A+BE
Architecture
and the
Built environment

#04
2016

Predicting energy consumption and savings in the housing stock

A performance gap analysis in the Netherlands

Daša Majcen

Predicting energy consumption and savings in the housing stock

A performance gap analysis in the Netherlands

Daša Majcen
Delft University of Technology, Faculty of Architecture and the Built Environment,
OTB - Research for the Built Environment

Architecture and the Built environment

abe.tudelft.nl

Design: Sirene Ontwerpers, Rotterdam

Cover image: 'Canal houses in Leiden', Hansenn, Nationale Beeldbank

ISBN 978-94-6186-629-5
ISSN 2212-3202

Predicting energy consumption and savings in the housing stock

A performance gap analysis in the Netherlands

Proefschrift

ter verkrijging van de graad van doctor aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.Ch.A.M. Luyben,
voorzitter van het College voor promoties,
in het openbaar te verdedigen op 12 april 2016 om 10:00 uur.
door Daša MAJCEN
Bachelor of Science in Environmental Sciences, University of Nova Gorica
geboren te Celje, Slovenië.

This dissertation has been approved by the

promotor: Prof.dr.ir. H.J. Visscher
copromotor: dr. L.C.M. Itard

Composition of the doctoral committee:

Rector Magnificus
Prof.dr.ir. H.J. Visscher promotor
dr. L.C.M. Itard copromotor

Independent members:

Prof.ir.P.G Luscuere, Delft University of Technology
Prof.dr.ir. J.L.M. Hensen, Eindhoven University of Technology
Prof.dr. P. Heiselberg, Aalborg University Denmark
Prof. A. Janssens, Ghent University, Belgium

Other members:

Dr.ir. Z. Lukszo, Delft University of Technology

Reserve member:

Prof.dr.ir. M.G. Elsinga, Delft University of Technology

Acknowledgements

The thesis that lays ahead was a complex research project that would not have materialised without the intellectual input of great thinkers and the support of several organisations.

Firstly, I would like to thank my supervisor Laure and promotor Henk, who have been closely involved in this project from the very beginning and were both unmissable. Laure contributed her endless creative enthusiasm in the moments when I most needed it and Henk ensured a smooth cooperation with all the parties involved. Both of them read my papers diligently and with the much appreciated criticism. Their attention to detail helped me polish my research skills and consecutively improved this thesis. With both of them, I deeply enjoyed the fruitful discussions about the research work as well as private topics and can hardly imagine a better tuned research team. Additionally, I would like to express my gratitude to Sylvia, who helped me with the questions regarding statistical methods. I am grateful to all my department colleagues as well as fellow PhD's from other departments for their academic input as well as slightly-less-academic discussions during lunches and coffee breaks.

Besides the colleagues mentioned above, the work would be much harder (if not impossible) without the help of the former AgenschapNL, Rekenkamer Amsterdam, AEDES and the Ministry of the Interior and Kingdom Relations who provided us with the much needed data and trusted us to use it for academic research. Additional thanks goes to the office of Statistics Netherlands who enabled the coupling of the mentioned data with the Dutch socioeconomic figures and provided me with support when carrying out the analysis. I am also indebted to Vabi Software B.V. for answering my numerous methodological questions.

Last but not least, I am very lucky to have an understanding family and a group of friends who stood by me through thick and thin, supported me and kept my spirits high. Without you guys, I wouldn't have made it!

Contents

Summary

Buildings are one of the key target sectors for energy and CO_2 reduction. Policy efforts in the past decades have resulted in technical improvements of the dwelling stock, but despite that, the energy consumption in the residential sector has not yet experienced a dramatic reduction necessary to achieve the set targets. In order to reduce the consumption of the dwelling stock, the European Commission implemented The Energy Performance of Buildings Directive (EPBD), which requires the member states to establish certification schemes for existing buildings. The Netherlands implemented energy labelling in 2008 and based the calculation of dwellings theoretical energy performance on a steady state method. Besides the indication of the label category (A++ to G), the certificate consists of the floor area, the type of dwelling and building related energy use (excluding the appliances), expressed in gas, electricity, heat and the total primary energy consumption. Since the theoretical gas and electricity consumption is portrayed on the front page of the Dutch energy label certificate as 'standard energy use of the dwelling' it became widely used as an indication of consumption and even included in policy goals.

Since the dwelling quality has been steadily improving, the primary reason for the reduction failure is the increasing demand from the side of the occupants. In response to that, the thesis stresses the occupant behaviour factor as crucial in actual energy consumption, accounting for as much as 50% of the variance in heating consumption. In order to improve existing policies and achieve real reduction, the occupant behaviour and its impacts on actual energy consumption needs to be understood better.

This thesis attempts to shine a light onto how well the theoretical consumption predicts the actual energy use and what the causes of discrepancies are, especially with regard to occupant behaviour. Namely, whether the mentioned reduction potential is realized, strongly depends on the actual energy use of the dwellings. Furthermore, the thesis explores what reductions can realistically be achieved by improving the thermal performance of the dwelling stock and whether or not these match the expectations of policy makers.

Research methods

The research used several large datasets, about dwellings theoretical energy performance, most of which were related to energy label certificates. All the datasets containing theoretical performance were merged with actual energy data. In addition to that, some were also enriched with socioeconomic and behaviour related data from Statistics Netherlands (CBS) or from surveys which were designed for the purpose of this research. Simple descriptive statistics were used to compare average theoretical and actual consumptions. Advanced statistical tests were used for detecting correlations, followed by several regression analyses. In a separate scenario study, the resulting averages of both theoretical and actual consumptions were extrapolated nation-wide in order to be compared with the existing policy targets.

Due to low predictive power of the variables in regression analyses, a sensitivity analysis of the theoretical gas use was performed on six assumptions made in the theoretical calculation to show how an increment in one of the assumptions affects the final theoretical gas consumption and whether this can explain the performance gap.

Last but not least, longitudinal data of the social housing dwelling stock between 2010 and 2013 was analysed, focusing on dwellings that had undergone renovation. The goal was to find out whether the theoretical reduction of consumption materialised and to what extent. A comparison of the actual reduction of different renovation measures was made in order to show what renovation practices lower the consumptions most effectively.

The discrepancies between actual and theoretical heating energy consumption in Dutch dwellings

Discrepancies between theoretical and actual gas and electricity consumption

On average, the total theoretical primary energy use seems to be in accordance with actual primary energy consumption but when looking at more detailed data, one can see that the contribution of gas to the actual primary energy is much lower than in the theoretical primary energy and that the contribution of electricity is opposite – higher in the actual than theoretical primary energy. The two effects cancel each other out so that in terms of total primary energy, the theoretical consumption seems to be well predicted. Furthermore, the analyses showed that the variation in electricity consumption is marginal across label categories. This together with the fact that most Dutch dwellings are heated with gas made us focus exclusively on gas consumption in the rest of the thesis.

Whereas it is clear that theoretical electricity consumption is much lower than actual since it does not account for appliances, however, it is much less obvious why gas consumption is on average so much lower in reality than according to theoretical calculation.

Performance gap in relation to energy label

The discrepancies in gas consumption were the largest in the poorest performing dwelling, where theoretical consumption surpassed the actual almost twice, which we also referred to as overprediction. On the other hand, well performing dwellings consume roughly 20% more gas than predicted. Theoretical electricity consumption was at least twice lower than actual in all label categories, due to the fact that actual consumption takes into account electricity use of appliances and theoretical does not. Actual and theoretical electricity consumptions seemed to be rather constant with regard to the label class. Primary energy consumption is a sum of consumption of gas and electricity in MJ for each label class where the efficiency of the electricity generation and of the network was taken into account as well as the heating value of gas burning. The theoretical primary energy use is dominated by gas consumption, since electricity is a relatively small fraction of primary energy use due to exclusion of the household appliances. The relation between actual and theoretical therefore remains similar as seen in gas consumption. For poor label classes, the theoretical consumption is overpredicted by about 30% and for good label classes it is underpredicted for roughly the same percentage.

Electricity consumption does not seem to depend on the energy performance of the dwelling. Moreover, the end uses of electricity included in actual and in theoretical consumption are different to an extent that renders a comparison meaningless (as the theoretical excludes appliances). Therefore the main focus of the thesis was gas consumption, which is also the predominantly used fuel for heating homes in The Netherlands.

Performance gap in different samples

The performance gap was analysed in four different datasets of varying size. All datasets provided very comparable results regarding average actual and theoretical consumptions across label categories. A closer analyses shows that the actual gas consumption has been dropping steadily within label categories A, E, F and G from 2010 till 2012. Theoretical gas consumption remained roughly the same in these years, which means that the performance gap has increased slightly.

Moreover, it was found that the dwellings which had no renovation measures applied and remained unchanged from year 2010 till 2012 still exhibit a 3,5% decrease in gas use between 2010 and 2012, which shows that the decrease detected in the fours studied samples is not due to sampling bias. This decrease could be a consequence of a changing household composition (smaller number of people per household) or a decreased use of gas for cooking, however, both these phenomena's occur at a pace smaller than 3,5%. Other factors which could be responsible for this decrease could be the changing calorific value of gas and/or the method for the calculation of standardized annual consumption.

Performance gap in relation to dwelling type, floor area and installation types

The analyses showed that floor area does not affect the performance gap strongly. In terms of dwelling type, semi-detached houses have the highest performance gap, followed by flats with a staircase entrance, detached houses and finally, gallery flats. The performance gap differed also in dwellings with different installation types. Dwellings with a local heater in the living room (gas stove) had the highest performance gap, followed by a combined boiler with $\eta<83\%$, and then each higher efficiency boiler had a smaller performance gap.

Energy reduction targets for built environment and actual reduction potential of the dwelling stock and of the individual dwelling renovation measures

Theoretical and actual achievability of the current targets

A scenario analyses was conducted in the third chapter. The baseline scenario was the scenario described in Covenant Energy Savings Housing Associations Sector' (Convenant Energiebesparing Corporatiesector, 2008), which aims is to save 20% gas consumption by 2018 by improving the dwellings to a B label or at least by 2 label classes. The refurbishment scenario of the mentioned agreement was one of the scenarios considered. Another, more radical refurbishment scenario was renovating the whole dwelling stock to label A. The two scenarios were tested on both baseline consumptions, actual and theoretical (Figure 4). It turned out that by using theoretical gas use as baseline, the least radical scenario is enough to ensure the potentials discussed in B.1 are fulfilled. However, if actual gas consumption is used as a baseline, most of these potentials seem unrealistic (exception is the 10% potential as defined by IDEAL project). This points to the fact that analysts as well as policy makers rely on theoretical gas consumption as a basis for future consumption estimates, which ultimately leads to unrealistic reduction targets and renovation plans.

Differences between the theoretical and actual reductions in dwellings where different renovation measures were applied

Longitudinal data of dwellings energy performance was used to identify renovated dwellings and analyse their energy consumption before and after the renovation. The results showed that most of the renovations are expected to yield larger reduction than what materialises, many times the realised saving is about half of the expected.

On average in all renovated dwellings, actual gas reduction is about a third lower than expected, however, there are big differences in the reductions of individual measures. Improvements in efficiency of gas boilers (space heating and hot tap water) yield the biggest energy reduction, followed by deep improvements of window quality. Improving the ventilation system yields a relatively small reduction compared to other measures, however, it is still much larger than theoretically expected. The measures achieving the most reduction are drastic improvements of window quality and an improvement of the efficiency of heating and hot tap water system (not a replacement of a local system). These are averages and the reductions for specific changes vary considerably. Measures that achieve an actual reduction higher that the theoretical seem to mostly be very modest improvements of insulation or window quality. Also notable is the underprediction of the reduction in dwellings where natural ventilation was replaced by mechanical exhaust and it is questionable whether such dwellings still have a sufficient quality of indoor air after the renovation.

Causes of the differences between actual and theoretical gas consumption

Explaining variation in gas use with dwelling, household and occupant characteristics

Regression based on socioeconomic data showed that explaining the actual gas consumption or the difference between the actual and theoretical with the publicly available variables yields a relatively low R^2 value (in view of existing literature these R values are not low) of 50,5% and 44,0%, respectively, meaning that 50,5% of the variance could be explained by these factors. Since our dataset contained many records, this relatively low explanatory power was thought to be due to the fact that many factor that do influence actual energy use, such as indoor temperature or presence of occupants, were not included. In the regression based on the survey data, these factors were included, but still not much more variation could be explained, probably due to a smaller sample size than was the case with socioeconomic data. The total R^2 values were 23,8% for actual gas use per m^2 as dependent variable and 40,9% for DBTA (difference between the theoretical and actual consumption) per m^2 as dependent variable. In both regression analyses, the majority of explanatory power for the DBTA came from dwelling characteristics. Household and occupant mattered less, although it was clear that the occupant behaviour data provided by the survey had a non-negligible predictive power for actual gas use per m^2 of 9,1%. The fact that dwelling characteristics dominate the performance gap emphasises the importance of the assumptions made in the calculation method.

Besides the regression analyses for the total sample, the model was tested on under and for overpredictions separately, since the hypothesis was that these two phenomenon would be explained by different variables. There was a large difference in the amount of variation that could be explained by all available variables in these two samples. In the underpredicted set of data 19,9% of variation could be explained by occupancy presence patterns, presence of a programmable thermostat and water saving shower head. On the other hand, in overpredictions as much as 50,8% of variation was accounted for by dwelling and installation type, age of the building, floor area, and indoor temperature. Furthermore, reported comfort was a significant predictor only in overpredictions.

The results demonstrate the difficulty of finding the right predictors for actual gas consumption. In the future both survey and socio demographic data could be combined to maximize the results, large samples should be used to ensure statistical significance and certain variables should probably be monitored in order to avoid survey bias. This includes variables like presence at home, indoor temperature an ventilation practices, since it seems that respondents might not be aware of their patterns well enough.

The relation between the performance gap and the normalised assumptions

Since the regression analyses did not cover the effect of variables such as indoor temperature, insulation quality, internal heat load etc. and this data was not available at that time, sensitivity of the theoretical calculation for certain parameters was conducted to fill this gap. Results showed, that an indoor temperature 2,7 degrees higher than assumed by the method currently (18 degrees) can explain the performance gap observed in label A and an indoor temperature 5,6 degrees lower than 18 degrees can account for the gap in label G. Both these temperature deviations are realistic, since people in well insulated dwellings probably heat their house more due to the small increment this causes in their monthly bill. Moreover, the installation system itself might be encouraging the occupants to heat more or less with for example low temperature floor heating installation in case of A labelled dwelling and with a local gas stove placed only in the living room in case of dwelling G. In the normalised calculation, all rooms are assumed to be heated. Heat resistance of the construction elements also had a big impact which demonstrates that in case of a poor inspection, the dwelling consumption could be very faulty due to an inaccurate estimation of insulation. This likely occurs in many old dwellings, where documentation is not available. Small increments in ventilation rates (up to 40% smaller or larger than current assumption) can also explain the performance gaps in label classes A to C. The two variables which had a smaller impact were the number of occupants and internal heat gains.

Longitudinal study confirmed the significant influence of insulation value by showing that the largest performance gaps appear in dwellings with poor envelope insulation, followed by those by poor window insulation. Considerable gaps appeared also in cases of heating installation of low efficiency.

A better model for theoretical gas consumption

Besides the exploratory regression analyses two other regression models were conducted in order to see whether the current theoretical consumption can be adapted with the new knowledge about the actual gas use. One model was made for under- and one for overpredicted consumptions. These models consisted of actual gas use as the dependent variable and theoretical gas consumption plus all other dwelling related features as predictors. The idea was to obtain the best possible theoretical consumption using only dwelling parameters so that the result could still be comparable among the dwellings. In the future, this could allow for determination of a more accurate dwelling consumption based only on dwelling parameters and average actual consumption data. For overpredictions, the model explained 33,8% of variation with installation and dwelling type being the significant variables (besides theoretical gas use). The explained variation was lower than for underpredictions, where it reached 60,0%, probably because the gap itself is much larger in overpredicted dwellings than in underpredictions.

The B coefficients obtained in these two models were then applied onto a different sample to see if a better predicted theoretical consumption could be obtained by adjusting the current theoretical use with the newly obtained parameters. The new theoretical consumption was indeed much closer to the actual gas use, which proves that this method could be used to obtain a better estimate of theoretical consumption.

Conclusion

There is a clear gap between actual and theoretical consumption in Dutch dwellings. Low performing dwellings tend to have a theoretical consumption much higher than actual, while high performing dwellings feature the opposite trend. These discrepancies are understandable at the level of individual dwellings and arise due to the standardizations made when calculating the theoretical consumption, however, on the level of the dwelling stock such a discrepancy is misleading and can lead to inaccurate policy reduction targets and sends wrong signals to several stakeholders (local governments, construction industry, renters and buyers etc.).

Regarding the causes of the discrepancies, they can party be explained by the features of the dwelling itself, meaning that the calculation model does not represent the reality accurately. However, a part of the discrepancy originates in the behaviour of the users and this part is difficult to quantify statistically. The results seem to indicate that underprediction is more difficult to explain and therefore probably more dependent on occupant practices than on the accuracy of the standardisation model. Overpredictions on the other hand, seem to have a lot in common with the fact that installation systems and the dwelling itself perform differently than expected. A methodological improvement seems to be more appropriate for the overpredicted cases while at the

TOC

same time tackling the fact that occupants of these dwellings are likely to feel cold. For underpredictions on the other hand, changes to the methodology would mean accepting that a higher heating intensity is inevitable in efficient dwellings. While this should be further researched in the future, behaviour incentives that would encourage people to use their homes more wisely and not waste energy could be more successful.

The label calculation is easy to use and can be, as shown in the thesis, a very valuable tool for following the energy efficiency of the dwelling stock. Since the accuracy of theoretical gas and electricity calculations can easily be improved, it is a pity to miss the opportunity to do so. Several recommendations for further research and policy development were proposed regarding the methodology for the calculation of theoretical consumption. Examples of this are a revision of several standardised factors, revision of method for determining the insulation values on-site and introduction of correction factors based on actual consumption statistics. Moreover, labels that are issued should be accurate and reliable, meaning that more attention should be paid to the quality of inspections and the robustness of the software used for label calculation.

This thesis demonstrated that research on the relationship between policy instruments and their effects is crucial to ensure the effectiveness and a continuous improvement of these tools. Theoretical models, such as energy labelling, are often used to support policy decisions. As was shown, such models do not always provide results that correspond to reality, and in the case of dwellings a big reason for this is disregarding the user, who seems to adapt to the thermal quality of the house itself. However, as was demonstrated, there is a clear need for a more accurate estimation of consumption on a broader, dwelling stock level in order to enhance the effectiveness of the current renovation policies. moreover, showed that a better estimation is feasible. The thesis showed that using the current knowledge and data availability, there is few reason not to reduce the performance gap and predict the dwelling consumption more accurately.

Samenvatting

Gebouwen vormen een van de belangrijkste sectoren waarop de energie- en CO_2-reductiedoelstellingen zich richten. Beleidsinspanningen in de afgelopen decennia hebben geleid tot technische verbeteringen in de woningsector, maar desondanks vertoont het energieverbruik in de woningsector nog niet de daling die noodzakelijk is om de gestelde doelen te behalen. Om het energieverbruik van woningen te verminderen heeft de Europese Commissie de Europese Richtlijn Energieprestatie Gebouwen (Energy Performance of Buildings Directive; EPBD) opgesteld. Volgens deze richtlijn moeten de lidstaten een energiecertificering voor bestaande gebouwen invoeren. Nederland heeft in 2008 het energielabel ingevoerd en baseerde de berekening van de theoretische energieprestaties van woningen op een gegeven set indicatoren. Het energielabel vermeldt de klasse (A++ tot G), het vloeroppervlak, het type en het bijbehorende standaard energieverbruik van het gebouw (exclusief apparaatgebruik), uitgedrukt in gas, elektriciteit, warmte en het totale primaire energieverbruik. Aangezien het theoretische gas- en elektriciteitsverbruik op het Nederlandse energielabel staat vermeld als 'standaard energieverbruik van de woning', werd het algemeen gebruikt als indicatie voor het verbruik en werd het zelfs opgenomen in beleidsdoelstellingen.

Aangezien de kwaliteit van woningen gestaag is verbeterd, moet de primaire oorzaak voor het niet realiseren van reductie gezocht worden in de toenemende energievraag van de gebruiker. Dit proefschrift stelt dat het gedrag van de gebruiker een cruciale factor is in het feitelijke energieverbruik en verantwoordelijk is voor tot wel 50% van de verschillen in energieverbruik voor verwarming. Om bestaand beleid te verbeteren en een daadwerkelijke reductie te realiseren, is meer inzicht nodig in het gedrag van de gebruiker en de invloed hiervan op het feitelijke energieverbruik.

Dit proefschrift belicht in hoeverre het theoretisch energieverbruik het werkelijke verbruik voorspelt, en wat de oorzaken zijn van discrepanties, met name waar het het gebruikersgedrag betreft. Of het genoemde reductiepotentieel wordt gerealiseerd is namelijk sterk afhankelijk van het werkelijke energieverbruik van de woning. Daarnaast onderzoekt het proefschrift welke reductie kan worden bereikt door de prestaties voor verwarming van de woningvoorraad te verbeteren en of deze al dan niet voldoen aan de verwachtingen van de beleidsmakers.

Onderzoeksmethoden

Bij het onderzoek is gebruikgemaakt van diverse grote datasets over de theoretische energieprestaties van woningen, waarvan de meeste gerelateerd waren aan energielabelcertificaten. Alle datasets met gegevens over theoretische prestaties zijn samengevoegd met gegevens over de werkelijke prestaties. Daarnaast zijn enkele datasets verrijkt met sociaal-economische en gedragsgerelateerde gegevens afkomstig van het Centraal Bureau voor de Statistiek (CBS) of afkomstig uit enquêtes die ten behoeve van dit onderzoek zijn opgesteld. Eenvoudige descriptieve gegevens zijn gebruikt om het gemiddelde theoretische en werkelijke verbruik te vergelijken. Voor het vaststellen van correlaties is geavanceerd statistisch onderzoek toegepast. Aansluitend zijn er diverse regressieanalyses uitgevoerd. In een afzonderlijk scenario-onderzoek zijn de gevonden gemiddelden van zowel theoretisch als werkelijk verbruik geëxtrapoleerd naar heel Nederland om een vergelijking te kunnen maken met de bestaande beleidsdoelstellingen.

Vanwege het geringe voorspellend vermogen van de variabelen in regressieanalyses, is een gevoeligheidsanalyse van het theoretisch gasverbruik uitgevoerd op basis van zes aannames in de theoretische berekening, om te laten zien hoe een toename in een van de aannames van invloed is op het uiteindelijke theoretische gasverbruik en of dit de discrepantie in energieprestaties kan verklaren.

Ten slotte zijn longitudinale gegevens over de sociale woningvoorraad tussen 2010 en 2013 geanalyseerd, speciaal gericht op gerenoveerde woningen. Het doel was na te gaan of en in welke mate de theoretische reductie in het energieverbruik was gerealiseerd. Een vergelijking is gemaakt met de werkelijke reductie bij verschillende renovatiemaatregelen om na te gaan welke renovaties het verbruik het meest effectief verminderen.

Discrepanties tussen theoretisch en werkelijk energieverbruik bij verwarming van Nederlandse woningen

Discrepanties tussen theoretisch en werkelijk gas- en elektriciteitsverbruik

Gemiddeld genomen lijkt het totale theoretische primaire energieverbruik in overeenstemming te zijn met het werkelijke primaire energieverbruik. Als de gegevens echter meer in detail worden bekeken, blijkt dat het aandeel gas in het werkelijke primaire energieverbruik veel lager is dan in het theoretische primaire energieverbruik en dat het aandeel elektriciteit in werkelijkheid juist veel hoger is dan in het theoretische primaire energieverbruik. De twee effecten heffen elkaar op, zodat op het totale primaire energieverbruik, het verbruik goed voorspeld lijkt te zijn. Verder bleek uit de analyses dat de verschillen in elektriciteitsverbruik tussen de verschillende energielabelklassen marginaal zijn. Dit en het feit dat de meeste Nederlandse woningen met gas worden verwarmd, heeft ons doen besluiten ons in de rest van het proefschrift uitsluitend op het gasverbruik te richten.

Hoewel het duidelijk is dat het theoretische elektriciteitsverbruik veel lager is dan het werkelijke verbruik, omdat er geen rekening is gehouden met het gebruik van apparaten, is het echter veel minder duidelijk waarom het werkelijke gasverbruik gemiddeld zo veel lager is dan het verbruik volgens de theoretische berekening.

Verschil in prestaties ten opzichte van het energielabel

De discrepanties in het gasverbruik waren het grootst in de slechtst presterende woningen, waar het theoretische verbruik bijna het dubbele was van het werkelijke verbruik. Dit duiden we ook wel aan met de term overschatting. Aan de andere kant verbruikten goed presterende woningen ongeveer 20% meer gas dan voorspeld. Het theoretische elektriciteitsverbruik was ten minste twee keer zo laag als het werkelijke verbruik in alle labelklassen, doordat bij het werkelijke verbruik ook het verbruik van apparaten is meegenomen en in het theoretische verbruik niet, zodat hier sprake is van onderschatting. Het werkelijke en theoretische elektriciteitsverbruik leek redelijk constant in de verschillende labelklasses. Primair energieverbruik is de som van het gas- en elektriciteitsverbruik in MJ voor elke labelklasse, waarbij ook rekening is gehouden met de efficiëntie van de elektriciteitsopwekking en van het netwerk, alsook de verbrandingswaarde van gas. De theoretische primaire energie bestaat grotendeels uit gasverbruik, aangezien het elektriciteitsverbruik door uitsluiting van huishoudelijke apparaten maar een klein deel uitmaakt van het primaire energieverbruik. De relatie tussen werkelijk en theoretisch verbruik is daardoor ongeveer gelijk aan die bij gasverbruik. Voor slecht presterende labelklassen is sprake van een overschatting van 30% bij het theoretisch verbruik en voor goed presterende labelklassen wordt het theoretische verbruik onderschat met ongeveer hetzelfde percentage.

Elektriciteitsverbruik lijkt niet samen te hangen met de energieprestaties van de woning. Bovendien verschilt het verbruik van elektriciteit in het werkelijke en theoretische verbruik dusdanig dat vergelijking geen zin heeft (aangezien bij het theoretisch verbruik apparaten niet zijn meegenomen). Daardoor ligt zoals eerder gezegd de focus van dit proefschrift op het gasverbruik. Gas is bovendien de belangrijkste brandstof voor het verwarmen van woningen in Nederland.

Verschil in prestaties in verschillende datasets

Verschillen in prestaties zijn geanalyseerd in vier afzonderlijke datasets van verschillende grootte. Alle datasets leverden zeer vergelijkbare resultaten op met betrekking tot het gemiddelde werkelijke en theoretische verbruik in de diverse labelklassen. Nadere analyse toont aan dat het werkelijke gasverbruik gestaag is afgenomen in de labelklassen A, E, F en G tussen 2010 en 2012. Het theoretische gasverbruik is in deze periode ongeveer gelijk gebleven, zodat het verschil in prestaties licht is toegenomen.

Daarnaast bleek dat woningen die niet zijn gerenoveerd en ongewijzigd zijn gebleven tussen 2010 and 2012 ook een daling van 3,5% in het gasverbruik vertoonden, hetgeen betekent dat de afname die in de vier onderzochte datasets is waargenomen, niet het gevolg is van selectievertekening (sample bias). Deze afname kan het gevolg zijn van verandering in de samenstelling van huishoudens (minder personen per huishouden) of een afname in het gebruik van gas voor koken. Beide doen zich echter in een geringere mate voor dan de gevonden 3,5%. Andere factoren die bij de afname een rol kunnen spelen zijn een verandering in de calorische waarde van gas en/of in de methode voor het berekenen van het standaard jaarlijkse verbruik.

Verschil in prestaties tussen verschillende woningtypen, vloeroppervlak en typen installaties

De analyses toonden aan dat vloeroppervlak niet van grote invloed is op het prestatieverschil. Als het gaat om het type woning, vertoonden twee-onder-een-kap-woningen het grootste verschil in prestaties, gevolgd door portiekwoningen, vrijstaande woningen en ten slotte galerijwoningen. Het verschil in prestaties varieerde ook tussen woningen met verschillende typen installaties. Woningen met een gaskachel in de woonkamer vertoonden het grootste prestatieverschil, gevolgd door woningen met een gecombineerde ketel met $\eta<83\%$, en vervolgens woningen met steeds efficiëntere HR-ketels met een steeds kleiner prestatieverschil.

Energiereductiedoelstellingen voor de gebouwde omgeving en het werkelijke reductiepotentieel van de woningvoorraad en de afzonderlijke woningrenovatiemaatregelen

Theoretische en werkelijke haalbaarheid van de huidige doelstellingen

In het derde hoofdstuk staat een scenario-analyse beschreven. Het basisscenario was het scenario zoals opgenomen in het Convenant Energiebesparing Corporatiesector (2008), met als doelstelling een besparing van 20% op het gasverbruik in 2018 door verbetering van de woningen naar een label B of een verbetering met ten minste twee labelklassen. Het renovatiescenario in het genoemde convenant was een van de scenario's die zijn bekeken. Een ander, meer drastisch renovatiescenario betrof renovatie van de hele woningvoorraad naar label A. De twee scenario's zijn getoetst aan de hand van de werkelijke en theoretische baseline-verbruiksgegevens (figuur 4). Hieruit bleek dat bij het theoretisch gasverbruik als uitgangspunt, het minst radicale scenario voldoende is om de potentiële besparingen zoals besproken in B.1 te behalen. Als echter het werkelijke gasverbruik als uitgangspunt wordt genomen, lijken de meeste potentiële reducties niet realistisch (met uitzondering van het

TOC

reductiepotentieel van 10% zoals aangegeven in het IDEAL-project). Dit duidt erop dat zowel analisten als beleidsmakers zich baseren op het theoretische gasverbruik bij voorspellingen van toekomstig verbruik. Dit leidt uiteindelijk tot onrealistische reductiedoelstellingen en renovatieplannen.

Verschillen tussen de theoretische en werkelijke reductie in woningen waar verschillende renovatiemaatregelen zijn toegepast

Voor de selectie en analyse van gerenoveerde woningen zijn de longitudinale gegevens van de energieprestaties van de woningen van voor en na de renovatie gebruikt. Uit analyse blijkt dat van de meeste renovaties een grotere besparing werd verwacht dan in werkelijkheid werd bereikt. Vaak bleek de gerealiseerde besparing maar de helft van de verwachte besparing. Gemiddeld bleek bij alle gerenoveerde woningen dat de werkelijke besparing op het gasverbruik ongeveer een derde lager lag dan verwacht. Er zijn echter grote verschillen in besparing tussen de verschillende maatregelen. Verbetering van de efficiëntie van combiketels (verwarming en heetwatervoorziening) leveren de grootste besparing op, gevolgd door grote kwaliteitsverbeteringen in beglazing. Verbetering van ventilatiesystemen levert een relatief kleine besparing op in vergelijking met de andere maatregelen, maar deze is altijd nog wel groter dan in theorie werd verwacht. De meeste besparing wordt behaald door drastische verbetering van de kwaliteit van de beglazing en verbetering van de efficiëntie van de verwarming en het warmwatersysteem (geen vervanging van een lokaal systeem). Dit zijn gemiddelden en de besparingen voor specifieke verbeteringen verschillen aanzienlijk. Maatregelen waarbij de werkelijke besparing hoger is dan de theoretische zijn doorgaans bescheiden verbeteringen in isolatie of kwaliteit van beglazing. Ook opmerkelijk is de onderschatting van de besparing bij woningen waar de natuurlijke ventilatie werd vervangen door mechanische ventilatie. De vraag is wel of de luchtkwaliteit van dergelijke woningen na renovatie nog voldoet.

Oorzaken van het verschil tussen werkelijk en theoretisch gasverbruik

Verklaring van verschillen in gasverbruik aan de hand van woning-, huishouden- en bewonerskenmerken

Regressieanalyse op basis van sociaal-economische gegevens toonde aan dat een verklaring van het werkelijke gasverbruik of het verschil tussen het werkelijke en theoretische gasverbruik met de publiek toegankelijke variabelen een relatief lage R^2-waarde opleverde van respectievelijk 50,5% en 44,0% (in het licht van de bestaande literatuur zijn deze R-waarden niet laag). 50,5% van de variantie kan dus door deze factoren verklaard worden. Aangezien onze dataset veel gegevens bevatte, leek het aannemelijk dat deze relatief lage verklarende waarde het gevolg was van het feit dat veel factoren die van invloed zijn op het werkelijke energieverbruik, zoals de binnentemperatuur of de aanwezigheid van bewoners, hierin niet zijn meegenomen. In de regressieanalyse gebaseerd op enquêtegegevens werden deze factoren wel meegenomen, maar nog steeds kon niet veel meer variantie worden verklaard, waarschijnlijk als gevolg van een kleinere steekproef dan bij de sociaal-economische gegevens. De totale R^2-waarden waren 23,8% voor het werkelijk gasverbruik per m^2 als afhankelijke variabele en 40,9% voor het verschil tussen theoretisch en werkelijk verbruik (DBTA; Difference Between Theoretical and Actual) per m^2 als afhankelijke variabele. In beide regressieanalyses werd het grootste deel van het verklarend vermogen voor het DBTA veroorzaakt door woningkenmerken. Huishouden en bewoners waren minder van invloed, hoewel het bewonergedrag dat uit de enquête naar voren kwam een niet verwaarloosbaar voorspellend vermogen had voor het werkelijke gasverbruik per m^2 van 9,1%. Het feit dat woningkenmerken de grootste rol spelen in het prestatieverschil, maakt het belang van de aannames die bij de berekeningsmethoden worden gebruikt, eens te meer duidelijk.

Naast de regressieanalyses voor de complete dataset, is het model ook afzonderlijk getest op onder- en overschatting, aangezien de hypothese was dat deze twee fenomenen door verschillende variabelen worden verklaard. Er was een groot verschil in de hoeveelheid variatie die door alle beschikbare variabelen in de twee steekproeven kon worden verklaard. In het geval van onderschatting kon 19,9% van de variatie worden verklaard door aanwezigheid van bewoners, een programmeerbare thermostaat en een waterbesparende douchekop. Bij overschatting daarentegen, kon tot 50,8% van de variatie worden verklaard door woning- en installatietype, ouderdom van het gebouw, vloeroppervlak en binnentemperatuur. Bovendien bleek gerapporteerd comfort alleen een significante voorspeller bij overschatting.

Deze resultaten laten zien aan hoe moeilijk het is de juiste voorspellers voor het werkelijke gasverbruik te vinden. In de toekomst zouden enquêtegegevens en sociaal-demografische gegevens kunnen worden gecombineerd om een optimaal resultaat te behalen. Daarnaast moet een grote steekproef worden gebruikt om statistische significantie te bereiken en bepaalde variabelen moeten waarschijnlijk worden gecontroleerd om vertekening bij enquêtegegevens te vermijden. Het betreft hier variabelen zoals aanwezigheid in de woning, binnentemperatuur en ventilatiegewoonten, aangezien het erop lijkt dat bewoners zich mogelijk onvoldoende bewust zijn van deze gedragspatronen.

De relatie tussen het prestatieverschil en de genormaliseerde aannames

Aangezien in de regressieanalyses niet het effect van variabelen als binnentemperatuur, isolatiekwaliteit, interne warmtelast, etc., was meegenomen en deze informatie op dat moment niet voorhanden was, is een gevoeligheidsanalyse van de theoretische berekening uitgevoerd om in deze omissie te voorzien. Hieruit komt naar voren dat een verhoging van de binnentemperatuur met 2,7 graden ten opzichte van de veronderstelde temperatuur van 18 graden in de berekening, het prestatieverschil in label A kan verklaren en een binnentemperatuur van 5,6 graden lager dan 18 graden het verschil in label G kan verklaren. Beide temperatuurafwijkingen zijn realistisch, aangezien bewoners van goed geïsoleerde woningen hun huis waarschijnlijk meer verwarmen omdat dit slechts tot een kleine verhoging in hun energierekening leidt. Bovendien is het mogelijk dat het installatiesysteem zelf uitnodigt tot meer of minder verwarming van de woning, met bijvoorbeeld lage-temperatuur vloerverwarming bij een woning met label A of met een kachel in alleen de woonkamer bij een woning met label G. In de genormaliseerde berekening wordt ervan uitgegaan dat alle kamers zijn verwarmd. De isolatiegraad van de woning was ook van grote invloed, zodat bij een onjuiste bepaling hiervan ook het energieverbruik van de woning onjuist wordt bepaald. Dit doet zich vooral voor bij oudere woningen, waarvan niet alle documentatie voorhanden is. Verschillen in de ventilatiegraad (tot 40% lager of hoger dan verondersteld) kunnen ook een verklaring zijn voor de prestatieverschillen in labelklassen A tot C. De twee variabelen met de minste invloed zijn het aantal bewoners en het voordeel van interne warmtelast.

Longitudinaal onderzoek bevestigde de significante invloed van de isolatiewaarde waarbij het grootste prestatieverschil zich voordoet in woningen met een slechte omhulsel(envelop)isolatie, gevolgd door slechte isolatie van ramen. Aanzienlijke verschillen werden ook gevonden bij minder efficiënte verwarmingsinstallaties.

Een beter model voor theoretisch gasverbruik

Naast de verkennende regressieanalyses zijn twee andere regressiemodellen uitgevoerd om na te gaan of het huidige theoretische verbruik kan worden geoptimaliseerd met de nieuwe kennis over het werkelijke gasverbruik. Het ene model is ontworpen voor onderschatting van het verbruik en het andere voor overschatting. Deze modellen omvatten het werkelijke gasverbruik als afhankelijke variabele en het theoretische gasverbruik en alle andere woning gerelateerde kenmerken als voorspellende variabelen. Het idee was om het optimale theoretische verbruik te verkrijgen op basis van alleen woningparameters, zodat het resultaat nog steeds vergelijkbaar zou zijn tussen de verschillende woningen. In de toekomst zou hiermee een accurater energieverbruik kunnen worden bepaald aan de hand van woningparameters en gemiddelde werkelijke gebruiksgegevens. In het geval van overschatting verklaarde het model 33,8% van de variatie, met installatie- en woningtype als de meest significante variabelen (naast theoretisch gasverbruik). De verklaarde variatie was lager dan bij onderschattingen, waar 60,0% werd behaald, waarschijnlijk omdat het verschil zelf veel groter is bij overschatte woningen dan bij onderschatte woningen.

De B-coefficiënten die in deze modellen werden verkregen zijn vervolgens toegepast op een andere dataset om na te gaan of een beter voorspeld theoretisch verbruik kon worden verkregen door het huidige theoretische verbruik aan te passen aan de hand van de nieuw verkregen parameters. Het nieuwe theoretische verbruik lag inderdaad veel dichter bij het werkelijke gasverbruik, wat aantoont dat deze methode kan worden gebruikt om een betere inschatting te maken van het theoretisch verbruik.

Conclusie

Er is een duidelijk verschil tussen het werkelijke en theoretische energieverbruik in Nederlandse woningen. Laag presterende woningen hebben doorgaans een theoretisch verbruik dat veel hoger ligt dan het werkelijke verbruik, terwijl dit bij hoog presterende woningen andersom is. Deze verschillen zijn verklaarbaar voor individuele woningen en zijn het gevolg van standaardisaties bij de berekening van het theoretisch verbruik. Voor de hele woningvoorraad is een dergelijk verschil echter misleidend. Bovendien kan het tot onnauwkeurige reductiedoelstellingen leiden en het verkeerde signaal afgeven aan de verschillende stakeholders (lokale overheden, bouwindustrie, huurders, kopers, etc.).

De oorzaken van de discrepanties kunnen deels verklaard worden door de kenmerken van de woning zelf, waarbij het rekenmodel de werkelijkheid niet accuraat weergeeft. Een deel van de discrepantie wordt echter ook veroorzaakt door het gedrag van de gebruikers – iets wat moeilijk statistisch te kwantificeren is. De uitkomsten lijken erop te wijzen dat onderschatting moeilijker te verklaren is en daardoor waarschijnlijk eerder bepaald wordt door bewonersgedrag dan een onjuist standaardisatiemodel. Overschattingen daarentegen lijken eerder veroorzaakt te worden doordat installatiesystemen en de woning zelf anders presteren dan verwacht. Verbetering van de methodologie lijkt meer voor de hand te liggen voor de overschatte gevallen. Het feit dat de bewoners van deze woningen het waarschijnlijk koud hebben, zal als gegeven moet worden geaccepteerd. Voor onderschattingen daarentegen zou aanpassing van de methodologie inhouden dat een hogere verwarmingsintensiteit als onvermijdelijk moet worden gezien voor energie-efficiënte woningen. Hoewel verder onderzoek in de toekomst nodig zal zijn, zouden prikkels die de bewoners aansporen hun woning slimmer te gebruiken en geen energie te verspillen effectiever kunnen zijn.

De labelberekening is gemakkelijk toe te passen en kan, zoals in het proefschrift wordt aangetoond, een zeer waardevol hulpmiddel zijn om de energie-efficiëntie van de woningvoorraad te volgen. Aangezien de nauwkeurigheid van de berekening van het theoretisch gas- en elektriciteitsverbruik eenvoudig kan worden verbeterd, zou het jammer zijn om deze mogelijkheid niet te benutten. Er zijn verschillende aanbevelingen gedaan voor verder onderzoek en beleidsontwikkeling betreffende de methodologie voor de berekening van het theoretisch verbruik. Voorbeelden hiervan zijn herziening van diverse gestandaardiseerde factoren, herziening van de methode om de isolatiewaarden ter plaatse te bepalen en introductie van correctiefactoren op basis van de werkelijke verbruiksgegevens. De labels die worden uitgegeven zouden nauwkeuriger en betrouwbaarder moeten zijn. Dit houdt in dat meer aandacht zou moeten uitgaan naar de kwaliteit van de inspecties en de software die voor de labelberekening wordt gebruikt.

Dit proefschrift toont aan dat onderzoek naar de relatie tussen beleidsinstrumenten en hun effecten cruciaal is voor de effectiviteit en continue verbetering van deze hulpmiddelen. Theoretische modellen, zoals energielabeling, worden vaak gebruik om beleidsbeslissingen te onderbouwen. Zoals is aangetoond, komen dergelijke modellen niet altijd overeen met de werkelijkheid. Bij energieverbruik in woningen is de belangrijkste reden hiervoor dat geen rekening wordt gehouden met de bewoners die zich lijken aan te passen aan de kwaliteit van het huis. Om echter de effectiviteit van het huidige renovatiebeleid te vergroten, is de noodzaak voor een accuratere inschatting van het verbruik voor de woningvoorraad als geheel, duidelijk gebleken. Bovendien blijkt dat een betere inschatting ook haalbaar is. Het proefschrift toont aan dat met gebruikmaking van de huidige kennis en beschikbare gegevens het prestatieverschil kan worden verkleind en het energieverbruik voor woningen beter kan worden voorspeld.

1 Introduction

§ 1.1 EPBD and energy use in dwellings

Buildings, residential structures in particular, consume about one-sixth of the total amount of energy used in Europe, and in the Netherlands about two-thirds of this is used for space heating. Policies and regulations to reduce the heating consumption have been formulated nationally and at the EU level. The Energy Performance of Buildings Directive (EPBD), which was first introduced in 2002, demands that member states establish minimum energy performance standards for new construction as well as a certification scheme for existing buildings. Since new construction has a marginal impact in terms of annual energy consumption (Yücel and Pruyt, 2011), certification (also called energy labelling) is regarded as an important tool in reducing the energy consumption of existing dwellings. The certification programme was inspired by the well-recognised directive for labelling the energy use of appliances instituted in 1992 (92/75/EEC). It uses the same A-G scale as an indicator of dwelling performance and requires that dwellings have a certificate when sale or rental transactions take place. The energy label certificate is issued by qualified inspectors who are licensed to carry out an inspection and label calculation. European member states were supposed to develop a label calculation methodology themselves according to a set of standards defined by the EU (see Chapter 2). By informing a potential buyer or tenant of a dwelling's energy efficiency, the directive is expected to accelerate renovation activities in the existing dwelling stock in order to reduce overall energy consumption.

The directive specifies that renovations should take place in a cost-effective way, ensuring that the savings achieved surpass the investment necessary within the lifetime of the newly installed component. The cost-optimal level is the 'energy performance level which leads to the lowest cost during the estimated economic lifecycle' (article 2.14 EPDB 2010/31/EU); however, a crucial point here is the level of energy consumption used as a baseline. To foster true cost-optimal measures, the calculation should be made using realistic levels of consumption.

With other EU member states, the Netherlands implemented energy labelling in 2008. It based the theoretical calculation of residential energy performance on a steady state method, assuming the same indoor conditions in all labelled dwellings for the sake of comparability. In addition to an indication of the label category (A++ to G), the Dutch energy certificate takes into account the floor area, the type of dwelling and building-related energy use (excluding appliances); more specifically, the consumption of gas [m^3], electricity [kWh], heat [GJ] and the total primary energy consumption [MJ]. The label categories are determined by using the energy index, which is calculated on the basis of total primary energy usage, summing up the primary energy required for heating, hot water, pumps/ventilators and lighting, and subtracting any energy gains from photovoltaic (PV) cells and cogeneration. Although the calculation includes a correction for the shape factor of the dwellings to correct for the dwelling type, it also contains many standardised values which are supposed to ensure the comparability of certificates, such as the efficiencies of boilers and distribution systems, transmission rates, assumptions about heated areas, heat gains and losses, standard heated floor area, number of occupants, normalisations of climate (indoor and outdoor), etc. In practice, many such assumptions may unrealistically account for certain thermal performance levels: for example, assuming that an entire dwelling is heated even if the dwelling only has a heater in the living room. These normalisations enable the cross-comparison of different dwellings at equivalent comfort levels, but may come at the expense of realistic estimates of theoretical energy consumption. A desire for comparability is understandable, but not at the cost of realistic use assumptions, if these turn out to relate strongly with dwelling performance.

Since this consumption has been displayed on the front page of the Dutch energy label certificate as 'standard energy use of the dwelling' (Figure 1) it has become widely accepted as a realistic consumption level and has even been included in policy goals (Chapter 2).

The Dutch energy label provides the following information on dwellings for the consumer: the label category (A++ to G), the floor area, the type of dwelling, the consumption of gas [m^3], electricity [kWh], heat [GJ] and the total primary energy consumption [MJ]. The label categories are determined using the energy index, which is calculated on the basis of total primary energy usage, summing up the primary energy required for heating, hot water, pumps/ventilators and lighting, and subtracting any energy gains from PV cells and/or cogeneration (ISSO, 2009). If no additional heat is consumed (from district heating for example), the total primary energy consumption (Q_{total}) can be expressed as described in Equation 1. The level of primary energy consumption is calculated according to the type of fuel used by the installations in the dwelling (Equation 2 and Equation 3). Since primary energy is a form of energy that is found in nature and has not been subject to any conversion or transformation process, appropriate heating values need to be taken into account when calculating its consumption. The assumed heating value for gas is 35.17MJ/m^3 (North Sea gas). The efficiency of the electricity network is considered to be 0.39. The theoretical gas and electricity consumption from Equation 2 are noted on the Dutch label certificate.

$$Q_{total}[MJ] = Q_{total\ gas}[m^3] \cdot 35.17 \left[\frac{MJ}{m^3}\right] + Q_{total\ el.}[kWh] \cdot 3.6 \left[\frac{MJ}{kWh}\right] : 0.39$$

EQUATION 1

$$Q_{total\ gas} = Q_{gas-\ space\ heating} + Q_{gas-tap\ water}$$

EQUATION 2

$$Q_{total\ el.} = Q_{el.-\ space\ heating} + Q_{el.-tap\ water} + Q_{el.-\ aux.energy} + Q_{el.-lighting} - Q_{el.-\ pv} - Q_{el.-\ cogeneration}$$

EQUATION 3

The energy used for heating depends on the demand for space heating, the efficiency of the distribution system and the efficiency of the heating installation equipment. The demand for space heating is the sum of losses through transmission and ventilation, taking into account solar and internal heat gains. The energy consumption of hot tap water takes into account the main hot water installation and the auxiliary kitchen boiler. Details of the space heating and hot tap water demand are available in Majcen et al., 2013b.

