Estimating the influence of rebound effects on the energy-saving potential in building stocks

Brøgger, Morten; Bacher, Peder; Madsen, Henrik; Wittchen, Kim B.

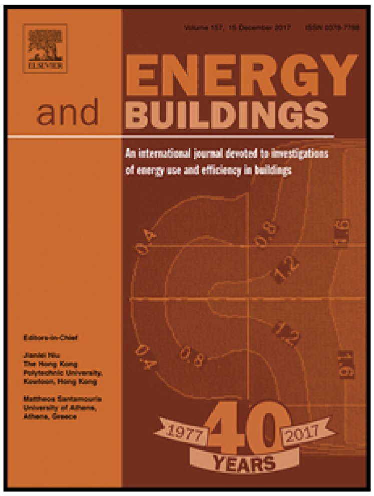
Published in:
Energy and Buildings

DOI (link to publication from Publisher):
10.1016/j.enbuild.2018.10.006

Publication date:
2018

Document Version
Accepted author manuscript, peer reviewed version

Accepted Manuscript



Estimating the influence of rebound effects on the energy-saving potential in building stocks

Morten Brøgger, Peder Bacher, Henrik Madsen, Kim B. Wittchen

PII: S0378-7788(18)30579-6
DOI: https://doi.org/10.1016/j.enbuild.2018.10.006
Reference: ENB 8835

To appear in: *Energy & Buildings*

Received date: 19 February 2018
Revised date: 17 August 2018
Accepted date: 7 October 2018

Please cite this article as: Morten Brøgger, Peder Bacher, Henrik Madsen, Kim B. Wittchen, Estimating the influence of rebound effects on the energy-saving potential in building stocks, *Energy & Buildings* (2018), doi: https://doi.org/10.1016/j.enbuild.2018.10.006

Estimating the influence of rebound effects on the energy-saving potential in building stocks

Morten Brøgger[a], Peder Bacher[b], Henrik Madsen[b], Kim B. Wittchen[a]

[a] *Danish Building Research Institute, Aalborg University, A.C. Meyers Vænge 15, DK-2450 Copenhagen*
[b] *DTU COMPUTE, Department of Applied Mathematics and Computer Science, Asmussens Allé, Building 303B, DK-2800 Kgs. Lyngby*

Abstract

The energy-saving potential in buildings (e.g. buildings proposed for an energy upgrade in an energy policy context) is often overestimated because implicit factors, such as rebound effects, are ignored. In order to get an accurate estimate of the *realisable* energy-saving potential in a building stock, these factors, as well as how they differ among buildings with different characteristics, must be accounted for.

On the building stock level, detailed information about the actual conditions in each building (e.g. indoor temperatures, domestic hot water consumption, internal heat loads, etc.) is rarely available. In its place, fixed assumptions are often made, usually disregarding the characteristics of the buildings. Therefore, a method that is based on available building stock data is needed for adjusting this *technical* energy-saving potential.

This study investigated how the heat consumption in residential buildings might be expected to change due to an energy upgrade, using a hybrid bottom-up model of the Danish residential building stock. Pseudo-rebound effects, inherent to the thermal standard of the building, were quantified in a sample of 134.000 buildings with different characteristics.

Results showed that estimating the heat-saving potential on the basis of the thermal characteristics alone (i.e. the technical heat-saving potential), would lead to a considerable overestimation of the realisable heat-saving potential in the residential building stock. However, the size of the realisable heat-saving potential was found to vary considerably among buildings with different characteristics, despite having the same technical potential. This indicated that the technical heat-saving potential should be corrected differently in buildings with different characteristics.

Keywords: Energy-saving potential, Buildings, Heat consumption, Building stock, Performance gap, Rebound effect

1. Introduction

With a view to comply with international agreements to reduce CO_2 emissions, energy-conservation measures provide cost-effective means to reduce the energy consumption in buildings [1, 2]. However, the realisable energy-saving potential is not easily assessed, because building's energy consumption is not easily calculated. This is demonstrated in several studies, where discrepancies between the calculated energy demand and the actual energy consumption are found [3, 4]. This discrepancy is commonly known as *the performance gap* [5, 6]. The existence of this gap has many possible causes including use of inappropriate modelling and simulation tools, lack of validation and assumptions that differ from the actual operation of the building among other things [7]. It should be noted that some of the potential causes of the performance gap are inherent to the particular application, i.e. whether the project is at the design stage or we are considering an existing building [7].

In the present study, we considered the heat consumption in a large sample of existing residential buildings in combination with the thermal standard of these buildings, in order to assess the realisable energy-saving potential while accounting for implicit effects.

1.1. Modelling energy consumption in existing buildings

It is commonly recognised that energy consumption in buildings is determined by four main components; namely the building envelope, building services (systems), the users of the building and external weather conditions. Modelling building's energy consumption requires knowledge about all four components; however, in most cases, information about these factors is incomplete. It should be noted that the the level of information available is very context dependent, however. Especially the user aspect tends to be difficult to get information about. This leaves the modeller with no choice but to make assumptions in place of the missing information, which was identified as one of the root causes of the performance gap by de Wilde [7].

Moreover, the same assumptions are made often about the users without taking the thermal standard of the building into account, e.g. before and after implementation of an energy-conservation measure. However, several studies suggest that the user's preferences are in fact related to the thermal quality of the building [8]. *The rebound effect* [9, 10] and *the prebound effect* [11] (which are closely linked) have been suggested as two of the primary explanations for the performance gap [3, 12]. In general terms, the former suggests that the building users tend to increase the comfort level after an energy upgrade, while the latter suggests that people tend to accept a poor indoor environment in poorly insulated buildings (sometimes referred to as 'fuel poverty'). Figure 1 illustrates these implicit effects conceptually.

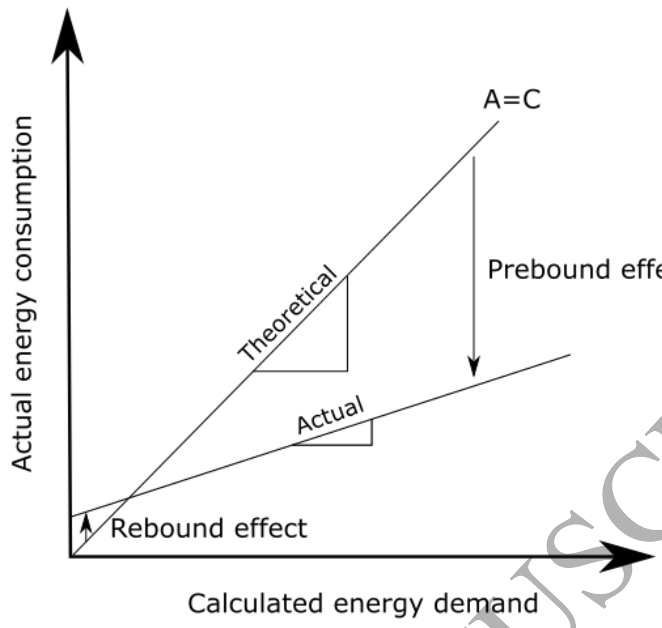


Figure 1: Several studies suggest that residents compromise on the indoor environmental quality in energy inefficient buildings (prebound effect) and improve it when energy upgrading (rebound effect). Conceptual illustration adapted from [8]

It is crucial to account for these phenomena when evaluating the energy-saving potential in buildings, e.g. in a policy context [8]. Furthermore, a recent study suggested that this rebound effect depends on more than just the thermal standard of the building, e.g. the household type [5], possibly because the same assumptions are made regardless of the non-thermal characteristics of each building. Likewise, other studies report on other factors, such as income, age and number of residents, that also have an effect on the heat consumption, but which are not accounted for in traditional heat demand calculations [13, 14, 15].

Therefore, it is necessary to study whether the same assumptions can be made across buildings with different characteristics and of different thermal standard. This way, we can estimate how much an energy-conservation measure is likely to actually reduce the heat consumption (by quantifying rebound effects), i.e. get better estimates of the *realisable* energy-saving potential.

1.2. Estimating the energy-saving potential in building stocks

In a building stock context, where many buildings with diverse characteristics are considered collectively, information about the users is seldom available. Therefore, indoor temperatures, air change rates due to opening of windows and doors, internal heat loads, among other things have to be assumed. This renders it particularly urgent to find a solution when considering large building stocks, e.g. at national scale [16, 17].

Different remedies have been proposed to work around the problem of the performance gap at scale, including adjusting assumptions based on expert knowledge [18] and empirically derived adaption factors [19].

In the present study, we investigated whether the same assumptions could be used across groups of buildings with different characteristics, as well as in groups of buildings of different thermal standard, in order to quantify the the rebound effect. This was assessed by examining how the heat consumption related to the thermal performance (in terms of the calculated heat demand) of buildings with different characteristics.

