Energy Efficiency (2019) 12:139–155
https://doi.org/10.1007/s12053-018-9664-2

ORIGINAL ARTICLE

Energy performance of buildings: A statistical approach to marry calculated demand and measured consumption

Michael Hörner · Markus Lichtmeß

Received: 22 November 2017 / Accepted: 8 April 2018 / Published online: 3 May 2018

Abstract In public debate, Energy Performance Certificates (EPCs) of buildings have been criticised for not reflecting the energy demand realistically. And indeed, measurement, as in energy bills, usually differs from the calculation, in particular, when simplified energy performance calculation models and standard specifications are applied, as in EPCs. Thus, energy-saving potentials of refurbishment recommendations and their cost-effectiveness tend to be over-estimated. Of course, this is not desirable. These effects were analysed in two sets of data, the Energy Performance Certificate Register for residential buildings in Luxemburg, run by the Luxemburg Ministry of the Economy (Lichtmeß, 2012) and a database gathered in the research project "Teilenergiekennwerte von Nichtwohngebäuden (TEK)" (Hörner et al., 2014a) funded by the German Federal Ministry of Economic Affairs and Energy. Multiple linear regression and error calculus were applied to study the gap between measurement and various calculation models in detail. A statistical procedure is proposed to estimate expectation value and variance of the future energy consumption of buildings in case of refurbishment, as a supplement to standard calculations in EPCs for example. Prerequisite is that for a sufficient number of buildings, data on both, measured energy consumption and calculated demand, are available.

Keywords Energy performance of buildings · Multiple linear regression · Errors-in-variables model · Calculated energy demand · Measured energy consumption · Calibration

M. Hörner (✉)
Institut Wohnen und Umwelt (IWU), Rheinstraße 65, 64295 Darmstadt, Germany
e-mail: m.hoerner@iwu.de

M. Lichtmeß
Goblet Lavandier & Associés Ingénieurs-Conseils S.A. (GoLav), 17, rue J.P. Sauvage, L-2514 Luxembourg, Luxembourg
e-mail: markus.lichtmess@golav.lu

Are we calculating wrong?

In 2002, the European Commission approved the first version of the Energy Performance of Buildings Directive (EPBD). The Directive required the EU Member States (MS) to introduce energy performance certification of buildings and regular inspections of heating and cooling systems. Energy Performance Certificates (EPCs) inform consumers on the energy efficiency of buildings and recommend improvements. Certificates are to provide this information in case of construction, sale or rental of buildings. All EU MS established control systems to ascertain compliance with the requirements of their corresponding energy savings ordinance.

But, to assess the energy-related quality of buildings is a complex task. There are two possible approaches: operational rating based on measured consumption taking the building as it is used or asset rating based on calculated demand appropriate to compare different buildings with standard users. Needless to say that calculated energy demand, in particular of heating and domestic hot water preparation, in buildings usually differs from the measured consumption. Deviations are substantial in many cases, though, and building owners and users do not really understand the difference in numbers between an EPC and, say, an energy bill. In the recent literature, this observation has been attributed to what is called prebound or rebound effects (Sunikka-Blank & Galvin, 2012).

Supposed, our meters work fine, what are we then doing wrong in the calculation? The physical processes of heat transfer are theoretically well understood, and validated simulation tools are at hand. However, these tools require quite some computing power and a considerable amount of input data. In particular for existing buildings, to specify the real variables of user behaviour and schedules, is almost impossible often times. Thus, simplified calculation models have been invented together with different approaches to parameters of the building, its use and weather depending upon the goal of the assessment. Standard specifications or rather assumptions on the safe side are to be applied in normative assessments, as in EPCs. In energy consulting, parameters of user behaviour and the physical building should be assumed realistically in order to calculate savings potentials reliably. Scenario calculations for building stocks on the other hand use building typologies and assumptions for the physical building and user parameters that represent averages within the usual spread of a building class.

Either way, however, deviation between measurement and calculation has been discussed in the corresponding literature (Erhorn, 2007; Knissel et al., 2006; Gruber et al., 2005) and it appears in a typical manner, as depicted in Fig. 1, the Luxemburg EPC Register (Lichtmeß, 2012), run by the Luxemburg Ministry of the Economy, taken as an example: There is a tremendous spread of measured values to a given calculated demand and for existing not yet modernised buildings measured consumption tends to be much lower than calculated, quite often more than factor 2. For refurbished or new buildings, on the other hand, the Luxemburg data are not indicative enough, but there is some evidence (Graf, 2016) that in very energy-efficient buildings with specific values of energy demand below 50 kWh/m^{2}a, this effect is reversed and calculation tends to be lower than measurement.

Provided our calculation scheme was ideal and all input variables had their real values, we would expect the data points to line up along the bisecting line. A thorough and realistic choice of input variables, in particular in energy-consulting, alleviates the problem of deviations but uncertainties on the input side remain. We shall demonstrate that later, analysing 92 non-residential building records in the TEK database, gathered in the research project "Teilenergiekennwerte von Nichtwohngebäuden (TEK)" (Hörner et al., 2014a) funded by the German Federal Ministry of Economic Affairs and Energy.

Probably, this deviation corresponds to both different user habits before and after refurbishment not adequately accounted for in the specification of user parameters and uncertain estimates of physical building parameters. Empirical data on user behaviour (Schröder et al., 2014) clearly show that occupants of older not yet modernised buildings do use energy more consciously and, e.g., set room temperatures in a way that the average temperatures in buildings result significantly lower than specified in standards, fuel costs assumably being a strong economic incentive.

Of course, this deviation between measurement and calculation is not desirable, since calculated energy savings of refurbishment recommendations and their cost-effectiveness are over-estimated. In public debate, the acceptance of EPCs and energy performance calculations, in general, is at stake.

Thus, simplified calculation models are not intrinsically flawed, but they are different in their approach to uncertain input variables and boundary conditions. There is no right or wrong but different goals. Still, like any model in physics, also energy performance calculations are supposed to be compared to measured and, supposedly, real values of energy consumption, as has been proposed earlier, e.g. in (Casties, 1997; Loga et al., 2003), and the range of uncertainty should be specified. A calibration procedure is needed.

Data bases

The Luxemburg Energy Performance Certificates Register

The EPC in Luxemburg defines a threefold labelling scheme for new and existing residential buildings with regard to energy need for heating, primary energy

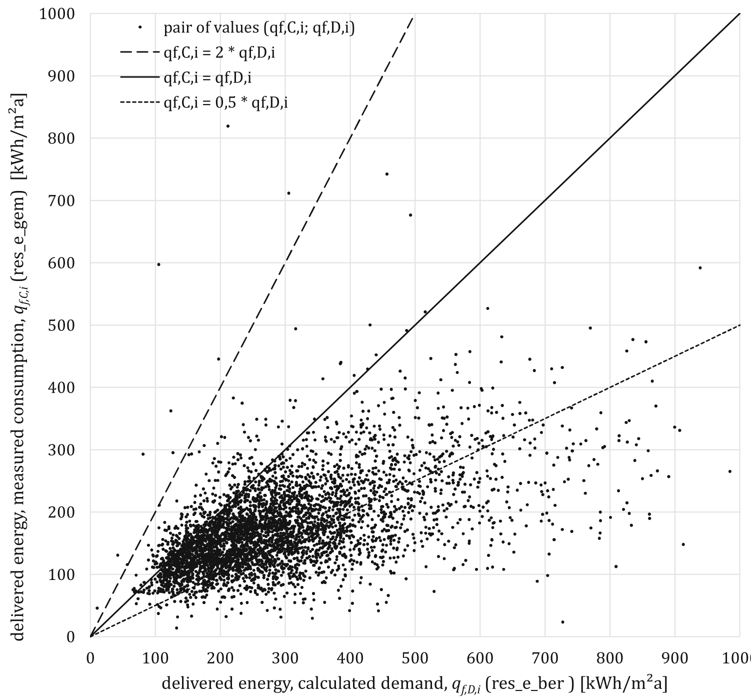


Fig. 1 Measured energy consumption (heating and domestic hot water) $q_{f,\,C,\,i}$ plotted against calculated demand $q_{f,\,D,\,i}$ with standardised specifications of the 4.407 EPCs selected from Luxemburg EPC Register. Data Source: Luxemburg EPC Register, Luxemburg Ministry of the Economy, IWU (reg2_DB6_LN)

demand and CO_2-emissions. Besides the certification, that legal requirements are met, its objective is to give reliable information on the energy standard of buildings, e.g. in sales ads. Only approved experts are eligible to issue EPCs. The EPC Luxemburg is based on a simplified demand calculation model with standardised parameters of user behaviour and climate. Remarkable is the fact, that for each existing building i, in addition to the calculated delivered energy demand $q_{f,D,i}$, the measured consumption $q_{f,C,i}$ has to be indicated as soon as an EPC is issued, for new buildings 4 years after commissioning. However, data on actual user behaviour and climate on the building site are not recorded in the database.