The energy index (EI, Equation 4) correlates directly with total primary energy consumption, but is corrected for the floor area of the dwelling and the corresponding heat transmission areas in order not to disadvantage larger dwellings and dwellings that have a greater proportion of their heat envelope adjoining unheated spaces (different building types) with constant insulation properties and efficiencies of the heating/ventilation/lighting system.

$$EI = \frac{Q_{total}}{155 \cdot A_{floor} + 106 \cdot A_{loss} + 9560}$$

EQUATION 4

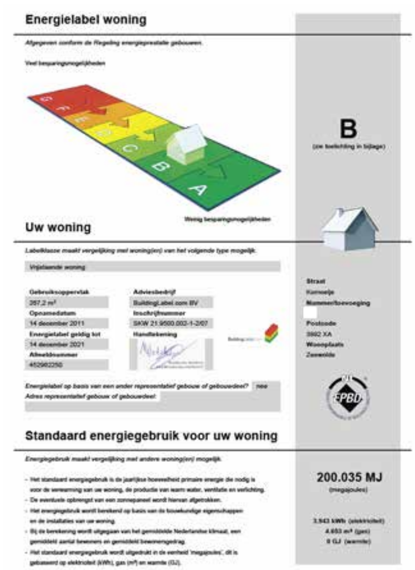


FIGURE 1 The Dutch energy label certificate

However, this was not the intention of the certificate's developers who have long stressed that the label certificate was only a tool meant to distinguish higher performing dwellings from those with lower performance. While it is clear that there will always be a variation in the actual amount of energy used in identical dwellings due to the fact that consumption levels are largely determined by users, for a broader level of the dwelling stock the average theoretical consumption should coincide with actual consumption. That is, whether a worse performing dwelling is consuming twice or four times more than a higher performing one determines the savings potential of renovation. When actual consumption differs significantly from the theoretical, the actual reduction in consumption may also differ from the theoretical in an absolute and a relative sense, leading to inaccurate estimations of cost efficiency and pay-back time. This thesis attempts to shine a light on how well theoretical consumption predicts actual consumption rates, the causes of discrepancies and the consequences for policy. It also characterises the actual consumption of Dutch dwellings at stock levels for the first time and attempts to propose measures for the improvement of the current situation.

§ 1.2 Scientific relevance

The EPBD directive was implemented across Europe by the end of 2009 and the process seems to have been well studied within the context of EU projects and the

TOC

EPBD Concerted Action initiative (Majcen et al., 2013a). However, an evaluation of the actual effects of the enforced certification has been hindered by the lack of publicly accessible databases (Perez et al., 2008) containing the information on label certificates on one hand and information about the actual energy consumption of the individual dwellings on the other. The studies that have so far been carried out have indicated a discrepancy between the actual and theoretical consumption rates of dwellings, in the Netherlands as well as elsewhere in Europe (Laurent et al, 2013). Recent studies by Cayre et al. (2011) in France, Hens et al. (2010) in Belgium, Sharpe and Shearer (2013) in Scotland and Guerra Santin (2010) in the Netherlands all showed that actual energy consumption levels were lower in reality than had been predicted theoretically in poor performing dwellings. The better the dwelling performs the smaller the difference between theoretical and actual energy consumption levels. However, in very well performing dwellings, actual energy consumption can be higher than theoretical levels. For example, Haas and Biermayr (2000) in Austria and Branco et al. (2004) in Switzerland showed that theoretical energy consumption rates in higher performing dwellings tend to be underpredicted, meaning, lower than is actually used. The disparity between the energy use predicted by the calculation model (theoretical consumption) and the energy use of those buildings in operation (actual consumption) is also referred to as the performance gap (de Wilde, 2014). A study by Pettersen (1994) showed that total heating energy consumption cannot be predicted more precisely than approximately 35-40%, which corresponds with the case study by Majcen et al. (2013b) and the other previously mentioned studies.

This gap may arise due to various uncertainties, which Ramallo-González (2013) classifies into three groups: environmental, workmanship and behavioural. Environmental relates to the climatic conditions being different than those assumed, which in this thesis is accounted for by correcting the calculated theoretical consumption to the actual number of degree days in a given year. The workmanship factor means that the performance of the dwelling components differs from what is documented. For example, the actual consumption differs from the theoretical if the execution of a renovation is sloppy or the installation systems underperform (the theoretical efficiency may be based on operation in laboratory conditions). Another example of workmanship is the quality of the inspection and calculation in the labelling process. The last group of uncertainties involved in the gap is related to the behaviour of the occupants and includes such variables as the indoor temperature settings, ventilation practices, showering and cooking habits. Additional complexity occurs because behaviours correlate with several parameters, such as dwelling characteristics, income, education, etc. For example, presence patterns and comfort correlate to energy performance itself (Gill et al., 2010, Guerra Santin, 2010, Haas et al., 1998). The fact that these parameters vary across different dwelling types and possibly relate to dwelling performance (especially comfort) is ignored in performance certification, most likely in order to ensure comparability. If differently performing dwellings are characterised by intrinsically different use practices and behaviours, disregarding these leads to incorrect estimates.

A widely researched example of how building use is related to performance is the so-called 'rebound effect', by which more efficient technologies cut energy bills but thereby encourage increased consumption, an effect that can take up to 30% of efficiency gains (Sorrell et al., 2009, Greening et al., 2000, Milne and Boardman, 2000). This is partly responsible for the phenomenon of underestimated theoretical consumption in high-performance dwellings. Many further studies address the correlations between actual energy use and potential influencing factors:

– Dwelling characteristics: Lindén et al. (2006), Guerra Santin et al. (2009), French et al. (2007), de Groot et al., (2008), Guerra Santin et al. (2010), Shipworth et al. (2009), Raynaud (2014)
– Household characteristics: Sardianou (2008) and Oreszczyn et al. (2006)
– Occupant characteristics: Guerra Santin (2010), Gill et al. (2010), Haldi and Robinson (2011)
– Occupant comfort: Hong et al., (2009), Ioannou and Itard (2015)

Further descriptions of these relationships are available in Chapter 4. All the above-mentioned factors need to be better understood in order to reduce the gap and will be discussed in the thesis. However, it is important to keep in mind that the performance gap, the main subject of the thesis, is caused as much by the factors influencing actual energy consumption as it is instigated by the calculation model itself. Unrealistic normalisation assumptions cause the theoretical consumption calculations to be severely flawed. As an example, the Dutch methodology assumes an indoor temperature of 18 degrees over the whole floor area during the entire heating period, while many older Dutch dwellings lack a heating unit in the bedrooms and cannot possibly maintain such a temperature over the winter. The standpoint of this thesis is that the current way in which the model represents reality is inaccurate and can be vastly improved by understanding the influencing parameters, which is why correlations between the above-mentioned parameters and their role in the theoretical calculation need to be thoroughly studied. As previously stated, a scientific model that does not accurately predict the energy consumption of dwellings at the stock level does not constitute a proper policymaking tool, since the actual effect of the improvement of the stock will on average be much lower than the predicted effect.

§ 1.3 Policy targets in the Netherlands

The policy targets for energy reduction are an important background to this research work, since the energy label is one of the main policy tools expected to lead to a reduction in the energy consumption of the existing dwelling stock. This thesis attempts to evaluate whether the current targets are realistic. At the national level, the Dutch federation of housing associations (Aedes) has adopted the 'Covenant Energy Savings Housing Associations Sector' (Convenant Energiebesparing Corporatiesector, 2008), which commits it to saving 20% on the consumption of natural gas (which is the main source of energy used to heat buildings in the Netherlands) in the existing social housing stock between 2008 and 2018. This agreement aims to achieve the set reductions by improving the dwellings for at least two label classes or until label B is reached. This implied a very high refurbishment rate that was of questionable feasibility; however, at the time the first paper was written this was the agreement currently in place. Later in 2012, a new target of 110PJ by 2020 was set by a new agreement (Koepelconvenant energiebesparing gebouwde omgeving, 2012), covering both residential and non-residential buildings as well as existing and new construction. Comparing these two targets reveals that the ambitions are now less focussed and apply more generally to the whole sector (also private housing and non-residential structures) which makes it impossible to estimate whether or not they will be attained within the timeframe of this research project, since the non-residential sector is not within the scope of this thesis. According to ECN (Energy research centre of the Netherlands) and as a result of publications by Majcen et al. (2014) and Fillipidou and Nieboer (2014), who were involved in the preparation of this new agreement, the new target is based on actual consumption data.

§ 1.4 Problem definition

It seems that the Dutch built environment is lagging behind other nations in improving its sustainability. Even though the energy efficiency of Dutch housing stock improved by almost a third since the 1990s (mainly due to the introduction of condensing boilers), the household use of gas dropped by a mere 5% (Majcen et al., 2013a). Looking at primary energy use the picture is even grimmer, as electricity consumption grew by 50% in the same period. Since new construction constitutes a marginal part of total dwelling stock, the focus of future energy reductions should be existing dwellings. The energy reduction potential of the built environment is discussed in several European and Dutch reports, setting a 20% rule of thumb by 2020 (Majcen et al., 2013a). Whether this is achievable strongly depends on the actual energy consumption of the dwellings. Therefore, this thesis explores which reductions can realistically be achieved by improving the thermal performance of the dwelling stock and whether or not these match the expectations of policy makers.

Existing studies of the performance gap (mentioned in 1.2) all have certain limitations. As previously stated, energy performance databases are difficult to access or do not exist and actual consumption data is not easy to obtain. Consequently, existing studies were based on small samples and were not truly representative. The discrepancies between the actual and theoretical energy consumption of labelled dwellings had, at the time of the study, not yet been studied in Dutch labelled dwelling stock. Therefore, the overarching idea in this thesis was to study the actual energy performance of the dwelling stock and to compare it with the theoretical performance in order to find out whether a performance gap exists. Furthermore, the extent of the performance gap is studied along with the factors that correlate with it. Such knowledge will enable better prediction of savings potential in the future.

§ 1.5 Research questions

The main focus of this work is the discrepancies between actual and theoretical energy consumption, the factors that cause them and their consequences for existing policies and existing energy reduction targets. In addition, recommendations have been developed on the basis of the insights gained and are presented in the conclusion chapter. The main research question of the thesis can therefore be summarised as follows:

What are the characteristics and consequences of the discrepancies between actual and theoretical heating energy use in Dutch dwellings?

The chapters of the thesis have been compiled chronologically, in terms of analytical work as well as their publishing timeline. Each chapter is essentially a journal article, either published (Chapters 2 to 4) or submitted (Chapter 5). The sixth and final chapter contains the overall conclusions of the thesis and summarises the answers to the research questions.

The research began with the initial idea to look into the extent and characteristics of discrepancies and what they mean for energy savings policies (Chapter 2). It became clear that the discrepancies were quite significant in some performance categories (more than a factor of 2). This also meant there would be a substantial impact on the energy savings targets set by the government, which were evaluated using a scenario study described in Chapter 2. The next logical step was to find out why the discrepancies occur and what can be done to reduce them (Chapters 3 and 4). In these chapters, the correlations between a large array of variables (described in 1.2) are evaluated in relation to actual and theoretical gas consumption. Chapter 5 provides an assessment of the actual energy reduction potential of various renovation measures, which also provides useful insights into the causes of the gap, particularly in terms of the standardisations used in the calculation method. There are some overlaps as well as synergies among the chapters, for example the first research question is answered in each chapter using a different dataset. The research questions arising from the main question are described as A, B and C.

A The discrepancies between actual and theoretical heating energy consumption in Dutch dwellings

The main goal of this section was to analyse the discrepancies in the total dwelling stock as well as across the label categories. Namely, previous research on smaller samples indicated over-predictions in lower performing dwellings and underprediction in energy efficient dwellings. Sub-question A.1 deals with consumption across the total stock and analyses gas, electricity and also total primary energy consumption. This was also done in Chapter 2, however there we show that the variation in electricity consumption is marginal across label categories. This together with the fact that most Dutch dwellings are heated with gas made us focus exclusively on gas consumption in the following sections as well as in Chapters 3, 4 and 5.

Sub-question A.2 – how the performance gap differs across different label categories – is one of the central topics of the thesis and is therefore discussed in all four thesis chapters. This sub-question discusses the relation between gas and electricity in different label categories since, as mentioned before, the rest of the thesis focuses on gas consumption only. Each chapter analyses a different sample, which ensures robust and highly representative results. Moreover, the datasets come from different years, which enables a longitudinal analysis of the trends in actual or theoretical gas consumption. A comparison of findings in different datasets (and years) is discussed in sub-question A.3. Sub-question A.4 analyses the performance gap in relation to other dwelling properties, such as dwelling type, floor area or installation.

A.1 What are the discrepancies between theoretical and actual gas and electricity consumption in the total dwelling stock?
A.2 What is the relation between actual and theoretical gas/electricity/primary energy/CO_2 emissions in dwellings with different energy labels?
A.3 Is the performance gap different among the studied samples and throughout the years?
A.4 How does the performance gap correlate with dwelling properties such as dwelling type, floor area and installation types?

B Energy reduction targets for the built environment and the actual reduction potential of the dwelling stock and of the individual dwelling renovation measures

The objective in this section was to analyse the current existing targets and compare them with the theoretical as well as the actual consumption of dwelling stock on the basis of the label data. The targets were reviewed in the second chapter, and since some of them changed later on, Chapter 4 contains some updates. Using scenario analyses, the existing targets are compared with the modelled potential (B.1). The results help to estimate whether or not Dutch dwelling stock is on a good path toward achieving the set energy reduction and CO_2 targets.

Besides evaluating the reduction in the total dwelling stock, this section also provides an analysis of the savings potential of different renovation measures at the dwelling level to see whether there is a big difference between the potential as calculated by the labelling methodology and the reduction achieved in reality after a dwelling is refurbished. Question B.2 is answered in Chapter 5 using large-scale longitudinal data of residential energy performance, in which the energy data before and after renovation are compared.

B.1 Are the current policy targets theoretically as well as actually achievable?
B.2 What are the differences between theoretical and actual reductions in dwellings where different renovation measures were applied?

C What causes the differences between actual and theoretical gas consumption?

As mentioned in section A, further analyses of energy consumption were narrowed down to gas use. Significant discrepancies were found (A.1) which in turn have a significant effect on current policies (B.1) and the next logical step was to find out why the discrepancies occur. Sub-question C.1 regarding the relative contributions of dwelling, household and occupant behaviour characteristics to the performance gap was answered in Chapter 3 as well as Chapter 4 of the thesis. The two chapters used different data (RVO data in Chapter 3 and Rekenkamer and WOON data in Chapter 4) and also a different methodological approach. In Chapter 3 we used a simple regression analysis of a larger database with fewer input variables and in Chapter 4 we used advanced regression in several subsamples and a larger number of predictors. In addition to the regression analysis of the data, Chapter 3 included a sensitivity analysis of the assumptions made in the theoretical calculations, which together with Chapter 5 studying the renovated dwellings provided an answer to question C.2. Finally, an attempt is made in Chapter 4 to develop an improved method for predicting theoretical consumption, relating to sub-question C.3.

C.1 How much of the variation in actual and theoretical gas use can we explain using dwelling, household and occupant behaviour characteristics?
C.2 What is the relation between the performance gap and the normalised assumptions made in calculation models?
C.3 Can a better model be obtained by using the available actual consumption data?

§ 1.6 Research outline and methods

This research studies the difference between actual and theoretical energy consumption in Dutch residential dwelling stock. The research utilised several large datasets, described in 1.7 and Table 1, about the theoretical energy performance of dwellings, most of which were related to energy label certificates. The one exception was the largest dataset – SHAERE, which contains data similar to that on the certificates but which was registered only by housing associations rather than by Dutch government authorities. All the datasets containing theoretical performance were merged with actual energy data. In addition, some were enriched with socioeconomic and behaviour-related data from Statistics Netherlands (CBS) or from surveys designed for the purpose of this research. Several statistical approaches were used to initially describe the gap itself and later to look into its causes and consequences. Simple descriptive statistics were used in Chapter 2 to compare average theoretical and actual consumption of gas, electricity and primary energy and CO_2 emissions. The resulting averages of both theoretical and actual consumption were then extrapolated nationwide in order to be compared with the existing policy targets. This enables an insight into whether or not the policy targets can be achieved using theoretical as well as actual energy consumption as a baseline. Electricity turned out to be constant in differently performing dwellings due to the fact that the use of appliances, which accounts for the largest amount of electricity end usage, depends little on residential thermal performance. Another reason for electricity to be constant is the fact that most Dutch dwellings (over 90%) use gas for heating and hot tap water, which is why the variation in thermal performance can best be detected looking at residential gas consumption. The scope of the thesis was therefore narrowed down to gas consumption.

The third chapter investigates the same dataset used in the second chapter, this time with the intention of gaining insight into the causes of the discrepancies discovered. National socioeconomic data were added to the studied sample and a regression analysis was carried out. Due to the low predictive power of the included variables, a sensitivity analysis of the theoretical gas usage was performed on the basis of the average row house taken from the WOON dataset. Sensitivity analysis was performed on six assumptions made in the theoretical calculation to show how an incremental difference in one of the assumptions affects the final theoretical gas consumption and whether this could explain the performance gap.

The unsatisfactory results of the first regression analysis (relatively low R^2 value of explained variance) based on socioeconomic data led to a survey carried out in Chapter 4 of the thesis. The survey was conducted on a subset of Amsterdam dwellings that had an official energy label. This 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. Initially, several statistical tests were performed. Non-parametric tests were chosen because the variables in question were not normally distributed and the attempt of transforming them into normally distributed functions was not successful. Spearman's rho was used for establishing correlations between continuous variables, the Kruskal-Wallis test for variables with more than two categories and Mann Whitney's U statistic was calculated for binary variables. Upon evaluating the results of these tests, several regression analyses were performed on different subsamples. The resulting B values were used to attempt to improve the predicted theoretical consumption in another independent sample.

As opposed to the samples studied in the first three papers, all of which were based on cross-sectional data, Chapter 4 was the first to analyse longitudinal data from the social housing dwelling stock between 2010 and 2013, meaning that the research was narrowed down to dwellings that had undergone renovations in order to see whether the theoretical reduction of energy consumption materialised and to what extent. Since in this sample the dwelling's geometry stays the same, the relation between performance gaps before and after renovations provides important insight into the accuracy of the normalisations used in the regulatory calculation model used in energy labelling. Moreover, a comparison of the actual reductions effected by different renovation measures was made in order to show which renovation practices lower energy consumption most effectively.

TOC

§ 1.7 Data

Four datasets were used in this thesis. The second and third chapters use the dataset provided by the Raad Voor Ondernemend Nederland (RVO, formerly AgentschapNL, a national government agency). The fourth chapter used a dataset acquired from Rekenkamer Amsterdam (an independent board that evaluates municipal policies). In the third and fourth chapters the WOON database, based on a national housing survey commissioned by the government, is used for validation. The last dataset used in the fifth chapter, called SHAERE, was provided by Aedes, the Dutch federation of housing associations. The datasets had some commonalities but also several differences.

The RVO dataset consisted of all energy labels issued in the Netherlands between January 2010 and December 2010 – a total of over 340,000 cases with 43 variables (regarding building location and technical characteristics, label certification, etc.). In the second chapter, which used this dataset, the following variables were used: energy index (transformed into energy label), theoretical electricity consumption, theoretical gas consumption. Apart from that, actual gas and electricity consumption data at the level of individual dwellings were acquired from the Statistics Netherlands office, the details about which are in Chapter 2. The final dataset included 193,856 records.

The third chapter essentially used the same data as in the second chapter; however, in addition to theoretical and actual gas and electricity consumption, other variables were included, such as floor area, dwelling type and installation. Furthermore, these data were coupled with the socioeconomic data available from Statistics Netherlands about the household and the occupants from the building register, municipal data administration and the employment database. Due to this coupling, the dataset analysed in the second chapter was smaller, including approximately 40,000 dwellings. Nevertheless, together with numerous variables describing the dwelling and household itself, this number of records was expected to yield interesting results when using regression analysis. For the second part of this chapter, the sensitivity analysis, the WOON survey conducted by the Dutch government in 2012 was coupled with the data. The WOON dataset is based on a survey of detailed energy performance (including an inspection) of Dutch dwelling stock carried out every 5 or 6 years by the Dutch Ministry of the Interior and Kingdom Relations (BZK) and the 5000 dwellings it contains are representative of the entire Dutch housing stock. This dataset was used for the formation of a reference building that was used for the sensitivity calculations, which were carried out by manipulating one of six assumptions made in the calculation method: average indoor temperature, number of occupants, internal load, ventilation rate, floor area and insulation values.

The fourth chapter was based on a dataset of 245,841 label certificates issued for the Amsterdam area from 2007 to 2013 and provided to us by the Rekenkamer Amsterdam. This data contained variables identical to those previously seen in the RVO data since they both originate from the same register of label certificates, managed by the BZK. To further enrich the studied dataset and due to the fact that regression analysis did not yield the desired results in Chapter 3 where publicly available data was used, an occupant survey was carried out on a subsample of about 1000 dwellings. The survey was carried out per label category, gathering the same data in each of the seven label categories, including 42 questions about dwelling properties that are not included in the label certificate (number of rooms, type of occupancy, thermostat type, water saving showerhead, etc.), household properties (number, age of occupants, ability to pay energy bill), behaviour of occupants (presence at home, heating and ventilation practices, showering habits, energy efficient behaviours, etc.) and comfort (temperature, air velocity, and humidity). Several records out of the 1000 were later found to have invalid actual energy data and had to be discarded, resulting in 460 records. The WOON survey from 2012, previously mentioned, was used for validation.

In the fifth chapter, we used the SHAERE database, which includes annual performance data (between 2010 and 2013) for almost all dwellings administered by social housing associations. Each year, the social housing associations record the state of most of their dwellings, including their energy performance, in the SHAERE register. SHAERE was set up by AEDES, the Dutch federation of housing associations, to be able to track the renovation pace of the stock in relation to the 2020 goals laid down in a covenant with the government and the tenants' organisation. The main difference between SHAERE and the RVO or the Rekenkamer sample is that the SHAERE dataset consists of pre-labels, which correspond to the complete thermal performance of the dwellings at the end of each calendar year, not just the label certificates registered with the authorities. Some of the pre-labels are registered with the authorities (an official certificate is later issued) but others are not. Therefore, the accuracy of the dataset could be inferior in comparison with the RVO and Rekenkamer sample, since those labels are registered and are legally required to comply with the actual state of the dwelling (however, no sanctions exist if the label is miscalculated). SHAERE is an export of the software the housing association uses for stock management and also for label registration and is therefore expected to have a certain degree of accuracy. The data it contains is much more detailed than the RVO and consists of the same dwellings over a period of four years (longitudinal). Apart from that, a significant advantage of this dataset is that more variables are available to describe the dwelling's thermal quality.

All samples used in the study were coupled with the actual energy data from Statistics Netherlands at the address level. Statistics Netherlands receives this information about gas and electricity use each year from the energy companies. Roughly one-third of every sample we used could not be coupled due to missing address or energy data or because of significant uncertainties about the quality of the data. An example is the removal of dwellings with collective heating installations or dwellings with shared cooking facilities, since actual consumption data for these is unreliable. Moreover, maximum actual consumption thresholds had to be set to exclude outliers. Furthermore, actual gas usage as available from Statistics Netherlands is standardised to a set number of degree days, so in order to compare this consumption data to the theoretical consumption where 2620 degree days are assumed, a correction factor was applied.

CHAPTER	2.	3.	4.	5.	3. AND 4.
Dataset	RVO	RVO	Rekenkamer Amsterdam	SHAERE	WOON Energy Module
Size (raw)	340,000	340,000	245,841	5,205,979	5000
Data type	All energy label certificates issued in 2010	All energy label certificates issued in 2010	Energy labels issued from 2007 on, dwellings renovated in last 3 years removed	All pre-labels available at Aedes from 2010 - 2013	Energy labels and survey data for 2012
Sample used	193,856	App. 40,000 (regression) and 713 (sensitivity)	460	644,586 (part B) and 81,740 (part C)	4800
Energy sources	Gas and electricity	Gas	Gas	Gas	Gas
Ownership type	Social housing and privately owned or rented dwellings	Social housing and privately owned or rented dwellings	Social housing	Social housing	Social housing and privately owned or rented dwellings

TABLE 1 Main properties of the datasets used in the thesis

§ 1.8 Limitations

Several limitations were encountered while analysing the data. The representativeness of the RVO dataset was found to be satisfactory, except when looking at ownership type (owner-occupied dwellings were underrepresented). This is a general problem when analysing Dutch energy label certificates, since the social housing sector is a frontrunner in labelling dwellings (more than two-thirds of the labelled dwellings are social housing dwellings). On the other hand, social housing constitutes about one-third of the Dutch market and is quite representative of the total dwelling stock. In the second chapter, its representativeness for various parameters is estimated. The sample is relatively representative of dwelling types, building years and most importantly, energy performance, however there is a clear difference between the overall Dutch stock and the studied sample when comparing ownership type and dwelling size. The

SHAERE dataset has properties similar to the RVO database, but does not include any private dwellings. In terms of its representativeness of all Dutch housing stock it is slightly less representative than the RVO sample due to the fact that private housing is not included. On the other hand, SHAERE is bigger than any other register of residential thermal performance analysed previously and the fact that several records are available for the same dwelling favourably affects its reliability.

Both the RVO and SHAERE samples encompassed all available energy performance records and the distribution of label classes followed a Gaussian curve (Figure 2 of Chapter 2). The Rekenkamer sample was different. In this sample, the survey deliberately sampled equal numbers of dwellings from each label category. Even though this deviates from the normal distribution of performance certificates, it was useful for the regression analysis, since it meant large enough proportions of very high and low performing dwellings. Because of that, it is easier to find significant predictors in those groups, which are usually less well represented.

Another limitation to consider is the reliability of the data. Energy label certificates that are registered with government authorities (RVO and Rekenkamer dataset) are less likely to contain errors than those that are not registered (SHAERE dataset). However, previous studies into the reliability of the label also demonstrated flaws contained in official certificates, since a 20% rate of error was detected and was attributed to poor inspection work. In some cases, the software used for label calculation permits illogical input, for example a combined high-efficiency system for hot tap water and a low-efficiency boiler for heating in the same house – even though a combined hot tap water boiler means it also heats water for heating. On the other hand, the survey we used in Chapter 4 introduces the bias of the respondent, which also affects reliability. Some questions were very specific (about occupancy patterns or indoor temperature) and it is possible that not all occupants answered these questions accurately.

The integrity of the data is another limitation to consider. The energy label datasets missed two crucial variables for analysis – hot tap water system and insulation of the dwelling. These two variables are strangely not a part of the official register of energy certificates. In the fourth chapter we improved this by including a question about a hot tap water system in the survey, and in the fifth chapter (SHAERE dataset) these variables were available.

The system boundaries of the analyses beyond the second chapter are confined to gas, and no longer to total energy consumption. There are several reasons for this. Most Dutch dwellings are heated by gas and we show that discrepancies arose entirely due to differing gas consumption. However, leaving electricity out in the following chapters meant that we automatically excluded all dwellings which have installation systems that are not gas-based. These are mostly very efficient installation systems (heat pumps and cogeneration) but also include some inefficient ones, such as electrical radiators. However, even with this exclusion over 90% of housing stock is covered.

§ 1.9 Thesis structure

Table 2 below summarises the research questions and sub-questions and relates them to the corresponding datasets (Table 1). The research questions A and C were partly analysed with all four datasets and answers to these questions are dispersed throughout the whole thesis. Question B, on the other hand, relies on two datasets, RVO and SHAERE, and is answered in Chapters 2 and 5.

RESEARCH QUESTION	DATASET	CHAPTER
A. The discrepancies between actual and theoretical heating energy consumption in labelled dwellings		
A.1 Discrepancies between theoretical and actual gas and electricity consumption	RVO	2
A.2 Performance gap in relation to energy label	RVO/RA/SHA/ WOON	2, 3, 4,5
A.3 Performance gap in different samples	RVO/RA/SHA/ WOON	
A.4 Performance gap in relation to dwelling type, floor area and installation type	RVO/RA/WOON	3, 4
B. Energy reduction targets for the built environment and actual reduction potential of the dwelling stock and of the individual dwelling renovation measures		
B.1 Theoretical and actual achievability of the current targets	RVO	2
B.2 Differences between theoretical and actual reductions in dwellings where different renovation measures were applied	SHA	5
C. Causes of the differences between actual and theoretical gas consumption		
C.1 Explaining variation in gas use with dwelling, household and occupant characteristics	RVO/ RA	2, 3
C.2 The relation between theoretical gas usage and the normalised assumptions	RVO/WOON/SHA	3, 5
C.3 A better model for theoretical gas consumption	RA/WOON	4
*RA stands for Rekenkamer Amsterdam and SHA for SHAERE dataset.		

**RA stands for Rekenkamer Amsterdam and SHA for SHAERE dataset*

TABLE 2 Data used per research question and relation to specific chapter

A more conceptual structure of the thesis is depicted in Figure 2. Chapters 3 and 4 deal with the correlation between a set of parameters (occupant behaviour, building and household characteristics), whereas Chapters 2 and 5 mainly tackle the consequences of the performance gap. Highlighted areas divide the content into three research questions described in 1.5.

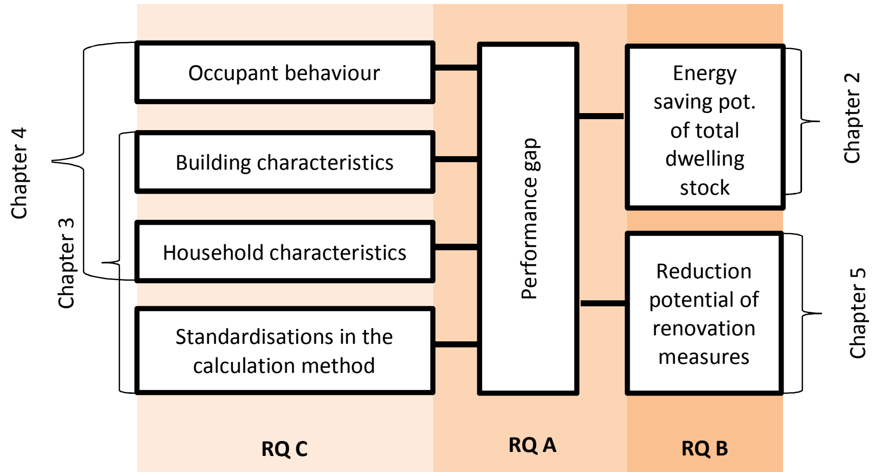


FIGURE 2 Research framework (RQ – research question)

§ 1.10 References

Branco, G., Lachal, B., Gallinelli, P., Weber, W., 2004. Predicted versus observed heat consumption of a low energy multifamily complex in Switzerland based on long-term experimental data, Energy and Buildings, Volume 36, Issue 6, June 2004, Pages 543-555.

Cayre, E., Allibe, B., Laurent, M.H., Osso, D., 2011. There are people in this house! How the results of purely technical analysis of residential energy consumption are misleading for energy policies, Proceedings of the European Council for an Energy Efficient Economy (eceee) Summer School, 6–11 June 2011, Belambra Presqu'île de Giens, France.

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.

Greening, L., Greene, D., Difiglio, C., 2000. Energy efficiency and consumption—the rebound effect—a survey, Energy Policy, Volume 28, Issue 6/7, June 2000, Pages 389–401.

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

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

Haas R., Biermayr, P., 2000. The rebound effect for space heating - Empirical evidence from Austria, Energy Policy, Volume 28, Number 6, June 2000 , Pages 403-410.

Hens, H., Parijs, W., Deurinck, M., 2010. Energy consumption for heating and rebound effects, Energy and Buildings, Volume 42, Issue 1, January 2010, Pages 105-110.

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

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

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

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.

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., Filippidou, F. Analysis of label changes in the social housing sector between 2008 and 2013, OTB-Research for the Built Environment, Faculty of Architecture, Delft University of Technology, May 2014

Filippidou , F., Nieboer, N. Energetische verbeteringsmaatregelen in de sociale-huursector, OTB-Research for the Built Environment, Faculty of Architecture, Delft University of Technology, September 2014

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.

Milne, G., Boardman, B., 2000. Making cold homes warmer: the effect of energy efficiency improvements in low-income homes, A report to the Energy Action Grants Agency Charitable Trust, Energy Policy, Volume 28, Issues 6–7, June 2000, Pages 411-424.

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

Perez, L., Ortiz, J., Gonzales, R., Maestre L.R., 2009. A review of benchmarking, rating and labelling concepts within the framework of building energy certification schemes, Energy and Buildings, Volume 48, Pages 272-278.

Pettersen, T.D., 1994. Variation of energy consumption in dwellings due to climate, building and inhabitants, Energy and Buildings, Volume 21, Issue 3, Pages 209 – 218.

Ramallo-González, A.P. 2013. Modelling Simulation and Optimisation of Low-energy Buildings. PhD. University of Exeter.

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

Sorrell, S., Dimitropoulos, J., Sommerville, M., 2009. Empirical estimates of the direct rebound effect: A review, Energy Policy, Volume 37, Issue 4, April 2009, Pages 1356-1371.

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.

Yücel, G., Pruyt, E., 2011. Energy Transition in The Dutch Dwelling Stock: Exploring the Extent of Inertia Against Change, Proceedings of International System Dynamics Conference 2011, July 24 – 28 2011, Washington DC, USA.

2 Theoretical vs. actual energy consumption of labelled dwellings in The Netherlands: Discrepancies and policy implications

Explanatory note

This research studies the difference between actual and theoretical energy consumption in Dutch residential dwelling stock. The research utilised the energy label certificates issued in The Netherlands in 2010, containing dwellings' theoretical performance. This dataset was merged with actual energy data on the level of individual dwelling. Simple descriptive statistics were used to compare average theoretical and actual consumption of gas, electricity and primary energy and CO_2 emissions. It became clear that the discrepancies were significant and related strongly to the performance category, which meant that there could be a substantial impact on the energy savings targets set by the government. Therefore, the resulting averages of both theoretical and actual consumption were used in a scenario study, where they are extrapolated nationwide in order to be compared with the existing policy targets. Results showed that while the targets can be achieved using the theoretical consumptions as baselines they are out of reach if projected on the basis of actual consumptions.

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

Abstract

In Europe, the Energy Performance of Buildings Directive (EPBD) provides for compulsory energy performance certification (labelling) for all existing dwellings. In the Netherlands, a labelling scheme was introduced in 2008. Certificates contain the energy label of the dwelling and corresponding theoretical gas and electricity consumption, calculated based on the dwellings physical characteristics, its heating, ventilation and cooling systems and standard use characteristics. This paper reports on a large-scale study comparing labels and theoretical energy use with data on actual energy use. A database of around 200,000 labels was coupled with data from

Statistics Netherlands on actual gas and electricity consumption provided by energy companies. The study shows that dwellings with a low energy label actually consume much less energy than predicted by the label, but on the other hand, energy-efficient dwellings consume more than predicted. In practice, policy targets are set according to the theoretical rather than the actual consumptions of the building stock. In line with identified discrepancies, the study shows that whereas most energy reduction targets can be met according to the theoretical energy consumption of the dwelling stock, the future actual energy reduction potential is much lower and fails to meet most of the current energy reduction targets.

§ 2.1 Introduction

Buildings are responsible for approximately 40% of the EU's energy consumption and accounted for 30% of EU's CO_2 emissions in 2005 (SERPEC-CC Summary Report, 2009). In 2002, the European Performance of Buildings Directive was put in place with the aim of reducing the amount of energy consumed by the residential and utility sectors by informing renters and buyers of the energy consumption of the buildings in which they live and setting an EU framework for energy performance certification (EPBD 2002/91/EC). The general requirements of the 2002 EPBD for residential buildings included the development of a system of energy certification for new and existing buildings, regular inspections of heating and air-conditioning systems and the introduction of minimum energy-performance standards for new and extensively renovated existing buildings with a useable floor area of over 1000m^2. Mandatory energy certification for residential buildings, which is the focus of this paper, was introduced for all properties constructed, sold or rented.

All member states had implemented the directive by the end of 2009, some more effectively than others (Andaloro et al., 2010). This process seems to have been well studied within numerous EU projects and initiatives (BPIE, 2011, ASIEPI, 2009, IMPLEMENT, 2010, IDEAL, 2009). Moreover, a joint initiative undertaken by the EU member states and the European Commission, the Concerted Action EPBD, enables member states to share their information and experiences of adopting and implementing this European legislation at the national level (www.epbd-ca.eu). The two major shortcomings of the directive as concluded in the EU project IMPLEMENT, are the looseness of the regulations in the directive, which leave ample room for interpretation, and the fact that no sanctions are imposed in cases where the rules of the EPDB are ignored (for example, failure to issue an energy certificate when selling a house). Additionally, the European Project IDEAL-EPBD was specifically designed to investigate why energy performance certificates hardly seem to motivate homeowners to take measures to improve the energy performance of their dwelling; it produced several policy proposals to improve the impact of the EPDB. However, all these projects deal with implementation of the EPDB strategically and overlook the accuracy and outcomes of the calculation methods used. It seems certain that this varies throughout the EU, since the methodology of the energy performance certificates (EPC) is not defined by the directive and is in hands of individual member states, which have developed very different approaches and methodologies (EPBD Concerted Action). However, in 2004 the EC appointed the CEN (mandate M/343) to develop a series of standards. These include the following: EN 15217 (energy performance of buildings - ways of expressing the energy performance of buildings and energy certification); EN15603 (the energy-efficiency of buildings – overall energy use and the definition of the energy rating); EN ISO 13790 (energy performance of buildings – calculating the energy used for heating and cooling). However, the methodologies do not comply fully with the standards in all member states (Andaloro et al., 2010), including the Netherlands.

Clearly, the theoretical values are merely an estimation of the actual consumption, since they are based on standard values and do not take account of the lifestyle of the occupants. However, the labels also provide homeowners and tenants with information on possible energy-saving measures, and the pay-back time for these measures is directly related to the theoretical energy consumption. Future targets for reducing energy consumption and feasible energy reduction policies are formulated according to the theoretical potential for energy reduction. If the label is to become an efficient tool with which to reduce household energy consumption in line with the targets set, the theoretical decrease in energy consumption when improving the energy label of a particular dwelling should closely reflect the actual decrease in energy consumption.

This study aims to identify the results of the energy performance calculation which was implemented in line with the EPBD directive, comparing it with the actual energy consumption of Dutch dwellings. In order to assess a broader efficacy of the energy label methodology as a policy tool for achieving reductions in household energy consumption, actual and theoretical energy consumption were examined in respect to the targets set for reductions in energy consumption and CO_2 emissions for the residential sector in the EU and the Netherlands.

This paper is organized as follows. Section 2.2 provides background information on the topic, a review of existing studies and energy and CO_2 reduction targets. In section 2.3, the energy-efficiency of Dutch households is presented together with an overview of the Dutch energy label calculation for dwellings. The results are presented in section 2.4, followed by a scenario study in section 2.5 and finally, the discussion and conclusions in sections 2.6 and 2.7.

§ 2.2 State of the Art

§ 2.2.1 Existing studies on actual energy consumption

According to Perez et al. (2008), the lack of a complete databases containing the information on energy performance coefficients of buildings in the national dwelling stock together with building type, size etc., impedes the evaluation of the policies at the national and EU levels. Poor availability and accessibility of energy label databases for researchers is probably the main reason that this subject has remained under-researched. The small amount of literature that is available relating the label of the dwellings with their actual performance is mostly based on small samples, with the intention of quantifying the role of occupancy in explaining differences. Guerra Santin (2012) compared the actual and expected energy consumption for 248 Dutch dwellings built after 1996. The dwellings were categorised according to their EPC value (the Dutch energy performance coefficient for new buildings). The EPC (NEN 5128) calculation method is broadly similar to the energy index calculation method, which is the basis for the energy label (see section 2.3.2).

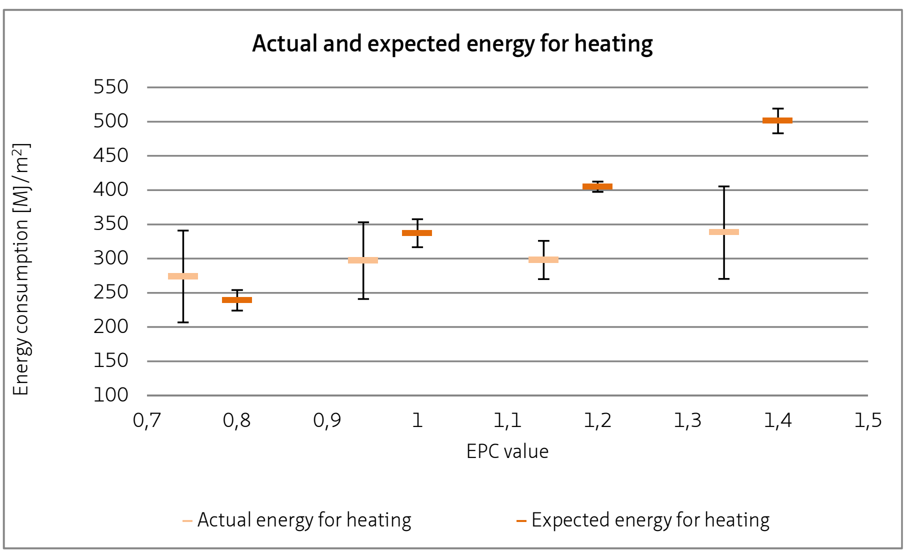


FIGURE 1 Mean and 95% confidence interval for the actual energy consumption (MJ/m²) and expected energy for heating (MJ/m²) per EPC value (Guerra Santin, 2012)

In energy-inefficient buildings with a high EPC, actual energy consumption for heating was almost half that expected, whereas in buildings with a low EPC (energy-efficient buildings), the actual and expected heating energy consumptions coincided much better. Due to the relatively small sample size, the differences between the actual heating energy of buildings with different EPC values were insignificant, although the mean actual consumption was consistently lower in buildings with a lower EPC (Figure 1).

In another study conducted in the Netherlands by Tigchelaar (2011), a 'heating factor' was calculated (the actual demand for heating is divided by the theoretical demand). The average heating factor in a sample of 4700 representative dwellings was found to be below one, meaning that the theoretical consumption was overestimated. Cayre et al. (2011) studied actual and theoretical energy consumption in 923 French dwellings and reached similar conclusions – the French EPC model overestimates the theoretical energy consumption in the sample, which was representative of the French dwelling stock as a whole. Hens (2010) arrived at similar findings when observing 20 low income, non-insulated dwellings in Belgium. There, the measured energy use was merely a fraction (on average approximately 50%) of the calculated consumption. These findings were extrapolated to a broader sample, showing that the difference between measured and calculated consumption is larger in non-insulated than in well-insulated homes. On the other hand, in 12 multi-family thermally retrofitted buildings in Austria, Haas and Biermayr (2000) found evidence that actual energy consumption significantly exceeded the expected. Similar results were obtained by Branco et al. (2004) in a multi-family complex in Switzerland and in a similar sample by Marchio en Rabl (1991) in France. On the basis of these results, it seems that the theoretical energy consumption tends to be overestimated when looking at average and less energy-efficient dwellings and underestimated when observing new or retrofitted buildings. The phenomenon of underestimated theoretical consumption can partly be explained by the 'rebound effect' (Berkhout et al., 2000), by which more efficient technologies (such as a low energy dwelling) cut energy bills but thereby encourage increased consumption. A typical example of rebound effect was found to be temperature control (Guerra Santin, 2010) - dwellings with a programmable thermostat turned out to consume more energy than households with a manual thermostat or manual valves on radiators. A similar phenomenon is described in previously mentioned study by Hens (2010), where the benefits of refraining from heating certain rooms in the dwelling are lower in well-insulated dwellings, since they are characterised by a more constant indoor temperature. Sorrell et al. (2009), provides an overview of the methods for calculating rebound effect and a summary of the studies available. Accordingly, he concludes that in OECD countries the mean value of the long-run direct rebound effect is likely to be below 30%. This means that up to 30% of the efficiency gained through the technical improvement of buildings and appliances result in increased consumption due to direct changes in user behaviour. In some cases, this can bring about increased comfort, but not always (for example, low energy bills may lead occupants to heat more rooms, which does not necessarily mean more comfort).

However, the size of the samples in the studies mentioned is relatively small, which sometimes leads to problems when assessing the statistical significance of the results. Moreover, the representativeness of the sample for the national dwelling stock is also not addressed at times. These factors are important when evaluating the accuracy of the energy label at a national level. Even in countries where energy label databases exist, few analyses of energy performance certificates are available.

§ 2.2.2 Energy and CO_2 reduction targets

As mentioned previously, buildings are an important sector in terms of the potential for reducing energy consumption and CO_2 emissions. The European Commission's Action Plan for Energy Efficiency, published in 2006, defines the full primary energy reduction potential of the residential buildings sector as around 27%. The EU's goal for overall primary energy is to reduce consumption by 20% by 2020 and, as stated in decision 406/2009/EC; a second goal is to reduce the total CO_2 emissions by 30% (including indirect emissions through the generation of electricity) by 2020 and by 50% by 2050. As part of this, the Netherlands has committed itself to reducing its total greenhouse gas emissions by 16% by 2020 (using 2005 as a baseline).

The SERPEC-CC report on the residential buildings and service sector was commissioned to identify the potential role of technology in reducing carbon emissions. It assumes the implementation of technologies which are available today or are likely to become economically viable in the near future, such as insulation, advanced heat supply technologies and more efficient electric appliances (lights, refrigerators, etc.). The reference level used was the standard practice and technology in 2005. The renewal of the buildings stock was assumed to occur at a pace of 1% per year and the renovation rate of buildings was assumed to occur at a maximum rate of 2.5% per year. Insulation measures and implementation of advanced heating systems were assumed to be implemented as part of a bigger project of buildings renovation, therefore the maximum implementation rate of these measures follows the rate of renovation. The future scenario, predicted for 2020, is comparable to a present-day energy-efficient house, which would now be labelled 'A' in the Netherlands. The study took account of technical measures rather than changes in behaviour (it assumed no rebound effect). It identified abatement costs, potential and reductions for the whole European Union within the built environment as 19% below 2005 emissions by 2020 and 29% by 2030. Reductions in the demand for heating are expected to result in a 61% decrease in CO_2 emissions by 2030, while electricity consumption is expected to increase by 5% due to the strong autonomous increase in electricity use. A similar study, which addressed member states separately, was also conducted within the European project IDEAL. On the basis of the results of the questionnaires relating to the building stock in the 10 participating countries, a preliminary estimate of the potential for energy savings was calculated. It was established that cost-effective energy savings of about 10% could be achieved by 2020 in most countries and 20% by 2030 – close to the goals set by the Netherlands.

As well as the laws and regulations concerning the energy performance of buildings at a national level in the Netherlands, several covenants have been made between the government and stakeholders, such as associations for the building sector, developers and housing associations. The Dutch federation of housing associations (Aedes) committed itself in the 'Covenant Energy Savings Housing Associations Sector' (*Convenant Energiebesparing Corporatiesector*, 2008) to save 20% on the consumption of natural gas (which is the main source of energy used to heat buildings in the Netherlands) in the existing social housing stock between 2008 and 2018. The social housing sector is set to achieve a 24PJ reduction in energy consumption between 2008 and 2020. The aim is to improve these dwellings to a B label or at least by 2 label classes. The so-called 'Spring Agreement' (*Lente-akkoord*, 2008) was signed by the Dutch government and other stakeholders, and states that all by 2015 newly constructed buildings will consume 50% less energy than in 2007. By 2020, all newly buildings should be 'energy-neutral'. However, at the time of writing of this paper it is still not clear what the exact definition of energy neutral building is, nor in The Netherlands nor in EU. However, rough guidelines are available in European Directive 2010/31/EU. Under the 'More with Less' (*Meer met Minder*, 2008) programme, the Dutch government and external stakeholders (corporations and external construction companies) are committed to achieving a reduction of 30% in the energy consumption (100PJ) of buildings by 2020.

§ 2.3 Household energy-efficiency and energy labels in the Netherlands

§ 2.3.1 Household energy-efficiency in the Netherlands

The energy-efficiency of the Dutch housing stock improved by 28% (Odyssee ECN, 2009) in the period between 1990 and 2008. The main reason for this significant improvement was the introduction of condensing boilers for heating and hot water.

