Energy & Buildings 288 (2023) 113024

Contents lists available at ScienceDirect

Energy & Buildings

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

The over-prediction of energy use by EPCs in Great Britain: A comparison of EPC-modelled and metered primary energy use intensity

Jessica Few *, Despina Manouseli, Eoghan McKenna, Martin Pullinger, Ellen Zapata-Webborn, Simon Elam, David Shipworth, Tadj Oreszczyn

UCL Energy Institute, University College London, 14 Upper Woburn Place, London WC1H 0NN, UK

ARTICLE INFO

Article history:
Received 13 December 2022
Revised 23 March 2023
Accepted 25 March 2023
Available online 29 March 2023

Keywords:
Energy Performance Certificate
Performance gap
Energy demand
Domestic buildings
Homes
Smart meter

ABSTRACT

This analysis compares the difference between the Energy Performance Certificate (EPC)-modelled and smart-meter measured annual energy use on a like-for-like basis in 1,374 gas-heated British households from the Smart Energy Research Lab (SERL) Observatory. EPCs and metered energy use were converted to primary energy use intensity (PEUI) to provide a comparison of the same quantity for the first time.

We show that EPCs predict significantly more energy use than metered in homes in Great Britain. EPC bands A and B show no statistically significant difference, but all other bands show a significant gap which increases as EPC rating worsens. The PEUI gap widens from −26 kWh/yr/m^2 (−8%) for band C to −276 kWh/y/m^2 (−48%) for bands F and G. Unlike previous research, we show that the difference persists in homes matching the EPC-model assumptions regarding occupancy, thermostat set-point and whole-home heating; suggesting that occupant behaviour is unlikely to fully explain the discrepancy.

EPCs are a core tool in the residential energy sector, and the gap between EPC-modelled and metered energy use could have a significant impact on policy, research, and industry. Future research should investigate disaggregated components of energy use, the underlying thermal model, and assumptions regarding building characteristics.

1. Introduction

Energy Performance Certificates (EPCs) were introduced to increase consumer awareness of building energy efficiency, and have been required at the point of sale, rent or construction of a building in Europe since 2013 under the European Building Performance Directive (EBPD) [17]. They have been in use in the UK since 2007 [31] and are still required following departure from the European Union. Since their introduction, 23 million EPC records have been generated in the UK, with over 1.5 million EPCs lodged in England in the 12 months to December 2021 [14].

In the UK, EPCs rate the energy efficiency of homes using an A – G rating system, where homes in band A are rated as the least expensive to run (most efficient). EPCs are increasingly used to support the net zero transition; for example rental properties in England and Wales are required to be rated EPC-E or above as part of the domestic Minimum Energy Efficiency Standard (MEES) [21], and the UK government has an aspiration that all homes, where feasible, are EPC-C rated by 2035 [31]. The model underlying the EPC is also used in official fuel poverty statistics in England, whereby homes in bands A-C (inclusive) cannot be classed as fuel poor since their energy expenditure is expected to be low [6].

EPCs are not intended to give a prediction of the actual energy use of a particular home as they assume normative consumption. However, to be as effective as possible given their current uses, on average they should provide a reasonably accurate measure of the regulated energy use. This leads to our first research question:

- Is there a significant difference between metered primary energy use intensity (PEUI) and EPC-modelled PEUI?

Additionally, any systematic difference between metered and modelled energy use should be independent of the modelled energy use, ideally there should be a 1:1 linear relationship between EPC modelled and actual energy use. This leads to our second research question:

- What is the relationship between metered PEUI and EPC-modelled PEUI?

* Corresponding author.
E-mail address: jessica.few@ucl.ac.uk (J. Few).

https://doi.org/10.1016/j.enbuild.2023.113024
0378-7788/

Existing research has been unable to control for factors that may plausibly contribute to discrepancies between modelled and actual energy use, this leads to our final research question:

- How does the relationship between metered PEUI and EPC-modelled PEUI change while controlling for EPC band and other key variables, including agreement with the EPC-modelled occupancy and heating assumptions?

We test these relationships using the detailed Smart Energy Research Lab (SERL) Observatory data [15]and present the first like-for-like comparison of the same quantity between the EPC-modelled and metered energy use in the UK. SERL is a UK research council funded project which brings together, for the first time, electricity and gas smart meter data alongside a participant survey, EPC, weather, and location data for a representative sample of over 13,000 households in Great Britain (GB). Webborn et al. [38] find that the SERL data provides a much wider array of energy demand co-variates and more detailed energy use data than had previously been reported in the literature, while McKenna et al. [30] find that the covariates within the SERL Observatory provide good explanatory power for the variation in daily energy demand.

The following section presents the EPC process in the UK and then explores the literature on the difference between metered and EPC-modelled energy use. Section 3 describes the SERL Observatory data and methods used for comparing the metered and EPC-modelled energy use in detail. The results are presented in Section 4 alongside a discussion of their implications, possible causes of the discrepancy and future research are explored in Section 5, with the main conclusions summarised in Section 6.

2. Literature review

2.1. EPC assessment in the UK

The primary EPC banding (A to G) is based on the modelled cost per floor area. In the UK, the model underlying the EPC is the Standard Assessment Procedure (SAP) [7]. SAP uses factors relating to the physical building (e.g. building form, insulation, solar gains, heating system efficiency) and assumptions regarding the occupancy and heating schedule to calculate the regulated energy use of the building. The 'regulated energy use' refers to demand from fixed elements in the building; in the UK this is energy for space and water heating and associated fans, pumps and controls, and fixed lighting, but excluding appliances and cooking [10].

The EPC reports the annual primary energy use intensity (PEUI) calculated using SAP with assumed *local weather* profiles, provided for each region within the SAP 2012 documentation (see Appendix U of SAP 2012) [7]. The EBPD requires that the energy use is reported as PEUI, rather than delivered energy use intensity. Primary energy is the amount of energy contained in the raw resources used to generate electricity or contained in gas, losses occur in transformation of the energy resource into electricity and in transmission of electricity and gas. The energy consumed by the end users in their home is known as delivered energy.

For the fuel cost output (which the EPC rating is based on) the regulated energy use is calculated assuming *average UK weather*, which is then converted to cost using the assumed fuel prices including standing charges. This cost is normalised by floor area and the most expensive homes to run receive a G rating, and the least expensive an A rating.

As a result, identical homes in different parts of the country would have the same EPC rating and be in the same band since these metrics use UK average weather, but the home in a colder location will report greater PEUI as this metric uses local weather. Moreover, two homes with the same delivered energy requirements could be in different bands and report different PEUI depending on the fuel used. For example, electricity is both more expensive and has a higher primary energy factor than gas, so electrically heated homes will tend to be in worse EPC bands and show higher PEUI.

In the UK, EPCs are generated for most new homes using a full SAP procedure whereby the energy assessor uses detailed characteristics as inputs to the SAP model, since these should be known from the building plans. However, for most existing buildings EPCs are generated using a Reduced Data Standard Assessment Procedure (RdSAP); details are provided in Appendix S of SAP 2012 [7]. RdSAP is used when all the inputs required for running the SAP model are not available, it defines defaults and conventions based on observable characteristics of the building to provide missing inputs. Thus, the calculation of an EPC using both the full SAP and RdSAP procedures use the same underlying model, but for RdSAP some input values are estimated. In the following, we use 'SAP model' or 'EPC model' to mean the underlying calculations used to generate the EPC rating and associated parameters, and 'RdSAP procedure' and 'SAP procedure' to refer to the processes of generating an EPC including the assessment by an energy assessor and assigning the values to use as inputs to the SAP model.

2.2. Should EPC-modelled and actual energy consumption match?

A discrepancy between modelled and actual energy use would be expected both because of complexities in real buildings, such as:

- Issues associated with build quality,
- Errors in the inputs to the model due to surveyor errors,
- Simplifications in the underlying model of the building physics,

And because SAP makes normative assumptions about the use of energy in homes. It does not attempt to calculate how much energy a particular home will use given its current occupant in a particular year. Thus, EPC-modelled annual energy use would be expected to be different to metered energy use for (at least) the following reasons:

- the occupants may behave differently than the model, for example they may heat to a different temperature than 21 ^{0}C (assumed by the SAP model in the living space), for a different number of hours, or for a different number of months in the year, and they may not heat the whole home,
- the weather may be different to that assumed in the model,
- the occupants may use unmetered energy (such as oil boilers or wood burning stoves).

Despite these reasons for disagreement on a household level, ideally, given a large enough sample to average out occupant variations, EPCs should provide a reasonably good agreement with metered energy use. This is important because, as noted above, EPCs underpin research and government modelling of energy demand, carbon emissions and fuel poverty in the residential sector. Moreover, since EPCs are primarily intended as a product rating tool, if homes are compared then the differences in EPC-modelled energy use should be representative of differences in actual regulated energy use.

2.3. The EPC performance gap

There have been several recent European papers exploring the difference between metered and modelled energy use according to local EPC processes (for compliance with the EBPD). In Flanders

(Belgium), Van Hove *et al.* [22] compared the primary energy use associated with metered gas and electricity in single-family homes with the modelled regulated primary energy use, finding an average overprediction of 147% relative to the metered energy use (i.e. almost two and a half times higher actual energy use than modelled). Moreover, they found the difference significantly increased with decreasing energy efficiency rating of the home. The overprediction was 217% for homes which were rated least energy efficient (band F), while homes rated most efficient (band A) underpredicted by 9%. Note that the regulated energy calculation in Belgium does not include appliances, cooking or lighting, so the quantities compared are not equivalent and if the total energy use was compared this would widen the discrepancy in the least efficient bands.