1.3. Study focus and delimitation

Focus of the present study was on enhancing current methods, in order to get more accurate estimates of the realisable heat-saving potential in a building stock perspective. Therefore, only existing buildings were considered. Likewise, only data from available sources, which were not subject to data privacy, where used. This entails that only building characteristics were considered. Hence, information about the occupants were not included. Likewise, information on indoor temperatures, DHW usage, heat gains from electrical appliances, etc. was not available. Therefore, causality (in terms of what factors that influenced the heat-saving potential) could not be studied in detail. Moreover, only residential buildings were considered.

Energy consumption was studied at a whole building level (e.g. a block of flats or a detached single-family house), since potential energy savings are often estimated at this level, rather than on a single unit (e.g. an apartment) level.

2. Data description

The analyses in the present study were based on data on 134,093 Danish residential buildings, obtained from two databases; the Danish Energy Performance Certificate (EPC) database, and the Danish Building and Dwelling Register (BBR).

The EPC database contains information about the physical properties of each building, down to a single component level, e.g. U-values and areas of all external walls in any given building. Moreover, the EPC database contains information about heated floor areas, types of heat supply, ownership and geographical location of each building, among other building related characteristics. Data in the EPC database was registered by energy experts, based on visual inspections made at an on-site visit upon issuing an EPC as required by the European energy performance of buildings directive (EPBD) [20].

The BBR database is a publicly available database, which is managed by the Danish tax authorities, that contains information about heated floors areas, year of construction and renovation, among many other variables related to each building and its technical systems. Moreover, the BBR contains registrations of the heat consumption for each building or property, reported by Danish utility companies. This entails that the registered heat consumption was the delivered (i.e. gross final) heat to each building. The heat consumption in each building was registered upon account, which implies that the length of the consumption periods varied from building to building. To get a corresponding annual heat consumption for each building, the registered heat consumption was normalised by the Danish tax authorities by means of the number of heating degree days in the given consumption period. It should be noted that the registered heat consumption included heat for both space heating and domestic hot water (DHW).

The characteristics in Table 1 were extracted from the two databases and used in the analyses of rebound effects in building with different characteristics.

Explanatory variable*	**Scale**	**Levels/Range**	**Abbreviation**
Calculated heat demand	Ratio	20 kWh/m^2 - 500 kWh/m^2	Q_{calc}
Building type	Nominal	Farmhouse (Farm), Detached SFH (SFH), Terraced house (Row) or Block of flats (MFH)	Type
Primary heat supply	Nominal	Individual boiler or District heating	PHS
Secondary heat supply	Nominal	None, Electrical heating, Stove or Both	SHS
Tenancy	Nominal	No/Yes	Rent
Municipality code**	Nominal	101 - 860 (98 levels)	Mun

* Building characteristics
** Nuisance variable

Table 1: Explanatory variables used in the statistical analyses

The building type was included to account for differences in building geometry (e.g. surface to volume ratios), as well as household characteristics, e.g. number of inhabitants, etc. that was otherwise assumed to be identical across building types. Information about the primary and secondary heat supply was included to account for heating system efficiency (which was not accounted for in the calculated heat demand), as well as supplementary heating (e.g. firewood) that was not registered by the utility companies. The ownership status (Tenancy) was included to investigate potential differences between owner occupied buildings and building that were rented out. Lastly, the municipality code was included as a nuisance variable to control for possible regional differences, but was not used for interpretation, since focus was on potential heat savings on a national scale.

In addition to the variables listed in Table 1, information about the heated floor area of the building, the year of construction and the EPC rating (among many other variables) was also available in the EPC database. However, because this information was already reflected in Q_{calc} (i.e. the variables were collinear), these variables were not included in the statistical model.

Unfortunately, causes of observed differences could not be determine because the available data did not include sufficient information about indoor temperatures, air changes rate, etc.

2.1. Calculated heat demands

The calculated heat demands (Q_{calc}) comprised heat losses due to transmission and ventilation, as well as heat losses from building services (i.e. heat- and DHW distribution pipes) and heating for DHW preparation, calculated in accordance with relevant European standards, ISO 13790 and ISO 15316 parts 2.3 and 3.2 [21, 22, 23]. The heat demand was calculated as the net energy demand; hence, not including the efficiency of the heating system (e.g. boiler efficiency) in the building.

Heat demands were calculated for each building individually, based on thermal properties, corresponding areas, as well as other relevant information, of each component registered in the Danish EPC database by energy experts. In addition to the physical properties of the building, a fixed average indoor temperature of 20°C was assumed by the authors. Assuming the same indoor temperature in all buildings allowed for detection of systematic differences among groups of buildings, by considering differences in heat consumption; i.e. evidence derived based on the available data.

All other variables were derived from the experts' inspections, including estimated air change rates, building heat capacity, internal heat loads, etc. However, it should be noted that many of these values were likely to be standard values, assumed by the energy auditors. In order to exclude faulty values, restrictions on maximum heat gains from persons and appliances were imposed according to national Danish calculation methods [24].

It should be noted that Q_{calc} was calculated anew for each building (and hence not extracted from the EPC database directly). This was done to exclude primary energy factors (as we considered delivered energy), as well as energy demands for electricity, which are part of the calculated energy demand in the EPC database [25]. For a detailed description of how the heat demands were calculated, refer to [26].

2.2. Data pre-processing

To exclude outliers in terms of buildings with extremely high or extremely low heat consumption, e.g. empty houses or faulty registrations, two thresholds were imposed; a minimum heat consumption of 20 kWh/m^2 and a maximum heat consumption of 500 kWh/m^2. Moreover, all registrations with a difference between Q_{calc} and Q_{reg} exceeding 300% were excluded. This criterion was imposed to discard erroneous registrations, while allowing for a large natural variation in the heat consumption. The specific threshold was chosen based on previous studies, which found the variation in energy consumption in buildings with identical characteristics, to be as large as 300% [27]. In addition, only buildings registered as constructed after 1600, and only buildings with a valid primary heat supply, were included in the analyses. Removing erroneous observations reduced the data set from 153,821 to 134,093 buildings (12.8%).

2.3. Visual data inspection

Plotting Q_{reg} against Q_{calc} for each building in the sample, served as a first inspection of data.

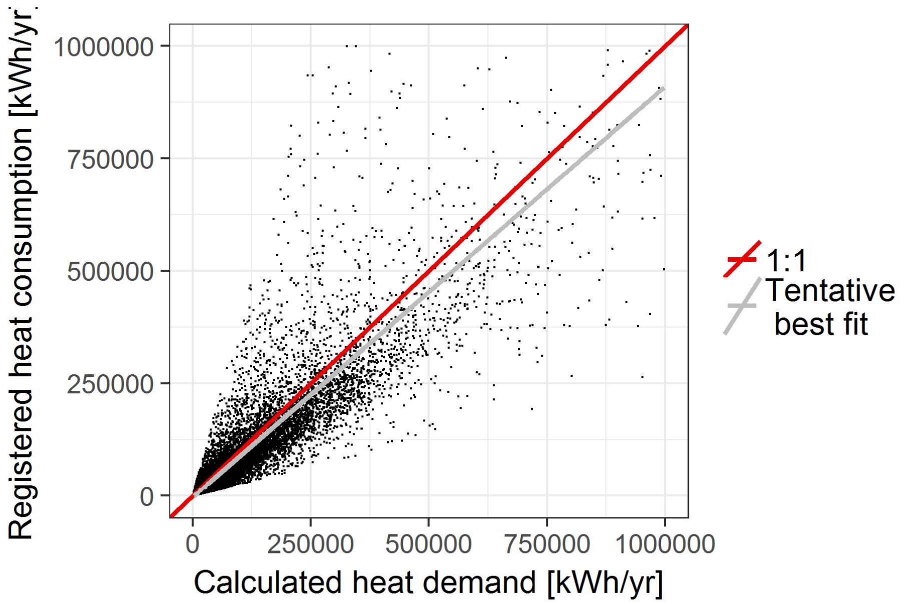


Figure 2: Registered heat consumption vs. calculated heat demand

Considering all buildings collectively, a modest systematic discrepancy between the calculated- and the registered heat consumption was observed (i.e. a performance gap), in addition to the natural variation. This discrepancy appeared as the difference between the red 1:1 (perfect agreement) and the grey tentative best-fit line. However, despite the discrepancy, it should be noted that the calculated heat demand clearly reflected the heat consumption.

3. Method

In order to get more realistic estimates of how much an energy-conservation measure would reduce the heat consumption (i.e. the realisable energy-saving potential), differences in heat consumption were studied conditional on the calculated heat demand. In this context, the calculated heat demand (Q_{calc}) was used to reflect the thermal properties of the building.

Pseudo-rebound effects were quantified by estimating differences between the calculated heat demand and the registered heat consumption in a linear regression model[1]. The building characteristics in Table 1 were included in this assessment, in order to clarify aspects that could have an effect on the rebound effect.

