Energy Policy 97 (2016) 82–92

Contents lists available at ScienceDirect

Energy Policy

journal homepage: www.elsevier.com/locate/enpol

Actual heating energy savings in thermally renovated Dutch dwellings

Daša Majcen*, Laure Itard, Henk Visscher

OTB – Research for the Built Environment, Faculty of Architecture and the Built Environment, Delft University of Technology, Julianalaan 134, 2628BL Delft, The Netherlands

HIGHLIGHTS

- Performance gap is lower in more efficient buildings.
- Replacements of gas boilers – the most energy reduction among renovation measures.
- Replacing the ventilation system yields a much larger reduction than expected.
- How well are the standard values of the calculation methods defined?
- Provide large public building performance databases including actual use data.

ARTICLE INFO

Article history:
Received 2 March 2016
Received in revised form
8 July 2016
Accepted 9 July 2016
Available online 20 July 2016

Keywords:
Theoretical heating consumption
Actual heating consumption
Performance gap
Renovation measure
Energy saving

ABSTRACT

The register of the Dutch social housing stock was analysed, containing 300.000 dwellings, renovated between 2010 and 2013. The main objective was twofold: to evaluate the performance gap in these dwellings before and after the renovation and to establish what renovation measures achieve the highest reduction of consumption, particularly in practice (actual savings). The results showed large performance gaps in dwellings with 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. Regarding the actual effectiveness of renovation measures, replacement of old gas boilers with more efficient ones yields the highest energy reduction, followed by deep improvements of windows. Installing mechanical ventilation yields a small reduction compared to other measures, but still much larger than theoretically expected. The paper shows once more that the calculation method currently in use cannot be considered accurate if compared to actual consumption. The study demonstrated that unrealistic theoretical efficiencies of heating systems and insulation values are causing a part of the performance gap. Nowadays, large datasets of buildings thermal performance and actual consumption offer an opportunity to improve these misconceptions.

1. Introduction

1.1. 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. 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 and to reflect on the adopted policies and apply amendments where necessary. 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. Previously published research

* Corresponding author. Present address: University of Geneva, Institute for Environmental Sciences and Forel Institute Office B611, Boulevard Carl-Vogt 66, 1205 Genève, Switzerland.
E-mail address: dasa.majcen@unige.ch (D. Majcen).

http://dx.doi.org/10.1016/j.enpol.2016.07.015

conducted on the mentioned data analysed the renovation pace between the years 2010 and 2013 (Filippidou et al., 2015, 2016). This paper builds upon these findings by observing theoretical and actual heating energy consumption before and after thermal renovation, which allows to compare the performance gap (difference between theoretical and actual gas consumption) before and after renovation, 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 to pinpoint the most effective renovation measure in terms of achieved savings. The outcomes obtained by using different analysis methods are compared, making the analysis robust 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. The later three are aggregated into one variable referred to as 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 pre-label is a complete thermal recording of the dwelling, including all dwellings energy labels, theoretical heating demand and dwelling properties, but not necessarily registered as an official label certificate. Label registration is the act of submitting the pre-label data to the competent authority, thereby obtaining an official label certificate. Energy index is calculated according to the national standards on the basis of total primary energy demand, 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 theoretical and actual gas consumption of a dwelling or group of dwellings.

1.2. State of the art

Filippidou et al. (2015, 2016) describes the annual frequencies of 7 renovation measures in The Netherlands. According to the author, 16.8% of the dwellings have improved their label class between years 2010 and 2013. Another study analyses the Dutch dwelling stock and the measures taken based on a survey of about 4000 representative dwellings (Tigchelaar and Leidelmeijer, 2013). The results show that 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. (2016), 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 but used cross sectional and not longitudinal data. The third study is a national monitoring carried out in The Netherlands (Hezemans et al., 2012) on the basis of a survey – managers of housing corporations were asked to report on to the changes in the stock retroactively. An assumption was made that implementing two saving measures (insulation of an envelope part or a replacement of heat 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 instead of 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 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, 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 rely 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, 2013b), Tigchelaar and Leidelmeijer (2013).

Up until recently, international research papers which evaluated operational energy consumption based their analysis on purely on theoretical heating energy (Adalberth, 1997; Winther and Hestnes, 1999; Dodoo et al., 2010; Thormark, 2002). As shown, theoretical consumption can diverge from the actual consumption by as much as 50% less or 30% more. In more recent years, the focus on actual consumption is increasing and studies of the performance gap are starting to appear all over Europe. They indicate results, similar to the Dutch ones – an overprediction of inefficient (Tigchelaar et al., 2011; Cayre et al., 2011; Hens et al., 2010) and underprediction (Haas and Biermayr, 2000; Branco et al., 2004; Marchio and Rabl, 1991) of efficient dwellings. The phenomenon of underestimated theoretical consumption is also referred to as 'rebound effect' (Berkhout et al., 2000). This means that an efficient technology (such as thermally renovated dwellings) cut energy bills but thereby encourage increased consumption. Next to the rebound effect, the term 'pre-bound effect' can also be found in the literature, describing the overprediction of consumption in old, inefficient dwellings (Sunikka-Blank and Galvin, 2012). A comprehensive overview of the literature can be found in Majcen et al. (2013a). International studies of adverse effects of the gap are presented in Majcen et al. (2013b), and a review of the international paper that examine the causes of the gap was done by Majcen et al. (2015).