Similarly, Coyne and Denny [11] compared the Irish EPC primary energy use and metered primary energy use. The pattern of the difference was very similar to Van Hove *et al* [22], with the most energy efficient homes, bands A and B, using on average 54% more primary energy than predicted (relative to the modelled energy use) and the least energy efficient homes, bands F and G, using 51% less. The model and process underlying the Irish EPC were derived from the UK system, and the two remain very similar. The Irish EPC does not report the modelled total energy use as it does not include cooking or appliance use.

Alongside these recent analyses, there is extensive evidence from across Europe regarding the difference between EPC-modelled energy use and actual metered energy use (for example, [9,37,2]. There are several common findings in this research. Firstly, overall, the modelled energy use is greater than the metered energy use. Secondly, the most energy-efficient homes frequently use more energy than predicted. Thirdly, the least energy efficient homes use far less energy than predicted. The combination of the second and third points means that the metered energy results show a much shallower change in energy use between band ratings than predicted. For example, in the UK, Summerfield et al. [35] show that across all EPC bands the metered gas consumption is almost always within the range estimated for EPC band C.

This body of literature suggests several common contributing causes of the difference:

a. **Differences in real and modelled occupant behaviour**, particularly with regard to heating set-point temperatures, schedules, and whether the whole home is heated (e.g. [22,28,35,36]). This is sometimes termed the rebound and pre-bound effect [36]. The rebound effect can refer to the suggestion that occupants in high-efficiency homes alter their space heating behaviour to provide higher internal temperatures than they would if they lived in a lower-efficiency home (even where a retrofit has not taken place). The pre-bound effect is the opposite whereby occupants seek lower temperatures, or heat fewer rooms, in lower-efficiency homes compared to higher-efficiency homes.
b. **Shortcomings in the regulated energy use calculation methods**. For example, Van Hove *et al.* [22] highlight that in Belgium the underlying model assumes 24-hour space heating, which is likely to result in an overestimate of energy use compared to reality.
c. **Shortcomings or inaccuracies in the quality of the input data**. Several authors note that the practical implementation of the building performance assessment is undertaken by assessors who are trained to make use of default input values for model parameters. The default values are pessimistic by design, with the intention that the resulting output: does not indicate a better-than-merited rating, highlights the advantage of carrying out a retrofit, encourages the identification of accurate information for an accurate rating and encourages the recording of energy efficiency upgrades [12,33]. Raushan et al. [33] explored the Irish EPC record and found that default U-values for elements such as walls and floors were almost always higher (suggesting worse insulating performance) than those recorded when more information was available to conduct a tailored U-value estimation. Although the assessors may be performing the assessment in accordance with guidance, the use of pessimistic default values could have the effect of increasing the modelled consumption compared to the real value. A separate issue related to the practical implementation by assessors was noted by Jenkins et al. [26] and Crawley et al. [13], who both found large uncertainties in the EPC rating, showing that when the same dwelling is rated repeatedly the resulting rating changes considerably.
d. **Difficulty in appropriately accounting for aspects of energy use not included in the model**. The energy use reported by EPCs is not the total energy consumption, but the regulated energy use which includes energy use related to space and water heating (and sometimes lighting, as in the UK), but excludes appliances and cooking. This means that the metered energy consumption is not directly comparable to the energy consumption reported by the EPC [11,22]. In some research this is presented as a known (and uncorrected) systematic bias towards lower modelled energy use compared to expected total energy use [22], while in others an estimation of the 'missing' energy use is provided and used to adjust either the metered or modelled energy use [11].
e. **Difficulty in appropriately accounting for aspects of energy use not included in the metered energy use.** Much previous analysis (e.g. [35,11,22] has used data which does not record whether homes were using unmetered energy (e.g. oil boilers, wood burning stoves) and therefore whether the metered energy use reflects the total energy consumption. In the UK, the EPC band rating is strongly influenced by fuel cost [28], meaning that F and G homes are more likely to be heated by more expensive fuels, such as (unmetered) LPG and oil. This may partially explain the large discrepancies between SAP calculations and metered energy use in low efficiency bands.

3. Research outline and methods

Fig. 1 presents an outline of the analysis presented in this work. Starting from the SERL metered energy use and EPC regulated primary energy use intensity, both quantities were converted to total primary energy use intensity (PEUI), thus facilitating comparison of the same quantity from both data sources. The conversion to the same quantity is discussed in Section 3.1. The weather conditions assumed in the SAP model are compared with the average conditions for SERL participants in 2021 in Section 3.2. Using the EPC and SERL survey we identify and filter out homes such that the metered and modelled PEUI reflect the same underlying energy uses. The criteria used to filter the data are described in Section 3.3. We then split the remaining households by key contextual variables, such as EPC band and agreement with the occupancy and heating-use assumptions; the approach to this is discussed in Section 3.4. Finally, Section 3.5 presents the approach to comparing the EPC-modelled and metered PEUI.

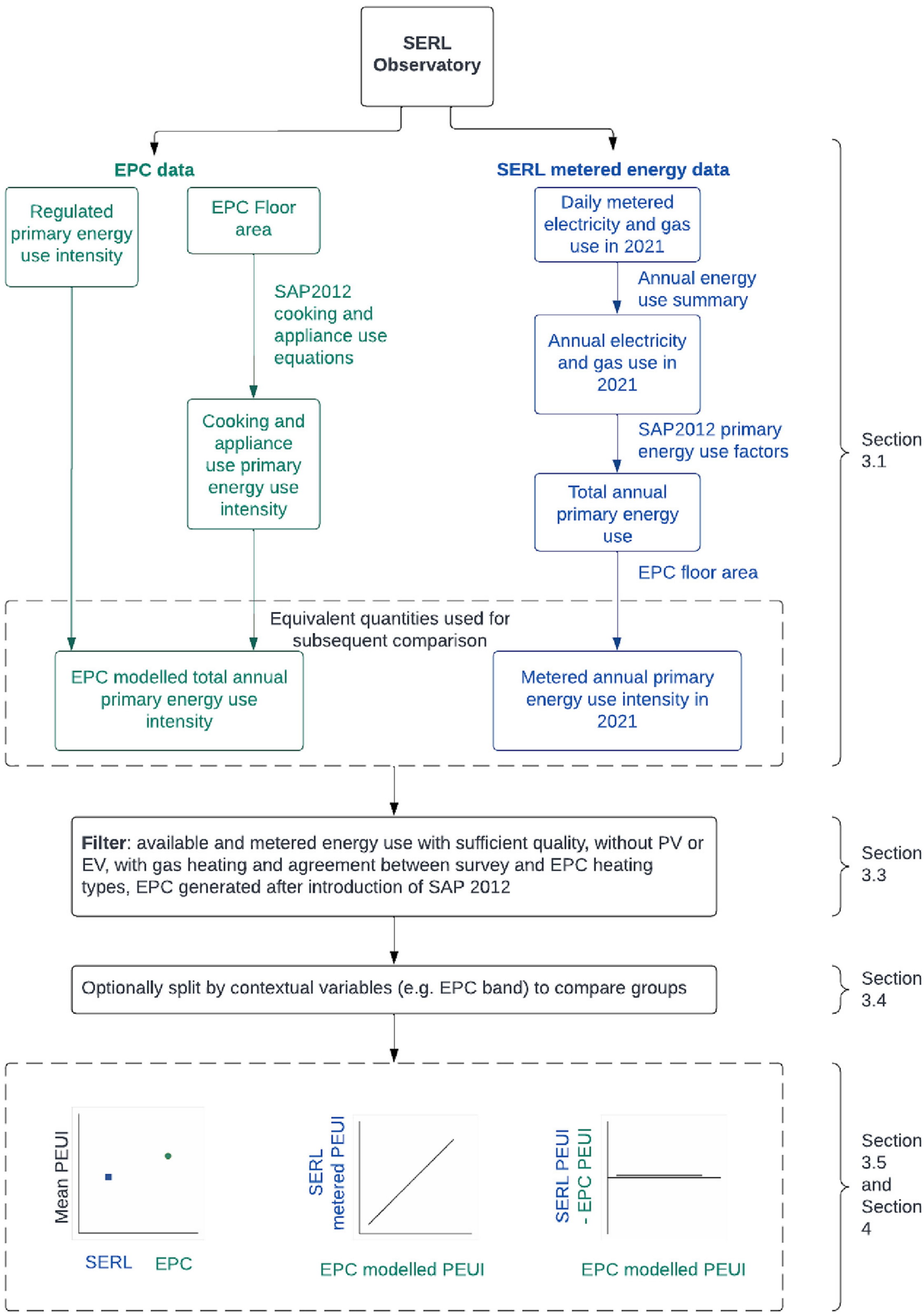

Fig. 1. Outline of analysis presented in this research and the sections in which they are discussed.

3.1. Metered and modelled primary energy use intensity

3.1.1. EPC modelled primary energy use intensity

The energy use reported by the EPC in the UK does not include appliance and cooking use, but this energy use can be calculated for each home using the SAP model. The heat gains from occupant metabolism, cooking and appliances are assumed to contribute to the energy required for space heating in the SAP model, with the remaining energy assumed to be provided by the heating system. The SAP calculation of energy use associated with appliances and cooking use are calculated using the assumed number of occupants in the dwelling (see Appendix L of SAP 2012), which is calculated using the floor area (see Table 1b in SAP 2012) [7].