Data included only physical properties of the building itself; hence, user behaviour was not modelled explicitly. This implies that the same boundary conditions (i.e. indoor temperatures, air change rates, etc.) were used in the calculations of the heat demand of each building, regardless of the thermal standard of the building. By making the same

[1]It should be noted that the rebound effect is normally defined only for the same building before and after undergoing an energy upgrade [8]. Therefore, the term the term *pseudo-rebound effects* was adopted here, where we considered *potential* energy upgrades.

assumptions about indoor conditions in all buildings (regardless of the characteristics and the thermal standard of the building) we let data suggest any differences between buildings with different characteristics, in terms of differences in heat consumption.

Therefore, systematic differences in the relationship between Q_{calc} and Q_{reg} in groups of buildings with different characteristics that could not be explained by the model were attributed to differences in the assumed boundary conditions in each group of buildings. This approach offered a way to account for differences among groups of buildings with different characteristics, which appeared to be associated with the thermal standard of the building.

The influence of each building characteristic could be seen as a measure of the relative importance of each characteristic, similarly to those in [13], or as correction factors similarly to those in [19].

3.1. Statistical model

The relationship between the calculated heat demand (Q_{calc}) and the registered heat consumption (Q_{reg}) was modelled statistically by means of ordinary least squares (OLS) multiple linear regression (MLR), in which the variables listed in Table 1 entered as explanatory variables. This allowed for an assessment of whether the relationship between Q_{calc} and Q_{reg} was systematically different, in buildings with different characteristics, in terms of the conditional mean heat consumption.

The multiple linear regression model included main effects, as well as first order interactions between each building characteristic and the calculated heat demand. This implies that second- and higher order interactions were omitted. The first order interaction terms were used to model pseudo-rebound effects because they represent the slopes of the regression lines in the model; i.e. given a reduction in heat demand (Q_{calc}), how much would the heat consumption change, see Figure 3.

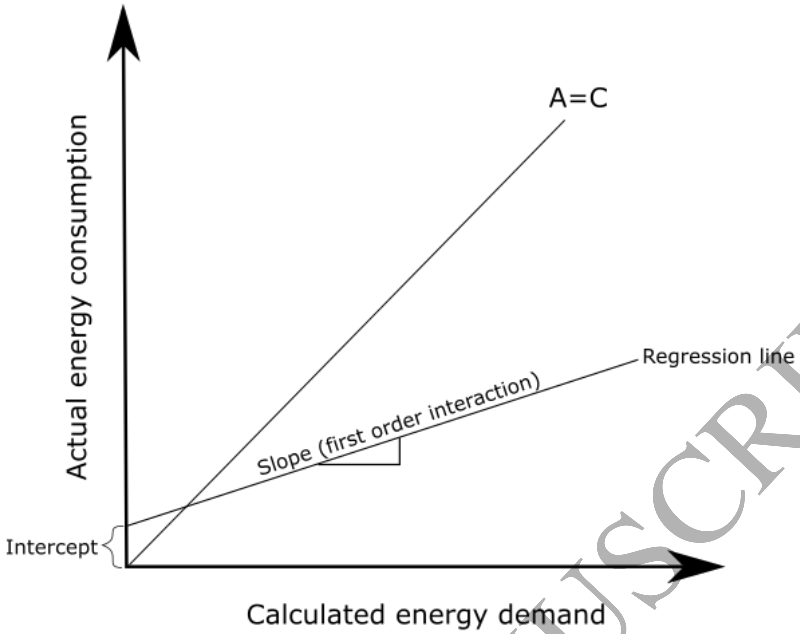


Figure 3: Using a linear regression model, the slope can be used for estimating the reduction in heat consumption given a reduction in heat demand due to an energy-conservation measure

A slope equal to one would indicate a 1:1 reduction in heat consumption due to an ECM (i.e. A=C on Figure 3). Slopes above one indicated an underestimation of the change in heat consumption whereas slopes below one indicated overestimates of changes in heat consumption.

The main effects, i.e. the differences in intercepts among groups of buildings, were not interpreted explicitly in the present study, because only the slopes of the regression lines were of interest. However, it should be noted that the intercept corrected for differences in the fraction of the heat consumption that the DHW accounted for; e.g. in new, well insulated buildings with a low demand for space heating compared with old buildings with a large space heating demand, see Figure 4.

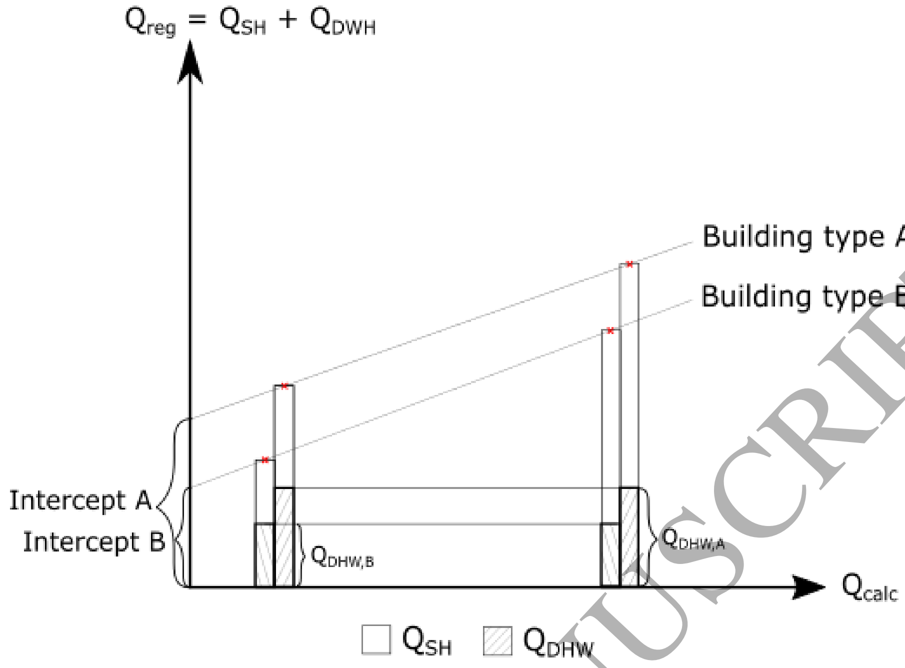


Figure 4: The fraction of the total heat consumption that preparation of DHW accounts for is taken into account in the model by means of the intercepts (conceptual illustration)

This implies that the DHW consumption was assumed to be unaffected by the thermal standard of the building.

3.1.1. Effect estimation

Whether there was a significant difference in rebound effects between buildings with different characteristics was evaluated statistically by means of analysis of variance (ANOVA) tests. The ANOVA test assesses whether the mean value is equal across different groups, in order to determine the statistical significance of observed differences [28]. In section 5, only the p-values are reported; however, the full ANOVA tables are listed in Appendix B.1.

In order to test whether each level of each characteristic (e.g. each building type) was significantly different from the other, we first reduced the model by merging the levels of each characteristic in turn, and tested whether the reduced model was significantly different from the full model. It should be noted that the model reduction approach tested the collective effect of each characteristic; i.e. the intercept and slope of the regression lines in combination.

Effect sizes were evaluated by considering the estimated slopes (given by the interaction terms), including the corresponding 99 % confidence intervals (CI), of the MLR model. Furthermore, in order to ensure that the parameter estimates of the MLR model were not distorted by potential outliers, the parameter estimates from the MLR model were compared with parameter estimates from a robust regression (RR) model. The RR model assigns a weight to each observation, depending on the corresponding residual. This way, highly influential observations that were not deemed errors in the data pre-processing, e.g. buildings with very high or very low heat consumption, were down-weighted [29]. The the RR model makes less restrictive assumptions which makes the parameter estimates more robust to outliers [30].

4. Model description

First, a simple linear regression model, in which the calculated heat demand (Q_{calc}) served as the only explanatory variable, was fitted. However, model validation plots indicated significant bias in the model, see Figure A.7 in Appendix A. Therefore, a multiple linear regression (MLR) model was fitted, in which all the building characteristics in Table 1 were included, in order to describe the variation in data, so that the residuals were unbiased (i.e. random).

In addition, considering Figure 2, it is evident that the variance in heat consumption was not constant. Therefore, the continuous variables were log-transformed for the linear model to fit the data structure and remedy any heteroscedasticity (i.e. inconstant variance). A Box-Cox analysis [31], performed using the MASS package [32] in R [33], suggested that a logarithmic transformation of the dependent variable (in this case the registered heat consumption) was appropriate. Moreover, choosing a model in which both the continuous variables were log-transformed offered a nice interpretation since *the interpretation is given as an expected percentage change in Y when X increases by some percentage* [34]. Considering the relationship between the two log-transformed variables depicted in Figure 5 the variance evidently became much more uniform after the transformation.