Additionally, EPC regulations were introduced in 1995 and were also strengthened periodically, which significantly increased the efficiency of newly constructed dwellings, meaning that their energy consumption had halved by 2008 compared to 1990. However, Guerra Santin (2010) argues that the trend of decreasing energy consumption for heating in new dwellings failed to continue post-1998, despite the strengthening of the system of EPCs. Even though the efficiency measures implemented in the Netherlands place it at the forefront of the European residential sector (Odyssee ECN, 2009), there is no evidence for consistent reduction in total household consumption of natural gas since 1990 (consumption in 2008 was only 5% lower than in 1990) and the electricity consumption of households grew by 50% in the same period. This means that the total energy consumed by household grew by 11% (looking only at gas and electricity, the most important sources of energy in Dutch households). The reduction of consumption in the residential sector was also low due to the continued growth of the housing stock. Between 2008 and 2010, there was no significant decrease in either gas or electricity consumption (*De Nederlandse Energiebranche* website*, 2012*) at the household level (taking temperature correction into account).

Yücel and Pruyt (2011) claim that new construction can only achieve a limited reduction of energy consumption within the sector, since its rates are between 0.9 and 1.5% of the total building stock annually with a small fraction of demolition of about 0.2%. According to Yücel and Pruyt (2011), new construction will account for only a very marginal reduction in energy consumption by 2020, assuming the expected periodic strengthening of regulation and demolition and new construction rates. The renovation of the existing housing stock together with increased turnover is seen as the solution for a significant reduction in energy consumption.

The Energy Label strives to promote renovation work and the creation of more efficient buildings. However, research conducted in Denmark (Kjærbye, 2008) regarding the renovation of labelled dwellings showed that in most label categories there was no significant energy reduction within 4 years of owners purchasing the house (and receiving the label). Dwellings with label A were an exception, because there has been some energy reduction in the first two years after purchase. Unfortunately, no similar research was available for the Netherlands at the time of writing this paper. On the other hand, increased turnover has been observed for more energy-efficient buildings in the Netherlands (Brounen and Kok, 2010).

The data obtained through this study gives us an insight into the real potential for future energy savings through the energy label scheme, and thereby enables us to assess whether the scheme will help achieve the objectives set for reducing energy consumption and CO_2 emissions.

§ 2.3.2 Method of calculating the Dutch energy label for dwellings

The energy labelling of dwellings plays a crucial role in European and national policies that aim to reduce energy use. The energy label in the Netherlands is based on the 'Decree on Energy Performance of Buildings' (BEG) and the 'Regulation on Energy Performance of Buildings' (REG) which came fully into force in 2008. The method for calculating the energy label is described in ISSO 82.3. The first goal of labels is to provide occupants and homeowners with information on the thermal quality of their dwellings. To increase the practical significance of the label, the expected (theoretical) energy usage of the dwelling is also mentioned on all Dutch labels issued after January 2010, expressed in kWh electricity, m^3 gas and GJ heat (in dwellings with district heating).

An energy label awards each dwelling a grade, ranging from 'A++' to 'G' (Table 1). The categories are determined on the basis of the energy index, which is calculated on the basis of total primary energy demand (Q_{total}). Q_{total} sums up the primary energy consumed for heating, hot water, pumps/ventilators and lighting, subtracting the energy gains from PV cells and/or cogeneration (Equation 1).

$$Q_{total} = Q_{space\ heating} + Q_{water\ heating} + Q_{aux.energy} + Q_{lighting} - Q_{pv} - Q_{cogeneration}$$

EQUATION 1 Calculation of total energy consumption (Q_{total})

The energy index correlates directly to the total primary energy consumption, but is corrected for the floor area of the dwelling and the corresponding heat transmission areas (Equation 2) in order to not disadvantage larger dwellings and those with a greater proportion of envelope adjoining unheated spaces (different dwelling types). A correction is also applied for the shape of the dwelling when considering infiltration losses within space heating demand – the air permeability coefficient depends on building shape factor. Such a correction for compactness is also common in other European countries, although it has previously been argued that not correcting could promote more energy-efficient architectural designs (PREDAC WP4 report, 2003). On the other hand, striving exclusively for energy efficient design could compromise the functionality of the dwelling.

$$EI = \frac{Q_{total}}{155 \cdot A_{floor} + 106 \cdot A_{loss} + 9560}$$

EQUATION 2 Calculation of energy index (EI)

The total primary energy demand can also be expressed as described in equation 3. Since primary energy is an energy form found in nature, that has not been subjected to any conversion or transformation process, appropriate heating values need to be taken into account when calculating it. The assumed heating value for gas is 35.17MJ/m^3. The efficiency of the electricity network is considered to be 0.39.

$$Q_{total}[MJ] = Q_{total,gas}[m^3] \cdot 35.17 \left[\frac{MJ}{m^3}\right] + Q_{total\ el.}[kWh] \cdot 3.6 \left[\frac{MJ}{kWh}\right] : 0.39$$

EQUATION 3 Calculation of total primary energy

The level of carbon dioxide emitted depends on which fuel is used. As stated in ISSO 82.3, for 1MJ of energy derived from gas, 0.0506kg CO_2 is emitted into environment and for 1MJ of electricity, 0.0613kg CO_2 is emitted (taking into account the network efficiency and the fuel mix of electricity production).

LABEL	A++	A+	A	B	C	D	E	F	G
Index values	< 0,50	0,51- 0,70	0,71-1,05	1,06-1,30	1,31-1,60	1,61-2,00	2,01-2,40	2,41-2,90	> 2,9

TABLE 1 Dutch energy labels and the corresponding energy index values

The total primary energy consumption, and consequently the energy label allocated, are based on average occupancy and the average outdoor climate, and do not take account of the lifestyle or behaviour of the occupants. The energy index reflects the thermal quality of the building. Ventilation, internal heat production, energy use for lighting and heat losses during water circulation all depend directly on the useful floor area, which is defined as the area inside the heated zone, including rarely heated areas such as halls, toilets, washing rooms and storage spaces. The loft is also included if it is heated and the roof is insulated. Cellars, garages or other large storage areas are not included, since they are normally outside the thermal envelope. During the heating season, losses through ventilation and infiltration are taken into account as well at the standard indoor and outdoor temperatures. Heat loss through ventilation is calculated using a standard ventilation coefficient, which depends on the type of ventilation and is multiplied by the floor area of the dwelling. Heat loss through infiltration depends on the type of dwelling, since for each type of dwelling, characteristic lengths of frames, joints etc. are assumed (ISSO 82.3). A correction is made in the ventilation and infiltration calculations when a heat recovery system is present. Efficiencies are also defined for all kinds of heating and hot water installation systems. Heat gains from the sun are taken into account during the heating season at a flat rate of 855MJ/m^2 on a south-facing vertical surface, accounting for frames and dirt on the glass.

Possible energy gains through PV cells or micro co-generation plants are also taken into account. The demand for hot water is determined on the basis of the assumed number of occupants, which is determined as shown in Table 2. The heat demand calculations are based on a 2620 degree days (212 heating days, where the average outdoor temperature is assumed to be 5.64°C and indoor 18°C).

	CATEGORY	NUMBER OF PEOPLE/M², ASSUMPTION OF ENERGY LABEL METHOD
Dwelling floor area [m²]	<50	1.4
	≥50 and <75	2.2
	≥75 and <100	2.8
	≥100 and <150	3
	>150	3.2
Degree days [degree days]		2620
Internal heat production [W/m²]		6
Internal heat gains, south vertical [MJ/m²]		855

TABLE 2 Assumptions used in calculation

§ 2.4 Research methods and data

§ 2.4.1 Energy label database

This research used all the Dutch energy labels issued between January 2010 and December 2010 – a total of over 340,000 cases with 43 variables (regarding building location and technical characteristics, the properties of the label itself etc.). This data set was provided by AgentschapNL – a public sector organisation appointed by the Dutch Ministry of the Interior and Kingdom Relations.

This data was, on the basis of the addresses of the households, linked to actual energy use data, which was provided by the CBS (Statistics Netherlands), which collected this data from the energy companies. The combined data file was then cleaned up (deletion of double addresses on the basis of the label registration date, deletion of missing addresses on the basis of missing value) leaving 247,174 cases. The CBS expressed doubts about the quality of the data obtained for the energy consumption of collective installations (a single installation system providing heats for more dwellings) because this type of installation is arbitrarily assigned to buildings with a heat consumption that is too high to be considered realistic for an individual system. It was therefore decided to omit households with collective installation systems from the analysis. Dwellings which have multiple installation systems were also eliminated, since these are very specific cases. Cases where electricity consumption was nil were also removed. At this point, the gas values which were defined as missing were investigated. It turned out that most of them belonged to dwellings with heating installations, which do in fact use gas. Such cases were deleted, and only those which used electricity as a power source for heating were retained in the database. Gas use was then redefined to 0 for these cases. When checked the theoretical energy use and area of the house, outliers were detected. The cases with a floor space of over 1000m^2 and primary energy use of over 500,000 MJ were discarded. Finally, the actual gas consumption values for 2009 were corrected according to the number of degree days used in the theoretical calculation. After all this, the sample included 193,856 cases.

In this study, the following variables were used: energy index (transformed into energy label), theoretical electricity consumption, theoretical gas consumption and actual electricity, and gas consumption. Other variables, such as household floor area, dwelling type, construction and renovation year will be reported in a subsequent paper.

§ 2.4.2 Theoretical vs. actual energy consumption

The theoretical calculation method only takes account of energy for certain end uses and omits those uses which are determined by the occupants' lifestyle. On the other hand, actual gas and electricity consumption are derived from the actual energy bills for the dwellings in question and reflect consumption for all possible purposes. An overview of differences can be seen in Table 3. One important variable in electricity consumption is household appliances, which are not taken into account in the theoretical calculation, but are of course reflected in electricity bills (and therefore in our database). Appliances account for 32.4% of household electricity consumption (Milieucentraal, 2012). The difference between theoretical and actual gas consumption comes from gas used for cooking, which is only reflected in the actual value. On average, gas consumption represents 67.3% of total primary energy use, while electricity consumption represents 32.7% (Milieucentraal, 2012).

	THEORETICAL CONSUMPTION	ACTUAL CONSUMPTION	SHARE OF THE END USE IN THE TOTAL ACTUAL HOUSEHOLD CONSUMPTION OF THE NETHERLANDS
Electricity	Hot tap water	Hot water heating	14.7%
	Heating/Cooling	Heating/Cooling	17.6%
	Auxiliary energy (pump/ electronics/ventilation in heating installa-tion, ventilation system)	Auxiliary energy (pump/ electronics/ventilation in heating installa-tion, ventilation system)	n/a
	(Negative) PV/WKK production	(Negative) PV/WKK production	n/a
	Lighting	Lighting	14.7%
		Household appliances	32.4%
Gas	Heating	Heating	72.7%
	Hot tap water	Hot tap water	23.3%
		Cooking	3.9%

TABLE 3 Comparison of the end uses of gas and electricity in actual and theoretical household consumption

§ 2.4.3 Representativeness of the sample

Europe's buildings under the microscope (BPIE, 2011) highlights that only 11 out of 28 member states have (at the national level) a database of energy performance certificates, the Netherlands being one of those. The total Dutch dwelling stock included 7,104,000 dwellings in 2009 (CBS Statline, 2012). The sample we researched therefore represents slightly under 3% of the total dwelling stock.

The data for the whole Dutch dwelling stock was acquired from the *Energiecijfers* database, the CBS (Statistics Netherlands) Statline and the Energie NED (*De Nederlandse Energiebranche*) database. The representativeness of the sample needed to be assessed in order to have a clear idea of the extent to which the results within the sample could be extrapolated to the Dutch dwelling stock as a whole.

Since there were only a few cases in categories A++ and A+, all the A label dwellings were aggregated into one category. The distribution of labels thus became more normal and the results statistically more significant. As can be seen from Figure 2, more than half the dwellings in the energy label database belong to the categories C and D. As for the rest of the dwellings, only 1% belong to either one of the three most efficient categories (A, A+ or A++) and around 4% to G, which is the label of the most energy-inefficient dwellings. In the total Dutch housing stock, a slightly lower percentage of dwellings are labelled B and C than our sample included (Figure 2).

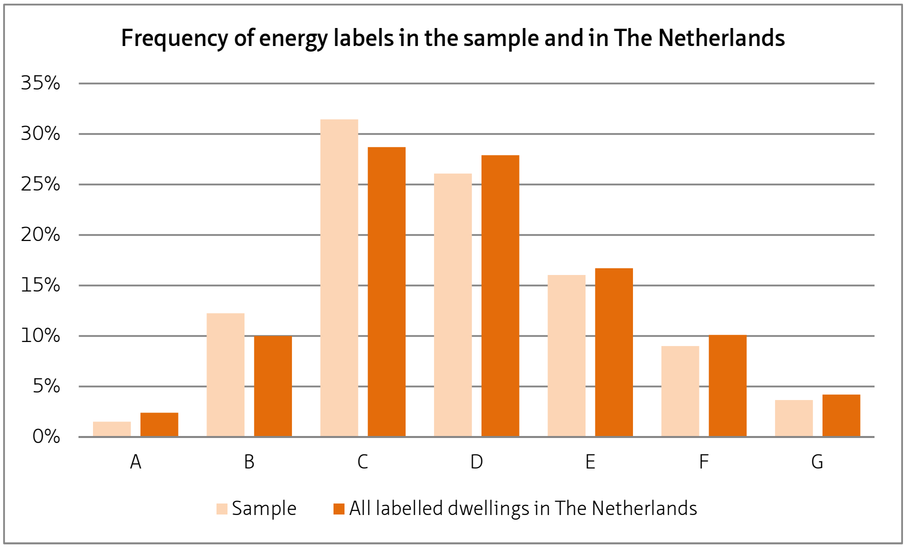


FIGURE 2 Shares of energy labels in the Dutch dwelling stock and in the sample

Almost half the dwellings in the sample were constructed in the 1970s, the 1980s, or the first half of the 1990s. Compared to the Dutch dwelling stock as a whole, one can see that the distribution in the dwelling stock is different to the sample (Figure 3).

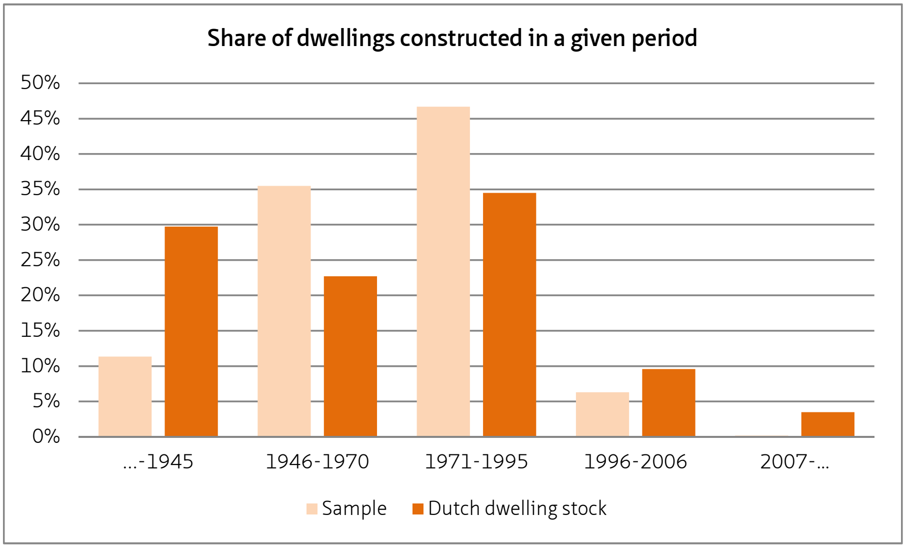


FIGURE 3 Share of the total Dutch dwelling stock and of the sample by period of construction/renovation

According to the *Energiecijfers* database, 62% of Dutch dwellings are terraced houses, 11% are detached (single family) houses and 27% are apartments. In our sample of dwellings, which was aggregated to the same four categories in Figure 4, this distribution was different. The discrepancies between the *Energiecijfers* database and our sample were the largest in the category of flats, which accounted for almost 36% of our sample but represented only just over 25% of the national housing stock in 2008, according to the *Energiecijfers* database. The below average number of detached dwellings in the sample is also reflected in the average size of a dwelling, which is over $10m^2$ smaller in the sample than the national average (Meijer & Itard, 2008).

The distribution of dwelling types according to the CBS in year 2009 is also shown in Figure 4, and this differs slightly from our sample as well as from the *Energiecijfers* database (the total stock is considered here to be 6,993,000 dwellings).

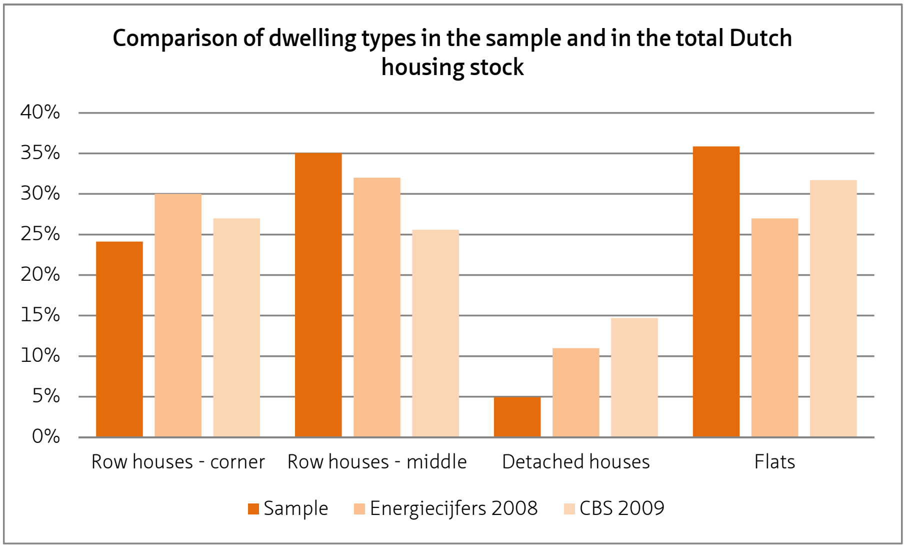


FIGURE 4 Representativeness of dwelling types of the Dutch housing stock in the sample, Energiecijfers 2008 and CBS 2009

In terms of ownership structure, the sample differs significantly from the national average (Energiecijfers database). Only slightly over 20% of the labelled dwellings are private owner occupied, while in the total housing stock this figure is 55%. Only one percent of dwellings in the sample were owner rental properties, whereas in the Netherlands as a whole, 12% of dwellings are owner rental properties. The third category is social housing, and this was over-represented in our sample (79% compared to 33% in the Netherlands as a whole), see Figure 5. The main reason for this was the absence of enforcement of the label scheme for owner occupants.

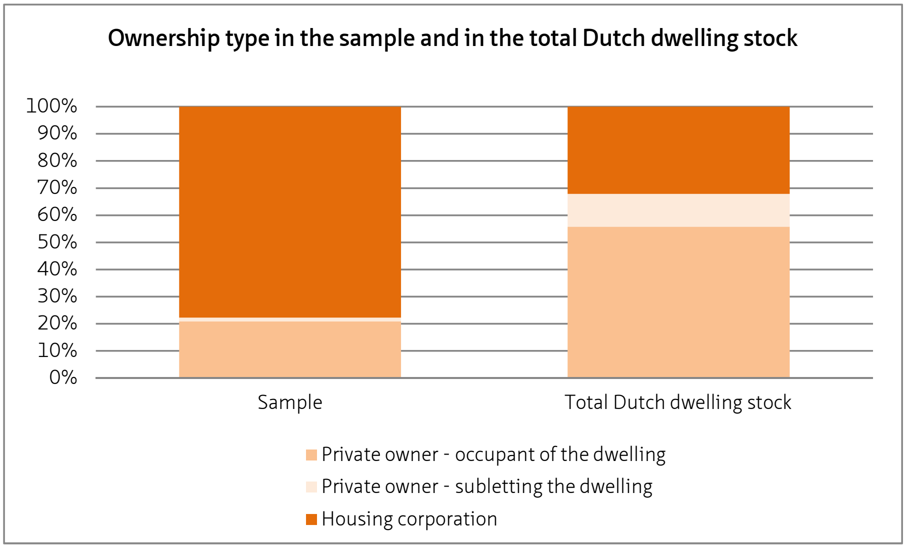


FIGURE 5 Ownership type distribution in the sample and in the Dutch housing stock as a whole

We can therefore conclude that our sample is well representative of all the labels issued in the residential Dutch dwelling stock. The construction years 1946-1995 are overrepresented. Flats and terraced houses are also overrepresented while detached houses are underrepresented. This is due to the fact that social housing is strongly overrepresented. The implications of this when interpreting the results are discussed in section 2.6.

§ 2.5 Results

§ 2.5.1 Actual vs. theoretical energy consumption

First of all, a comparison was made between the actual and theoretical primary energy consumption in the sample described above. The values appeared very similar, as can be seen from Figure 6. However, since it is known that theoretical consumption does not take into account end uses such as household appliances, which account for about 22% of total household energy consumption and the use of gas for cooking, which contributes 1.3% (calculated from the data in section 2.4.2), one might reasonably expect the theoretical consumption to be lower. Because gas and electricity are the two main energy sources for Dutch households and are also mentioned specifically on the energy label, they are examined separately in this study.

On average within the analysed sample, the theoretical primary energy use relating to gas consumption in a dwelling is on average much higher than the actual one, and the theoretical primary energy use relating to electricity consumption is significantly lower than the actual consumption of the same dwellings (Figure 6). In the case of electricity consumption, the fact that the amount of electricity used by appliances is not taken into account caused a part of the underestimation in theoretical consumption. However, judging from the values in Table 3, this is not the only cause (appliances account for an average of 32.4% of electricity consumption; if the overestimation in our sample was only due to appliances, these would contribute 64%). This may indicate that either the estimated electricity consumption of household appliances is inaccurate, or that electricity consumption for hot tap water and heating is higher than predicted. In contrast to electricity consumption, gas consumption was over-estimated. Since there is only one end uses for gas, with the exception of cooking, the difference in consumption reflects either a deviation from the assumed user behaviour or discrepancies in the assumptions used to estimate the demand for fuel for heating (air infiltration, U-values, floor area, transmission areas etc.) and the real values. However, this study does not aim to identify where these discrepancies come from, but rather their effect on the outcomes of energy policy targets in future.

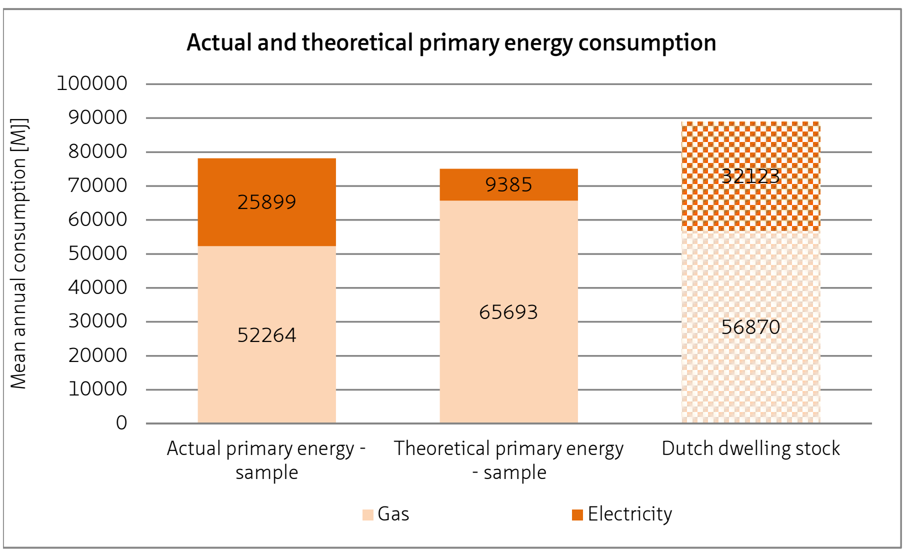


FIGURE 6 Actual and theoretical mean primary energy consumption per dwelling in the sample (N=193,856) and in the Dutch housing stock

In the Dutch housing stock as a whole (Figure 6), 3480 kWh of electricity (corresponding to 32123 MJ of primary energy) was consumed in a dwelling on average in 2010 according to *Energie Nederland*. This is around 700 kWh (6224 MJ of primary energy) more than the average in our sample. The same applies to gas: the average consumption in 2010 according to *Energie Nederland* was 1617m^3 (56870 MJ of primary energy), whereas consumption in our sample was 1487m^3 (52264 MJ of primary energy). This discrepancy is likely to have been caused by the smaller average size of the dwellings in our sample compared to the housing stock as a whole (see section 2.4.3).

§ 2.5.2 Energy consumption vs. energy label

The energy consumption for each label category is first presented separately for gas and electricity. Later, it is also presented together as total primary energy consumption.

§ 2.5.2.1 Gas

To understand how the energy label relates to the discrepancies described in the previous section, we examined gas and electricity consumption in various label categories. The plots in this report are presented with +/- 1 standard deviation. Because of the extremely large size of the sample, it is not relevant to plot the 95% confidence interval, which is always very small, meaning that the location of the mean value is known to a high degree of certainty and that all the differences were statistically significant on a 95% interval.

Figure 7 shows actual and theoretical gas use for each dwelling and Figure 8 shows the energy consumption per square metre of floor area of dwelling. Almost no difference can be discerned between either, except the difference in actual gas use between label A and label B. At the level of individual dwellings, the actual consumption was identical, but at the level of square metres of floor area, dwellings in category A use less gas than dwellings in category B. This may relate directly to the fact that dwellings in label category A were found to be considerably larger than all other dwellings (Figure 9). From these figures it is clear that although lower labels lead to increased actual gas consumption, there is a clear difference between the mean theoretical and mean actual gas consumption for each label.

For the most energy-efficient categories (A, A+ and A++) and for category B, Figure 7 and Figure 8 show that the theoretical calculation underestimated the actual annual gas consumption, in contrast to the rest of the categories for which the theoretical calculation largely overestimated the actual annual gas consumption. The theoretical and actual values only coincided for label C. It is worth noting that in label category G, actual gas consumption was only half theoretical consumption. Theoretical gas use predicts a much larger difference between an energy-efficient dwelling (A) and an energy-intensive dwelling (G) than we observed in our analysis of actual gas use. If the two consumptions are thought of as a linear function, they would differ significantly in the angle of their slope.

When standardizing the consumption per dwelling to consumption per square metre of floor space in the dwelling, we expected a better match between actual and theoretical levels of gas consumption because the dwellings could have different mean sizes in different categories. However, Figure 8 shows that this was not the case. The difference therefore does not arise because the dwellings are of different sizes, except for a small effect due to size among labels A and B (as is discernable from Figure 9). It is noticeable that the standard deviation of theoretical consumption decreases in Figure 8, meaning that the variation in terms of floor area is responsible for a large part of the variation in theoretical gas consumption at the level of individual dwellings (in Figure 7 the standard deviation is 40.7% of mean value for label G and in Figure 8 standard deviation is 20.8% for the same label).

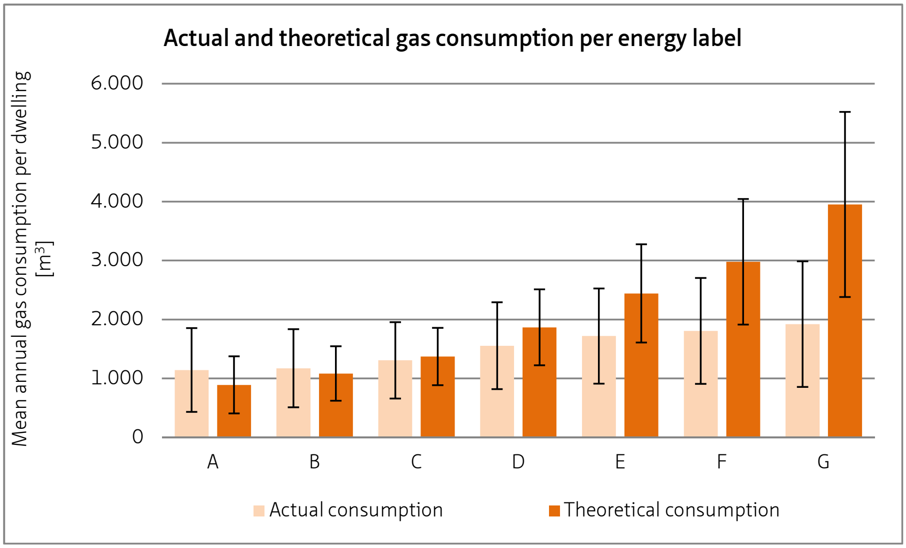


FIGURE 7 Actual and theoretical gas consumption per dwelling per label

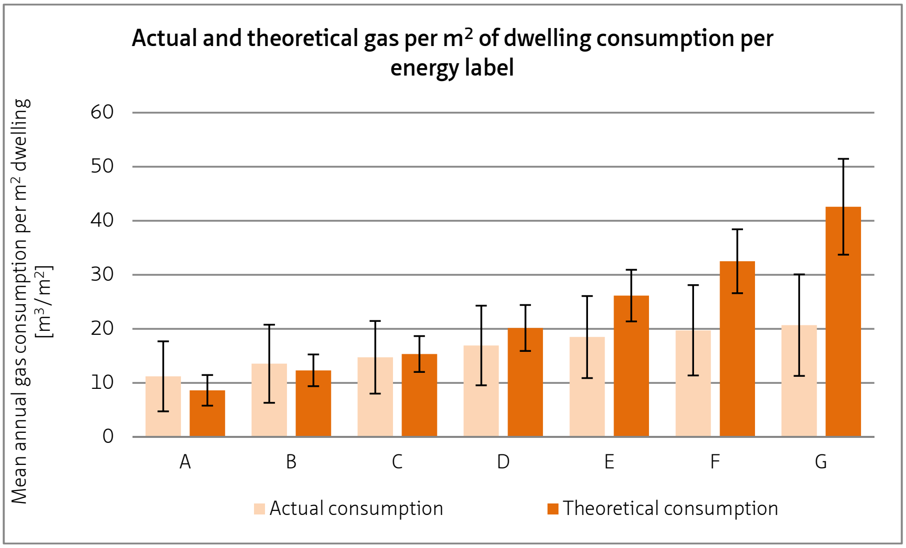


FIGURE 8 Actual and theoretical gas consumption per m^2 of dwelling area per label

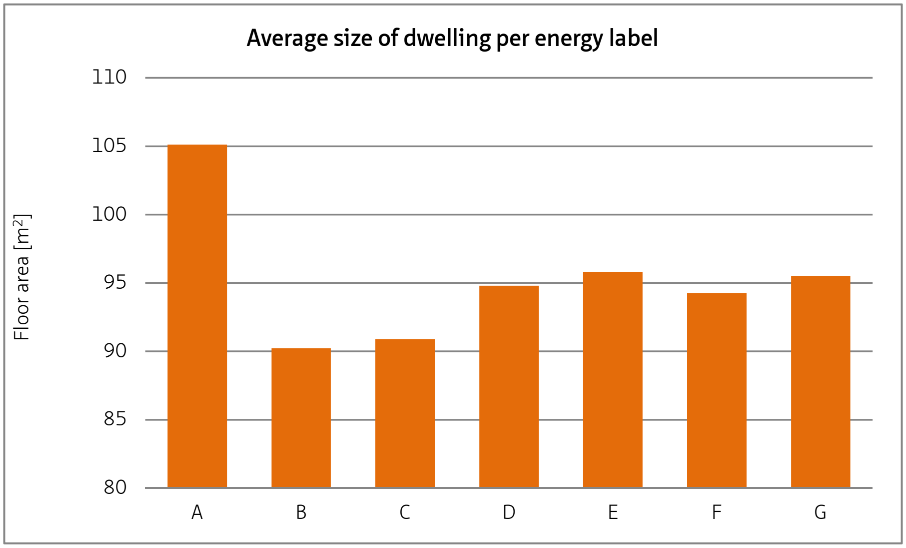


FIGURE 9 Average dwelling size (m^2 floor area) per label

§ 2.5.2.2 Electricity

In contrast to what we observed for gas consumption in the previous section, the theoretically calculated electricity consumption underestimated the actual consumption (Figure 6). Figure 10 shows that both actual and theoretical electricity consumption bear little relation to the label allocated. There is a very slight trend towards higher consumption in dwellings graded A, D and E which could be attributable to the electricity that is used for space and water heating or mechanical ventilation in certain more efficient dwellings (a larger proportion of heat pumps) and the larger floor areas. Figure 11, which shows electricity consumption per square metre of floor area, shows that the higher consumption for label A relates to larger floor areas. However, the curve still shows a slightly convex shape for the actual electricity consumption and a concave shape for the theoretical consumption, but ultimately the label does not appear to play a major role in the difference in electricity consumption. In fact, the differences between labels are very small compared to what was observed for gas consumption.

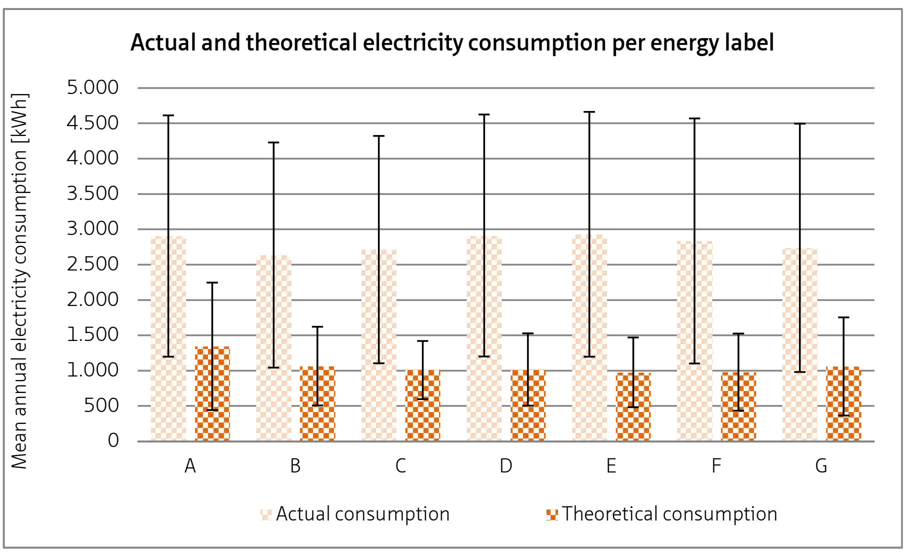


FIGURE 10 Actual and theoretical electricity consumption per label

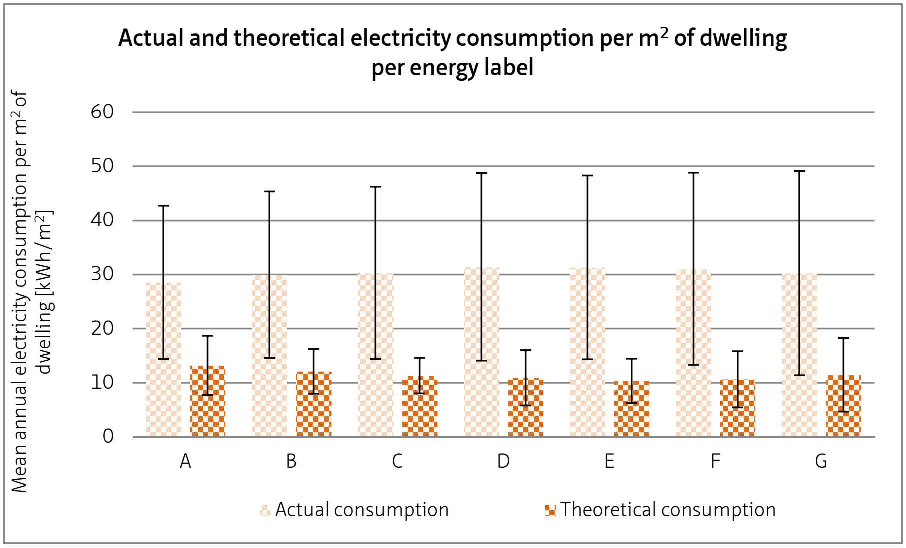


FIGURE 11 Actual and theoretical electricity consumption per m² of dwelling per label categories

§ 2.5.3 Total primary energy and CO_2 emissions per label category

An interesting insight into total primary energy consumption (Figure 12) can be gained by summing up the gas and electricity consumption data according to equation 3. From this figure, the occupants in dwellings with labels A – D can expect to consume more than the label certificate indicates. This will mainly be a consequence of higher gas consumption and will be offset by the fact that the household appliances are not a part of the label.

However, the difference in theoretical consumption is here again much greater between labels A and G than is the case in reality (looking at the actual values). This may have a very strong influence on the pay-back times and on the achievable savings. Dwellings with labels E, F or G seem to consume a similar amount of actual primary energy, even though the technical characteristics are much better in E than in G. The label may thus reflect the technical characteristics of a dwelling, but because actual primary energy consumption seems almost identical in each of the three categories, it might not be worth improving the technical specifications of houses labelled as G. From this figure it is clear that the savings which are expected to be achieved by improving the technical characteristics of a house, do not actually occur in practice. The theoretical primary energy consumption of a dwelling with an A label is 70% lower than that of a G label, but the actual primary energy consumption of an A label is only 28% lower than a G label.

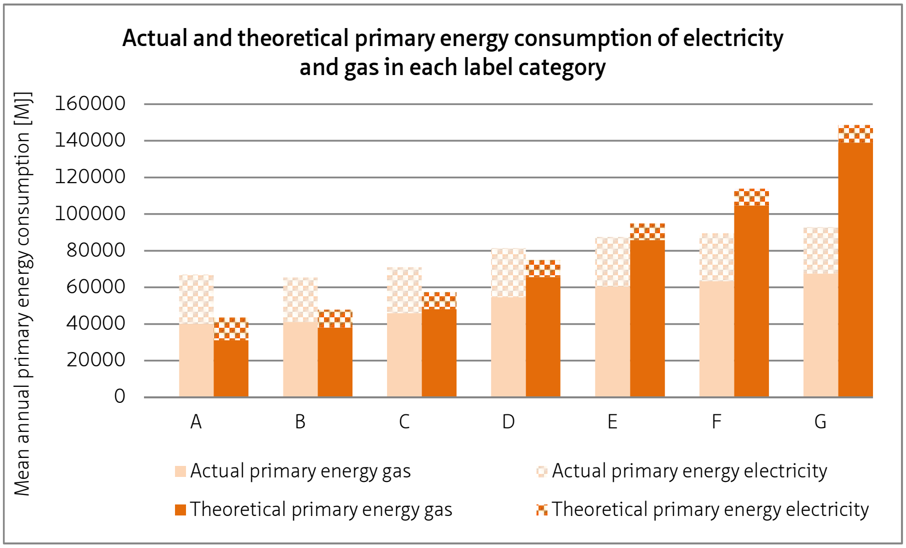


FIGURE 12 Actual and theoretical primary energy consumption per label

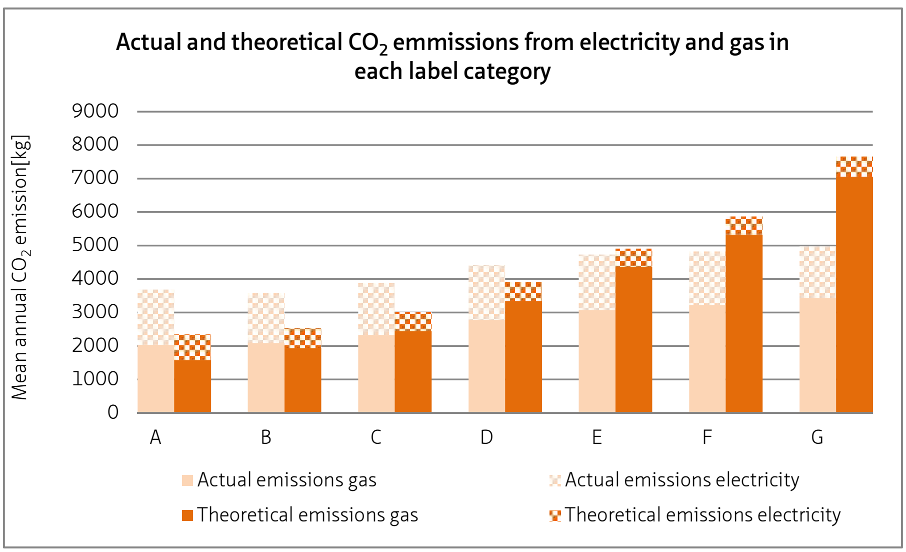


FIGURE 13 Actual and theoretical CO_2 emissions per label

Since European targets are not solely meant to reduce energy consumption but also CO_2 emissions, it is useful to look to what the energy label means in relation to CO_2 emissions. One megajoule of electricity produced in the Netherlands causes more CO_2 emissions than burning a megajoule of gas (0.0613kg vs. 0.0508kg of CO_2 per MJ). The

CO_2 emissions were calculated on the basis of this data. Theoretical CO_2 emissions are lower than actual emissions in labels except A – D. Interestingly, there is no significant decrease in CO_2 emissions for labels G, F and E and the label A is responsible for more CO_2 than label B. It is predicted that CO_2 emissions will decrease by 70% when moving from a G label to an A label, but in reality, looking at the actual consumption, this decrease is only 26%.

§ 2.6 Scenario study

An examination of Figure 12 and Figure 13 has cast doubt on the feasibility of the expected energy savings, as described in section 2.2.2, since these rely widely on theoretical estimates of consumption rather than on actual consumption data. As it was shown, actual and theoretical consumption differ considerably.

In order to determine what savings are actually possible by improving the energy label of dwellings already labelled, three different scenarios were tested. The analysis of consumption in the three scenarios is particularly interesting because this not only predicts the potential savings on the basis of the theoretical values but also on the basis of the actual consumption data from our sample. The average values for a particular label are extrapolated to the Dutch dwelling stock as a whole according to the distribution of labels all over Netherlands (Figure 2) and not only in the studied sample, thereby ensuring greater representativeness.

The first scenario equals the one proposed in the 'Covenant Energy Savings Housing Associations Sector', which aims to improve dwellings for at least by 2 label classes until the label B is achieved (so that dwellings with C labels are only improved by one label, dwellings labelled with B or A would not get improved, and all other dwellings are improved by 2 label classes) by the year 2018 (see section 2.2.2). In the covenant they assume that the entire housing stock that is labelled with C or lower will get refurbished by 2018. This implies a very high refurbishment rate and its feasibility is questionable. However, it is the target that Dutch housing associations have set and therefore it is tested in this paper. The second scenario assumes improving all labelled dwellings to label A, while the third assumes refurbishment to label B (dwellings currently labelled with A or B do not get improved). The first scenario is the least radical, while the second would require the most drastic refurbishment of the housing stock.

The differences in potential saving obtained through label calculation method (section 2.3.2) or by using the actual energy consumption data is clear (Table 4). According to the theoretical consumption, most of the targets would already be achievable with the implementation of the least stringent scenario – the only exception is the 100 PJ decrease in energy consumption as defined under the 'More with Less' Agreement (see section 2.2.2). However, this target can be achieved in the other two more radical scenarios.

However, the picture is completely different when the average actual consumption in each label category is used. The only target achievable with the first scenario is the 24PJ reduction in the energy consumption of social housing. There might be some bias here due to the fact that our sample contains both social and private dwellings (Figure 5), but in any case, social housing represents the majority (80%) of the sample. The 20% reduction in gas consumption throughout the whole dwelling stock, also proposed under the 'Covenant Energy Savings Housing Associations Sector', is also achievable with the implementation of scenario 2 or 3. All other targets regarding primary energy consumption reduction except the target of European project IDEAL, do not appear to be achievable (Table 4), regardless of the refurbishment scenario chosen. Interestingly, according to primary energy savings and CO_2 emission reductions, it seems better to aim for scenario 3 than scenario 2, since this scenario offers higher actual reductions of primary energy consumption and CO_2 emissions (but not gas consumption). This is a consequence of the phenomenon evident from Figure 10, which predicts a higher actual consumption of electricity for label A than for label B. The primary energy in one kWh of electricity is so high that it outweighs the impact of primary energy derived from gas consumption (which is indeed lower in dwellings with an A label).

		ACTUAL			THEORETICAL		
	AGREED SAVINGS	SCENARIO 1	SCENARIO 2	SCENARIO 3	SCENARIO 1	SCENARIO 2	SCENARIO 3
Convenant Energie-besparing Corporatiesector	-24PJ primary energy	70PJ	85PJ	96PJ	72PJ	146PJ	117PJ
	-20% gas use	16%	24%	22%	34%	54%	44%
Meer met minder	-100PJ primary energy	70PJ	85PJ	96PJ	72PJ	146PJ	117PJ
	-20-30% primary energy	12%	15%	17%	30%	43%	38%
SERPEC-CC	-19% primary energy	12%	15%	17%	30%	43%	38%
IDEAL	-10% primary energy	12%	15%	17%	30%	43%	38%
Dutch government	-16% CO_2	6%	9%	12%	21%	24%	27%
EC Action Plan for Energy Efficiency	-27% primary energy	12%	15%	17%	30%	43%	38%

TABLE 4 Energy and CO_2 savings in the three scenarios. he values in red are not achievable.

§ 2.7 Discussion

As mentioned in section 2.2.1, the strength of this study lies in the very large sample of households and energy certificates included (193,856). Figure 2 showed that the sample was representative in terms of the frequency of label categories, which was important since this study aimed to compare actual and theoretical energy consumption within label bands and extrapolate the predictions made within the energy label calculation to the whole Dutch dwelling stock (section 2.6). However, other characteristics of the sample, such as the type of dwellings or the ownership type showed poorer representativeness and we cannot exclude the possibility that this influenced some of the findings of this study to a certain extent. For instance, it may be the case that actual energy consumption in houses with poor label categories is higher in the (as yet) unlabelled housing stock than it is in our sample, which includes more social housing. This may therefore also influence the results of the scenario study (section 2.6).

Two additional points concerning the quality of the data used should also be noted. First, there are some concerns about the quality of the inspections on which the input data for the energy index calculations are based. A study carried out by the Inspection Service of Public Housing reported that in a sample of 120 labels issued in 2009, 60.8% of the inspected labels were incorrect, meaning that their energy index deviated more than 8% (Rapportage Gebruik en betrouwbaarheid energielabels bij woningen, 2009). In 2010 only 26.7% were incorrect, however the investigated sample contained only 30 houses (Betrouwbaarheid van energielabels bij woningen, 2010). In 2011, 16.7% of labels deviated more than 8% in their energy index in a sample of 48 dwellings (Derde onderzoek naar de betrouwbaarheid van energielabels bij woningen, 2011). There seems to be a trend of improvement, although the studied samples are very small. Most faults occur due to inaccurate input data and do not seem to correlate with the label of the dwelling. However, analyses of the data available in these studies show that the deviations are not symmetrical, in particular in label A, where the recalculated energy index is on average higher for 10% systematically, meaning that these dwellings were less efficient as demonstrated by their original certificate. In dwellings labelled with E and F the original index was higher than the recalculated one (2 and 1% respectively), meaning that the dwellings actually performed better. This is a small contribution to the performance gap detected in poor label classes but a significant one in dwellings with an A label.

Second, during the study some concerns arose concerning the quality of the actual energy data as given by energy companies to CBS. Because energy companies are required by law to check energy consumption at the meters only once every three years, it is possible that the consumption data used in the study are not the actual data for 2009, but contain some averages from the years 2006-2009. There is therefore also a possibility that thermal renovation of the dwellings at the end of this period (e.g. placing a heat pump) would then not be borne out by the actual data (measuring the old gas boiler). A sensitivity analysis on the sample showed that only slightly more than 300 cases may be concerned, and as such a small proportion of the total sample. In any case, these data were the best available, because the direct metering of energy use for such a large sample cannot be achieved.

Notwithstanding these limitations, we believe that for the first time this study provides useful information from a very large sample and gives an indication of the further research required and the effectiveness of energy-saving policies.

§ 2.8 Conclusion

It appears from this research that the energy label has some predictive power for the actual gas consumption. However, according to the labels, dwellings in a better label category should use on average significantly less gas than dwellings with poorer labels, which is not the case. The actual heating energy consumption is on average lower than theoretical consumption levels for most buildings (in our study for dwelling with labels C to G) as was observed previously by Guerra Santin and Itard (2012), Tigchelaar et al. (2011), Cayre et al. (2011) and Hens et al. (2010). Guerra Santin already pointed out that at a lower EPC value, the difference between the expected and actual consumption will be smaller. Our study has proved this, and showed that even in very energy-efficient buildings actual gas consumption can exceed the predicted levels (Figure 7). On the other hand, less energy-efficient dwellings are predicted to use more gas than they actually do: theoretical gas consumption seems to be around twice the actual levels. Unlike gas consumption, the discrepancies between theoretical and actual consumption for electricity are relatively constant for all the different categories (Figure 10) and part of the difference is probably caused by electricity consumption by household appliances. The fact that labelled dwellings vary in terms of gas consumption but not much when it comes to electricity consumption proves that the energy label can (on a large scale) only be efficient in reducing gas consumption, at least as long as gas remains the main source of heating energy. However, in Figure 13 one can see the importance of electricity in the carbon footprint of households – it accounts for more than one third of all CO_2 emissions, which is why efforts should be made in the future to reduce not just the demand for heating from households, but also the demand for electricity.