The publicly-available EPC data within the SERL Observatory does not include all the input data to calculate the EPC rating or the predicted delivered energy. However, it does provide the annual primary energy use intensity (PEUI) for the regulated energy uses. The primary energy is calculated from the modelled energy consumption with the energy use for each fuel multiplied by a 'primary energy factor'; 3.07 for electricity and 1.22 for mains gas in SAP 2012 Version 9.92 (the version used in this analysis) [7]. Note that it is not straightforward to convert the PEUI to delivered EUI because the ratio of electricity to gas energy use is not reported. To facilitate comparison with the metered energy use, the calculated energy use for appliances and cooking according to the SAP model was converted to PEUI (assuming all cooking use was electric) using the above primary energy factors. This was added to the regulated PEUI to provide the total SAP modelled PEUI per household.

3.1.2. SERL metered primary energy use intensity

The SERL Observatory contains six datasets linked at the household level: electricity smart meter data, gas smart meter data, weather data, location data, participant survey data and EPC data (for full details of the data see [38], and for details of participant recruitment see [39]). The latest edition contains data from over 13,000 homes with smart meter data from August 2018 to August 2022 [15]. Data collection is ongoing and future editions will provide updated smart meter data. Note that some homes have historic EPC lodgements, but only the most recent is used for this analysis.

We calculated the 2021 annual gas and electricity energy consumption for homes in the SERL Observatory using the method described by Few et al. [18]. Briefly, the method is as follows:

- Calculate the mean daily energy use per month if at least 50% of the data was available and flagged as valid (according to standard SERL data quality processing, see documentation associated with [15])
- If the monthly mean was available for all 12 months in 2021 then calculate the mean daily energy use for the year (weighted by number of days in each month).

This method ensures that any period of missing data, or participant withdrawal from the Observatory part-way through the year, does not skew the annual result. For each home, the mean daily energy consumption for 2021 was converted to metered PEUI using the SAP 2012 primary energy factors [7] and the floor area from the EPC data.

3.2. Comparison of modelled weather with 2021 weather

Differences in actual weather compared to the assumed weather in the SAP model would be expected to result in a discrepancy between metered and modelled energy use. Fig. 2 shows the mean monthly temperature and solar radiation for SERL homes in 2021 and the average UK weather used in the SAP model. The yearly mean temperature was 10.2 ^{o}C for SERL and 10.0 ^{o}C for SAP, and the yearly mean solar irradiance was 120 W/m^2 for SERL and 108 W/m^2 for SAP. Although some months show significant deviations (for example April 2021 was particularly cold and sunny), the overall conditions were similar. Moreover, taking only October to March (the assumed months for space heating in the SAP model), the SERL mean temperature was 7.0 ^{o}C and 6.3 ^{o}C for SAP, and the SERL mean solar irradiance was 53 W/m^2 and 49 W/m^2 for SAP. Deviations in weather conditions are therefore not thought to have contributed considerably to any differences between metered and modelled energy use in the following analysis.

This analysis of weather data was repeated for each region, but patterns were similar so only the average was presented here for brevity. Note that in the subsequent analysis Scotland is excluded as their EPC process is slightly different, but for this comparison of the average UK weather Scotland was included.

3.3. Household selection criteria

As noted in Section 2.3, previous analysis of EPC-modelled and metered energy use has not been directly comparable. Using the linked contextual data within the SERL Observatory a sub-group of homes in which the metered energy use is robustly comparable to the SAP calculated value was selected, the following criteria were used:

- Homes with an EPC available through the SERL Observatory
- Homes with a recorded floor area between 20 m^2 and 500 m^2, which were self-contained according to the SERL survey[1].
- Homes with sufficient data (at least 50% in each month) to calculate an annual metered energy consumption value for gas and electricity using the method described above (see Section 3.1.2).
- The central heating was mains gas, based on SERL survey responses regarding the main heating system[2]. This was required because electrically heated homes would have a much higher primary energy consumption. Comparing homes with gas heating means that the energy uses associated with each energy vector are similar, increasing the reliability of the comparison.
- The EPC and SERL survey agreed over the main heating fuel in the home. Some homeowners reported different heating systems to that reported by an EPC assessor and this could result in large PEUI differences.
- Any secondary heating was either mains gas or electricity (i.e. the energy use was metered), this was determined using the SERL survey question regarding secondary heating[3].
- Homes without an electric vehicle (EV) according to the SERL survey. Excluding EVs avoids an energy load not used in the home and not included in the SAP model.
- Homes without photovoltaic (PV) electricity generation. Discrepancies between the way the SAP model treats electricity generation and the way the meters record electricity consumption and generation mean comparison of these two groups is not trivial.
- Homes in England and Wales; the Scottish EPC system is slightly different, so these were excluded.
- Homes where the EPC was created after the introduction of SAP 2012 (from May 2014 in England and from August 2014 in Wales) [7]. Significant changes to the primary energy factors were introduced in this version, and since the ratio between electricity and gas use is not reported in the publicly available EPC database it is not straightforward to convert PEUI calcu-

[1] SERL survey question B2: "Is your accommodation self-contained? By this we mean all the rooms, including the kitchen, bathroom and toilet, are behind a door that only your household can use".

[2] SERL survey question A3: "What type of central heating does your accommodation have? By central heating we mean a central system that generates heat for multiple rooms".

[3] SERL survey questions A6: "Other than central heating, does your household use any standalone heaters in your accommodation? This could be an electric heater, fireplace and so on" and A7: "Some standalone heaters use mains gas or electricity supply, such as gas fires or plugged-in electric bar heaters, while others have their own source of fuel. Do any of your standalone heaters have their own source of fuel (e.g. from logs, coal or bottled gas)?".

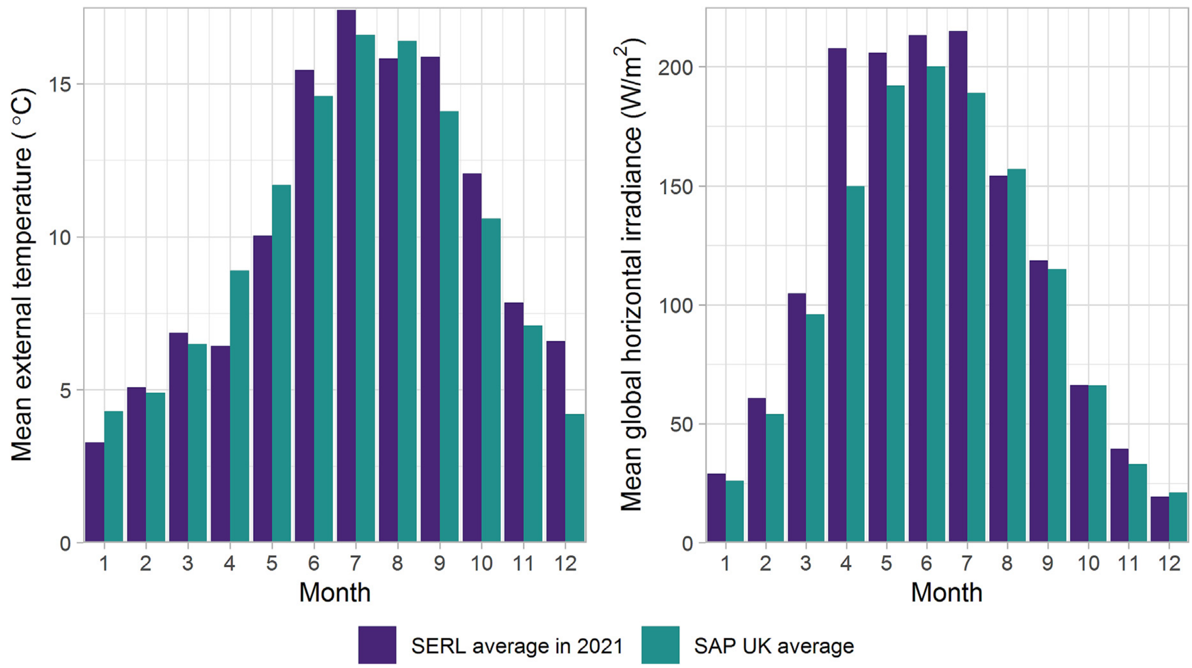

Fig. 2. Mean UK average monthly external temperature and solar irradiance for the SAP model and using SERL data for 2021.

lated under different versions of SAP. The extent to which the primary energy factors affect the discrepancy will depend on the degree of agreement between the ratio of electricity to gas use as metered and as modelled. Note that this filter was applied last and almost halved the sample. The analysis was also performed without this filter and the findings were almost identical, with smaller confidence intervals due to the increased sample sizes.

These filters resulted in a final subset of 1,374 homes.

3.4. Comparison of subgroups

We subdivide the homes according to the following characteristics to explore the agreement between metered and modelled PEUI for different subgroups:

- **EPC current energy rating (A to G)**: As noted in the literature review, previous research has identified less energy efficient EPC bands under-consuming compared to the SAP model and more energy efficient EPC bands over-consuming. Note that for this analysis only gas-heated homes were included, so the expense correlates very well with efficiency.
- **Number of occupants**: previous research has identified that number of occupants has a significant effect on energy use [23,27], particularly electricity. SAP assumes the number of occupants based on floor area, which may be different to the actual number of occupants; this may be related to the discrepancy.
- **Thermostat set point**: SAP assumes a thermostat set point of 21 ^{0}C in the living zone, although previous research has found a wide range of thermostat set points in homes[4,38]. Differences in thermostat set point would be expected to contribute to a discrepancy.
- **SAP occupancy and heating behaviour agreement**: previous research has suggested that differences in how occupants heat their home compared to the modelled assumptions could significantly contribute to the observed over estimation of energy use by the EPC model. See Section 3.4.1 for the details regarding what was considered to match the assumptions.
- **Purpose of EPC rating**: the transaction type variable in the EPC register records the reason for generating an EPC. We group this into 'new homes' and 'other EPC reason', because new build homes will be rated using the full SAP process rather than the RdSAP process used for most existing homes. As noted in Section 2, RdSAP provides default values for characteristics of a dwelling based on its age, whereas the full SAP process does not. This may impact the outcome of the assessment and could result in differences between groups of homes rated using each procedure.