In 2014, the Luxemburg EPC Register was launched. EPCs issued after that date have to be registered. Since then, a database of ten thousands of EPC records has grown, each one consisting of 174 parameters. The EPC Register increasingly provides for an overview of the energy standard in the Luxemburg building stock and supports statistical analysis.

For the statistical analysis, here, only anonymised records of existing buildings were taken into account, for which calculated demand, as well as measured consumption for heating and domestic hot water, was available. A comprehensive quality control of EPCs has not yet been established, in particular concerning measured consumption, where not even weather correction can be assured. Plausibility checks have been built into the EPC software tool so that grossly deficient or incomplete EPCs could be sorted out beforehand. Thus, 4.407 EPC records of residential buildings in Luxemburg were selected from the Luxemburg EPC Register for the statistical analysis.

The TEK database

The research project "Teilenergiekennwerte von Nichtwohngebäuden" (TEK) (Hörner et al., 2014a) within the ENOB research program of the Federal Ministry for Economic Affairs and Energy (BMWi) delivered a database consisting of 92 records of existing non-residential buildings, and the TEK tool (Hörner et al., 2014b) for demand calculations was developed, based on the general terms of German standard DIN V 18599 (DIN Deutsches Institut für Normung e.V.: DIN V 18599, 2016), though implementing various simplifications in the algorithms. The TEK database contains detailed information on properties of the building envelope, energy efficiency of technical installations, user behaviour, measured energy consumptions and calculated energy demands.

Table 1 Distribution of the 92 non-residential buildings in the TEK database for different classifications

Age band	Number of buildings	Net floor area	Number of buildings	Use	Number of buildings
Before 1918	10	Up to 1.000 m^2	3	Office	23
1919–1948	5	1.001 to 5.000 m^2	36	Trade	11
1949–1977	38	5.001 to 10.000 m^2	29	University	19
1978–1994	26	10.001 to 30.000 m^2	20	Hotel	8
1995–2001	7	> 30.000 m^2	4	School	15
After 2002	6			Event	16

The data cover a wide variety of age bands, net floor area and building uses as shown in Table 1.

In order to study the consequences of standard versus realistic specifications of input variables, three different calculation schemes were applied to the building sample:

- *Real-real* with both, data on user behaviour and physical building parameters, collected on site as realistic as possible,
- *Std-real* with standard specifications on user behaviour as of German standard DIN V 18599-10: 2011-12 (DIN Deutsches Institut für Normung e.V.: DIN V 18599-10: 2011-12, 2011) and
- *Std-simplified* with a simplified model of the building envelope geometry in addition to standard specifications of user behaviour.

The results in Fig. 2(B-1 to B-3) demonstrate how the data points regroup gradually towards the bisecting line the more realistic the input data are chosen. In Fig. 2(B-4), this effect is made more visible, the gradients of the best fit straight lines of the corresponding data approaching the bisecting line the more the “better” data are.

Thus, it seems to be worthwhile to gather more realistic input data in order to get more valid calculation results. But a considerable effort on site is required not easy to provide for many buildings. In the course of this paper, we shall develop another approach to “marry” calculation to measurement, taking advantage of detailed information on the building from the first and extracting more information of the effects of user behaviour intrinsically contained in the latter.

Methodology

In the first part of this paper, a normative calculation scheme, the Luxemburg EPC for residential buildings (Le gouvernement du grand-duché de Luxembourg, 2007), will be amended with information on measurement results from the EPC Register. We shall apply statistical methods to allow for the prognosis of future energy consumption as an expectation value, the variance specifying the range of uncertainty.

As mentioned before, descriptive approaches in consulting and scenarios will take advantage of these procedures as well, as will be demonstrated in the second part, when we undertake regression analyses of 92 non-residential building records in the TEK database (Hörner et al., 2014a). We will derive calibration factors for demand calculations in the Std-simplified calculation scheme of the TEK tool (Hörner et al., 2014b), exemplarily, to be used in energy consultancy and scenario calculations in the German non-residential building stock.

Regression analysis

We want to forecast delivered energy consumption of heating and domestic hot water for new buildings or buildings after refurbishment. Regression analysis is a statistical method that generally allows predicting the value of a dependent variable depending upon one or more independent variables.

We consider the area-related delivered energy demand $q_{f,D}$ of a building as a function of U values U_i and areas A_i of i different parts of the envelope, expenditure factors e_j of j heat generators and distributions, k parameters of user behaviour B_k, such as room temperatures, and l climate parameters C_l, e.g. the ambient temperature.

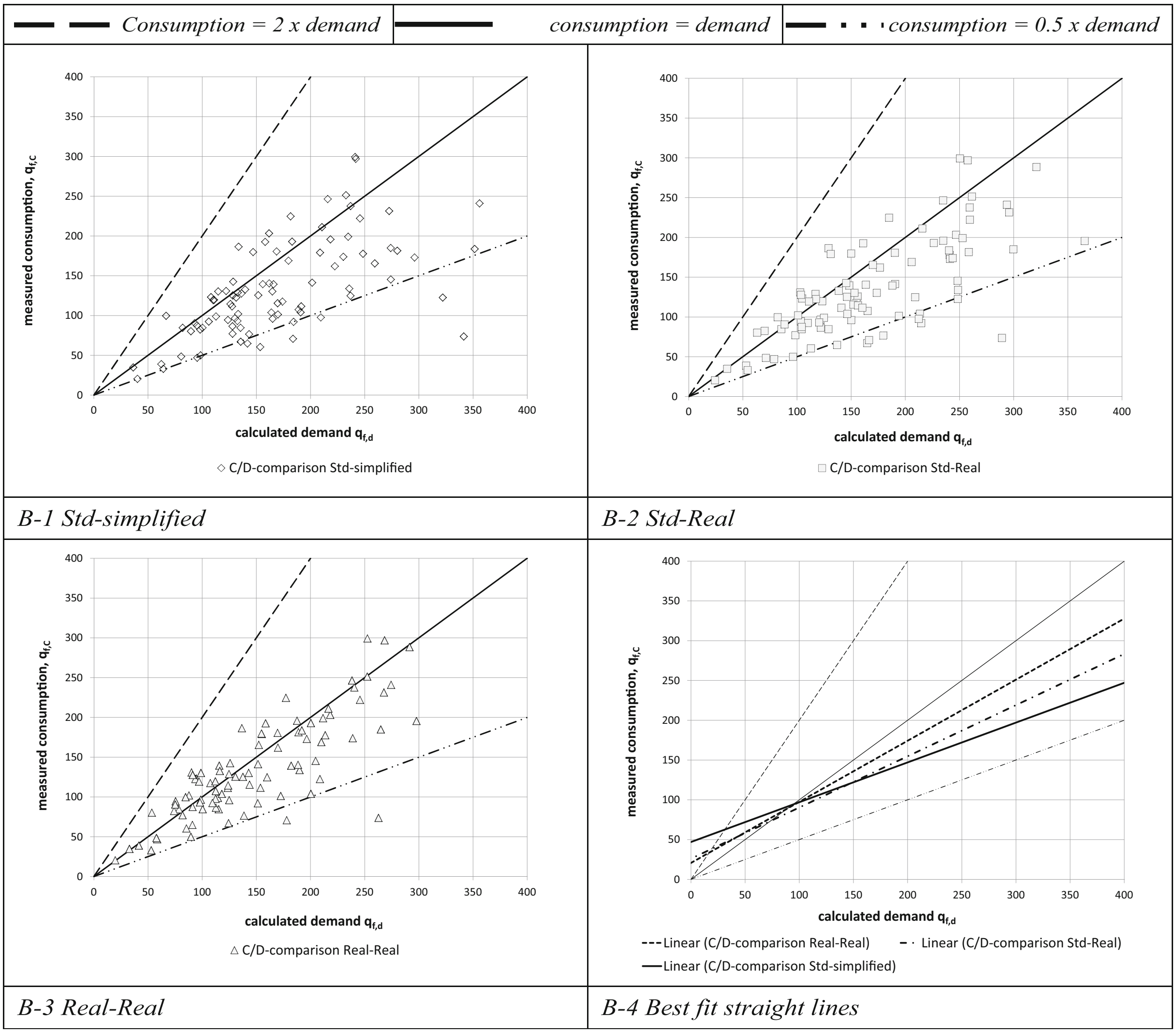


Fig. 2 Plot of measured consumption over calculated demand of the buildings' delivered energy for heating + domestic hot water for three calculation schemes: Std-simplified (B-1), Std-real (B-2), real-real (B-3) and best fit straight lines (B-4). Data Source: TEK-Database, IWU (2015625_QSA-Regression-E/D-C-comp)

$$q_{f,D} = q_{f,D}\left(U_i, A_i, e_j, B_k, C_l\right) \quad (1)$$

The demand calculated according to EPC specifications may then be considered as a special subclass of function (1) with standardised parameters of usage and weather

$$q_{f,D}^{EPC} = q_{f,D}^{EPC}\left(U_i, A_i, e_j, B_k^{std}, C_l^{std}\right) \quad (2)$$

We further suppose that measured delivered energy consumption of buildings $q_{f,\,C}$ is a function of the same variables as demand many of which are well represented in a standardised demand calculation like the EPC, $q_{f,D}^{EPC}$, and other independent variables X_m, to be identified yet

$$q_{f,C} = q_{f,C}\left(q_{f,D}^{EPC}, X_1, \ldots, X_m\right) \quad (3)$$

Since the exact function (3) is unknown, we guess from Fig. 1 that $q_{f,\,C}$ is directly proportional to $q_{f,\,D}$ and from physical reasoning of cause and effect we expect a linear dependence. Thus, we model $q_{f,\,C}$ to be a linear function of $q_{f,D}^{EPC}$

$$q_{f,C}\left(q_{f,D}^{EPC}, X_1, \ldots, X_m\right) = b_0 + b_1 \cdot q_{f,D}^{EPC} + u \quad (4)$$

plus a quantity of uncertainty u, which accounts for all influences of user behaviour, weather and the calculation model so far unknown or not quantified. Since the two coefficients b_k are unknown, Eq. (4) cannot be solved.

