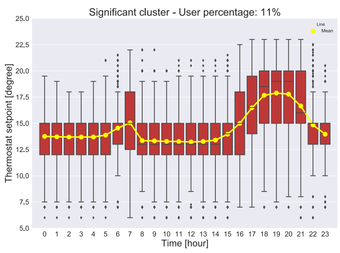


(h) Cluster 8

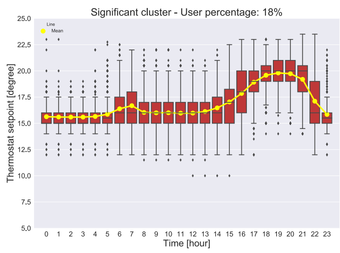
Figure B.7: Cluster method: Ward, Cluster size 15

C
Cluster size

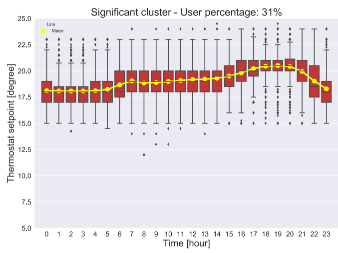
In this appendix, the clustering output of the different clusters sizes is shown for the chosen ward clustering method. As described in chapter 3 the clusters are compared for the sizes 4,6,10 and 14. Both the significant and non-significant clusters outputs are shown in this appendix.