Analysis of some groups has been combined due to small sample sizes and the need to comply with statistical disclosure control (SDC). EPC bands A and B, and F and G, and households with 6 or 7 occupants were grouped together for this reason. The thermostat variable was also grouped into 2 ^{0}C bands. Finally, not all participants answered every question in the SERL survey, so each variable based on the survey also includes a 'no answer or don't know' option (these are suppressed in the results that follow).

3.4.1. SAP occupancy and heating behaviour agreement

Homes categorised as matching SAP occupancy and heating behaviour have the following characteristics:

- The number of occupants according to the SERL survey matches the SAP assumed number of occupants to the nearest integer. The SAP assumed number of occupants is based on the floor area, see Table 1b in SAP 2012 [7]. The number of occupants influences the modelled appliance and cooking energy use, and heat gains associated with these are considered alongside metabolic gains from the occupants in calculation of the energy that the heating system must deliver. The mean difference between the number of occupants according to the SAP model and the SERL survey in the matching group was less than 0.1.

- The self-reported thermostat set point according to the SERL survey is between 20.5°C and 21.5°C. SAP assumes a set point of 21°C for the living room, with the temperature of the rest of the home modified by the heating control and the heat loss parameter, see Table 9 in SAP 2012 [7].
- Participants report that they heat their whole home. SAP assumes the whole home is heated, although it treats the living room differently to the rest of the dwelling if there is sufficient heating control, see Table 4e in SAP 2012.
- Participants report that they can keep comfortably warm in the winter, this gives greater confidence that, as well as their thermostat being set to 21 °C ± 0.5 °C, they also experience reasonably warm indoor conditions.
- Participants report that they are not struggling financially[4], this gives some confidence that these homes are unlikely to deliberately curtail their energy use to save money on their bills.

51 homes match the criteria above, with 1,323 not matching. The above criteria match the SAP occupancy and assumed heating behaviours as closely as possible with the available data, but not all SAP assumptions could be included. SAP assumes 9 h of heating during the week and 16 h during the weekend, but we were not able to select homes based on heating hours. The implications of the occupancy and heating assumptions are discussed in Section 4.2.2.

3.5. Comparing EPC and metered primary energy use intensity

To compare the metered and modelled PEUI, we report the absolute and percentage differences, as well as constructing linear models of the discrepancy and investigating the regression coefficients and their significance. Where we present the percentage difference, we follow the term referred to as the performance gap by Cozza et al. [12]:

$$\Delta PEUI\% = \frac{PEUI_{metered} - PEUI_{modelled}}{PEUI_{modelled}} \times 100 \tag{1}$$

A paired *t*-test was used to compare the means of the metered and modelled PEUI, with the null hypothesis that there is no difference in the mean metered and modelled PEUI, accepting that there was sufficient evidence to reject the null hypothesis if the p-value was less than 0.05.

A linear model of metered PEUI against EPC-modelled PEUI was fitted to further explore the relationship between metered and EPC-modelled energy use. If the EPC model of PEUI provides a good estimation of the difference in energy use between homes on average, then a gradient of one and intercept of zero would be expected. A gradient significantly different from one would indicate that the level of the discrepancy is related to the EPC-modelled energy use. For all homes, and for each group (e.g. new home EPCs and EPCs generated for other reasons), linear models were generated with the form:

$$PEUI_{metered} = \alpha PEUI_{modelled} + \beta \tag{2}$$

Where, α is the gradient and β the intercept, expected to be one and zero respectively if the EPC model accurately reflects metered energy use. Residuals were also examined to explore the consistency of the trend across the range of EPC-modelled PEUI.

Equivalently, a linear model of the *difference* in PEUI against the SAP modelled PEUI should show a gradient of zero if the EPC provides a good estimation on average. In fitting a model of this form, whether the gradient is significantly different from zero can be assessed using a *t*-test with the null hypothesis that there is no association between $\Delta PEUI$ and $PEUI_{modelled}$, accepting a p-value less than 0.05 to indicate that the gradient is significantly different from zero [25]. To explore this, models were generated with the form:

$$\Delta PEUI = \hat{\alpha} PEUI_{modelled} + \hat{\beta} \tag{3}$$

where $\hat{\alpha}$ is the gradient and $\hat{\beta}$ is the intercept, both expected to be zero if the EPC model provides a good estimation.

Analysis was implemented in R version 4.0.1 [32], using packages dplyr [42], reshape [40], ggplot2 [41]and cowplot [43].

4. Results and discussion

In the following section we present the overall discrepancy between PEUI predicted by the EPC and calculated from metered energy use, in Section 4.2 we then discuss the results split by EPC band, agreement with the SAP occupancy and heating assumptions, number of occupants, thermostat set point, and by reason for EPC generation. In Section 4.3 we present the factors associated with having an A or B EPC rating.

4.1. Overall primary energy use intensity difference

Fig. 3 presents a histogram of the difference between the PEUI calculated from the SERL smart meter data and the PEUI calculated using the EPC model. The distribution is weighted towards negative values, with a mean difference of – 66 kWh/year/m^2 and a *t*-test showing that the difference is significant (p-value $\ll$ 0.05), indicating that the EPC model tends to overpredict the PEUI compared to the metered value. In response to the first research question, this shows that there is a significant difference between the metered and EPC-modelled PEUI.

Fig. 4 and Table 1 show a linear fit of the SERL metered PEUI to the EPC-modelled PEUI, showing that the discrepancy tends to increase with increasing EPC-modelled PEUI. As noted in Section 3, if the SAP model was providing a good estimation of the metered PEUI on average a linear model of metered to modelled PEUI should find a gradient of one. The RMSE and confidence intervals around the gradient and intercept remain relatively large, suggesting that the EPC PEUI is not a good predictor of metered PEUI. The residuals were largely normally distributed, although there was a tendency for negative residuals at both low and high EPC-modelled PEIU values, indicating that the deviation of metered energy use from the EPC-modelled energy use is not fully linear. Nonetheless, the results clearly show that the SAP model does not accurately reflect the metered PEUI on average.

While this pattern of increasing difference with decreasing home energy efficiency is similar to previous findings across Europe and the UK, there are important differences in the present analysis. These results show that the discrepancy is not (solely) caused by the presence of homes with unmetered energy use, comparisons of energy metrics which do not include the same energy services, or comparisons of different energy units, as the pattern of differences persists in this analysis despite accounting for these factors.

4.2. Primary energy use intensity difference by key characteristics

4.2.1. Does the EPC band explain the difference between modelled and measured PEUI?

Fig. 5 shows the mean and standard error of the metered and modelled PEUI by EPC band and Table 2 presents the mean difference and their statistical significance according to a paired *t*-test. All bands except A and B have statistically significantly lower

[4] In response to SERL survey question D4: How well would you say you yourself are managing financially these days?" participants did not respond "finding it quite difficult" or "finding it very difficult".

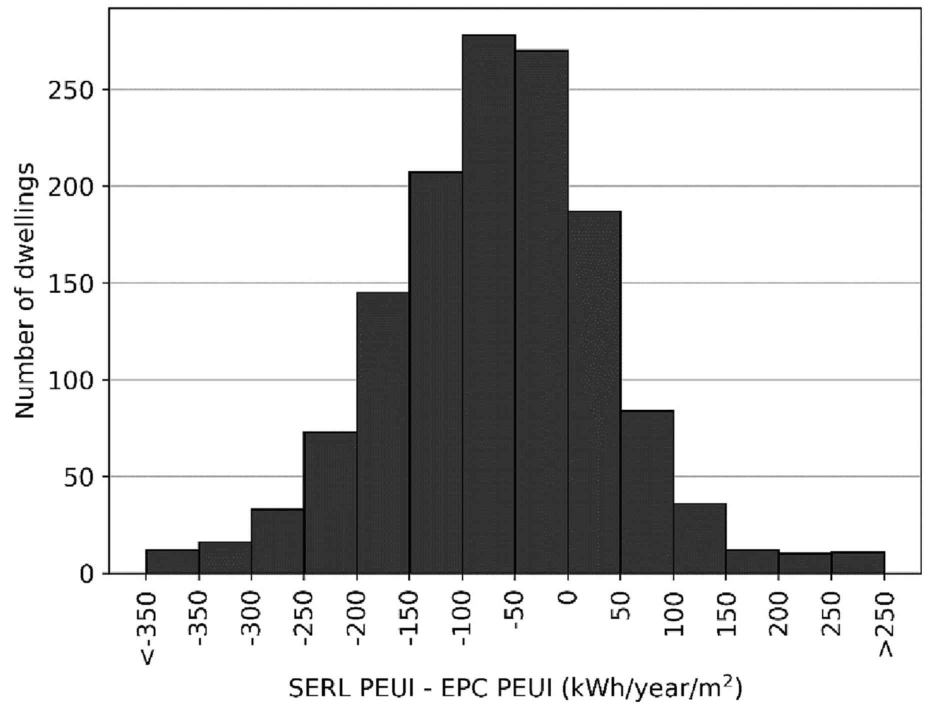

Fig. 3. Histogram of the per household difference between the PEUI using SERL metered data and the EPC modelled PEUI.