Based on a sample of N pairs of values $\left(q_{f,D}^{EPC}, q_{f,C}\right)$ out of a principally infinite population, a linear function

$$\hat{q}_{f,C} = \beta_0 + \beta_1 \cdot q_{f,D}^{\mathrm{EPC}} \quad (5)$$

can be estimated with regression coefficients β_k being estimates of the unknown coefficients b_k. This leaves the difference between the measured consumption and its estimate, called residue,

$$\left|q_{f,C,i} - \hat{q}_{f,C,i}\right| = \hat{u}_i,\ i = 1, \cdots, N \quad (6)$$

representing the amount that cannot be explained by the function $\hat{q}_{f,C}$ alone. The regression coefficients β_k are most commonly determined by the method of ordinary least squares (OLS), minimising the sum of squares of all residues $\hat{u}_i$ of the sample.

For any building, with the calculated demand $q_{f,D,i}^{EPC}$ given, Eq. (5) then renders the best possible estimate $\hat{q}_{f,C,i}$ of the so far unknown "true" consumption $q_{f,C,i}$. For the sample of building values from the Luxemburg EPC Register, the estimation function is

$$\hat{q}_{f,C} = 92 + 0,28 \cdot q_{f,D}^{EPC} \quad (7)$$

as illustrated in the graph in Fig. 3. For an examplary building with a calculated demand $q_{f,D,i}^{EPC} = 200 \left[\frac{kWh}{m^2a}\right]$, the estimated consumption would turn out to be $\hat{q}_{f,C,i} = 148 \left[\frac{kWh}{m^2a}\right]$. Three facts are remarkable about this function: The slope is considerably lower than expected, the axis intercept is unreasonably high and the range of uncertainty, illustrated by the standard error σ, seems to not represent the spread of the data points very well. This estimate seems to not approximate the data very well.

Heteroscedasticity

A special feature of the distribution in Fig. 3 is heteroscedasticity, that means the spread in residues $\hat{u}_i$ correlates with the independent variable $q_{f,D}^{EPC}$, as is illustrated by the two-sided arrows and as, indeed, the results of a Goldfeld-Quandt-Test[1] confirm. This breaches one of the premises of linear regression and, subsequently, the standard error is estimated too large for low demands and too small for high demands.

This deficiency of the data can be cured by a non-linear, mostly logarithmic, transformation of the variables, the result of which is plotted in Fig. 4. The data points are spread much more homogenously, the sample of logarithmic pairs of value $\left(Ln\ \hat{q}_{f,C,i}, Ln\ q_{f,D,i}^{EPC}\right)$ is homoscedastic and thus appropriate for linear regression

$$Ln\ \hat{q}_{f,C} = \beta_0 + \beta_1 \cdot Ln\ q_{f,D}^{EPC} \quad (8)$$

We get the estimate of the future consumption $\hat{q}_{f,C}$ as the numerus by raising the base e, Euler's Number, to the power $Ln\ \hat{q}_{f,C}$ as in the following equation

$$\hat{\mathrm{q}}_{\mathrm{f,C}} = \mathrm{e}^{\beta_0 + \beta_1 \cdot \mathrm{Ln}\ \mathrm{q}_{\mathrm{f,C}}^{\mathrm{EPC}}} = \left(\mathrm{q}_{\mathrm{f,D}}^{\mathrm{EPC}}\right)^{\beta_1} \cdot \mathrm{e}^{\beta_0}. \quad (9)$$

For the sample of building values from the Luxemburg EPC Register, the estimation equation turns into

$$\hat{\mathrm{q}}_{\mathrm{f,C}} = \left(\mathrm{q}_{\mathrm{f,D}}^{\mathrm{EPC}}\right)^{0,49} \cdot \mathrm{e}^{2,37} \quad (10)$$

plotted in Fig. 5 and a building delivered energy demand of $q_{f,D,i}^{EPC} = 200 \left[\frac{kWh}{m^2a}\right]$ corresponds to a supposed consumption of $\hat{q}_{f,C,i} = 142 \left[\frac{kWh}{m^2a}\right]$.

Testing the fit

The reliability of this prediction may be assessed by various tests (Backhaus et al., 2005). For our purposes, the standard error $\sigma\left(\hat{q}_{f,C,i}\right)$ is of particular importance since it is a measure for the standard deviation of the estimated consumption, corresponding to the dispersion of the estimators of different samples about the "true" consumption. Unfortunately, the latter is not known until it has been measured. However, regression analysis can not only give an estimate for the unknown consumption but furthermore, the standard error can also be estimated $\hat{\sigma}\left(\hat{q}_{f,C,i}\right)$ meaning that: With a probability

[1] The Goldfeld-Quandt-Test compares sample-variances s^2 of two subsamples, e.g. s^2_{low} of the lower and s^2_{up} of the upper half of the observations. If $s^2_{low} / s^2_{up} > F_{crit}$, a critical value of the F-distibution, then a sample is considered heteroscedastic.

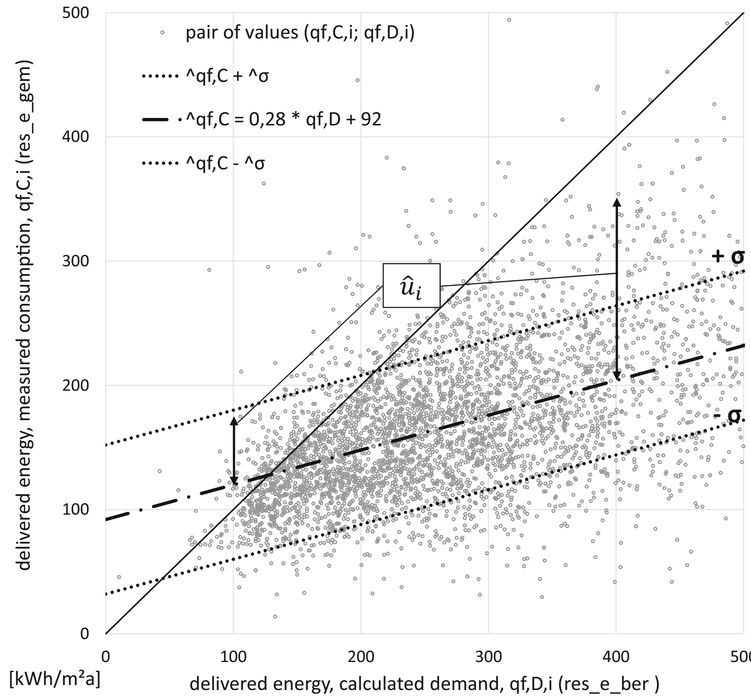


Fig. 3 Plot of the estimation function (7) using the 4.407 EPCs selected from the Luxemburg EPC Register. The fact that the spread of residues $\hat{u}_i$ depends upon the independent variable $q_{f,D,i}^{EPC}$ is illustrated by the two-sided arrows. Data Source: Luxemburg EPC Register, Luxemburg Ministry of the Economy, IWU (reg2_DB6_LN)

of 68% the "true" but unknown future consumption $q_{f,C,i}$ of building i will be in the interval $\hat{q}_{f,C,i} \pm \hat{\sigma}\left(\hat{q}_{f,C,i}\right)$, given the calculated demand $q_{f,D,i}^{EPC}$ with standard specifications from an EPC of the building.

