4 Statistical model of the heating prediction gap in Dutch dwellings: Relative importance of building, household and behavioural characteristics

Explanatory notes

The unsatisfactory results of the first regression analysis based on socioeconomic data led to a survey carried out in this Chapter 4 of the thesis. The survey was conducted on a subset of Amsterdam dwellings that had an official energy label, which provided a deeper understanding of the performance gap, since in addition to the more extensive household and economic profile of each household that was presented in Chapter 3, occupant behaviour was also included. Upon evaluating descriptive results of several statistical tests, several regression analyses were performed on different subsamples. Aside from the in depth analyses of the causes for the discrepancies, this chapter also demonstrates a possible solution for better predictions of consumption in the future.

Published as: Majcen, D., Itard, L., Visscher, H., 2015. Statistical model of the heating prediction gap in Dutch dwellings: Relative importance of building, household and behavioural characteristics, Energy and Buildings 105, October 2015

Abstract

The European Performance of Buildings Directive (EPBD) set the regulatory framework for a cost-effective improvement of the existing dwellings in 2002. The transformation of the stock towards higher efficiency is expected to be stimulated by labelling of the dwellings. The certificate itself is required to contain a list of potential cost-effective measures for the dwellings' thermal retrofit. However, the theoretical heating consumption provided in the certificate is not a good baseline for the calculation of cost effectiveness, as it is based on normalised dwelling conditions. Normalised conditions include a constant occupancy, constant indoor temperature and normalisations of other parameters, which in reality differ in different types of dwellings. The discrepancies between the normalised theoretical and actual heating consumption are also referred to as the performance gap. In this paper, we examined

these discrepancies using the example of The Netherlands. Using descriptive statistics and multiple regression, we investigated several parameters thought to have a different effect on actual and theoretical heating energy use – dwelling, household, occupant behaviour, as well as comfort – in order to propose improvements to the current theoretical consumption calculation. Aside from analysing the total sample, the data is regarded separately for overpredicted and underpredicted consumption records.

§ 4.1 Introduction

Dwellings represent a great potential for future energy savings. Several policy measures have been undertaken in the EU and nationally to encourage the transformation of the dwelling stock towards lower energy consumption. The European Performance of Buildings Directive (EPBD) has set the guidelines for dwelling performance certification, called the energy label, since 2002 and label certificates in The Netherlands have been issued since 2007. The Dutch energy label assesses dwellings' energy performance based on a steady-state energy model (detailed methodology is described in Majcen et al., 2013b), resulting in an energy label that ranges from A (good thermal performance) to G (poor thermal performance). Dwelling owners are required to possess a label at the moment of sale or rent, although non-compliance is currently still not sanctioned. Still, the number of performance certificates in The Netherlands reached 2,5 million by April 2014 (Compendium voor de Leefomgeving website, 2014), slightly over a third of the dwelling stock.

The target for dwelling stocks energy savings in the Netherlands is 110PJ by 2020 (Koepelconvenant energiebesparing gebouwde omgeving, 2012), using 617PJ as a baseline for the year 2008. This target covers residential and non-residential dwellings as well as existing and new construction. However, preceding this target, The Dutch federation of housing associations (Aedes) committed itself in the 'Covenant Energy Savings Housing Associations Sector' (Convenant Energiebesparing Corporatiesector, 2008) to achieve a 24 PJ reduction of the consumption of natural gas in the existing social housing stock (represented by roughly a third of the country's stock) between 2008 and 2018. Under the 'More with Less' (Meer met Minder (Convenant Energiebesparing bestaande gebouwen, 2008)) programme, the Dutch government and external stakeholders (corporations, real estate companies, and other stakeholders) have committed themselves to achieving a reduction of 30% of the energy consumption (100 PJ) of buildings by 2020. Comparing these two targets with the 90PJ target from 2012, which contains the residential as well as the non-residential sector, reveals that the ambitions have dropped significantly in the past. The new target is finally based on actual consumption data, which is important, since numerous research projects in the recent past highlighted the fact that the actual energy use in individual dwellings deviates from the predicted consumption. In poor performing dwellings, the heating energy use is overestimated (Sharpe and Shearer, 2013; Majcen et al., 2013a) and in well-performing dwellings, the trend is the opposite (Laurent et al., 2013, Majcen et al., 2013a), therefore using theoretical data as baseline which compromises the effectiveness of policy measures (Majcen et al., 2013a).

The phenomenon of discrepancies also called the performance gap (de Wilde, 2014), is shown on the example of Netherlands in Figure 1. This discrepancy is of crucial importance for the success of EPBD in the long run, since the directive states (Article 1 of EPBD) that it promotes the improvement of the energy performance of buildings within the Union, taking into account cost-effectiveness and to successfully estimate the cost effectiveness one needs to be certain of the baseline consumption. This study as well as in Figure 1 analyses the heating component of the total primary energy consumption, which is the basis for the label certificate. The average total primary energy consumed in each label category, is available in Majcen et al., 2013a).

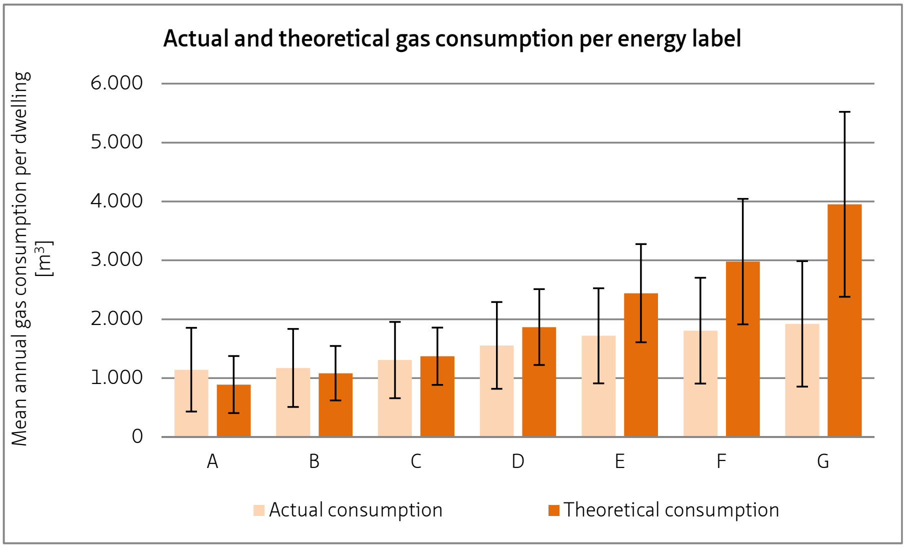

FIGURE 1 Actual and theoretical gas consumption in dwellings across label categories with ± 1 standard deviation (Majcen et al., 2013a).
Note that the two bars differ from each other in each category, this difference is in this paper referred to as the DBTA (difference between theoretical and actual gas use).

§ 4.1.1 Theoretical vs. actual gas and primary energy use

The discrepancy between theoretical and actual heating consumption observed in Figure 1 has already been studied extensively all over Europe (Laurent et al., 2013) as well as in the Netherlands (Santin and Itard 2012, Majcen et al. 2013a, Majcen et al., 2013b, Tigchelaar, 2011). However, the label certificate in the Netherlands does not specify heating energy use, but rather gas (in m^3), electricity (in kWh), and total primary energy (in MJ). Gas use in the Netherlands corresponds almost entirely to heating (space and water) and is also the scope of this paper. In The Netherlands, dwellings are predominantly heated with gas and heating is necessary for roughly 200 days in the year, and since there is rarely any cooling demand (nor are the majority of dwellings equipped with air conditioning), heating represents the majority of the dwellings' energy use. A small fraction of dwellings is heated by electricity, but in our sample they were excluded. From the data used, one could not distinguish gas for cooking from gas for heating; therefore it was included in the analysis. However, cooking represents a small fraction, less than 5% on household level, and is constant regardless of dwellings performance. Therefore it does not skew the analysis.

It is important to note that If we correlate theoretical gas consumption with actual, we do get a significant result (albeit correlation is weaker in reality than one might expect). In other words, dwellings with a more efficient label do have significantly lower actual gas consumption (Figure 3). In that sense, the label correctly predicts dwellings' thermal performance. To illustrate, Guerra Santin (2010) found the Pearson's correlation between actual and theoretical energy use for space heating within a sample of 185 dwellings to be 0,391 and the correlation in the two samples studied in this paper was 0,532 (N=4106) and 0,320 (N=468) respectively. However, at the same time, neither the 185-dwelling sample of Guerra Santin (2010) nor a larger sample from the same study of 563 dwellings demonstrated a correlation between the theoretical and actual total primary energy consumption, meaning that better performing dwellings do not necessarily have lower total primary energy consumption. This is logical because the actual total primary energy use includes the total electricity use of the dwellings (including all household appliances) while the theoretical primary energy use includes only the electricity use relating to the building (lighting, pumps, & ventilators but no household appliances). It was also shown that electricity use remains rather constant regardless of the label class (Figure 12 in Majcen 2013a), which decreases the correlation strength. To prevent that, the present paper focuses on gas consumption only.

§ 4.1.2 What causes the discrepancies?

The differences between theoretical and actual gas consumption (DBTA) are thought to arise from a multitude of factors. Theoretical gas consumption is based on normalized conditions such as indoor temperature of 18 degrees and 2620 degree days, heating of the entire floor area, a standardised number of occupants (which is a function of the floor area), infiltration rate assumed on the basis of the characteristics of the construction elements (for example length of window frames), etc. (Tables 7 and 4 in Majcen, 2013b). The way that occupants use the building in reality probably differs from these assumptions. According to several authors (Gill et al., 2010, Guerra Santin, 2010, Haas et al., 1998), occupant behaviour and lifestyle is thought to be a key factor in the discrepancy between theoretical and actual heating energy use and is correlated to energy performance itself. To elaborate, it is believed that in poor performing dwellings, the occupants are encouraged to conserve by the intrinsic poor performance of the dwelling itself (for example – never heat unoccupied bedrooms), while the situation in well-performing dwellings is opposite since a small increase in overall indoor temperature causes only a small change in the total energy bill. Sometimes the physical properties of the dwelling cause a certain type of behaviour; for example, occupants in dwellings with floor heating often do not have a choice but to condition the entire floor area, a practice opposite to the one in many poor performing dwellings with a sole heating element in the living room. Since the theoretic calculation normalises many parameters that inherently differ in dwellings' with different performance, a mismatch appears. The fact that behaviour and dwellings are so intertwined makes the causality analysis of the difference between theoretical and actual gas consumption (DBTA) very challenging.

Looking at different performance classes, the DBTA seems to be positive in poor performing dwellings (later on referred to as overpredition), meaning that theoretical gas use is higher than actual. In the most extreme cases the theoretical gas use can be as high as double of the actual consumption. This phenomenon seems to arise from the fact that poor performing dwellings are in fact under heated. On the other hand, underpredictions are characterised by an actual consumption higher than the theoretical, which occurs in well performing dwellings. In literature the expression 'rebound effect' is also used (Sunikka-Blank and Galvin, 2012), meaning that the consumption of energy increases when applying a saving measure. In the same paper, the overprediction of theoretical heating energy consumption is referred to as the pre-bound effect.

§ 4.2 Research objective

§ 4.2.1 State of the art

Many studies address the correlations between actual energy use and potential influencing factors (Wei et al., 2010). Among those, one can find dwelling-related factors such as type of the dwelling or its age, but also a multitude of occupant- and behaviour-related factors. In this paper, we distinguish four groups of influencing factors: dwelling, household, occupant behavioural characteristics, and comfort. The first three are generally thought to be the cause of the discrepancy seen in Figure 1, whereas the last one is actually a performance indicator, which is neglected most of the time.

Regarding the dwelling characteristics, Linden et al. (2006) found that occupants in detached houses adopt a lower set point temperature than those in apartments. Hunt and Gidman (1982), Santin et al. (2009) and French et al. all found a negative correlation between dwelling age and set point temperature. Furthermore, dwellings with a programmable thermostat seem to be correlated with a higher heating demand than those without (de Groot et al., 2008) and Santin et al. (2010). Also the relation between aspects of building quality and indoor temperature has been previously quantified in the papers from Haas et al. (2010) as well as Shipworth et al. (2009) and Raynaud (2014), all of whom found that more insulated dwellings have a higher indoor temperature. Raynaud (2014) also found that the difference between theoretical and actual consumption strongly depend on the theoretical thermal characteristics of the building itself and little on the theoretical performance (efficiency) of heating energy systems. Another important factor was whether the heating system was centrally controlled and the surface area of the dwelling.