Despite the multitude of monitoring studies and studies on the performance gap, there seems to be a lack of studies analysing the efficiency of 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. The objective of this paper is therefore to fill this data gap by studying actual consumption of dwellings on a large scale.

1.3. 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. This enables a reflection on the usability of SHAERE dataset and provide guidelines for future setup of data registers in different European countries.

The core of the paper is the improvement of thermal performance of the dwellings and the actual energy savings following thermal renovations, in the results section only the dwellings which have undergone changes are selected and the theoretical as well as actual reduction of energy consumption before and after renovation is analysed.

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 the trends noted regarding the effectiveness of different thermal renovation measures, the performance gap and the validation of the calculation method are summarised.

2. Methodology

2.1. Dataset properties

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. The dataset contained about one million dwellings in each of the four years between 2010 and 2013, thereby offering a great opportunity to gain insight into the dynamics of energy performance of the stock. The SHAERE register is a raw, full export of the entire energy performance certificate calculation according to the Dutch standard (ISSO 82.3, 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, 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, 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 are available pre-and post-renovation (also called longitudinal data), which greatly improves the variance between groups due to the changes in conditions we do not control for (different household and occupant properties in different groups etc.). The dwelling information available in this paper was also more detailed than in previous studies, including detailed information on the quality of insulation and hot tap water installation.

2.2. Variable extraction

From the original 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. Using R value for heat transfer coefficient of wall elements and U value for windows is not common everywhere around the world, but is typical for the Netherlands, therefore the same terminology was used in this paper.

$$R_{average} = (A_1 + A_2 + (\ldots) + A_n) / \left(\frac{A_1}{R_1} + \frac{A_2}{R_2} + (\ldots) + \frac{A_n}{R_n} \right)$$

$$U_{average} = \left(\frac{U_1 \cdot A_1 + U_2 \cdot A_2 + (\ldots) + U_n \cdot A_n}{A_1 + A_2 + (\ldots) + A_n} \right)$$

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 ($\eta = 65\%$) is a local gas heater, where local means that the heater – a gas stove – is situated in one place 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 lowest efficiency ($65\% < \eta < 83\%$), regarded as $\eta < 83\%$, this kind of heater is non-condensing. A conventional

Table 1
Categories of insulation values used.

R envelope [K m²/W]	Categorised R value	U window [W/K m²]	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

condensing boiler has a $\eta > 83\%$, and there are several high(er) efficiency condensing boilers with efficiencies of 90%, 94% and 96%, referred to as $\eta > 90\%$, $\eta > 94\%$ and $\eta > 96\%$. 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 ($\eta < 83\%$), improved ($83\% < \eta < 90\%$) and high efficiency condensing boiler ($\eta > 90\%$).

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, demand controlled decentralised mechanical exhaust ventilation.

2.3. 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. Sample selection is summarised also in Table 2.

Table 2
Sample selection and the size at each stage.

Stage of selection	Size (N)
Initial SHAERE sample	5,205,979
Removing duplicate records	4,612,020
Removing outliers (floor area)	4,606,749
Removing non-gas based and collective systems	3,729,256
Removing non-independent dwellings	3,728,143
Removing records where actual consumption is not yet available	2,726,600

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 15 m² and above 500 m² 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 utilise either district heating (4%) or electricity (6%) for hot tap water and about 7% of the space heating installations are electrical systems. Besides the variable about the installation type, information was available whether the whole system of hot tap water and space heating was collective or individual (8,3% and 16,6% of the total sample, respectively). District heating systems had to be removed due to the inaccurate 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 the Statistics Netherland is collected from the energy companies, which base it on the annual metre readings done by the occupants. The data are 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, 2013b, 2015). The actual gas consumptions corresponded to the climatic year regarding the degree days, therefore corrections were applied to compare these consumption values with the theoretical ones (Majcen et al., 2013b).

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 (longitudinal data). The number of pre-label certificates from different years is gathered in Table 3.

Dwellings with at least two pre-labels (sum of row 4 till 7 in Table 3) 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. For instance, when studying changes in space heating and hot tap water, all dwellings with a replacement of 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

Table 3
Number of dwellings having a pre-label in a given year.