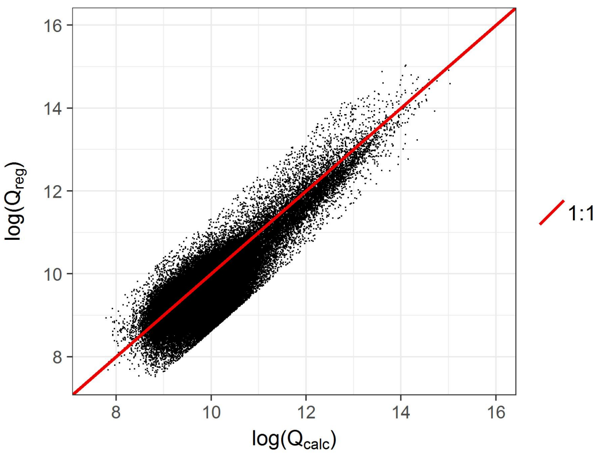


Figure 5: Log-transformed variables

The clear-cut edge of the data point parallel to the red line were caused by the 300% difference restriction in the data pre-processing, see subsection 2.2.

The MLR model is outlined in equation 1 below.

$$
\begin{aligned}
\log(Q_{\mathrm{reg},i}) = {} & \beta_0 + \beta_1 \cdot \log(Q_{\mathrm{calc},i}) \\
& + \beta_2 \cdot \mathit{Type}_i + \beta_3 \cdot \mathit{PHP}_i + \beta_4 \cdot \mathit{SHS}_i \\
& + \beta_6 \cdot \mathit{Rent}_i + \beta_7 \cdot \mathit{Mun}_i \\
& + \beta_8 \cdot \mathit{Type}_i \cdot \log(Q_{\mathrm{calc},i}) + \beta_9 \cdot \mathit{PHS}_i \cdot \log(Q_{\mathrm{calc}}, i) \\
& + \beta_{10} \cdot \mathit{SHS}_i \cdot \log(Q_{\mathrm{calc},i}) + \beta_{12} \cdot \mathit{Rent}_i \cdot \log(Q_{\mathrm{calc},i}) + \epsilon_i
\end{aligned} \tag{1}
$$

The parameter estimates β_2 through β_7 denote differences in intercepts from one group of buildings to another. Likewise, β_8 through β_{12} (i.e. the first order interactions) denote the differences in slopes between groups of buildings with different characteristics.

This model showed no notable sign of dependence between the fitted values and the residuals; i.e. they appeared to be randomly distributed, see Figure A.8 in Appendix A. However, some observations did not follow the normal distribution perfectly, in the upper right corner of Figure A.8b, indicating that these observations could be outliers.

4.1. Robust regression

In order to ensure that potential outliers, in terms of observations with large residuals and leverage, did not compromise the parameter estimates from the MLR model, a robust regression (RR) model was fitted in which these observations were down-weighted.

By comparing the parameter estimates from the MLR model with the parameter estimates from the robust regression model, extreme observations were found not to affect the parameter estimates of the MLR model unacceptably, see Table C.11 in Appendix C. Therefore, the parameter estimates from the MLR model were used for effect size estimation throughout the study.

5. Results

This section evaluates whether potential heat savings due to an energy-conservation measure depends on the characteristics of the building, in addition to the thermal standard, including how much.

5.1. Effect size estimation - pairwise comparisons

Effect sizes were assessed by means of the parameter estimates. However, before comparing the parameter estimates, we corrected for multiple comparisons by adjusting the width of the confidence interval, using Bonferroni correction [28]. With an initial significance level of $\alpha = 0.01$ and $n_{\mathrm{test}} = 14$ tests (the four building types tested against one another, two types of primary heat supply tested against each other, three types of secondary heat supply and rental/not rental, considering only main effects), this left us with adjusted significance level of 0.001:

$$
\alpha'' = \frac{\alpha}{n_{\mathrm{test}}} = \frac{0.01}{14} = 0.0007 \approx 0.001 \tag{2}
$$

Which corresponds to increasing the confidence interval to 99.9993 %.

In the subsequent sections, the effect size of each characteristic is considered in turn, in order to evaluate how much each characteristic affected the realisable heat saving potential.

5.2. Effects on heat consumption

To test whether pseudo-rebound effects were significantly different in buildings with different characteristics, we first considered the ANOVA table (Table B.8) in Appendix B.1. The p-values are listed in Table 2 below.

Interaction	$\log(Q_{\mathrm{calc}})$:Type	$\log(Q_{\mathrm{calc}})$:PHS	$\log(Q_{\mathrm{calc}})$:SHS	$\log(Q_{\mathrm{calc}})$:Rent
p-value	0	0	0	0

Table 2: p-values from ANOVA test for significant effects of each building characteristic. The full ANOVA table is listed in Table B.8

The p-values provided strong evidence against the null hypothesis of no effect (because p-value $< \alpha''$); i.e. the interaction terms indicated that the relationship between the Q_{calc} and Q_{reg} depended on each of the characteristics of the building.

It should be noted that the assumption of common variance among the groups of buildings was not satisfied in the present case, see Appendix E.1. However, since the test is robust with respect to violation of this assumption, we proceeded using this test.

5.3. Effects of building type

Having established an effect of the building type, we considered which of the four building types that were different from the other. In Table 3 the reduced models were compared with the full model.

Model	Full model	Farm + SFH	Farm + Row	Farm + MFH	SFH + Row	SFH + MFH	Row + MFH
p-value	-	$1.427 \cdot 10^{-12}$	$9.039 \cdot 10^{-41}$	$1.761 \cdot 10^{-54}$	0	0	0

Table 3: ANOVA table testing whether the model could be reduced by merging the building types indicated in the Model-row. The full table may be found in Table B.9

The p-values indicated that we could not reduce the model by merging building types, because they were lower than the α'' level in Equation 2, meaning that each building type was significantly different from the others. Therefore, we considered the estimated slopes (i.e. the estimated intercepts of the model) related to each building type to evaluate the effect size; i.e. how different the relationship between Q_{calc} and Q_{reg} was in the four building types.

In Table 4, the estimated slopes are given together with the corresponding confidence intervals. These denote the average (i.e. conditional mean) effect of each building type, when correcting for the other building characteristics in the model. To get the slope of a building with particular characteristics, the estimated slopes may be adjusted consecutively using the parameter estimates in Table C.11.

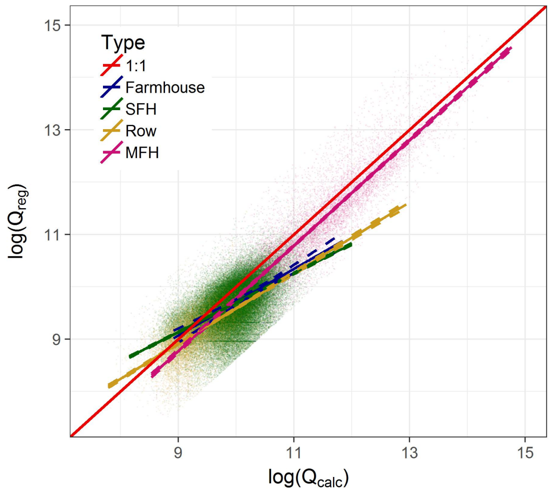
Building type	**lwr. CI**	**Estimated slope**	**upr. CI**	**Std. Error**
Farm	0.524	0.613	0.702	0.028
SFH	0.484	0.497	0.509	0.004
Row	0.582	0.600	0.617	0.006
MFH	0.897	0.914	0.932	0.005

Table 4: Estimated slopes of the regression lines (i.e. interaction terms of the MLR model in Equation 1) for each building type. The regression lines are depicted in Figure 6

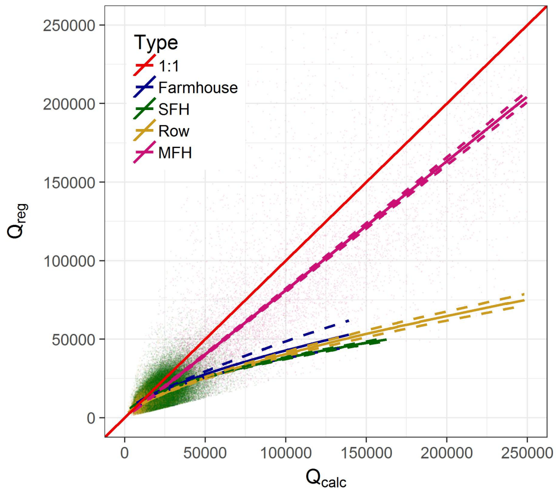
The estimated slopes indicated that potential heat-savings were likely to be different for the four building types, except for farmhouses and detached single-family houses, as well as farmhouses and the row houses (because their confidence intervals overlapped).
In practice, this suggested that multi-family houses were much less subject to rebound effects than singe-family houses, thereby offering a larger energy-saving potential due to energy-upgrading of the building envelope. The parameter estimates suggested that changing the calculated heat demand by 1 % in a block of flats would result in a 0.914 ± 0.18 % change in heat consumption on average, whereas the same change in a detached single-family house would only change the heat consumption by some 0.497 ± 0.13 %. Hence, the reduction in heat consumption in a single-family house, given a unit reduction in heat demand, could be expected to be only about half that in a block of flats!