Furthermore, studies also explore a multitude of household related characteristics that could influence actual energy use, such as number of occupants, which tend to be correlated with a higher energy consumption (Sardianou, 2008 and Oreszczyn et al., 2006). In this paper, household characteristics relate to occupants' demographic properties (age, household type, etc.) while occupant behaviour signifies occupants' lifestyle practices and their habits. Apart from the direct influence of the household feature on heating practices, it might also be that dwellings in different performance classes host certain characteristic households (for example, lower income occupants in dwellings with a poorer performance), which would in turn also cause a difference in energy use. Past studies have also shown that older occupants prefer a higher indoor temperature and that people with lower income tend to have a lower indoor temperature (Guerra Santin, 2010).

Though difficult to describe statistically, occupant behaviour seems to be one of the reasons for actual energy use not coinciding with theoretical. Under the term behaviour, we understand factors such as: presence at home, setpoint temperature, ventilation practices, number of showers number of heated bedrooms, heating of halls etc. Gill et al. (2010) showed that a composite variable describing efficient vs. inefficient behaviour would account for more than half (51%) of the variation in heating energy use. Occupant behaviour is also strongly dependent of the characteristics of the dwelling and at the same time clearly has a significant impact on dwellings actual performance. Behavioural practices are also expected to cross correlate with a multitude of characteristics of the household (their age, income, type of employment, etc.). Also in a bottom-up study, Haldi and Robinson (2011) showed that explicit consideration of occupants behaviour enables a more accurate prediction of energy demand. They also concluded that behaviour accounts for a greater variability in heating demand than building characteristics.

Last but not least, dwelling energy performance also relates to occupants 'comfort –the better the performance, the higher the comfort (Hong et al., 2009). On the other hand, it was previously shown in a sensitivity analysis of a dynamic simulation of a dwelling's energy use (Ioannou, 2015) that even occupants in very well performing dwellings are not comfortable during the heating season at a temperature of 20^{o}C. The author therefore questions the validity of PMV as an index for comfort measure. However, as formulated by Mishra et al. (2013), conditioned spaces (these are generally well performing) have narrower comfort zones compared to naturally ventilated buildings (generally poorer performing). To explore these phenomena, some comfort variables were included in the analysis in this paper.

§ 4.2.2 Motivation and goal

The fact that the relationship actual-theoretical heating energy use remains of middle size and not larger is related to the discrepancies we find between actual and theoretical consumption on a categorical level (between label classes). Even though it is clearly unrealistic to expect a correlation of 1, which would mean a perfect linear relationship on the level of individual dwellings, the correlation should be strong enough to ensure an accurate prediction within a certain label category on average, which is currently not the case. Without this, it is deceiving to portray the theoretical heating consumption of each individual dwelling on the label certificate. Policy implications of the poor correlations can be found in Majcen et al. (2013a) and Tigchelaar et al. (2011). It has been proven that without a more accurate determination of theoretical use prior to renovation, a better estimation of consumption after the renovation is not possible (Raynaud, 2014). Existing performance certificates are designed to be used solely to compare dwellings performance with other labelled dwellings and therefore policy makers, investors, researchers, homeowners, and other parties for whom payback time of a measure is relevant should understand that for any kind of future projections actual consumption has to be considered instead of theoretical consumption. To name an example, the European commission claims that old buildings consume 5 to 7 times the amount of heating energy of new buildings and that the saving potential of buildings is 5% of total European energy consumption (DG Energy website, 2015). Looking at Figure 1, the statement might be true looking at theoretical gas consumption as baseline, but far from it if we look at actual gas use in Dutch houses. Since acquiring actual energy data is costly, difficult (privacy laws), and sometimes even impossible (in case we want to renovate an existing building and accurately predict the savings), one should be able to model the consumption better. With dynamic modelling of individual dwellings and the occupants, one can estimate the consumption much more accurately. However, this is complex, expensive, and does not work on a dwelling stock level. This paper tries to understand what influences actual energy consumption and to what extent, so that in the future, more accurate projections can be made. To find this out, we use label certificate data coupled with actual energy data.

Therefore, this paper has a twofold objective: to offer insight into the relation between dwelling energy performance and dwelling, household, behavioural, and comfort characteristics and to study how different dwelling, household, behavioural, and comfort characteristics relate to the actual energy consumption. Last but not least, analysis of these two points enables us to propose a way of improving the current theoretical gas consumption towards a better fit with the actual gas use.

§ 4.2.3 Research design

§ 4.2.3.1 Correlations

Based on previously conducted studies, we expected to discover certain patterns between the four parameters observed in this study (Figure 2). In the first part of this paper, we looked for correlations between several parameters. The factors investigated in this paper are summarized in Table 1 in four groups and the nature of the correlations is shown in Figure 2, where the thickness of arrows in Figure 2 demonstrates the expected effect size. The hypotheses about the correlations are presented below.

TYPE OF DWELLING	EXPLANATION
Terraced house – corner	The last house in a row of houses. Can also be a semi-detached house.
Terraced house – middle of terrace	A terraced house surrounded by another house on its left and right.
Flat – middle – roof	A flat surrounded by two other flats on its left, right and underneath side, with a roof exposed to the air.
Flat – corner – roof	A flat, surrounded by two other flats underneath and on one of the sides, with an external wall and a roof exposed to the air (corner of the building).
Flat – middle – middle floor	A flat, surrounded by other flats above, below and on both sides.
Flat – corner – middle floor	A flat, surrounded by two other flats above, below and on one side, with an external wall on the other side (corner of the building).
Flat – middle – ground floor	A ground-floor flat, surrounded by other flats above and on both sides.
Flat – corner – ground floor	A ground-floor flat, surrounded by two others above and on one side, with an external wall on the other side.
Detached house	A detached house.
Dwelling characteristics	Label class (cat.), dwelling type (cat.), heating type (cat.), ventilation type (cat.), electrical boiler presence (cat.), heating of the hall yes/no (cat.), programmable thermostat presence (cat.), floor area (cont.), number of rooms (cont.), age of the building (cont.)
Household characteristics	Ownership type (cat.), household composition (cat.), education (cat.), ability to pay the energy bills (cat.), age of respondent (cont.), spendable income (cont.), number of occupants (cont.)
Occupant behaviour	Perception dwellings/households energy performance (cat.), awareness of the label certificate (cat.), ventilation practices - living room/kitchen/bathroom/bedrooms (cat.), ventilation habits weekends (cat.), perceived household energy behaviour (cat.), presence of water saving shower head (cat.), not setting thermostat too high (cat.), not ventilating while heating (cat.), no energy saving measures taken (cat.), number of weekdays of presence – morning/midday/evening/night separately (cont.), average temperature during the day - day/evening/night/nobody at home separately (cont.), showers per week (cont.)
Comfort	Perception of heat/cold, dry/humid and draft separately (cat.), unpleasant long waiting time for hot water (cat.)

* 'cat.' means a variable was categorical and 'cont.' that it was continuous

TABLE 1 Parameters investigated in this paper

1 In the category of dwelling characteristics, one expects to find a strong correlation with the theoretical gas consumption, but the correlation with actual consumption will probably be much weaker. This is because theoretical gas use depends mostly on dwelling characteristics (and a little bit on normalised household characteristics), other groups of parameters can of course also turn out to have an effect but it will be an indirect one.
2 Household characteristics will, on the other hand, have a large effect on actual gas consumption, but a much smaller one on theoretical gas consumption, since the theoretical calculation assumes standardised behaviour. However, just like in the previous category, it might be that household characteristics are different in different label categories and that's why a correlation could be detected with theoretical gas use.
3 Regarding occupant behaviour, theoretical gas consumption is based on a normalized occupancy and should therefore not correlate with these parameters; but again,

some effect will probably be found, since there is a correlation with actual gas consumption, which, as said previously, does correlate with the theoretical gas consumption. In theory, one can expect relatively strong correlations with actual gas use; however, one of the questions here remains how well we can actually capture the behaviour by using a survey.

4 The fourth parameter besides occupant behaviour, household, and dwelling characteristics is perceived comfort. In this paper, we look at comfort in a simplified way as an independent variable. It undoubtedly correlates also with other three groups of parameters, but apart from the cross correlation testing required for the regression analysis, these relationships were outside the scope of this paper. In Figure 2 it is depicted as an extension of gas consumption boxes, since our hypothesis was that this is in fact another output of the studied system. We believe comfort to be yet another performance indicator just like energy use. One can expect differently performing dwellings to have a different percentage of people dissatisfied with the temperature, humidity or air velocity conditions in the house. Comfort is likely to have a stronger correlation with theoretical gas use, since worse performing dwellings are probably less comfortable. Poor performing dwellings are often draughty, have non-centralised heating (only in the living room) and single glazing, whereas well performing dwellings are conditioned to a more constant temperature, giving occupants fewer reasons to feel uncomfortable. A smaller correlation might be found between comfort and actual gas use due to an indirect correlation with theoretical gas use. It could also be that households who consume little gas can in fact not afford more – such occupants would probably also feel uncomfortable.

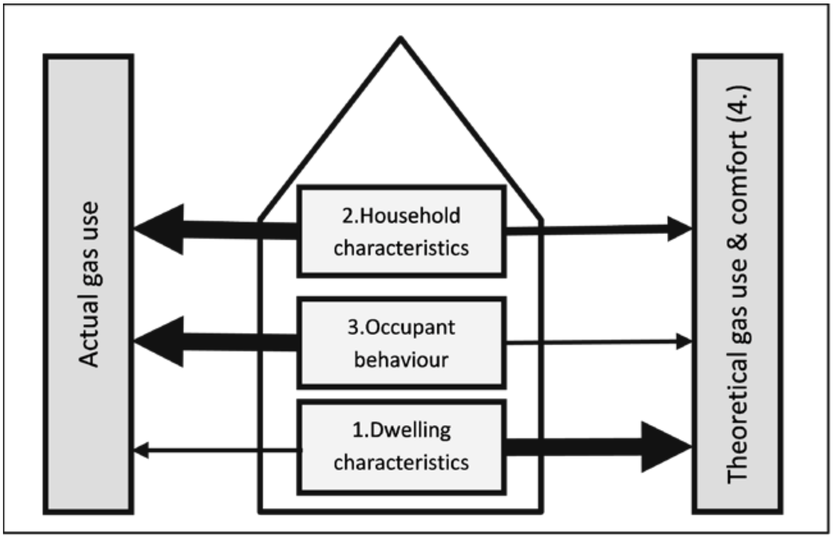

FIGURE 2 Effects of different parameter groups on actual and theoretical gas consumption.

§ 4.2.3.2 Regression analysis

After examining the correlations between all available variables belonging to any of the four mentioned groups, the results were revised. All variables that were significantly correlated to either actual or theoretical gas consumption were included in the regression analysis later on. Since as was said, some variables, such as occupant parameters have effect on actual as well as the theoretical gas consumption, and the objective of this paper was in fact to examine the causes for the discrepancy, we also look at correlations between variables and the difference between theoretical and actual gas consumption (further in this paper referred to as DBTA). It can be that a variable has an effect on actual gas consumption, but it is compensated for also in theoretical gas consumption and consequently there is no effect on DBTA. For example, dwelling type might have a significant impact on actual gas consumption but that can be true also for correlation with theoretical gas consumption and consequently there is no effect of dwelling type on DBTA. If the effect is not taken into account as strongly in theoretical as in actual gas consumption we can expect there will still be an effect of that variable on DBTA.

Regression was done on the dependent variables (actual and theoretical gas use, DBTA) in order to evaluate which of the variables is really causing a difference in consumption. For example, if both income and presence at home had a correlation with actual gas consumption, it could still be that this is due to a correlation between income and presence at home. Regression tells us which of the variables adds independent information about gas consumption in presence of other variables. Before the regression analysis multicollinearity was checked using a correlation matrix and no problematic (above 0.4) cross correlations were detected.

Additionally, we have observed the regression of DBTA separately for cases where theoretical gas use is overpredicted and where it is underpredicted. These two seem like two different phenomena; therefore these regressions might give different results than regression of the total sample. We thought about conducting regressions separately for dwellings in each label class, but there was not enough records to assure significant results and this was a good compromise.

§ 4.2.3.3 Improving the existing theoretical gas use

Last but not least, in this paper we tried to develop a new model for determining theoretical gas consumption based on the actual consumption data. In this section, we used actual gas use as dependent variable and theoretical gas use together with only dwelling characteristics as predictors. The rationale behind using only dwelling characteristics and not behavioural or comfort parameters is that it is the only information available when making the performance certificate and we do want to keep the theoretical consumption valid even if the occupancy changes. We believed that by using the actual data of a smaller sample, coefficients could be developed with which we could modify the current theoretical consumption of labelled dwellings (on a stock level) in order to get a better fit. Therefore, we modified the theoretical gas use of a larger sample (WOON sample see 4.3.1.2) based on the beta values obtained from the regression analysis in a smaller sample (Rekenkamer sample 4.3.1.1) and looked at how well the new value fits actual gas consumption.