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

	Label changes	Space heating and hot tap water	Ventilation	U value windows	R value envelope
Deteriorations	5%	2%	1%	6%	10%
No change	78%	87%	95%	77%	74%
Improvements	17%	12%	4%	18%	15%
Total sample	109,278	79,241	25,783	116,025	96,688

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

Year	2010	2011	2012	2013
Average actual gas use [m^3/year]	1,054*	1,034*	1,017*	1,016*
Average theoretical gas use [m^3/year]	1,113	1,113	1,113	1,113
Gas reduction relative to 2010 [m^3/%]	/	20 [1.9]	37 [3.5]	38 [3.6]

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

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. 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.

2.4. Uncertainties

In the section before, we showed that deteriorations of properties were observed in a small part of the sample (1–10%) 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 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 (cut-off points for outliers being 15 and 6,000 m^3 gas). 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. From this data alone it is difficult to say what is causing this autonomous decrease, but the decrease of gas used for cooking and the decrease of the number of occupants could partly explain the reduction in demand. Further reasons could be inaccurate degree day corrections (only correcting for temperature and not solar gains), and inaccuracies in the corrections of energy quality of gas. Such a high degree of autonomous and continuous annual reduction is surprising and should be further researched in the future.

3. Results

3.1. Change in only space heating and hot tap water

This section shows the actual and theoretical reduction of dwellings where space heating and hot tap water installation was replaced. 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. We therefore show the results for dwellings which had a replacement in both, heating and hot tap water systems. This way the amount of results is manageable and the most interesting combinations are studied. To ensure statistical significance, groups with less than 30 cases are omitted from the figures.

In this sample of 30,749 cases, heating and hot tap water installation was replaced according to the information in SHAERE database.

A large performance gap before the renovation does not signify a large performance gap after the renovation in this Fig. 1. Visually, there does not seem to be a correlation between the size of the performance gap before and after the renovation. It does seem that dwellings are better predicted after renovation than before, meaning that theoretically better performing installations are better predicted. It also seems that replacements within the category of non-condensing boilers (all efficiencies below 90%) are reasonably well predicted as well as replacements within the categories of condensing boilers (all efficiencies above 90%).

3.2. Change in ventilation only

This section shows the actual and theoretical reduction of dwellings which benefited from a replacement of in the ventilation installation. We excluded the groups of dwellings which contained less than 30 cases to ensure statistical significance.

Fig. 2 seems to suggest the savings when replacing a natural ventilation system with mechanical exhaust ventilation to be at least three times as high as expected. The theoretical gas consumption barely reduces after the renovation. When looking at the