It should be noted that we found all building types to be significantly different from one another in subsection 5.2; this is most likely because each effect (i.e. the intercept and the slope) was tested collectively in the former, while we tested each effect separately, when considering the parameter estimates.

To illustrate the difference in heat consumption between different building types, we plotted the calculated heat demand against the registered heat consumption for each of the building types, see Figure 6.



(a) Q_{calc} vs. Q_{reg} in the log-transformed variable space



(b) Q_{calc} vs. Q_{reg} in the back-transformed variable space

Figure 6: Relationship between Q_{calc} and Q_{reg} for the four building types. The dashed lines represent the respective confidence intervals.

Figure 6 clearly illustrates the difference in energy-saving potential (i.e. rebound effect) between the four buildings types.

It should be noted that the regression lines in Figure 6 were based on a simplified model, in which the calculated heat demand and the building type were the only predictors. This means that subgroups were not separated (e.g. farmhouses with boilers could not be distinguished from farmhouses with district heating). To overcome this simplification, all subgroups should be considered separately, which was not done here. However, for illustration purposes, Figure 6 clearly illustrated a difference in the relationship between Q_{calc} and Q_{reg}, among the four building types.

5.4. Effects of the primary heat supply

Considering the results in Table 2, we found a significant effect of the heating system, i.e. whether there was a difference between buildings with an individual boiler and buildings with district heating.

Since only two types of primary heating systems were included in this study, we did not reduce the model by merging the two types into one, as this would be the same as excluding the primary heat supply from the model, which was tested in Table 2. Instead, we turned directly to the parameter estimates in order to investigate the size of the effect.

The *difference in slope* between the two groups of buildings are given in the table below.

Building type	lwr. CI	Estimated difference in slope	upr. CI	Std. Error	p-value
$\log(Q_{\text{calc}})$:PHS_{DH}	0.04	0.054	0.068	0.004	$5.716 \cdot 10^{-41}$

Table 5: Difference in the relationship between Q_{reg} and Q_{calc} (denoted by the difference in slopes of the regression lines) between buildings with individual boiler and buildings with district heating; taken from Table C.11.

The p-value indicated that the two slopes were in fact statistically significantly different, which was also what we found in Table 2. The difference in slope suggested that a difference in Q_{calc} was accompanied by a slightly greater difference in Q_{reg} in buildings with district heating than in buildings with an individual boiler, i.e. 0.054 ± 0.014 percentage points (given a 1 % change in Q_{calc}). Hence, buildings with district heating were subject to a slightly lower rebound effect (thus possessing a slightly larger energy-saving potential) than buildings with an individual boiler. However, considering the full table of intercepts in Table C.11, it is evident that buildings with district heating were consistently more energy-efficient than buildings with a boiler.

5.5. Effects of secondary heat supply

As with the two previous characteristics, the secondary heat supply turned out to have a significant effect on the relationship between Q_{calc} and Q_{reg}, see Table 2. Testing a simplified model, in which we merged all the three types of secondary heat supply (SHS), and testing it against the full model provided evidence of a significant effect of having a secondary heat supply, see Table B.10 in Appendix B.3. However, considering the parameter estimates in Table 6, neither having electric heating, nor having both a stove and electrical heating, as a secondary heat supply had a significant effect on the relationship between Q_{calc} and Q_{reg}.

Building type	lwr. CI	Estimated difference in slope	upr. CI	Std. Error	p-value
$\log(Q_{\text{calc}})$:SHS_{El}	-0.025	0.012	0.048	0.011	0.2769
$\log(Q_{\text{calc}})$:$\text{SHS}_{\text{Stove}}$	0.017	0.039	0.062	0.007	$2.856 \cdot 10^{-09}$
$\log(Q_{\text{calc}})$:SHS_{Both}	-0.001	0.065	0.131	0.02	0.0009371

Table 6: Effects of two types of secondary heating on the relationship between Q_{calc} and Q_{reg}

Having a stove as secondary heat supply, on the other hand, affected the relationship by 0.039 ± 0.023 percentage points (given a 1 % change in Q_{calc}). Hence, a similar conclusion could be drawn as that in the case of the primary heat supply: buildings with a stove were generally somewhat more energy efficient (i.e. had a lower registered heat consumption) compared with other buildings; however, the marginal effect of having a stove (on the rebound effect), was negligible.

5.6. Effects of rental status

The effect rental status was evaluated in a similar manner to that in the previous subsections. The parameter estimates are outlined in Table 6.

Building type	lwr. CI	Estimated difference in slope	upr. CI	Std. Error	p-value
$\log(Q_{\text{calc}})$:Rent_{Yes}	0.039	0.053	0.066	0.004	$2.92 \cdot 10^{-40}$

Table 7: Effects of rental status on the relationship between Q_{calc} and Q_{reg}

The difference in the slopes of the two regression lines indicate that there was in fact a significant difference in the relationship between Q_{calc} and Q_{reg} in buildings that were rented out and owner-occupied buildings (with rented buildings having a slightly higher potential for energy-savings). However, there could be differences between different types of ownership, e.g. private land lords and non-profit housing associations, which were not assessed here.

5.7. Combined effect

In the previous sections, all characteristics turned out to have a significant effect on the energy-saving potential (i.e. the relationship between Q_{calc} and Q_{reg}); however, the effect size varied considerably. Therefore, one effect could seem unimportant on its own; however, considering the effects in combination, the effect could be of a considerable size. For instance, the total effect of having district heating (versus having a boiler), having stove as secondary heating, in a rented house was 29.4 % higher than in a corresponding house with an individual boiler, with no secondary heating, in which the owners live (i.e. 0.497 vs. 0.643). All effect sizes are listed in Table C.11.

6. Discussion

In the present study, the influence of pseudo-rebound effects on the heat-saving potential in Danish residential buildings was assessed. This was done in order to assess the effect of an energy-conservation measure on the *realisable* energy-saving potential in buildings with different characteristics. Using a hybrid bottom-up model (i.e. a model that combines a building-physical model with a statistical model), the relationship between the calculated heat demand (Q_{calc}) and the registered heat consumption (Q_{reg}) was studied, in order to identify pseudo-rebound effects.

In the study, all differences in the relationship between (Q_{calc}) and (Q_{reg}) were attributed to pseudo-rebound effects, i.e. differences in user behaviour related to the thermal standard of the building, as well as other relevant building characteristics. This could cause the implicit effects to be overestimated (i.e. underestimating the energy-saving potential), if the thermal standard of the building fabrics had been misjudged across buildings of different thermal performance, as proposed elsewhere in the literature [35].

Furthermore, this study considered groups of buildings in order to account for implicit effects on the energy-saving potential in a building stock. Therefore, the present analyses offered only a superficial treatment of the cause of the observed discrepancies in different groups of buildings, rather than an in-depth analysis of one particular group of buildings. However, this method gave an indication of where discrepancies arise, providing a foundation for further research. Explaining the observed differences would require additional data as well as an in-depth analysis of the various subgroups, i.e. one group of buildings with a specific set of characteristics. This could include considering different subsets (e.g. different building types) in different models.

6.1. Data considerations

The present study was based on utility data, i.e. meter data, for which reason energy used for space heating and domestic hot water could not be considered separately. This could jeopardise the validity of the model, if energy for DHW preparation was related to the thermal standard of the house. However, given that the energy used for DHW preparation could be determined as a fraction of the heated floor area, this was accounted for correctly in the model in terms of the intercept, which was not analysed explicitly in this study.

Therefore, the proposed model could be used for analysing the energy-saving potential, with the available energy consumption data, even though energy consumption for DHW preparation may constitute a larger fraction in a new well-insulated house than in an old energy inefficient house. Likewise, electricity consumption was not included in the analyses (due to lack of data), except as an assumed part of the internal heat loads. Therefore, systematic differences in electricity consumption between groups of buildings could affect the results.

In addition, the thermal properties of the building, as registered by the energy auditors, used for calculating the heat demand relied on tabulated values, which in most cases were assumed based on the construction year of the building. Therefore, if these tabulated values were consistently incorrect, it would cause the calculated heat demands to be biased. This would in turn cause the model to be biased, because it was not possible to include the year of construction in the model. However, an assessment of the residuals showed no sign of any dependence between the residuals and the construction year, see Appendix E.

Lastly, different consumption periods were normalised to comprise an entire year. However, varying lengths of the consumption periods, as well as different types of fuels, could affect the variance of the heat consumption. Therefore, weighting the variance according the the actual length of the consumption period and the type of fuel might contribute to a better understanding of the heat consumption.

6.2. Explanatory variables

Since the present study only included building characteristics, behavioral and socio-economic characteristics such as disposable income and number of inhabitants were left unaccounted for. However, in a study of the effects of occupancy and buildings on the energy used for space heating and DHW in the Dutch residential building stock, the authors found the building characteristics to explain the majority of the variation in heat consumption

[13]. Moreover, including occupancy characteristics in their model turned out not to affect the results notably.