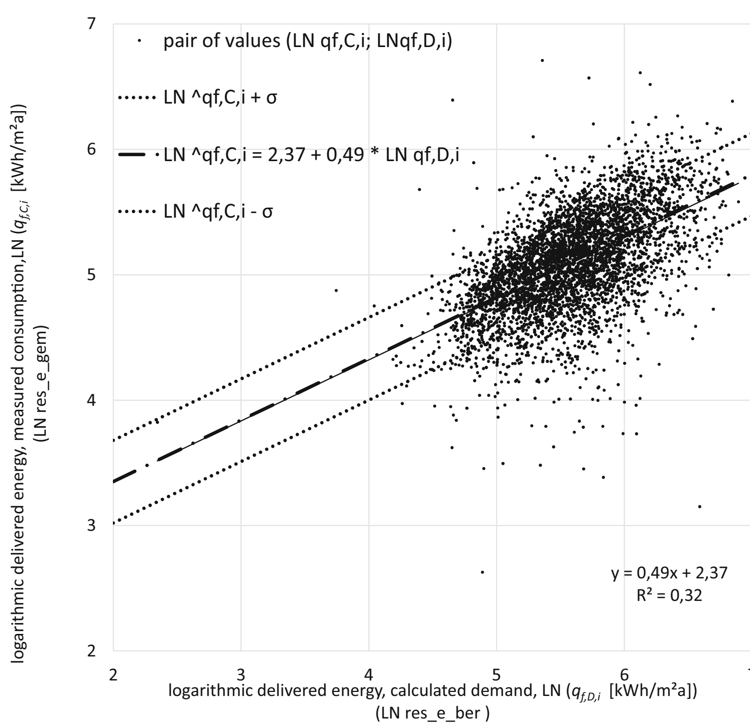


Fig. 4 Plot of the logarithmic pairs of values and of the estimation function (8) including transformed standard error σ. Data Source: Luxemburg EPC Register, Luxemburg Ministry of the Economy, IWU (reg4_DB6_LN)

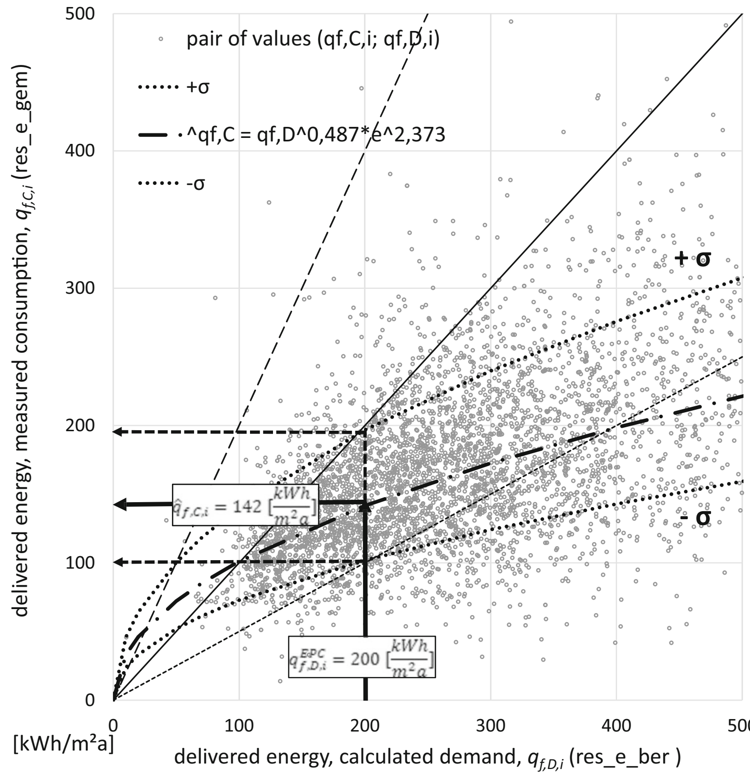


Fig. 5 Plot of the reversely transformed estimation function (10) and standard error illustrating case example (12). Data Source: Luxemburg EPC Register, Luxemburg Ministry of the Economy, IWU (reg4_DB6_LN)

In the above example of a building with $q^{EPC}_{f,D,i} = 200 \left[\frac{kWh}{m^2 a}\right]$, ($Ln\ q^{EPC}_{f,D,i} = 5,30$), the future logarithmic consumption will be in the range of

$$\begin{aligned} \mathrm{Ln}\ \hat{q}_{f,C,i} &= \beta_0 + \beta_1 \cdot \mathrm{Ln}\ q^{EPC}_{f,D,i} \pm \hat{\sigma}\left(\mathrm{Ln}\ \hat{q}_{f,C,i}\right) \\ &= 4,97 \pm 0,33 \end{aligned} \tag{11}$$

which reversely transforms into an asymmetric error band

$$\begin{aligned} \hat{q}_{f,C,i} &= \left(q^{EPC}_{f,D,i}\right)^{\beta_1} \cdot e^{\beta_0} \cdot e^{\pm\hat{\sigma}\left(\mathrm{Ln}\hat{q}_{f,C,i}\right)} \\ &= 142 \begin{cases} +56 \\ -40 \end{cases} \left[\frac{kWh}{m^2 a}\right] \end{aligned} \tag{12}$$

as illustrated in Fig. 5 showing an error band seemingly in more accordance with the growing spread in measured consumption values.

Another most common test is the coefficient of determination R^2, giving the ratio of the spread explained by the regression model and the total dispersion of the sample. In Fig. 4, an $R^2 = 32\%$ reflects the fact that a considerable amount of the spread cannot be explained by this model.

The p value test tells whether a variable is statistically significant. The F test assesses the significance of the whole regression model.

Regression in the errors-in-variables model

Measurement error in the independent variables

As the deviation between measurement and calculation appears in the typical manner, described above and in many other publications, we look for a more quantitative explanation than just saying it is the user. Physical variables can only be measured with finite precision; the calculation model then contains measurement error. What does that mean for the above regression analysis?

Sometimes measurement error is considerable, for example when we look at the standardised parameters of user behaviour. Of course, there are numerous other factors contributing uncertainties to the calculated demand like heat transfer coefficients of the building envelope, floor and envelope areas, weather conditions, etc. (Chari et al., 2017). For most of the input variables, frequency distributions have been established characterising their range of uncertainty (Brohus & Heiselberg, 2009; Santos

Silva & Ghisi, 2014). Of particular importance are, according to (Brohus & Heiselberg, 2009; Corrado & Mechri, 2009), heating room temperatures, natural ventilation in winter, hot water consumption and equipment heat gains, i.e. parameters of user behaviour.

On the other hand, EPCs are available for many buildings and their number is supposed to be growing until at some point in the future an EPC has been issued for every building. For building stock modelling, EPCs are a good source of information. Thus, for the sake of this analysis, we focus on the case of an EPC calculation and assume that, ideally, standard specifications of user behaviour parameters B_k^{std} in EPCs should have been defined as mean values of frequency distributions from random samples of measured values. Thus, we identify the standard specifications with the true mean of $\overline{B}_k$ and its known error $\sigma(B_k)$

$$B_k^{std} = \overline{B}_k \pm \sigma(B_k) \tag{13}$$

We estimate the error propagation of these uncertainties on calculated demand from the Gaussian Law of Error Propagation, which is strictly speaking valid only for normal distributions of user behaviour parameters, an assumption being supported by other authors (Brohus & Heiselberg, 2009) for most of these parameters.

$$\sigma\left(q_{f,D}^{EPC}(B_k)\right) = \sqrt{\sum_k \left(\frac{\partial q_{f,D}^{EPC}}{\partial B_k}\cdot\sigma(B_k)\right)^2}, k = 1, \cdots, K \tag{14}$$

In order to account for the spread in user behaviour variables within our regression model, we consider an errors-in-variables model in the independent variables (Wooldridge, 2003) at least for the calculated EPC demand. We assume that $q_{f,\,D,\,i}$ was the true demand of building i. It is unknown though, since actual user behaviour is unknown. We further assume $q_{f,D,i}^{EPC}$ to be our best measure of the true demand, but it comes with a "measurement" error σ_i

$$q_{f,D,i}^{EPC} = q_{f,D,i} + \sigma_i \tag{15}$$

due to the spread in user behaviour. Let us consider regression Eq. (8) with $Ln\left(q_{f,D,i}^{EPC}\right)$ being the only independent variable. Using the asymptotic of the ordinary least squares estimate, the amount of inconsistency due to measurement errors can be determined as

$$\lim_p(\beta_1) = b_1\cdot\frac{\sigma^2(Ln(q_{f,D,i}))}{\sigma^2(Ln(q_{f,D,i})) + \sigma^2(\Sigma_i)} < b_1 \tag{16}$$

Hence, the estimator of the slope coefficient, β_1, is always smaller in magnitude than the true value b_1 since variances are non-negative and the reliability ratio $\lambda = \frac{\sigma_{true}^2}{\sigma_{true}^2+\sigma_\Sigma^2} < 1$ always is smaller than 1. This is called the attenuation bias. This explains partly what we observe in Fig. 4: A shallow slope of the estimation function instead of the physically expected bisecting line with $b_1 \sim 1$. Thus, a great variance in the independent variable $Ln(q_{f,\,D,\,i})$ compensates partly for the "measurement" error, from the value of the slope estimator $\beta_1 = 0.49$ in this case we can infer about the same size of both quantities.