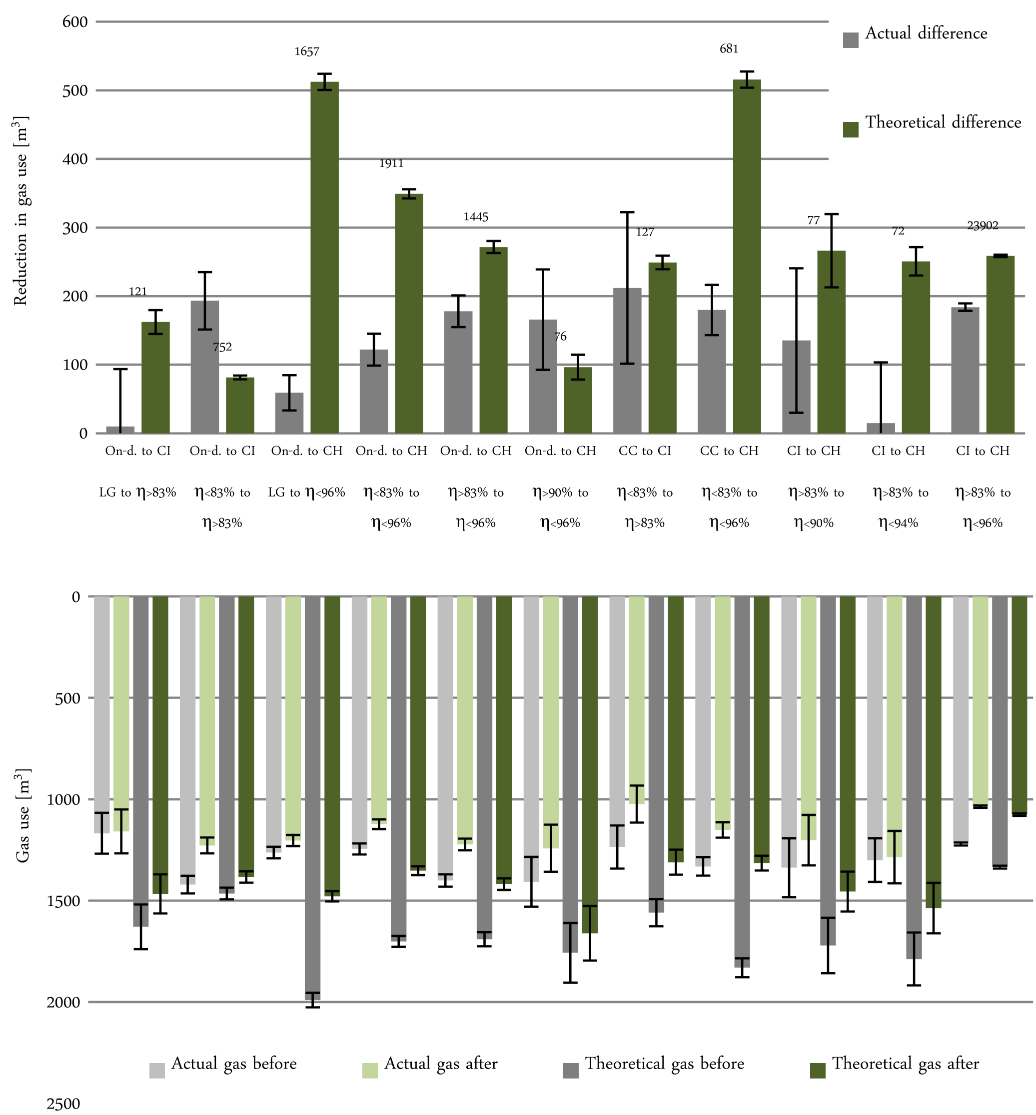

Fig. 1. Actual and theoretical consumption difference before and after renovation in dwellings with changed hot tap water and heating installation system ($N > 30$). On-d. = on-demand tankless boiler, CI/CC/CH = combined conventional/improved/high efficiency boiler. Actual reduction of the first and before last column is below the background reduction.

calculation method this makes sense, since mechanical and natural ventilation both use exactly the same 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. 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 mechanical exhaust is not interpreted by the inspectors correctly due to its complexity leading to frequent input errors.

3.3. Changes in window quality only

This section shows the actual and theoretical reduction of dwellings which had an improvement in the window quality. We excluded the groups of dwellings which contained less than 30 cases to ensure statistical significance. In this section insulation quality as described in Table 1 are used.

Fig. 3 reveals that dwellings that had a drastic change in window quality (U8 to U2, U7 to U1) tend to have an actual gas reduction lower than the theoretical. 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). It is questionable whether such small improvements are real changes or administrative corrections, since one would imagine that in most cases when windows are replaced, the improvement is bigger. However, it could also be that only one or a few windows were replaced. One also needs to keep in mind the background gas reduction, since in some cases the actual gas reduction seems to be smaller than that (for example U4 to U1). Looking at the absolute gas consumption before and after renovation one can see (bottom graph in Fig. 3) some overpredictions.