§ 4.2.3.4 Boundaries

The two important factors that fall beyond of the scope of this study are the errors in the energy label certificates and uncertainties in actual consumption data quality. Regarding the first, it seems that many times the inspection is not carried out as accurately as it should be and the certificate doesn't correspond to the real state of the dwelling. A 2011 study has proved a rate of inaccuracy of 16,7% (Derde onderzoek naar de betrouwbaarheid van energielabels bij woningen , 2011) and in 2013 the inaccuracy was 21,2% (Herhalingsonderzoek betrouwbaarheid energielabels bij utiliteitsbouw, 2013), although the research in 2013 only looked at non-residential buildings. However, there was a trend of improvement in preceding years, so the certificate accuracy in the sample used should be sufficient as it is not substantially different from the accuracy in our former studies Nevertheless, one should note that certificates of poor performing dwellings carry a greater risk of uncertainty since determining their construction features is a more tedious and error prone process due to a lack of documentation and many of the characteristics are assumed on the basis of the construction year of the dwelling. On the other hand, newer dwellings are usually much easier to inspect as all the construction properties are well known.

The second important factor that is, to some extent, beyond the scope of this paper is the quality of energy data. The data originates from Statistics Netherlands, a governmental organisation that collects this data from energy companies. The companies report the billing data, which are calculated on the basis of meter readings. In some cases the occupants do not report the meter reading and in such instances, the consumption is based on the average consumption of dwellings in the region managed by one network management company, corrected for climatic variations (Informatiecode Elektriciteit en Gas, 2014). It has been said by government officials (Kamp, 2014) that the data is estimated in 10 to 20% of the cases annually for both gas and electricity. The mentioned code, however, obligates the network managing company to collect the meter readings by themselves at least once in 36 months, which ensures at least some basic actualisation of the data.

§ 4.3 Methodology

§ 4.3.1 Data

The paper is based on a dataset gathered for a study commissioned by the Rekenkamer Amsterdam, the audit office of Amsterdam municipality with the objective of evaluating the subsidies given to social housing corporations by the municipality in previous years. Since it was not possible to get reliable longitudinal data on the dwellings that were actually renovated, the study was based on analysing consumptions of dwellings in different label categories and comparing them among each other (Majcen and Itard 2014). This paper is based on the same dataset. However, to strengthen the findings of this study, cross checks were made using WOON 2012 dataset. Both Rekenkamer and WOON data are presented below.

§ 4.3.1.1 Rekenkamer dataset

The dataset initially contained 245.841 label certificates issued for the Amsterdam area since 2007. To avoid coupling the certificate data with an outdated energy consumption data (as mentioned before this is in some cases estimated), dwellings which have been renovated or had more than one certificate issued in the years 2010 – 2012 have been removed from the dataset, leaving 140.480 certificates. This was done using a dataset of all major dwelling renovations provided by the Rekenkamer Amsterdam. This deletion ensures that the coupling with actual gas use is done as correctly as possible (and we do not couple a renovated dwelling with a pre-renovation gas use). Statistics Netherlands could find a match for 116.744 addresses, the rest could not be linked due to either unknown address or missing data about actual energy use.

9.473 dwellings with heat supplied from outside (district heating), were left out due to the fact that their actual energy use is not individually metered. Furthermore, records in which actual electricity or gas data was missing or zero (10192 for electricity and 9047 for gas) were removed. Last but not least, records where dwelling type was an apartment building with not-independent units (student houses, retirement homes) were removed (32) leaving 87.946 dwellings. The sample at this point contained certificates dating from 2007 to 2012. However, it was discovered that the years 2007 – 2009 had many problems; theoretical gas and electricity were not reported separately and there seemed to be a misplaced decimal comma in all 2009 data. Due to these uncertainties a choice was made to only analyse dwellings from 2010 onwards (50.156). To avoid extreme outliers, apartments with a floor area above 1000m^2 were discarded leaving a final sample of 48.929 dwellings.

Parallel to certificate data which contains the theoretical energy use, coupled with actual energy from the statistics office, an occupant survey was carried out (the full survey is an annex of the report written by Broekhuizen and Jakobs, 2014). This was done on a much smaller sample of about 1000 dwellings, selected from the sample of 140.480 dwellings mentioned before. As a result, some of the survey results could not be coupled with the actual energy use and the sample turned out to be well below 1000 after it underwent the steps described in paragraph above. The survey was carried out per label category, gathering the same amount of dwellings in each of the 7 label categories. Although this means that the sample is not representative for label distribution, it is much easier to find significant correlations and predictors in regression analysis since it offers a high share of data also in extreme label categories, such as A and G.

The survey was short (12 minutes time to fill out the online version) but was designed in a way to capture information as condensed as possible. It included 42 questions about dwelling properties that are not present in the label certificate (number of rooms, type of occupancy, thermostat type, water saving shower head etc.), household properties (number, age of occupants, ability to pay energy bill), behaviour of occupants (presence at home, heating and ventilation practices, showering, energy efficient behaviours etc.) and comfort (temperature, air velocity, and humidity). Variables obtained from the survey are gathered in Table 2 and Table 3.

§ 4.3.1.2 WOON dataset

The Dutch Ministry of the Interior and Kingdom Relations carries out a study of energy performance of the Dutch dwelling stock (Woon Energy) every 5 to 6 years as a part of a larger survey of Dutch dwellings (Woon – Woon Onderzoek Nederland, which stands for Housing survey Netherlands). For the validation and comparison of the results obtained in the Rekenkamer survey, the Woon survey from 2012 was used, which was done on a sample of 4.800 representative Dutch dwellings. A general report using this data is publicly available (Tigchelaar and Leidelmeijer, 2013), however, the survey was much richer than described in the mentioned report and is of excellent quality to validate and provide depth to the Rekenkamer data. Variables obtained from the survey are gathered in Table 2 and Table 3.

§ 4.3.1.3 Actual energy data standardization

Both Rekenkamer and WOON datasets were coupled to standardise actual energy consumption data from the CBS. To enable a comparison between the Statistics Netherlands data and theoretical gas consumption data, a standardisation had to be applied. The Statistics Netherlands data corresponded to climatic year of 2012, which had 2878,8-degree days. The energy label calculation, on the other hand, assumes 2620-degree days (for method description see Majcen, 2013), therefore a correction factor of 2620/2878,8 had to be applied to the actual gas consumptions supplied by the CBS.

§ 4.3.2 Statistical analysis

The use of parametric vs. non-parametric tests remains controversial in statistics. The common procedure is to first assess normality of the data and carry out analysis using parametric tests if normality is met. Data analysis of the Rekenkamer sample showed that most continuous variables were not normally distributed. An attempt was made to transform them, but this yielded little success using the most common transformation functions such as log, ln, square, square root etc. After this step it was decided to rather avoid very tedious interpretation of complexly transformed variables so we did not proceed with transformations.

However, regarding the normality, significance can be detected easily in large samples (Lantz, 2013 and Lin, 2014) and also normality tests detect non-normality very easily in large samples. There is no easy answer as to where the cut-off between small and large sample lies, although N>30 is in most cases considered as 'large enough' to detect a normal distribution, but the cut-off for not finding a normal distribution due to large sample size is not known just as it is not known at what sample size parametric tests are usable. However, robustness of parametric tests increases with sample size and non-parametric tests are in general thought to be useful for smaller samples (Fagerland, 2012) where the probability distribution is not known or non-normal. In a previous study conducted for the Rekenkamer Amsterdam, in which the same data was used, we have used parametric tests considering all the mentioned arguments. However, although the sample size is relatively large, the data is non normal, which is why we have decided to use non-parametric tests for this study.

Therefore, Spearman's rho was used for establishing correlations between continuous variables (Table 2). Spearman's correlations revealed a lot of significant correlations between continuous variables and gas consumptions with more detectable correlations coming from the WOON dataset. This was to be expected due to the larger sample size.

However, the fact that most correlations found in the Rekenkamer data were present also in WOON data adds strength to our analysis.

Table 3 shows results of categorical and binary variables, where correlation coefficients could not have been computed. Instead, we observed whether or not the groups differ from each other significantly. Kruskal Wallis's non parametric test for independent measures was used for variables with more than two categories and Mann Whitney's U statistic was calculated for binary variables. Since the Kruskal Wallis's test only tells us whether or not there is a significant difference between at least two of the categories and not where the difference is, means with 95% confidence intervals are depicted in several plots in 4.4.1. Based on these graphics one can see which categories are significantly different from each other.

The general finding is that WOON data complies with the smaller Rekenkamer sample. Presumably due to a larger sample size WOON does demonstrates slightly more significant results than Rekenkamer dataset. Descriptive statistics for the variables can be found in Table 3 below and are depicting mean, standard deviation and also median, since the variables are not normally distributed. Table 2 and Table 3 are both divided into four sections, just like the following paragraphs of the paper, according to the groups of parameters as described in Figure 2.

§ 4.4 Results

§ 4.4.1 Single variable correlations

First of all, it is important how the new datasets relate to previously conducted research in The Netherlands. Theoretical and actual consumptions of all three datasets are therefore plotted in Figure 3 together with their corresponding 95% confidence interval. The confidence interval is the smallest in 2010 label dataset (studied in Majcen, 2013a and Majcen, 2013b b), since it contained the most records (ca. 200.000). It is also notable that this dataset had the highest actual energy consumption (dating to year 2009) in poor performing label categories. In newer datasets, WOON (from 2012, using energy data from 2010) and Rekenkamer (using energy data from 2012), where sample sizes were much smaller (4.800 and 460 respectively), despite the fact that equal degree day standardization was applied, the actual energy consumption is lower. This could be due to sample properties or due to the fact that degree days method does not account efficiently for annual variations, which is out of the scope of this paper.

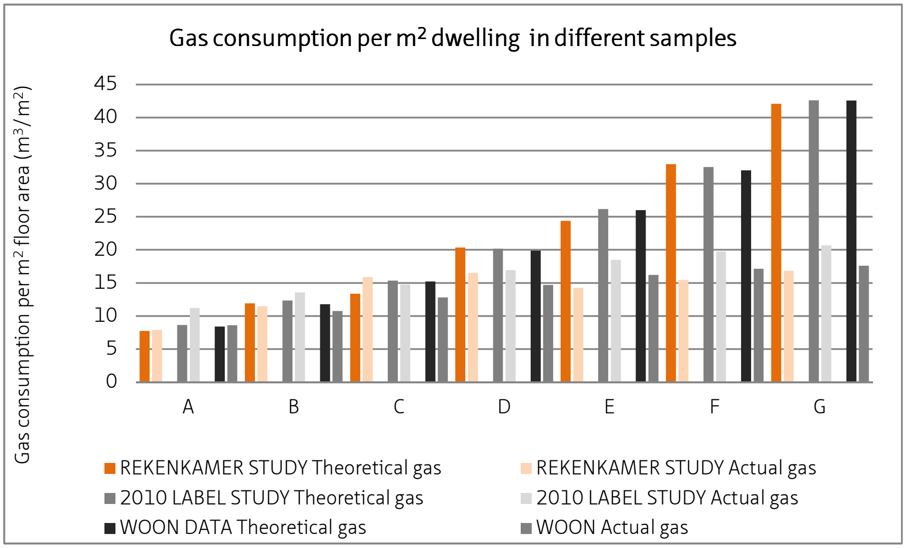

FIGURE 3 Average actual and theoretical gas consumption per m² dwelling including the 95% confidence interval.

Despite small differences, the phenomenon of over and underpredicted actual gas use remains the same in all three datasets, which makes the two selected samples appropriate for analysis.

In the following sections, data from Table 2 and Table 3 are described per group of parameters. Each group is separated further into continuous (Table 2) and categorical variables (Table 3). For categorical variables, we show some descriptive graphics with means and confidence intervals for better understanding; however, due to the amount of data, we only show the most interesting graphics. All means, medians, standard deviations, and sample sizes for WOON and Rekenkamer data, are nonetheless shown in Table 4.

§ 4.4.1.1 Dwelling characteristics

A Continuous variables

Woon data suggest that a larger number of rooms leads to a bigger discrepancies between actual and theoretical gas use; however, this was not confirmed using Rekenkamer data. This could be due to the fact that the Rekenkamer sample contains no dwellings with a number of rooms larger than eight and also fewer dwellings with six or seven rooms.

Both datasets show strong correlations of consumptions with building year. The older the building, the higher the actual and theoretical consumptions, where the theoretical consumptions correlate almost twice as strongly as the actual. Older dwellings also correlate with a larger DBTA (Table 2).

In the Rekenkamer sample, floor area remains a good predictor of actual gas use even though the consumptions are corrected for the dwellings floor area. It seems that even with the correction, larger dwellings consume less gas per m^2. WOON sample does not demonstrate this correlation, but there is a correlation in this sample between floor area and theoretical gas use/DBTA.