In addition, since no model assumptions were violated (the residuals being randomly distributed in particular) this was likely only to have a disguising effect, i.e. increase the width of the confidence intervals. Furthermore, including the municipality code as a nuisance parameter in the model should control for some of these effects to some extent.

6.3. Model considerations

Choosing an OLS regression model entailed excluding of certain explanatory variables, such as the heated floor area and the year of construction. However, some of the information these variables could contribute with was of course already reflected in the calculated heat demand. Moreover, an analysis of the residuals showed no sign of violation of the model assumptions.

7. Conclusion

In the present paper, we found that estimating the heat-saving potential in residential buildings on the basis of the thermal properties of the building alone, sometimes referred to as the *technical* heat-saving potential, would result in an overestimation of the *realisable* heat-saving potential.

Using data available from the Danish EPC scheme and publicly available records, it was possible derive factors for adjusting the technical heat-saving potential, in order to get more accurate estimates of the realisable heat saving potential in the residential building stock.

The proposed model made it possible to identify pseudo-rebound effects in the building stock under consideration. These pseudo-rebound effects were found in all parent groups of buildings. The size of these implicit effects varied considerably between buildings with different characteristics, ranging from no difference to a factor of two in the estimated heat-saving potential. However, even small effects mounted up in combination with each other, when considering a group of buildings with very specific characteristics.

Though not analysed explicitly in this paper, the differences in realisable heat saving potentials could suggest that the assumptions, in terms of average indoor temperatures, DHW consumption, air change rates, time of occupancy and electricity use (among other occupant related factors), used for calculating the technical heat-saving potential could be systematically different in buildings with different characteristics. Therefore, the technical energy-saving potential should be adjusted, in cases where assumptions were made due to limited access to data about the actual conditions in the building(s) under consideration.

Thermal properties and system efficiencies, which could have been misjudged in the EPC, also pose a potential cause for the observed discrepancy between the calculated heat demand and the registered heat consumption, which could not be captured by this model. Likewise, socio-economic characteristics, which were not included in this study, could also have an effect on the size of the heat-saving potential.

8. Further work

Since heat consumption for space heating and DHW could not be separated in the present study, it should be considered to do a sensitivity around the influence of heat consumption for DHW on the total heat demand.

In order to address what caused the observed differences in this study, individual factors should be investigated. Measuring the indoor temperature, air change rate, etc. across a broad sample of buildings could help justify (or falsify) some of the assumed input values. Likewise, an investigation of the actual transmittance values of the building envelope elements would be valuable in checking for systematic biases in the tabulated values that are normally used for EPC ratings.

9. Acknowledgements

This study was sponsored by the Danish Innovation Fund (Innovationsfonden) (4106-00009A) and should be seen as a contribution to the SAVE-E project, as well as the IEA Annex 70 and the IEA Annex 71.

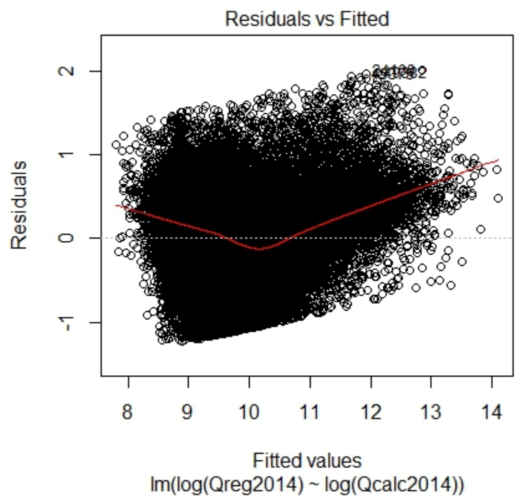
All plots in the present study were made in R [33] using either the built in plot function or the 'ggplot2' package [36].

Appendix A. Model validation plots

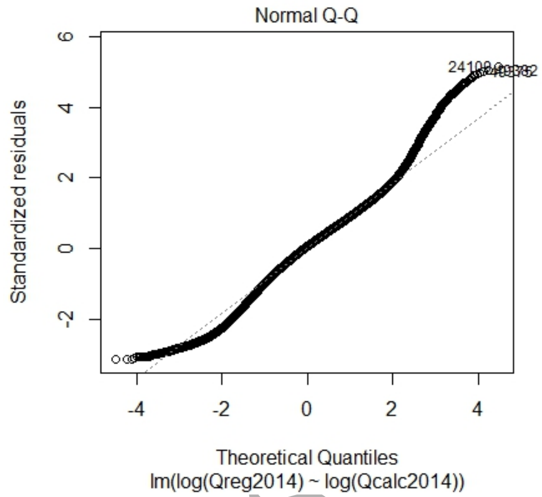
The present appendix presents the model validation plots for the simple linear regression (SLR) model, as well as for the multiple linear regression (MLR) model.

Appendix A.1. Simple linear regression model

In Figure A.7, the residuals of the SLR model are plotted against the fitted values of the model (Figure A.7a). The quantile-quanlie (Q-Q) plot (Figure A.7b) displays the distribution of the residuals.



(a) Residuals vs. fitted values



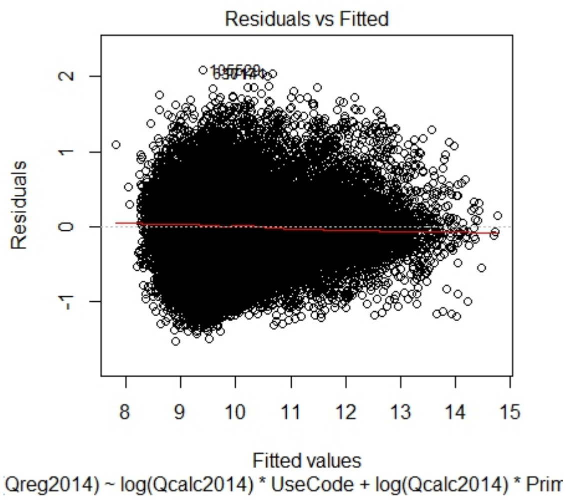
(b) Normal quantile-quantile plot of the residuals

Figure A.7: Distribution of the SLR model residuals

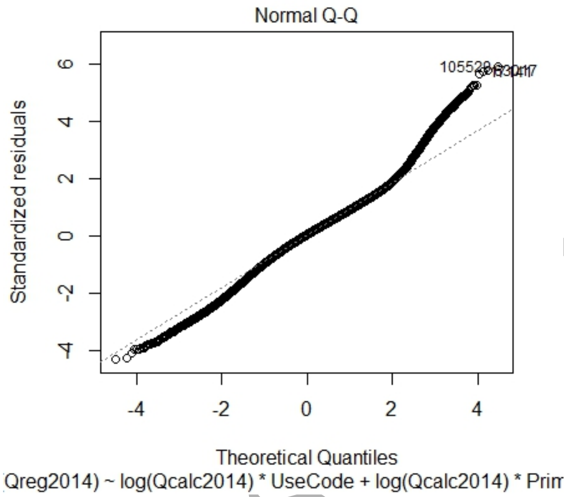
Figure A.7a showed considerable dependence between the residuals and the fitted values. Likewise, Figure A.7b indicated that the distribution of the residuals deviated somewhat from a normal distribution. Therefore, the model was deemed unsuitable.

Appendix A.2. Multiple linear regression model

The model validation plots of the MLR model (Figure A.8) showed considerably less bias in the residuals.



(a) Residuals vs. fitted values



(b) Normal quantile-quantile plot of the residuals

Figure A.8: Distribution of the MLR model residuals

Some observations, with large standardised residuals (in the upper right corner of Figure A.8b), did not follow a normal distribution, indicating that these could be outliers. However, a comparison of the parameter estimates from the MLR model with those from a robust regression model indicated that these observations did not compromise the parameter estimates unacceptably, see Appendix D.

Appendix B. Tests for statistical significance - ANOVA

In the present appendix, all ANOVA tables, used in testing for significant differences between groups of buildings, are listed.

Appendix B.1. Effects on heat consumption

In subsection 5.2, we tested for effects of the various building characteristics in Table 1. Below the ANOVA table is listed with the test for significance of the interaction terms.

Parameter	Df	Sum Sq	Mean Sq	F value	p-value
$\log(Q_{\text{calc}})$:Type	3	1137.5	379.17	3017.29	0
$\log(Q_{\text{calc}})$:PHS	1	21.46	21.46	170.76	0
$\log(Q_{\text{calc}})$:SHS	3	4.86	1.62	12.88	0
$\log(Q_{\text{calc}})$:Rent	1	22.19	22.19	176.54	0
Residuals	133951	16832.99	0.13		

Table B.8: ANOVA table MLR

In all cases the p-values were lower than the significance level in Equation 2 indicating that there was a significant effect of all building characteristics on the relationship between Q_{calc} and Q_{reg}.

Appendix B.2. Effects of building type

In the ANOVA table below, we compared each of the reduced models with the full model. The model names indicate which building types that were merged.