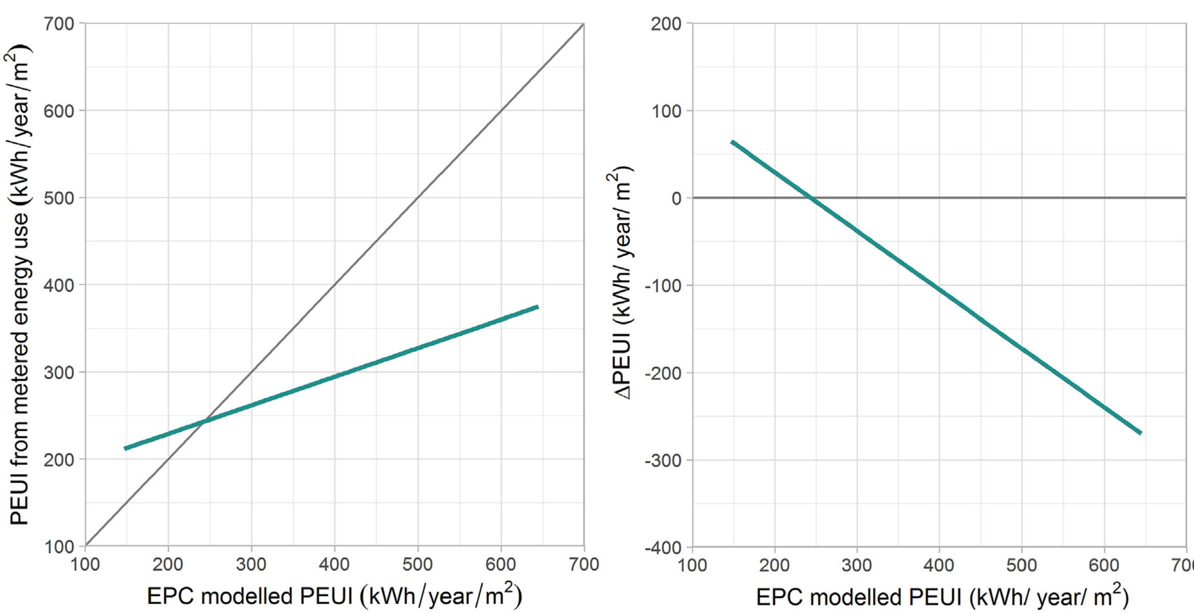

Fig. 4. Left panel: linear best fit of metered compared to EPC modelled PEUI, with the y = x line shown in black for comparison. Right panel: Linear best fit of the difference in PEUI against modelled PEUI. Note that for this and all following figures of linear models, the model is displayed over the range between the mean of the 10 lowest EPC modelled PEUI values and the mean of the 10 highest modelled PEUI values (the full range is not used for statistical disclosure control reasons).

Table 1
Results of the linear regression of the difference in PEUI according to metered energy use and EPC model against the EPC modelled PEUI (see equation (3) and right panel of Fig. 4).

Gradient [95% confidence interval]	Gradient p-value	Intercept (kWh/yr/m^2) [95% confidence interval]	Intercept p-value	RMSE (kWh/yr/m^2)	Number of participants
−0.67 [-0.74, −0.61]	<0.01	164 [-141, 186]	<0.01	95.4	1,374

metered PEUI than modelled, with the difference increasing as the EPC Energy Efficiency rating decreases to F and G. On average, F and G rated homes show a metered PEUI almost half that expected by the EPC model (298 kWh/yr/m^2 metered vs 573 kWh/yr/m^2 modelled). Note that a Wilcoxon test (suited to testing small samples) also indicated a significant difference for the F and G group.

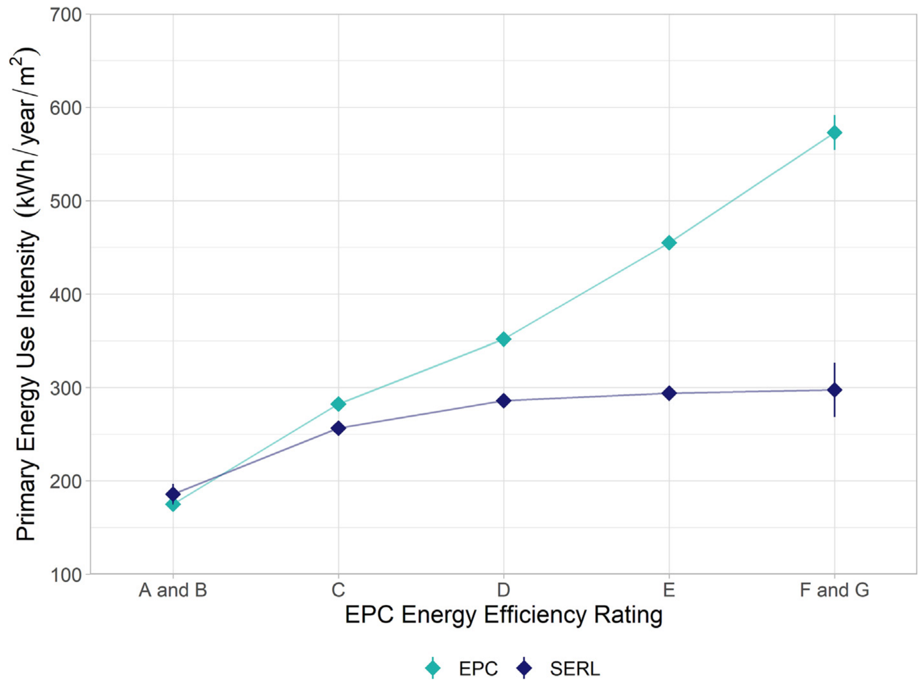

Fig. 5. Mean primary energy use intensity according to the SERL metered energy data in 2021 and the EPC model. Points represent the mean, and the bars represent the standard error.

Table 2
Results of a paired *t*-test of the mean difference between the EPC modelled and SERL metered primary energy use intensity in 2021.

EPC Band	Mean difference (kWh/yr/m^2) [95% confidence intervals]	P-value	Number of dwellings
A and B	10.6 [−10.7, 31.9]	0.32	32
C	−25.9 [−17.9, −33.9]	<0.001	439
D	−65.8 [−58.3, −73.4]	<0.001	705
E	−161 [−144, −178]	<0.001	185
F and G	−276 [−206, −345]	<0.001	13

The difference in metered PEUI between bands is much smaller than modelled; the difference in average metered PEUI for homes in band D compared to band E is 8.1 kWh/yr/m^2, compared to the modelled 103 kWh/yr/m^2, less than 10% of the expected difference.

Fig. 6 shows the linear best fit of metered PEUI against modelled PEUI and difference in PEUI against modelled PEUI for each EPC band. This shows that only in band A and B are changes in the modelled PEUI matched by equivalent changes in the metered PEUI (the residuals are reasonably normal and show no clear pattern). All other bands show a much shallower increase in metered PEUI with increasing modelled PEUI. Table 3 shows the regression coefficients and their significance for the linear models of the difference in PEUI. The only insignificant gradients are for bands A and B and F and G. The F and G group has a small sample size and wide confidence intervals for the gradient and intercept, with the central estimate of the gradient for the F and G group much further from zero than the A and B group. Only the A and B group show no significant gradient or intercept, and no significant difference between the metered and EPC modelled PEUI according to the paired *t*-test above.

Previous analysis has frequently found that homes rated as most efficient use more energy than the EPC predicts, whereas our analysis shows that on average EPC-modelled PEUI is fairly accurate for A and B rated homes. This could be because previous analysis has not accounted for appliance and cooking energy use which are not present in the model, whereas this analysis uses the EPC model of cooking and appliance energy use to provide an expected total PEUI (see Section 3.1.1). Since the most efficient homes are likely to be more thermally efficient, the proportion of their energy use associated with cooking and appliances will be greater than for low efficiency homes where energy use is dominated by heating. Excluding cooking and appliance use would decrease the modelled energy use proportionally more in high efficiency homes, and when the model is compared to metered energy use (which includes cooking and appliances) could be interpreted as suggesting that high efficiency homes use more energy than expected.

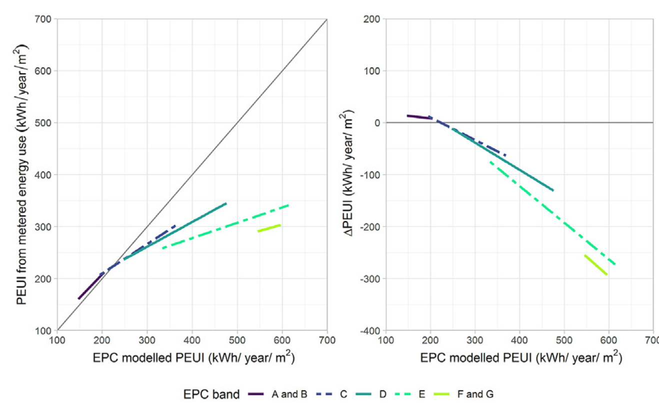

Fig. 6. Left panel: linear best fit for each EPC current energy rating band of metered compared to EPC modelled PEUI, with the y = x line shown in black for comparison. Right panel: Linear best fit for each EPC current energy rating band of the difference in PEUI against modelled PEUI.