An important finding of this study is that the reduction in primary energy consumption, which is assumed to happen when improving a building from label G towards label A, turns out to be much lower in reality than expected. This could easily lead to inaccurate estimations of the payback times for measures taken to improve the energy-efficiency of dwellings and achieve the targets that have been set for primary energy as well as for reducing CO_2 emissions. From our calculations based on actual energy consumption, it seems that these targets may be unrealistic. Calculations were conducted in order to assess the broad feasibility of the energy (and CO_2) reduction targets set for the built environment, with the assumption that the Dutch housing stock as a whole was labelled and the average consumption values described in section 2.5 apply. It was discovered that even if the whole Dutch housing stock were refurbished and upgraded to an A label (which would in itself be an unrealistically ambitious undertaking), the actual primary energy savings would not meet most of the current targets (Table 4). However, if the theoretical levels of consumption are used, most of the targets seem (misleadingly) achievable. The targets for gas consumption and reduction in CO_2 emissions turned out to be similarly problematic. In the future, the actual energy consumption of houses should be taken into account when formulating targets. This way, measures developed to meet the targets will have a better chance of success.

The question remains of whether it makes sense to indicate the theoretical gas and electricity consumption on the label as has been done in the Netherlands since 2010. This may cause confusion instead of assisting the occupant, because it is not representative of actual values. A dwelling with a good label does not necessarily mean low energy usage. The label gives an approximate indication of the thermal quality of the dwelling but cannot predict the real energy consumption.

As a final remark, more research on the relationship between policy instruments and their effects is needed to validate the efficiency of these instruments and improve them. Simulation tools (such as the Dutch energy labelling method) are often used to support policy. However, these simulation tools do not always provide results that correspond to reality. This is not surprising because much is still unknown, especially in the field of statistically valid and standardized dwelling use and the relationships between dwelling use, dwelling type and occupant characteristics. However, the alternatives to simulation methods (as used in some countries), such as energy labels calculated on the basis of the actual energy consumption of the former occupant or based solely on insulation values, are not expected to produce more accurate results.

§ 2.9 References

Andaloro, A. P. F., Salomone, R., Ioppolo, G., Andaloro, L., 2010. Energy Certification of Buildings: A Comparative Analysis of Progress Towards Implementation in European Countries, Energy Policy

Volume 38, Issue 10, October 2012, Pages 5840-5866.

Berkhout, P. H. G., Muskens, J. C., Velthuijsen, J. W., 2000. Defining the rebound effect, Energy Policy, Volume 28, Issues 6–7, June 2000, Pages 425-432.

Betrouwbaarheid van energielabels bij woningen, VROM-Inspectie, June 2010.

Branco, G., Lachal, B., Gallinelli, P., Weber, W., 2004. Predicted versus observed heat consumption of a low energy multifamily complex in Switzerland based on long-term experimental data, Energy and Buildings, Volume 36, Issue 6, June 2004, Pages 543-555.

Brounen D., Kok, N., 2010. On the economics of energy labels in the housing market, Journal of Environmental Economics and Management, Volume 62, Issue 2, September 2011, Pages 166–179.

Cayre, E., Allibe, B., Laurent, M. H., Osso D., 2011. There are people in this house! How the results of purely technical analysis of residential energy consumption are misleading for energy policies, Proceedings of the European Council for an Energy Efficient Economy (eceee) Summer School, 6–11 June 2011, Belambra Presqu'île de Giens, France.

CBS (Statistics Netherlands) database, accessed on 9th April 2012, http://statline.cbs.nl/statweb/

Convenant Energiebesparing Corporatiesector, October 2008, accessed on 9th April 2012 on http://www.aedesnet.nl/binaries/downloads/2008/10/20081009-convenant-energiebesparing-corporatiesect.pdf

Country specific factors – Report of Findings in WP3 (Deliverable 3.1), IDEAL EPBD European Project on Consumer response to energy labels in buildings, May 2009.

Convenant Energiebesparing bestaande gebouwen ("Meer met Minder"), January 2008, accessed on 15th May 2012 on http://www.vng.nl/PDO/PDO_ConvenantMmM.pdf

Decision No. 406/2009/EC of the European Parliament and of the Council, 23 April 2009.

Derde onderzoek naar de betrouwbaarheid van energielabels bij woningen, VROM-Inspectie, August 2011.

Europe's buildings under the microscope, country-by-country review of the energy performance of Europe's buildings, Building Performance Institute Europe (BPIE), 2011, accessed on 9th April 2012 on http://www.buildup.eu/publications/19446

Energiecijfers database, accessed on 9th April 2012, http://senternovem.databank.nl/

Energie NED, De Nederlandse Energiebranche website with data resources: www.energiened.nl , accessed 28th March 2012.

Energy Efficiency Policies and Measures in The Netherlands Monitoring of Energy Efficiency in EU 27,

Norway and Croatia, ODYSSEE-MURE, ECN, 2009, accessed 16th March 2012 on http://www.odyssee-indicators.org/publications/PDF/netherlands_nr.pdf

Energy Performance Certificates across Europe: From design to implementation, Building Performance Institute Europe (BPIE), December 2010.

Directive 2002/91/EC of the European Parliament and of the Council of 16 December 2002 on the energy performance of buildings

Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings

Evaluation of the impact of national EPBD implementation in MS, ASIEPI EU project, December 2009.

Guerra Santin, O., 2010. Actual energy consumption in dwellings; The effect of energy performance regulations and occupant behaviour, PhD thesis.

Guerra Santin, O., Itard, L., 2012. The effect of energy performance regulations on energy consumption, Journal of Energy Efficiency, 8 February 2012.

Guide for a buildings energy label: Promoting bioclimatic and solar construction and renovation, PREDAC WP4 Report, 2003, accessed on 9th April 2012, http://www.cler.org/info/IMG/pdf/WP4_guide.pdf

Haas R., Biermayr, P., 2000. The rebound effect for space heating Empirical evidence from Austria, Energy Policy, Volume 28, Number 6, June 2000 , Pages 403-410.

Hens, H., Parijs, W., Deurinck, M., 2010. Energy consumption for heating and rebound effects, Energy and Buildings, Volume 42, Issue 1, January 2010, Pages 105-110.

Implementing the Energy Performance of Buildings Directive (EPBD) – Featuring country reports 2010, EPBD Concerted Action, April 2011.

ISSO 82.3 Publication Energy Performance Certificate - Formula structure (*Publicatie 82.3 Handleiding EPA-W (Formulestructuur*'), Senternovem, October 2009.
Campaigning for the Future: Different approaches, unexpected results, Presentation of the experiences and outcomes of the project IMPLEMENT, March, 2010.
Kjærbye, V., 2008. Does energy labelling on residential housing cause energy savings? Copenhagen: AKF, Danish Institute of Governmental Research.
Lente-akkoord Convenant, accessed on 9th April 2012 on http://www.lente-akkoord.nl/wp-content/uploads/2009/04/Convenant-Energiebesparing-in-nieuwbouw.pdf
Marchio, D., Rabl, A., 1991. Energy –efficient gas heated housing in France: Predicted and observed performance, Energy and Buildings, Volume 17, Pages 131 – 139.
Milieucentraal website, accessed on 12th May 2012 on http://www.milieucentraal.nl
NEN 5128 norm (2001).
Perez, L., Ortiz, J., Gonzales, R., Maestre L.R., 2009. A review of benchmarking, rating and labelling concepts within the framework of building energy certification schemes, Energy and Buildings, Volume 48, Pages 272-278.
Rapportage Gebruik en betrouwbaarheid energielabels bij woningen, VROM-Inspectie, May 2009.
Sectoral Emission Reduction Potentials and Economic Costs for Climate Change (SERPEC-CC), Summary report, B. Wesselink, Y. Deng, October 2009, accessed on 9th April 2012 on http://www.ecofys.com/files/files/serpec_executive_summary.pdf
Sorrell, S., Dimitropoulos, J., Sommerville, M., 2009. Empirical estimates of the direct rebound effect: A review, Energy Policy, Volume 37, Issue 4, April 2009, Pages 1356-1371.
Tigchelaar, C., Daniëls, B., Maenkveld, M., 2011. Obligations in the existing housing stock: Who pays the bill?, Proceedings of the European Council for an Energy Efficient Economy (eceee) Summer School, 6–11 June 2011, Belambra Presqu'île de Giens, France.
Yücel, G., Pruyt, E., 2011. Energy Transition in The Dutch Dwelling Stock: Exploring the Extent of Inertia Against Change, Proceedings of International System Dynamics Conference 2011, July 24 – 28 2011, Washington DC, USA.

3 Actual and theoretical gas consumption in Dutch dwellings: What causes the differences?

Explanatory note

The results of chapter 2 show that the discrepancies between theoretical and actual gas consumptions were quite significant and have a substantial impact on the energy savings targets set by the government. Therefore, the next logical step was to find out why the discrepancies occur. The third chapter investigates the same dataset used in the second chapter, this time with the intention of gaining insight into the causes of the discrepancies discovered. National socioeconomic data were added to the studied sample and a regression analysis was carried out. Due to the low predictive power of the included variables, a sensitivity analysis of the theoretical gas usage was performed on the basis of the average row house taken from the WOON dataset. Sensitivity analysis was performed on six assumptions made in the theoretical calculation to show how an incremental difference in one of the assumptions affects the final theoretical gas consumption and whether this could explain the performance gap.

Published as: 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.

Abstract

Energy labels in buildings are awarded based on theoretical gas and electricity consumption based on dwelling's physical characteristics. Prior to this research, a large-scale study was conducted in The Netherlands comparing theoretical energy use with data on actual energy use revealing substantial discrepancies (Majcen et al., 2012). This study uses identical energy label data, supplemented with additional data sources in order to reveal how different parameters influence theoretical and actual consumptions gas and electricity. Analysis is conducted through descriptive statistics and regression analysis. Regression analysis explained far less of the variation in the actual consumption than in the theoretical and has shown that variables such as floor area, ownership type, salary and the value of the house, which predicted a high degree of change in actual gas consumption, were insignificant (ownership, salary, value) or had a minor impact on theoretical consumption (floor area). Since some possibly fundamental

variables were unavailable for regression analysis, we also conducted a sensitivity study of theoretical gas consumption. It showed that average indoor temperature, ventilation rate and accuracy of U-value have a large influence on the theoretical gas consumption; whereas the number of occupants and internal heat load have a rather limited impact.

§ 3.1 Introduction

Buildings account for approximately 40% of the EU's total energy consumption. One way of achieving a significant reduction in energy demand of the residential sector is to inform tenants and homeowners of the energy consumption of their dwelling. The European Performance of Buildings Directive was passed in 2002, setting up an EU framework for energy performance certification. The directive introduced mandatory energy performance certification (labelling) for all residential buildings at the time of construction, sale or rental. The Netherlands' energy label is based on the 'Decree on Energy Performance of Buildings' (BEG) and the 'Regulation on Energy Performance of Buildings' (REG) national requirements which came into force in 2008 (Beerepoot, 2007). The Dutch energy label certificate allocates each home into a category, ranging from 'A++' to 'G', and states its expected (theoretical) energy consumption.

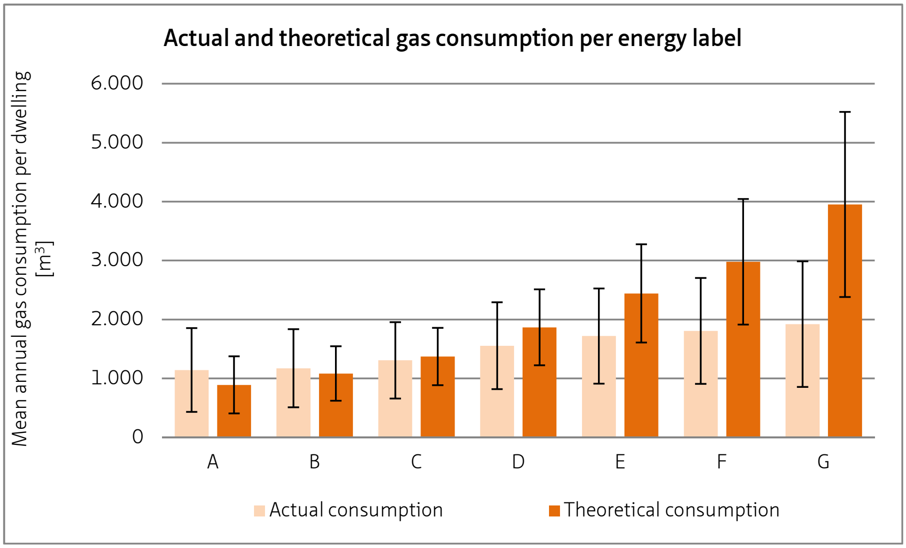


FIGURE 1 Actual and theoretical gas consumption in dwellings across label categories with ± 1 standard deviation (Majcen et al., 2012)

The motivation for the present study was a previous paper by Majcen et al. (2012), which compared the theoretical energy consumption stated on nearly 200,000 energy label certificates issued in the Netherlands with the actual consumption of those dwellings. The results showed that in energy-inefficient dwellings (labelled F or G), predicted gas consumption (gas is the chief energy source for heating in the Netherlands) was much higher than the actual rates of consumption, while energy-efficient dwellings (labelled A or B) consumed slightly more than predicted. For label C dwellings, actual and theoretical gas consumption match relatively well (Figure 1).

While it is clear that the calculation method implemented to certify dwellings is simplified and therefore deviates from actual dwelling consumption on the level of individual dwelling due to assuming zero variation in climate and occupant characteristics, the average actual consumptions of a certain label category should coincide with the theoretical consumptions declared on the certificate. If it is not the case, it disables an estimation of actual energy savings when improving the label of the dwelling (Majcen et al., 2012), which is the final aim of such an energy label.

Actual vs. theoretical heating energy consumption

Results similar to those shown in Figure 1 were obtained in numerous studies across Europe, including those by Guerra Santin and Itard (2012), Tigchelaar et al. (2011), Cayre et al. (2011) and Hens (2010) about the overestimation of heating energy consumption in energy-inefficient dwellings and Haas and Biermayr (2000), Branco et al. (2004) and Marchio and Rabl (1991) concerning the underestimation in energy-efficient dwellings. These examples and the study by Majcen et al. (2012) seem to show that the theoretical consumption, which is calculated using various design and policy-based calculation tools, often fails to represent the actual energy consumption of residential buildings accurately. A study in Norway (Pettersen, 1994) showed that total heating energy consumption cannot be predicted more precisely than approximately 35-40%, which corresponds with the case-study by Majcen et al. (2012) and others previously mentioned cases of discrepancies. The causes for these discrepancies are complex. One of them is the variation in presence patterns and comfort. Under many calculation methods, in particular those used for certification, this variation is deliberately ignored in order to produce a standardised measure of the thermal properties of the dwelling. Nevertheless, in many countries, including the Netherlands, the theoretically estimated consumption shown on the label certificate is the basis on which the energy savings of potential renovation measures are calculated. This calls for a theoretical consumption that corresponds to a dwellings' actual consumption better than demonstrated in Figure 1. To arrive at a more accurate theoretical consumption, Gaceo et al. (2009) calculated energy consumption by what he called 'specific user' profiles. Unlike the 'average user' profiles that are usually used for energy performance calculations, using the specific profiles resulted in a much more accurate estimate of energy consumption. However, the effects of occupant behaviour are complex and depend on environmental factors such as climate (Pettersen, 1994) and the characteristics of the building (Guerra Santin, 2010). For example, households with a programmable thermostat are more likely to keep the heating on for longer than households with a manual thermostat (Guerra Santin, 2010). It is therefore not only occupant preferences, but also the characteristics of the dwelling that can explain the variation in the accuracy of predictions across the range of label categories (Figure 1). Furthermore, evidence shows that occupants tend to increase their comfort demands when the efficiency increases, which in the literature is referred to as the 'rebound effect'. An overview of studies regarding the rebound effect in residential heating was conducted by Greening et al. (2000) and according to Haas and Biermayr (2000), the rebound effect can amount to 20-30% of the energy savings gained through a retrofit. A study conducted in the UK by Milne and Boardman (2000) estimates that at an indoor temperature of 16.5°C, 30% of the benefits gained through energy-efficiency improvements are offset because the residents are likely to want to raise the temperature of the dwelling further, meaning that the full energy saving will only be gained while implementing saving measures at an average indoor temperature of 20°C.

Furthermore, the results presented by Majcen et al. (2012) raise questions about the methods in place for predicting theoretical levels of consumption. Even now, there is little information available regarding the reliability of energy performance certificates, how they relate to the state of the building and the accuracy of the calculation methods. No validation of the calculation methods used in the Netherlands or elsewhere in Europe has been found in literature.

Inaccurate estimates of spending on energy can also hamper the process of estimating the potential savings, which seems to be a problem across the EU. In Ireland, a 20% reduction target was set for 2002, relative to the old regulations in place from 1997, but a reduction of only 10% was achieved, according to Rogan and Gallachóir (2011). Majcen et al. (2012), examined the discrepancies between the actual and theoretical energy consumption with respect to the national targets set for energy and CO_2 reduction in the residential sector in the Netherlands. It was established that most policy targets for energy and CO_2 emissions can be achieved by extrapolating the theoretical consumptions of the dwelling stock, but if actual consumptions are used, almost none of the reduction targets for the next 20 years are achievable.

This study aims to gain a better understanding of the major discrepancies between theoretical and actual gas consumptions by looking at the influence of building and household characteristics on theoretical and actual gas consumption rates. A regression analysis explores the predictors of theoretical and actual rates of gas consumption and the differences between them. We then seek to gauge the impact of the quality of the input and of the assumptions made in the calculation method by analysing the sensitivity of the calculation model. The results will give us a better insight into actual household energy consumption and the sensitivity of the calculation models, and will therefore help us to improve labelling certificates.

The paper is structured as follows. Section 3.2 provides a brief overview of the Energy Labelling Framework in the Netherlands. Section 3.3 presents the sample data, the research methods and the regression analysis. The results and methods of the sensitivity analysis are given in Section 3.4. Finally, a discussion follows in Section 3.5 and our conclusions are presented in Section 3.6.

§ 3.2 The method used to calculate the energy label and the data used

§ 3.2.1 Calculation method

The Dutch energy label provides the following information on the dwelling for the consumer: the label category (A++ to G), the floor area, the type of dwelling, the consumption of gas [m³], electricity [kWh], heat [GJ] and the total primary energy consumption [MJ]. The label categories are determined using the energy index, which is calculated on the basis of total primary energy usage, summing up the primary energy required for heating, hot water, pumps/ventilators and lighting, and subtracting any energy gains from PV cells and/or cogeneration as shown in equation 1 (ISSO, 2009). Any energy needed for cooling is not included in this calculation method.

$$Q_{total} = Q_{space\ heating} + Q_{water\ heating} + Q_{aux.energy} + Q_{lighting} - Q_{pv} - Q_{cogeneration}$$

EQUATION 1

If no additional heat is consumed (from district heating for example), the total primary energy consumption can also be expressed as described in equation 2. The primary energy consumption is calculated according to the type of fuel used by the installations in the dwelling (Equation 3 and Equation 4). Since primary energy is a form of energy that is found in nature and has not been subject to any conversion or transformation process, appropriate heating values need to be taken into account when calculating it. The assumed heating value for gas is 35.17MJ/m³ (north sea gas). The efficiency of the electricity network is considered to be 0.39. On the Dutch label certificate, the theoretical gas and electricity consumption from equation 2 are presented.

$$Q_{total}[MJ] = Q_{total,gas}[m^3] \cdot 35.17 \left[\frac{MJ}{m^3}\right] + Q_{total\ el.}[kWh] \cdot 3.6 \left[\frac{MJ}{kWh}\right] : 0.39$$

EQUATION 2

$$Q_{total\ gas} = Q_{gas-\ space\ heating} + Q_{gas-tap\ watwer}$$

EQUATION 3

$$Q_{total\ el.} = Q_{el.-\ space\ heating} + Q_{el.-tap\ water} + Q_{el.-\ aux.energy} + Q_{el.-lighting} - Q_{el.-\ pv} - Q_{el.-\ cogeneration}$$

EQUATION 4

The energy used for heating ($Q_{space\ heating}$, Equation 5) depends on the demand for space heating, the efficiency of the distribution system and the efficiency of the heating installation equipment. The efficiency of the distribution system ($\eta_{dist,system}$) ranges up to 1 for a dwelling where the temperature setting is optimal, there is individual metering and there is insulation on the ducts. The efficiency of the installation system ($\eta_{installation}$) may be higher than 1 in case of heat pumps, however. The potential contribution of a solar boiler (using a table of standard gains per m^2 of collectors) is accounted for. The energy needed for the pilot flame is assumed to be 2,500MJ. The space heating energy does not depend on the number of occupants.

$$Q_{space\ heating} = \frac{\frac{Q_{space\ heat\ demand}}{\eta_{dist.\ system}} - Q_{solar\ boiler}}{\eta_{installation}} + Q_{pilot\ flame}$$

EQUATION 5

The demand for space heating (Equation 6) is a sum of losses through transmission and ventilation, taking into account solar and internal heat gains. The equations below are simplified for a system without heat recovery and with natural ventilation. Transmission rates ($Q_{transmission\ loss}$) are calculated on the basis of an annual heating period of 212 days and a constant average indoor temperature of 18°C. The useful floor area consists of the heated rooms (bedrooms, living room, kitchen), plus some areas that are occasionally heated (halls, toilet, washing room, storage). Basements, attics and garages are generally not included. Heat losses through ventilation

($Q_{ventilation\ loss}$) are calculated using standard ventilation coefficients (f_1, f_2), which depend on the ventilation type and the infiltration rates. Ventilation losses are relative to the type of dwelling ($q_{reference}$) since for each type of dwelling, characteristic lengths of frames, joints etc. are assumed (ISSO, 2009). Internal gains ($Q_{internal\ gain}$) are assumed to be 6W/m² of floor area during the heating season (212 days) and solar gains ($Q_{solar\ gain}$) are based on the g-value of the glass. The assumed gains for lighting are 6 kWh/m². Heat gains from the sun are taken into account during the heating season at a constant rate of 855MJ/m² on a south vertical surface, accounting for frames and dirt on the glass.

$$Q_{space\ heat\ demand} = Q_{transmission\ loss} + Q_{ventilation\ loss} - Q_{internal\ gain} - Q_{solar\ gain}$$

$$Q_{transmission\ loss} = (\sum_{k=1}^{K} a_k \cdot A_k \cdot U_k) \cdot (T_{indoor} - T_{outdoor}) \cdot t_{duration\ heating\ season}$$

a_k – weigh factor for each surface, 0 if it borders on heated space, 1 if unheated

A_k – area of each surface$[m^2]$

U_k – U value of each surface $[W/m^2K]$

$$Q_{ventilation\ loss} = c.f \cdot (T_{indoor} - T_{outdoor}) \cdot t_{duration\ heating\ season}) \cdot \rho_{air} \cdot c_{air} \cdot q_{v,i}$$

c. f – correction factor, set to 1 in EPA

ρ_{air} – air density 1,2 $[kg/m^3]$

c_{air} – air heat capacity 1000 [J/kgK]

$$q_{v,i} = f_1 \cdot A_{floor} + f_2 \cdot q_{reference} \cdot \frac{A_{floor}}{A_{reference}}$$

$f_1 - 0{,}47\ in\ case\ of\ natural\ ventilation\ [dm^3/s \cdot m^2]$

$f_2 - 0{,}13\ in\ case\ of\ natural\ ventilation\ [dm^3/s \cdot m^2]$

$A_{reference} - 120m^2 for\ a\ detached\ house$

$q_{reference} - 310\ dm^3$

$/s\ for\ a\ detached\ house\ with\ a\ pitch\ roof\ a\ with\ draft\ proofing$

EQUATION 6

The energy consumption for hot tap water (Qwater heating) takes into account the main hot water installation and the auxiliary kitchen boiler (which, if present, is assumed to have a standard consumption of 8164.1MJ/year). Again, the standard efficiency of the installation system is applied (see Equation 7). The equations given below relate to a condensing boiler. As stated in ISSO (2009), standard hot water consumption is determined on the basis of the national average.

$$Q_{water\ heating} = Q_{main\ boiler} + Q_{auxiliary\ kitchen\ boiler}$$

$$Q_{main\ boiler} = \frac{c_f \cdot TAP}{\eta_{boiler}} \cdot r_{tap} + Q_{standing\ still} + Q_{stand.circulation\ loss} \cdot \frac{A_{floor}}{100} \cdot (1 - \eta_{useful})$$

c_f – conversion factor [MJ · day/l · year]

TAP – quantity of water [l · day]

η_{boiler} – boiler efficiency – 0.9 in case of a condensing boiler

r_{tap} – correction factor for short piping – 0.9 if $< 5m$, *else* 1

$Q_{standing\ still}$ – 4220,2MJ in case of a condensing boiler

$Q_{stand.circulation\ loss}$ – 10000MJ if non insulated, 4000MJ if insulated

η_{useful} – used part of the loss (0.44)

$$TAP = c_{kitchen} + c_{basins} + N_{people}(c_{per\ person} + c_{shower} \cdot F_{saving\ head} \cdot D + c_{bath} \cdot B \cdot B_{yes/no}$$

$c_{kitchen}$ – 13,03 for a condensing boiler [l · day]

c_{basins} – 3,97 for a condensing boiler [l · day]

$c_{per\ person}$ – 7,1 for a condensing boiler [l · day]

c_{shower} – 20,8 for a condensing boiler [l · day]

$F_{saving\ head}$ – saving shower head, if present 0.9 else 1

D – number of showers /person/day – 0.61

c_{bath} – 41,5 in case of condensing boiler [l · day]

B – number of baths /person/day – 0.096

$B_{yes/no}$ – presence of bath, if present 1 else 0

EQUATION 7

The auxiliary energy needed for the kitchen boiler is also determined using standard values. The number of people in equation 7 is determined using the following table:

FLOOR AREA	NUMBER OF PEOPLE (EPA)
<50 m^2	1.4
≥50 m^2 and <75 m^2	2.2
≥75 m^2 and <100 m^2	2.8
≥100 m^2 and <150 m^2	3.0
>150 m^2	3.2

TABLE 1 Number of people in a dwelling according to EPA calculation

The energy index (EI, Equation 8) correlates directly with the total primary energy consumption, but is corrected for the floor area of the dwelling and the corresponding heat transmission areas (Equation 2) in order not to disadvantage larger dwellings and dwellings with a greater proportion of their heat envelope adjoining unheated spaces (different building types) with constant insulation properties and efficiencies of the heating/ventilation/lighting system. Shape correction is also applied when considering infiltration losses within demand for space heating – the air permeability coefficient depends on the building shape factor.

$$EI = \frac{Q_{total}}{155 \cdot A_{floor} + 106 \cdot A_{loss} + 9560}$$

EQUATION 8

§ 3.2.2 Assumptions in the calculation method and accuracy of the inspection data

Many assumptions are made in the theoretical calculations which could lead to inaccuracies in the estimates of theoretical gas consumption. The calculation of theoretical values assumes that the whole floor area of a dwelling is heated, which may in some cases result in a significant overestimation of the demand for heating. In newer dwellings the whole surface area is likely to be heated and the indoor temperature is likely to be more uniform than in older dwellings (Guerra Santin et al., 2009). In older dwellings, especially where only the living room is heated with an old-fashioned stove (powered by wood, oil or gas), the heated surface area may be lower leading to a much lower average indoor temperature than assumed in the calculation method.

Moreover, it is possible that the estimated insulation values for walls in new dwellings is closer to the actual values than those for old dwellings. Inspecting older dwellings is often difficult and instead of measuring U-values, a guess is made as to whether the cavity walls were insulated at the time of construction and what the quality of that insulation may be after many years. In some cases, therefore, it is possible that older buildings are better insulated than is assumed.Compared to insulation of the wall, distinguishing between U values of windows is relatively easy when deciding between single, double or triple glazing but equally complicated when it comes to determining the exact type of double glazing (for example in case of low emissivity coating or gas filled cavity). Differences may also occur due to different assumptions when it comes to rates of ventilation and infiltration. It may be that less air enters older dwellings through natural ventilation than is assumed, and this may also influence the accuracy of the estimated demand for heat. On the other hand, it is possible that air flows with mechanical ventilation are underestimated, explaining the underestimates for labels A and B.

In addition, behaviour also influences temperature preferences, heated floor area, ventilation preferences and the internal heat gains of a dwelling, but in the theoretical calculations these are all assumed to be constant or a function of floor area.

The behaviour assumptions in itself are not problematic, since they were introduced in order to make dwellings comparable within the dwelling stock. However, the fact that the energy consumption calculated under these assumption is nearly double than actual in label G and roughly a third lower in label A (Figure 1) suggests that the assumptions used might not fit every label category equally well. As stated previously, if a label certificate is to inform about the quantity of the dwellings' heating energy consumption, these discrepancies should not occur.

§ 3.2.3 Energy label dataset

The Energy Label database, the core database used for this report, was provided by NL Agency – a public sector organisation that serves the Netherlands Ministry of the Interior and Kingdom Relations. The database contains all the energy labels issued from beginning of January until December 2010, including information about the installation for space heating, the dwelling type, its theoretical energy consumption (gas and electricity), floor area, construction and renovation year, date of labelling and the coded address variable to enable matching with other data. One limitation of this study was the missing information about hot tap water installation and ventilation systems.

This dataset was linked to actual energy use data for the year 2009, which was provided by CBS (Statistics Netherlands). CBS collects this data from the energy companies; however, it is important to note that the annual data is sometimes an extrapolation of monthly values. Unfortunately, not much is known about the reliability of this data, but our assumption was that the data yields reasonably accurate averages. The combined dataset was cleaned up (doubled addresses and incomplete cases were deleted), leaving 247,174 cases. The CBS expressed its doubts about the quality of the data obtained for the actual energy of collective installations due to the fact that this type of installation is arbitrarily assigned to buildings with a heat consumption that is too high to be considered realistic for an individual system. It was therefore decided to omit households with collective installation systems from the analysis. Dwellings which have multiple installation systems were also omitted since these are very specific cases. Cases where electricity consumption was null were also removed. At this point, the gas values which were defined as missing were investigated. It turned out that most of them belonged to dwellings with heating installations, which in fact do use gas. Such cases were deleted, with only those dwellings that use electricity as power source for heating being kept in the database. Gas consumption was then redefined to 0 for those cases. On checking the theoretical energy use and the areas of the house, outliers were detected. Cases with a floor area of over $1000m^2$ and primary energy use of more than 500,000 MJ were discarded. Finally, the actual gas consumption values for 2009 were corrected to the number of degree days used in the theoretical calculation. At the end of this process, the sample contained 193,856 cases.

The actual gas consumptions available from CBS corresponded to the climatic year of 2009. To be able to compare these values with the theoretical ones, the ratio between the actual degree days in the year 2009 and the degree days assumed in the theoretical calculation was calculated (factor f in Equation 9). The factor f was then applied to actual gas value.

$$f = \frac{DD_{theo}}{DD_{2009}} = \frac{2620}{2804} = 0.934$$

$D_{theo} = number\ of\ degree\ days\ in\ theoretical\ calculation\ (reference\ year\ De\ Bilt\ 1964 - 1965)$

$D_{2009} = number\ of\ degree\ days\ in\ 2009$

EQUATION 9 Equation 9

A possible limitation of the study is the differences between the end uses included in predicted and actual gas consumption. They both contain gas for hot tap water and space heating, but actual gas consumption also includes gas for cooking (see

Table 3 in Majcen et al., 2012). However, gas used for cooking is less than 3% of the total gas consumption.

The sample used represents slightly less than 0.3% of the total dwelling stock in the Netherlands (CBS Statline, 2012). An overview of the representativeness of the sample is available in a study from Majcen et al. (2012), which shows that the sample is representative for label categories, but less so for dwellings or ownership type. It is therefore important to note that while the results of this study are valid for this large sample of dwellings, some of the variables used might have a different predictive power when applied to the Dutch dwelling stock as a whole.

§ 3.2.4 Other datasets used in the paper

To account for the differences between theoretical and actual energy use as accurately as possible, more datasets were obtained from the CBS and matched with the basic database of 193,856 cases mentioned earlier. Up to date housing register (*Woonruimteregister*), municipal records (*Gemeentelijke Basisadministratie*), employment database (*Social Statistisch Bestand Banen*), and the 'Woon' survey conducted by the Dutch government in 2009, were coupled with the energy label data.

For the regression analysis presented in the section 3.3, the energy label dataset was coupled with the first three mentioned databases, leaving a total of approximately 40,000 dwellings. In section 3.4, in which a sensitivity analysis was performed on the methodology, the Woon database was also incorporated. Consequently, the sample used for that purpose included around 700 terraced houses.

§ 3.3 Regression analysis

§ 3.3.1 Methodology

The goal of the regression analysis was to see how much variation can be explained and which variables have the best predictive power for theoretical and which for actual gas consumption. In addition to the variables used, many variables which could have been relevant to our analysis, such as the presence of hot water taps and ventilation systems, were not available. Variables that could relate indirectly to occupants' behaviour, are grouped under the category of 'household characteristics'. All other variables belong to the 'building characteristics' group (Table 2). Variables that describe occupant behaviour directly, such as indoor set temperature, the presence of thermostat, time spent at home, heating bedrooms, and so on, could have been very relevant but these are typically survey questions, and a survey large enough to give relevant regression results when coupling with our database was not available.

After preliminary analysis, the data was found to be suitable for parametric analysis. Forced entry regression analysis was conducted on actual and theoretical gas consumption per dwelling. Dummy variables were defined for the categorical variables in order to include them in the regression analysis (Table 2). If all dummy values of a categorical variable were insignificant, this variable was not retained in the regression analysis (in the cases of dwelling and installation type, only a few dummy values are insignificant, therefore these variables were retained). The variables which were found to be insignificant according to the criteria mentioned were omitted and forced entry regression was repeated without these variables. In Table 2, the dummy values that are not significant (sig. above 0.01) are highlighted.

Multicolinearity among the predictors was generally not an issue, with a slightly higher correlation detected between label category E, F and G and the vintage of the dwellings. However, these correlations were in the range of 0.2 – 0.25, which is still considered a weak correlation and did not disturb the regression analysis (Field, 2009).

§ 3.3.2 Results

The results for gas consumptions can be found in Table 2. A much higher degree of response variation of the theoretical gas consumption can be explained by the regression model (87.9%) than is the case for the actual gas consumption, for which only 50.5 % of response variation can be explained.

Floor area, label and vintage

Floor area is a good predictor of theoretical and actual gas consumption (Table 2). We can interpret these results as meaning that for every $10m^2$ added to the size of the dwelling, theoretical gas consumption increases by $12.1m^3$, but the actual increase is only about $6.7m^3$. This means that in larger dwellings, the difference between theoretical and actual gas consumption is relatively larger than in small dwellings. It can be concluded that a larger floor area does not raise actual energy use as much as the estimates would lead us to believe. This could be due to occupant behaviour: in large houses, it is unlikely that all rooms will be heated evenly.

The age of the dwellings is a significant predictor of theoretical and actual consumption, predicting a similar increase in both. Each higher label is a stronger predictor in actual gas consumption, meaning that the label accounts for the thermal quality in the correct order. However, the beta values are smaller than those for the theoretical consumption, meaning that the label correlates with the theoretical consumption more strongly than with the actual.

Dwelling and installation type

Terraced houses located on corners and ground-floor flats surrounded by two others have higher gas consumption in both actual and theoretical terms, if a detached house is used as the reference dummy variable. Considering the geometry of these dwellings, this result was unexpected (detached houses have the least favourable shape in terms of heat conservation). This phenomenon is probably due to the fact that some of the variation is taken on by other predictors used in the analysis. For the other dwelling types, the variation was as expected according to their geometry. The predictive power of dwelling types was very similar for both actual and theoretical gas consumption, which suggests that dwelling type is not responsible for the large discrepancies seen in Figure 1.

Regarding the installation types used in the regression analysis, the reference dummy is an improved efficiency boiler. A negative beta power would be assumed for higher-efficiency condensing boilers, which is the case for the actual consumption, but strangely not for the theoretical consumption. However, regression analysis is only valid for the specific combination of predictors and does not necessarily mean that gas consumption will be higher for high-efficiency boilers than in improved-efficiency boilers. In general, the installation type seems to be a considerably worse predictor for actual gas consumption than for theoretical gas consumption. Since many dummy values are not significant predictors of actual gas consumption, the mean gas consumption at different installation types is also presented in Figure 2 below. It is notable that dwellings with central electrical heating and heat pumps consume gas in non-negligible quantities. Unfortunately, the gas installation systems for hot tap water were not included in the available data, which is a limitation for this study and could explain this discrepancy, together with the insignificant results in the regression analysis. However, a more detailed investigation of the dwellings with local electrical heating showed that all 98 of these dwellings have actual gas consumption, while only 37 dwellings have theoretical gas consumption. Heat pumps are a similar case: slightly less than half of the cases have theoretical gas consumption, while all the cases have actual gas consumption (hence the large standard deviation in Figure 2). Even though the information on hot tap water was missing in the database, this clearly indicates either flaws in the inspection phase or generically inaccurate actual gas consumption data. Due to the assumptions made by energy companies, a thermal renovation of the dwellings may not be reflected in the actual data (Majcen et al., 2012). The dwellings with other installation systems also show a large difference between the mean theoretical and actual gas consumption, which is difficult to analyse because the installations for hot tap water were unknown. It seems that the less efficient the installation system, the higher the overestimation, which is a similar trend to the one seen in Figure 1. Lower labels do indeed have a higher proportion of inefficient installation systems than more efficient labels (Table 2), although there was no significant correlation. It is also possible that the heating surface area is overestimated when a gas/oil stove is in use (with these systems, probably only one or two rooms are heated rather than the whole dwelling, as assumed in the theoretical calculation). This could explain some of the theoretical overestimation in labels E, F and G.

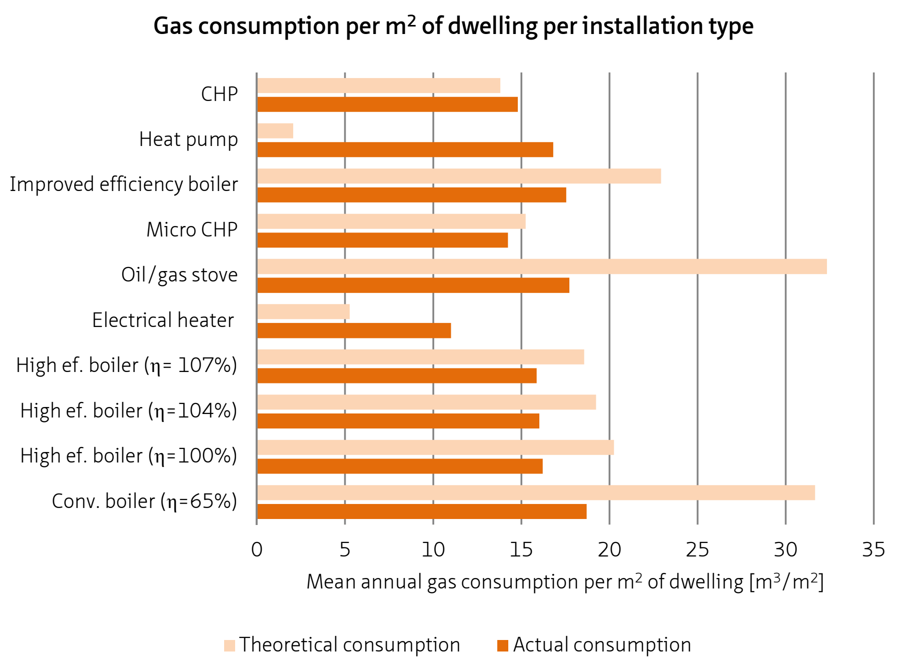


FIGURE 2 Mean annual gas consumptions per m^2 dwelling per installation type with ± 1 standard deviation (see Appendix for definitions of the installation types)

Household characteristics

The beta values for the 'private rental' dummy are insignificant, but owner-occupied dwellings seem to have a slightly higher theoretical gas consumption than the social dwellings. On the other hand, actual gas consumption in owner-occupied dwellings is about 50m^3 less than in social housing, which could be attributable to better insulation in owner-occupied dwellings or to different behaviour.

Interestingly, for a dwelling worth €100,000 or more, there will be 38 m^3 more theoretical consumption, whereas actual consumption will be 97m^3 higher. Results for salary per person in the household are similar – this predictor is insignificant for theoretical gas consumption, but an income that is €10,000 higher annually is associated with an increase in actual gas consumption of 8 m^3. Salary is not a significant predictor for theoretical gas consumption.

Whereas one additional occupant means an increase of 45m^3 in actual gas consumption, this variable fails to explain the variation in theoretical consumption, which is logical since the number of occupants is a function of floor area only, see Table 1.

Other household predictors were not significant.

	Independent Variables	GAS CONSUMPTION PER DWELLING [M^3]						
		DUMMIES	THEORETICAL [R^2=87.9%]			ACTUAL [R^2=50.5%]		
			B	BETA	SIGN.	B	BETA	SIGN.
	Constant					-143.710		0.046
BUILDING CHARACTERISTICS	Floor Area	Ratio variable	12.100	0.438	0.000	6.670	0.313	0.000
	Vintage of building	Ratio variable	1.800	0.047	0.000	1.720	0.058	0.000
	Label [ref. dummy variable is A]	B	406.440	0.124	0.000	220.890	0.087	0.000
		C	719.630	0.322	0.000	366.470	0.212	0.000
		D	1146.12	0.489	0.000	539.500	0.299	0.000
		E	1672.40	0.617	0.000	655.940	0.314	0.000
		F	2237.10	0.632	0.000	735.060	0.270	0.000
		G	3146.00	0.565	0.000	802.000	0.187	0.000

>>>

TOC

	Independent Variables	GAS CONSUMPTION PER DWELLING [M³]						
		DUMMIES	THEORETICAL [R^2=87.9%]			ACTUAL [R^2=50.5%]		
			B	BETA	SIGN.	B	BETA	SIGN.
	Constant					-143.710		0.046
BUILDING CHARACTERISTICS	Dwelling type dummy [ref. dummy variable is detached house]	Flat – corner – roof	227.800	0.036	0.000	48.330	0.010	0.031
		Flat – corner – ground floor	139.500	0.019	0.000	143.760	0.025	0.000
		Flat – corner – middle floor	-104.600	-0.017	0.000	-69.470	-0.014	0.002
		Terraced house – corner	388.400	0.166	0.000	288.700	0.160	0.000
		Terraced house – middle	64.380	0.030	0.000	16.490	0.010	0.290
		Flat – middle – roof	69.660	0.016	0.000	-101.040	-0.030	0.000
		Flat – middle – middle floor	-96.230	-0.031	0,000	-136.340	-0.057	0,000
		Flat – middle – ground floor	919.770	0.221	0,000	578.220	0.181	0,000
	Installation type dummy [ref. dummy variable is improved efficiency boiler (η=83%)]	Conv. boiler (η=65%)	-90.970	-0.013	0,000	-39.190	-0.007	0.061
		High efficiency boiler (η=100%)	36.340	0.009	0.000	-20.380	-0.006	0.120
		High efficiency boiler (η=104%)	28.800	0.003	0.060	-12.750	-0.002	0.593
		High efficiency boiler (η= 107%)	23.540	0.011	0.000	-22.450	-0.013	0.004
		Electrical heater	-1266.30	-0.038	0.000	-375.010	-0.015	0.000
		Oil/gas stove	-206.600	-0.038	0.000	-236.130	-0.056	0.000
		Micro CHP	317.500	0.005	0.008	185.780	0.004	0.320
		Heat pump	-1210.20	-0.048	0.000	150.890	0.008	0.031
		CHP	-18.000	0.000	0.832	22.410	0.001	0.865
HOUSEHOLD CHARACTERISTICS	Ownership type [ref. dummy variable is social housing]	Private rental	-2.060	0.000	0.905	18.930	0.003	0.480
		Owner-occupied	-5.540	-0.002	0.282	-48.610	-0.028	0.000
	Value (2009)	Ratio variable	0.000	0.040	0.000	0.001	0.125	0.000
	Number of people	Ratio variable	-1.380	-0.002	0.685	45.480	0.074	0.000
	Number working population per household	Ratio variable	-7.090	-0.006	0.009	-3.561	-0.004	0.400
	Salary per person	Ratio variable	0.000	-0.005	0.023	0.001	0.017	0.001
	Household type	Couple/elderly/ family	Insignificant variable					
	Household type	Ratio variable	Insignificant variable					
	Number of children in household.	Ratio variable	Insignificant variable					
	Days worked per person in household	Ratio variable	Insignificant variable					
	Overtime per person in household	Ratio variable	Insignificant variable					
	Salary per person in household	American Institute of Architects, 2002	Insignificant variable					

TABLE 2 Regression analysis of gas consumption (see Appendix for definitions). The orange values are insignificant on a 99% confidence interval scale.

§ 3.4 Sensitivity of the calculation method

In addition to the variables used for the regression analysis, parameters such as temperature preferences, time spent at home and other behavioural characteristics of the occupants could also affect the discrepancy seen in Figure 1. Moreover, it is said that the thermal qualities of the dwelling are often assessed inaccurately during the inspection and this could be another plausible explanation for the overestimation of the energy consumption of low-efficient dwellings (see section 3.1). However, no trustworthy data was available for matching with the large sample used in the regression analysis. The second part of this paper will seek to bridge this data gap by examining how changes in behaviour and assumptions related to the dwelling influence the theoretical gas consumption and whether more accurate assumptions could lead to a better match between actual and theoretical rates of gas consumption.

§ 3.4.1 Reference dwellings

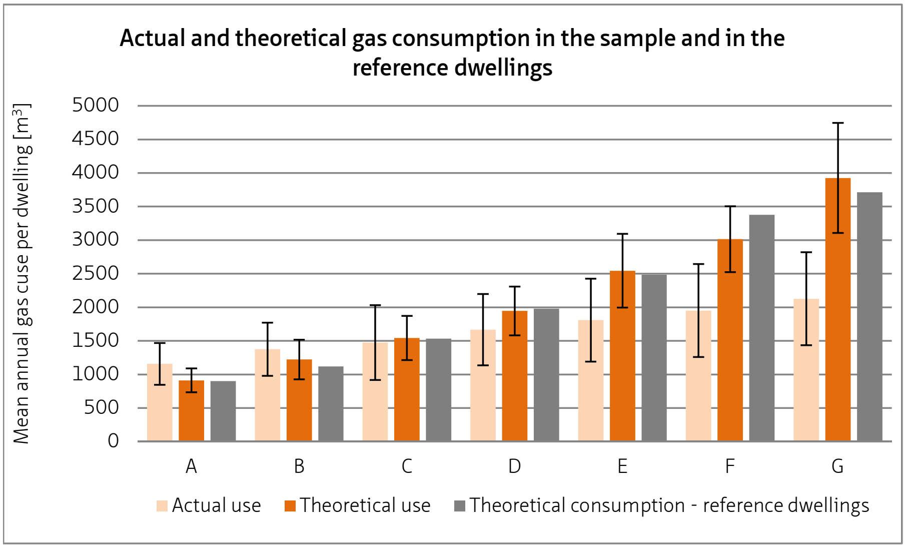


FIGURE 3 Actual and theoretical gas consumption in the sample of terraced houses together with the theoretical consumption of reference dwellings

The energy label database does not include complete information about the geometry of dwellings. In order to test how adjustments to the calculation assumptions could influence theoretical gas consumption rates, data such as the number of floors, floor area, shape of dwelling and roof type is needed; without this, the theoretical gas consumption cannot be calculated. To overcome this lack of data, the calculations were performed using reference dwellings that were representative of an average dwelling in each label category of the sample.

In order to reduce uncertainty due to the dwelling type, terraced houses (the most common type of dwelling in the Netherlands) were chosen as the subsample to be investigated in the sensitivity analysis. Because the Woon database was matched with the terraced houses from the original energy label database, the newly composed database consisted of 713 matched cases. Figure 3 shows the actual and theoretical consumption in this smaller sample, in order to ensure that this sample is representative of the trend described previously in Figure 1.