Model	Res DF	RSS	DF	Sum of Sq	F-value	p-value
Full model	133951	16832.989				
Farm + SFH	133953	16839.846	-2	-6.857	27.281	$1.427 \cdot 10^{-12}$
Farm + Row	133953	16856.179	-2	-23.19	92.268	$9.039 \cdot 10^{-41}$
Farm + MFH	133953	16864.126	-2	-31.137	123.888	$1.761 \cdot 10^{-54}$
SFH + Row	133953	17297.042	-2	-464.053	1846.385	0
Row + MFH	133953	17383.61	-2	-550.621	2190.824	0
SFH + MFH	133953	18423.652	-2	-1590.663	6328.968	0

Table B.9: ANOVA table Building types

These tests provided strong evidence that each of the four building types were significantly different from each other.

Appendix B.3. Effects of secondary heat supply

To test whether there was an effect of the individual secondary heating supply type, all levels were merged and tested aginst the full model.

Model	Res DF	RSS	DF	Sum of Sq	F-value	p-value
Full model	133951	16832.989				
SHS vs. no SHS	133955	16847.111	-4	-14.122	28.095	$2.315 \cdot 10^{-23}$

Table B.10: ANOVA table secondary heat supply

The p-value in Table B.10 provided strong evidence against no effect; i.e. the effect of having a secondary heat supply did depend on which type it was.

Appendix C. Parameter estiamtes MLR

The slope ($\log(Q_{calc})$) indicated a farmhouse with an individual boiler, no secondary heating, which was not rented out. All other parameters indicated a difference associated with a different building characteristic.

Parameter	0.5 % CI	Estimate	99.5 % CI	Std. Error	p-value
Intercept	2.749	3.732	4.715	0.291	$9.539 \cdot 10^{-38}$
$\log(Q_{\text{calc}})$	0.518	0.613	0.708	0.028	$4.992 \cdot 10^{-105}$
Type_{SFH}	0.137	1.125	2.112	0.292	0.0001166
Type_{Row}	-1.035	-0.038	0.958	0.294	0.8967
Type_{MFH}	-3.957	-2.959	-1.961	0.295	$1.137 \cdot 10^{-23}$
PHS_{DH}	-0.681	-0.545	-0.41	0.04	$1.877 \cdot 10^{-42}$
SHS_{El}	-0.544	-0.175	0.194	0.109	0.1085
$\text{SHS}_{\text{Stove}}$	-0.623	-0.397	-0.172	0.067	$2.583 \cdot 10^{-09}$
SHS_{Both}	-1.386	-0.715	-0.043	0.198	0.0003152
Rent_{Yes}	-0.713	-0.577	-0.441	0.04	$9.518 \cdot 10^{-47}$
$\log(Q_{\text{calc}})$:Type_{SFH}	-0.212	-0.116	-0.021	0.028	$3.933 \cdot 10^{-05}$
$\log(Q_{\text{calc}})$:Type_{Row}	-0.11	-0.013	0.083	0.029	0.64
$\log(Q_{\text{calc}})$:Type_{MFH}	0.205	0.301	0.398	0.029	$4.466 \cdot 10^{-26}$
$\log(Q_{\text{calc}})$:PHS_{DH}	0.04	0.054	0.068	0.004	$5.716 \cdot 10^{-41}$
$\log(Q_{\text{calc}})$:SHS_{El}	-0.025	0.012	0.048	0.011	0.2769
$\log(Q_{\text{calc}})$:$\text{SHS}_{\text{Stove}}$	0.017	0.039	0.062	0.007	$2.856 \cdot 10^{-09}$
$\log(Q_{\text{calc}})$:SHS_{Both}	-0.001	0.065	0.131	0.02	0.0009371
$\log(Q_{\text{calc}})$:Rent_{Yes}	0.039	0.053	0.066	0.004	$2.92 \cdot 10^{-40}$

Table C.11: Parameter estimates from the MLR model and the robust regression (RR) model respectively

Appendix D. Parameter estimates - model comparison

This appendix lists all parameter estimates from the MLR model and the robust regression (RR) model, as well as corresponding standard errors.

	MLR		**RR**	
Parameter	Estimate	Std. Error	Estimate	Std. Error
$\log(Q_{\text{calc}})$	0.613	0.028	0.587	0.027
$\log(Q_{\text{calc}})$:Type_{SFH}	-0.116	0.028	-0.1	0.027
$\log(Q_{\text{calc}})$:Type_{Row}	-0.013	0.029	-0.019	0.027
$\log(Q_{\text{calc}})$:Type_{MFH}	0.301	0.029	0.324	0.027
$\log(Q_{\text{calc}})$:PHS_{DH}	0.054	0.004	0.049	0.004
$\log(Q_{\text{calc}})$:SHS_{El}	0.012	0.011	0.016	0.01
$\log(Q_{\text{calc}})$:$\text{SHS}_{\text{Stove}}$	0.039	0.007	0.038	0.006
$\log(Q_{\text{calc}})$:SHS_{Both}	0.065	0.02	0.084	0.019
$\log(Q_{\text{calc}})$:Rent_{Yes}	0.053	0.004	0.044	0.004

Table D.12: Parameter estimates from the MLR model and the robust regression (RR) model respectively

Comparing the two models, the confidence intervals on the parameter estimates overlapped in all cases, indicating that the parameter estimates were not significantly different.

Appendix E. Residual analysis

This appendix is devoted to an analysis of the residuals, including homogeneity of the residual variance among groups and an analysis of potential bias caused by explanatory variables that were not included in the MLR model.

Appendix E.1. Residual variance

In an ANOVA the variance is assumed to be equal in the groups under consideration, an assumption which was violated in this case; see Table E.13.

Due to the large number of groups, we considered the smallest- and the largest variance in groups with less than ten observations, groups with more than 100 observations and groups with observations in between.

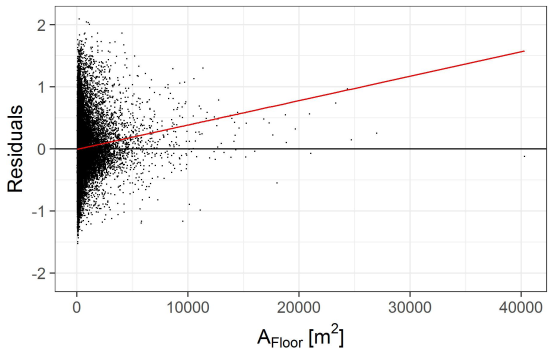
Group size	Min	Max
< 10	0.03	1.52
> 100	0.10	0.28
10-100	0.07	0.25

Table E.13: Min and max variance of the residuals of the MLR model in groups of different size

In particular the large variance in some of the groups with few observations compared with the small variance in some groups with many observations could cause an ANOVA to be invalid. However, because the test is robust with regards to violation of this assumption, we proceeded using it throughout the study.

Appendix E.2. Analysis of excluded explanatory variables

To ensure that variables not included in the model, in this case the heated floor area and the year of construction, did not cause a bias in the model, we plotted the residuals of the MLR model against each of these two variables.



(a) Residuals vs. heated floor area

(b) Residuals vs. construction year

Figure E.9: Distribution of the MLR model residuals

Despite a slight positive trend in Figure E.9a, the visual inspection of the residuals did not indicate any notable dependence.

In addition to the visual inspection, we fitted two simple linear regression models, in which we include the residuals as the dependent variable and each of the the heated floor area and the year of construction (YOC) as the independent variables respectively to quantify any bias, see Equation E.1 and Equation E.2.

$$Resid_{\text{MLR}} = \beta_0 + \beta_1 \cdot A_{\text{floor}} \tag{E.1}$$

$$Resid_{\text{MLR}} = \beta_0 + \beta_1 \cdot YOC \tag{E.2}$$

Model	**Intercept**	**Slope**	**Adj.** R^2
Equation E.1	$9.08 \cdot 10^{-3}$	$3.93 \cdot 10^{-5}$	0.004
Equation E.2	$7.21 \cdot 10^{-2}$	$3.68 \cdot 10^{-5}$	$5.97 \cdot 10^{-6}$

Table E.14: Regression line coefficients to quantify any dependence

Considering the parameter estimates, both regression lines suggested almost no dependence between the variables that were excluded from the analyses and the residuals, i.e. neither the heated floor area nor the year of construction explained much of the variation in the residuals (0.4 % and 0.0006 % respectively).

References

[1] H. Lund, J. Z. Thellufsen, S. Aggerholm, K. B. Wittchen, S. Nielsen, B. V. Mathiesen, B. Möller, Heat Saving Strategies in Sustainable Smart Energy Systems, International Journal of Sustainable Energy Planning and Management 04 (2014) 3–16. doi:10.5278/ijsepm.2014.4.2.