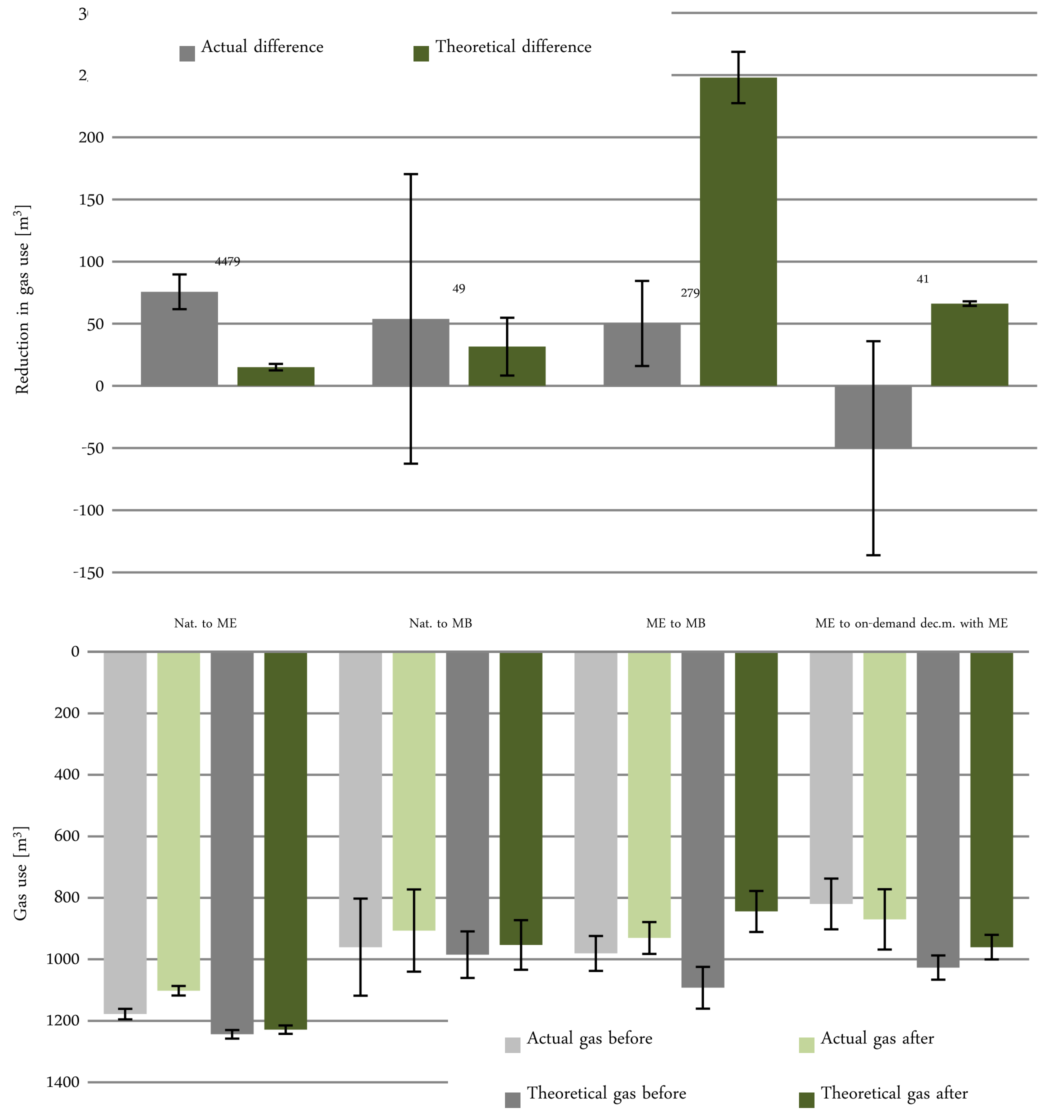

Fig. 2. Actual and theoretical consumption difference before and after renovation in dwellings with changed ventilation system (N > 30).

3.4. Changes in envelope quality only

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.

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, overprediction occurs in R5 to R10 and underprediction in R1 to R4.

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 for 1 step (1 category), and changes for one or two categories are the most numerous. This might seem drastic, but a short calculation shows, that a dwelling with envelope of 300 m^2 and an R value of 0.4 insulates the roof (10% of total area) with R=2.5, the new R value is 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 windows – small changes are well predicted and actual reduction is close or surpassing the theoretical whereas deeper changes result in a lower actual reduction. The better insulated the dwelling is, 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 Fig. 4).

3.5. 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 6 taking into account all the groups containing more than 30 records. Interestingly, the measure which achieves the largest actual cumulative as well as individual saving is the replacement of heat and hot tap water system. Envelope improvement is in the second place and ventilation system replacement the last. The most remarkable considering individual measures, is the reduction in dwellings with an improved ventilation systems achieving a 2.5 times higher reduction than predicted.