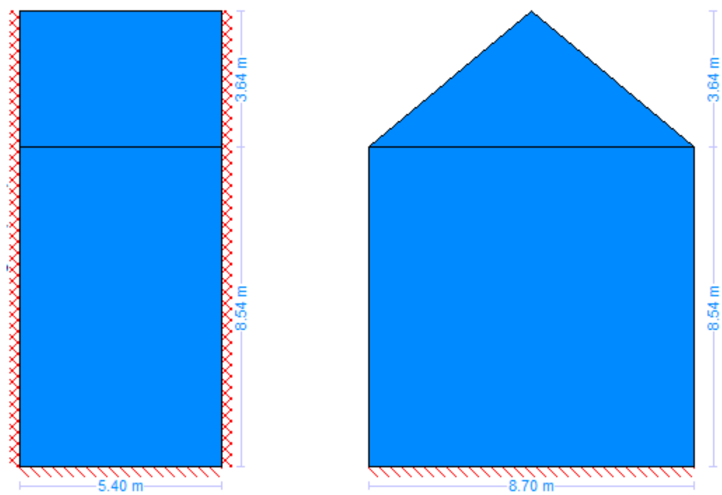


FIGURE 4 Side and front view of the reference terraced house

The Woon database provided the average number of rooms in the sample, which was used – together with the floor area, type and construction year – as the basis for the choice of the geometry of the reference dwelling (Novem, 2002). The reference geometry is shown in Figure 4. Both the front and back of the house are characterised by approximately $8m^2$ of window area. Both the side walls are shared with another heated house. The indoor floor area of the dwelling is $105m^2$. The chosen geometry was based on averages from these 713 dwellings and was used as a reference in all label categories. Although this might introduce a slight error – because in practice the geometry does correlate slightly with the energy label – the purpose was to test the sensitivity of the calculation method in different label categories due to the different thermal quality of the dwelling and not due to the varying geometry. We therefore deliberately selected a single reference geometry.

Appropriate envelope U-values were applied to the reference geometry in order to get 7 thermally representative reference dwellings, one for each label category (AgentschapNL, 2011). The installation and ventilation type in each of these dwellings was determined using the average of the available dataset (713 dwellings). The properties of the seven reference dwellings are summarised in Table 3. According to their theoretical consumption (Figure 3), reference dwellings are well representative of the sample. The assumed infiltration rate was 23.3 dm^3/s (ISSO, 2009) and the assumed ventilation flow rate was 49.4 dm^3/s (ISSO, 2009) for all the reference dwellings regardless of the label class.

	A	B	C	D	E	F	G
Installation system	HE boiler	HE boiler	HE boiler	HE boiler	HE boiler	IE boiler	IE boiler
Ventilation system	Natural	Natural	Natural	Natural	Natural	Natural	Natural
Supply water temperature of heating system [°C]	>55	>55	>55	>55	>55	>55	>55
U value wall [W/m²K]	0.2	0.36	0.5	0.64	1.6	2.0	2.4
U value window [W/m²K]	1.8	2.9	5.2	5.2	5.2	5.2	5.2
U value floor/roof [W/m²K]	0.4	0.4	0.5	1.3	1.3	1.72/1.54	1.7

TABLE 3 Properties of the seven reference dwellings for each label

§ 3.4.2 Calculation method

The gas consumption (Q) of a dwelling is a function of many parameters such as the average indoor temperature (T) (averaged out over the heated floor area and time) number of occupants (N_{people}), internal load (Q_{int}), ventilation rate (F_{vent}), floor area (A) and insulation values (U). In the theoretical gas consumption calculation, these are set at nominal values (Table 4), which are shown in superscript '^{0}'.

The first four variables in Table 4 show the behavioural assumptions made in energy label calculations. As well as occupant behaviour, poor quality of inspection could also lead to an inaccurate estimation of theoretical gas consumption, which can result in an erroneous label and contribute to discrepancies between theoretical and actual consumption, because the dwelling should actually be in another category. A sensitivity test was therefore also conducted on the insulation quality of the dwelling and the heated floor area. In section 3.4.3.1, behavioural assumptions are addressed and section 3.4.3.2 relates to the sensitivity of the floor area and the quality of the insulation.

VARIABLE	ASSUMPTIONS IN EPA CALCULATION METHOD	RESULT OF INSPECTION
T^0	18°C	N/a (fixed in method)
N^0_{people}	3.2	N/a (fixed in method)
Q^0_{int}	$6W/m^2$	N/a (fixed in method)
F^0_{vent}	Standard correction factor c.f=1	N/a (fixed in method)
A^0	All surface area is heated	$105\ m^2$
U^0	Default values vary throughout label categories (see Table 3)	Estimations from inspection or default values are used

TABLE 4 Assumptions in the EPA calculation method

According to our educated guess (see also section 0), inaccurate estimates concerning these six core parameters are very likely to be the cause of the discrepancies between the theoretical and actual rates of gas consumption. Since the software used for the energy label calculations (EPA-W by Vabi, 2011) is also used for a broad custom advice on dwelling energy consumption with which occupants are advised on how to reduce energy consumption in their particular situation, modifications could be made to the parameters from Table 4 and gas consumption was recalculated.

The new values for the parameters mentioned in Table 4 were fed into the calculation software which yielded altered gas consumption values.

$$\delta Q = Q(P^0 + \delta P) - Q(P^0)$$

EQUATION 10 Equation 10

Inversely, we also looked for the change in parameter which could explain the discrepancy and whether the gas consumption changes in a linear fashion with the altered parameter.

$$\Delta P(Q_{diff}) = \frac{\Delta Q_{diff}}{\frac{\delta Q}{\delta P}}$$

EQUATION 11 Equation 11

$$Q_{diff} = Q_{theoretical\ gas} - Q_{actual\ gas}$$

EQUATION 12 Equation 12

Each change in parameters ΔP was introduced back into the calculation software (Equation 11) to test whether the change in gas consumption was linear at that increment. Sometimes this was not the case and in such instances Equation 11 is not an accurate way to calculate ΔP, since the δQ is not a monotonous function of δP as described in Equation 10, but depends on other additional parameters. However, the purpose of the exercise was to begin to understand whether the flaws in estimations could realistically be the culprit for the discrepancies seen in Figure 1. In most cases, the calculation model behaved in a linear fashion where the solution ΔP which explained Q_{dif} was realistically possible. In cases where ΔP would have to be relatively large (for example with more than 5 occupants in a dwelling or a floor area larger than $100m^2$), gas consumption was not linear; however, such solutions are not likely to occur anyway.

§ 3.4.3 Results

§ 3.4.3.1 Behavioural parameters

Table 5 shows the sensitivity of the theoretical gas consumption model to the four behavioural parameters. The second column shows the difference between the theoretical and actual rates of gas consumption, as seen in Figure 3. The values in the δQ columns are highlighted whenever the difference exceeds the Q_{dif}. This means that the δP change in parameter would explain the difference Q_{dif}. The highlighted values in the columns $\Delta P(Q_{dif})$ signify that the gas consumption is a monotonous function and therefore the ΔP is valid.

For greater clarity, the theoretical rates of gas consumption at δP are also presented in Figure 5 together with the theoretical rate of consumption for the reference dwellings and the actual rate of gas consumption of the sample.

Indoor temperature

In the third and fourth columns, Table 5 shows the differences in theoretical gas consumption if the indoor temperature is raised or lowered from the assumed 18 degrees C by ±2°C. However, Table 5 shows that such increments can only explain the discrepancy (Q_{dif}) between the theoretical and actual rates of consumption in dwellings with labels B, C and D (highlighted). All the values in the fifth column are highlighted, because gas consumption is linear within the δT in all label categories.

TOC

On the basis of these results, indoor temperature would have to be 12.4°C in order to explain the discrepancy between theoretical and actual rates of gas consumption in dwellings with a G-label certificate. With an outdoor temperature of 5.64°C during the heating season, a heated area of 57.43m^2 would yield such an average indoor temperature (assuming very poor insulation). This is a realistic value if only the living room and some other smaller room (such as a kitchen or bedroom) are heated. On the other hand, an average indoor temperature of 20.7°C would explain the discrepancy for dwellings with an A-label certificate, which is realistic, considering these are very efficient houses with fewer temperature fluctuations and in which the occupants tend to adjust their comfort preferences upwards. However, it is likely that an inaccurate temperature estimate is not the only culprit for the difference and that a handful of factors are involved. On the basis of this table, one can say that the accuracy of temperature estimation has a major impact on the accuracy of the theoretical gas consumption.

Number of occupants

The difference in gas consumption remains fairly constant across label categories, which is because the demand for gas for space heating is independent of the number of occupants; only the demand for gas for hot tap water changes with this parameter. The demand for hot tap water is not related to the thermal properties of the dwelling (section 3.2, Equation 3). As such, theoretical gas consumption does not respond in a linear way to any change in the number of occupants (values in column $\delta N_{people}(Q_{dif})$ are not highlighted).

The inaccurate estimation assumption regarding the number of occupants is therefore unlikely to be the cause of the discrepancies. It could explain minor differences in middle-ranking labels, such as label C. In higher labels, the assumption about the number of occupants would have to be very inaccurate (over 10 occupants too many) in order to account for the overestimation. This conclusion, however, does not exclude the possibility that the algorithm for hot tap water gas demand is not representative of the actual state of the dwellings. Flaws in more complex assumptions such as the standard efficiency of boilers, average losses through piping, standard rates of hot water consumption and so on could also be the cause of inaccuracies when estimating theoretical gas consumption.

Internal heat gains

The third part of Table 5 is about internal heat gains, nominally set at 6 W/m^2. Internal heat gain influences the gas consumption for space heating, while the consumption of gas for hot tap water remains constant. The increment of 2 W/m^2 can only explain the difference in label C. If we derive the internal heat gain which would explain the difference, the value in G-label dwellings is very high, around 27 W/m^2, which again is unrealistic. Gas consumption is linear even for all positive increments in internal heat gain, but is not linear for negative increments. An inaccurate estimation of internal heat gains could therefore be responsible for moderate discrepancies.

Ventilation rate

In the energy label calculation, the correction factor for heat demand which occurs as a consequence of ventilation is set at 1. The ventilation is determined as a function of the ventilation system and infiltration rate, which is determined on the basis of nominal rates for the dwelling type and corrected for the floor area. In the sensitivity analysis, the factor was modified for increments of 0.5 upwards and downwards. Such a change in gas demand due to ventilation explains the discrepancies in labels B to D. For label A, a correction factor of 0.6 would explain the discrepancy. For label E and below, the correction factor would have to be negative to explain the difference, which is not possible (gas demand due to ventilation cannot be negative in the model).

LABEL	Q_{DIFF} [M³ GAS]	INDOOR TEMPERATURE			NUMBER OF OCCUPANTS		
		δQ*		δT (Q_{DIF})* [°C]	δQ*		$δN_{PEOPLE}$ (Q)(Q_{DIF})*
		δT = +2°C	δT= 2°C		$δN_{PEOPLE}$ =+2	$δN_{PEOPLE}$ =-2	
A	**-232.0**	175.3	-170.3	2.7	101.7	-154.2	5.1
B	-116.3	211.4	-208.6	1.1	101.7	-154.2	2.6
C	72.2	279.5	-278.7	-0.5	101.7	-153.6	-1.0
D	272.0	354.7	-354.6	-1.5	101.7	-153.6	-3.9
E	738.6	437.8	-437.6	-3.4	101.7	-153.6	-10.5
F	1081.3	583.8	-583.8	-3.7	101.7	-153.6	-15.4
G	1815.7	635.4	-644.0	-5.6	97.4	-157.9	-24.3

TABLE 5 (I) Sensitivity of gas consumption for behaviour parameters

LABEL	INTERNAL HEAT GAINS			VENTILATION RATE		
	δQ*		$δQ_{INT}$ (Q_{DIF})* [W/M²]	δQ*		$δQ_{VENT}$ (Q_{DIF})*[C.F**]
	$δQ_{INT}$= -2W/M²	$δQ_{INT}$ T= +2W/M²		$δQV_{EN}$ =0.5 C.F**	$δQ_{VEN}$ =-0.5C.F**	
A	104.3	-112.2	-4.3	279.1	-265.1	0.4
B	111.2	-115.5	-2.1	282.3	-277.1	0.2
C	116.3	-117.7	1.2	284.2	-283.3	-0.1
D	117.9	-118.3	4.6	284.6	-329.5	-0.4
E	118.3	-118.6	12.5	284.8	-284.7	-1.3
F	140.7	-140.8	15.4	338.2	-338.2	-1.6
G	145.1	-136.6	27.5	395.1	-403.7	-2.2

* Highlighted values in the columns δQ mean that the difference in gas consumption meets the Q_{diff}, highlighted values in the columns δP (P=parameter) signify that the gas consumption responds linearly to this change of parameter.
** C.f stands for ventilation correction factor. The assumed ventilation rate (a function of dwelling type and door and window frame length) is multiplied with the ventilation factor. The factor is assumed to be 1 in energy label calculations.
The orange values are insignificant on a 95% confidence interval scale.

TABLE 5 (II) Sensitivity of gas consumption for behaviour parameters

Small changes in ventilation (10% less for label C or 20% more for label B) can already explain a large part of the discrepancies. This is an indication that the ventilation rate is very important in the theoretical calculation, and is a potential culprit if the assumptions are not matched by the reality. The validation of all the nominal rates and factors used could confirm or refute that.

§ 3.4.3.2 Floor area and insulation quality

The left-hand section of Table 6 shows how gas consumption changes when the heated floor area is increased or reduced by 20m². Such a reduction would only explain the discrepancy for label C. With a slightly larger inaccuracy in the estimation of floor area (approximately 30m²), the discrepancies can also be explained for labels B and D. For other labels, gas consumption no longer changes in a linear manner when the floor area changes; however, it would seem feasible that the actual heated floor area could be as little as half the assumed heated floor area, especially in thermally poor dwellings. On this basis, one can claim that the assumed heated floor area does indeed influence gas consumption to a significant extent. Furthermore, the heated floor area influences the average indoor temperature.

LABEL	U VALUE [W/M²K]	Q_{DIF}	FLOOR AREA			INSULATION VALUE			
			δQ*		δA (Q_{DIF})*	δQ*		δU (Q_{DIF})* [% ASSUMED U VALUE]	δU (Q_{DIF})* [W/M²K]
			δA=+20M²	δA=-20M²		δU =+20%	δU =-20%		
A	0.200	-232.0	91.6	-89.7	50.6	103.0	-103.0	45.0	0.09
B	0.360	-116.3	100.8	-99.0	23.1	104.9	-105.3	22.2	0.08
C	0.500	72.2	121.6	-119.1	-12.1	106.3	-106.3	-13.6	-0.07
D	0.640	272.0	191.8	-186.2	-29.2	106.6	-106.5	-51.1	-0.33
E	1.600	738.6	246.4	-240.6	-61.4	106.6	-106.6	-138.6	-0.89
F	2.000	1081.3	341.8	-333.9	-64.8	126.6	-126.6	-170.8**	-3.42**
G	2.400	1815.7	383.8	-375.2	-96.8	126.6	-126.6	-286.8**	-6.88**

* Highlighted values in the columns δQ mean that the difference in gas consumption meets the Q_{diff}, highlighted values in the columns δP (P=parameter) signify that the gas consumption responds linearly to this change of parameter.
The red values are insignificant on a 95% confidence interval scale.

TABLE 6 Sensitivity of gas consumption for floor area and insulation quality

Unlike all five parameters mentioned previously, the U-value was more complex to test since it was impossible to use the same increment in all label categories. We only changed the U-value of the dwellings' walls because we assumed that this was the most frequent cause of errors during the inspection process. Inspecting the windows, floor or roof insulation is usually much more straightforward. Since an increment of ΔU=0.1 W/m²K would be a very significant amount for label A and very few in label G, we used percentage increments of 20% of the initial U-value for that label category (see second column of Table 6). Such an increment only explains the difference for label C, however, when deriving the increment $\delta T(Q_{dif})$ [%] which would explain the difference (Q_{dif}), values from 45% (label A) to -287% (label G) were obtained. These values are then translated to absolute increments of U $\delta U(Q_{dif})$ $[W/m^2K]$) by multiplying them by the initial U-values (second column). The necessary increment of U-value in label categories F and G yielded a negative U-value (they are marked with a double asterisk). For all other labels, an increment in U value can explain the discrepancy seen in relation to actual consumption.

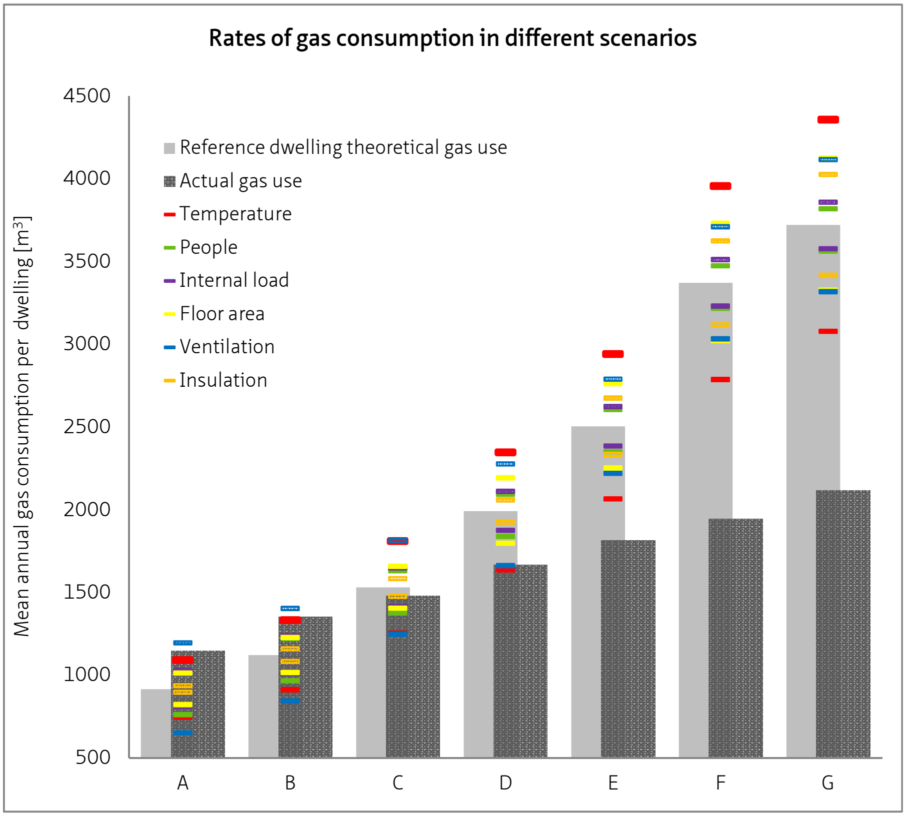


FIGURE 5 Rates of gas consumption when changing the values of the six assumptions considered in the sensitivity analysis

TOC

§ 3.4.3.3 Combined scenario

So far, the influence of each parameter on theoretical gas consumption has been investigated separately. However, it is likely that in reality several of the assumptions made are inaccurate. In this section a combined scenario involving all 6 parameters are modified as in Table 5 and Table 6 and applied to gas consumption simultaneously. Table 7 shows how the parameters changed under the two scenarios.

	δT	δN_{PEOPLE}	δQ_{INT}	δQ_{VEN}	δA	δU
Spending scenario	+2°C	+2	-2 W/m²	0.5 c.f	+20m²	+0.2%
Conserving scenario	-2°C	-2	+2 W/m²	-0.5 c.f	-20m²	-0.2%

TABLE 7 Parameters in the two combined scenarios

Figure 6 shows graphically the variation in gas consumption which arose due to the changed parameters. The actual gas consumption is somewhere within the two scenarios in all label categories. For labels A to E, the actual gas consumption falls fully within the variation range. For labels F and G, actual gas consumption can only be predicted if all the assumptions take extreme values. However, it is likely that if more extreme but still realistic assumptions (such as an average indoor temperature of 12.4°C) were made, actual energy use would have fallen within the range of the variations.

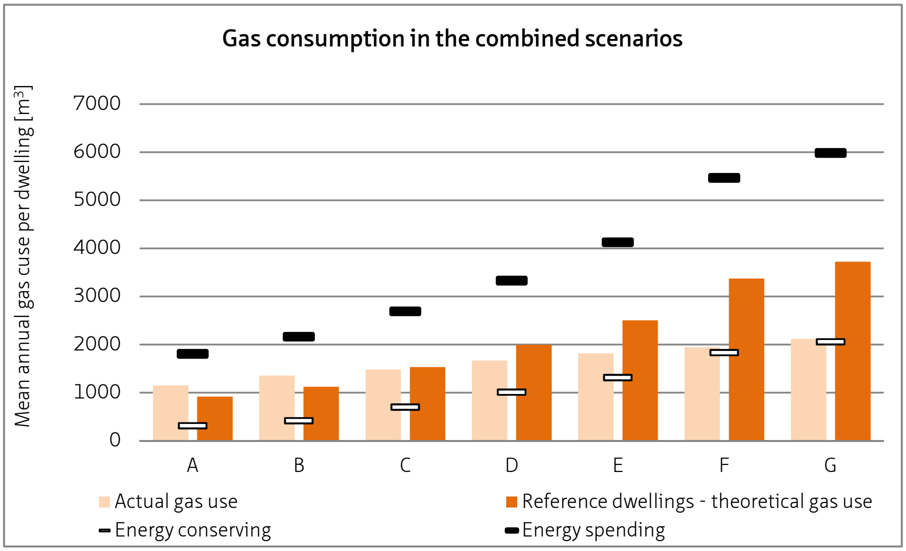


FIGURE 6 Rates of gas consumption in the combined scenarios

§ 3.5 Conclusions

Regression analysis using the make-up of the household and physical characteristics of the dwelling revealed that variables such as floor area, ownership type, salary and the value of the house, which predicted a high degree of change in actual gas consumption, were not significant (ownership, salary, value) or had a minor impact on theoretical consumption (floor area). This is most likely a consequence of occupant behaviour influencing actual gas use. Besides that, the installation system predictors showed that there was more overestimation in less energy-efficient systems, which can again be attributed to occupant preferences, or better yet, the relationship between the systems and the way the dwelling is heated. When a heat pump is present, there is usually under-floor heating so that the whole floor area is heated; meanwhile in older dwellings heated using a gas stove, the nature of the installation prevents the occupant from heating all the rooms. Moreover, there seem to be inspection faults in the current energy databases.

In the sensitivity analysis, average indoor temperature was found to have a large influence on the theoretical gas consumption together with the ventilation rate. The number of occupants together with internal heat load have a more limited impact on theoretical gas consumption. The accuracy of the U value estimation was also addressed, showing that slight deviations from the assumed U value can account for a large part of the discrepancy.

§ 3.6 Discussion and future work

The energy label calculation as the basis for the energy label is a simplified, static model, which does not take into account variations in occupant preferences. There are significant discrepancies between the actual and theoretical rates of gas consumption at the level of the Dutch housing stock and this has detrimental consequences for pay-back time calculations, estimates of potential savings and last but not least, people's confidence in the added value of certificates. This paper has sought to identify the source of these discrepancies.

The behaviour of the occupant undoubtedly has a major influence, even though it is sometimes difficult to quantify. Due to the fact that occupant preferences affect actual gas consumption, regression analysis explained much less of the variation in the actual consumption than in the theoretical. However, not all the variables that we wanted to examine were available for the regression analysis, so a sensitivity study was carried out on theoretical gas consumption to help to fill in the data gap. The variables examined in the analysis require further validation study in the future to find out whether assumptions regarding the factors and the reference dwellings used in the calculation (Equation 6) are actually representative of the Dutch dwelling stock. Moreover, the gas required for hot tap water, which is a variable influenced by the number of occupants, should be studied more thoroughly in the future since there are many assumptions involved which could reduce representativeness vis-à-vis the Dutch housing stock as a whole. The effect of the number of occupants on the number of rooms heated (and thus on the average indoor temperature) should also be studied. The fact that even slight changes in U-value resulted in a very different energy demand, again emphasises the importance of thorough inspections if label certificates are to represent the thermal quality of dwellings accurately.

In order to implement the best possible assumptions regarding the behaviour of occupants and the characteristics of installation systems, a thorough validation study would be needed, comparing all the assumptions used in the calculations with real values from a sample of dwellings representative of the Netherlands as a whole. It could be that different occupant profiles are required depending on the characteristics of the dwelling itself, whether in terms of thermal quality, installation system, dwelling type or some other quality it remains yet to be studied as well.

On the other hand, a perfect calculation method cannot reduce the inaccuracies that occur due to poor inspection of the dwelling; there should therefore be more emphasis on accuracy in the inspection phase.

At the same time, a question arises of whether the certificate in its current form is really the best possible option. The relevance of the theoretical rates of gas and electricity consumption on the label certificate is certainly open to question if the actual consumption rates deviate by more than 50% from the theoretical.

§ 3.7 References

Agentschap NL, 2011. Voorbeeldwoningen 2011 Bestaande bouw.

Branco, G., Lachal, B., Gallinelli, P., Weber, W., 2004. Predicted versus observed heat consumption of a low energy multifamily complex in Switzerland based on long-term experimental data, Energy and Buildings, Volume 36, Issue 6, June 2004, Pages 543-555.

Beerepoot, M., 2007. Energy policy instruments and technical change in the residential building sector, OTB Research Institute, December 2007.

Guerra Santin, O., L. Itard, H. Visscher, 2009. The effect of occupancy and building characteristics on energy use for space and water heating in Dutch residential stock, Energy and Buildings, Volume 41, Issue 11, November 2009, Pages 1223-123.

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., 2012. The effect of energy performance regulations on energy consumption, Journal of Energy Efficiency, 8 February 2012.

CBS Statline, 2012. CBS Statistics Netherlands database, http://statline.cbs.nl/statweb/CBS, accessed on 9th April 2012.

Cayre, E., Allibe, B., Laurent, M.H., Osso, D., 2011. There are people in this house! How the results of purely technical analysis of residential energy consumption are misleading for energy policies, Proceedings of the European Council for an Energy Efficient Economy (eceee) Summer School, 6–11 June 2011, Belambra Presqu'île de Giens, France.

Field, A., 2009. Discovering Statistics Using SPSS, Third Edition, January 2009.

Gaceo, S. C., Vázquez, F. I., Moreno, J. V., 2009. Comparison of standard and case-based user profiles in building's energy performance simulation, Proceedings of IBPSA Conference, July 27-30 2009, Glasgow, Scotland.

Greening, L., Greene, D., Difiglio, C., 2000. Energy efficiency and consumption—the rebound effect—a survey, Energy Policy, Volume 28, Issue 6/7, June 2000, Pages 389–401.

Haas R., Biermayr, P., 2000. The rebound effect for space heating - Empirical evidence from Austria, Energy Policy, Volume 28, Number 6, June 2000 , Pages 403-410.

ISSO, 2009. 82.3 Publication Energy Performance Certificate - Formula structure, (Publicatie 82.3 Handleiding EPA-W - Formulestructuur), Senternovem, October 2009.

Majcen, D., Itard, L., Visscher, H., 2012. Theoretical vs. actual energy consumption of labelled dwellings in the Netherlands: Discrepancies and policy implications, Accepted for the International Journal Energy Policy on November 2nd, 2012.

Milne, G., Boardman, B., 2000. Making cold homes warmer: the effect of energy efficiency improvements in low-income homes, A report to the Energy Action Grants Agency Charitable Trust, Energy Policy, Volume 28, Issues 6–7, June 2000, Pages 411-424.

Novem, 2002. Referentiewoningen bestaande bouw.

Pettersen, T.D., 1994. Variation of energy consumption in dwellings due to climate, building and inhabitants, Energy and Buildings, Volume 21, Issue 3, Pages 209 – 218.

Rogan, F., Ó Gallachóir, B.P., 2011. Ex-Post Evaluation of a Residential Energy Efficiency Policy Measure Using Empirical Data, 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., Daniëls, B., Maenkveld, M., 2011. Obligations in the existing housing stock: Who pays the bill?, Proceedings of the European Council for an Energy Efficient Economy (eceee) Summer School, 6–11 June 2011, Belambra Presqu'île de Giens, France

Vabi, 2011. EPA-W Stand Alone Software Package 3.01, Vabi Software BV.

§ 3.8 Appendix : definitions and abbreviations

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.
TYPE OF INSTALLATION	**EXPLANATION**
Conventional boiler (η=65%)	Central heating, gas boiler, efficiency above 65%.
Improved efficiency boiler	Central heating, gas boiler, efficiency above 83%.
High efficiency boiler (η=100%)	Central heating, condensing gas boiler, efficiency above 100%.
High efficiency boiler (η=104%)	Central heating, condensing gas boiler, efficiency above 104%.
High efficiency boiler (η= 107%)	Central heating, condensing gas boiler, efficiency above 107%.
Electrical heater	Small electrical heaters, portable electrical radiators etc.
Oil/gas stove	Oil- or gas-burning stove, usually located in the living room.
Micro CHP	Single-family home cogeneration of heat and power (electricity).
Heat pump	Transfers thermal energy from outside air or water to the inside of the house.
CHP	Cogeneration of heat and power (electricity).

TABLE 8 Appendix: Definitions and abbreviations

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).

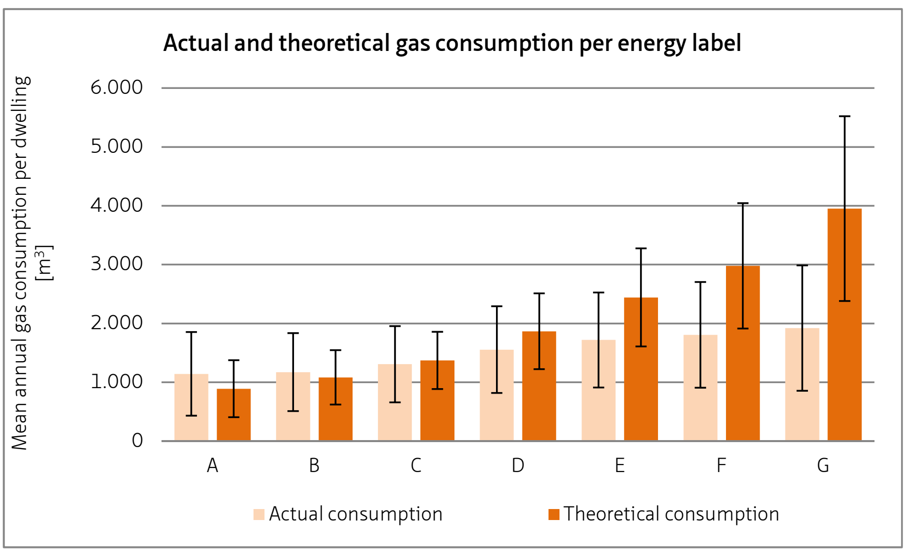


FIGURE 1 Actual and theoretical gas consumption in dwellings across label categories with ± 1 standard deviation (Majcen et al., 2013a).
Note that the two bars differ from each other in each category, this difference is in this paper referred to as the DBTA (difference between theoretical and actual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

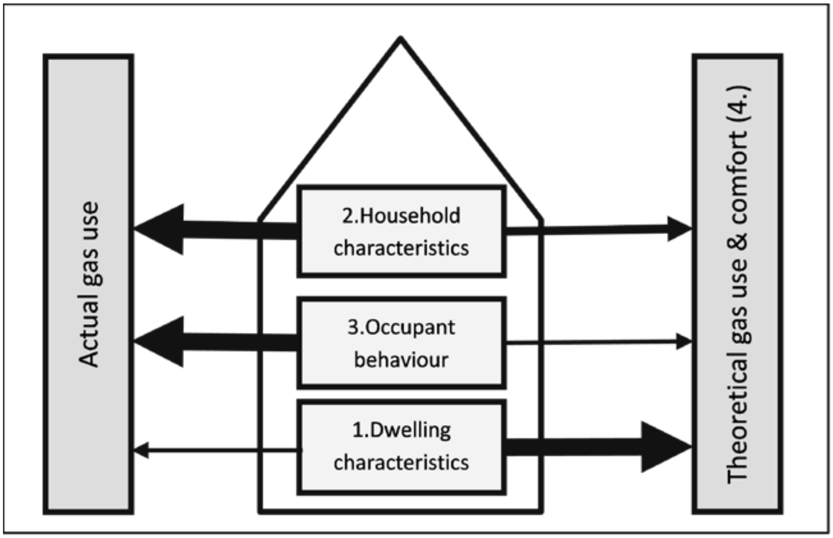


FIGURE 2 Effects of different parameter groups on actual and theoretical gas consumption.

§ 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

TOC

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.

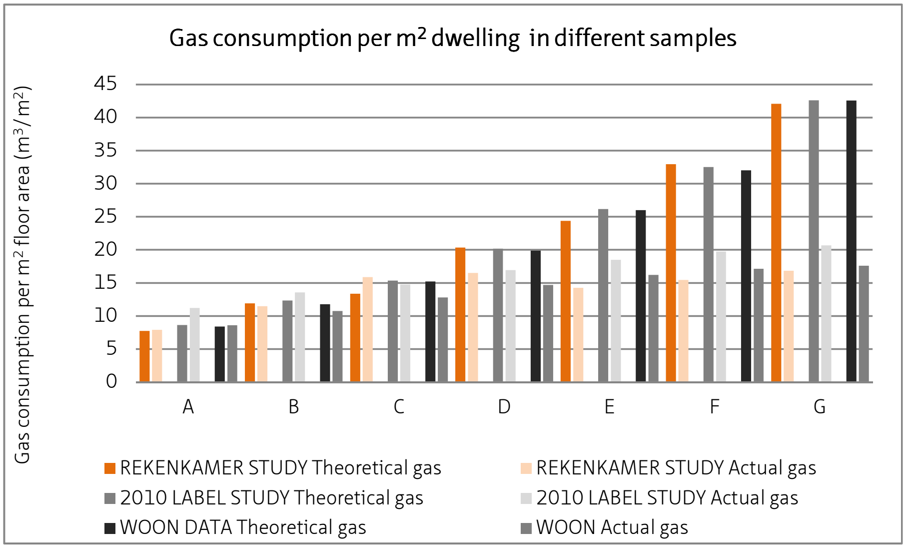


FIGURE 3 Average actual and theoretical gas consumption per m² dwelling including the 95% confidence interval.

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.

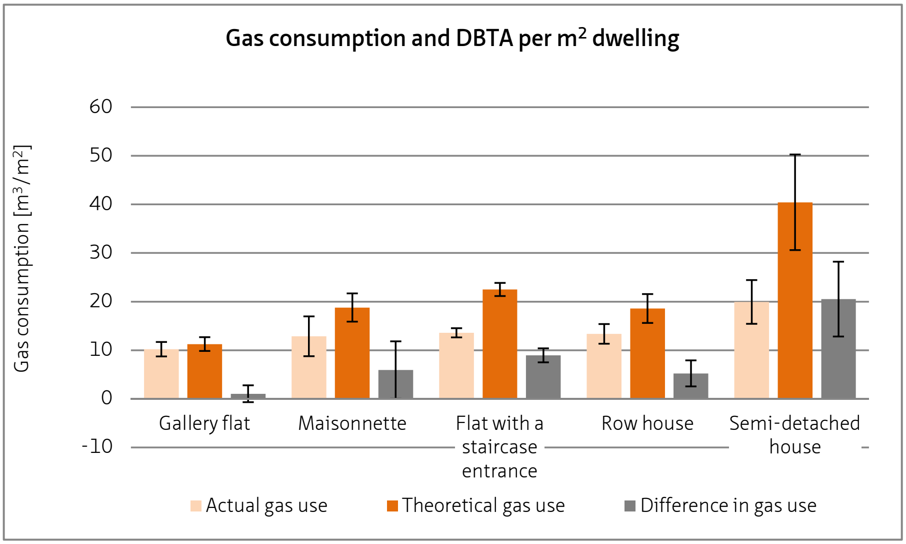


FIGURE 4 Actual consumption, theoretical consumption and DBTA per m² floor area of different dwelling types in the Rekenkamer sample.

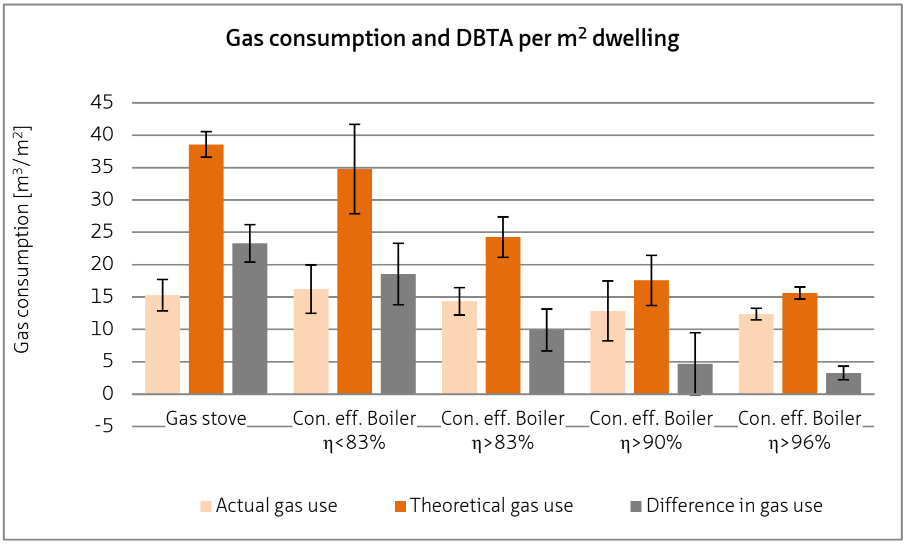


FIGURE 5 Actual consumption, theoretical consumption and DBTA per m² floor area of different installation types in the Rekenkamer sample

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

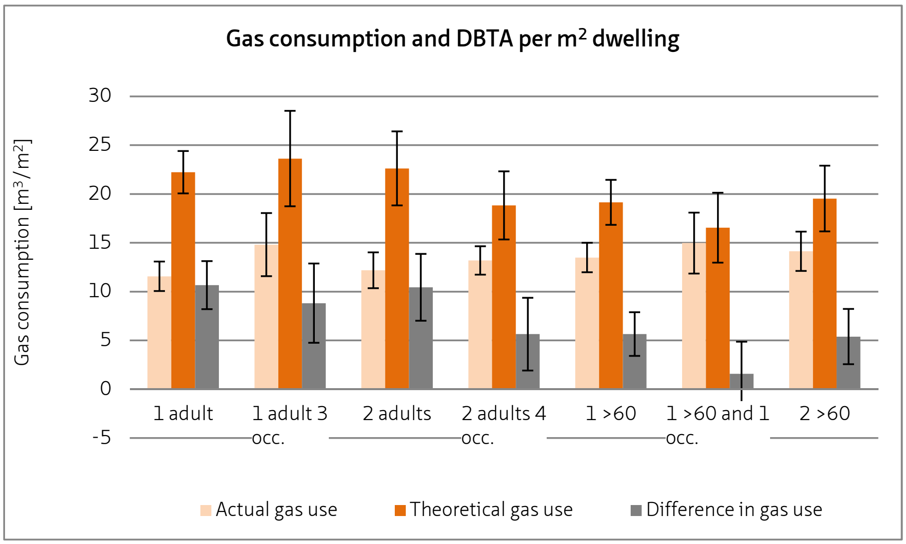


FIGURE 6 Actual consumption, theoretical consumption and DBTA per m^2 floor area of different household compositions
**>60 = occupant over 60 years of age*

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.

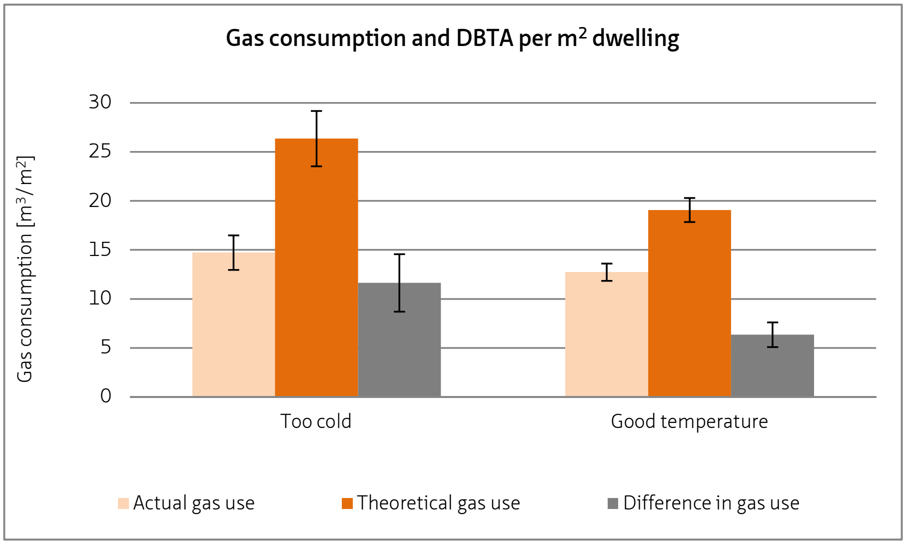


FIGURE 7 Actual consumption, theoretical consumption and DBTA per m^2 floor area in dwellings with difference temperature perceptions

		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 CHARACTER-ISTICS	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 M²	THEORETICAL GAS USE PER M²	DBTA	ACTUAL GAS USE PER M²	THEORETICAL GAS USE PER M²	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	THEORETICAL 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

>>>

		REKENKAMER					WOON				
			ACTUAL GAS USE PER M^2		THEORETICAL GAS USE PER M^2			ACTUAL GAS USE PER M^2		THEORETICAL GAS USE PER M^2 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.
DWELLING CHARACTERISTICS	(Constant)	8,901	3,108		,004
	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 m^2	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)

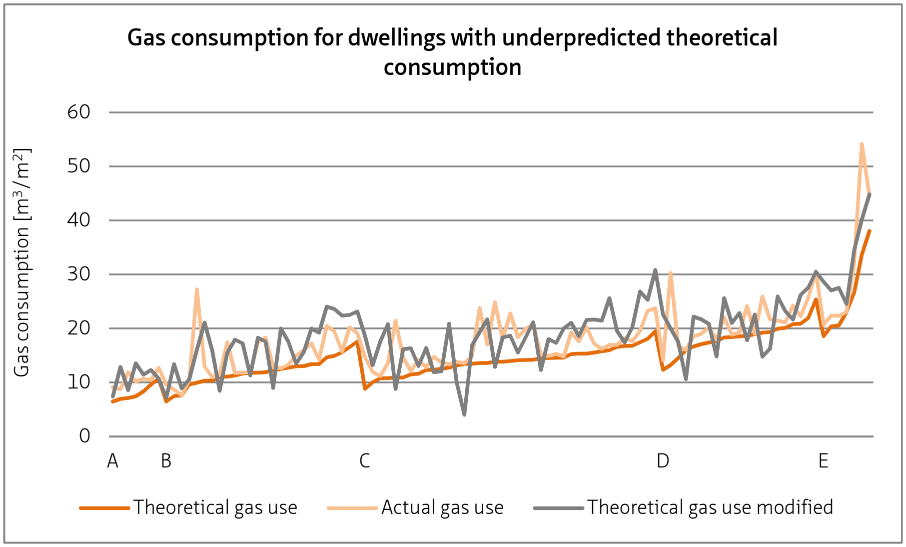


FIGURE 8 Theoretical, actual and modified theoretical gas consumption for dwellings with underpredicted theoretical consumption, a random sample of 100 dwellings from WOON sample

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.

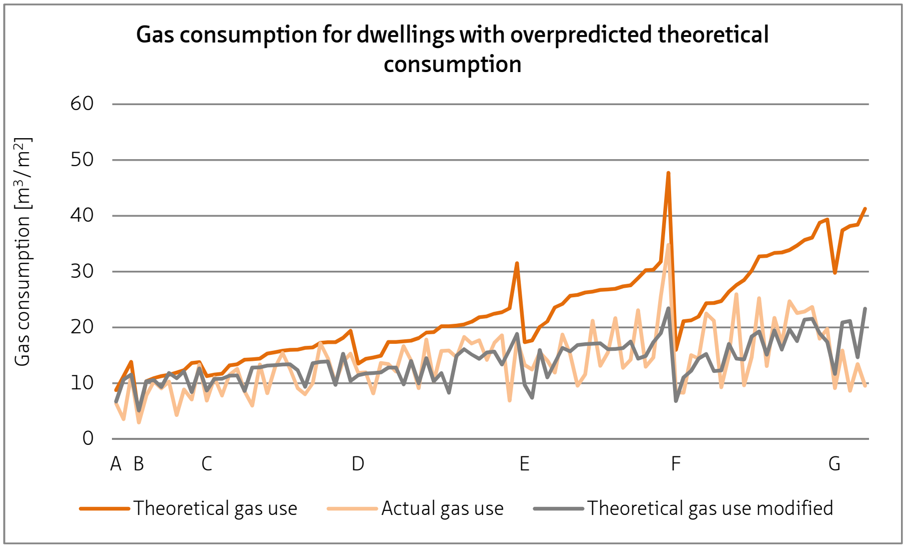


FIGURE 9 Theoretical, actual and modified theoretical gas consumption for dwellings with overpredicted theoretical consumption, a random sample of 100 dwellings from WOON sample

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

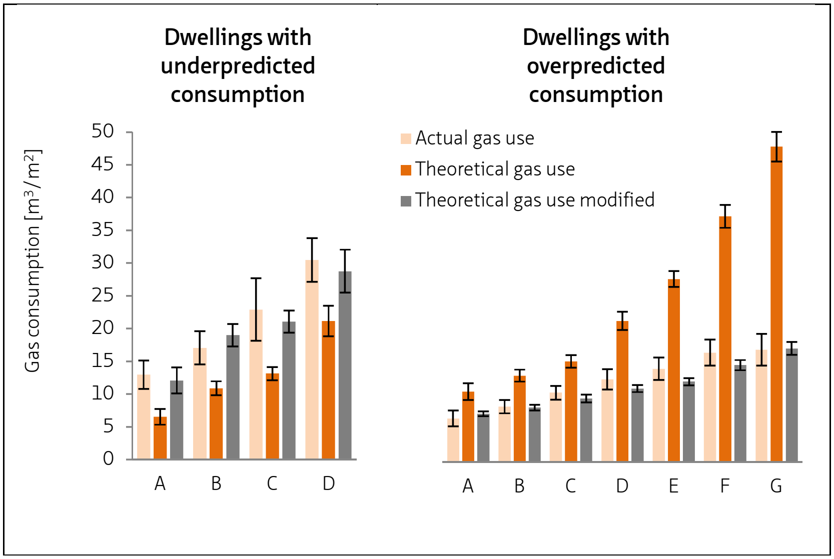


FIGURE 10 Mean and 95% confidence interval of the theoretical, actual and modified theoretical consumption

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 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

TOC

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.

5 Actual heating energy savings in thermally renovated Dutch dwellings

Explanatory notes

As opposed to the samples studied in the first three papers, all of which were based on cross-sectional data, Chapter 5 was the first to analyse longitudinal data from the social housing dwelling stock between 2010 and 2013, meaning that the research was narrowed down to dwellings that had undergone renovations in order to see whether the theoretical reduction of energy consumption materialised and to what extent. Since in this sample the dwelling's geometry mostly stays the same, the relation between performance gaps before and after renovations provides important insight into the accuracy of the normalisations used in the regulatory calculation model used in energy labelling. Moreover, a comparison of the actual reductions effected by different renovation measures was made in order to show which renovation practices lower energy consumption most effectively.

Submitted for publication to Energy Policy in October 2015.

Abstract

Since previous research has indicated large discrepancies between the theoretical and actual heating consumption in dwellings, it is important to know what savings renovations achieve in reality. The register of the Dutch social housing stock was analysed, containing dwelling thermal performance information of ca. 2 million dwellings between 2010 and 2013. Renovated dwellings were identified, providing insight into the performance gap before and after the renovation and the actual vs. the theoretical energy reduction of renovation measures. Improvements in efficiency of gas boilers (space heating and hot tap water) yield the highest energy reduction, followed by deep improvements of windows. Improving the ventilation yields a small reduction compared to other measures, however, it is still much larger than theoretically expected. High R and low U values of insulation are well predicted, as well as efficient heating systems whereas low R and high U values, local heating systems, changes from a non-condensing into a condensing boiler and upgrades from a natural ventilation system are not well predicted. The study therefore demonstrated that unrealistic theoretical efficiencies of heating systems and insulation values are causing a part of the performance gap.