B Categorical variables

From Table 3 above one can see that label category has a significant correlation with all consumption variables, as illustrated also by Figure 3. However, the minimal but steady decrease of actual gas use per m^2 when improving the label category as seen in the WOON 2012 and energy label data in 2010 (Figure 3) is much less evident in the Rekenkamer sample. This could be related to poor representativeness of this sample for Dutch dwelling stock.

Type of ownership was not a significant variable in the Rekenkamer sample, as opposed to the WOON 2012 study. The Amsterdam sample was meant to represent mostly social housing and is therefore not representative for ownership type, since owner occupant dwellings are underrepresented. Dwelling, heating and ventilation categories are significantly different in their actual as well as theoretical consumption. In both samples, gallery apartments have the lowest theoretical and actual gas consumption and flats with a staircase entrance are significantly higher in both (Figure 4). Corner row houses are probably not a representative group in the Rekenkamer sample, since they are only 9 dwellings and their consumption deviates significantly from the consumption in WOON sample. Again, the Rekenkamer sample does not contain a representative population of dwelling types in the Netherlands due to the specific architecture of the city.

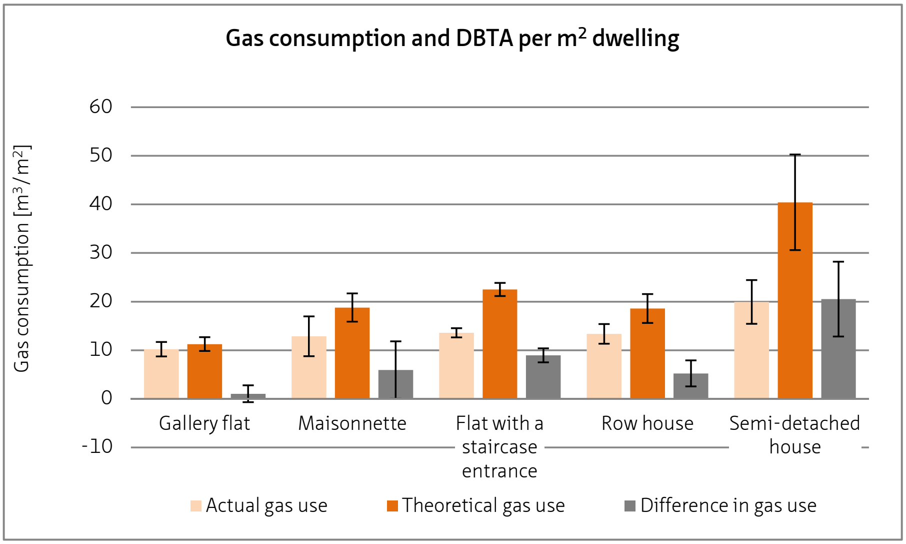

FIGURE 4 Actual consumption, theoretical consumption and DBTA per m² floor area of different dwelling types in the Rekenkamer sample.

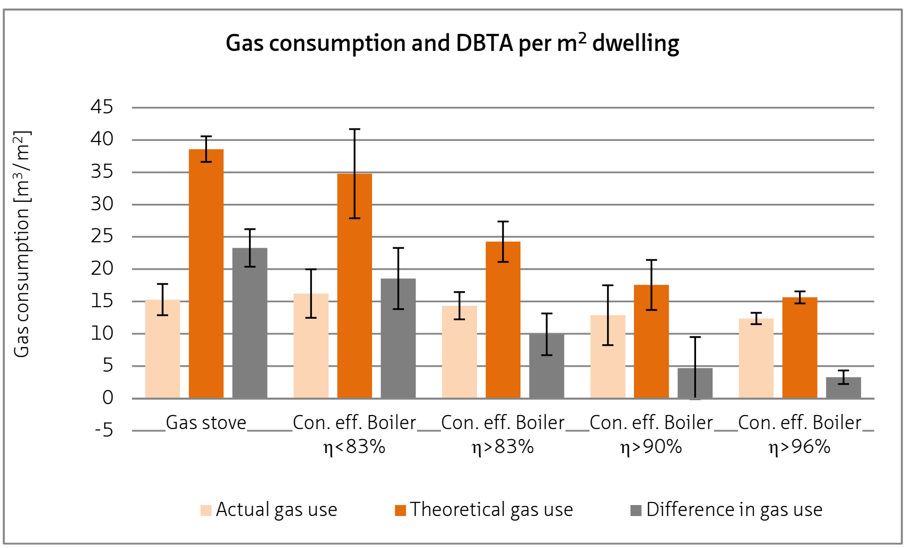

FIGURE 5 Actual consumption, theoretical consumption and DBTA per m² floor area of different installation types in the Rekenkamer sample

As one can see on Figure 4, dwelling type plays a role regarding the theoretical gas use and the DBTA. Gallery apartments seem to have a smaller DBTA than other types.

According to the Kruskall Wallis test, dwellings with more efficient installation systems have a lower theoretical and actual gas use with Figure 5 confirming the phenomenon. However, similarly to Figure 3, the differences in actual consumption between different systems are small – much lower than the theoretically anticipated. From the theoretical point of view there is a significant difference between lower efficiency boilers / boilers with η >0.93 /boilers with η >0.9. However, when looking at the actual consumption, the only significant difference is between very high efficiency (>0.96) and very low (gas stove). From this picture it is also very clear that—despite a 95% confidence interval overlap—the lower the theoretical efficiency the larger the DBTA which could mean that the efficiency of 'poor' heating systems is underestimated.

Similar to the above, dwellings with a mechanical ventilation fare better than the ones with natural ventilation in theoretical as well as actual gas use. The overprediction seems to be higher in dwellings with less efficient systems in general.

The presence of an electric boiler, programmable thermostat, and type of tap water heating also seems to affect theoretical gas consumption and consequently the difference. Dwellings with an electrical boiler or a programmable thermostat have a significantly lower theoretical gas consumption and DBTA than those without. When it comes to hot tap water installation, a gas boiler without hot water reserve has the lowest theoretical gas use followed by an electrical boiler and finally a boiler with hot water storage and the same goes for actual gas user and DBTA. Woon confirms these results although presence of a boiler was also significant with regard to actual gas use and not just theoretical consumption and DBTA as in the Rekenkamer sample. The significance was however, lower than significance for theoretical gas use and difference which is in compliance with the findings in Rekenkamer data.

§ 4.4.1.2 Household characteristics

A Continuous variables

A larger number of occupants correlates with higher actual gas use in the case of Rekenkamer data. This was not confirmed using WOON data, however, the difference and the theoretical gas use in WOON data did correlate with number of occupants and were smaller in dwellings with more occupants.

Older respondents are correlated with a higher actual gas use in WOON dataset. There is no significant correlation between these variables in the Rekenkamer data; however, there is a negative correlation between age and theoretical gas use and the difference.

Another interesting correlation which is present in both data's is the amount of spendable income and theoretical gas use; people with more money use less gas, probably because people with a higher income tend to occupy better performing dwellings. Furthermore, from WOON data it also seems that there is a smaller overprediction in households which are better off and lower actual gas use, which probably confirms the fact that richer people occupy better labelled dwellings.

B Categorical variables

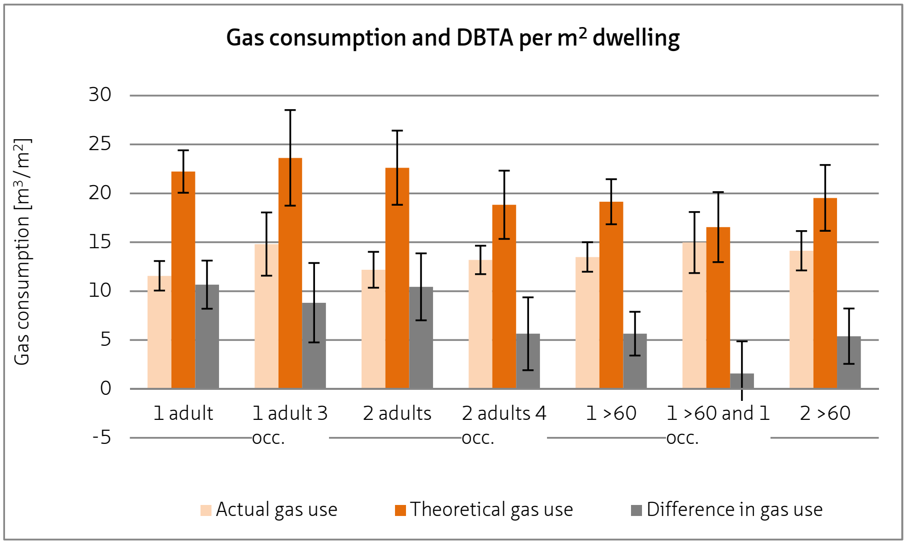

FIGURE 6 Actual consumption, theoretical consumption and DBTA per m² floor area of different household compositions
\>60 = occupant over 60 years of age

The three household-related variables—household composition, ability to pay energy bills, and education—also have a significant impact on actual gas consumption or on the difference between them. The findings are largely confirmed by the WOON sample, although there are more significant differences found in the theoretical gas use. Figure 6 shows that households with elderly persons do have a smaller DBTA than households where only adults or children are present. This has to do with lower theoretical gas use in these groups and also a higher actual use. The fact that elderly correlate with higher gas consumption means that they probably have higher comfort standards or/and maybe spend more time at home. We can also note that households with more members have a higher actual gas consumption. However, the variable household composition was tricky to recode. In the survey, ages of all occupants were collected. We then recoded these ages into 4 categories – elderly, adults (above 24), teenagers (above 16) and children. In the end, there were few dwellings with teenagers in the sample (15) and their presence did not make a significant difference, so they were considered in one category together with children. We also tried simplifying the categories into presence of children-elderly, but it did not yield more significant results so we stuck with the more detailed version.

The lower gas use of people who find it really easy to pay the bill might mean that they live in better performing houses.

§ 4.4.1.3 Occupant behaviour

A Continuous variables

Both datasets demonstrate a negative correlation between presence at home in several parts of the day and the difference. The more days people are present, the lower the overprediction. The size of the effect is larger in the Rekenkamer data then in WOON dataset.

In the average temperature setting, both datasets demonstrate a similarly sized correlation between higher temperature and smaller DBTA. Both datasets also demonstrate a positive correlation between actual gas use and higher temperature; however, only in WOON data is there also a negative correlation between theoretical gas use and temperature. Since the temperature assumption is the same in all dwellings when we look at theoretical gas use, the only possible explanation is that there is some other indirect correlation that relates to a higher temperature (for example the heated surface area).

The amount of showers taken in a week correlated positively with a higher actual gas consumption in both datasets, but only in WOON dataset there was also a correlation with theoretical and the difference between the consumptions.

B Categorical variables

Regarding occupant behaviour, few categorical variables were significant. As expected, occupants' perception of dwellings and households energy performance is a good predictor of dwellings actual and theoretical gas use.

Ventilation practices did not yield any significant results in the Rekenkamer data and but a few in the WOON dataset. Significant impact was recorded on gas use when examining presence of shower head, thermostat setting, ventilating while heating and implementing energy measures.

§ 4.4.1.4 Comfort perception

Regarding comfort, perception of temperature was related with differences in gas consumption in the Rekenkamer sample (Figure 7). Actual gas use as well as DBTA seemed to be lower in dwellings where occupants thought the temperature was satisfactory than in those where people were too cold. We suspected there could be a correlation between the setpoint temperature and the perception of cold, but the Spearman's test revealed no significant correlations. Unfortunately, there was no variable in WOON to compare this result to.