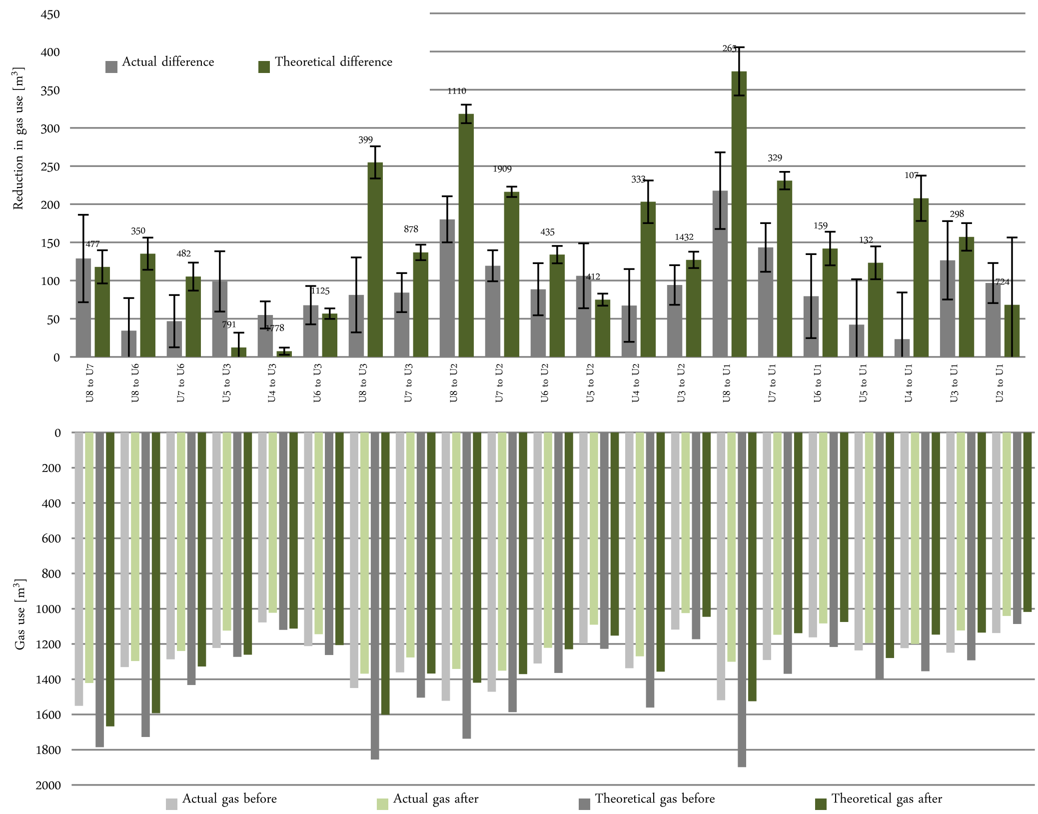

Fig. 3. Actual and theoretical difference between the first and second pre-label in dwellings with changed windows (U-value). Confidence intervals are omitted in the bottom graphic for better readability. Actual reduction of U4 to U1 is below the background reduction.

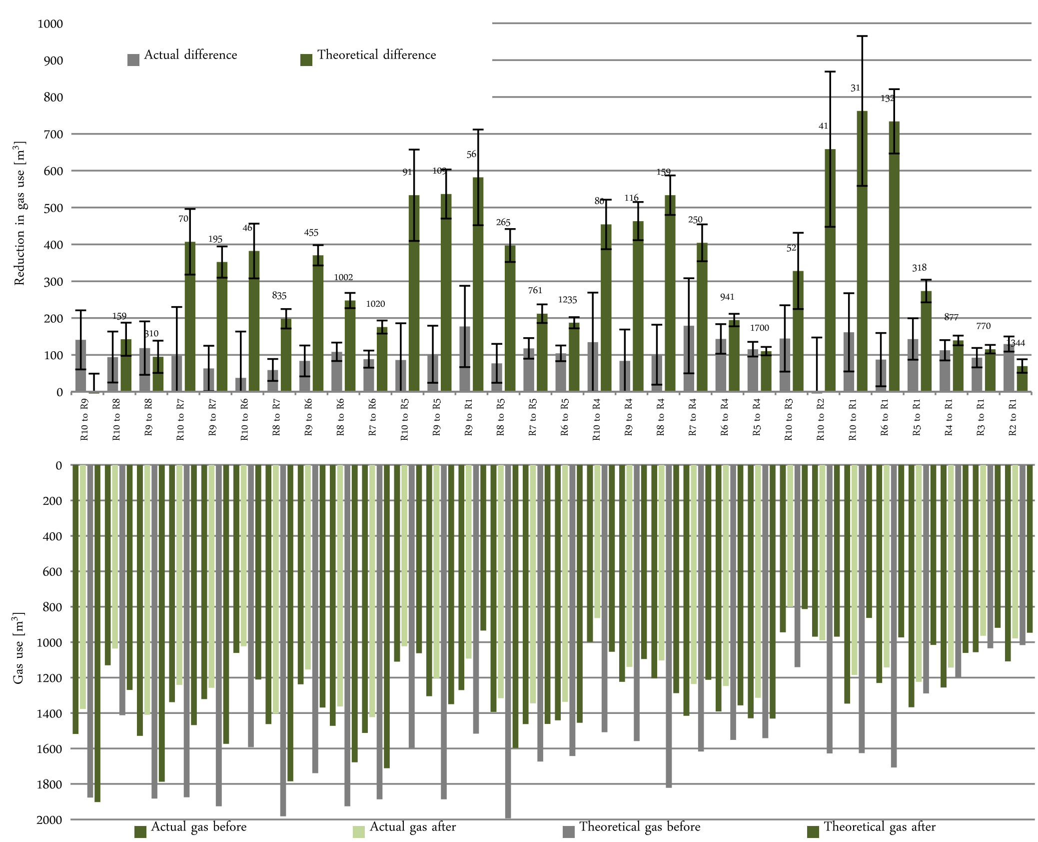

Fig. 4. Actual and theoretical difference between the first and second pre-label in dwellings with changed envelope insulation (*r*-value). Confidence intervals are omitted in the bottom graphic for better readability.

Table 6
Totals and averages of actual and theoretical gas reduction in different measure groups using sample C – unique measure (groups with N > 30).