§ 5.1 Background

Energy Performance of Buildings Directive is, since its first adoption in 2002, the main policy driver in reducing energy consumption in buildings in Europe. By proposing several actions such as a national performance calculation methodology (Article 3), performance certification of new and existing buildings (Article 11 and 12), cost optimality calculation (Article 5), the directive strives to raise awareness and increase investments leading to an accelerated transformation of the dwelling stock. In May 2010, a recast EPBD was drafted as a response to the more ambitious 2020 targets - 20% reduction of energy consumption and CO_2 emissions set by the commission in 2007 and 2009, respectively. To ensure that the directive is paving the way towards achievement of the set goals, monitoring of the dwelling stock efficiency is paramount on the national and European level to prove whether or not the improvements in efficiency are driving towards the desired targets. Monitoring would thus enable member states and the EU to reflect on the adopted policies and apply amendments where necessary. In 2011, registers of performance certificates were established nationally in 11 member states (Economidou et al., 2011) with the share of dwellings it contains ranging up to 24% in both The Netherlands and UK. For this study, we used a non-public register called SHAERE, which includes the annual performance of almost all dwellings of social housing associations between 2010 and 2013. In The Netherlands the social housing stock represents about a third of the total dwelling stock and is supposed to set nation-wide example for lowering the stock's energy consumption. Each year, the associations record the state of most of their dwellings, including their energy performance in the SHAERE register. SHAERE was set up by AEDES, the national organisation of housing associations, to be able to report the progress of energy renovations and improvement of the energy performance of their stock in relation to the 2020 goals laid down in a covenant with the government and the tenants organisation.

The dataset contained about one million dwellings in each of the four years, thereby offering a great opportunity to get insight into the changing energy performance of the dwelling stock. Previously published research conducted on the mentioned register, analysed the renovation pace of the dwellings between the years 2010 and 2013 (Filippidou et al. 2015a, Filippidou et al. 2015b). This paper, builds upon the findings of those papers by observing theoretical and actual heating energy consumption before and after the thermal renovation, which allows to compare performance gap (difference between theoretical and actual gas consumption) before and after renovation, thereby providing a much needed validation of the current label calculation method. Moreover, the theoretical reductions in dwellings where specific measures have been taken are compared with the actual metered reductions.

This helps establish the highest saving of the most commonly implemented thermal measures and enables a comparison of their effectiveness. The outcomes obtained by using different analysis methods are compared, making the analysis robuster and offering an insight into the accuracy of the methods.

Several definitions are used throughout the paper. *Dwelling properties* include 5 dwelling characteristics: type of space heating installation, hot tap water system, ventilation system, window thermal quality and the quality of insulation of roof, floor and wall aggregated as one variable called the insulation of the envelope. *A renovation measure* is defined as a change in at least one of these 5 parameters from one category into another (the continuous properties for insulation and window quality have been categorised). A *validated renovation measure* is a measure that yields the actual energy reduction comparable to the one predicted. A p*re-label* is a complete thermal recording of the dwelling, including all dwellings energy labels, theoretical heating demand and dwelling properties, which was reported to Aedes at least once in the period 2010 – 2013. *Label registration* is the act of submitting the pre-label data to the government thereby obtaining an official label certificate. *Energy index* is calculated according to the national standards on the basis of total primary energy usage, summing up the energy required for heating, hot water, pumps/ventilators and lighting, and subtracting any energy gains from PV cells and/or cogeneration and finally correcting this sum for the floor and envelope area. *The performance gap* is the difference between (average) theoretical and actual gas consumption of a dwelling or group of dwellings.

§ 5.2 State of the art

The SHAERE register was established in 2010 and includes complete thermal performance of the majority of the Dutch social housing dwelling stock, bringing the much anticipated data required for dwelling stock monitoring. First analyses of this dataset, encompassing over 1,2 million dwellings annually have been conducted by Filippidou et al. (2015a and 2015b). Filippidou et al. (2015a) describes the frequencies of 7 renovation measures as recorded in SHAERE in each available year. According to the author, 35,5% of the dwellings had a change in their energy label, 15% had an improvement of a single dwelling property and 12,7% had a change in more than one dwelling property. The author further breaks down the measures among the 757.614 dwellings which had a change in the energy label (the mentioned 35,5%) and established that 16,8% of the dwellings have improved their label class between years 2010 and 2013 resulting in an increased share of A and B labels (well performing) and decreased share of C-G labels. The remaining 18,7% had a deteriorated label class, which was thought to be a consequence of poorly executed dwelling inspection, which led to re-inspection and recalculation. Another study analyses the Dutch dwelling stock and the measures taken based on a survey of about 4000 representative dwellings (Tigchelaar and Leidelmeijer, 2013) who examine the frequency of various dwelling properties in the samples over the years. Based on the studied sample, however, the energy index of dwellings has improved from 2.09 to 1.89 (label E to label D) in the years 2006-2012, which is comparable to the pace of improvement as described by Filippidou et al (2015b), where the index dropped from 1.81 in 2010 to 1.69 in 2013. The sample analysed in the study by Tigchelaar and Leidelmeijer was relatively large, representative, and not limited to social housing associations. However, unlike the study of Filippidou, it did not follow renovations but samples of representative dwellings in each year. The third study is a national monitoring carried out in The Netherlands (Hezemans et al., 2012) on the basis of surveyed label improvements made in a sample stock of specific housing associations. An assumption was made that by implementing two saving measures (insulation of an envelope part or improvement in installation) coincides with 20% reduction in energy use. In the mentioned years together it was established that about 950.000 dwellings were made 20 – 30% more energy efficient. This monitoring was indirect (the assumption that two measures correspond to 20% energy reduction is a very rough one), used survey and not measured data and analysed relatively small samples which affects representativeness. However, it was the best available at that time and the assumption about two measures coinciding with a 20% reduction has been made due to serious gaps in existing knowledge about actual energy saving of renovation measures.

These three studies delivered information about the thermal measures taken in the housing stock but not on their effectiveness to achieve energy savings. Studying the actual energy savings of thermal renovation measures enables a precise evaluation of renovation strategies and subsequently policy effectiveness. Previous research showed that in The Netherlands, well performing dwellings consume more than expected and that poor dwellings consume up to half less than expected (Majcen et al., 2013a, Majcen et al, 2013b) causing the actual energy savings to be smaller in reality than expected. One of the causes of this performance gap is the fact that theoretical calculations are based on the same normalised conditions (for example average indoor temperature) regardless of the dwelling quality, even though in practice it turns out that the indoor environment differs greatly in poor performing dwellings from the one in efficient dwellings. The gap seems to be difficult to explain statistically, mostly due to the complex nature of the variation in actual gas consumption. However, differences in average indoor temperature and in the quality of estimation of insulation and ventilation flow rates in dwellings of different quality and socioeconomic factors were shown to be important factors in explaining this gap (Majcen et al., 2015). Menkveld studies the relation between the energy saving measure taken and the actual energy reduction using the national energy label database, which is dominated by social housing associations (about 70% of social housing and 30% of private dwellings, Majcen et al., 2013a). However, this study observes cross sectional dwelling data (only one record in time available for each dwelling), comparable also with previous analysis done by Majcen et al., 2013a and Majcen et al., 2013b and Tigchelaar and Leidelmeijer, 2013.

Numerous scientific papers have evaluated individual dwellings operational energy use, such as Adalberth, 1996, Winther et al., 1999, Dodoo at al., 2010, Thormark, 2001. However, as a rule these studies are based purely on theoretical operational heating energy, which as shown before can diverge from the actual consumption by as much as 50% less or 30% more. Karlsson et al. (2006) did base their operational energy consumption on real monitoring data of a reference dwelling, but still based energy calculations for different renovation scenarios on the exact same indoor temperature assumptions, which might not yield realistic results. Small scale projects are usually not that interesting for scientific audience since they lack representativeness and the results shown in non-scientific sources (construction companies, housing associations, even local governments) are likely to be skewed with an emphasis on successful examples.

Therefore, there seems to be a lack of studies analysing the efficiency of thermal renovation measures at the stock level. However, the gap in the literature is understandable since no large scale data about the dwelling stock's energy performance and actual energy use was available previously.

Despite this, an objective and representative evaluation of the undertaken saving measures is paramount in order to evaluate and improve the effect of current retrofit policies. This paper complements the results described above.

§ 5.3 Goal and scope

Using the detailed energy performance register coupled with annual actual energy consumption data gathered by Statistics Netherlands at address level, this paper offers an in-depth insight into longitudinal dwelling stock transformations. By studying a large sample of dwellings that underwent thermal renovation we aim to answer two research questions:

1. What is the actual heating energy saving in renovated dwellings for different thermal renovation measures?

2 What is the performance gap (difference between theoretical and actual gas consumption) in thermally renovated dwellings before and after the renovation? This way, we can not only provide data on actual energy savings but also offer a validation of the calculation method used to calculate the label. Additionally, the various samples studied (see methods section) will enable a comparison of different analytical approaches. Through the use of these methods we can comment on the usability of SHAERE dataset and provide guidelines for future setup of data registers in different European countries. In the results section we present the first results for the total changes in dwelling performance. Each of the thermal renovation properties is then divided in two sections – B and C. Until now (Majcen et al., 2013b and Majcen et al. 2015), the influence of dwelling properties on actual and theoretical gas use was determined cross sectionally, mostly with the use of both descriptive statistics and regression analysis. Since this is the first study using longitudinal data, section A provides cross sectional statistics of data used in longitudinal analysis (B and C). This enables a comparison of cross sectional and longitudinal analyses and validates the results.

In section A we present the actual and theoretical consumptions of dwellings in different label classes cross-sectionally, in the whole available dwelling stock in year 2010 (first available SHAERE record) and 2012 (last useful SHAERE record). This is done in order to place the results among the existing literature on the subject, since existing studies of the performance gap have invariably focussed on cross sectional data. Moreover, this first section gives an idea how the total thermal performance of the whole stock changes through time (how many label changes there are and how much energy consumption changes in each label class).

However, the core of the paper is the efficiency improvement of the dwellings and the actual energy savings following thermal renovations, therefore in parts B and C of the results we select only dwellings which have undergone changes and analyse the theoretical as well as actual reduction of energy consumption before and after renovation. In section B, all dwellings having a change in one specific dwelling property are studied, regardless of whether the other properties have changed or not. This may seem illogical, but in the past, such an approach was applied often in order to obtain significant results despite the small sample sizes. In section C, the dwellings having only a change in this specific property while all others are constant, are studied.

In the methodology section which follows, the process of data handling and subsample selection is outlined and the way of dealing with the data accuracy is explained. The results are presented separately for each examined dwelling property (space heating, hot tap water, ventilation, window quality and insulation). In the discussion section we first compare the three different methods, followed by a discussion of trends noted regarding the effectiveness of different thermal renovation measures, the performance gap and the validation of the calculation method.

As the Netherlands have an oceanic climate with cool summers and moderate winters, most of the energy consumption comes from heating demand. Natural gas is used as a source of heating in most Dutch dwellings and therefore also label certificates express heating energy consumption in m^3 gas. The actual consumption data is available at Statistics Netherlands in the same units, which is why we chose to study gas consumption as a measure of dwellings thermal performance. This means, however, that the dwellings that make use of electrical installation systems (e.g. heat pumps) were excluded from the analysis.

§ 5.4 Methodology

§ 5.4.1 Dataset properties

The SHAERE register is a raw, full export of the entire energy performance certificate calculation according to the Dutch standard (ISSO, 2009) on the level of dwellings for each year from 2010 on. The data differs significantly from the certificate data stored by the Ministry of the Interior and Kingdom Relations of The Netherlands (label certificates registered by the authorities as used in the studies by Majcen et al. 2013a and 2013b), since it includes all detailed properties required for the calculation of the energy label. However, the data in SHAERE does not consist of registered label certificates, but of so-called pre-labels. A pre-label is a label certificate of a dwelling that may have not been registered at the authorities yet but has nevertheless been recorded internally by a housing association. According to Aedes, pre-labels are updated whenever a renovation measure takes place and are considered accurate because housing associations report to use these pre-labels as an asset management tool (Visscher et al., 2013). Aedes provided the data from 243 Dutch housing associations (in 2011 there were a total 289 associations in The Netherlands) in years 2010, 2011, 2012 and 2013. It is important to note, that social housing represents 33% of the Dutch dwelling stock (Energiecijfersdatabase) and even though some properties differ with the private sector (Majcen et al., 2013a) such a larger sample does offer a great deal of representativeness. The database included dwellings geometry, envelope and installation system characteristics (including detailed information on the quality of insulation, ventilation and heating and hot tap water installation), as well as the theoretical heating energy consumption calculated according to the Dutch ISSO standard (ISSO 82.3, 2009).

In the present paper the dwelling data is available pre-and post-renovation (also called longitudinal data), which probably greatly decreases the variance between groups due to the changes in conditions we do not control for (different household and occupant properties in different groups etc.).

§ 5.4.2 Variable extraction

From the MSSQL SHAERE database, the tables about dwelling information, heating and hot tap water installation information, ventilation and envelope characteristics were merged for analysis, based on the dwelling ID. The type of each construction element (floor, roof, wall, window or door), area, U-value (heat transfer coefficient for windows) or R value (thermal resistance for all other constructions) is known.

To simplify the analysis we computed the average R value for the whole envelope and U value for windows using the formulas below using basic thermodynamic principles.

Insulation values for floor, roof, wall, windows and doors were available as continuous values. To simplify the detection of changes in insulation in between years, these variables were discretised into a finite number of categories. We first considered using the commonly encountered categories of insulation (as described in the Dutch standard ISSO 82.1), but since this yielded distributions highly dominated by the average value, we rather decided to rank the data into 10 categories and use the top and bottom value of each rank class as a basis for the category. We aimed for 10 categories within each label (each containing 10% of records). That way we capture more changes than by using the commonly used insulation groups. The categories are described in Table 1. The categories for R-value may seem to have strange ranges: the maximum R-value is 1,36 which is relatively low. One should keep in mind that an old Dutch dwellings may often have an R-value of 0,19 and insulation is generally brought only on a part of the house (e.g. the roof only or the wall between the window and the floor only) leading to average values that are still low.

The heating installation systems were all gas powered. The least efficient system (η=65%) is a local gas heater, where local means that the heater – a gas stove - is situated in one or two places in the apartment, most commonly the living room. The rest of the bedrooms are in this case not heated. An upgraded version of this system is a gas stove that is used to also heat the bedrooms, this is the gas heater with efficiency between 65% and 83%, regarded as η<83%, this kind of heater is non-condensing. A conventional non-condensing boiler has an efficiency between 83 and 90%, in named in this paper as η>83%. And several high(er) efficiency condensing boilers with efficiencies of 90, 94 and 96%, are referred to as η>90%, η>94% and η>96%.

The heaters for hot tap water are similar, in most cases the heater for space and water is combined, and in cases where it is not combined, the households use a tankless gas boiler for water heating. The methodology predicts several water efficiencies of water heaters – conventional (η<83%), improved (83%<η<90%) and high efficiency condensing boiler (η>90%).

Regarding ventilation, most dwellings in The Netherlands only have natural ventilation. In the data we also encountered several types of mechanical ventilation, such as, central mechanical exhaust, central demand controlled mechanical ventilation (DCV) controlled by CO_2 sensors, mechanical balance ventilation with heat recovery, decentralised mechanical ventilation with heat recovery, and finally, demand controlled decentralised mechanical exhaust ventilation.

R ENVELOPE EXCLUDING SURFACE RESISTANCE [M²K/W]	CATEGORISED R VALUE	U-WINDOW [W/M²K]	CATEGORISED U-VALUE
-0.19	R10	/	
0.19-0.21	R9	/	
0.21-0.25	R8	>4	U8
0.25-0.28	R7	3.7-4.0	U7
0.28-0.34	R6	3.1-3.7	U6
0.34-0.45	R5	2.93-3.1	U5
0.45-0.68	R4	2.9-2.93	U4
0.68-1.01	R3	2.6-2.93	U3
1.01-1.36	R2	1.8-2.6	U2
1.36-	R1	≤1.8	U1

TABLE 1 Categories of insulation values used)

§ 5.4.3 Sample selection

In theory, all dwellings should be pre-labelled and reported to Aedes each year, therefore ideally, each dwelling would have one record for each year of observation starting with 2010 up to 2013, adding up to four records. However, due to several reasons such as changes in associations reporting on the stock (some may cancel or start their cooperation with Aedes), purchases and/or sales of dwellings and demolition and new construction many dwellings have less than 4 records. In principle, more and more dwellings are pre-labelled and reported each year, since more associations decide to participate and the reported dwellings stock continues to grow. If one dwelling had several records in one given year and in case all dwelling properties were equal, we deleted the copies to leave only one record per dwelling. In some instances, not all properties were identical in both records and in that case we deleted both cases as we could not determine which one is more recent (the only time reference in the database is the year of the pre-label, no day or time stamp is available). After deleting those, our dataset was reduced from the initial 5.205.979 to 4.612.020 cases over four years.

After examining frequencies it became clear that the dataset contained a number of dwellings with an unrealistically small or large floor area. Therefore cases where floor area is below 15m^2 and above 500m^2 were deleted, resulting in a further reduced sample of 4.606.749 cases.

Most Dutch dwellings are heated by gas, and in the SHAERE sample almost 90% of the dwelling records (over all four year together) had a gas-powered hot tap water system and 93% had a gas-powered heating system. The rest of the dwellings utilize either district heating (4%) or electricity (6%) for hot tap water and about 7% of the space heating installations are electrical systems. District heating systems had to be removed due to the inaccurate actual annual consumption data for such installations. Electrical heating systems, mostly heat pumps, have been omitted to keep the scope limited and results more accurate. Removing non-gas based and collective systems left us with a sample of 3.729.256 reported pre-labels and further deletion of non-independent dwellings (student rooms, rooms in elderly homes etc.) resulted in a dataset of 3.728.143 pre-labels. As the actual energy consumption data from Statistics Netherlands was not yet available for the year 2014, we narrowed the sample further to the period of 2010 – 2012, resulting in 2.726.600 pre-label reports. For the measures that were taken in 2013 we would namely not be able to find a corresponding actual consumption (see also further in this section).

The actual energy use data provided by Statistics Netherland is collected from the energy companies, which base it on the annual meter readings done by the occupants. The data is therefore sometimes missing and averaged on the basis of similar households and sometimes an extrapolation of monthly values (if the reading are less than a year apart). This can cause inaccuracies that have already been discussed in previous papers (Majcen et al. 2013a, Majcen et al. 2013b, Majcen et al. 2015). The actual gas consumptions were corrected with degree days of the theoretical gas use (Majcen et al. 2013b).

Three types of subsamples were used in order to demonstrate trends with as much accuracy as possible. The abovementioned SHAERE sample of 2.726.600 reported pre-labels corresponds to 1.234.724 individual dwellings. In this dataset, every dwelling contained one or several pre-labels. The number of pre-label certificates from different years is gathered in Table 2.

2010 only	93.797	8%
2011 only	104.959	9%
2012 only	126,599	10%
2010 and 2011 only	151,467	12%
2010 and 2012 only	64,140	5%
2011 and 2012 only	111,255	9%
2010, 2011 and 2012	582,507	47%
Total	**1,234,724**	**100%**

TABLE 2 Number of dwellings having a pre-label in a given year

A Performance gap in the total stock

To show what changes occurred in the social housing stock data globally (section A of the results see Goal and scope), we first analysed the entire sample by coupling it with the corresponding annual actual gas consumption on address level (pre-labels from 2010 were coupled with 2010 actual gas data, 2011 pre-labels with 2011 and so on…). Reports with missing actual gas data were removed using outlier thresholds of 15 and 6000 m^3 gas (Table 3) per year. Part A analyses the theoretical and actual gas consumption in all pre-labels at the end of 2010 (835.313 pre-labels remained after the 891.911 total records were coupled with actual energy use) and in all pre-labels at the end of 2012, which includes also the years prior (1.152.320 coupled records out of the 1.234.724 total data, see Table 2). This means that for 2012, only the latest reports were taken into account. If there are no labels in 2011 and 2012 for example, we assume that there was no modification to the 2010 situation. In this section we compare all available records in 2010 and 2012, meaning that the dwelling that we observe are not identical (nor is the size of the sample). However, this gives a good idea of the changes made in SHAERE dataset globally over the years.

However, a sample of 835.313 (2010), representing 35% of the total social housing stock, can be considered to be well representative. Former studies (Majcen et al. 2013a, Majcen et al. 2013b) were based on such samples. The sample from 2012 is even more representative (ca. 50% of the stock). Therefore, under these assumptions of representativeness a comparison between 2010 and 2012 should lead to valid results about the changes in the dwelling stock.

YEAR	2010	2012
Total pre-label reports	891,911	1,234,724
Valid actual consumption data	835,313	1,152,320

TABLE 3 Pre-label reports with available actual gas consumption data

B Dwellings with a change in at least one dwelling property

In this section, dwellings with at least two pre-labels (sum of row 4 till 7 in Table 2) were selected, in total they amount to 909.369 dwellings. Due to missing actual gas consumption data and the fact that some categories contained less than 30 dwellings (which leads to high 95% confidence intervals and low statistical significance), the sample was reduced to 644.586 dwellings. Sample B is for each property, a subsample of these 644.586 dwellings. For instance, when studying changes in space heating and hot tap water, all dwellings with an improvement in space heating between the first and the last pre-label were selected, leading to a sample of 79.241 dwellings (Table 4). For dwellings with more than two pre-labels, the first and the last one were selected. Since dwelling observations were annual, last actual gas consumption before the first pre-label report year was used as baseline and the first available consumption data after the last pre-label report year. For example, for dwellings having the first pre-label report in 2010, gas data from 2009 was used and for dwellings having their last pre-label report in 2012, gas data for 2013 was used. Another condition was that both actual and theoretical consumptions have to be valid before and after the renovation (between 15 and 6000 m^3).

As Table 4 shows, the database reveals that some of dwellings in the sample have improved, most stayed the same and a fraction even deteriorated. Since all stock should be reported each year, it is logical that a large fraction remained unchanged as most dwellings do not undergo any change. Deteriorations are more surprising at first sight, but appear to occur due to a re-inspection of dwelling leading to a re-calculation of the label. This occurred due to changes in the inspection procedure or faults in the first inspection. All three installation variables observed have rather few deteriorations – between 1 and 2% whereas insulation values have slightly more (Table 4). Since we suspect these are administrative corrections, we do not show these changes in the graphics and consider only the improvements.

	LABEL CHANGES	SPACE HEATING AND HOT TAP WATER	VENTILATION	U-VALUE WINDOWS	R-VALUE ENVE-LOPE
Deteriorations	5%	2%	1%	6%	10%
No change	78%	87%	95%	77%	74%
Improvements	17%	12%	4%	18%	15%
Total sample size A	835,313 cases for 2010 and 1,152,320 for 2012				
Total sample size B	109,278	79,241	25,783	116,025	96,688
Total sample size C	/	30,749	4,866	15,744	21,035

TABLE 4 Share of improvements and deteriorations of various dwelling properties and sizes of analysed subsamples

C Dwellings with a change in only one dwelling property

The drawback of the sample selection in the previous paragraph is, that a change in for example heating installation system doesn't mean all other dwelling parameters remain constant. In fact, in most cases, more aspects of the dwelling have changed. In section C renovated dwellings were selected like in section B, but in addition all dwellings having more than one property changed were eliminated, meaning that dwellings have one and only one property changed. Categories with a number of records below 30 were discarded and Table 4 shows the amount of dwellings observed. While the samples in this section are much smaller than in section B, they offer valuable results about the effect of one single measure, which have to our knowledge not been previously described in scientific literature.

§ 5.4.4 Uncertainties

There was one difference between the end uses of theoretical and actual gas consumption, which is gas used for cooking. Actual gas consumption takes it into account and theoretical does not. However, cooking constitutes less than 2% of total gas consumption and it should therefore not affect the results too much.

In the section before, we showed that deteriorations of properties were observed in a small part of the sample (1 to 10%) due to re-inspection and re-calculations. We cannot exclude a comparable amount of improvements being caused by re-inspection and re-calculations rather than by real improvements. This will be taken into account in the analysis of the results. Moreover, also degree days calculation applied to actual gas consumptions (see section 5.4.3) and socioeconomic factor could influence the results (varying household size or composition, economic crisis, changing energy source for cooking etc.). To test these impacts, a control group consisting of unchanged dwellings was studied. Dwellings with 4 pre-label reports (497.088 dwellings) were selected out of the 2010-2013 SHAERE database containing 3.728.143 cases, after removing dwellings with missing actual gas data. From these 497.088 dwellings only the ones which had identical theoretical gas consumption four times were selected. These dwellings had no changed in any of the properties considered in this paper. This subsample contained 15.602 dwellings where no renovation measures took place. Table 5 shows a slight decrease of actual gas consumption of about 1,6% annually. In the identified sample of 15.602 dwellings their standardised actual gas use has decreased with 3,6% in years 2010 – 2013, which means that energy savings below 38 m^3 should not be considered as real improvement but as background noise. The numbers of degree days in the studied years were 3321, 2622, 2879 and 3078 from 2010 up to 2013.

YEAR	2010	2011	2012	2013
Average actual gas use [m³/year]	1054*	1034*	1017*	1016*
Average theoretical gas use [m³/year]	1113	1113	1113	1113
Gas reduction relative to 2010 [m³/%]		20 [1,9]	37 [3,5]	38 [3,6]

*The differences in actual consumption between the four years are significant on a 95% confidence interval.

TABLE 5 Reduction in actual gas consumption between 2010 and 2013 in non-renovated dwellings (N=15,602)

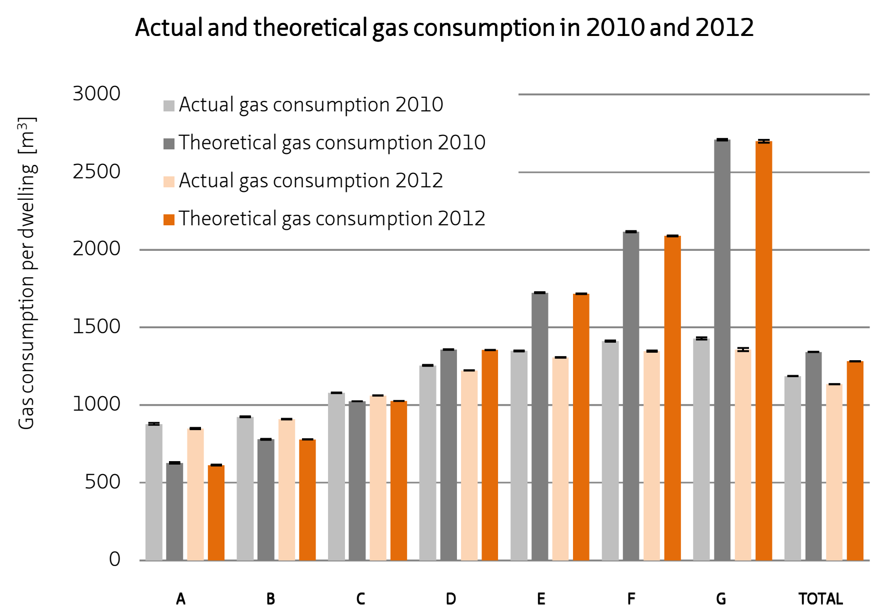
§ 5.5 Results

For an easier overview, the results are shown in sections 5.5.1 for label calculation and 5.5.2 to 5.5.7 per renovation measure. Section 5.5.1 consist of part A and B and later sections consist of B and C (like described in methodology section). Finally, section 5.5.8 compares the actual reduction of different measures investigated with method C and comments on their performance gap.

The results are presented in m^3 gas consumption per dwelling and not per m^2 floor area, since previous research demonstrated that although there are some slight differences in average floor sizes of dwellings in different label classes (size of A labelled dwellings is on average 105 m^2 and in other label classes the size is between 90 and 96 m^2), the performance gap does not change significantly whether observed per m^2 dwelling or not (Majcen et al., 2013). Furthermore, samples B and C represent renovations, therefore the floor area remains constant.

§ 5.5.1 Total thermal performance of the dwellings – comparison of label categories

This section shows the actual and theoretical reduction of dwellings which had their energy label improved, meaning that their total energy performance is observed.



		A	B	C	D	E	F	G	TOTAL
N	2010	16295	100718	265625	224814	117435	77767	32659	835313
	2012	36553	173351	383472	286707	145280	89387	37570	1152320
F	2010	2%	12%	32%	27%	14%	9%	4%	100%
	2012	3%	15%	33%	25%	13%	8%	3%	100%

FIGURE 1 Actual and theoretical gas consumption in 2010 and 2012 with 95% confidence intervals

Figure 1 reveals that while actual gas consumption drops from year 2010 to 2012 within all label categories (a drop between 14 and 71 m³), theoretical remains more constant (the drop in range of +3 and -26 m³) whereby the differences between theoretical consumption in categories B, D and G are also not statistically significant. Very similar results can be found in previous studies on this subject, using different samples (Majcen et al. 2013a, Majcen et al. 2013b, Majcen et al. 2015), with the results being comparable in terms of annual trends as well as the performance gap across categories. Overpredictions occur in labels D to G and underpredictions in the rest of the categories. While a difference of 2086 m³ gas can be noted when comparing theoretical consumptions of category A and G, the difference in actual consumption is a mere 508 m³, almost 4 times less. The difference between the two consumption of category F and G is the most drastic, 609 m³ for theoretical and only 10 m³ for actual consumption. Despite the changes noticeable in the performance gap in different label categories, there is only a slight decrease in the performance gap of the total sample – from 156 m³ in 2010 to a 148 m³ in 2012. This is because of the increasing number of better performing dwellings and the decreasing number of poor performing dwellings.

The frequencies of label classes change throughout the years (table below Figure 1). Frequencies of well performing dwellings (labels A – C) have increased (in total from 46% to 51%) and there are fewer D – G labels (from 54% to 49%).

Figure 1 shows that in total both average actual and theoretical gas use are lower in 2012 than in 2010. The absolute difference in the actual gas use is 52 m^3 and in the theoretical gas use it is even higher, 60 m^3. The theoretical reduction is a reflection of an improving dwelling stock (as said before, the frequency of good labels is increasing) and the actual gas use probably partly reflects that as well, however, the 38 m^3 background reduction should be disregarded leading to an actual gas reduction is only 19 m^3, three times less than expected. These 19 m^3 are either due to a different sample (many new dwellings were added) or performance improvements within one label category.

B Dwellings with a change in label class

For the results in this section, the sample of 644.586 records described in 4.3 was used. To show how this sample relates to the one in 5.5.1.A, gas consumptions in 2012 of both samples, A as well as B are plotted on Figure 2. Even though the confidence intervals are not plotted for better readability, the differences in consumptions between the two samples are negligible (not significant). This means that in terms of actual and theoretical consumptions on average, sample B is representative for sample A (which is larger).

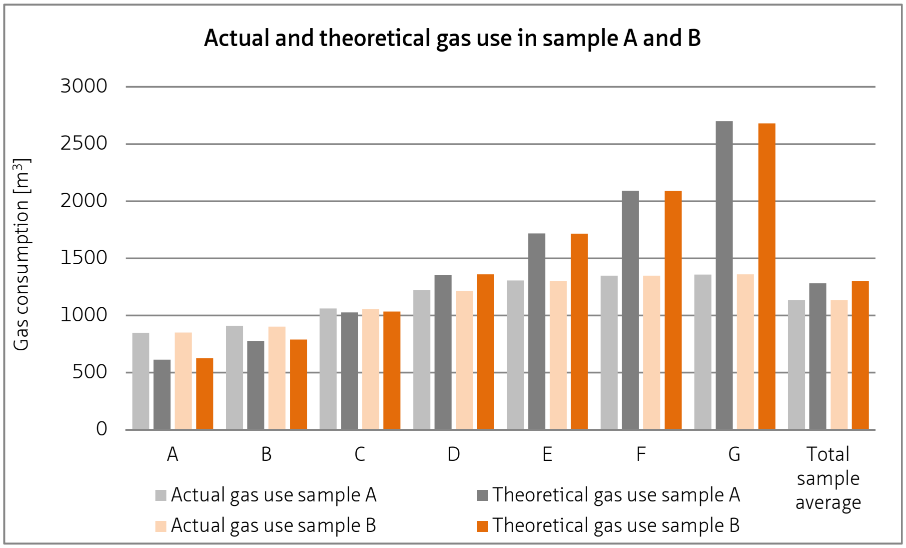


FIGURE 2 Actual and theoretical gas use in sample A and B

Table 6 shows the actual and theoretical gas reduction between years 2010 and 2012 in the selected sample (dwellings having a change in label class). As shown in Table 4, majority of the dwellings that had two labels reported in this period did not change label class, 17% has been improved and 5% have deteriorated. In this section, we focus on the sample of the 17% that have improved showing the actual and theoretical reduction in each of the label changes together with the ratio between them.

When looking at the changes in actual and theoretical gas consumptions at the time of first and second label (sample B) two possibly related phenomena can be noted (Table 6). Firstly, the actual improvement corresponds with the theoretical the best in dwellings that were well performing already before the measure (for example the improvement B to A has a theoretical reduction of 125 m^3 and actual of 129 m^3 whereas a dwelling that went from F to E has a theoretical reduction of 374 m^3 and actual of 136 m^3). Secondly, smaller improvements seem to be better predicted than deep renovations (for example B to A or C to B achieve 103 and 95% of the expected theoretical reduction), while F to B achieves only 27%. Renovations of very poor performing dwellings such as G or F result achieve a smaller % of theexpected reduction, 36% when improving from F to E and only 26% when improving from G to F. Renovating such poor dwellings to an even higher standard is even less well predicted (G to A dwellings achieve 21% (2075 m^3 theoretical and 446 m^3 actual) and G to F realize 26% of expected savings (508 m^3 theoretical and 133 m^3 actual). However, the absolute values prove that deep renovations nevertheless yield a higher saving in m^3 than minor renovations. These findings are in line with previously mentioned cross sectional studies.

Table 6 also shows the comparison of cross sectional data (section A) vs. the longitudinal data (section B) for renovated dwellings where the label class has changed. Whereas relatively comparable results were obtained when observing larger changes in thermal performance (more than 2 label classes), in changes for only one or two classes (A to B or G to F and G to E) cross sectional methods (section A) seem to strongly underestimate the actual gas saving (G to F 133 m^3 vs. 10 m^3). Longitudinal data (section B) results in actual reductions larger than those of cross sectional data. Dwelling characteristics, which correlate with a particular label class in sample A (for example, more apartments efficient label classes) whereas poor label classes are more dominated by detached and row houses, present a limitation of cross sectional data use, as they cause a comparison of two entities that are essentially very different. Moreover, there is a possibility that behaviour and lifestyles of the occupants in cross sectional data are different in different label classes. The longitudinal data on the other hand, assures that the same dwellings are compared before and after the renovation, which reduces these uncertainties. The occupants could still have moved during this time, but this probably only happened in a fraction of the dwellings (whereas in cross sectional data, the occupants are always different). Moreover, the performance gap expressed as the ration between the actual and the theoretical gas consumption in generally much smaller in sample B than in sample A.

	DWELLINGS WITH AN IMPROVEMENT OF LABEL CLASS, SAMPLE B				WHOLE DWELLING STOCK STATISTIC, SAMPLE A				
	ACTUAL [M³]	THEO-RETICAL [M³]	N	RATIO ACTU-AL/	ACTUAL [M³]	THEORETICAL [M³]	N BEFORE	N AFTER	RATIO ACTUAL/
G to F	133	508	3,576	0.26	10	609	37,570	89,387	0.02
G to E	153	846	2,090	0.18	51	983	37,570	145,280	0.05
G to D	215	1,415	934	0.15	135	1,345	37,570	286,707	0.10
G to C	301	1,742	730	0.17	297	1,672	37,570	383,472	0.18
G to B	354	1,871	348	0.19	449	1,921	37,570	173,351	0.23
G to A	446	2,075	78	0.21	509	2,086	37,570	36,553	0.24
F to E	136	374	2,090	0.36	41	373	89,387	145,280	0.11
F to D	135	674	934	0.20	125	735	89,387	286,707	0.17
F to C	227	1,091	730	0.21	287	1,063	89,387	383,472	0.27
F to B	371	1,379	348	0.27	439	1,312	89,387	173,351	0.33
F to A	510	1,688	78	0.30	499	1,477	89,387	36,553	0.34
E to D	127	323	934	0.39	84	362	145,280	286,707	0.23
E to C	187	626	730	0.30	246	690	145,280	383,472	0.36
E to B	342	920	348	0.37	398	938	145,280	173,351	0.42
E to A	392	1,107	78	0.35	458	1,104	145,280	36,553	0.42
D to C	150	242	730	0.62	161	328	286,707	383,472	0.49
D to B	217	473	348	0.46	313	577	286,707	173,351	0.54
D to A	318	718	78	0.44	374	742	286,707	36,553	0.50
C to B	157	165	348	0.95	152	249	383,472	173,351	0.61
C to A	137	310	78	0.44	213	414	383,472	36,553	0.51
B to A	129	125	2,499	1.03	61	165	173,351	36,553	0.37

*The orange highlights signify a more than twice as high ratio of method B compared to method A.

TABLE 6 Actual and theoretical heating energy savings corresponding to different label steps made

§ 5.5.2 Space heating and hot tap water

This section shows the actual and theoretical energy reduction in dwellings which had an improvement in the space heating and hot tap water installation. The two systems are viewed together despite the fact that in SHAERE database, these were two separate variables. However, during the preliminary analyses many illogical combinations of space heating and hot tap water were observed, such as a combined high efficiency hot tap boiler together with local gas heater. Such an installation is impossible in practice, since 'combined' boiler means that it is used also for heating. Because of this hot tap water and heating were analysed together, only looking at the dwellings with a logical combination of the two systems. Furthermore, for better readability we only show the results for dwellings which had an improvement in both heating and hot tap water systems and not just in one. To ensure statistical significance, groups with less than 30 cases are omitted from the figures.

A Dwellings with a change in heating and hot tap water system

The most common change among this measure in the observed sample is the replacement of the space heating boiler from improved $\eta>83\%$ efficiency boiler to a condensing boiler with $\eta<96\%$ efficiency and at the same time changing a combined improved (CI) efficiency tap water boiler with high efficiency (CH) one (this is in fact one system, last column in Figure 3). More than half of the studied dwellings within this measure have undergone such a renovation which makes this result very robust. The actual reduction is about two thirds of the theoretical.

Decrease in gas consumption is much smaller than expected in most dwellings with renovated heating and hot tap water systems (Figure 3). Roughly, the results can be divided into two groups, one group being the dwellings with a hot tap water boiler improved from an on-demand tankless boiler to a combined boiler and the other group where hot tap water combined boiler has been improved in efficiency (five last columns of Figure 3). In the first group, the difference between theoretical and actual reduction is in general larger than in the second group. If we look at the changes in heating installation, there seems to be few correlation between the extent of efficiency improvement and the actual gas reduction. Changing a local gas boiler has an actual gas consumption far below the theoretical. A pattern can be detected if one keeps in mind that boilers with efficiencies $\eta<83\%$ and $\eta>83\%$ are non-condensing and other boilers ($\eta<90\%$, $\eta<94\%$, $\eta<96\%$) are condensing. It seems that changes towards a higher efficiency within the category of non-condensing boilers are well predicted (second and tenth column). Similarly, also improvements in efficiency within the category of condensing boilers are reasonably well predicted (eighth and ninth column). In the changes of efficiency between non-condensing and condensing group the predictions are worse (in the last four columns). In some cases the reduction in actual consumption seems to be negative despite the large theoretical reduction (fourth column). It could also be that such group contains a complex of apartments which were recorded at the same time and contain a systematic error.

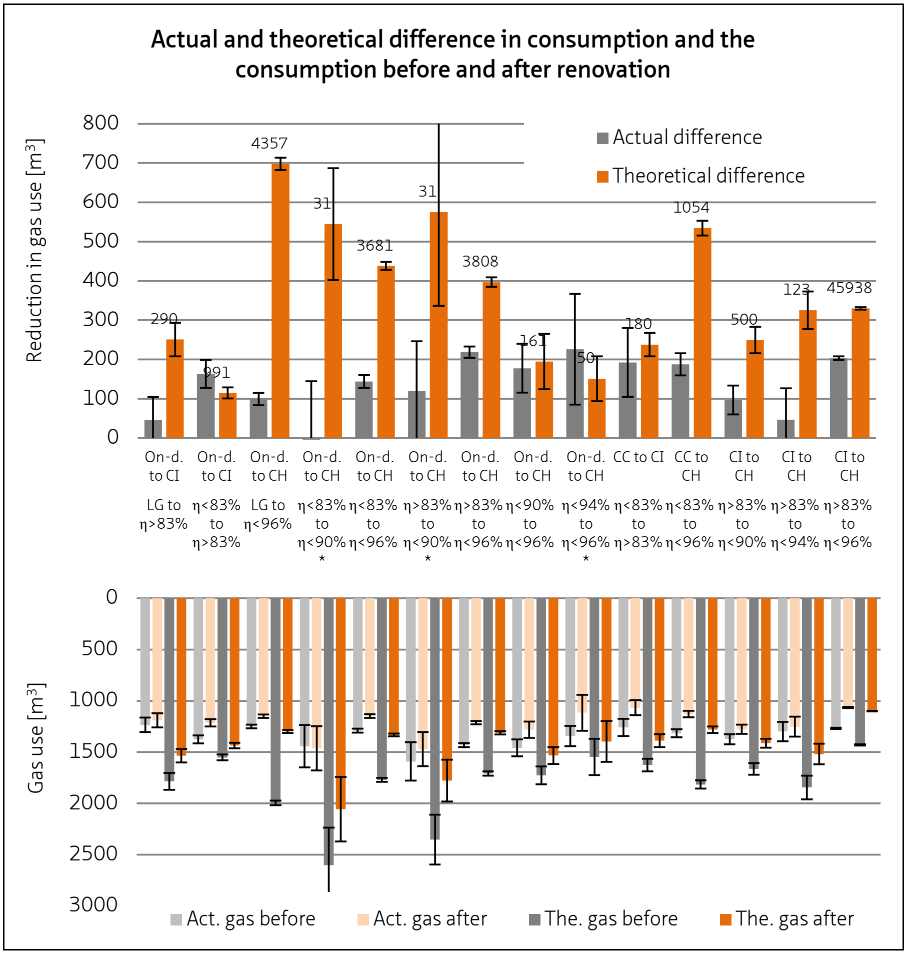


FIGURE 3 Actual and theoretical reduction and number of cases (parenthesis upper figure) and consumption before and after renovation (below) in dwellings with a renovated hot tap water and heating installation system – sample method B (N>30). On-d.= on-demand tankless boiler, CC/CI/CH = combined conventional/ improved/high efficiency boiler , LG = local gas heater * This group is not present in the results of section C.

Figure 3 does not show whether for example, a high actual performance gap is more a consequence of poorly predicted consumption before after the implementation of the measure. This means in practice, that one cannot tell whether a low performance gap is indicative of a low performance gap after renovation, which was observed previously by Reynaud (2014). Therefore consumptions before and after are plotted on the bottom of Figure 3. It seems that the dwellings which were poorly predicted before the renovation remain poorly predicted after renovation, however, on average all dwellings seem to be better predicted after the renovation, which is in accordance with the previously noted fact that better performing dwellings are better predicted. Also, the heating systems with η<96% efficiency seem to be well predicted (see light orange and light grey bars), especially where the samples are larger (third and last column). Moreover, it shows once again that in cases where local gas stove was changed with a more efficient system, the gap generally decreases - this is due to mentioned ill-assumption of heated floor area in case of local gas heaters that was shown in (Delghust et al. 2015 and Majcen et al. 2013b): in houses with local gas heaters, generally only one or two rooms are heated, whereas the calculation are based on heating of all rooms. In general, the more efficient the heater the better the prediction.

B Dwellings with a change in only heating and hot tap water change

In this sample of 30.749 cases only the heating and hot tap water installation had changed according to the information in SHAERE database. Among all the studied measures, heating and hot tap water have the most similar samples in section B and C (79.241 and 30.749 dwellings), which means one can expect the most comparable results: when the heating system is changed, there are usually no other measures taken. The difference between the theoretical and actual reduction seems slightly less drastic (see last column of Figure 3 and Figure 4) but despite from that, the results are indeed comparable.

It seems that dwellings are again better predicted after renovation than before. Visually, there does not seem to be a correlation between the size of the performance gap before and after the renovation.

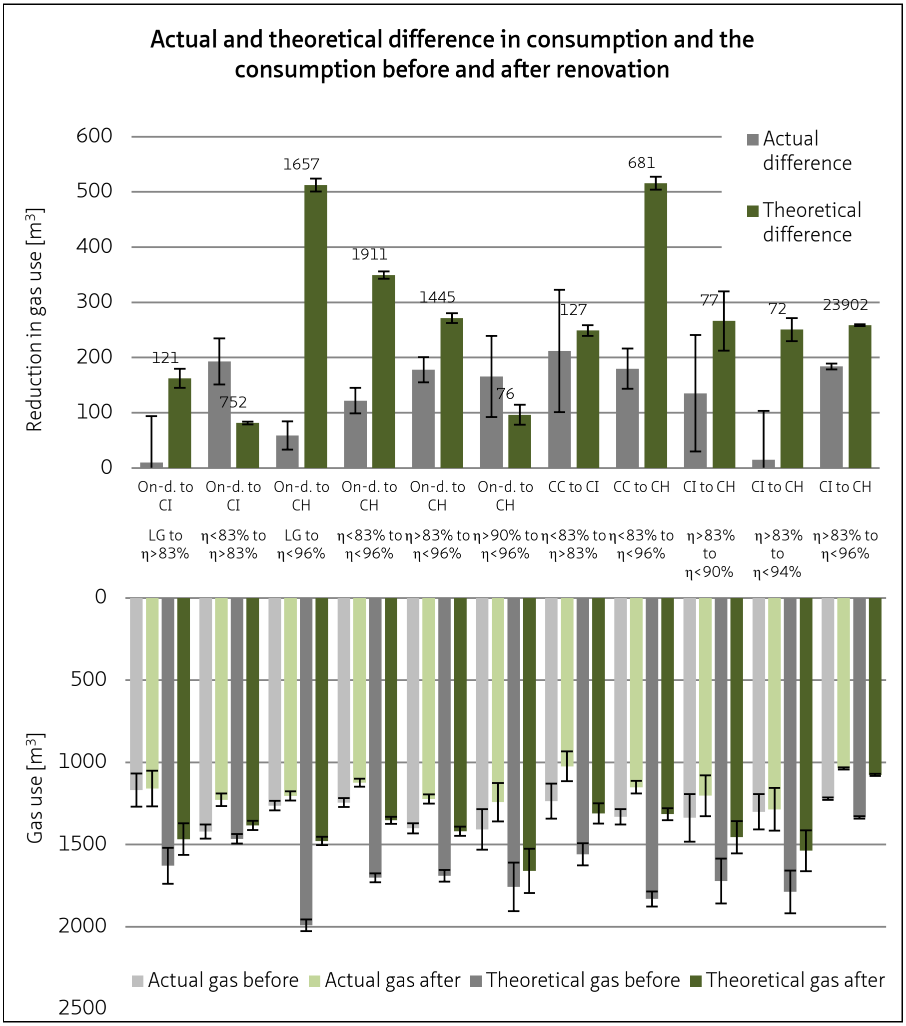


FIGURE 4 Actual and theoretical reduction and number of cases (parenthesis upper figure) and consumption before and after renovation (below) in dwellings with a renovated hot tap water and heating installation system – sample method B (N>30). On-d.= on-demand tankless boiler, CC/CI/CH = combined conventional/improved/high efficiency boiler , LG = local gas heater). Actual reduction of the first and before last column is below the background reduction.

§ 5.5.3 Ventilation

This section shows the actual and theoretical reduction of dwellings which had an improvement in the ventilation installation. We excluded the groups of dwellings which contained less than 30 cases to ensure statistical significance.