Finally, both the A and B and F and G sample sizes were much smaller than the other groups. This increases the uncertainty asso-

Table 3
Results of linear regressions of the difference in PEUI according to metered energy use and EPC model against the EPC modelled PEUI (see equation (3) and right panel of Fig. 6).

EPC Band	Gradient [95% confidence interval]	Gradient p-value	Intercept (kWh/yr/m^2) [95% confidence interval]	Intercept p-value	Number of participants
A and B	−0.10 [−1.0, 0.80]	0.83	28 [−131, 187]	0.73	32
C	−0.43 [−0.65, −0.21]	<0.001	97 [34, 160]	0.002	439
D	−0.52 [−0.69, −0.36]	<0.001	119 [61, 177]	<0.001	705
E	−0.70 [−0.93, −0.48]	<0.001	160 [57, 263]	0.002	185
F and G	−0.76 [−1.78, 0.27]	0.13	157 [−434, 749]	0.57	13

ciated with their results, and future research on EPCs may benefit from purposively sampling to increase the proportion of less common bands in the sample.

4.2.2. Does the SAP modelled occupancy and heating use explain the difference between modelled and measured PEUI?

Fig. 7 and Table 4 present the best fit lines of metered PEUI and difference in PEUI against modelled PEUI for homes where the SERL survey responses suggest that the home is operated similarly to the SAP model assumptions (see Section 3.4.1), and for homes where this is not the case. Despite the relatively small number of homes matching the SAP model assumption (51), the gradient of the differences is significantly below zero for both groups (p less than 0.001), meaning they both show overprediction by the SAP model.

There could be remaining differences between the way the occupants use their homes and the SAP model assumptions. Occupants may heat their home for more or fewer hours than assumed by SAP (9 h per day in the week and 16 h per day at the weekend), even if they have the same set-point. Information on length of heating is not available for the majority of SERL Observatory participants. However, Huebner et al. [24] reported on the responses of 1,016 participants who took part in the SERL Covid-19 survey, these participants were drawn from the SERL Observatory participants who had been recruited by September 2019. They found a mean of 7.5 heating hours during the first lockdown and that less than 15% of respondents reported increasing their heating hours during this time. Additionally, the recent Energy Follow-Up Study (EFUS, a large interview and measurement survey of household heating patterns, thermal comfort and energy consumption) found that the median number of heating hours was 7 h during the week and 8 h at the weekend, and found no significant difference with EPC band [4].

While EFUS and the SERL Covid-19 survey results suggest that the SAP model is likely to overestimate the number of heating hours on average, note that the effect of this on the outcome of the SAP model is not straightforward. The heating hours are used to estimate a monthly mean internal temperature from which a semi-steady-state calculation is used to calculate the energy required for heating. The difference in mean internal temperature between the SAP model and reality would be required to understand whether the modelled heating hours results in a realistic mean internal temperature from which to estimate heating use. Calculating the modelled mean internal temperature would require the EPC input data.

Although SERL homes included in this analysis might be heated for fewer hours than the SAP model assumes, the occupants nonetheless report being comfortably warm, that they heat the whole house, and they report thermostat temperatures between 20.5 ^{0}C and 21.5 ^{0}C (higher than the 20 ^{0}C median found by EFUS [4]. Even in these homes with seemingly comfortable conditions, the increasing discrepancy with EPC-rated inefficiency remains an apparent trend. It thus seems unlikely that the trend is primarily caused by occupants deliberately 'under-consuming' by choosing less comfortable conditions in inefficient homes, suggesting that there are other reasons, either part of the SAP model or the SAP process, which are important contributors to the discrepancy.

4.2.3. Does the number of occupants or thermostat set point explain the difference between modelled and measured PEUI?

Fig. 8 and Fig. 9 show the best fit lines of metered against modelled PEUI by occupant number, and by reported thermostat setpoint, respectively. Previous analysis suggests that increases in these parameters are associated with increases in home energy consumption. This analysis shows increases in metered energy use with these parameters as expected. SAP models occupant number based on the floor area and assumes heating to 21 ^{0}C in the living room (other rooms depend on heating control and thermal efficiency), these assumptions are not always accurate, but across all values the gradient of the best fit lines remains very similar, suggesting that these do not account for the discrepancy between the model and the metered PEUI across the spectrum of energy efficiency ratings.

4.2.4. Does the reason for EPC generation (SAP or RdSAP procedure) explain the difference between modelled and measured PEUI?

As noted in Section 2, when EPCs are generated for a new home the full SAP procedure is performed, for which detailed inputs are provided using the details available at the time of construction. In contrast, when EPCs are generated for almost any other reason (normally sale or rent of existing property) an RdSAP procedure is performed, which can use default values for building parameters based on the building's age, unless evidence is provided that the defaults are not appropriate. We explored this by comparing homes in which the EPC was generated for a new home, and homes in which the EPC was generated for any other reason as shown in Fig. 10 and Table 5. The gradient of the difference against the modelled PEUI is not statistically significantly different from zero (with residuals reasonably normal and no clear pattern) in new homes. As noted for the A and B rated homes above, there was a relatively small sub-sample of new homes and future research may benefit from purposively sampling this group.

Young [44] reported on dwellings that were designed to be highly efficient receiving EPC A or B ratings when rated via the SAP process as new buildings, and subsequently receiving EPC C or D ratings when they were later reassessed via the RdSAP process. Cases such as this would result in homes inaccurately being placed in less-efficient EPC bands, with the result that the modelled energy use increases and possibly exceeds the metered energy use. Inputs to the model underlying the EPC rating may not always be accurate [26,13,33], and this may influence EPC outcomes.

While these results add weight to the argument that the process of calculating an EPC rating using the RdSAP procedure might result in systematic overprediction of energy consumption, it is important to note that there are other confounding factors associated with the group of new homes, we explore this further in the following section.

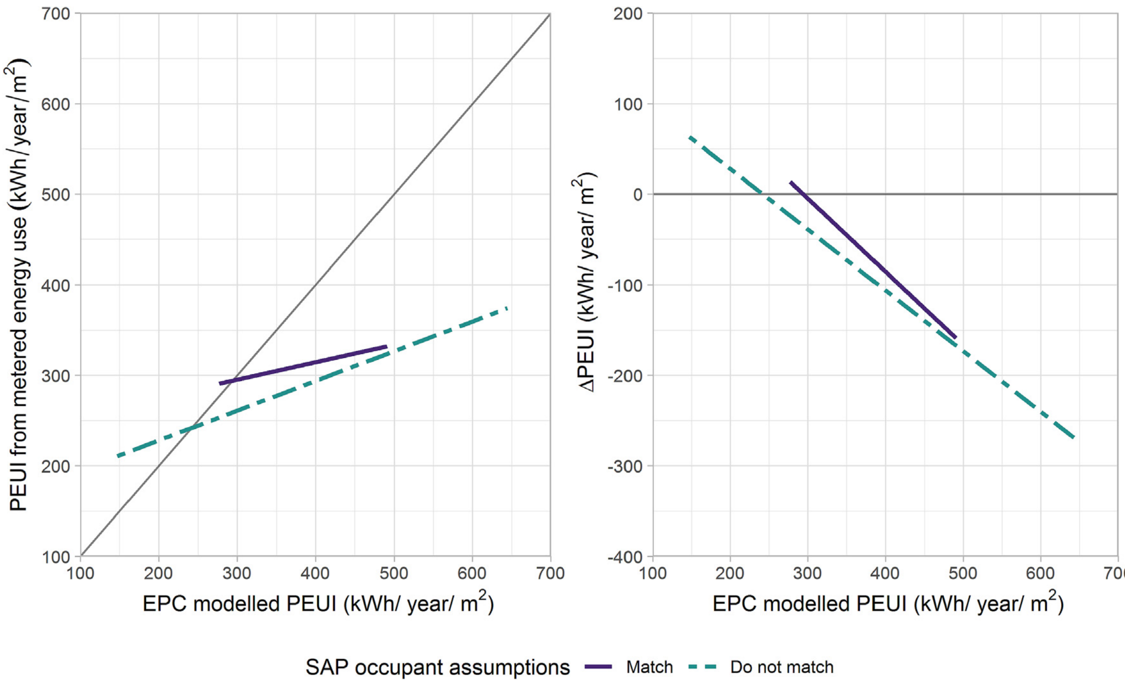

Fig. 7. Left panel: linear best fit of metered to EPC modelled PEUI for homes that match EPC model occupancy and heating use and homes that do not match. The y = x line shown in black for comparison. Right panel: linear best fit of the difference in PEUI against modelled PEUI for homes that match EPC model occupancy and heating use and homes that do not match.

Table 4
Results of linear regressions of the difference in PEUI according to metered energy use and EPC model against the EPC modelled PEUI (see equation (3) and right panel of Fig. 7).

Match EPC model occupancy and heating use	Gradient [95% confidence interval]	Gradient p-value	Intercept (kWh/yr/m^2) [95% confidence interval]	Intercept p-value	Number of participants
Match	−0.81 [−0.48, −1.13]	<0.001	237 [113, 361]	<0.001	51
Do not match	−0.67 [−0.61, −0.74]	<0.001	162 [139, 185]	<0.001	1323

4.3. Factors associated with a and B rated homes

The analysis in Section 4.2 showed that only homes in EPC bands A and B, and new homes rated via the full SAP process showed good agreement between the EPC modelled and metered PEUI. These are related because new homes are much more likely to be A or B rated than older homes. To investigate whether other factors are also associated, the factors in the publicly-available EPC data and the SERL survey responses associated with EPC bands A and B were explored using Fisher's exact test [1], with full results shown in the appendix. Fisher's exact test is used instead of the χ-squared test in examples where the number of cases is small. In the present case contingency tables of two variables have small numbers of cases in some groups, so Fisher's exact test was used to ensure a fair test for all variables. The null hypothesis is that the relative proportions of one variable are independent of the second variable. We accept a p-value of 0.05 as sufficient to reject the null hypothesis.