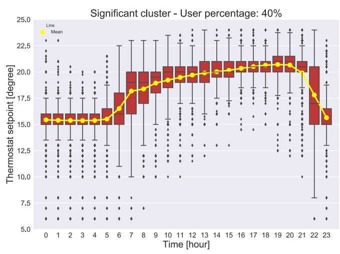
(a) Cluster 1



(b) Cluster 2

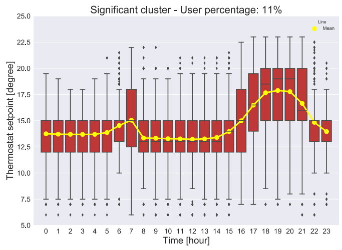


(c) Cluster 3

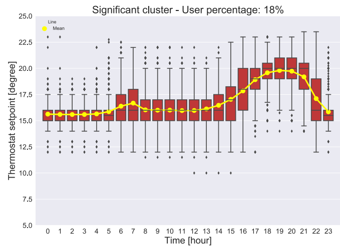


(d) Cluster 4

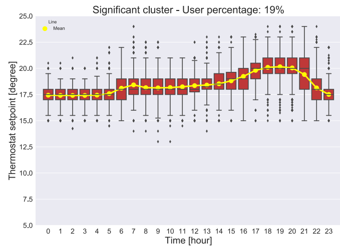
Figure C.1: Significant Clusters size 4



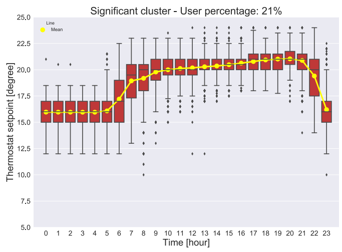
(a) Cluster 1



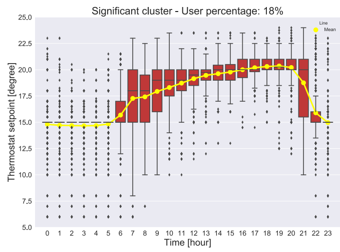
(b) Cluster 2



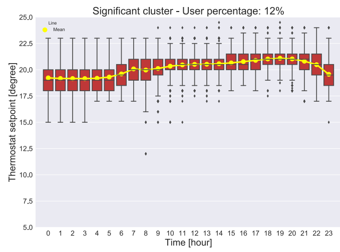
(c) Cluster 3



(d) Cluster 4

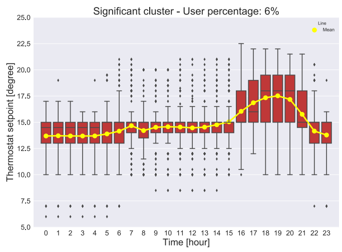


(e) Cluster 5

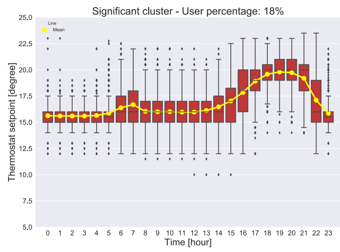


(f) Cluster 6

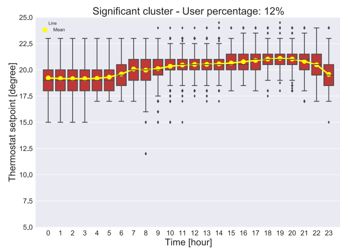
Figure C.2: Significant Clusters size 6



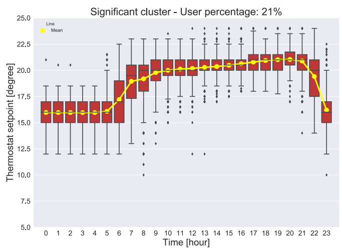
(a) Cluster 1



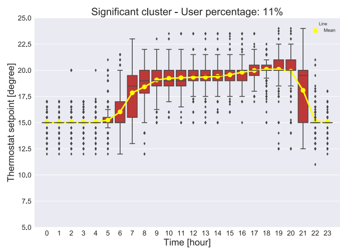
(b) Cluster 2



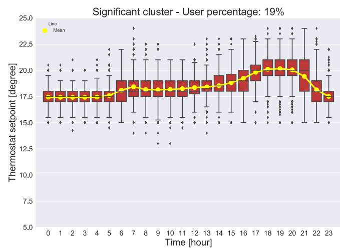
(c) Cluster 3



(d) Cluster 4

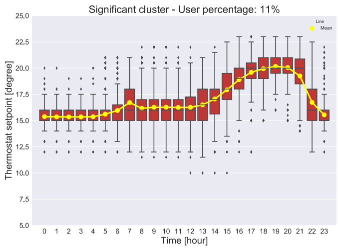


(e) Cluster 5

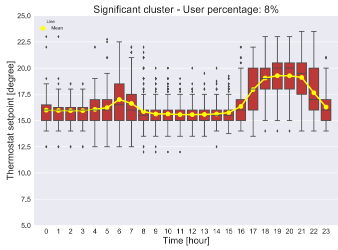


(f) Cluster 6

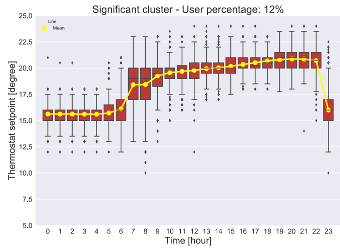
Figure C.3: Significant Clusters size 10



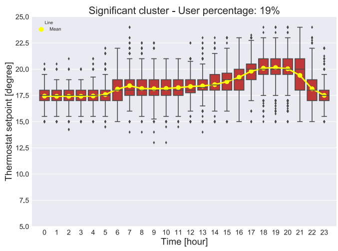
(a) Cluster 1



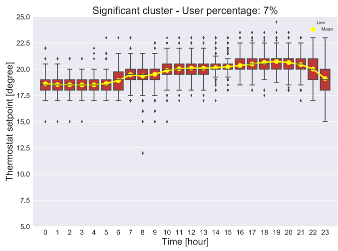
(b) Cluster 2



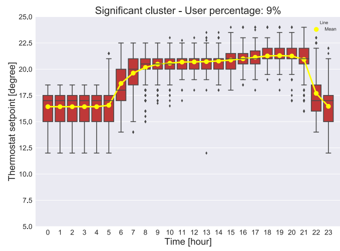
(c) Cluster 3



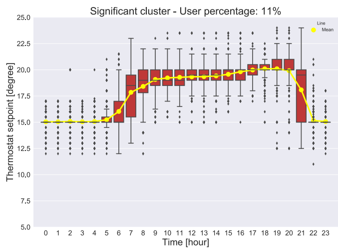
(d) Cluster 4



(e) Cluster 5

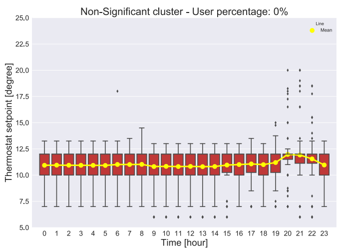


(f) Cluster 6

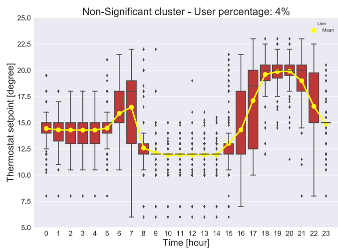


(g) Cluster 7

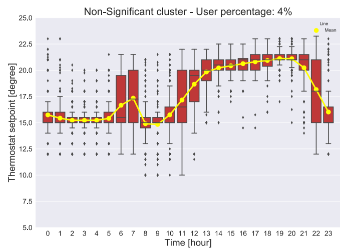
Figure C.4: Significant Clusters size 14