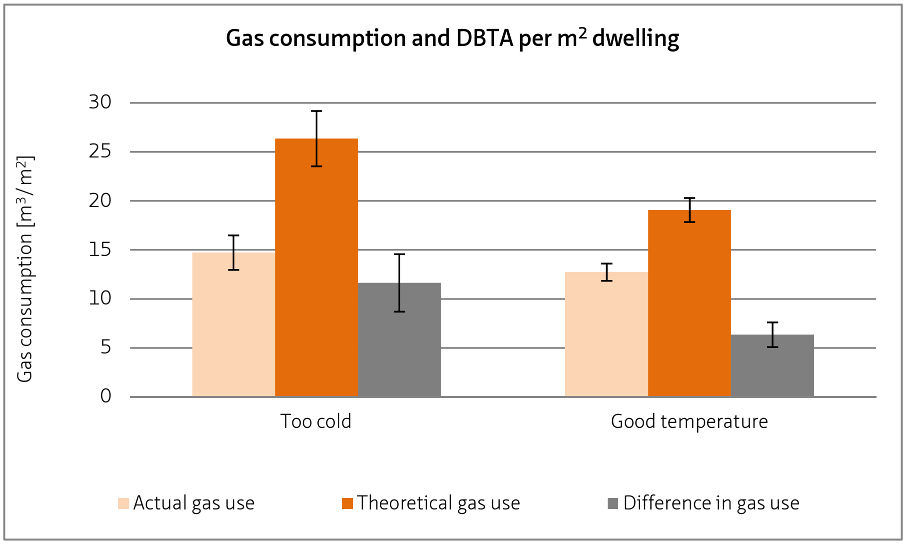

FIGURE 7 Actual consumption, theoretical consumption and DBTA per m^2 floor area in dwellings with difference temperature perceptions

		REKENKAMER DATASET - CORRELATION (N)			WOON DATASET - CORRELATION (N)		
		ACTUAL GAS USE PER M²	THEORETICAL GAS USE PER M²	DBTA	ACTUAL GAS USE PER M²	THEORETICAL GAS USE PER M²	DBTA
DWELLING CHARACTERISTICS	Floor area	-0,210 (460)	-0,407 (460)	-0,250 (460)	-0,235 (4110)	-0,227 (4262)	-0,069 (4106)
	Number of rooms						0,034 (4106)
	Age of the building	0,277 (460)	0,663 (460)	0,465 (460)	0,393 (4110)	0,779 (4262)	0,564 (4106)
HOUSEHOLD CHARACTERISTICS	Age of respondent		-0,164 (426)	-0,193 (426)	0,058 (4110)		
	Spendable income		-0,122 (304)		-0,088 (4110)	-0,151 (4262)	-0,089 (4106)
	Number of occupants	0,128 (434)				-0,106 (4262)	-0,098 (4106)
OCCUPANT BEHAVIOUR	Number of weekdays present – in the morning			-0,122 (460)			
	Number of weekdays present – during midday	0,170 (460)		-0,208 (460)		-0,031 (4262)	-0,044 (2126)
	Number of weekdays present – in the evening			-0,105 (460)		-0,062 (2209)	-0,047 (2126)
	Number of weekdays present – at night						
	Average reported temperature during the day	0,192 (415)		-0,193 (415)	0,125 (3838)	-0,099 (3971)	-0,205 (3834)
	Average reported temperature in the evening	0,171 (402)		-0,184 (402)	0,075 (3838)	-0,127 (3971)	-0,195 (3834)
	Average reported temperature at night	0,256 (402)		-0,166 (402)	0,067 (3838)	-0,096 (3971)	-0,148 (3834)
	Average reported temperature when nobody is at home	0,245 (398)	-0,104 (398)	-0,248 (402)	0,093 (3838)	-0,090 (3971)	-0,165 (3834)
	Showers per week	0,145 (314)			0,039 (4110)	-0,056 (4262)	-0,104 (4106)

*Highlighted fields are significant on a 95% confidence interval.

TABLE 2 Spearman correlation coefficients and number of cases in each group*

		Rekenkamer			Woon		
		Chi-square/Mann-Whitney U			Chi-square/Mann-Whitney U		
		Actual gas use per m^2	Theoretical gas use per m^2	DBTA	Actual gas use per m^2	Theoretical gas use per m^2	DBTA
Dwelling characteristics	Label class	51	388	260	3516	768	2160
	Dwelling type	22	81	43	142	324	137
	Heating type	14	180	137	86	531	377
	Electrical boiler presence	1	8	9	865712	914348	795248
	Heating of the hall	1485	1083	1508	184768	120571	116053
	Ventilation type	30	100	52	482	1730	814
	Tap water heating type	10	90	62	53	432	344
	Programmable thermostat presence	9771	7814	7653	1962208	1954475	1847913
	Ownership type	0	2	2	27	38	15
Household characteristics	Household composition	19	12	27	20	61	44
	Education	27	17	13	16	36	6
	Ability to pay the en. bills	13	4	2			
Occupant behaviour	Perception of dwellings/household energy performance	50	75	36	225	57	47
	Awareness of the label certificate	6	2	4			
	Ventilation practice in the living room	3	6	9	34	11	1
	Ventilation practice in the kitchen	7	13	7			
	Ventilation practice in the bathroom	8	14	12			
	Ventilation practice in the bedrooms	10	8	6	28	4	3
	Ventilation habits during weekends	5	2	8			
	Perception of household energy behaviour	20	6	5	377	293	50
	Presence of a water saving shower head	21620	19044	19312	21	47	13
	Not setting the thermostat too high	12198	11117	14381			
	Not ventilating while heating	19342	22916	20210			
	No energy saving measures taken	1349	1514	2009			

>>>

		REKENKAMER			WOON		
		CHI-SQUARE/MANN-WHITNEY U			CHI-SQUARE/MANN-WHITNEY U		
		ACTUAL GAS USE PER M²	THEORET-ICAL GAS USE PER M²	DBTA	ACTUAL GAS USE PER M²	THEORETICAL GAS USE PER M²	DBTA
COMFORT	Perception of heat-cold/heat	5	23	12	401922	440122	403956
	Perception of cold				697417	648306	732037
	Perception of dry/humid air	6	16	8	886199	806931	865960
	Perception of draft	14830	14014	15293	1331444	1220532	1280748
	Unpleasant long waiting time for hot water	21292	19480	20171			

*Highlighted fields are significant on a 95% confidence interval.

TABLE 3 Chi-square from Kruskal-Wallis test and U statistic from Mann-Whitney test together with significance (for a description of the categories, see Table 4.4)

		REKENKAMER					WOON				
			ACTUAL GAS USE PER M²		THEORETICAL GAS USE PER M²			ACTUAL GAS USE PER M²		THEORETICAL GAS USE PER M² DBTA	
		N	MEAN	SD	MEAN	SD	N	MEAN	SD	MEAN	SD
LABEL	A	64	8,4	5,4	8,4	3,7	146	8,6	3,5	8,4	2,1
	B	93	11,2	7,1	11,4	3,1	596	10,7	4,9	11,8	2,7
	C	80	13,8	9,6	13,5	2,8	1108	12,8	7,6	15,2	5,0
	D	54	14,5	8,3	19,6	3,9	806	14,7	5,7	19,9	4,1
	E	53	13,5	6,5	24,5	4,2	621	16,2	6,4	26,0	4,8
	F	59	15,5	7,7	33,3	6,0	502	17,1	6,5	32,0	5,3
	G	57	16,4	9,9	42,6	7,6	329	17,6	7,3	42,6	10,3
OWNERSHIP TYPE	Social rent	412	13,2	8,4	21,1	12,4	1342	14,8	6,9	21,3	9,9
	Private rent	12	14,3	10,9	18,4	14,1	265	14,9	7,2	25,0	12,7
	Owner-occupant	36	12,0	5,8	13,6	7,0	2503	13,7	6,8	20,7	10,5
DWELLING TYPE	Gallery	81	10,2	6,8	11,3	6,5	209	12,2	7,3	15,6	9,7
	Maisonette	13	12,9	7,5	18,8	5,3	198	13,2	6,7	22,8	11,5
	Flat with staircase entrance	321	13,6	8,6	22,5	12,3	334	15,0	7,8	24,5	12,0
	Row house - between	33	13,4	6,0	18,6	8,7	1272	13,0	5,2	19,2	8,2
	Semi-detached	9	19,9	6,9	40,4	15,0	552	14,7	5,6	21,4	9,9
	Row house - corner						684	15,6	6,4	23,7	10,0
	Detached						568	15,0	6,5	23,7	11,7

>>>

TOC

		REKENKAMER					WOON				
			ACTUAL GAS USE PER M²		THEORETICAL GAS USE PER M²			ACTUAL GAS USE PER M²		THEORETICAL GAS USE PER M² DBTA	
		N	MEAN	SD	MEAN	SD	N	MEAN	SD	MEAN	SD
HEATING TYPE	Gas stove	64	15,3	9,9	38,6	8,0	152	14,8	7,5	37,7	13,1
	Gas boiler η<83%	16	16,2	7,7	34,8	14,1	86	17,7	8,2	32,6	11,0
	Gas boiler η>83%	51	14,3	7,7	24,3	11,4	178	15,9	7,0	25,3	11,5
	Gas boiler η>83% electric flame						344	16,0	7,1	25,5	11,5
	Gas boiler η>90%	13	12,9	8,5	17,6	7,1	288	14,3	6,0	21,8	9,4
	Gas boiler η>94%	1	12,4		18,8		44	14,9	5,2	20,8	7,8
	Gas boiler η>96%	314	12,4	7,9	15,6	8,3	3014	13,7	6,8	19,3	9,1
ELECTRIC BOILER	Electric boiler	452	13,0	8,1	20,3	12,2	3596	14,3	7,1	21,1	10,7
	No electric boiler	8	17,9	13,3	29,2	12,9	514	13,3	5,3	21,5	8,8
HAL HEATED	Hall not heated	275	12,9	8,5	21,5	12,7	305	15,4	7,3	28,5	12,2
	Hall heated	103	13,4	7,8	17,9	11,5	1271	14,7	6,9	20,2	9,0
MECHANICAL VENTILATION	Mechanical ventilation	167	11,4	7,7	15,8	9,6	1640	12,5	5,8	16,0	7,7
	No mechanical ventilation	170	15,1	8,7	27,3	12,5	2130	15,3	6,4	25,1	10,2
TAPWATER TYPE	Gas boiler without hot water reserve	338	12,6	8,0	17,8	10,0	3002	13,9	6,3	19,7	9,1
	Gas boiler + hot water reserve						712	14,3	8,7	21,5	10,8
	Kitchen boiler	46	15,8	9,3	37,5	10,0	161	16,6	7,0	36,9	14,3
	Shower boiler	21	14,6	7,6	27,8	16,0	115	16,6	8,9	29,8	9,4
	Gas boiler						46	14,2	6,9	27,3	14,3
	Electric boiler	8	17,9	13,3	29,2	12,9	55	13,0	5,6	30,1	11,3
THERMOSTAT	None	411	13,1	8,4	20,9	12,2	2504	14,3	6,8	21,8	10,7
	Programmable	49	13,0	7,5	16,8	11,7	1606	13,9	7,1	20,2	10,2

TABLE 4 Descriptive statistics of categorical variables

§ 4.4.1.5 Comparison of Rekenkamer and WOON data

Some interesting observations could be made when comparing Rekenkamer data results with WOON results. In general, WOON dataset managed to confirm most significant correlations with actual gas use (15 out of 17), an equal number (15 out of 17) of correlations detected with theoretical gas use and 18 out of 22 detected correlations with DBTA. Hereby we do not count the variables that were not present in both datasets, such as 'not ventilating while heating, not setting thermostat too high etc.' and we only look for significance in WOON where there has been a significant correlation in the Rekenkamer dataset.

However, WOON dataset contained almost 10 times as many records; therefore, several additional correlation were found. In particular, we detected more correlations between theoretical gas use and behaviour variables such as presence and set point temperature. This means these correlations might also exist also in the Rekenkamer, but it could be that our sample is too small for to detect them.

Another problem in dealing with two datasets which are based on a different survey is that some variables are not exactly the same and hence difficult to compare. This is the case especially in some behaviour variables and to some extent also in comfort variables.

§ 4.4.2 Regression analysis

§ 4.4.2.1 Whole sample

Regression analysis of the total sample showed that with the variables used one can explain 23,8% variance in actual energy use, 65,1% in theoretical and 40,9% in the DBTA. Regression analysis for the total sample is further broken down in the next section. Regressions were also performed per group of characteristics, to see how much variance in total gets explained by a single group (Table 5). One can see that dwelling characteristics and occupant behaviour explain a roughly equal amount of variation in the actual gas consumption, whereas in other two consumption categories dwelling characteristics explain much more, in case of theoretical gas use even a majority of variation. Household characteristics explain small variations (up to 5%) in all three consumption categories.

R_2 VALUES	DWELLING CHARACTERISTICS	HOUSEHOLD CHARACTERISTICS	OCCUPANT BEHAVIOUR	COMFORT	TOTAL
Actual gas use per m^2	8,6	3,1	10,7	0	23,8
Theoretical gas use per m^2	64,3	4,3	7,5	0	65,1
DBTA	39,3	4,3	9,1	2,5	40,9

TABLE 5 R^2 values in each group of predictors separately and in the total regression (all predictor groups)

For actual consumption, each additional 10 years to building age results in 0,39 m^3/m^2 more gas consumption (Note: this is only true in the exact combination of predictors used in the regression analysis) (Table 6). Conversely, 10m^3 less floor area causes a decrease in consumption for about 1,18 m^3/m^2 (Table 6). Both these variables were also significant predictors for theoretical gas use, building age about twice as strong and floor area about a third half less. Age of the dwelling remains a good predictor for DBTA – for each 10 additional years, dwelling has a DBTA larger for 0,67 m^3/m^2 (Table 8).