A Dwellings with a change in ventilation system

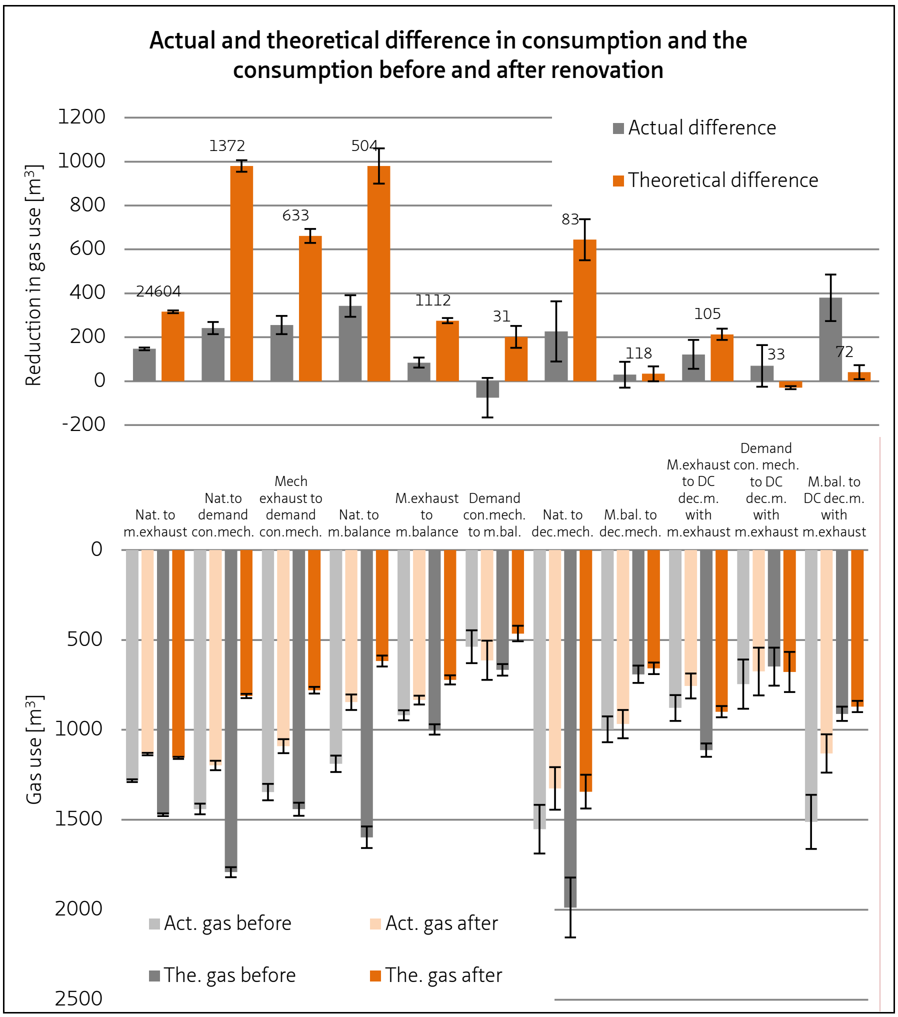


FIGURE 5 Actual and theoretical reduction and number of cases (above) and consumption before and after renovation (below), in dwellings with renovated ventilation system – sample method B (N>30)

The most common change in this category is replacement of natural ventilation with mechanical exhaust ventilation. In this and also most other categories, the decrease in gas consumption is much smaller than expected with the exception of converting a mechanical balanced ventilation system to a demand controlled decentralised mechanical ventilation with mechanical exhaust. Converting a naturally ventilated dwelling into one with mechanical exhaust (the most common renovation) ventilation yielded 147 m^3 of the expected 316 m^3 gas reduction. Other renovation from a natural ventilation system also yielded half to a third of the expected savings.

It is rather interesting, that many categories go from overprediction of gas use (this is typically the case for natural and mechanical exhaust ventilation), to underprediction after the renovation. This creates the large 3-4 fold ratio between theoretical and actual reduction (Figure 5 above). Like Figure 3 and Figure 4, also Figure 5 does not show a correlation between performance gap before and after renovation. Rather this gap seems to correlate well with the type of system (the energy performance of less efficient systems is overpredicted and efficient ones are underpredicted).

B Dwellings with a change in only ventilation system

As opposed to Figure 5, Figure 6 seems to suggest the savings when changing from natural to mechanical exhaust ventilation to be at least three times as high as expected. In Figure 5 we have seen the performance gap in dwellings that changed from natural to mechanical exhaust ventilation system to decrease substantially and the actual gas consumption was overpredicted both before and after renovation. Both these phenomena are not observed in Figure 6. The theoretical gas consumption barely reduces after the renovation, which is logical, since in the calculation method mechanical and natural ventilation both use exactly the same air flow rates. In practice it could be that the savings are achieved at the expense of the air flow rates. Mechanical balance ventilation makes use of heat recovery, which explains the theoretical reduction in the third column, however, the fact that the actual reduction is so much less could mean that heat recovery does not work at the rate assumed by the calculation method. Since in the second column the ventilation is also upgraded to a balance system, it is not clear why the two theoretical consumption are so different, that may relate to project specific data and the small amount of cases. Column three states with statistical significance that actual reduction when replacing mechanical exhaust with balance ventilation is less than a quarter of the expected. Also the last column gives an interesting result, since there is an actual increase in consumption of the systems which are expected to have a reduction. The implemented demand ventilation system does have lower theoretical air flow rate, which explains the theoretical reduction. A validation of air flow rates could solve these problems in the future. A possibility is also that this last category of on-demand decentralised ventilation with exhaust ventilation is not interpreted by the inspectors correctly due to its complexity which could lead to frequent input errors.

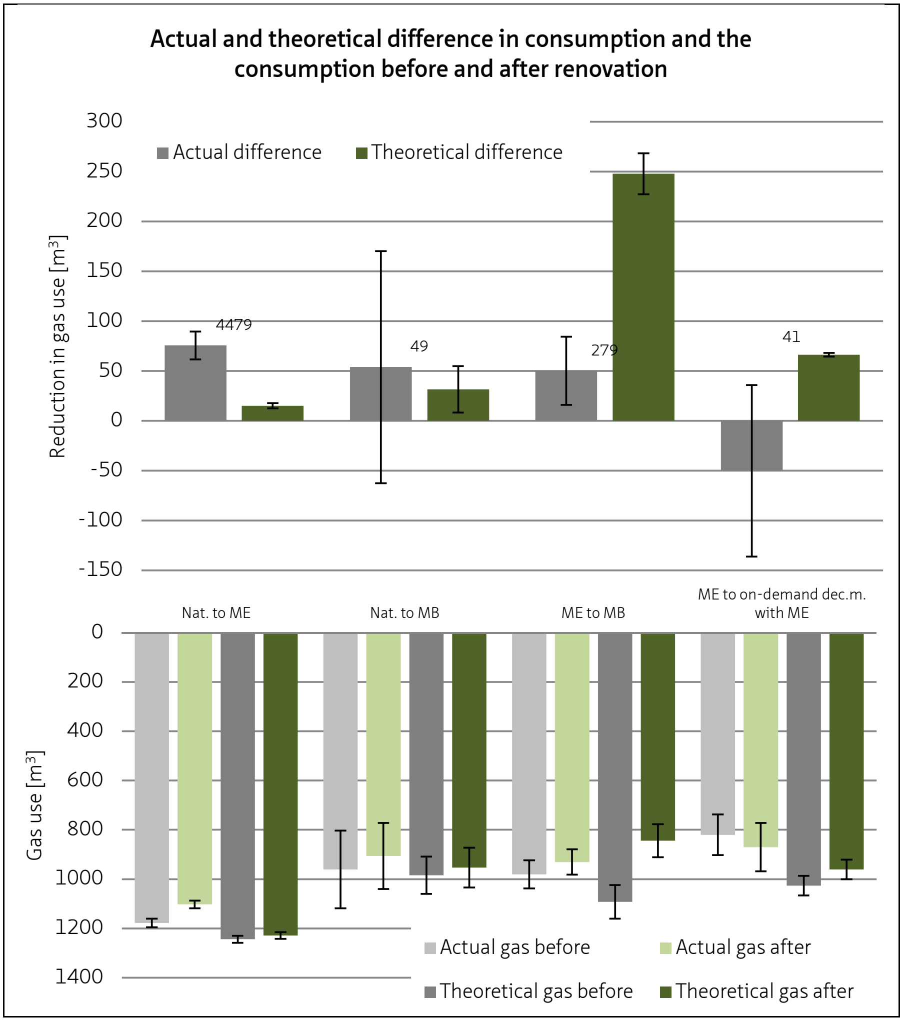


FIGURE 6 Actual and theoretical reduction and number of cases (above) and consumption before and after renovation (below), in dwellings with replaced ventilation system – sample method C (N>30). Nat. to ME = Natural to mechanical exhaust ventilation, Nat. to MB = Natural to mechanical balanced ventilation, ME to MB = Mechanical exhaust ventilation to mechanical balanced ventilation, ME to on-demand dec.m. with ME = Demand-controlled mechanical ventilation with mechanical exhaust

§ 5.5.4 Changes in window quality

This section shows the actual and theoretical gas reduction of dwellings which had an improvement in the window quality. In this section insulation quality as described in Table 1 are used. To keep the results in Figure 7 readable, changes of windows to window insulation category U5 and U4 are not shown. These do follow the same pattern and they have been included in the results of section 5.5.6.

A Dwellings with a change in window quality

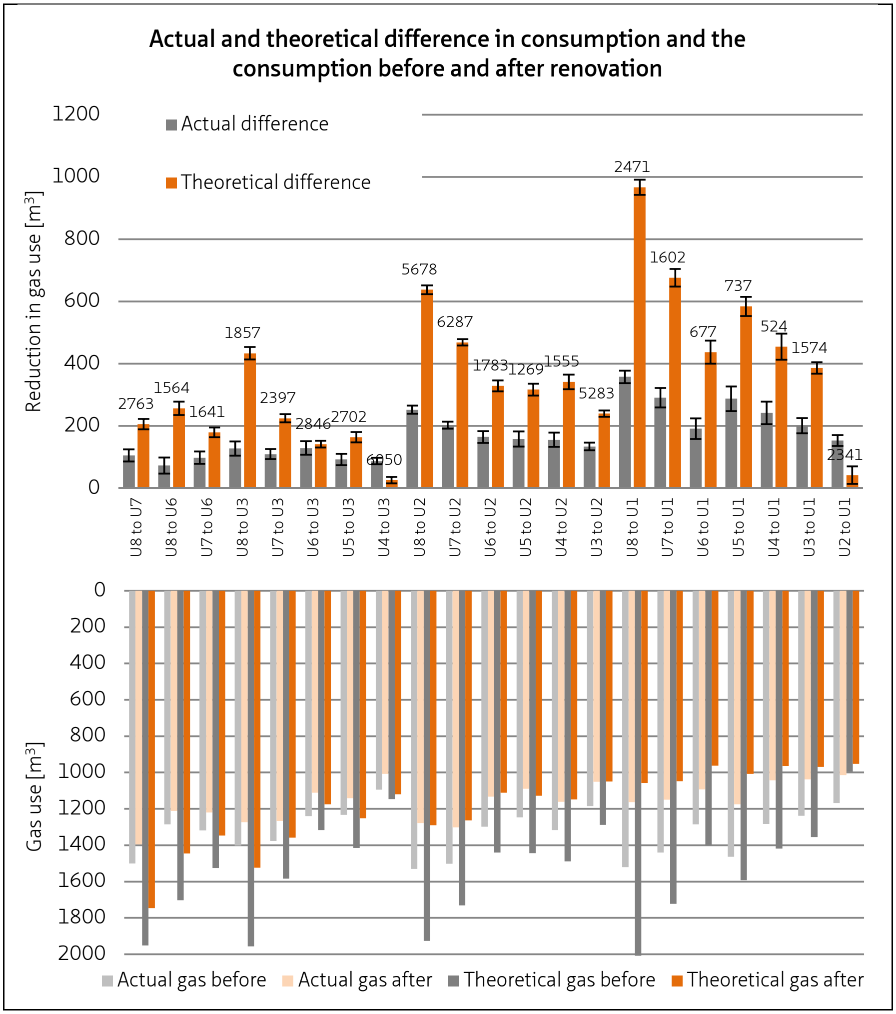


FIGURE 7 Actual and theoretical reduction and number of cases (above graphic) between the first and second pre-label in dwellings with renovated windows (U-value) – sample method B (N>30) Confidence intervals in the bottom graphic are omitted in the bottom graph for better readability

As opposed to previous measures, in this section there is no specific measure that stands out in terms of frequency. This is a feature of window as well as envelope replacements, probably partly because average insulation values were analysed (section 5.4.2). Replacing the glazing never comes close to the expectations, but rather to about half of the predicted saving. Dwellings which were subject to a deeper renovation of windows exhibit a larger reduction in actual gas use (U8 to U1 yielded 357 m^3 reduction out of the expected 966 m^3 reduction). U8 to U7 yielded 105m3 reduction out of the expected 206 m^3 reduction and U2 to U1 290 m^3 out of the theoretical 676 m^3. There are, however, some inconsistencies, such as the group of dwellings which had windows improved from U5 to U1, which saved more than the group with more drastic renovation of U6 to U1. It is possible that a certain group of dwellings contains a large residential dwelling block which had specific renovation characteristics which skews the result of a particular category.

Another thing noticeable from Figure 7 is, that dwellings which had their windows replaced to a more moderate standard (U3-U5) and did not start out with the worse window quality (U8), but rather a U6-U7, exhibit the best match between actual and theoretical reduction. It is nevertheless questionable whether these changes were real renovations or administrative corrections, since such windows are these days not considered standard anymore.

The bottom graphic in Figure 7 shows that the positive performance gap (overprediction) observable before renovation everywhere except in the very last column(category U2 to U1) is just as present after the renovation for all categories except U2 and U1. It seems that dwellings with U value U3 and higher always consume less than predicted whereas others consume more.

B Dwellings with a change in window quality only

Figure 8 reveals that dwellings that had a drastic change in window quality (U8-U2,U7-U1) tend to have an actual gas reduction lower than the theoretical. This phenomenon was seen before in Figure 7 – where just like in Figure 8, the least drastic changes were the best predicted. Some more moderate changes have an actual reduction closer or exceeding the predicted one (U6 to U3, U5 to U2), which is also the case for some small improvements (U2 to U1 or U8 to U7). One also needs to keep in mind that in some cases the actual gas reduction seems to be smaller than the background gas reduction (see section 5.4.4), for example U4 to U1.

Looking at the absolute gas consumption before and after renovation one can see (bottom graph in Figure 8) that the overpredictions observed in bottom chart of Figure 7 in categories U3 and larger is less visible (in some categories they are still notable but in much smaller scale than previously). Also the underpredictions noted previously for U2 and U1 no longer appear consistently. One can therefore hypothesise that the trends seen in Figure 7 were mostly a consequence of a high correlation of window insulation value with other measures taken.

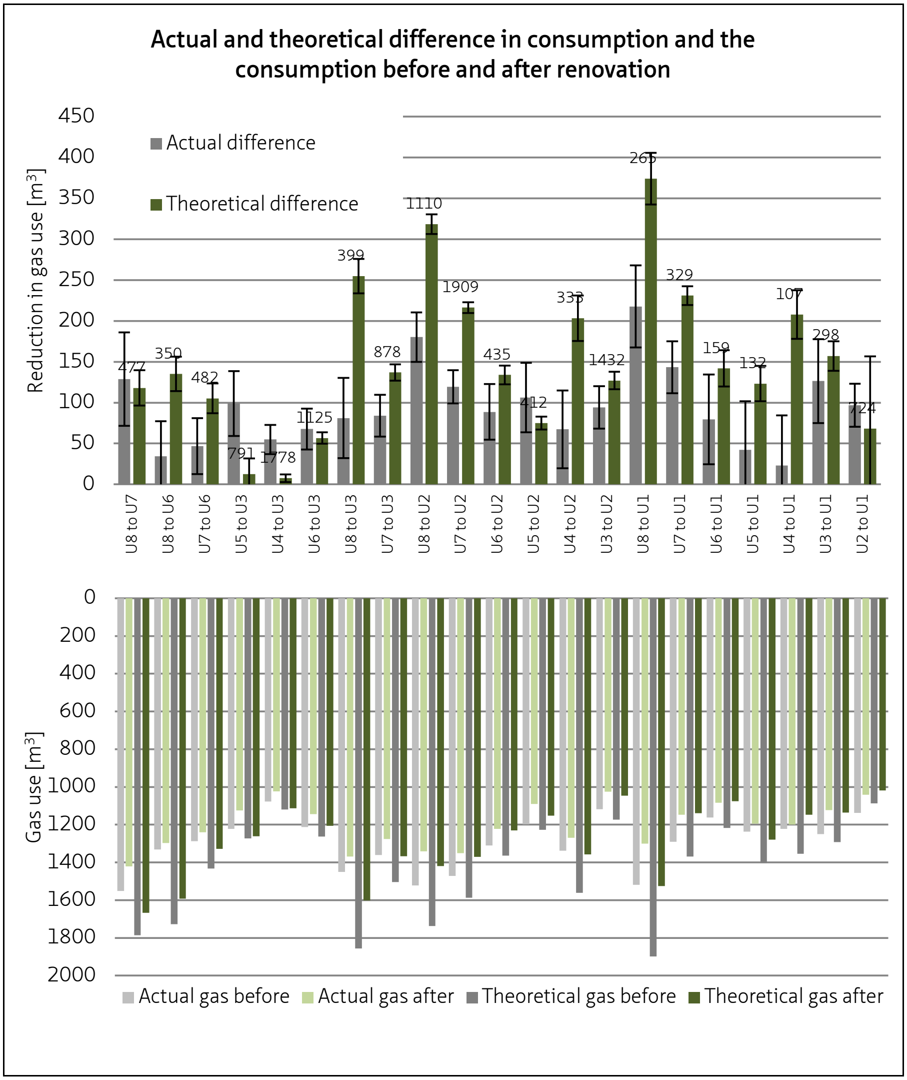


FIGURE 8 Actual and theoretical reduction and number of cases (above) between the first and second pre-label in dwellings with replaced windows (U-value) – sample method C (confidence intervals are omitted in the bottom graph for better readability), N>30. Confidence intervals in the bottom graphic are omitted in the bottom graph for better readability

§ 5.5.5 Changes in envelope quality

This section shows the actual and theoretical reduction of dwellings which had an improvement in the envelope, excluding the groups of dwellings which contained less than 30 cases to ensure statistical significance. The insulation values as described in Table 1 are used. To keep the results in the Figure 9 below readable, we do not show changes of envelope to insulation category R2 and R3. These results follow the same pattern so not much is lost by not conveying those results, which are included in section 5.5.6.

A Dwellings with a change in envelope quality

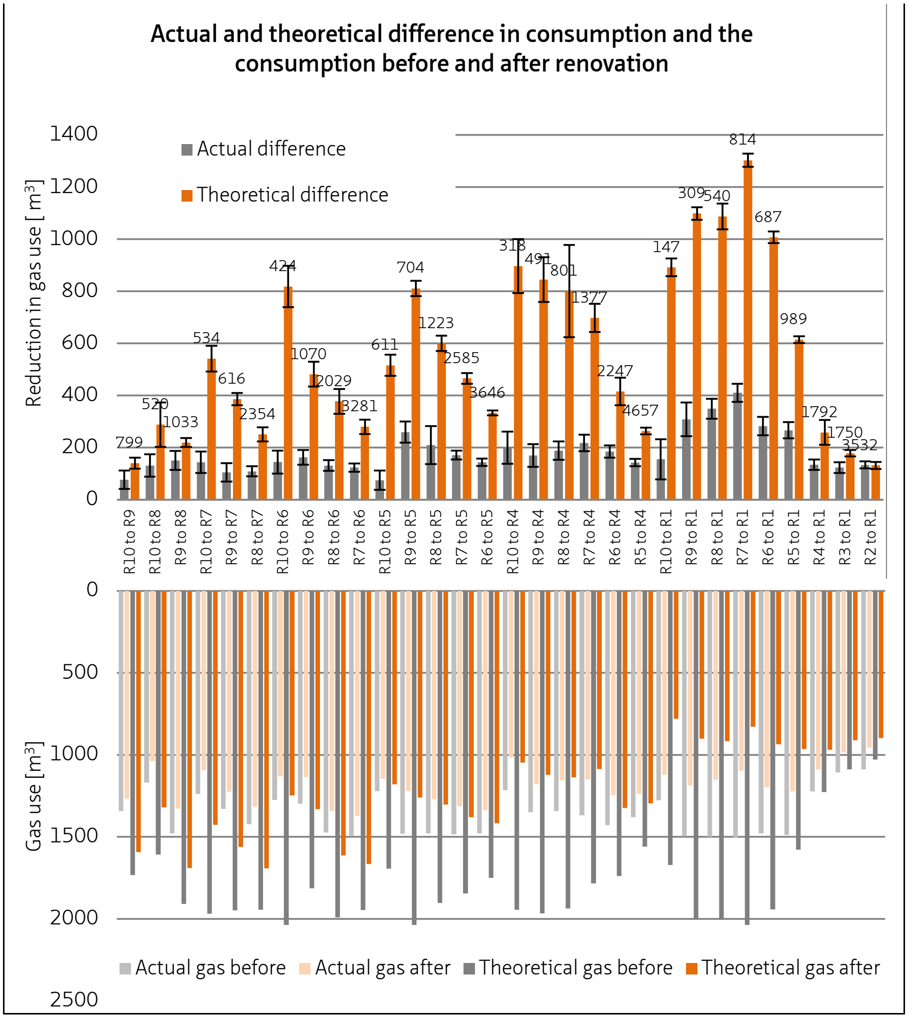


FIGURE 9 Actual and theoretical reduction and number of cases (above) between the first and second pre-label in dwellings with replaced envelope insulation (R-value) – sampling method B, N>30. Confidence intervals are omitted in the bottom graph for better readability

Just as in case of window renovations, there is no measure that stands out in terms of frequency like in the installation measures. The least drastic changes again result in the actual reduction closest to the theoretical, just like in window insulation measure. Even drastic changes yield at most about a third of the expected saving. Roughly, strong overprediction occurs in R5 to R10 and slight underprediction in R1 to R4.

B Dwellings with a change in envelope quality only

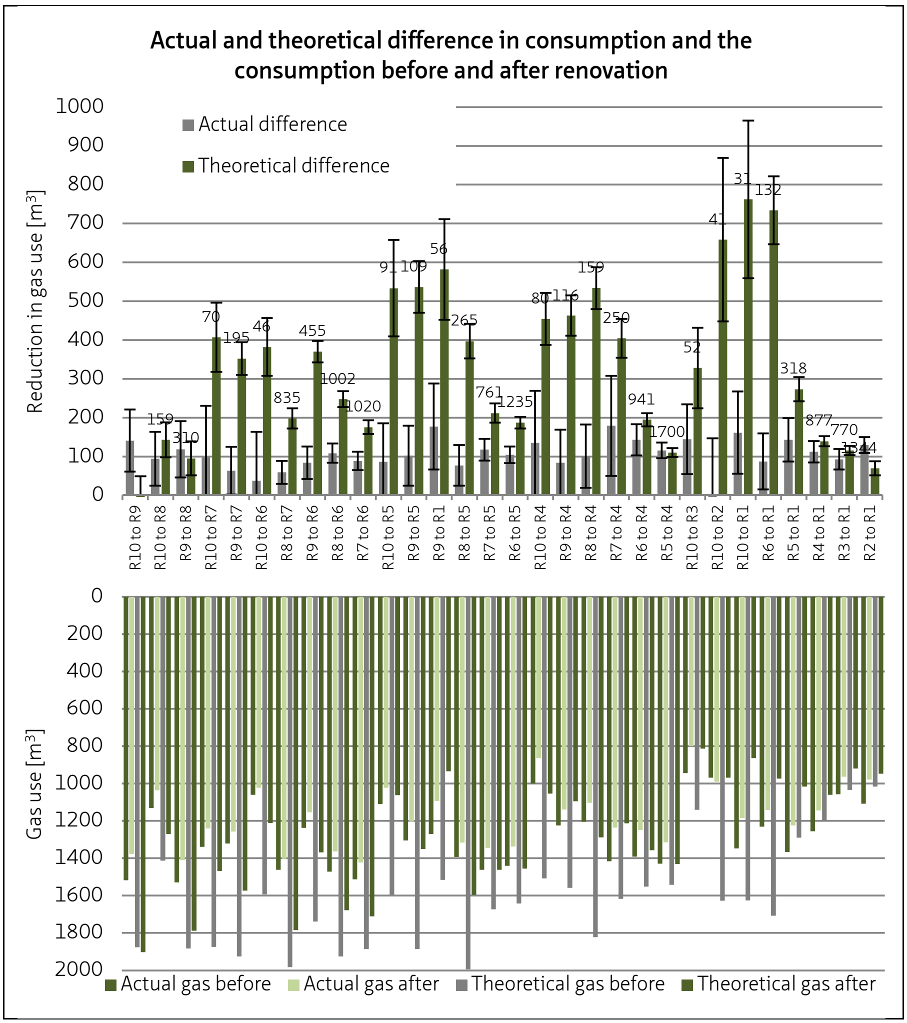


FIGURE 10 Actual and theoretical difference between the first and second pre-label in dwellings with changed envelope insulation (R-value) – sampling method B (N>50). Confidence intervals are omitted in the bottom graphic for better readability

The R value of the envelope is an average value of floor, wall and roof and due to averaging there are fewer dwellings with drastic improvements of the envelope, mostly they only improve by one or two categories. This might seem dissapointing, but in a dwelling with envelope of 300m^2 and an R value of 0,4 insulating the roof (10% of total area) with R=2,5, leads to a new R value of 0,31, which corresponds to a change for one category only (R5 to R6).

The results are similar to those for improving U value of the envelope – small changes are well predicted and actual reduction is close or surpassing the theoretical whereas deeper changes result in actual reduction being much lower than predicted. The better the dwelling is insulated, the easier it is to achieve the envisioned saving, as in general, the gap between predicted and actual consumption is larger in insulations R5 and higher (bottom graphic of Figure 10).

§ 5.5.6 Actual consumption savings among different measures

One of the objectives of the paper was to see which measures are most effective in achieving energy savings. Several tables in this section demonstrate average reduction rates for separate measures. First of all, averages of various measures are calculated in

Table 7 and Table 8 taking into account all the groups containing more than 30 records, the first summing up the results of sample B and the later of sample C. Sample B studies a larger sample, therefore the totals and numbers of dwellings within a measure are, logically, higher. Interestingly, the measure which achieves the largest actual cumulative as well as individual saving in sample B is window replacement and in sample C it is the replacement of heat and hot tap water system. In both samples envelope improvement is in the second place and ventilation improvement the last. Looking at savings in the two tables, both actual as well as theoretical consumption reductions are higher in sample B than in C, which makes sense, since there is a large chance that dwellings in sample B had another renovation measure taken. Comparing the numbers of dwellings in each measure group (last column) reveals that the group of heating and hot tap water has the highest similarity in both samples, since sample B had 60.960 dwellings in this group and sample C 30.749, which is more than half. This means that more than half of the dwellings with a change in heating and hot tap water had no other dwelling change, whereas the other smaller half, did. About two thirds of dwellings with envelope improvement also had other measures taken (21.035 in sample C vs. 62.955 in sample B) and about three quarters of dwellings with window improvement also had other measures taken (15.744 in sample C vs. 61.233 in sample B). The measure which was most usually combined with others was ventilation improvement, which also explains the drastic difference in reduction of this group in sample B and C. When comparing the ratios of actual vs. theoretical reduction, one first notices a higher average ratio in sample C than that of sample B. This means that dwellings with a single renovation measure have on average a better predicted reduction than those with combined measures. The most remarkable considering individual measures, is the gas reduction in dwellings with an improved ventilation systems, achieving a 2,5 times higher reduction than predicted.

RENOVATION MEASURE	CUMULATIVE SAVING (TOTAL SAMPLE)		INDIVIDUAL SAVING (PER DWELLING)			N
	TOTAL ACTUAL GAS REDUCTION	TOTAL THEORETICAL GAS REDUCTION	AVERAGE ACTUAL GAS REDUCTION [M³]	AVERAGE THEORETICAL GAS REDUCTION [M³]	RATIO ACTUAL /THEORIETICAL GAS REDUCTION	
Ventilation	11%	11%	148	327	0.45	26,325
Windows	33%	30%	203	363	0.56	61,233
Envelope	25%	29%	147	352	0.42	62,955
Heating and hot tap water	31%	30%	190	365	0.52	60,960
Total [m³]	37,177,026	75,269,315				211,473
Average			176	356	0.49	

TABLE 7 Totals and averages of actual and theoretical gas reduction for different renovation measure using sample B - non-exclusive measure (groups with N>30)

RENOVATION MEASURE	CUMULATIVE SAVING (TOTAL SAMPLE)		INDIVIDUAL SAVING (PER DWELLING)			N
	TOTAL ACTUAL GAS REDUCTION	TOTAL THEORETICAL GAS REDUCTION	AVERAGE ACTUAL GAS REDUCTION [M³]	AVERAGE THEORETICAL GAS REDUCTION [M³]	RATIO ACTUAL /THEORIETICAL GAS REDUCTION	
Ventilation	4%	1%	73	29	2.52	4,848
Windows	16%	14%	96	134	0.72	15,744
Envelope	23%	25%	104	180	0.58	21,035
Heating and hot tap water	57%	56%	172	279	0.62	30,749
Total [m³]	9,367,264	14,622,945				72,376
Average			131	188	0.70	

TABLE 8 Totals and averages of actual and theoretical gas reduction for different renovation measure using sample C – unique measure (groups with N>30)

	ACTUAL REDUCTION [M^3]	N	RATIO
U8 to U1	218	265	0,6
η<83% to η>83% and CC to CI	212	127	0,9
η<83% to η>83% and On-d. to CI	193	752	2,4
η>83% to η<96% and CI to CH	184	23,902	0,7
U8 to U2	180	1,110	0,6
η<83% to η<96% and CC to CH	180	681	0,3
η>83% to η<96% and On-d. to CH	178	1,445	0,7
η>90% to η<96% and On-d. to CH	166	76	1,7
U7 to U1	143	329	0,6
R5 to R1	143	318	0,5
η>83% to η <90% and CI to CH	135	77	0,5
U8 to U5	133	253	0,5
R2 to R1	130	1,344	1,9
U8 to U7	129	477	1,1
R8 to R3	128	90	0,2
U3 to U1	126	298	0,8
η<83% to η<96% and On-d. to CH	122	1,911	0,3
R4 to R1	113	877	0,8
R8 to R6	109	1,002	0,4
R8 to R4	101	159	0,2
U8 to U4	99	111	0,4
U2 to U1	97	724	1,4
R3 to R1	93	770	0,8
R6 to R1	87	132	0,1
U8 to U3	81	399	0,3
U6 to U1	80	159	0,6
R8 to R5	77	265	0,2
Natural to mechanical exhaust	76	4,479	5,0
LG to η<96% and On-d. to CH	59	1,657	0,1
R8 to R7	59	835	0,3
Natural to mechanical balance	54	49	1,7
Mechanical exhaust to mechanical balance	50	279	0,2
U5 to U1	42	132	0,3
U8 to U6	34	350	0,3
U4 to U1	23	107	0,1
η>83% to η <94% and CI to CH	15	72	0,1
LG to η>83% and On-d. to CI	10	121	0,1
Mechanical exhaust to on-demand decentralised mechanical with mechanical exhaust	-50	41	-0,8

TABLE 9 Actual consumption reduction per dwelling of various single renovation measures

Table 9 shows the actual gas reduction, the number of dwellings and the ratio between actual and theoretical consumption reduction. The highest reduction is achieved by drastically improving the U value of the windows (U8 to U1). The actual reduction of such a change (Table 9 first row left) is below the theoretical and the number of dwellings in this category is rather low. The category containing the most dwellings, is the one where heating systems were replaced from a η>83% to η<96% and hot tap water installation renovated from improved to high efficiency. The actual reduction of this group is also below the expected. The measures achieving the most reduction are therefore drastic improvements of window quality and an improvement of the efficiency of heating and hot tap water system (not a replacement of a local system).

Measures that achieve an actual reduction higher that the theoretical seem to mostly be less drastic changes, such as insulation improvement from R2 to R1 or window improvement from U8 to U7 or U2 to U1. Also notable is the underprediction of the reduction in dwellings where natural ventilation was replaced by mechanical exhaust and it is questionable whether such dwellings still have a sufficient quality of indoor air after the renovation. The two heating installation improvements that yielded a reduction higher than theoretical (third and eight row of Table 9) are both within a certain boiler type (in first case non-condensing and in the second, condensing), improvements in between these categories have an actual consumption lower than the theoretical one. This probably means that some of the calculation factors used for efficiencies of gas boilers do not reflect the real efficiency correctly.

MEASURES RESULTING IN HIGHEST CUMULATIVE SAVING	ACTUAL GAS REDUCTION *N [M³]	N	% OF TOTAL REDUCTION IN STUDIED SAMPLE
Heating boiler η<83% to η>96% hot water from improved to high-efficiency boiler	4,396,716	23,902	38%
Natural to mechanical exhaust ventilation	340,404	4,479	3%
Heating boiler η<83% to η<96% hot water from on demand to high-efficiency boiler	257,204	1,445	2%
Heating boiler η>83% to η<96% hot water from on demand to high-efficiency boiler	233,094	1,911	2%
U8 to U2	199,800	1,110	2%
R2 to R1	174,720	1,344	2%
Heating boiler η<83% to η>83% hot water from on demand to improved efficiency boiler	145,277	752	1%
Heating boiler η<83% to η>96% hot water from conv. to high-efficiency boiler	122,457	681	1%
R8 to R6	109,218	1,002	1%
R4 to R1	99,100	877	1%

TABLE 10 Cumulative actual gas consumption reduction of the studied sample

Results in Table 9 are informative in terms of the efficiency of individual measures, however, the problem is that many of these results have poor statistical significance due to the low sample size (the confidence bands can be seen in previous sections). To emphasise the measures which yield the most savings in the studied sample, Table 10 sorts the measures according to the cumulative saving – the sum of the savings of all dwellings in a particular category. This is of course strongly dependent on the sample, but if we consider the studied sample representative it is impressive how much actual gas reduction (38%) comes from replacing the heating and hot tap water system and that 3% of savings come from upgrading the natural ventilation system. Probably the popular measures are the most cost-effective ones.

§ 5.6 Discussion

The results section showed results using three sampling methods. Cross sectional method (A) was only used for dwellings total thermal performance (energy label) and comparison with method B yielded similar results in terms of performance gap (see ratio column in Table 6) unless looking at small changes (mostly one label step) of very poor or very well performing dwellings (e.g. G to F or B to A). Summarizing, longitudinal data is essential when examining the effect of single renovation measures. Albeit carefully, cross sectional data can be used for estimating deep improvements in overall performance (roughly, more than one label class).

The reason could be that in those extreme labels (G or A), cross sectional method compares entities that are not comparable – for example, dwellings in A label are significantly larger than B dwellings (Majcen et al. 2013a), or they could be characterised by a much larger number of occupants. Longitudinal methods do not ensure that analysed dwellings have not undergone a change in household – the chance is, however, much smaller than in cross sectional data, where we know households to be different in each dwelling group. However, even though the ratio of the performance gap across label classes is roughly similar, the actual gas consumption reduction is consistently larger using longitudinal data than cross-sectional data. This highlight the importance of longitudinal data collection for better estimation of actual gas reduction.

If the theoretical consumption before and after renovation would be comparable using method B and C, it would mean that sample B represents well the theoretical consumption of the observed measure. This is, however, almost never the case, since sample B includes a number of cases where also other measure have occurred. Comparing method B and C for renovation measures in fact yielded roughly comparable results when it comes to dwelling insulation (window and envelope) and very different results when looking at installation systems. It seems that better performing systems in general exhibit a smaller performance gap, such as boilers with a higher efficiency, mechanical ventilation and better insulation. Two very notable performance gaps were the one in local gas heater and on-demand tankless water boilers and naturally ventilated buildings. The most extreme example are dwellings with a changed ventilation system where the performance gap ratio in method C is 4 times the ratio of method B. This proves that when analysing single measures, one should definitely ensure other properties are constant making the results of method C are therefore a better basis for conclusions regarding performance gap and actual reduction of the measures. The problem of this method is, however, that we (currently) cannot find enough data to provide significant results for many of the possible combinations of measures, which should be improved in the future with expansion of SHAERE.

The average actual gas reduction in sample B is 176 m^3, which represents 15,5% of the total consumption (see Figure 2) and corresponds to one or several implemented measures. The theoretical reduction of this same sample, 356m^3 makes 27,4% of the theoretical total consumption (Figure 2). For single measures (sample C) the actual and theoretical gas reductions are 131 and 188m^3 which makes up for a reduction of 11,6% and 16,9%. Hezemans et al. from 2012, who assumed that two measures coincide with a 20% reduction, was therefore quite close to reality, although the actual average value is somewhere between 11,6 and 15,5%.

There are some uncertainties regarding the results. According to Aedes, pre-labels are updated whenever a renovation measure takes place and are considered accurate, however, the fact that a number or deteriorations were identified within SHAERE demonstrates that this is not entirely true. This could probably improve in the future as the database grows, however, it is a major uncertainty in this study. This study was done purely on social housing sector and moreover excluded certain heating types (heat pumps), which has consequences for representativeness of the results. Another situation in which a dwelling was not considered in this paper is if during the renovation, its address changes, which is the case in a number of deep renovations. At the time of the study, it was not possible to find out the extent to which this occurs. Moreover, certain parameters such insulation of wall, floor and roof have been aggregated in this paper and would be interesting to analyse independently using continuous instead of categorical values. In section C we analysed the change in one of the dwelling properties, however, we neglected the impact of others (even though constant). For example, it might be significantly different whether the dwellings which had a renovated installation system was very well or poorly insulated. In the future, other statistical methods (correlation tests, regression analysis) should be tested on similar large data, since this allows to include more variables and also enables the use of control variables. In the upcoming studies, one could also limit oneself to deeper performance changes. Here we observed all changes (also small ones, within one label category), however, the results might be more robust selecting a subsample where one or even two label steps have been taken – especially in line with the uncertainties regarding administrative corrections in the data.

§ 5.7 Conclusions

To conclude, several main findings are summarized below.

- In terms of gas reduction by single measures, improvements in efficiency of gas boilers (space heating and hot tap water) yield the biggest energy reduction, followed by deep improvements of window quality. Improving the ventilation system yields a relatively small reduction compared to other measures, however, it is still much larger than theoretically expected.

- In terms of the performance gap between actual and theoretical consumption, high R and low U values of insulation are well predicted, as well as efficient heating systems. On the other hand low R and high U values, local heating systems, changes from a non-condensing into a condensing boiler and upgrades to a natural ventilation system are not well predicted. In Majcen et al., 2013b, it was shown that departures from the standard average dwelling temperature were causing a part of the performance gap and in the present paper it is shown that efficiencies of heating systems and insulation values are also causing a part of the gap.

- This poses the question of how well the standard values are really defined in the calculation method. It could be that excessively low efficiencies have been attributed to inefficient systems simply because of misconception and lack of knowledge, as from an economical point of view, it is more logical to invest effort into estimating the performances new systems. However, not knowing the real efficiencies of older systems results in a performance gap.

- However, since actual consumption data on the level of individual dwelling is becoming available these inconsistencies become visible. The standard values should either be revised or alternatively, one should utilise the available actual gas consumption values in order to make better estimates (Majcen et al., 2015).

- Large datasets such as the SHAERE investigated in this paper are now arising across Europe and few experience is available about how to handle them. The results of large samples are statistically robust and representative, however selecting subsamples from the data offers insight into specific combinations of measures and allows identification of best practices. Energy performance registers should be made publicly available, possibly already coupled with actual consumption data.

- It is of utmost importance to ensure that building performance databases are of sufficient quality and have trustworthy input data. Ensuring such level of quality is not simple, even if dwellings are used for asset management by large housing companies (associations). This paper has highlighted the importance of analysing dwelling stock registers for both the validation and evaluation of energy label calculation. However, in The Netherlands, a simplified label came into force in 2015 next to the existing, complete label. This changed a lot in this field, since the simplified label requires no inspection at all and can be filled in online by the owner of the house himself. The implications of this simplified label are not yet clear, just as it is not clear yet, whether housing associations will continue to inspect a dwelling and get a complete energy label or not.

- Further study should also include costs of the different renovation measure. The results of this paper showed that windows and installation system upgrades provide a high actual reduction, and the remaining question is which of the two is more viable economically. This question is relevant also in the framework of cost effectiveness of nZEBS according to EPBD.

Overall, this paper has shown once more that the calculation method currently in use cannot be considered accurate if compared to actual consumptions. The question that remains is how to, under these circumstances, determine the effectiveness of a specific renovation measure, which is of importance on dwelling level and even more so on the level of the whole stock. If theoretical methodology is to be used as baseline without the use of actual consumption at some point in the process, realistic standard values have to be prescribed.

§ 5.8 References

Adalberth, K., 1997. Energy use during the life cycle of single-unit dwellings: Examples, Building and Environment, Volume 32, Issue 4, Pages 321-329.

Delghust, M., Roelens, W., Tanghe, T., Weerdt, Y.D., Janssens, A. 2015. Regulatory energy calculations versus real energy use in high-performance houses, Building Research & Information. Pages 1-16.

Dodoo, A., Gustavsson, L., Sathre, R. 2010. Life cycle primary energy implication of retrofitting a wood-framed apartment building to passive house standard, Resources, Conservation and Recycling, Volume 54, Issue 12, Pages 1152-1160

Economidou, M., Atanasiu, B., Despret, D., Ingeborg, J.M., Rapf, N.O. 2011. Europe's buildings under the microscope, Country-by-country Review of the Energy Performance of Europe's Buildings, BPIE, 2011.

Filippidou, F., Nieboer, N., Visscher, H., 2015a. Energy efficiency measures implemented in Dutch non-profit housing sector, ECEEE 2015 Summer Study proceedings, Hyeres, France.
Filippidou, F., Nieboer, N., Visscher, H., 2015b. The energy renovation pace of the Dutch non-profit housing sector, Submitted to Energy Policy in July 2015
Hezemans A., Marquart E., Monné T., Monitor Energiebsparing Gebouwde Omgeving 2012, Agentschap NL, Juni 2012.
ISSO 82.3 Publication Energy Performance Certificate—Formula Structure (Publicatie
82.3 Handleiding EPA-W (Formulestructuur'), Senternovem, October 2009.
Karlsson, J.F., Moshfegh, B., 2007. A comprehensive investigation of a low-energy building in Sweden, Renewable Energy, Volume 32, Issue 11, Pages 1830-1841
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., Visscher, H., 2015. Statistical model of the heating prediction gap in Dutch dwellings: Relative importance of building, household and behavioural characteristics, submitted to Energy and Buildings in June 2015
Menkveld, M., Leidelmeijer, K., Vethman, P., Cozijnsen., E. 2012. Besparingsgetallen energibesparende maatregelen, ECN, May 2012
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.
Thormark, C. 2002. A low energy building in a life cycle—its embodied energy, energy need for operation and recycling potential, Building and Environment, Volume 37, Issue 4, Pages 429-435
Tigchelaar, C., Leidelmeijer, K. 2013. Energiebesparing: Een sampenspel van woning en bewoner – Analyse van de module Energie WoON 2012, ECN, August 2013
Visscher, H., Majcen, D., Itard, L. 2013. Gebruik van de SHAERE-database voor het monitoren van het Convenant Energiebesparing Huursector, Technische Universiteit Delft, Faculteit Bouwkunde, OTB - Onderzoek voor de Gebouwde Omgeving.
Winther, B.N., Hestnes, A.G., 1999. Solar Versus Green: The Analysis of a Norwegian Row House, Solar Energy, Volume 66, Issue 6, Pages 387-393.

6 Conclusions

§ 6.1 Introduction

This thesis has described the actual and theoretical energy consumption of Dutch residential dwelling stock and how they differ across varying energy performance levels. Furthermore, it analysed the causes for these discrepancies and the implications they have for policy. The study examined the existing dwelling stock, with a focus on dwellings that have been labelled with an energy label certificate, a tool introduced by the European Energy Performance for Buildings Directive (EPBD) in 2002. In The Netherlands the energy label certificate includes a steady state thermal model of the dwelling, which is the basis for the scale indicated on the label. The scale spans from A to G, with G indicating the lowest performing dwelling. Energy labelling is directed at reducing energy consumption in existing dwelling stock by informing potential buyers/renters about the energy performance of the house or flat, which should encourage the sale of higher performing dwellings and therefore lead to more thermal renovations. However, at the beginning of the study not much was known about what the actual energy consumption of the labelled dwellings was, as this was a relatively new regulation (introduced in The Netherlands in 2008) and the government register of the certificates had only recently been established. Since potential energy savings depends on actual consumption levels and not on the theoretical consumption levels assumed by the energy label certificate itself, the extent of actual energy consumption in these dwellings is important for policy makers as well as the construction industry.

As work on the thesis progressed it was found that the difference between theoretical and actual consumption, also referred to as the 'performance gap', arises due to a normalisation of the indoor conditions of the dwelling as well as due to assumptions about infiltration rates, efficiencies of the systems, etc. Incorrect assumptions occur because of a lack of knowledge about the real performance of the dwelling, such as building air tightness or the actual efficiency of boilers. Furthermore, specific dwelling systems seem to encourage particular behaviours. For example, in a dwelling with floor heating the heating is on even when occupants are not present. Moreover, occupants of lower performing dwellings seem to realize the wastefulness of excessive heating more readily than those living in higher performing dwellings, where an incremental change in temperature leads to only a slight increase in the energy bill. As a result, actual indoor temperature depends strongly on the type of heating system and the dwelling itself, whereas the theoretical calculation method assumes an equal indoor temperature in all heating systems. The discrepancy between actual and theoretical consumption is therefore logical and by itself not problematic. However, as the thesis shows, existing policies do not take these discrepancies sufficiently into account. The theoretical reductions in consumption attributed to dwelling renovations turned out to be significantly higher than the actual reductions. If policies are based on erroneous theoretical baselines and assume no behavioural changes take place in the households when renovations are completed, the expected reduction in consumption will fail to materialise.

Energy label certificates originating from the national RVO register were analysed in several large stock samples. The certificates containing the theoretical heating consumption levels of dwellings were coupled with actual consumption data for individual dwellings, obtained from the national statistics office (Statistics Netherlands). Finally, the merged data was enriched with additional socioeconomic information about the dwellings, households and the behaviour of the occupants. The objective was to establish whether theoretical consumption deviates from actual consumption to the extent that it creates adverse consequences for the effectiveness of policies, why these discrepancies arise (e.g. what makes the theoretical calculation differ from the actual) and how to mitigate them in the future.

The main research question of the thesis was defined as:

What are the characteristics and consequences of the discrepancies between the theoretical and the actual use of heating energy in Dutch dwellings?

In order to answer this question, the discrepancies (also referred to as the performance gap) between the theoretical and actual gas consumption were analysed thoroughly. At the time the study was begun, these discrepancies had not yet been studied in labelled dwelling stock. However, existing research into energy consumption levels in newly constructed dwellings in The Netherlands as well as experience in other countries with existing dwelling stock (see Section 1.2 of the Introduction) suggested that a performance gap might also exist in the labelled dwelling stock. To confirm this, an analysis of the discrepancies was carried out in five datasets, ranging in size from several hundred to several million dwellings with an energy label certificate (Table 2 of the Introduction). Most of these datasets were quite large and fit a new trend of the availability of big data. Their large size permits a well-representative population to be drawn instead of the small samples which used to dominate this type of research. Additionally, the content of the datasets was exceptional, since it included complete sets of building characteristics. The scope of the research was narrowed down to gas consumption, since electricity was constant in all label classes and almost all dwellings in The Netherlands are heated with gas.

The detected performance gap was indeed significant, with dwellings with a G label consuming around half as much as expected and dwellings with an A label consuming around a quarter more than expected. These were exciting results and encouraged a further investigation of consequences for this phenomenon. A scenario study was performed (Chapter 2), comparing the current policy targets with the renovation agreements set by the Dutch government and private stakeholders (construction industry, housing associations, etc.). The renovation scenarios were then applied to two baselines: the theoretical consumption levels and the actual consumption figures from the dwellings. It turned out that the current policy targets for energy reduction are unreachable using actual consumption as the baseline, although they can be achieved using the theoretical levels. This proved that actual consumption is not taken into account by policy makers, which inspired further investigation of real reductions in renovated dwellings (Chapter 5). Here it turned out that indeed, on average, renovated dwellings achieve savings that are one-third smaller than theoretically expected. However, depending on the individual measure, significant variations were found (for example, upgrading a dwelling's natural ventilation to mechanical exhaust yielded an actual reduction far above the theoretical level), which is a very valuable insight for industry as well as for policy makers. Since as stated, the performance gap turned out to have great implications for policy, the thesis also analysed the causes for the discrepancies. This was done by a sensitivity analysis of theoretical consumption described in Chapter 3, where several scenarios of the assumptions used in the label calculation were tested, showing that by slightly modifying the indoor temperature or the building's envelope quality, theoretical consumption levels change drastically and can in some cases account for the detected performance gap. In addition to the sensitivity analysis, two extensive regression analyses were made (Chapter 3 and 4), where the influence of occupant, building and household characteristics on the actual and theoretical gas consumption numbers were quantified.