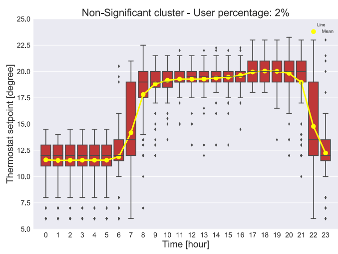
(a) Cluster 1



(b) Cluster 2

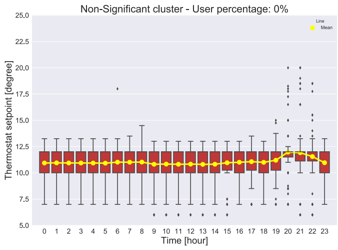


(c) Cluster 3

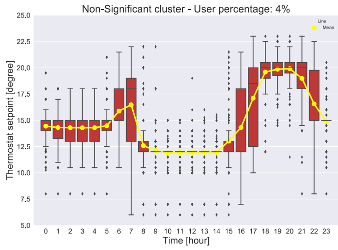


(d) Cluster 4

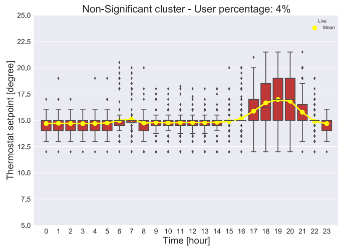
Figure C.5: Non-significant Clusters size 10



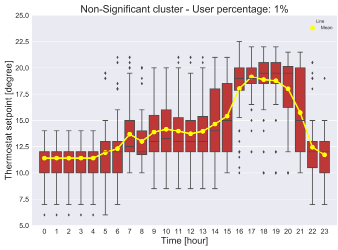
(a) Cluster 1



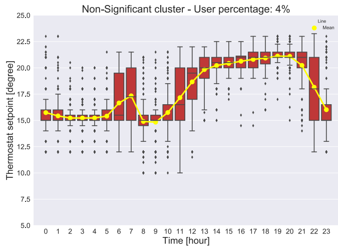
(b) Cluster 2



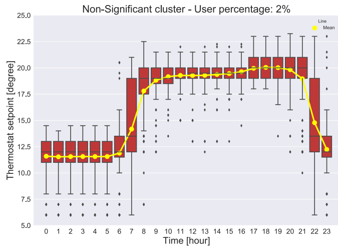
(c) Cluster 3



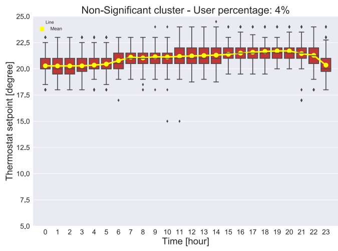
(d) Cluster 4



(e) Cluster 5



(f) Cluster 5

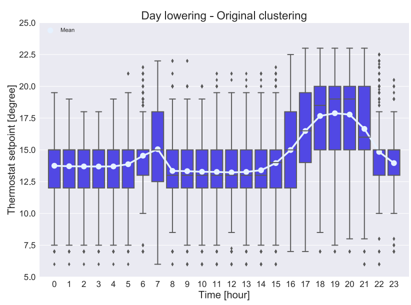


(g) Cluster 5

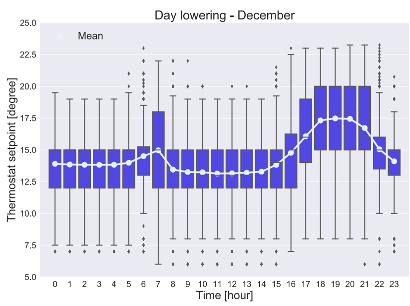
Figure C.6: Non-significant Clusters size 14