The results showed, as expected, that there was sufficient evidence to reject the null hypothesis and suggest an association between whether a home is A- or B-rated and the energy efficiency of all major building components (e.g. wall energy efficiency, floor energy efficiency), and whether the home is newly built. Building age and property type are also associated, with newer homes and flats more likely to be A and B rated. There was an association with region, with comparatively more A and B rated homes in the East of England and Yorkshire, and fewer in Greater London and the Northwest in our sample. Some EPC variables were not associated, such as the hot water efficiency, heating control and number of habitable rooms. Almost all occupant-related parameters from the SERL survey were not associated with the whether the home was A or B rated (whether participants were managing financially, the number occupants, number of appliances, whether heating controls were adjusted when the dwelling was empty, how frequently lights were switched off and windows opened, whether the whole house was heated and whether the occupants made an effort to reduce their energy use). The only exception was the temperature set point, where 21.5 ^{o}C to 23.5 ^{o}C was more likely in A and B rated homes. Finally, the floor area band (in ranges of 50 m^2) and the Index of Multiple Deprivation (IMD) quintile were not associated with A and B rated homes.

While it is unclear which factors are responsible for the good agreement between modelled and metered PEUI in A and B rated homes, it is worth noting several points. Firstly, many of the associated parameters are intrinsically related; for example, highly energy efficient homes tend to have highly efficient components (similarly with more recently built homes; current building regulations require a far higher efficiency than was required for old homes). Secondly, while it is interesting to note the association with temperature set point, Fig. 9 shows that homes with this higher set point have a higher metered PEUI, but the gradient between bands is still far shallower than suggested by the EPC process, so this is unlikely to account for the systematic difference across the range of predicted energy use. Similarly, the regional difference in distribution of EPC bands is unlikely to contribute to the causes of the discrepancy in metered and modelled energy use, and best fit lines for each region show little difference in gradients (results for this were omitted since they are similar to the results presented).

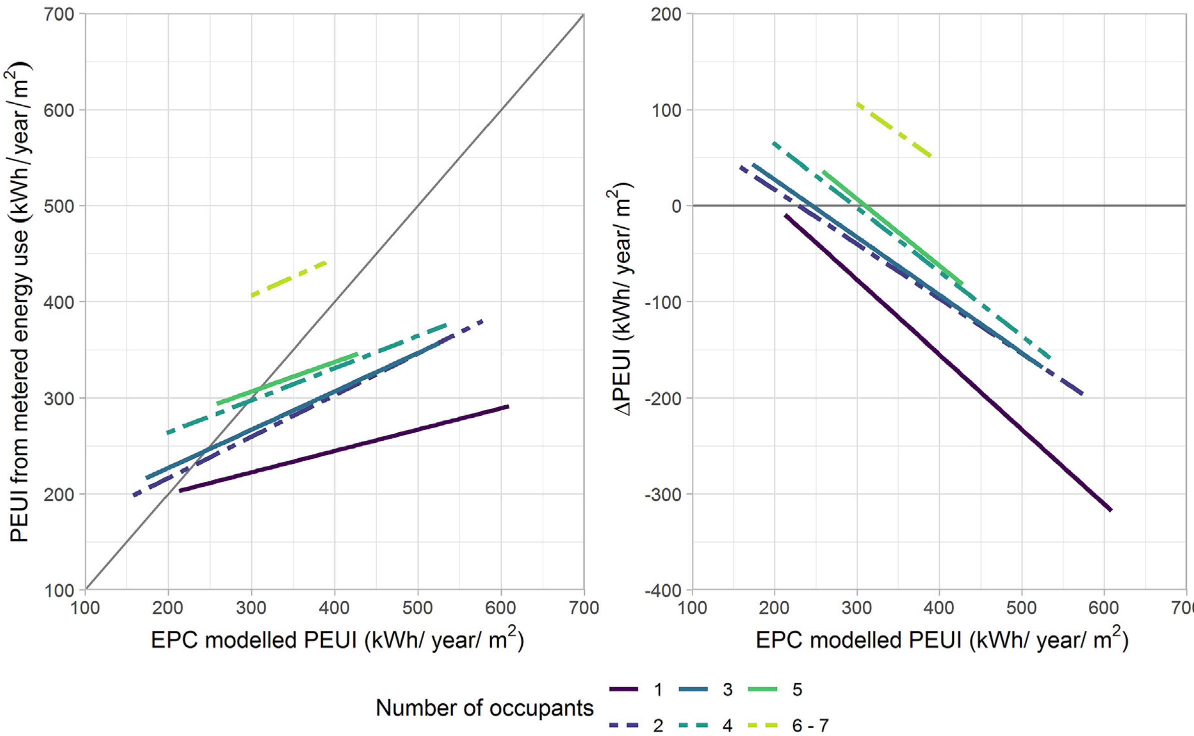

Fig. 8. Left panel: linear best fit of metered to EPC modelled PEUI for homes with different numbers of occupants. The y = x line shown in black for comparison. Right panel: linear best fit of the difference in PEUI against modelled PEUI for homes with different numbers of occupants.

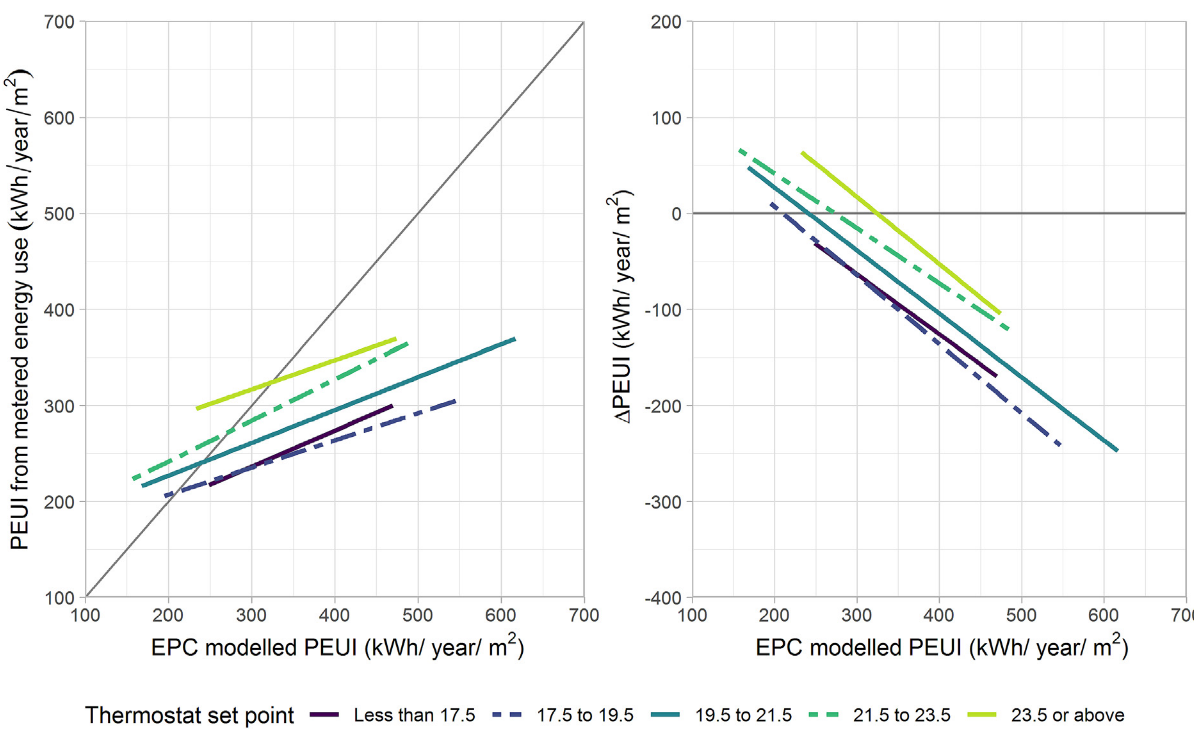

Fig. 9. Left panel: linear best fit of metered to EPC modelled PEUI for homes with different thermostat set point temperatures. The y = x line shown in black for comparison. Right panel: linear best fit of the difference in PEUI against modelled PEUI for homes with different thermostat set point temperatures in degrees Celsius.

5. Future work: What could cause the discrepancy between metered and EPC-modelled primary energy use intensity?

At present it is unclear which factors are responsible for the increasing discrepancy as energy efficiency rating worsens. However, alongside the issues discussed above, there are several plausible factors which will be explored in future work.