Matters of the errors-in-variables model become more intricate with more independent variables, details lie beyond the scope of this paper.

Independent studies have to be undertaken in order to reduce uncertainty in user behaviour parameters and subsequently in calculated demand, thereby reducing the attenuation bias. Unbiased regression functions are useful tools to calibrate calculated demands from simplified models to measured consumptions, as shown in the following chapter on non-residential buildings. This is particularly important in scenarios when reliable consumptions including uncertainties are supposed to be predicted considering different paths of refurbishment action in the building stock in the future.

Measurement error in the dependent variable

To measure the energy consumption of a building that is "true" in the sense that it may be compared to the corresponding demand calculation often is quite difficult too, due to inappropriate metering equipment, different refurbishment status, unknown weather conditions, temporary vacancies, etc. Thus, we are confronted with measurement error in the dependent variable as well.

Let $q_{f,C}^*$ be the true consumption and $q_{f,\,C}$ our best measure with a "measurement" error e_0, such that $q_{f,C}^* = q_{f,C} - e_0$, we obtain the estimable model

$$q^*_{f,C} = b_0 + b_1 \cdot q^{EPC}_{f,D} + u \Rightarrow q_{f,C} = b_0 + b_1 \cdot q^{EPC}_{f,D} + u + e_0 \quad (17)$$

It is reasonable to assume that the measurement error e_0 is statistically independent of each explanatory variable. Then, the estimation in (17) is unbiased and consistent. It may also be assumed that e_0 and u are independent since there are many reasons for the calculated demand to deviate from the measured consumption not only measurement error in the latter. Given this, the measurement error in the dependent variable results in a larger error variance

$$\mathrm{Var}(u + e_0) = \sigma^2_u + \sigma^2_{e_0} > \sigma^2_u \quad (18)$$

And this larger variance is propagated to all regression estimators. The only thing one can do about it is to measure energy consumption very carefully, for example in well-designed empirical surveys focussing on measured consumption and its interdependence with user behaviour.

Statistical analysis of energy performance certificates for residential buildings in Luxemburg

Multiple linear regression analysis of calculated demand and measured consumption

The requirement to indicate measured consumption in addition to the calculated demand in the Luxemburg EPC, of course, turned public awareness to the fact that often times there is a considerable delta between these two quantities. An empirically well-based estimation function would allow for the prediction of presumed consumption and its standard error in the EPC pointing out to building owners that individual user behaviour may cause a considerable deviation of the actual consumption from the calculated demand. Not to mention the effects of different ambient temperatures and other weather parameters.

Hypotheses

With the typical phenomena concerning measurement and calculation discussed and the tools prepared, hypotheses regarding further variables in the EPC Register were tested with regard to whether they could contribute to explain the still tremendous spread in the sample in Fig. 1. As mentioned before, actual values of the parameters of user behaviour are not documented, building geometry and physical properties of building components and the technical plant are well accounted for within the demand calculation. Thus only variables remain, that are not well considered in the whole spread of their occurrence by the simplified calculation model of the EPC.

Several regression functions were tested and analyses conducted. The following variables turn out to be significant, according to the corresponding p value: Number of dwelling units n_{DU}, reference area $A_n[m^2]$, air tightness $n_{50}[1/h]$ and compactness $A/V_e[1/m]$.[2]

Luxemburg EPC

Keeping in mind that certain arbitrariness is allowed, the final regression function for the Luxemburg EPC was defined as

$$\mathrm{Ln}(\hat{q}_{f,C}) = \beta_0 + \beta_1 \cdot \mathrm{Ln}\left(q^{EPC}_{f,D}\right) + \beta_2 \cdot n_{DU} + \beta_3 \cdot A_n + \beta_4 \cdot n_{50} + \beta_5 \cdot {}^{A}/_{V_e} \quad (19)$$

rendering the estimation function

$$\hat{q}_{f,C} = \left(q^{EPC}_{f,D}\right)^{\beta_1} \cdot e^{\beta_0 + \beta_2 \cdot n_{DU} + \beta_3 \cdot A_n + \beta_4 \cdot n_{50} + \beta_5 \cdot {}^{A}/_{V_e}} \quad (20)$$

to determine the estimated "measured" consumption with regression coefficients and regression statistic as in Tables 2 and 3. Defying immediate perception the coefficients may be interpreted as generalised slopes of a hyperplane in a five-dimensional space, β_1 meaning that with $q^{EPC}_{f,D}$ increasing by 1% $\hat{q}_{f,C}$ is doing so by β_1% whereas the other β_j mean that with the corresponding variable increasing by 1 unit $\hat{q}_{f,C}$ is doing so by $\beta_j \cdot 100\%$.

There are some interesting properties to this estimation function. While the logarithms in Eq. (19) are to heal heteroscedasticity, the estimation function meets the origin and the standard error turns into a percental quantity. Correlation between the variables is reasonable and the F test indicates a statistically well-confirmed regression model.

[2] Here A means the area of the building envelope in m^2 and V_e the enclosed volume in m^3.

Table 2 Values of the regression coefficients

Coefficient	Variable	Value
β_0	Intersect	2.427378825
β_1	$\mathrm{Ln}(q_{f,D}^{\mathrm{EPC}})$	0.473667306
β_2	n_{DU}	0.029001539
β_3	A_n	− 0.000343169
β_4	n_{50}	− 0.014787772
β_5	A/V_e	0.164350872

Data Source: IWU (reg2_DB6_LN)

Based on this analysis, the Luxemburg EPC has been supplemented: From the year 2016 on, the EPC depicts the estimate $\hat{q}_{f,C,i} \pm \hat{\sigma}\left(\hat{q}_{f,C,i}\right)$ besides the calculated demand and, if already available, measured consumption $q_{f,\ C,\ i}$, indicating to the owner of the building that from calculations with standard specifications only limited accuracy can be expected. The correct interpretation is: With a probability of 68%, the measured consumption $q_{f,\ C,\ i}$ will be in the range of the estimate's uncertainty $\hat{\sigma}\left(\hat{q}_{f,C,i}\right)$.

Figure 6 illustrates how this estimate relates to measured consumption in the sample, depicting pairs of values $\left(q_{f,C,i}, \hat{q}_{f,C,i}\right)$ together with $\left(q_{f,C,i}, q_{f,D,i}^{EPC}\right)$ as before. The slope has improved considerably towards the bisecting line but still, the cloud is quite spread out and there are extreme outliers. Further analysis is needed.

Table 3 Regression statistics and analysis of variance

Regression statistics	
Multiple correlation-coefficient	0.59
Coefficient of determination	34%
Adjusted coefficient of determination	34%
Standard error	32%
Observations	4407
ANOVA	
F test	458
F_{crit}	0

Data Source: IWU (reg2_DB6_LN)

Estimation of the influence of user behaviour on calculated energy demands

As pointed out before, the actual user behaviour is not being recorded in the Luxemburg EPC. Presumably, it is the predominant cause of the spread of the data, which we still observe in Fig. 6.

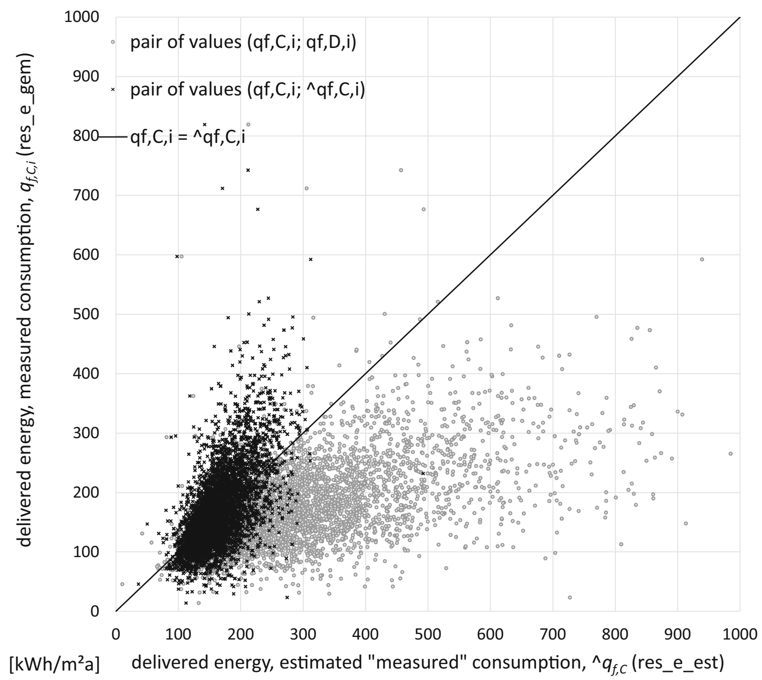
The distributions of "true" user behaviour parameters B_k must be estimated from other sources (Schröder et al., 2014; Building Research Establishment Ltd, 2013; Schröder, 2017; Brohus & Heiselberg, 2009). We take them as estimates of the error propagated when calculating the demand with standard specifications. Figure 7, for example, shows frequency distributions of winter room temperature measurements, mean room temperatures $\overline{\vartheta}_{int}$ increasing by 4 °C from buildings built prior to 1978 to passive houses while the temperature spread between rooms within the same dwelling unit is reduced considerably.