D
Cluster validation

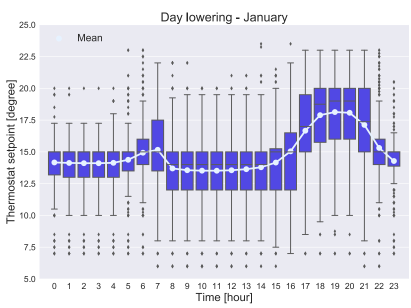
In this appendix, the clustering results of the separated data set into 3 different months and original clustering is presented. The performed clustering is in line with the clustering performed in section 3.3. The cluster outcomes are grouped on the basis of visual similarity.



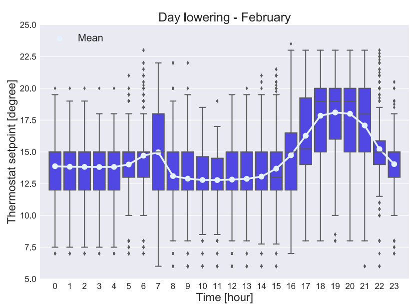
(a) Day lowering - original clustering



(b) Day lowering - December clustering

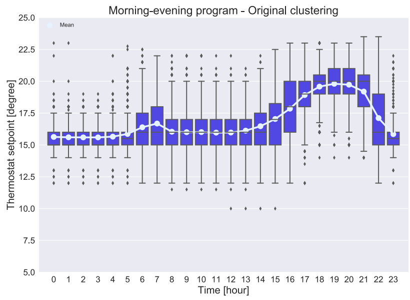


(c) Day lowering - January clustering

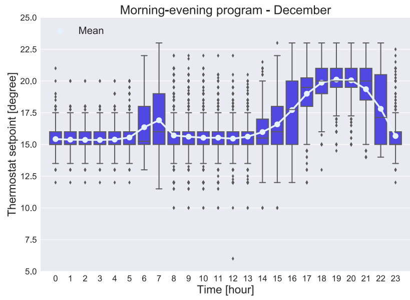


(d) Day lowering - February clustering

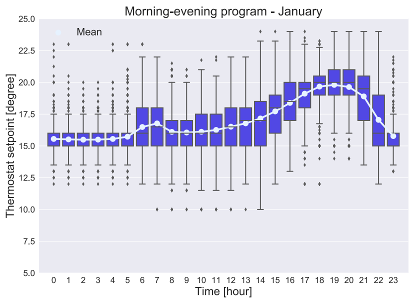
Figure D.1: Visual validation of clustering - Day lowering



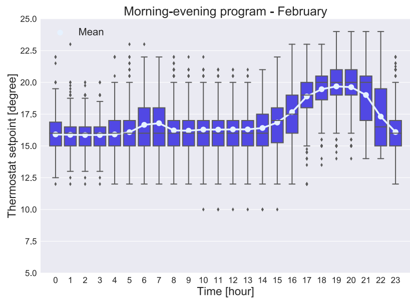
(a) Morning-evening program - original clustering



(b) Morning-evening program - December clustering

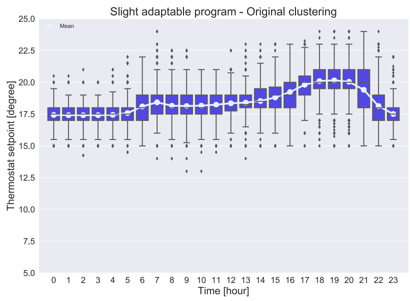


(c) Morning-evening program - January clustering

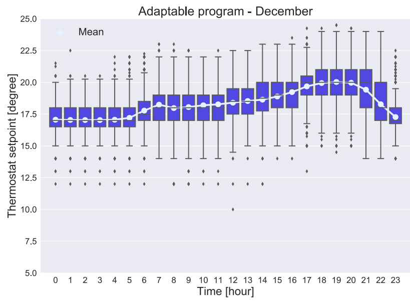


(d) Morning-evening program - February clustering

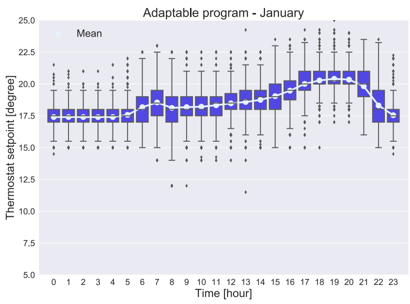
Figure D.2: Visual validation of clustering - Morning-evening program