Renovation measure	Cumulative saving (total sample)		Individual saving (per dwelling)			N
	% of total actual gas reduction	% of total the gas reduction	Average actual red. [m^3]	Average the. red. [m^3]	Ratio actual/theoretical gas use	
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 [m^3]	9,367,264	14,622,945	131	188	0.70	72,376

Table 7
Actual consumption reduction per dwelling of various single renovation measures.

	Act. [m^3]	N	Ratio		Act. [m^3]	N	Ratio
U8 to U1	218	265	0.6	R8 to R4	101	159	0.2
$\eta < 83\%$ to $\eta > 83\%$ and CC to CI	212	127	0.9	U8 to U4	99	111	0.4
$\eta < 83\%$ to $\eta > 83\%$ and On-d. to CI	193	752	2.4	U2 to U1	97	724	1.4
$\eta > 83\%$ to $\eta < 96\%$ and CI to CH	184	23,902	0.7	R3 to R1	93	770	0.8
U8 to U2	180	1,110	0.6	R6 to R1	87	132	0.1
$\eta < 83\%$ to $\eta < 96\%$ and CC to CH	180	681	0.3	U8 to U3	81	399	0.3
$\eta > 83\%$ to $\eta < 96\%$ and On-d. to CH	178	1,445	0.7	U6 to U1	80	159	0.6
$\eta > 90\%$ to $\eta < 96\%$ and On-d. to CH	166	76	1.7	R8 to R5	77	265	0.2
U7 to U1	143	329	0.6	Natural to m. exhaust	76	4,479	5.0
R5 to R1	143	318	0.5	LG to $\eta < 96\%$ and On-d. to CH	59	1,657	0.1
$\eta > 83\%$ to $\eta < 90\%$ and CI to CH	135	77	0.5	R8 to R7	59	835	0.3
U8 to U5	133	253	0.5	Natural to m. balance	54	49	1.7
R2 to R1	130	1,344	1.9	M. exhaust to m. balance	50	279	0.2
U8 to U7	129	477	1.1	U5 to U1	42	132	0.3
R8 to R3	128	90	0.2	U8 to U6	34	350	0.3
U3 to U1	126	298	0.8	U4 to U1	23	107	0.1
$\eta < 83\%$ to $\eta < 96\%$ and On-d. to CH	122	1,911	0.3	$\eta > 83\%$ to $\eta < 94\%$ and CI to CH	15	72	0.1
R4 to R1	113	877	0.8	LG to $\eta > 83\%$ and On-d. to CI	10	121	0.1
R8 to R6	109	1,002	0.4	M. exh. to on-d. dec.m. with m.exh.	−50	41	−0.8

Note 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 and about three quarters of dwellings with window improvement also had other measures taken.

Table 7 shows the actual gas reduction, number of dwellings in a category and the ratio between actual and theoretical consumption reduction, that is, how many times larger the actual saving is from the theoretical. The highest reduction is achieved by drastically improving the U value of the windows (U8 to U1). The actual reduction of such a change (6 first row left) is below the theoretical and the number of dwellings in this category is rather low. A category that contains the most dwellings, is the dwellings where heating system was replaced from a $\eta > 83\%$ to $\eta < 96\%$ 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 a replacement of heating and hot tap water system (usually from a rather poor performing system prior to the renovation). Improvements of the envelope follow, however, the actual reduction there is in general lower than expected. An exception is improvement from R2 to R1. Changes of the ventilation system achieve a lower actual gas reduction, however, it is important to note that an upgrade from a natural ventilation to mechanical exhaust ventilation still yields a saving five times higher than expected. Other changes in ventilation system yield less saving and are also mostly overpredicted (except upgrading natural system to a mechanical balance, where the prediction is relatively good).

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 replacements that yielded a reduction higher than theoretical (third and eight row of Table 7) are both within a certain boiler type (in first case non-condensing and in the second, condensing), other replacements of heating systems have an actual consumption lower than the theoretical. This probably means that some of the calculation factors used for efficiencies of gas boilers do not reflect the real efficiency correctly.

4. Discussion

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.

On average a single a single measure leads to 131 m^3 gas savings while two measures lead to 188 m^3 savings which makes up for a reduction of 11.6% and 16.9%. Considering the report by Hezemans et al. (2012), which assumed that two measures coincide with a 20% reduction, this value now seems quite realistic to mildly overpredicted on the basis of this paper as well.

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 was identified within SHAERE demonstrates that this is not entirely true. This will probably improve in the future as the database grows, however, it was 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 the fact that during the renovation, the address of the dwelling sometimes changes, especially in the case 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 an insulation of wall, floor and roof have been simplified in this paper and would be interesting to analyse independently using continuous instead of categorical values. Also, 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. Conclusions and policy implications

To conclude, several main findings can be summarised. As stated in the introduction, the objective of the paper was to first evaluate the performance gap in renovated buildings on a large scale and second to use these results in order to evaluate what renovation measures seem to be the most effective in terms of actual savings. 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. One can now see that not only is the indoor temperature not well predicted (Majcen et al., 2013b), but also the efficiencies of systems and insulation values.