Obviously, real room temperatures in buildings tend to spread significantly about the Luxemburg EPC standard specification of 20 °C, the mean value of the total sample in Fig. 7 has been estimated to $\overline{\vartheta}_{\mathrm{int}}$= 19.0 °C and the sample standard deviation $\sigma(\vartheta_{\mathrm{int}})$ = 3.3 °C.

There is no comparable data for Luxemburg but it seems to be acceptable to estimate the error in calculated EPC demands from uncertainties in room temperatures in German dwellings. Ideally, standard specifications of user behaviour parameters in EPCs should have been defined as mean values of frequency distributions from random samples. Thus, we identify the standard specifications with the true mean of B_k and its known error $\sigma(B_k)$ taking advantage of the standard deviation's invariance under changes in the location of the random variable $\sigma(B_k + const) = \sigma(B_k)$.

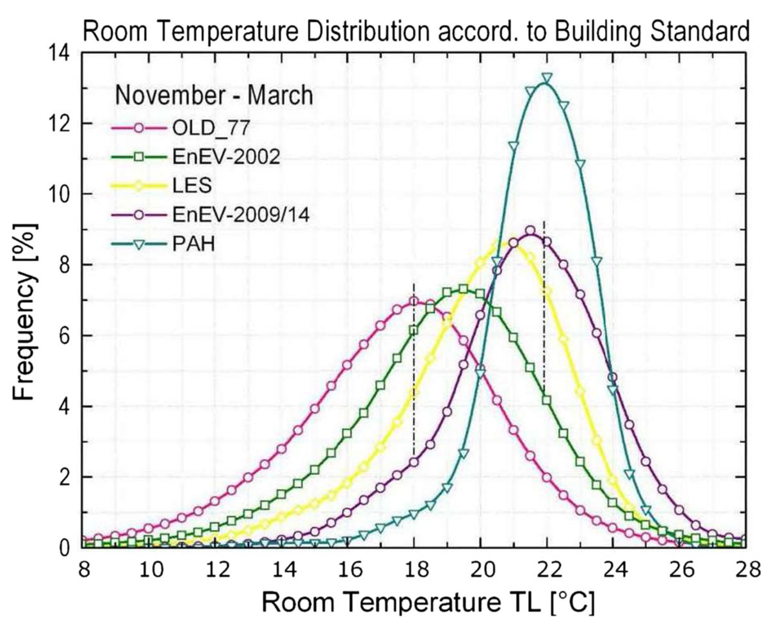
Besides room temperature $\vartheta_{\mathrm{int}}[^{\circ}C]$, we include three other parameters of user behaviour, that are to be specified in the Luxemburg EPC, namely the thermally effective window ventilation rate $n[1/h]$, specific internal heat gain $q_{int}[W/m2]$ and the specific domestic hot water demand q_{DHW}[kWh/m2a]. Calculated demand, as in Eq. (2), may then be considered as a function of these variables of user behaviour and other variables x_i, $i = 1, \ldots$, M describing the physical building, supposed to have been specified with reasonable accuracy, which is why uncertainties were neglected for this analysis. Provided that, the uncertainty of the calculated demand is given by equation

Fig. 6 Measured energy consumption (heating and domestic hot water) $q_{f,\,C,\,i}$ plotted against estimated "measured" consumption $\hat{q}_{f,C,i}$ as compared to calculated demand $q^{EPC}_{f,D,i}$ of the 4.407 EPCs selected from Luxemburg EPC Register. Data Source: Luxemburg EPC Register, Luxemburg Ministry of the Economy, IWU (reg2_DB6_LN)



$$\sigma\left(q^{EPC}_{f,D}\right) = \sqrt{\left(\frac{\partial q^{EPC}_{f,D}}{\partial \vartheta_{int}} \cdot \sigma(\vartheta_{int})\right)^2 + \left(\frac{\partial q^{EPC}_{f,D}}{\partial n} \cdot \sigma(n)\right)^2 + \left(\frac{\partial q^{EPC}_{f,D}}{\partial q_{int}} \cdot \sigma(q_{int})\right)^2 + \left(\frac{\partial q^{EPC}_{f,D}}{\partial q_{DHW}} \cdot \sigma(q_{DHW})\right)^2} \quad (21)$$

Fig. 7 Frequency distributions of measured winter room temperatures within German rental flats, developing (from left to right) with advancing energy efficiency standards from buildings built prior to 1978 (OLD_77), various requirement levels of the German Energy Savings Ordinance (EnEV) to passive houses (PAH). A 4 K increase in mean room temperatures and a decline in variance is evident. Data Source: METRONA (2017) (Schröder, 2017)



The uncertainties in user behaviour variables are estimated from the above mentioned external sources (e.g. as shown in Fig. 7) or from educated guess as depicted in Tables 4 and 5. We further assume that these values of user behaviour parameters apply for a single dwelling unit (DU), while in multi-family dwellings they average to values of the building, that approximate the standard specifications better with the standard error of the mean $\sigma\left(\overline{B}_k\right) = {}^{\sigma(B_k)}/_{\sqrt{n_{DU}}}$.

A considerable average uncertainty in the independent variable ${}^{\sigma\left(q_{f,D}^{EPC}\right)}/_{q}f, D^{EPC} = 25\%$ arises from the assumed spread in user behaviour, due to the room temperature averaging the error is considerably smaller for multi-family than for single-family houses.

Statistical analysis of calculated energy demand and measured consumption for non-residential buildings in Germany

Similar regression analyses were applied within the TEK-project (Hörner et al., 2014a) to the sector of non-residential buildings in Germany in order to analyse the observed discrepancy between calculated energy demand and actual energy consumption for heating and domestic hot water and electrical energy.

Heating and hot water

As can easily be seen from the best fit straight lines in a simple regression model in. Fig. 2 (B-4) calculation of the delivered energy demand for heating and domestic hot water is in better accordance with measurement the more realistic values of the input variables have been chosen. Data points in the Std-simplified-scheme in Fig. 2 (B-1) assuming standard specifications of user parameters and a simplified model of the building envelope geometry are similarly spread as the Luxemburg EPC data, also showing heteroscedasticity.

Table 4 Assumed uncertainties of user behaviour variables

Variable		$\sigma(x)$
Room temperature	ϑ_{int} [°C]	± 3.3 °C
Thermally effective window ventilation rate	n [1/h]	± 30%
Specific internal heat gain	q_{int} [W/m^2]	± 30%
Specific domestic hot water demand	q_{DHW} [kWh/m^{2}a]	± 30%

Data Source: Luxemburg EPC Register, Luxemburg Ministry of the Economy, IWU (DB4_eceee)

Following our rationale, for the further derivations, we will focus on the Std-simplified-scheme which is our preferred candidate for fast calculations of many buildings in scenarios of a building stock for example. From physical reasoning and following the procedure outlined in the previous chapters, we propose the multiple regression equation

$$\mathrm{Ln}\left(\hat{q}_{f,C}\right) = \beta_0 + \beta_1\cdot f_{winvent,area} + \beta_2\cdot\Delta q_{int,std-real} + \beta_3\cdot\Delta t_{use,std-real} + \beta_4\cdot\Delta\vartheta_{int,std-real} + \beta_5\cdot\mathrm{Ln}\left(q_{f,D}^{Std-simple}\right) \quad (22)$$

We take advantage of the on-site building assessments in the TEK-project delivering momentary values of user parameters and their deviation from standard specifications in DIN V 18599-10: differences of standard and real heat gain $\Delta q_{int,\ std-obj}$, use time $\Delta t_{nutz,\ std-obj}$ and room temperature $\Delta\vartheta_{Raum,\ std-obj}$. $f_{winvent,\ area}$ is a percentual value of the building's net floor area with window ventilation, while $q_{f,D}^{Std\text{-}simpl.}$ denotes calculated delivered energy demand in the Std-simplified-scheme and $\hat{q}_{f,C}$ the estimator of measured consumption.

Thus, we figure the estimation function

$$\hat{q}_{f,C} = \left(q_{f,D}^{Std-simple}\right)^{\beta_5}\cdot e^{\beta_0+\beta_1\cdot f_{winvent,area}+\beta_2\cdot\Delta q_{int,std-real}+\beta_3\cdot\Delta t_{use,std-real}+\beta_4\cdot\Delta\vartheta_{int,std-real}} = q_{f,D}^{Std-simple}\cdot f_{C/D}\left(q_{f,D}^{Std-simple}\right) \quad (23)$$

rendering a calibration function $f_{C/D}$

$$f_{C/D} = \frac{\hat{q}_{f,C}}{q_{f,D}^{Std-simple}} = \left(q_{f,D}^{Std-simple}\right)^{\beta_5-1}\cdot f_{use} \quad (24)$$

that is plotted over $q_{f,D}^{Std-simple}$ in Fig. 8.