In accordance with the progression of the research work on one hand and data availability on the other, the thesis consists of four chapters, presented in three sections: A, B and C. Three chapters have been published, and one has been submitted for publication. Section A analyses the discrepancies, B examines their consequences and C investigates the causes for the performance gap. The research questions, which were presented in the Introduction, reflect the data covered in each of the three sections.

A The discrepancies between actual and theoretical heating energy consumption in Dutch dwellings

Before beginning the thesis research, an extensive literature review was carried out, a summary of which can be found in Section 1.2 of the Introduction and also at the beginnings of Chapters 2 to 5. After examining the existing studies, it seemed that actual and theoretical heating consumption levels can vary greatly in different performance categories. This section discusses whether or not such differences were found in the studied samples of labelled dwellings (sub-question A.2). The preceding section A.1 explains whether or not there are differences between theoretical and actual heating consumption in the total stock on average. Section A.3 describes which other dwelling characteristics were found to correlate with the discrepancy between theoretical and actual gas use (besides the thermal performance of the dwelling mentioned in A.2). Note, however, that whereas sub-questions A.1 and A.2 explore gas and electricity consumption (which together constitute primary energy consumption, also expressed as CO_2 emissions), sub-question A.3 focuses only on gas use. This focus was decided upon based on the outcomes of sub-questions A.1 and A.2.

A.1 What are the discrepancies between theoretical and actual gas and electricity consumption in the total dwelling stock?

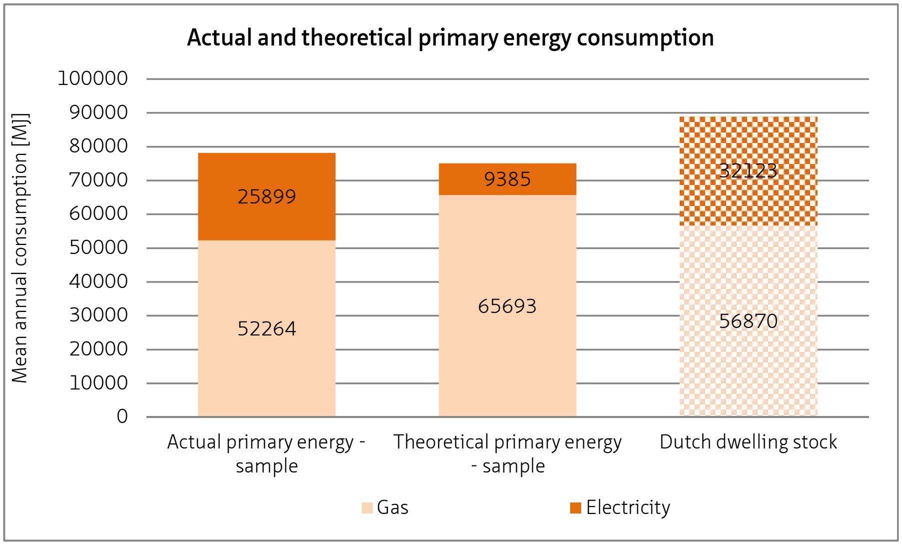


FIGURE 1 Actual and theoretical mean primary energy consumption per dwelling in the sample and in the Dutch dwelling stock (N=193,856)

A comparison of actual and theoretical primary energy consumption in the total dwelling stock in Chapter 2 (Figure 6) showed that, on average, the total theoretical primary energy use seems to be in accordance with actual primary energy consumption. However, when looking at more detailed data, one can see that the contribution of gas to the total actual primary energy use is much lower than is reflected in theoretical primary energy use and that it is the opposite for contribution of electricity – higher in the actual numbers than theoretical primary energy. These two effects cancel each other out so that in terms of total primary energy use, theoretical consumption seems to be accurately predicted (Figure 1).

Although it is clear that theoretical electricity consumption is much lower than actual usage since it does not account for appliances, it is much less obvious why gas consumption is on average so much lower in reality than it is in theoretical calculations. This relation was suspected to be different in dwellings with different performance levels, which is why analysis per label class was conducted in the next step.

A.2 What is the relation between actual and theoretical gas/electricity/primary energy/CO_2 emissions in dwellings with different energy labels?

These discrepancies were first studied in the second chapter, which used the RVO database, the first available sample that enabled the analysis of theoretical as well as actual energy consumption in dwelling stock level in The Netherlands. In this sample theoretical and actual energy consumption were broken down into gas (Figure 7 of Chapter 2) and electricity (Figure 10 of Chapter 2). Moreover, cumulative primary energy consumption (Figure 12 of Chapter 2) as well as CO_2 emissions (Figure 13 of Chapter 2) were analysed. The discrepancies were analysed across the label categories.

The discrepancies in gas consumption were largest in the lowest performing dwellings (label category G), where theoretical consumption surpassed the actual by nearly 200%, which we referred to as over-prediction. On the other hand, higher performing dwellings consume roughly 20% more gas than predicted (gas consumption is under-predicted). This corresponds to the phenomena observed in the existing literature on the space heating of dwellings (1.2 of Introduction), which is logical since most Dutch dwellings (and most dwellings in the RVO sample) are heated with gas. An important difference is that theoretical gas consumption does not include gas used for cooking, which is included in the actual gas consumption figures. Since gas used for cooking contributes marginally to overall gas consumption, at roughly $50m^3$ annually per dwelling, it does not affect the discrepancies significantly. Both actual and theoretical gas consumption do take into account gas used for heating and hot tap water.

Theoretical electricity consumption was at least two times lower than actual consumption in all label categories, due to the fact that actual consumption takes into account the electricity use of appliances and theoretical consumption does not. Actual and theoretical electricity consumption seem to be rather constant with regard to the label class. There did not seem to be a coherent relation between label category and theoretical or actual electricity use, except for a slightly higher theoretical electricity consumption in label A, probably due to a few dwellings that were heated with electricity (heat pumps).

Both primary energy consumption and CO_2 emissions are essentially sums of the consumption of gas and electricity in MJ for each label class where the efficiency of both the electricity generation and the network (η=0.39) are taken into account, as well as the heating value of gas burning (35.17MJ/m^3). The theoretical primary energy use is dominated by gas consumption, since electricity constitutes a relatively small fraction of primary energy use due to the exclusion of household appliances. The relation between actual and theoretical use therefore remains similar, as seen in gas consumption. For lower performing label classes, theoretical consumption is over-predicted by about 30% and for higher performing label classes it is under-predicted by roughly the same percentage.

For analysing CO_2 emissions, emission factors of 0.0506kg CO_2 per MJ gas and 0.0613kg CO_2 per MJ electricity were applied, meaning that these results were dominated more by the constant values of electricity than by gas. Over-prediction in labels for lower-performing dwellings was therefore only slight, about 5%, and under-prediction in labels for energy-efficient dwellings was almost 50%.

Even though the results for primary energy consumption and CO_2 emissions were interesting, the main problem is that electricity consumption does not seem to depend on the energy performance of the dwelling, which was the subject of our investigation. Moreover, the end uses of electricity included in actual and theoretical consumption figures differ to an extent that renders a comparison meaningless (because the theoretical figures exclude appliance usage). On the other hand, a strong relation between gas consumption and the energy label was detected and the end usage numbers for theoretical and actual gas consumption were comparable (with the exception of gas for cooking, which is, as mentioned previously, negligible). Therefore, the scope of the study was narrowed to gas consumption. Gas consumption, for the purpose of this thesis, arises from heating and hot tap water consumption, since the systems which do not utilise gas for heating and hot tap water were removed from the studied samples.

A.3 Is there a difference in the performance gap among the studied samples and throughout the years?

The trends in gas consumption discussed in sub-question A.2 remained very similar in each of the studied samples (Figure 2) despite some differences between the datasets (Table 1, Introduction).

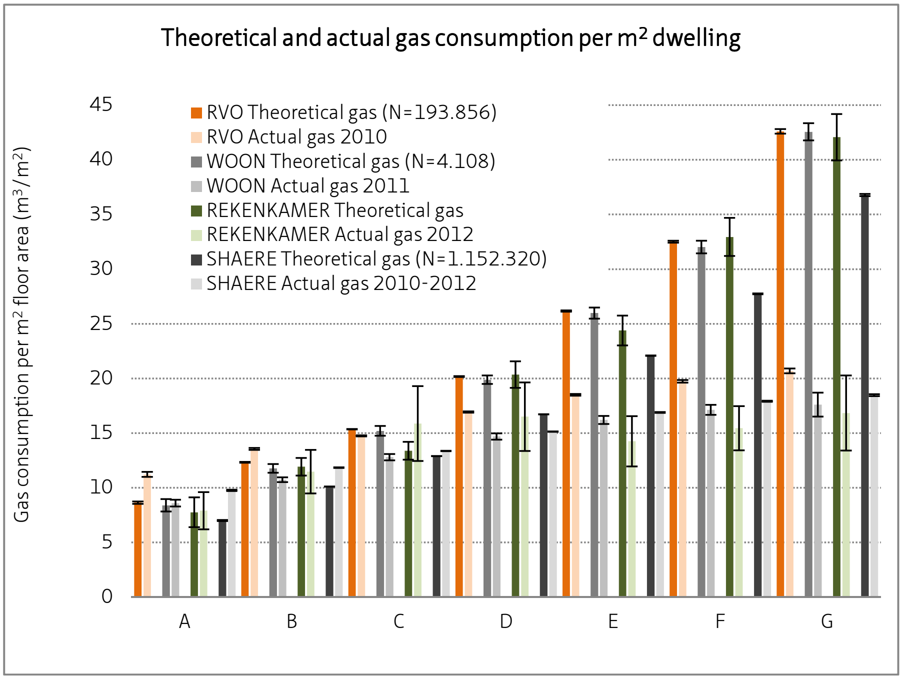


FIGURE 2 Theoretical and actual gas consumption per m^2 dwelling in all samples examined

Comparing the results of different datasets in Figure 2 reveals that actual gas consumption dropped steadily within label categories A, E, F and G from 2010 to 2012. Theoretical gas consumption remained roughly the same in these years, which means that the performance gap increased slightly. The SHAERE sample contains a mix of these years and should therefore, logically, come somewhere in between. In most label categories the difference between the actual gas usage in a label category is statistically significant between RVO and WOON, but not significant between Rekenkamer and others. The reason for this is the small sample size of the Rekenkamer dataset. Also, the Rekenkamer data consist of very similar dwellings from a narrow geographical area (Amsterdam only), which probably affects the results.

In the 'Uncertainties' section of Chapter 5 one can also see that even dwellings that underwent no renovation measures and remained unchanged from the year 2010 to 2012 exhibit a 3.5% decrease in gas use between 2010 and 2012. This decrease could be a consequence of changing household compositions (a smaller number of people per household) or a decrease in the amount of gas used for cooking; however, both of these phenomena occur at a pace slower than 3.5%. Other factors that could be responsible for this decrease could be the changing calorific value of gas and/or the method used to for the calculation of standardised annual consumption.

Apart from this, confidence intervals in Figure 2 also differ in the samples, since the size is very different. They are the largest in the Rekenkamer sample, which contains only 277 records, and the smallest in the SHAERE sample. Theoretical gas consumption remains rather constant in all of the studied samples, except in the SHAERE sample, where it is slightly smaller. In the two most recent samples, SHAERE and Rekenkamer, there is also a noticeable turning point between under-prediction and over-prediction - in label category C, whereas in the other two samples it occurs in label B.

A.4 How does the performance gap correlate with other dwelling properties such as dwelling type, floor area and installation type?

This question is first addressed in the second chapter, but the predictive power of floor area and other dwelling-related parameters for actual and theoretical consumption is further studied in the third and fourth chapter in the regression analyses. Initially, an attempt was made to look at descriptive statistics to find out whether dwelling or installation type would offer a clear explanation for the performance gap. Floor area could potentially have a great influence on the performance gap, since the area of dwellings with label A is around 15m^2 larger than in other categories (Figure 9 of Chapter 2). Therefore, gas and electricity consumption was analysed per dwelling and also per m^2 dwelling. It turned out that it does not affect the performance gap strongly. Detailed descriptive statistics for other parameters were conducted prior to regression analyses of Chapter 4, (section 4.4.1.1 of Chapter 4). It seemed that semi-detached houses have the highest performance gap, followed by flats with a staircase entrance, detached houses and finally, gallery flats (Figure 4 of Chapter 4). The performance gap differed also in dwellings with different installation types. Dwellings with a local heater in the living room (gas stove) had the highest performance gap, followed by those with a combined boiler with $\eta<83\%$, and then each higher efficiency boiler had a smaller performance gap. Each dwelling has a specific combination of these properties (and a variety of others not mentioned here) and they all affect heating consumption to a certain extent so the descriptive results of a single property might provide a distorted impression (for example, that the low performance gap in gallery flats is due to the ventilation system). To find out if this influence differs in theoretical and actual gas consumption, and if so, to what extent, a regression analysis was performed. The regression analysis is described in Chapters 3 and 4: in Chapter 3 we regressed actual and theoretical gas consumption separately and in Chapter 4 we conducted the regression analyses for actual and theoretical consumption as well as the difference between them (referred to as DBTA – difference between theoretical and actual gas consumption). These results will be further discussed in section 6.4.

B Energy reduction targets for the built environment and actual reduction potential of the individual dwelling renovation measures

B.1 Are the current policy targets achievable theoretically as well as actually?

National monitoring showed that between the years 2008 and 2011 (Hezemans et al., 2012) about 950,000 dwellings were made 20-30% more energy efficient. The monitoring was indirect, assuming that two measures corresponded to a 20% energy reduction. It used survey data rather than measured data and analysed relatively small samples which adversely affects representativeness. However, it was the best option at that time since large datasets were not yet available. The assumption about two measures coinciding with a 20% reduction was made because of serious gaps in the existing knowledge of the actual energy savings of renovation measures.

To see how realistic the set targets really were, a scenario analysis was conducted in the third chapter. The baseline scenario was described in the Energy Savings Housing Associations Sector's covenant (Convenant Energiebesparing Corporatiesector, 2008), which aims to reduce gas consumption by 20% by 2018 by improving its dwellings to a B label or at least improving them by two label classes. This was a valid agreement in effect at the time of the study, although it changed later in 2012. A refurbishment scenario was one of the alternatives described in this agreement. Another, more radical, refurbishment scenario involved renovating all dwelling stock to label A. These two scenarios are both rather optimistic, since they assume all dwellings that currently have label C or less will be renovated, but the intention was to identify the maximum savings potential. The two scenarios were tested on both actual and theoretical baseline consumptions (Figure 35). It turned out that by using theoretical gas use as a baseline, the least radical scenario is enough to ensure that the potentials discussed in B.1 are fulfilled. However, if actual gas consumption is used as a baseline, most of these potentials seem unrealistic (an exception is the 10% potential defined by the IDEAL project). This points to the fact that analysts as well as policy makers rely on theoretical gas consumption as a basis for future consumption estimates, which ultimately leads to unrealistic reduction targets and renovation plans.

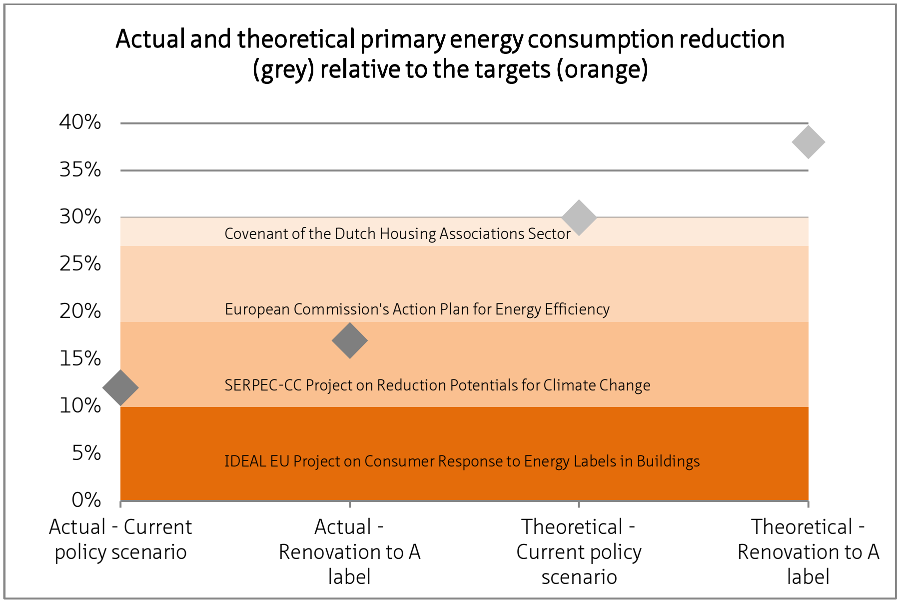


FIGURE 3 Energy saving potential of two policy scenarios based on actual or theoretical consumptions compared to the existing targets

B.2 What are the differences between the theoretical and actual reductions in dwellings where different renovation measures were applied?

To look more broadly into the consequences of the discrepancies, the theoretical reduction in energy use in renovated dwellings was also studied in relation to the actual reduction. Question B.2 is answered in Chapter 5 using large-scale longitudinal data of residential energy performance. In this longitudinal data, dwellings that were renovated are identified and energy consumption before and after the renovation is studied. A reduction of gas use in renovated dwellings was observed in two ways. First, by selecting dwellings that had undergone a change in properties (for example window quality), referred to as sample B, and second, by selecting dwellings that had undergone a change in *only* a *single* dwelling property (sample C).

In the first case (sample B), chances are that the renovated dwelling had undergone other renovations besides the one we noted; in the second sample, all other properties remain constant. With the exception of dwellings that were renovated from label B to label A, these changes always yield a smaller actual savings than expected, with the actual savings being mostly 2 to as much as 6 times lower than the theoretical change. This corresponds well to the performance gaps observed in previous chapters.

Separate renovation measures are best observed in the selection of dwellings that had a change in only a single selected property. In Chapter 5, we actually observed these measures in both sample selections for several reasons. First of all, sample B is significantly larger, possibly making it easier to generate significant results, and second, it is useful to compare the results in order to see whether the two methods result in different conclusions. Below, only the results of sample C are summarised, meaning that only individual measures are compared and analysed.

As seen in the graphics in Chapter 5, most of the renovations are expected to yield a larger reduction than what materialises (see ratios in Table 8 of Chapter 5). Many times the realised savings is about half of what was expected, however in some specific measures it is higher. On average in all renovated dwellings, actual gas reduction is about one-third lower than expected (based on average reductions, Chapter 5, Table 13); however, there are big differences in the reductions of individual measures. To explore which measures achieve the highest actual and theoretical reductions, averages of various measures observed in Chapter 5 are calculated in Table 13. Improvements in the efficiency of gas boilers (space heating and hot tap water) yield the biggest energy reduction, followed by significant improvements to window quality. Improving the ventilation system yields a relatively small reduction compared to other measures, however, it is still much larger than theoretically expected. These are averages and the reductions for specific changes can be found in Chapter 5, Table 9. The measures achieving the most reduction are drastic improvements of window quality and an improvement of the efficiency of the heating and hot tap water system (not the replacement of a local system). Measures that achieve an actual reduction higher than the theoretical reduction seem to consist mostly of very modest improvements of insulation or window quality.

RENOVATION MEASURE	AVERAGE ACTUAL GAS REDUCTION PER DWELLING [M³]	AVERAGE THEORETICAL GAS REDUCTION PER DWELLING [M³]	N	RATIO ACTUAL/ THEORETICAL GAS USE
Ventilation	73	29	4848	2.52
Windows	120	195	4714	0.62
Envelope	105	197		0.53
Heating and hot tap water together	172	279	5792	0.62
Average	147	210	30,749	0.70

TABLE 1 Totals and weighted averages of actual and theoretical gas use reductions in groups where different renovation measures were applied (N>30)

The prediction of reduction seemed better when renovation was made within the group of condensing or within non-condensing boilers. If the boiler changed from non-condensing to condensing, the prediction was poorer. Due to the larger performance gap of mechanical ventilation systems, the actual reduction of improving a natural ventilation system to a mechanical exhaust was higher that theoretically expected. It is questionable whether such dwellings still have a sufficient quality of indoor air after the renovation. These findings point at bad estimations regarding the insulation quality in poorly insulated dwellings and poor predictions regarding poor installation (efficiencies) and mechanical ventilation systems (air flow rates).

C Causes of the differences between actual and theoretical gas consumption

C.1 How much of the variation in actual and theoretical gas use can be explained by dwelling, household and occupant behaviour characteristics?

This question was answered in Chapter 3 as well as in Chapter 4, and each of the chapters used different data and a different approach. In Chapter 3 a regression model and sensitivity analysis of the theoretical calculation method were conducted on the basis of a large dataset (approximately 40,000 records) with 15 variables based on publicly available socioeconomic data. In Chapter 4 the dataset used was less numerous (460 records). The regression model in Chapter 4 was performed on several smaller subsamples of the total and contained more variables (a total of 44) originating from a survey designed to fill the remaining knowledge gaps about occupant behaviour. Regression was made separately for actual and for theoretical gas consumption as the dependent variable, but also for the difference between them (DBTA or performance gap). Furthermore, the regression in Chapter 4 was conducted separately for dwellings with over-predicted and under-predicted consumption. An important difference was also the fact that in Chapter 3 we regressed total dwelling gas consumption and in Chapter 4 we regressed gas consumption per m^2 which reduced the predictive power of floor area. In terms of the predictors used, label class was included as a predictor in Chapter 3 but not in Chapter 4 so it would not take over the predictive power of other, more precise variables related to the thermal performance of dwellings.

Regression in Chapter 3 showed that explaining actual gas consumption or the difference between the actual and theoretical with the publicly available variables yields a relatively low R^2 value (an examination of the existing literature shows that these R values are not low) of 50.5% and 44.0%, respectively, meaning that 50.5% of the variance could be explained by these factors. Since our dataset contained many records, this relatively low explanatory power was thought to be due to the fact that many of the factors that influence actual energy use, such as indoor temperature or the presence of occupants, were not included. This gap was then filled in Chapter 4, including occupant behaviour as well as comfort perception variables, but was of little use, since the total R^2 values were even lower: 23.8% for actual gas use per m^2 as a dependent variable and 40.9% for DBTA per m^2 as a dependent variable. This was probably due to the smaller sample size compared with that used in Chapter 3. Out of these R^2 values, both chapters demonstrated that the majority of explanatory power for the DBTA came from dwelling characteristics (Figure 4). Household and occupancy mattered less, although it was clear that the occupant behaviour data provided by the survey had a non-negligible predictive power for actual gas use per m^2 of 9.1%. The fact that dwelling characteristics dominate the performance gap emphasises the importance of the assumptions made in the calculation method.

In addition to the regression analyses for the total sample, the model in Chapter 4 was constructed from several samples, first with all data and then separately for under- and over-predictions, since the hypothesis was that these two phenomena would be explained by different variables. There was a large difference in the amount of variation that could be explained by all available variables in these two samples. In the under-predicted set of data 19.9% of variation could be explained by occupancy presence patterns, the presence of a programmable thermostat and a water-saving showerhead. On the other hand, in over-predictions as much as 50.8% of variation was accounted for by dwelling and installation type, the age of the building, floor area, and indoor temperature. Furthermore, the level of reported comfort was a significant predictor only in over-predictions.

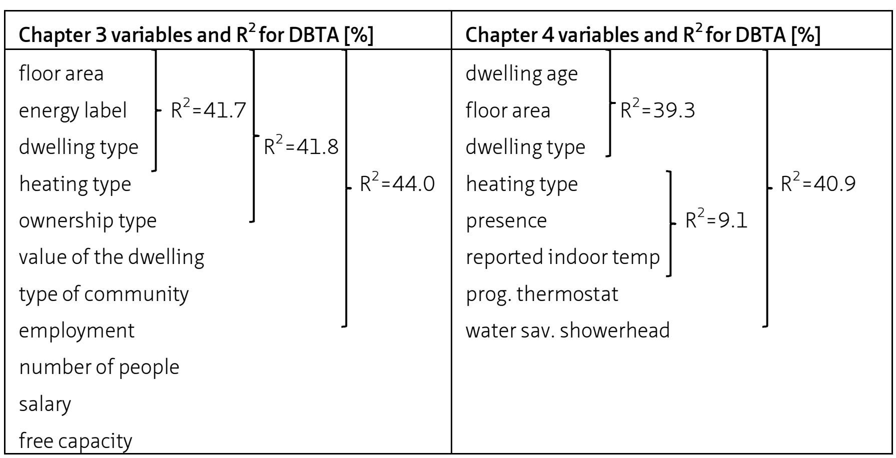


FIGURE 4 Variables included and R^2 values for regression analyses in Chapters 3 and 4

These two chapters emphasise the difficulty of finding the right predictors for actual gas consumption. In the future both survey and sociodemographic data could be combined to maximise results, large samples should be used to ensure statistical significance and certain variables should probably be monitored in order to avoid survey bias.

These include variables such as occupant presence at home, indoor temperature, and ventilation practices, since it seems that respondents might not be adequately aware of their own patterns.

C.2 What is the relation between the performance gap and the normalised assumptions made in calculation models?

Since the regression analyses in Chapter 3 did not cover the effect of variables such as indoor temperature, insulation quality, internal heat load, etc. and this data was not available when the analyses were performed, a sensitivity analysis of the theoretical calculation for certain parameters was made to fill the gap. Six variables (Table 5 of Chapter 3) were modified in the reference buildings of different label classes (A to G) in order to observe the effect on theoretical gas consumption. We found out that an indoor temperature 2.7 degrees higher than assumed by the current method (18 degrees) can explain the performance gap observed in Chapter 2 and that an indoor temperature 5.6 degrees lower than 18 degrees can account for the gap in label G. Both these temperature deviations are in fact quite realistic, since people in well-insulated dwellings probably heat their house more due to the small increase this causes in their monthly bill. Moreover, the installation system itself might encourage the occupants to heat more or less, for example with low temperature floor heating installation in the case of A-labelled dwellings and with a local gas stove in the living room (bedrooms left unheated) in the case of G-labelled dwellings. The average temperature also relates to the heated floor area and could easily explain the more moderate performance gaps in label classes A to C. This is because in the normalisations, all rooms are assumed to be heated, which is more likely to be the case in A to C labelled dwellings (modern dwellings with central heating) than in F or G dwellings (often equipped with local heating where not all rooms are heated). The heat resistance of the construction elements also had a big impact which demonstrates that in cases of poor inspection, the dwelling's estimated consumption could be faulty due to an inaccurate estimate of the insulation. This probably occurs in many old dwellings, where documentation is not available. Moreover, small increments in ventilation rates (up to 40% smaller or larger than the current assumption) can also explain the performance gaps in label classes A to C. The two variables which had a smaller impact and are not likely to be a major cause of the DBTA were the number of occupants and internal heat gains. Results very similar to these were obtained by a study commissioned by Velux in Denmark (Worm, 2012). Theoretical heating consumption of a reference building there was originally 3 times higher than the actual consumption level, however, after correcting the standardised values (indoor temperature, internal heat load of occupants and appliances, building condition, ventilation efficiency and solar radiation) with actual values, the resulting consumption was almost the same as the actual (with a discrepancy of only a few percentage points).

The performance gap of dwellings with certain characteristics also provides insight into how well the calculation method fits reality. Chapter 5 showed that the largest performance gaps appear in dwellings with poor envelope insulation, followed by those with poor window insulation. Similar gaps appeared in installation systems, where lower systems with a lower efficiency had a larger gap. Regarding ventilation, a mechanical systems turned out to have a larger performance gap than a natural system.

C.3 Can a better model be obtained by using actual consumption data?

Besides the previously mentioned exploratory regression analyses in Chapter 4, two other regression models were conducted in order to see whether current theoretical consumption figures could be adapted with the new knowledge about actual gas use. One model was made for under-predicted and one for over-predicted consumption. These models consisted of actual gas use as the dependent variable and theoretical gas consumption plus all other dwelling-related features as predictors. Household and occupant variables were not included, since the idea was to obtain the best possible theoretical consumption figure using only dwelling parameters so that the result would still be comparable with the other dwellings. In the future, this could allow the determination of more accurate dwelling consumption models based only on dwelling parameters and the average actual consumption data. For over-predictions, the model explained 33.8% of variation, with installation and dwelling type being the significant variables (in addition to theoretical gas use). The explained variation was lower than for under-predictions, where it reached 60.0%, probably because the gap itself is much larger in over-predicted dwellings than in under-predicted ones. Significant variables in under-predicted dwellings were floor area, programmable thermostat and water-saving showerhead. The fact that floor area was significant demonstrates that larger dwellings tend to be more under-predicted than smaller ones. This means that, most likely, the whole floor area is heated in smaller dwellings, which is not the case in over-predictions. The B coefficients obtained in these two models were then applied to a different sample (WOON dataset, see Table 1 of the Introduction) to see if a better prediction of theoretical consumption could be obtained by adjusting the current theoretical use with the newly obtained parameters. The new theoretical consumption was indeed much closer to the actual gas use (Table 11, Chapter 4), which proves that this method could be used to obtain a better estimate of theoretical consumption.

§ 6.2 Data quality: Limitations and recommendations

§ 6.2.1 Theoretical consumption data limitations

The most notable limitation of the theoretical consumption data concerns its accuracy. There has been some improvement in the quality of the label certificates in recent years. However, even though the percentage of erroneous certificates, which was 26.7% in 2010, decreased further in 2011 to 16.7%, there is no information available for the most recent years. A short investigation done in the chapter 2 showed that the poor quality of the inspection itself seems to cause the performance gap in dwellings with label A, whereas in other label classes this influence of inspection quality is negligible. A report by Kuindersma and Ruiter (2007) established that the most common mistake was an incorrect estimation of U value, which coincides with some findings of the sensitivity analysis in Chapter 3. Besides the quality of the inspection itself, the assumptions made during the inspection process in order to simplify it could lead to erroneous label estimation, even if the process follows the standardised protocol closely. For example, a dwelling without documentation available will be assumed to have an insulation value typical of its construction year, which could be a faulty assumption. Similar problems could occur with the efficiencies of heating and hot tap water installation system, which can perform differently in practice than assumed in theory. Data on real, measured insulation values or the efficiencies of heating installations could help explain the gap between theoretical and actual gas use but was unfortunately not available during the research for this thesis.

Another disadvantage of the theoretical consumption data from the national registers was a lack of information about the hot tap water systems and exact window insulation quality. For some reason, this data is not available for export from the national register of the certificates, which hinders the analyses made using this data. Fortunately, the SHAERE database used in Chapter 5 did not have this limitation.

In Chapter 5, another aspect of accuracy was discussed. The data analysed there was not registered with governmental authorities but is the so-called 'pre-label' data collected by Aedes. Even though this organisation believes the data to be robust and pre-labels are updated with every renovation, our analyses identified a considerable number of dwellings which deteriorated over the years instead of being improved. Since this cannot possibly occur in practice, such deteriorations are thought to be a consequence of 'administrative' updates of the dwelling, meaning that the first record of the dwelling in the system was faulty. The quality of this dataset has been improving over the years but it is still difficult to say what percentage of improvements are actual renovations of the dwellings.

§ 6.2.2 Actual consumption data limitations

The actual energy data acquired from Statistics Netherlands (CBS) has been collected annually from utility companies since 2009. However, meter readings in The Netherlands are only obligatory every three years, which means that an estimated 10 to 20% of households are charged on the basis of the average consumption of similar households and not on their real meter reading. Even though this does introduce a certain amount of noise to our analyses, we believe that our results are still accurate since we never analyse individual dwellings, where such estimated data might indeed pose a significant problem.

Actual consumption data was always corrected to the number of degree days used in the calculation method to enable a comparison. However, this method might not be perfect, since individual heating practices do not only depend on the outside temperature, but also the time of year when heating use begins. The number of heating degree days are based on an indoor temperature of 18 degrees Celsius, however it is unlikely that the occupants of poorly insulated houses begin using heat at the same time as occupants of better insulated houses. The corrections by degree day may therefore introduce an additional error leading to a slight overestimation of the actual energy use as well.

§ 6.2.3 Limitations related to other data sources

The biggest limitation regarding other data sources was their availability. As seen in Chapter 3, publicly available datasets offer only limited additional information about occupants and the dwellings themselves. In many instances the reliability of the data is also questionable: for example, whether the occupants registered at a certain address during a certain period really live there or are only registered at that address. This uncertainty can be somewhat improved by using survey data about the occupants, which was done in Chapter 4, however in a survey one deals with the bias of the respondents which can also be considerable. When analysing survey data it seemed that the questions about presence at home and heating practices were sometimes too complicated or time-consuming for the occupants to respond to. However, had the questions been further simplified, valuable information would have been lost.

§ 6.2.4 Representativeness of the selected samples

As mentioned previously, a major strength of this thesis, in particular Chapters 2, 3 and 5, lies in the large sample sizes, since similar analyses were previously always based on a smaller number of records. However, representativeness remains one of the limitations, since many more social housing dwellings are labelled in the Netherlands compared to owner-occupied dwellings and these two groups do have some different characteristics, as Chapter 2 showed. However, it seemed that ownership type did not make a difference in the performance gap. Apart from ownership type, the analysis in this thesis was limited to dwellings heated by gas, which prevents the findings from being representative of other heating types. Similarly, dwellings heated by district heating installations and dwellings with shared facilities were all excluded from the samples. These heating and dwelling types should be included in future research, especially since some of these systems are considered theoretically very efficient (heat pumps), but few large-scale studies of actual consumption are available.

§ 6.3 Overall conclusion

The overall research question of the thesis was:

What are the characteristics and consequences of the discrepancies between actual and theoretical heating energy use in Dutch dwellings?

The main conclusion of the thesis is that there is a clear gap between actual and theoretical energy consumption in Dutch dwellings. Lower performing dwellings tend to have a theoretical consumption that is much higher than their actual level, while higher performing dwellings demonstrate the opposite trend. These discrepancies are understandable at the level of individual dwellings and arise due to the standardisations made when calculating theoretical consumption. However, when broadened to the level of the dwelling stock such discrepancies are misleading and can lead to inaccurate policy reduction targets and send the wrong signals to several stakeholders (local governments, the construction industry, renters and buyers, etc.).

The causes of the discrepancies can be partly explained by dwelling features, meaning that the calculation model does not accurately represent reality. However, part of the discrepancy originates in the behaviour of energy users and this part is difficult to quantify statistically. The results seem to indicate that under-prediction is more difficult to explain and therefore probably more dependent on occupant practices than on the accuracy of the standardisation model. Over-predictions on the other hand seem to have a lot to do with the fact that installation systems and the dwelling itself perform differently than expected. A methodological improvement seems to be more appropriate for the over-predicted cases while at the same time tackling the problem that occupants of these dwellings are likely to feel cold. For under-predictions on the other hand, changes to the methodology would mean accepting that a higher heating intensity is inevitable in efficient dwellings. While this should be further researched in the future, behaviour incentives that would encourage people to heat their homes more wisely and not waste energy could still be successful.

§ 6.4 Recommendations

Several recommendations can be based on the outcomes of this thesis. With regard to the actual consumption of labelled dwellings, this thesis concluded that dwellings with a lower label class consume more gas than dwellings with a higher label class. This is a good effect for the label methodology, since it means that it effectively assigns a thermal quality certificate to a dwelling. However, in addition to the label class, the label certificate also depicts a dwelling's theoretical energy consumption. Dwelling stock averages comparing theoretical and actual gas consumption across different thermal performance levels showed that lower performing dwellings consume about fifty percent less than the figure displayed on the certificates and that higher performing dwellings consume about a quarter more than the amount displayed.

The gap is a consequence of a discrepancy between the two entities, theoretical and actual gas consumption. Theoretical gas use is determined by the regulatory calculation model, and actual gas consumption is influenced by the dwelling's occupants. Therefore, the recommendations we made refer either to insights into the calculation method, or into occupant behaviours that explain the performance gap. As mentioned in the Introduction, theoretical gas consumption is a model, and as such it is an imperfect representation of reality. By the improved understanding of reality afforded by the work done in this thesis on household characteristics, occupant behaviour and perceptions of comfort, the author reflected on the quality of the calculation method which considers all these parameters to be standard. If one can better understand the behaviours and varieties of the occupants that inhabit certain types of dwellings, better predictions can be made about the theoretical energy consumption of those same dwellings.

Even though it is clear that on the level of an individual dwelling theoretical energy consumption cannot correspond with what is actually used, this thesis showed that large discrepancies between the two at a global level, the dwelling stock level, have detrimental consequences for policy, especially considering the cost effectiveness of possible renovation measures.

§ 6.4.1 Recommendations for policy

In terms of policy, on the basis of this thesis work, recommendations can be made regarding the energy label methodology and regarding the potential for reductions in gas usage in Dutch dwelling stock.

Energy label methodology

Starting with the energy label calculation, a revision of several standardised factors used in the calculation method should be made. This calculation method was developed before actual energy consumption data was available on a large scale and was therefore not validated on a dwelling stock scale. Since such validation is now possible it should be carried out. Based on this work, it seems that the theoretical efficiencies of boilers, especially those with a lower efficiency, might not be representative of reality. Similarly, the air flow rates utilised for different ventilation systems seem far removed from reality. Furthermore, the quality of insulation should not be based on the construction year of the dwelling, but instead on a simple measurement of its conductivity. Although critics might say that such tools are expensive, new techniques have recently been developed that enable a quick and reliable determination of conductivity (Rasooli et al., 2014). This thesis also concluded that the way in which the dwelling is used, in particular the amount of surface area that is heated, depends strongly on its thermal performance. Therefore, correction factors should be incorporated into the label calculation method, similar to the existing factors for the type of e dwelling (corrections for surface heat loss). This is particularly urgent in order to reduce the over-predictions in lower label classes.

Moreover, the energy labels that are issued should be accurate and reliable, meaning that more attention should be paid to annual re-inspections of a sample number of dwellings. Such a control is necessary to motivate the qualified inspectors to issue trustworthy, high-quality labels instead of hastily produced approximations. However, the latest policy developments have resulted in the cancellation of the mandatory inspection as a part of the development of what is known as a 'simplified label'. This poses further questions about the quality of the energy label in the future.

In terms of the current form of the label certificate, the question remains whether it makes sense to indicate the theoretical amount of energy consumed on the label, as has been done in the Netherlands so far. This seems to cause confusion since it is totally unrepresentative of reality. The label seems to correctly estimate the average thermal quality of the dwelling but cannot predict actual energy consumption. On the other hand, the label calculation is easy to use and can be, as shown in the thesis, a very valuable tool for tracking the energy efficiency of the dwelling stock. Since the accuracy of theoretical gas and electricity usage calculations can easily be improved, it would be a pity to miss the opportunity to do so.

Last but not least, care should be taken to ensure that the software used for label calculation does not allow illogical input. In the current version, it is for example possible to input a combi boiler (meaning one that is also used for space heating) for hot tap water and a separate boiler for heating, which in practice is not possible.

Usage reduction potential of the dwellings

In this thesis, usage reduction potential was evaluated in two ways: first, globally, at the dwelling stock level and later also at the level of specific renovation measures. Regarding the global reduction potential, it is essential that actual consumption values are taken into account when formulating targets. In 2008, a goal was set to reduce gas consumption by 20% by 2018 in the social housing sector only, by improving the dwellings in two label classes or until label B was reached. The goal for the built environment was updated in 2012; however at the time Chapter 2 was written, the 2008 goal was still valid. A scenario study showed that while the target might be achieved if theoretical gas consumption was used as the basis for the reduction calculation, it was far beyond reach if actual gas usage was taken as the baseline. Therefore, this thesis highlights the importance of considering actual consumption figures when formulating policy targets in order to invest effort in a realistic roadmap.

In addition to the recommendation regarding the potential for energy reduction in the dwelling stock, the thesis also provided conclusions about renovation measures. It clearly demonstrated the great difference between basing the reduction potential calculations on theoretical or on actual energy use and emphasised that the use of actual consumption figures should be encouraged. Therefore the insights offered regarding the potential for actual energy reduction are very useful for policy makers, since such evaluations do not yet exist on a large scale. It was shown that, on average, a single renovation measure brings about an 11.6% reduction in actual gas consumption while a 16.9% reduction was expected. On average, the highest yield is achieved when replacing a heating system, followed by the improvement of windows, the building's envelope, and finally, a ventilation system. Policies should be developed according to these findings, encouraging the measures that are most effective in reality and not just in theory. While there must be opportunities for innovative technologies (for example, heat pumps), it is also important that the real performance of these, too, is closely followed in practice, since this thesis has shown the tendency to deviate from theoretical performance. By doing so, the models can be improved to better fit reality.

§ 6.4.2 Recommendations for practical application

It was demonstrated that when examining renovation possibilities, the theoretical consumption level provided by the energy label methodology (and still widely used by engineers and consultants) is not a good baseline and leads to erroneous results. The savings potentials included in renovation scenarios should take actual consumption figures into account.

This research also showed that while occupants in lower performing dwellings seem to heat their homes less than expected, the higher performing dwellings are heated above the predicted amount. When regressing the over-predicted cases in Chapter 3, it was established that the feeling of cold was a significant predictor for the performance gap This indicates that despite the fact that lower performing dwellings perform much better than expected, renovations should still be undertaken since they are the best way of improving people's comfort and preventing excessive energy use.

On the other hand, significant predictors for the performance gap in higher performing dwellings were a programmable thermostat and water-saving showerheads. Dwellings with a manual thermostat tend to consume less actual gas and have a lower performance gap. Similar results are found in dwellings where water-saving showerheads are installed. Such simple measures should therefore be implemented in all dwellings with good thermal performance.

§ 6.5 Recommendations for future research

Based on the outcomes of this thesis, two lines of research could contribute to more accurate predictions of theoretical energy consumption and thereby decrease the performance gap. The first is improving the standardised values used in the calculation method, the second is using individual gas consumption data at the dwelling stock level to make better predictions.

Regarding the first line, this thesis concluded that several standardised factors are responsible for the discrepancy between theoretical and actual gas consumption. By finding out the real values of these parameters in different performance categories and feeding them into the calculation model, the performance gap would be greatly reduced. These parameters are: R and U values, heated floor area and indoor temperature, efficiency of the heating system, air flow rates and, to a lesser extent, internal heat generation (number of appliances and people). A validation study of all these parameters in dwellings with different thermal performance should be carried out in the future. As we concluded earlier with regard to heated floor area and indoor temperature, surveys and publicly available socioeconomic data is not sufficient for validation, therefore monitoring of occupants practices should be carried out in real time. For validations of other parameters, detailed inspections should be carried out.

The other option for reducing the gap, would be applying correction factors based on the actual consumption data that is now available. Depending on basic dwelling properties such as dwelling type, heating installation, insulation and the presence of a ventilation system, several correction factors could be calculated on the basis of the actual gas consumption in such dwellings. The same was done in Chapter 4 and was shown to be effective at calculating a more realistic level of theoretical consumption. As was already shown in this chapter, a positive performance gap has completely different causes than a negative one and a good way to research the gap in the future would be to analyse different thermal performance (label) classes separately. With the larger datasets that are becoming more readily available this will be increasingly more feasible.

As mentioned in the Introduction, the performance gap is a consequence of poor modelling of the actual situation, and therefore a strategy targeted at 'improving reality' could reduce the performance gap as well. This is a realistic solution to the performance gap in high efficiency dwellings, where a rebound effect and the high indoor temperature play a large role in the under-predictions. Further research should explore the possibilities of reducing the indoor temperature of selected high efficiency dwellings. At the same time, it is unrealistic and undesirable to expect the occupants of low efficiency dwellings to consume as much as the model has predicted, meaning that the solution to the gap reduction on this side lies purely in better predictions by the model described above.

The reliability of actual consumption data should also be further studied. Except for a few estimates, there is no empirical research available into how many of the consumption figures are based on real meter readings and how many are estimated. This thesis showed that on average, actual energy consumption has dropped slightly, even in dwellings that have not been renovated in the period of the past four years; however, the reasons for this were not investigated. Further studies should also look into how accurately the degree day method corrects for heating intensity, especially between the categories of well and poorly insulated dwellings. Moreover, there are many more uncertainties about specifications of older systems with lower levels of performance (such as heating installations or building envelope insulation) and these lead to further miscalculations. The actual performance of older dwellings, which is currently determined on the basis of the construction year, needs to be more carefully studied in order to better predict the reductions generated by renovations.

However, even a perfect calculation method cannot reduce the inaccuracies that occur due to poor inspection of the dwelling; therefore more attention should be given to the accuracy of the inspection phase in the form of re-inspections, the improved training of experts, etc.

Regarding the data types used in this thesis, further research should be done using cross-sectional data, which is becoming more and more abundant. Cross-sectional data gives decent reduction estimations for deeper renovations, however, for the small changes that are the most common it has proved to be inaccurate. Future research should be carried out using longitudinal data looking at combined renovation measures and employing advanced statistical methods which allow for control variables and allow the data to be used in continuous form as well (for example, for the insulation value of walls and windows).

§ 6.6 Final remarks

This thesis demonstrated that research on the relationship between policy instruments and their effects is crucial to ensure the effectiveness of these tools and their continuing improvement. Theoretical models, such as energy labelling, are often used to support policy decisions. As has been shown, such models do not always provide results that correspond to reality, and in the case of dwellings a big reason for this is a disregard of the user, who seems to adapt to the thermal quality of the house itself. However, as was demonstrated, there is a clear need for a more accurate estimation of consumption on the broader level of dwelling stock in order to enhance the effectiveness of current renovation policies. Moreover, the thesis showed that better estimation is feasible, and that, using the current knowledge and data available, there are few reasons not to reduce the performance gap and more accurately predict the energy consumption of Dutch dwellings.

§ 6.7 References

Convenant Energiebesparing Corporatiesector, October 2008, accessed on 9th April 2012 on http://www.aedesnet.nl/binaries/downloads/2008/10/20081009-convenant-energiebesparing-corporatiesect.pdf

Kuindersma, P., Ruiter, C.J.W. 1007. Eindrapportage Woonkwaliteit Binnenmilieu in Nieuwbouwwoningen, VROM-Inspectie.

Rasooli, A., Itard, L., Infante Ferreira, C. 2016. A response factor based method fort he rapid in-situ determination of thermal resistance in existing buildings, Energy and Buildings, 2016.

Worm, A. 2012. Bolig for livet - Energi og indeklima i måleperioden 2010-2011, Velux, January 2012.

Biographical Note

I was born in 1985 in Celje, Slovenia. Having always had a green thumb and being interested in everything related to nature I chose to study Environmental Sciences at the University of Nova Gorica. I always enjoyed travelling and when an opportunity arose to spend a year as an Erasmus student at Delft University of Technology I did not think twice. Upon completing all the courses I decided to stay at the TU Delft to work on my diploma thesis on the topic of life cycle assessment of flat roofs at the OTB Research Institute. Little did I know that Delft would eventually become my second home. After graduating in 2009 I returned to Slovenia to work as a consultant, however, the job there failed to challenge me. This motivated me to do an internship in the European Commission in Brussels in the cabinet of the Commissioner for Environment. After that, I returned to Delft and started a PhD at the already familiar OTB Research Institute. I completed my dissertation on the topic of theoretical and actual heating consumption in Dutch dwellings in 2015.

Residential buildings are one of the key target sectors for energy and CO_2 reduction. Research on the relationship between policy instruments and their effects is crucial for the continuous improvement of these tools. This thesis focuses on The Energy Performance of Buildings Directive (EPBD) as an EU flagship policy for technical improvements of the existing dwelling stock. At the core of the directive, an energy certificate is prescribed for all existing dwellings and is required to be present at moments of sale or rent. The certificate includes a theoretical consumption of each individual dwelling. However, theoretical models do not always correspond to reality with the desired accuracy. This large scale study explores the relation between the consumption assumed by the label certificate and the real consumption of the dwellings. It demonstrates that low performing dwellings have a theoretical consumption roughly two times the actual, while well performing dwellings consume one third less. These discrepancies are shown to have adverse effects on policy targets. Therefore, the thesis quantifies the causes of the discrepancies, looking into dwelling and household, as well as behavioural factors.

There is a clear need for a more precise estimation of heating consumption on a broader, dwelling stock level in order to enhance the effectiveness of the current renovation policies. The thesis showed that using the current knowledge along with the growing amount of available data, there is enough motivation to reduce the performance gap by improving the predictions of actual dwelling consumption.

Delft University of Technology,
Faculty of Architecture and the Built Environment,
OTB – Research for the Built Environment