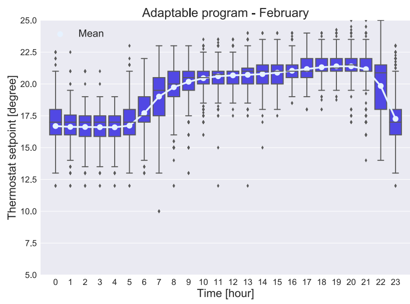
(a) Adaptable program - original clustering



(b) Adaptable program - December clustering

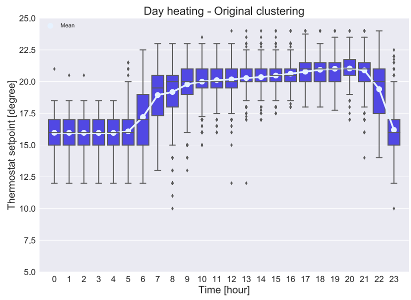


(c) Adaptable program - January clustering

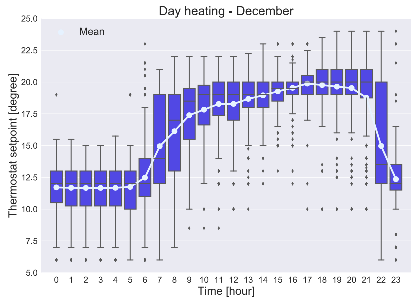


(d) Adaptable program - February clustering

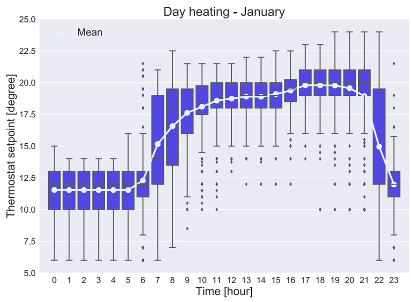
Figure D.3: Visual validation of clustering - Adaptable program



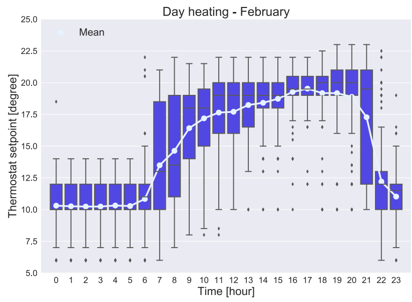
(a) Day heating - original clustering



(b) Day heating - December clustering

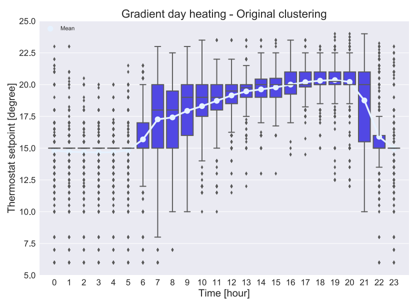


(c) Day heating - January clustering

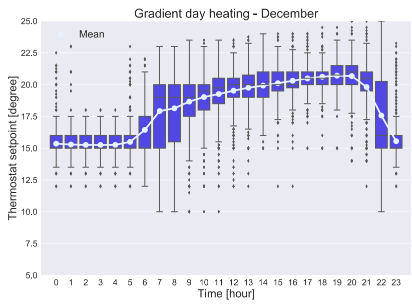


(d) Day heating - February clustering

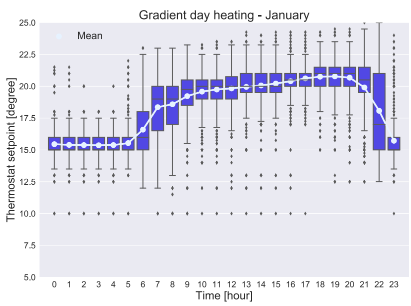
Figure D.4: Visual validation of clustering - Day heating