Thus with a simplified calculation model of TEK and a handful of parameters characterising actual usage of a building, we get a pretty good estimation of the building's presumable consumption, the adjusted coefficient of determination $R^2 = 63\%$ and an estimated standard error $\hat{\sigma}\left(\hat{q}_{f,C}\right) = 31\%$. The slope estimator β_5 in the multiple regression model becomes $\beta_5 = 0.81$, which comes pretty close to 1.

Table 5 Average relative uncertainties of the calculated demand according to Luxemburg EPC Δq/q for different dwelling types due to uncertainties in user variables

Dwelling	Δq/q (%)	Number of observations
Single-family houses (SFH)	30	2.788
Small multi-family houses (sMFH)	22	938
Multi-family houses (MFH)	10	681
All houses	25	4.407

Data Source: Luxemburg EPC Register, Luxemburg Ministry of the Economy, IWU (DB4_eceee)

We apply this calibration procedure to a refurbishment example (Fig. 8): An existing building with a calculated energy demand for heating and hot water of $q_{f,D}^{\text{Std-simple}} = 200\ [kWh/m^2a]$ and a calibration factor $f_{C/D} = 0,\ 72$ is considered to have an estimated consumption of $\hat{q}_{f,C} = 200{\cdot}0,72 = 144\ [kWh/m^2a]$. Supposed, after refurbishment calculated demand decreases to $q_{f,D}^{\text{Std-simple}} = 50\ [kWh/m^2a]$, yields a calibration factor $f_{C/D} = 0,\ 94$ and an estimated consumption of $\hat{q}_{f,C} = 47\ [kWh/m^2a]$. An estimated savings of $\Delta\hat{q}_{f,C} = 97\ [kWh/m^2a]$ results, instead of an uncorrected savings of $\Delta q_{f,D}^{\text{Std-simple}} = 150\ [kWh/m^2a]$.

Errors-in-variables

Let us return to a simple regression model with only ln $\left(q_{f,D,i}^{\text{std-simpl.}}\right)$ as an independent variable. Results in the real-real-scheme may be considered as true values of the demand figured independently of the Std-simplified-scheme, since advantage has been taken of on-site measurements of parameters of user behaviour, though momentarily only and thus error prone also, and a more realistic geometry model. As a plausibility check, we adopt the errors-in-variables model from equation 15 ln

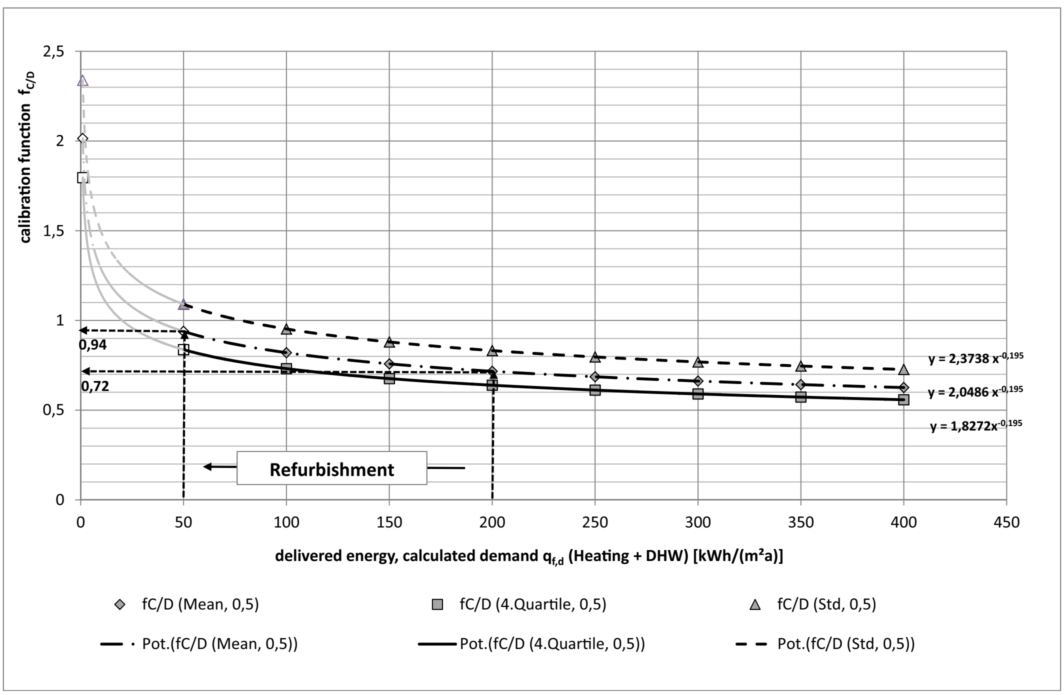


Fig. 8 Calibration function $f_{C/D}$ over calculated demand in Std-simplified-scheme with 50% window ventilated area and various deviations of user behaviour parameters from standard specifications assumed: No deviation (Std, 0.5), mean of deviations (mean, 0.5) and mean of the 4th quartile of deviations (4th quartile, 0.5). The grey parts of the graphlines mark the range of values of calculated demand where the Luxemburg EPC register does not provide a sufficient number of cases for the analysis. Exemplary illustration of refurbishment and the different effects of calibration on standardised demand calculations before and after. Data Source: TEK-Database, IWU (2015625_QSA-Regression-E / calib-func)

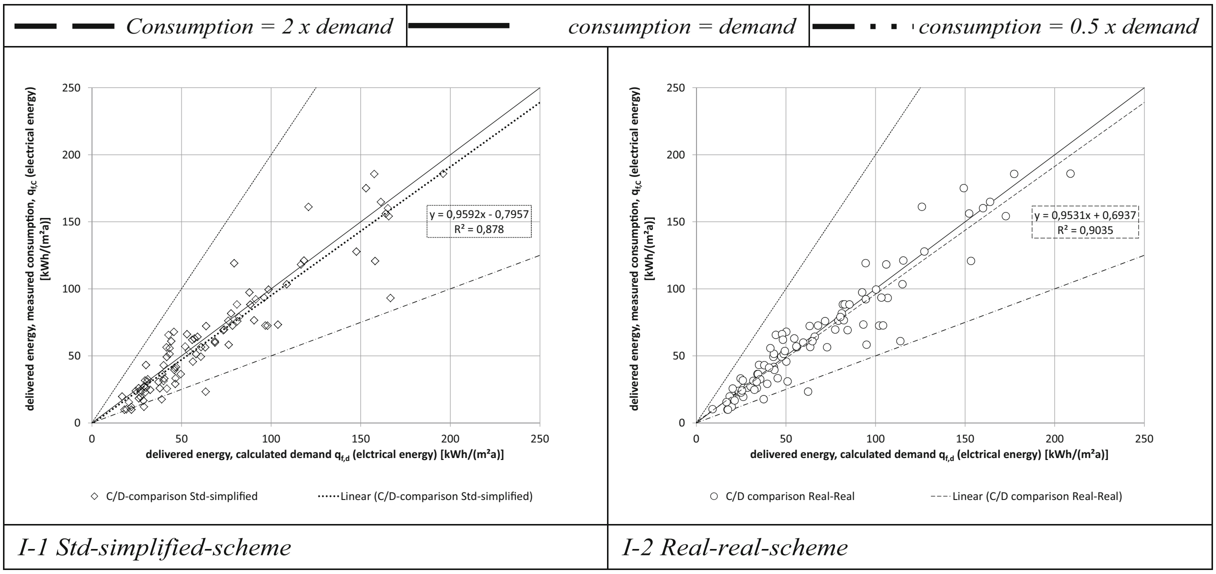


Fig. 9 Plot of measured consumption over calculated demand of the buildings' delivered electrical energy for the Std-simplified- and the real-real-scheme

$\left(q_{f,D,i}^{\text{std-simpl.}}\right) = \ln\left(q_{f,D,i}^{\text{real-real}}\right) + \Delta_i$. With the Pearson correlation coefficient $\rho\left(\ln\left(q_{f,D,i}^{\text{real-real}}\right), \Delta_i\right) = 0{,}39$, we assume the two quantities to be uncorrelated, the classical errors-in-variables assumption. The reliability ratio then turns out to be

$$\lambda = \frac{\sigma^2\left(Ln\left(q_{f,D,i}^{\text{real-real}}\right)\right)}{\sigma^2\left(Ln\left(q_{f,D,i}^{\text{real-real}}\right)\right) + \sigma^2(\Delta_i)} = 0,81 \tag{25}$$

Thus, the error of the calculated demand in the Std-simplified-scheme seems to cause most of the observed attenuation in the slope coefficient β_5.