Presence and indoor temperature are two variables that have effect on actual consumption and the DBTA. For each additional day of midday presence, actual gas use is 0,631 m^3/m^2 (Table 6) higher, whereas night-time presence has the opposite effect of lowering gas use by 0,995 m^3/m^2 (Table 6). Each additional degree night time temperature also increases the gas use for 0,123 m^3/m^2 (Table 6) and midday temperature for 0,242 m^3/m^2 (Table 6). When looking at the DBTA, midday presence has the effect of reducing the difference by -0,942 m^3/m^2 (Table 8), but when indoor temperature in occupants absence is lower, the difference is also lower(-0,189 m^3/m^2) (Table 8).

Dwelling type is a variable significant only when regressing theoretical gas consumption (Table 7). Flats with staircase entrance, semidetached houses and row houses seem to consume more theoretical gas use than gallery flats, which is line with the consumptions in Figure 4.

When it comes to heating type, all types have a significantly lower DBTA consumption than gas stove. An even better predictive power is however encountered looking at theoretical gas consumption; all systems relate to a lower theoretical gas use than gas stove. Installation system has few effect on actual gas consumption; however, there is a difference between the least efficient gas stove and the most efficient boiler (η>96%), which can also be seen in Figure 5.

Regarding household composition, it can be noted that all household types with an elderly occupant have higher gas consumption. Furthermore, people who find it really easy to pay the energy bill seem to consume less gas in reality than the people who find it 'only' easy. The occupants with only averagely efficient behaviour and the ones that set thermostat too high turned out to consume more gas. All these variables were not significant regarding the theoretical gas use and DBTA.

	ADJ. R^2=65,1%	B	STD. ERROR	BETA	SIG.
	(Constant)	8,901	3,108		,004
DWELLING CHARACTERISTICS	Age of the building	,039	,010	,181	,000
	Floor area	-,118	,021	-,302	,000
	Age of the respondent	,084	,029	,166	,004
	Number of occupants	1,195	,467	,142	,011
HOUSEHOLD CHARACTERISTICS	Missing vs. very easy to pay energy bill	3,502	4,072	,039	,390
	Relatively easy vs. very easy to pay energy bill	-2,136	,830	-,135	,010
	A bit hard vs. very easy to pay energy bill	,002	1,100	,000	,999
	Very difficult vs. very easy to pay energy bill	1,054	1,957	,026	,590
OCCUPANT BEHAVIOUR	Number of weekdays of presence - midday	,631	,207	,168	,002
	Number of weekdays of presence - night	-,995	,360	-,134	,006
	Average reported temperature during the day	,242	,104	,110	,021
	Average reported temperature at night	,123	,051	,116	,015
	Missing vs. energy efficient behaviour	7,545	4,946	,068	,128
	Average vs. energy efficient behaviour	2,125	,751	,133	,005
	Inefficient vs. efficient behaviour	3,715	1,874	,090	,048

TABLE 6 Regression analysis of actual gas consumption per m^2 floor area

	ADJ. R^2=65,1%	B	STD. ERROR	BETA	SIG.
	(Constant)	30,656	2,752		,000
DWELLING CHARACTERISTICS	Age of the building	,097	,012	,287	,000
	Floor area	-,079	,019	-,134	,000
	Maisonette vs. gallery house	3,314	2,434	,044	,174
	Flat with a staircase entrance vs. gallery house	2,650	1,082	,098	,015
	Row house vs. gallery house	3,621	1,666	,074	,030
	Semidetached vs. gallery house	18,661	2,851	,204	,000
	Missing data vs. gallery house	2,125	7,372	,008	,773
	Heating with η<83% boiler vs. gas stove	-4,427	2,225	-,066	,047
	Heating with η>90% boiler vs. gas stove	-11,717	2,773	-,136	,000
	Heating with η>96% boiler vs. gas stove	-14,530	1,321	-,546	,000
	Heating with η>83% vs. gas stove	-6,478	1,624	-,162	,000
	Heating other vs. gas stove	-16,705	5,359	-,092	,002
	Shower boiler vs. combined gas boiler (no hot water reserve)	5,814	1,737	,099	,001
	Kitchen boiler vs. combined gas boiler (no hot water reserve)	5,039	1,437	,126	,001
	Electric boiler vs. combined gas boiler (no hot water reserve)	1,328	2,691	,015	,622
	Other vs. combined gas boiler (no hot water reserve)	-1,710	3,186	-,016	,592
OCCUPANT BE-HAVIOUR	Ventilating in the week missing data vs. week-ends more ventilation	6,285	2,123	,090	,003
	Ventilating in the week equal vs. week-ends more ventilation	1,336	,878	,050	,129
	Ventilating in the week less vs. week-ends more ventilation	3,709	1,732	,068	,033
COM-FROT	Draft yes/no	-1,910	,847	-,065	,025

TABLE 7 Regression analysis of theoretical gas consumption per m^2 floor area

		ALL DATA (R^2=40,9%)			UNDERPREDICTIONS (R^2=19,9%)			OVERPREDICTIONS (R^2=50,8%)		
		B	SE	BETA	B	SE	BETA	B	SE	BETA
	(Constant)	21,28	2,11		-4,21	2,02		23,70	2,31	
DWELLING CHARACTERISTICS	Age of the building	0,07	0,01	0,20				0,06	0,01	0,21
	Floor area							-0,07	0,02	-0,14
	Maisonette vs. gallery house							5,35	2,77	0,09
	Flat with a staircase entrance vs. gallery house							0,84	1,27	0,04
	Row house vs. gallery house							-0,24	1,92	-0,01
	Semidetached vs. gallery house							10,11	2,77	0,16
	Missing data vs. gallery house							2,51	7,06	0,01
	Heating with η<83% boiler vs. gas stove	-2,97	3,00	-0,04				-4,66	2,33	-0,09
	Heating with η >90% boiler vs. gas stove	-14,89	3,28	-0,19				-10,86	2,86	-0,16
	Heating with η>96% boiler vs. gas stove	-16,24	1,49	-0,62				-12,82	1,20	-0,62
	Heating with η>83% boiler vs. gas stove	-10,46	1,98	-0,27				-8,81	1,59	-0,29
	Heating other vs. gas stove	-12,95	6,78	-0,08				-13,99	7,05	-0,08
OCCUPANT BEHAVIOUR	Number of weekdays of presence in the morning				1,27	0,54	0,30			
	Number of weekdays of presence during midday	-0,94	0,23	-0,16	-1,78	0,56	-0,42			
	Average reported temperature when nobody is at home	-0,19	0,06	-0,12				-0,23	0,05	-0,18
	Programmable thermostat				5,49	1,79	0,29			
	Water saving shower head				-4,93	1,39	-0,34			
COMFORT	Missing data vs. average temperature							-5,51	7,00	-0,03
	Too cold vs. average temperature							2,18	0,97	0,09

The orange values are insignificant on a 95% confidence interval scale.

TABLE 8 Regression analysis of the DBTA per m^2 floor area for all data, only underpredictions and only overpredictions

§ 4.4.2.2 DBTA—Separate analysis for under and overprediction

Considering the fact that under and overprediction are also in literature described separately (Sunikka-Blank and Galvin, 2012), we also made a regression model for each of the two phenomenons separately (besides the regression model for the total sample). Here, cases where theoretical gas use per m^2 is higher than actual (overprediction) were analysed separately from underpredictions (theoretical consumption is lower than actual). We found out that underpredictions seemed to be harder to explain with our set of variables, only 23% of variance was explained. The factors explaining underpredictions were completely different from overpredictions (Table 8). For underprediction, all explanatory variables relate to occupant behaviour: presence at home seemed to matter, together with the presence of a programmable thermostat and water-saving shower head. Overpredictions could be explained more than twice as well, R^2 was 50,8%. Here, dwelling characteristics (dwelling and installation type) play the main role, although average temperature and perception of indoor temperature were significant as well. This seems to indicate that the building parameters are responsible for most of the discrepancy in overpredictions; however, occupancy patterns are more significant in underprediction.

§ 4.4.3 Improved theoretical model based on the regression analysis

In this section, a regression analysis was made using theoretical gas consumption per m² floor area together with all other available dwelling characteristics as predictors and actual gas consumption per m² floor area as a dependent variable. This way we were able to tell how much of the variation in the actual gas consumption we can account for by using theoretical gas use and how much by additional information about the dwelling.

R^2=33,8%	B	STD. ERROR	BETA
(Constant)	1,224	1,438	
Theoretical gas use per m²	0,305	0,032	0,611
Maisonette vs. gallery house	-1,183	1,863	-0,03
Flat with staircase entrance vs. gallery house	0,787	0,844	0,056
Row house vs. gallery house	3,083	1,308	0,124
Semidetached vs. gallery house	4,167	2,015	0,107
Missing data vs. gallery house	-1,02	5,142	-0,009
Heating with η<83% boiler vs. gas stove	2,219	1,552	0,073
Heating with η>90% boiler vs. gas stove	2,6	2,116	0,059
Heating with η>96% boiler vs. gas stove	2,417	0,993	0,187
Heating with η>83% boiler vs. gas stove	3,529	1,11	0,183
Heating other vs. gas stove	4,644	5,17	0,04

*Highlighted values are significant on a 90% confidence interval.

TABLE 9 Regression of actual gas use using theoretical gas use and dwelling characteristics as predictors in dwellings where actual consumption is lower than theoretical (overprediction)

As shown in Table 9, for overpredictions, dwelling type and installation type are significant variables apart from theoretical gas consumption. The R^2 value is relatively low, meaning that only a few variation in actual gas use can be explained using these variables. Table 10 shows that for underpredictions, variations are more easily explainable (also because the discrepancies are smaller). Here, one can explain about 60% using the additional variables of thermostat type and presence of water saving shower head.

R^2=60,0%	B	STD. ERROR	BETA
(Constant)	12,747	3,837	
Theoretical gas use per m²	0,94	0,106	0,656
Floor area	-0,075	0,039	-0,144
Programmable thermostat	-5,246	1,871	-0,191
Water saving shower head	4,008	1,429	0,188

*Highlighted values are significant on a 90% confidence interval.

TABLE 10 Regression of actual gas use using theoretical gas use and dwelling characteristics as predictors in dwellings where actual consumption is higher than theoretical (underprediction)

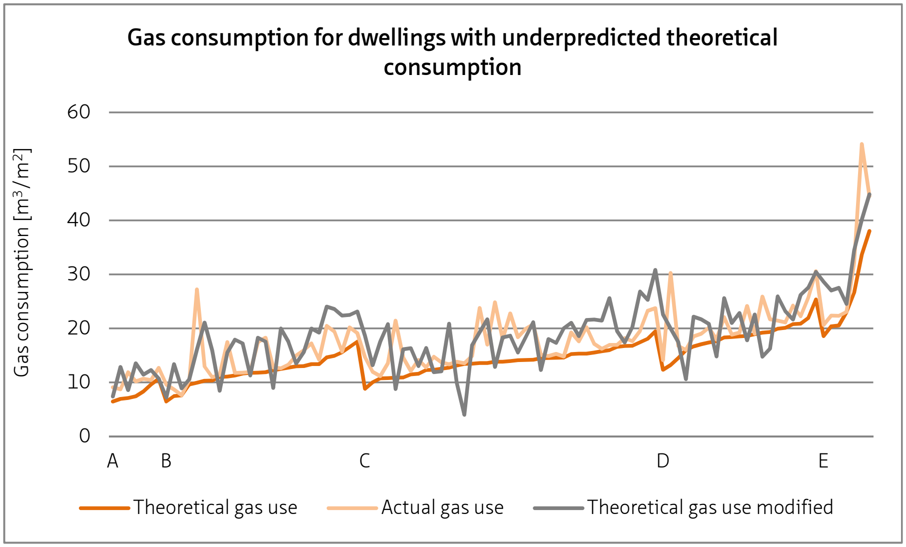

FIGURE 8 Theoretical, actual and modified theoretical gas consumption for dwellings with underpredicted theoretical consumption, a random sample of 100 dwellings from WOON sample

The values (B coefficients) acquired in these regression analyses used the Rekenkamer dataset which were then used on the larger WOON dataset. Figure 8 ad Figure 9 show that by using actual energy data for a regression analysis and modifying the theoretical consumption according the regression results can result in values, much closer to actual ones.