These results can be directly translated into actual and theoretical savings achieved per renovation measure. In terms of single measures reductions, replacements of gas boilers with more efficient ones (heat and hot tap water) yields the biggest energy reduction, followed by deep improvements of window quality. Replacing the ventilation system yields a relatively small reduction compared to other measures, however, it is still much larger than theoretically expected. A shortcoming of this study was not analysing the buildings with a combination of measures, as these might in fact be the most common. Analyses of these complex renovations will follow in subsequent publications and will help clarify the actual effectiveness of renovations further.

Regarding the recommendations for policy, the results obtained pose a question of how well the standard values of the method used for the calculation of the theoretical consumption are defined, in particular the indoor temperature, the ventilation rates, internal heat gains and infiltration losses, but also the efficiencies of the heating installation systems. It could be that excessively low efficiencies have been attributed to inefficient systems simply because of misconception and lack of knowledge and/or validation on a large scale. Moreover it is difficult to avoid the suspicion that such standard values have been in use in order to seemingly increase the innovation rate and falsely inflate the theoretical potential for improvement. However, now that actual consumption data on the level of individual dwelling is available, the inconsistencies are no longer concealed and there is no excuse to continue this status quo. The standard values should either be revised or alternatively, one should utilise the available actual gas consumption values in order to make better estimates (as suggested already in Majcen et al. (2015)). More realistic standard values would result in a more accurate estimation of consumption on the dwelling stock scale. An accurate baseline on the dwelling stock level would enable policy makers to implement measures which realistically lead to agreed national targets for reduction of energy consumed in the built environment.

It is of utmost importance to ensure building performance databases of sufficient quality and 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.

Energy performance registers should be made publicly available, possibly already coupled with actual consumption data. The availability is an issue in many European country and even in The Netherlands, which is generally progressive in this field, privacy restrictions are the main reason for refraining from opening such data. Several options exist for solving this problem, such as anonymization and aggregation of records.

Moreover, large datasets such as the one investigated in this paper are now arising across Europe, however, few experience is available about how to handle them. Experience with the use of such data should be shared and made available to the public. 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.

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.

References

Adalberth, K., 1997. Energy use during the life cycle of single-unit dwellings: examples. Build. Environ. 32 (4), 321–329.

Berkhout, P.H.G., Muskens, J.C., Velthuijsen, J.W., 2000. Defining the rebound effect. Energy Policy 28 (6–7), 425–432.

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 Build. 36 (6), 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. In: Proceedings of the European Council for an Energy Efficient Economy (ECEEE) Summer School, 6–11 June 2011, Belambra Presqu'île de Giens, France.

Dodoo, A., Gustavsson, L., Sathre, R., 2010. Life cycle primary energy implication of retrofitting a wood-framed apartment building to passive house standard. Resources, Conserv. Recycl. 54 (12), 1152–1160.

Filippidou, F., Nieboer, N., Visscher, H., 2015. Energy efficiency measures

implemented in Dutch non-profit housing sector. In: Proceedings of ECEEE 2015 Summer Study, Hyeres, France.
Filippidou, F., Nieboer, N., Visscher, H., Energy efficiency measures implemented in the Dutch non-profit housing sector, Energy and Buildings, June 2016, ISSN 0378-7788.
Haas, R., Biermayr, P., 2000. The rebound effect for space heating empirical evidence from Austria. Energy Policy 28 (6), 403–410.
Hens, H., Parijs, W., Deurinck, M., 2010. Energy consumption for heating and rebound effects. Energy Build. 42 (1), 105–110.
Hezemans A., Marquart E., Monné T., 2012. Monitor Energiebsparing Gebouwde Omgeving 2012, Agentschap NL, June 2012.
ISSO 82.3, 2009. Publication Energy Performance Certificate—Formula Structure Publicatie 82.3 Handleiding EPA-W (Formulestructuur'), Senternovem, October 2009.
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. Energy and Buildings. 105, 43–59.
Marchio, D., Rabl, A., 1991. Energy–efficient gas heated housing in France: predicted and observed performance. Energy Build. 17, 131–139.
Sunikka-Blank, M., Galvin, R., 2012. Introducing the prebound effect: the gap between performance and actual energy consumption. Build. Res. Inf. 40 (3), 260–273.
Thormark, C., 2002. A low energy building in a life cycle—its embodied energy, energy need for operation and recycling potential. Build. Environ. 37 (4), 429–435.
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 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. Sol. Energy 66 (6), 387–393.