Electrical energy

Interesting is the fact that for electrical energy no such problems were found in the TEK-project. As Figure 9 shows, even with standard specifications of user behaviour and a simplified geometry model for the building envelope calculated demands and measured consumptions accord in a reasonable manner. The trendline is very close to the bisecting line and to the real-real-trendline.

It seems that there is less influence of user behaviour on electrical energy consumption, since room temperatures and window ventilation are of less concern. Unlike the German EPC procedure for non-residential buildings based on DIN V 18599 (DIN Deutsches Institut für Normung e.V.: DIN V 18599-10: 2011-12, 2011), TEK is taking into account not only the electrical energy demand of technical installations like air handling units and lighting but also all sorts of appliances used in the building, e.g. elevators, IT-equipment, kitchen appliances. The electrical energy consumed by these appliances is not only regarded as internal heat gain in the heating energy balance but is considered in the electrical energy balance also. Furthermore, the modelling of the ventilation system in TEK is different to the approach of DIN V 18599. In TEK, the exact values of fan power, air flow and operation time are used, typically as displayed on the nameplate, and the ventilation system is calculated as a whole.

Thus, it seems that confusing deviations between calculation and measurement can be easily avoided in balancing electrical energy in the non-residential building domain.

Conclusions

EPCs are made to inform about the energy-related quality of buildings and to certify compliance with Energy Performance Ordinances irrespective of user behaviour and climate parameters. Thus, specific values of calculated energy demand in particular for heating and domestic hot water deviate considerably from measurement. Many users and owners of buildings do not understand that. However, EPCs are an important source of information on energy efficiency parameters of buildings, the benefits should be made use of.

The methods discussed here serve as an approach to reconcile standardised demand calculations, as readily available from EPCs, with measurements as documented, e.g. in energy bills. There is no need to modify standardised calculation schemes in EPCs, except for a more consistent consideration of electrical energy in non-residential buildings. But findings from empirical data on consumption should be supplemented in an additional step. In the paper, a procedure to generate an estimation of future consumption including the range of uncertainty from a standardised demand calculation has been defined regressing measured consumption on calculated demand. The Luxemburg EPC has already been improved by adding the characteristic value of estimated consumption including uncertainty to the characteristic values of standardised demand and carbon dioxide emissions.

In energy consulting and the economic assessment of energy savings measures, on the other hand, most often similar calculation schemes are used as in EPCs but with actual values of user behaviour and climate parameters and adjustment of physical building parameters within their usual ranges of uncertainty. A considerable effort comes along with this, but calculation results are much more consistent with measurements, as we have demonstrated with the results of the TEK-project. Either way discrepancies between measured energy consumption and calculated demand cannot be avoided. Calibration procedures as derived in the paper will help to yield more realistic results.

As our analysis in the errors-in-variables model suggests, uncertainties in dependent and independent variables affect the estimators in regression analysis in a typical way, attenuating slope parameters and increasing variances. The distribution of user behaviour parameters especially should be quantitatively analysed with the objective to realistically define standard specifications as the mean values of typical distribution functions (Brohus & Heiselberg, 2009), thereby reducing standard errors as a measure of their spread (Santos Silva & Ghisi, 2014) by using building typologies for typical categories of usage.

For analyses in building stocks, simplified calculation models with well-defined standard specifications including uncertainties and calibration functions are the method of choice to predict the future energy consumption of the building sector in scenario calculations. Quantifiable uncertainties are an essential in order to use these scenario results as a base for decision-making in the political arena. Statistical methods as exemplified above should be considered as standard in scenario calculations to achieve this.

Prerequisite are databases of representative samples of building data on measured consumption and calculated demand including user behaviour parameters based on the model of the Luxemburg EPC Register.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

K. Backhaus, K. et al. (2005): Multivariate Analysemethoden – Eine anwendungsorientierte Einführung, Springer-Verlag, Berlin 2005, 11. Auflage.

Brohus, H., & Heiselberg, P. (2009). *Application of partial safety factors in building energy performance assessment* (pp. 1014–1021). Glasgow: Eleventh International IBPSA Conference.

Building Research Establishment Ltd. (2013): Energy follow-up survey 2011—report 2: mean household temperatures, BRE report number 283078 prepared by BRE on behalf of the Department of Energy and Climate Change (DECC),UK, December 2013.

Manfred Casties (1997): Untersuchung zum Zusammenhang zwischen Nutzerverhalten und Heizenergieverbrauch, -bedarf von Wohngebäuden; Verlag für Wissenschaft und Forschung, Berlin 1997.

Chari, A., Xanthos, S., & Christodoulou, S. (2017). Stochastic assessment of the energy performance of buildings, S.E. In *Energy efficiency* (Vol. 10, pp. 1573–1591). https://doi.org/10.1007/s12053-017-9545-0.

Corrado, V., & Mechri, H. E. (2009). Uncertainty and sensitivity analysis for building energy rating. *Journal of Building Physics, 33*(2), 125–156.

DIN Deutsches Institut für Normung e.V.: DIN V 18599: 2016-10 Energy efficiency of buildings—calculation of the net, final and primary energy demand for heating, cooling, ventilation, domestic hot water and lighting, Beuth Verlag, Berlin 2016.

DIN Deutsches Institut für Normung e.V.: DIN V 18599-10: 2011-12 Energy efficiency of buildings—calculation of the net, final and primary energy demand for heating, cooling, ventilation, domestic hot water and lighting - Part 10: boundary conditions of use, climatic data, Beuth Verlag, Berlin 2011.

Hans Erhorn (2007): Bedarf – Verbrauch: Ein Reizthema ohne Ende oder die Chance für sachliche Energieberatung?, Fraunhofer-Institut für Bauphysik, Stuttgart (available at: http://www.buildup.eu/sites/default/files/Manuskript_Messe_Bau_2007_Erhorn_Bedarf-Verbrauch_15_01_07_p.pdf).

Alicia Graf (2016): Analyse des Energieverbrauchs wärmetechnisch modernisierter Mehrfamilienhäuser - Entwicklung von Verbrauchsbenchmarks zur Beurteilung der Energieeffizienz, Masterthesis Institut für Massivbau, Technische Universität Darmstadt, 2016, S. 38–43.

Edelgard Gruber et al. (2005): Energiepass für Gebäude – Evaluation des Feldversuchs. Schlussbericht an die Deutsche Energie-Agentur, Karlsruhe 2005.

Michael Hörner et al. (2014a): Teilenergiekennwerte – Neue Wege in der Energieanalyse von Nichtwohngebäuden im Bestand, IWU 05/14, ISBN 978–3–941140-38-7.

Michael Hörner, Jens Knissel et. al. (2014b): Berechnungsgrundlagen des TEK-Tools Version TEK 6.2, IWU 03/14, Darmstadt 2014, ISBN 978-3-941140-36-3 (TEK-Tool (V 6.3-DB4.34), downloadable at http://www.iwu.de/forschung/energie/laufend/teilenergiekennwerte-von-nicht-wohngebaeuden/).

Jens Knissel et al. (2006): Vereinfachte Ermittlung von Primärenergiekennwerten – zur Bewertung der wärmetechnischen Beschaffenheit in ökologischen Mietspiegeln, IWU, Darmstadt 2006 (available at http://www.iwu.de/fileadmin/user_upload/dateien/energie/werkzeuge/Vereinfachte_Ermittlung_von_Primaerenergiekenwerten-1.0.pdf).

Le gouvernement du grand-duché de Luxembourg (2007): Règlement grand-ducal du 30 novembre 2007 concernant la performance énergétique des bâtiments d'habitation (14.12.2007). Luxembourg : Service central de legislation.

Markus Lichtmeß (2012): Forschungsprojekt Gebäudedatenbank und Plausibilitätscheck im LuxEeB-Tool, interne Ergebnisse, Goblet Lavandier & Associés Luxemburg.

Loga, T. et al. (2003): Der Einfluss des Gebäudestandards und des Nutzerverhaltens auf die Heizkosten, IWU on behalf of Viterra Energy Services AG, Essen, 2003, Page 50.

Santos Silva, A., & Ghisi, E. (2014). Uncertainty analysis of user behaviour and physical parameters in residential building performance simulation. *Energy and Buildings, 76*, 381–391, ISSN 0378-7788. https://doi.org/10.1016/j.enbuild.2014.03.001.

Franz Schröder (2017): Note, METRONA Wärmemesser Union GmbH, München, 2017.

Franz Schröder et al. (2014): Reale Trends des spezifischen Energieverbrauchs und repräsentativer Wohnraumtemperierung bei steigendem Modernisierungsgrad im Wohnungsbestand, Bauphysik 36 (2014), Heft 6, p. 309 - 324.

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

Jeffrey M. Wooldridge (2003): Introductory econometrics: a modern approach, 2e, Thomson learning – South-Western, Mason, Ohio 45040, USA, p. 305 ff., 503 ff.