[2] S. Aggerholm, Cost-optimal levels of minimum energy performance requirements in the Danish Building Regulations, 1st Edition, SBi, Statens Byggeforskningsinstitut, Aalborg Universitet, Danish Building Research Institute, Aalborg University A. C. Meyers Vænge 15, DK-2450 Copenhagen SV E-mail sbi@sbi.aau.dk www.sbi.dk, DK-2450 Copenhagen SV, 2013.
URL http://sbi.dk/Pages/Cost-optimal-levels-of-minimum-energy-performance-requirements-in-the-Da aspx

[3] D. Majcen, L. Itard, H. Visscher, Theoretical vs. actual energy consumption of labelled dwellings in the Netherlands: Discrepancies and policy implications, Energy Policy 54 (2013) 125–136. doi:10.1016/j.enpol.2012.11.008.
URL http://www.sciencedirect.com/science/article/pii/S0301421512009731

[4] M. Delghust, A. Janssens, W. Roelens, T. Tanghe, Statistical study on the link between real energy use, official energy performance and inhabitants of low energy houses, 10th Nordic Symposium on Building Physics.

[5] P. van den Brom, A. Meijer, H. Visscher, Performance gaps in energy consumption: household groups and building characteristics, Building Research and Information 46 (1) (2018) 54–70. doi:10.1080/09613218.2017.1312897.
URL https://doi.org/10.1080/09613218.2017.1312897

[6] K. Gram-Hanssen, S. Georg, Energy performance gaps: promises, people, practices, Building Research and Information 46 (1) (2018) 1–9. doi:10.1080/09613218.2017.1356127.
URL https://doi.org/10.1080/09613218.2017.1356127

[7] P. de Wilde, The gap between predicted and measured energy performance of buildings: A framework for investigation, Automation in Construction 41 (2014) 40–49. doi:10.1016/j.autcon.2014.02.009. URL http://www.sciencedirect.com/science/article/pii/S092658051400034X{%}5Cnhttp://www.sciencedirect.com/science/article/pii/S092658051400034X/pdfft?md5=ba264f47082ab346b71eb8b403b4dd1f{&}pid=1-s2.0-S092658051400034X-main.pdf

[8] R. Galvin, The Rebound Effect in Home Heating, Routledge, Taylor & Francis Group, 2016.

[9] R. Haas, H. Auer, P. Biermayr, The impact of consumer behavior on residential energy demand for space heating, Energy and Buildings 27 (2) (1998) 195–205. doi:10.1016/S0378-7788(97)00034-0. URL http://linkinghub.elsevier.com/retrieve/pii/S0378778897000340

[10] H. Hens, W. Parijs, M. Deurinck, Energy consumption for heating and rebound effects, Energy and Buildings 42 (1) (2010) 105–110. doi:10.1016/j.enbuild.2009.07.017.

[11] M. Sunikka-blank, R. Galvin, Introducing the prebound effect : the gap between performance and actual energy consumption, Building Research & Information 3218 (October 2013) (2012) 260–273. doi:10.1080/09613218.2012.690952.

[12] H. Visscher, D. Majcen, L. Itard, Energy Saving Policies for Housing Based on Wrong Assumptions?, Open House International 39 (2) (2014) 78–83.

[13] O. Guerra Santin, L. Itard, H. Visscher, The effect of occupancy and building characteristics on energy use for space and water heating in Dutch residential stock, Energy and Buildings 41 (11) (2009) 1223–1232. doi:10.1016/j.enbuild.2009.07.002.

[14] D. Majcen, L. Itard, H. Visscher, Actual and theoretical gas consumption in Dutch dwellings: What causes the differences?, Energy Policy 61 (2013) 460–471. doi:10.1016/j.enpol.2013.06.018.

[15] D. Majcen, L. Itard, H. Visscher, Statistical model of the heating prediction gap in Dutch dwellings: Relative importance of building, household and behavioural characteristics, Energy and Buildings 105 (2015) 43–59. doi:10.1016/j.enbuild.2015.07.009.
URL http://linkinghub.elsevier.com/retrieve/pii/S0378778815301262

[16] V. Corrado, I. Ballarini, Refurbishment trends of the residential building stock: Analysis of a regional pilot case in Italy, Energy and Buildings 132 (March 2013) (2016) 91–106. doi:10.1016/j.enbuild.2016.06.022.
URL http://dx.doi.org/10.1016/j.enbuild.2016.06.022

[17] T. Csoknyai, S. Hrabovszky-Horváth, Z. Georgiev, M. Jovanovic-Popovic, B. Stankovic, O. Villatoro, G. Szendrő, Building stock characteristics and energy performance of residential buildings in Eastern-European countries, Energy and Buildings 132 (2016) 39–52. doi:10.1016/j.enbuild.2016.06.062.

[18] J. Kragh, K. Wittchen, Development of two Danish building typologies for residential buildings, Energy and Buildings 68 (2014) 79–86. doi:10.1016/j.enbuild.2013.04.028.
URL http://linkinghub.elsevier.com/retrieve/pii/S037877881300604X

[19] C. A. Balaras, E. G. Dascalaki, K. G. Droutsa, S. Kontoyiannidis, Empirical assessment of calculated and actual heating energy use in Hellenic residential buildings, Applied Energy 164 (2016) 115–132. doi:10.1016/j.apenergy.2015.11.027.
URL http://dx.doi.org/10.1016/j.apenergy.2015.11.027

[20] European Commision, Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings (recast) (2010). doi:doi:10.3000/17252555.L_2010.153.eng.

[21] Dansk Standard, DS/EN 15316-2-3 - Heating systems in buildings - Method for calculation of system energy requirements and system efficiencies - Part 2-3: Space heating distribution systems (2007).

[22] European Standard, EN 15316-3-2 Heating systems in buildings - Method for calculation of system energy requirements and system efficiencies - Part 3-2: Domestic hot water systems, distribution (2007).

[23] E. N. ISO, 13790: Energy performance of buildings–Calculation of energy use for space heating and cooling (EN ISO 13790: 2008), European Committee for Standardization (CEN), Brussels 2006 (50).

[24] S. Aggerholm, K. Grau, Buildings' energy demand: calculation guide (only available in Danish: Bygningsers energibehov (Be10)). SBi Direction 213, Danish Building Research Institute, Aalborg University, Copenhagen, Denmark, 2014.

URL https://www.statsbiblioteket.dk/au/{#}/search?query=recordID:{%}22sb{_}6144776{%}22
[25] J. Kragh, J. Rose, H. N. Knudsen, O. M. Jensen, Possible explanations for the gap between calculated and measured energy consumption of new houses, Energy Procedia 132 (2017) 69–74. doi:10.1016/j.egypro.2017.09.638.
URL http://linkinghub.elsevier.com/retrieve/pii/S1876610217347859
[26] M. Brøgger, K. B. Wittchen, Flexible building stock modelling with array-programming (in press), in: Fifteenth International IBPSA Conference, International Building Performance Simulation Association, San Francisco, 2017, p. 10.
[27] K. Gram-hanssen, D. Building, Technology and Culture A s Explanations for Variations in Energy Consumption Social Construction of Technology, in: Proceedings of ACEEE 2002 Summer Study on Energy Efficiency in Buildings, Asilomar, Pacific Grove, CA, 2002, pp. 79–90.
[28] D. M. Diez, C. D. Barr, OpenIntro Statistics Third Edition, third edit Edition, openintro.org, 2015.
URL https://www.openintro.org/stat/textbook.php?stat{_}book=os
[29] UCLA: Statistical Consulting Group., ROBUST REGRESSION — R DATA ANALYSIS EXAMPLES.
URL https://stats.idre.ucla.edu/r/dae/robust-regression/
[30] P. E. College of Science, Robust Regression Methods.
URL https://onlinecourses.science.psu.edu/stat501/node/353
[31] G. E. P. Box, D. R. Cox, An Analysis of Transformations Revisited, Rebutted, Journal of the American Statistical Association 77 (377) (1982) 209. arXiv:arXiv:1011.1669v3, doi:10.2307/2287791.
URL http://www.jstor.org/stable/2287791?origin=crossref
[32] W. N. Venables, B. D. Ripley, Modern Applied Statistics with S, 4th Edition, Springer, New York, 2002, iSBN 0-387-95457-0.
URL http://www.stats.ox.ac.uk/pub/MASS4
[33] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria (2013).
URL http://www.R-project.org/
[34] K. Benoit, Linear Regression Models with Logarithmic Transformations (2011).
URL http://www.kenbenoit.net/courses/ME104/logmodels2.pdf
[35] F. G. Li, A. Z. Smith, P. Biddulph, I. G. Hamilton, R. Lowe, A. Mavrogianni, E. Oikonomou, R. Raslan, S. Stamp, A. Stone, A. J. Summerfield, D. Veitch, V. Gori, T. Oreszczyn, Solid-wall U -values: Heat flux measurements compared with standard assumptions, Building Research and Information 43 (2) (2015) 238–252. doi:10.1080/09613218.2014.967977.
URL http://dx.doi.org/10.1080/09613218.2014.967977
[36] H. Wickham, ggplot2: Elegant Graphics for Data Analysis, Springer-Verlag New York, 2009.
URL http://ggplot2.org