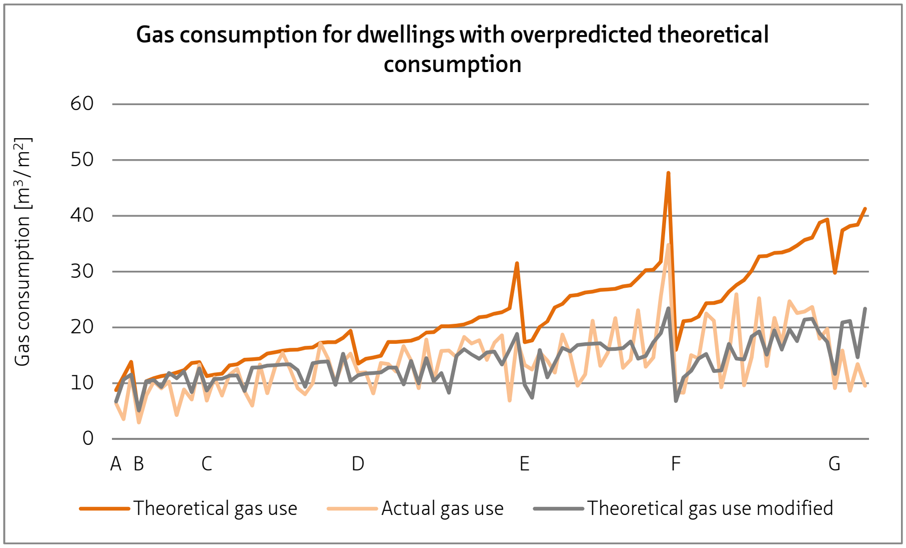

FIGURE 9 Theoretical, actual and modified theoretical gas consumption for dwellings with overpredicted theoretical consumption, a random sample of 100 dwellings from WOON sample

Just like the figures above, Table 11 and Figure 10 prove that the modified values are indeed closer to actual gas use than the original values. The standard deviations remain comparable, and in case of overpredictions they are even smaller (relative SD of 27% vs. 45 in the original theoretical consumption), which means that adapting the values for the B coefficients does not create extreme outliers.

	UNDERPREDICTIONS	OVERPREDICTIONS
N total	505	2691
Mean theoretical gas consumption (m^3/m^2)	15,1	22,3
Mean actual gas consumption (m^3/m^2)	18,5	13,1
Mean theoretical gas consumption modified (m^3/m^2)	19,0	14,1
SD theoretical gas consumption (m^3/m^2)	5,7	10,1
SD actual gas consumption (m^3/m^2)	7,4	5,5
SD theoretical gas consumption modified (m^3/m^2)	7,6	3,9
N (%) better fitting prediction	412 (82%)	2567 (95%)
N (%) poorer fitting prediction	93 (18%)	124 (5%)

TABLE 11 Descriptive statistics of the entire WOON sample

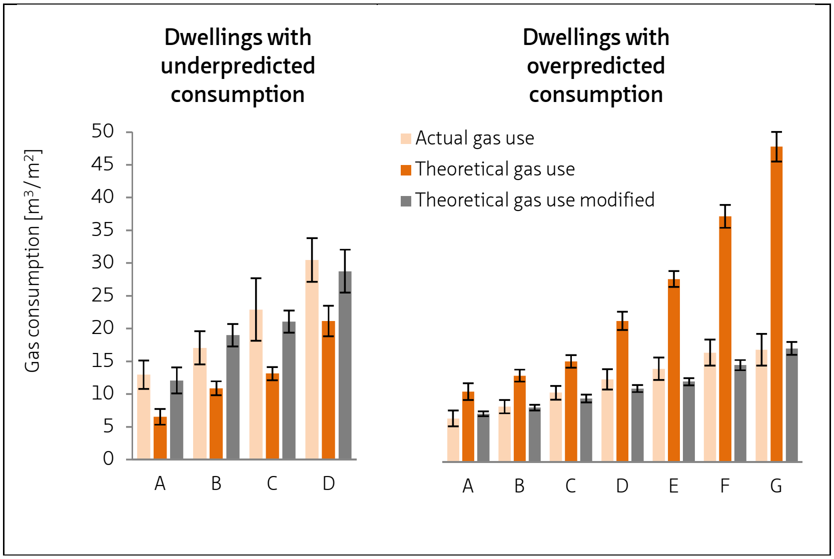

FIGURE 10 Mean and 95% confidence interval of the theoretical, actual and modified theoretical consumption

Figure 8 and Figure 9 show that by using actual gas consumption data, much better estimates of theoretical gas consumption can be obtained. The results are undoubtedly better regarding the average within a label category (Figure 10). For the individual dwelling, the new prediction is sometimes very good, but there are still some outliers. In the future, these should be investigated more closely to see which features cause these consumptions to fit the actual use poorly; it could be dwelling, household, or behaviour related.

§ 4.5 Are the results in line with expectations?

Table 12 shows the variables that were significant in the Rekenkamer dataset. The general outcome largely corresponds to correlations we expected to obtain (4.2.3.1). Dwelling characteristics seem to dominate the correlations with the theoretical gas use, whereas household and occupant characteristics are more relevant in actual gas use. Comfort played no role in actual gas consumption, but did have a correlation with theoretical gas use, which shows that our hypothesis of differently performing dwellings having different levels of comfort was correct. We found the temperature perception to be significantly correlated with dwellings performance. This is an important finding, since it proves that heating demand is not the only difference between performance classes, but that albeit forgotten, comfort is also an output

TOC

that should be measured. These findings were similar in both, individual correlation data as well as regression results. It is notable though, that it is much easier to find significant variables looking at individual correlations. In regression analyses, less factors are significant.

It is also extremely important not to take the precise results out of context – the heating system for example was significant regarding actual gas use, but as seen from Figure 5, only the gas stove and the most efficient boiler were in fact significantly different in their actual consumptions. Precise analysis of categorical variables is therefore imperative in such studies, as well as a multiple regression analysis which puts individual variables into context.

	DWELLING CHARACTERISTICS	HOUSEHOLD CHARACTERISTICS	OCCUPANT BEHAVIOUR	COMFORT
Actual gas use per m^2	Floor area, Age of the building, Dwelling type, Heating type, Ventilation type	Number of occupants, Household composition, Education, Ability to pay the energy bills	Number of weekdays of presence - midday, Average reported temperature during the day/evening/night/nobody at home, Showers per week, Perception dwellings/household energy performance, Not setting thermostat too high, Not ventilating while heating, No energy saving measures taken	
Theoretical gas use per m^2	Age of the building, Dwelling type, Heating type, Ventilation type, Electrical boiler presence, Tap water heating type, Programmable thermostat presence	Age of respondent, Spendable income, Ownership type	Average reported temperature, nobody at home, Presence of water saving shower head, Not setting thermostat too high, No energy saving measures taken	Perception of heat-cold/heat, Perception of dry/humid air, Perception of draft
DBTA	Age of the building, Dwelling type, Heating type, Ventilation type, Electrical boiler presence, Tap water heating type, Programmable thermostat presence	Age of respondent, Ownership type, Household composition	Number of weekdays of presence - morning/midday/evening, Average reported temperature during the day/evening/night/nobody at home, Perception dwellings/ household energy performance, Ventilation habits weekends, Presence of water saving shower head, Not setting thermostat too high, Not ventilating while heating	Perception of heat-cold/heat, Perception of dry/humid air

TABLE 12 Summary of significant variables from correlation results for the Rekenkamer sample

The regression results in 4.4.1.5 comply largely with the hypothesis in 4.2.3.1 with occupant behaviour explaining the most variance in actual gas use and comfort being relevant only for DBTA. Dwelling characteristics play the most prominent role in theoretical consumption. Also the fact that in total we can explain less variance in actual (23,8%) than in theoretical consumption (65,1%) and DBTA (40,9%) is logical, since theoretical depends only on the parameters considered in the calculation method.

Regarding regression of the total sample, the fact that floor area is a significant predictor for actual and theoretical gas use but not for the DBTA implies that floor area is well corrected for across different label categories. However, our hypothesis was that dwelling-related parameters would correlate more with the theoretical gas use than with actual; in this case, actual gas use had a slightly higher correlation. In both cases, a larger floor area means lower gas consumption per m^2. However, floor area is no longer a good predictor when we regress the difference between the consumptions, meaning that floor area plays no role in over/underpredictions when we look at consumption per m^2 dwelling.

Age of the building complies with the hypothesis and has a smaller impact on actual than on theoretical gas use, just like dwelling type and installation system. This makes sense, since age is known to relate well to dwellings performance. However, actual heating consumption depends also on other factors. Age remains relevant also in regression of DBTA – an older dwelling has a higher difference between consumptions.

Furthermore, our hypothesis was also correct in predicting a higher correlation of household and behavioural variables with actual gas use, which was detected in household composition, the ability to pay energy bills, presence at home, set point temperature and efficiency of behaviour. Presence and indoor temperature are two very important parameters in determining real gas use of a dwelling. The fact that midday presence relates to a decreased DBTA could mean that households who spend more time at home somehow match conditions assumed by the theoretical calculations better (because they probably heat their house longer). On the other hand, occupants who spend more time at home during the night tend to have an increased DBTA. It seems that people who are not often sleeping elsewhere tend to have a larger DBTA. Conversely, the ones that often sleep elsewhere (they should in fact be heating their house less) have a smaller DBTA. There could however, be an indirect relationship between people in houses with a smaller DBTA (better performing) and the weekends spent away (wealthier people, more work-related travel, etc.) that was not captured in the multicollinearity tests.

Dwelling and installation type were both relevant predictors of actual gas consumption, however, as hypothesised in the beginning, both were more strongly correlated with theoretical gas use. Semidetached correlate with a larger DBTA, which could be caused by houses a larger outside wall area. Moreover, they have a larger floor area out of which some bedrooms are often not heated – this occurs less in gallery apartments. A correction could be applied towards a better fitting of the theoretical gas consumption. Similar could be done with installation types, since better installation systems seem to perform worse than theoretically expected. This would decrease the difference between the DBTA.

§ 4.6 Conclusions

§ 4.6.1 New insights

Occupant behaviour proved once more to give a large effect on heating consumption, in particular actual where it accounts for almost half of the variance. Also in theoretical consumption and in the DBTA the behaviour accounts for over 7,5 and 9,1% of variance, which is still remarkable.

Moreover, significant differences were found in the separate analysis of under and overpredictions that have not been documented before. Regarding the DBTA and the separate regression for under and for overprediction it seems that whereas in overpredictions (poor performing dwellings) a big role is played by the installation system, dwelling type, floor area and age (all these are parameters that correlate well with theoretical gas use), in underpredictions this is not the case at all. Water saving shower head and programmable thermostat are the two factors that seem to effect DBTA in underpredictions but these two were not significant with regard to theoretical gas use. Underpredictions seem more complex to understand, the effect of significant variables in underprediction is much smaller than in overprediction (R^2=19,9% vs. R^2=50,8%). Some presence variables (morning and midday) were significant predictors, but are also difficult to interpret, since the results are conflicting (positive predictive power for morning and negative for midday presence). Another remarkable finding is that in underprediction, no difference in comfort perception is detected whereas in overpredictions it can be found.

Similar results were obtained in the section 4.4.2; dwelling characteristics play a bigger role in overpredictions. Using the results from this section, one can see which dwelling features should be given a bigger/different weight in the theoretical consumption calculation, to get closer to real, actual values. The results of this section cannot be extrapolated on the whole Netherlands, a much larger and very well representative sample should be used for this purpose, but the results do give an idea of what is possible. The problem with the normalised theoretical calculation is namely, that it was never tested against actual consumption data. Data is now available that enables us to make better predictions. However, for the use of factors as described above in practice, better data would be needed. In fact, a regression analysis would have to be done per label category to obtain the appropriate factors for each label class. After the theoretical calculation of dwellings label certificate using the existing methodology, the factor for the specific label category would be applied.

§ 4.6.2 Implications

Our study confirmed the previously discrepancies between theoretical and actual gas use across different performance classes (in our case label categories) shown in previous studies. Normalising building use with default values such as indoor temperature, heated floor area, occupancy etc. does not yield accurate predictions about heating energy use. To avoid confusion among users of dwellings' performance certificates, this has to be improved. We showed that as hypothesised, dwelling characteristics play a big role in the variation of theoretical gas consumption, whereas occupant behaviour related better to actual gas consumption, which is also summarized in Table 13. This table highlights some interesting results, such as the fact that the influence of building age, and dwelling and installation type probably comes from the overpredicted cases. It also demonstrates that by narrowing down the sample to underpredicted dwellings, variables such as water saving shower head and programmable thermostat become significant. Similar methods should be used in the future to obtain more refined results, for example to find out in which specific subgroup the presence of elderly influences the actual gas use significantly (first column Table 13). In terms of practical results, it turns out that flats with a staircase entrance, semi-detached dwellings and dwellings with a less efficient heating installation system are characterised by a larger performance gap (Table 7, Table 8 and Table 9) and this is due to the overpredicted records (Table 8 and Table 9). On the basis of the results, a correction factor could be applied to the theoretical gas consumption of these groups of dwellings in order to reduce the performance gap. Similar corrections could be applied if a similar study would be repeated on a larger sample (where also less well-represented dwelling groups, such as detached houses would be more numerous).