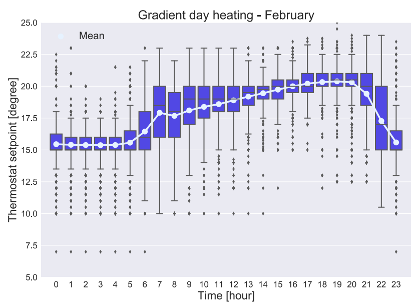
(a) Gradient day heating - original clustering



(b) Gradient day heating - December clustering

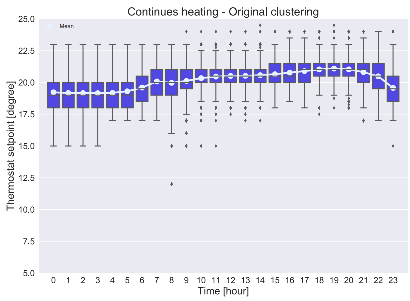


(c) Gradient day heating - January clustering

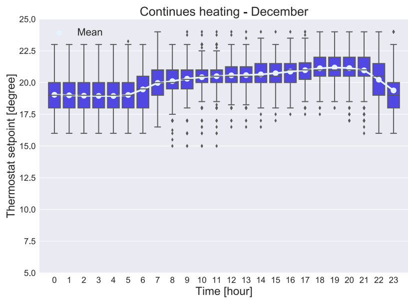


(d) Gradient day heating - February clustering

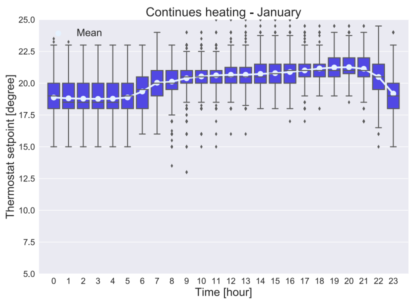
Figure D.5: Visual validation of clustering - Gradient day heating



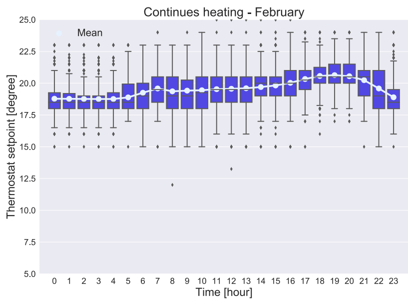
(a) Continues heating - original clustering



(b) Continues heating - December clustering



(c) Continues heating - January clustering



(d) Continues heating - February clustering

Figure D.6: Visual validation of clustering - Continues heating

E

Potential saving of dwelling characteristics

In this appendix, the potential saving for each of the saving options is calculated for separate household and dwelling characteristics. The size of the household accounts for the household characteristic and the type of the dwelling and size account for the dwelling characteristics. The potential saving for different characteristics generate insight in the impact of the different characteristics on the potential saving.

Table E.1: Potential saving household size

Household size	*Active lowering*	*Absent lowering*	*Before sleep lowering*	*Overnight lowering*
1 person	0.60%	1.82%	0.06%	2.36%
2 persons	0.83%	1.23%	0.05%	1.96%
3 persons	0.74%	1.34%	0.06%	2.15%
4 persons	0.88%	1.21%	0.06%	1.92%
5 or more persons	0.91%	1.13%	0.06%	1.95%

Table E.2: Potential saving dwelling type

Dwelling type	*Active lowering*	*Absent lowering*	*Before sleep lowering*	*Overnight lowering*
Apartment	0.86%	0.87%	0.05%	1.79%
Detached	0.96%	1.16%	0.06%	2.17%
Semi-Detached	0.67%	1.47%	0.06%	2.15%
Terraced	0.85%	1.26%	0.06%	1.96%

Table E.3: Potential saving dwelling size

Dwelling size m^2	*Active lowering*	*Absent lowering*	*Before sleep lowering*	*Overnight lowering*
15-50	1.25%	1.49%	0.07%	2.93%
50-75	0.74%	1.31%	0.06%	2.03%
75-100	1.14%	1.15%	0.06%	2.26%
100-125	0.80%	1.36%	0.06%	2.02%
125-150	0.77%	1.14%	0.05%	1.88%
150-200	0.79%	1.21%	0.05%	1.90%
200-300	0.90%	1.08%	0.05%	1.68%
300-500	1.07%	0.61%	0.04%	1.54%
500 or more	1.22%	0.36%	0.06%	2.29%