However, variation in actual gas use is very complex and difficult to explain even by using detailed survey data. In the future this could be improved by monitoring of occupants presence and practices real-time which would give more detailed and realistic information, since surveys are always prone to biases. By the use of monitoring data, a great deal of the uncertainty would be improved.

	ACTUAL GAS USE	THEORETICAL GAS USE	DBTA TOTAL	DBTA UNDER-PREDICTION	DBTA OVER-PREDICTION	MODIFIED THEORETICAL GAS USE UNDER-PREDICTION	MODIFIED THEORETICAL GAS USE OVER-PREDICTION
Theoretical gas use	/	/	/	/		Theoretical gas use	Theoretical gas use
Dwelling characteristics	Building age, floor area, dwelling type, installation type	Building age, floor area, dwelling type, installation type	Building age, dwelling type, installation type	Water saving shower head, programmable thermostat	Building age, floor area, dwelling type, installation type	Water saving shower head, programmable thermostat	Dwelling type, installation type
Household characteristics	Elderly, ability to pay the bill						
Occupant characteristics	Midday presence, night temperature presence, efficiency of behaviour, thermostat setting		Presence midday and morning, temperature when nobody is home	Presence midday and morning			
Comfort		Temperature perception			Temperature perception		

TABLE 13 Summary of all regression results per parameter group for all independent variables

Furthermore, the paper has proven that a positive DBTA has completely different causes than a negative one. The two issues should be addressed separately also in the future. If enough data is present it might also be a good idea to analyse the DBTA in different label classes separately.

Also, the paper shows that by using aggregated actual heating energy data, it is very well possible to calculate a more accurate predicted heating consumption on the level of an individual dwelling by using regression analysis. Already by modifying dwelling and/or household characteristics only, we obtain a much more accurate prediction. Expanding the prediction to variable occupant behaviour and comfort perception might also be useful for some applications (like tailored advice about efficient energy saving measures for a specific household), but not for a performance certificate, since this would mean that a certificate is no longer valid when occupied by a different user.

In the paper we found dwelling and household characteristics to be relatively easy to record via a survey if compared to the other two parameter groups. The two slightly more complex parameters among household characteristics were household composition and education. A clever survey design is needed here to really capture groups that demonstrate differences when it comes to gas use. Since so far, few detailed research is available, our survey questions might have been too granulated (for example, it does not seem to matter whether there are three children and two adults and three children and three adults). This was even more of a problem in occupant behaviour variables such as presence at home, where it seemed as if presence in the morning and midday were the only ones significant. It might be better to have a good composite variable for presence, like was done in the Majcen and Itard (2014b).

Besides clever design of survey questions, results of regression analysis might also depend on sample selection. Our studies sample was not selected randomly which has some disadvantages (less chance of a good representatively) and some advantages (enough data points to show correlations also in extreme consumptions). We have seen in this paper that in dwellings where theoretical consumption is higher than actual completely different predictors were relevant than in the ones where theoretical consumption was lower. Underprediction seems to be more complex and more behaviour dependent; however, the variation in the actual consumption in these dwellings is more easily explained by a normalised theoretical consumption since the discrepancy is relatively smaller than in dwellings with overpredictions. The fact that differently performing dwellings correlate with predictors differently has to be considered in future studies as well.

Furthermore, some uncertainties were encountered. It remains unclear how well the degree day method really corrects for the heating intensity, and in these paper we showed some uncertainties regarding actual use of different samples in The Netherlands. At the same time, there are no official references proving how much of the actual data is based on real meter readings and how much is estimated.

§ 4.7 References

Broekhuizen, J., Jakobs, E., 2014. Energielabel, energiegedrag, energiearmoede en wooncomfort Amsterdamse corporatiewoningen. Amsterdam municipality, Department Research and Statistics, Amsterdam, September 2014. Accessible via http://www.rekenkamer.amsterdam.nl/wp-content/uploads/2014/11/Bureau-O-S-Energielabel-energiegedrag-energiearmoede-en-wooncomfort.pdf

Compendium voor de leefomgeving website http://www.compendiumvoordeleefomgeving.nl/indicatoren/nl0556-Energielabels-woningen.html?i=9-53 (Accessed on 5th December 2014)

Convenant Energiebesparing Corporatiesector, October 2008, accessed on 9th January 2015 on http://www.rvo.nl/onderwerpen/duurzaam-ondernemen/gebouwen/woningbouw/beleidskader/convenant-energiebesparing-corporatiesector.

Convenant Energiebesparing bestaande gebouwen (''Meer met Minder''), January 2008, accessed on http://www.rijksoverheid.nl/documenten-en-publicaties/convenanten/2012/06/28/convenant-energiebesparing-bestaande-woningen-en-gebouwen.html

Derde onderzoek naar de betrouwbaarheid van energielabels bij woningen, VROM Inspectie, Ministry of Infrastructure and Environment, August 2011.

DG Energy website, http://ec.europa.eu/energy/en/topics/energy-efficiency/buildings, accessed on 20th January 2015

Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the Energy Performance of Buildings

Fagerland, B.M.C., 2012. t-tests, non-parametric tests, and large studies—a paradox of statistical practice? MC Medical Research Methodology 12 (78)

French, L.J., Camilleri, M.J., Isaacs, N.P., Pollard, A.R. 2007. Temperatures and heating energy in New Zealand houses from a nationally representative study—HEEP, Energy and Buildings, 39 (7), pp. 770–782

Gill, Z., Tierney, M., Pegg, I., Allan, N., 2010. Low-energy dwellings: the contribution of behaviours to actual performance, Building Research & Information, 38 (5), 491-508.

de Groot, E., Spiekman, M., Opstelten, I. 2008. Dutch research into user behaviour in relation to energy use of residences, PLEA 2008 – 25th Conference on Passive and Low Energy Architecture, Dublin, Ireland, 22–24 October 2008

Guerra-Santin, O., Itard, L. 2010. Occupants' behaviour: determinants and effects on residential heating consumption Building Research and Information, 38 (3), pp. 318–338

Guerra Santin, O., 2010. Actual Energy Consumption in Dwellings: the Effect of Energy Performance Regulations and Occupant Behaviour. OTB Research Institute, October 2010.

Guerra Santin, O., Itard, L. and Visscher, H. 2009. The effect of occupancy and building characteristics on energy use for space and water heating in Dutch residential stock. Energy and Buildings 41(11), pp. 1223-1232.

Guerra Santin, O., Itard, L., 2012. The effect of energy performance regulations on energy consumption. Energy Efficiency, 5(3), 1-14.

Haldi, F. Robinson, D. 2011. The impact of occupants' behaviour on building energy demand. Journal of Building Performance Simulation, 4 (4), pp. 323-338.

Haas, R., Auer, H., Biermayr, P. 1998. The impact of consumer behavior on residential energy demand for space heating Energy and Buildings, 27 (2), pp. 195–205

Herhalingsonderzoek betrouwbaarheid energielabels bij utiliteitsbouw, Inspectie Leefomgeving en Transport, Ministry of Infrastructure and Environment, November 2013.

Hunt, D., Gidman, M., 1982. A national field survey of house temperatures, Building and Environment, 17 (2), pp. 107–124

Hong, S.H., Gilbertson, J., Oreszczyn, T., Green, G., Ridley,I. 2009. the Warm Front Study Group, A field study of thermal comfort in low-income dwellings in England before and after energy efficient refurbishment, Building and Environment, Volume 44, Issue 6, Pages 1228-1236

Informatiecode Elektriciteit en Gas, 2014 https://www.acm.nl/download/documenten/acm-energie/informatiecode-19-februari-2014.pdf

Ioannou, A., Itard, L.C.M., 2015. Energy performance and comfort in residential buildings: Sensitivity for building parameters and occupancy, Energy and Buildings, Volume 92, 2015, Pages 216-233

Ioannou, A., Itard, L. 2015. Energy performance and comfort in residential buildings: sensitivity to building parameters and occupancy, Energy and Buildings, accepted in January 2015.

Kamp (2014). Answer from Minister Kamp (Economic Affairs), in the name of Ministry of Infrastructure and Environment, dated from 17 January 2014, accessed on 29th January 2015 via https://zoek.officielebekendmakingen.nl/ah-tk-20132014-987.html

Koepelconvenant Energiebesparing gebouwde omgeving, June 2012, accessed on 10th January 2015 on http://www.rijksoverheid.nl/documenten-en-publicaties/convenanten/2012/06/28/koepelconvenant-energiebesparing-gebouwde-omgeving.html

Lantz, B., 2013. The large sample size fallacy, Scandinavian Journal of Caring Sciences 27 (2), 487–492.

Laurent, M., Allibe, B., Oreszczyn, T., Hamilton, I., Tigchelaar, C., Galvin, R., 2013. Back to reality: How domestic energy efficiency policies in four European countries can be improved by using empirical data instead of normative calculation, In: Proceedings of the European Council for an Energy Efficient Economy (ECEEE) Summer School, 3–8 June 2013, Belambra Presqu'île de Giens, France.

Lin, M., Lucas, H., Shmueli, G. 2013. Too Big to Fail: Large Samples and the p-Value Problem, Information Systems Research, 24(4), pp.906-917.

Lindén, A., Carlsson-Kanyama, A., Eriksson, B., 2006. Efficient and inefficient aspects of residential energy behaviour: what are the policy instruments for change? Energy Policy, 34 (14), pp. 1918–1927)

Majcen, D., Itard, L., Visscher, H., 2013a. Actual and theoretical gas consumption in Dutch dwellings: What causes the differences? Energy Policy 61, 460–471.

Majcen, D., Itard, L., Visscher, H., 2013b. Theoretical vs. actual energy consumption of labelled dwellings in the Netherlands: Discrepancies and policy implications, Energy Policy 54, 125–136.

Majcen, D., Itard, L. 2014b. Relatie tussen huishoudenskenmerken en gedrag, energielabel en werkelijk energiegebruik in Amsterdamse corporatiewoningen, September 2014, OTB Research Institute

Majcen, D., Itard, L. 2014a. Relatie tussen energielabel, werkelijk energiegebruik en CO_2-uitstoot van Amsterdamse corporatiewoningen, August 2014, OTB Research Institute

Mishra, A.K., Ramgopal, M. 2013. Field studies on human thermal comfort — An overview, Building and Environment, Volume 64, Pages 94-106

Oreszczyn, T., Hong, S.H., Ridley, I., Wilkinson, P. 2006. Determinants of winter indoor temperatures in low income households in England Energy and Buildings, 38 (3), pp. 245–252

Raynaud, M. 2014. Evaluation ex-post de l'efficacité de solutions de rénovation énergétique en résidentiel, Doctoral thesis, MINES ParisTech Centre Efficacité énergétique des Systèmes.

Sardianou, E. 2008. Estimating space heating determinants: an analysis of Greek households Energy and Buildings, 40 (6), pp. 1084–1093

Sharpe, T.R., Shearer, D. 2013. Adapting the Scottish tenement to twenty-first century standards: An evaluation of the performance enhancement of a nineteenth century "Category B" listed tenement block in Edinburgh, Journal of Cultural Heritage Management and Sustainable Development; 3(1), 2013.

Shipworth, M., Firth, S.K., Gentry, M.I., Wright, A.J., Shipworth, D.T., Lomas, K.J. 2009. Central heating thermostat settings and timing: building demographics, Building Research and Information, 38 (1), pp. 50–69

Sunikka-Blank, M., Galvin, R. 2012. Introducing the prebound effect: the gap between performance and actual energy consumption, Building Research & Information, 40 (3), pages 260-273

Tigchelaar, C., Daniëls, B., Maenkveld, M., 2011. Obligations in the existing housing stock: who pays the bill? In: Proceedings of the European Council for an Energy Efficient Economy (ECEEE) Summer School, 6–11 June 2011, Belambra Presqu'île de Giens, France.

Tigchelaar, C., Leidelmeijer, K., 2013. Energiebesparing: Een samenspel van woning en bewoner – Analyse van de module Energie WoON 2012, RIGO and ECN, August 2013.

Tigchelaar, C., Leidelmeijer, K., 2013. Energiebesparing: Een samenspel van woning en bewoner - Analyse van de module Energie WoON 2012, a report published by ECN (Energy research Centre of the Netherlands) and RIGO Consultancy.

Wei, S., Jones, R., de Wilde, P., 2014. Driving factors for occupant-controlled space heating in residential buildings, Energy and Buildings, Volume 70, Pages 36-44.

de Wilde, P., 2014. The gap between predicted and measured energy performance of buildings: A framework for investigation, Automation in Construction, Volume 41, Pages 40-49.