Previous research has suggested that aspects of the EPC assessment process could be subject to bias and measurement uncertainty. Hårsman et al. [20] studied EPC outcomes from individual

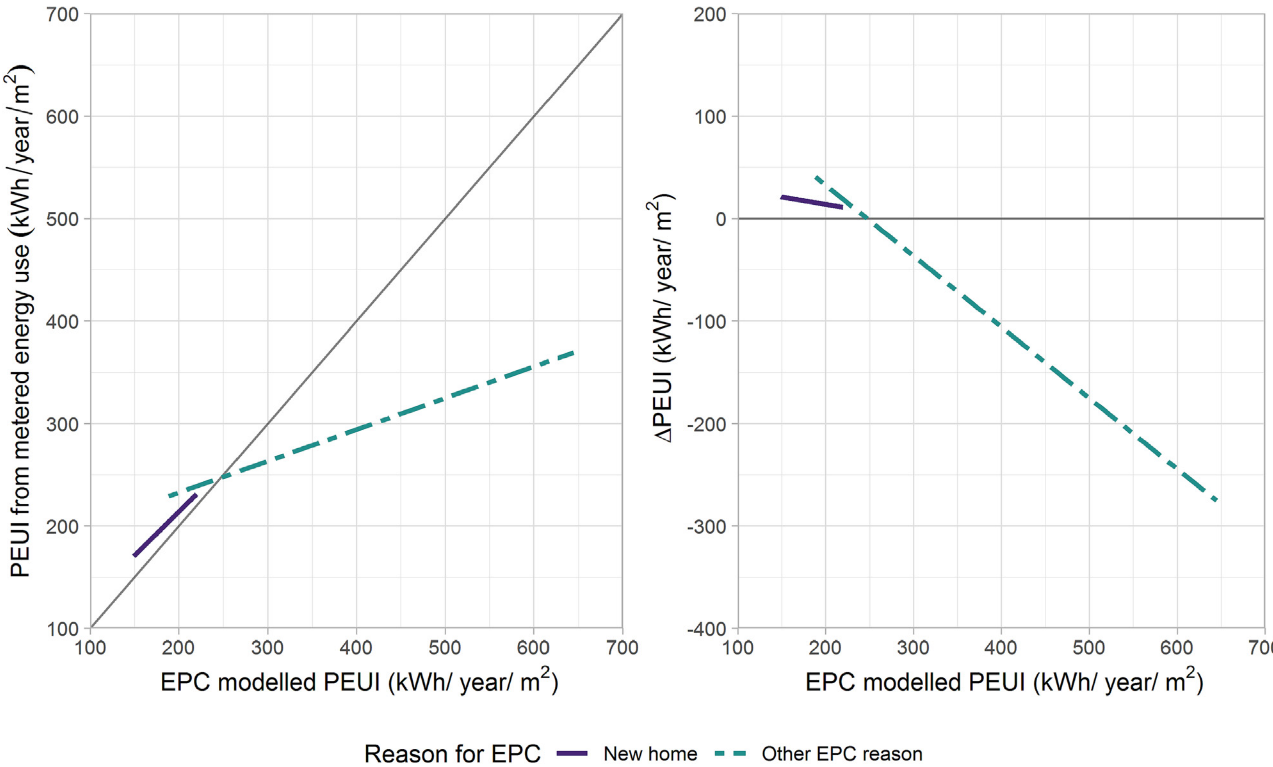

Fig. 10. Linear best fit of metered to EPC modelled PEUI for new homes (EPC generated via a full SAP process) and all other homes (EPC likely generated via an RdSAP process). The y = x line shown in black for comparison. Right panel: linear best fit of the difference in PEUI against modelled PEUI for new homes and all other homes.

Table 5
Results of linear regressions of the difference in PEUI according to metered energy use and EPC model against the EPC modelled PEUI (see equation (3) and right panel of Fig. 10).

EPC transaction type	Gradient [95% confidence interval]	Gradient p-value	Intercept (kWh/yr/m^2) [95% confidence interval]	Intercept p-value	Number of participants
New home	−0.14 [−0.80, 0.47]	0.65	42 [−78, 163]	0.48	33
Any other reason	−0.69 [−0.76, −0.62]	<0.001	171 [147, 195]	<0.001	1341

EPC assessment firms in Sweden, finding that individual firms systematically rated homes as more or less efficient than the central estimate. At present it is not known whether similar findings would be replicated in other countries using EPCs. In the UK, concerns have been raised over the repeatability of EPC ratings [26,13], finding that the same home rated multiple times does not produce the same outcome. Additionally, the possibility that assessors may experience pressure to provide a particular EPC assessment outcome has been raised [19]. The cumulative effect of these biases and uncertainties on the relationship between modelled and metered energy use is currently unclear.

EPCs are not automatically updated, so while installing more-efficient heating systems, or adding wall insulation, would improve the energy rating of the dwelling [34], it would not be reflected in the rating until a new EPC was generated [13]. This is unlikely to affect A and B rated homes since they are unlikely to have energy efficiency retrofits. The extent to which undocumented upgrades have an impact on the observed discrepancy will depend on the rate and degree of energy efficiency upgrades in the stock.

Small differences which are not recorded via the EPC process may make a significant difference. For example, a thick carpet on an uninsulated floor can reduce the heat loss considerably [29], and this would have most impact in an inefficient home.

Changes to the underlying SAP model or process are also not reflected in existing EPCs. Such changes include primary energy factors and default U-values, which could result in significant changes in the EPC outcome [3]. These are not reflected until a new EPC is generated, which complicates the comparison of EPCs which were generated using different versions of SAP. Moreover, the publicly available EPC data cannot easily be used to construct the values that would have been obtained had a different version been used. The present analysis attempts to limit this effect by only selecting EPCs generated using SAP 2012 (adopted from May 2014 in England and August 2014 in Wales) [7].

Distortionary effects associated with comparing the primary energy use intensity could impact the extent of the discrepancy between metered and modelled energy use. If participants have shifted much of their energy use to electricity after they completed the SERL survey or after the EPC was generated (e.g. due to installing a heat pump or using portable electric heater), then the PEUI could change considerably.

Moreover, different components of the SAP model (water heating, space heating, lighting, etc) are likely to have differing agreement with actual disaggregated energy use. For example, in a thermally efficient home with many occupants the proportion of energy used for hot water is likely to be high compared to heating demand, and vice versa for a thermally inefficient home with a small number of occupants, which may not be reflected in the model. It is important to note that good agreement between metered and modelled PEUI in A and B rated homes, and new homes, does not necessarily provide evidence that SAP accurately models disaggregated energy uses. Future research into the disaggregated components of the SAP model would help to identify such issues.

Limitations in the underlying thermal model within SAP could also be responsible for some of the discrepancy between metered and modelled energy use. The space heating calculation relies on an estimation of the mean internal temperature (MIT) [7]. The underlying model assumes set heating hours during which the temperature is modelled to be equal to the assumed set point (21 ^{0}C). When the heating switches off, the model assumes a linear cool-down, with the cooling rate set according to the thermal characteristics of the building. The resulting MIT is used to calculate the energy required for space heating. The analysis presented here was able to select homes in which the set-point temperature matched the assumption, but real homes may be unable to reach the temperature set-point depending on heating system output and building envelope thermal efficiency. Moreover, real homes will not heat up to the set point instantaneously and will cool down exponentially rather than linearly. As a result, the modelled MIT may be significantly different to the 'real' MIT. This may particularly affect less-efficient homes likely to heat slowly and cool quickly, with these homes likely to have a lower MIT than assumed. A lower MIT could decrease the amount of energy that these types of home would be expected to use according to the model. While there is some evidence that homes with different EPC ratings have different MIT [5], the extent to which this matches the EPC model, and so the extent of the impact on the discrepancy, are unclear as yet.

Finally, we note that a version of SAP, SAP 10.2, was introduced in 2022 and is now used for the full SAP procedure to generate EPCs for new homes [8]. Further changes are proposed for SAP11 and future versions of RdSAP [16]. EPCs for individual homes remain valid for 10 years so the version of SAP analysed in this research, SAP 2012, will remain the dominant rating method for most existing EPCs in the near to medium term. Moreover, until a new version of RdSAP is released the current version will be used to generate EPCs for most existing homes. Nonetheless, we plan to investigate the impact of new versions of SAP as the underlying models are made available.

6. Conclusions

This research compared EPC modelled primary energy use intensity with primary energy use intensity (PEUI) calculated from metered energy use using the EPC primary energy use factors in gas heated homes in Great Britain. While there has been considerable research into the difference between EPC modelled energy use and metered energy use, we believe the present analysis uniquely addresses three important shortcomings in the existing literature:

- Using the SERL survey we were able to exclude homes using unmetered forms of energy in the home (e.g. oil boilers)
- We compared quantities where the modelled and metered energy use were expected to be comprised of the same underlying elements. The publicly available EPC results for PEUI do not include appliance or cooking use; we calculated this 'missing' energy use with the equations used in the underlying EPC model.
- We compared the same quantity. The publicly available EPC results report PEUI, so we converted the metered energy use to primary energy use intensity using the same primary energy factors as SAP 2012 [7].

This meant the analysis was able to robustly compare the EPC modelled and metered PEUI, and in doing so we draw four main conclusions:

- On average, EPC modelled PEUI is significantly larger than metered PEUI (see Section 4.1)
- The discrepancy is statistically insignificant in A and B rated homes, but the discrepancy increases with decreasing EPC energy efficiency rating, to a mean difference of –276 kWh/yr/m^2, or –48%, in F and G homes (see Section 4.2.1).
- The gradient of the difference in measured PEUI against the modelled PEUI is significantly less than zero in bands C to E (between -0.4 and -0.7). Categories of homes which showed good agreement between the EPC model and metered PEUI were A and B rated homes and new homes (note the high overlap between groups) (see Section 4.2.1 and Section 4.2.4)
- The discrepancy remained even in homes where the occupancy and heating behaviours were matched as closely as possible with the SAP model assumptions (see Section 4.2.2)

It is important to note that this study is observational, what happens following energy efficiency retrofit in terms of EPC ratings, actual energy use and modelled energy use according to a new EPC assessment cannot be inferred from this work. Nonetheless, in the current context in which EPCs are widely used to support policy interventions regarding building energy efficiency, it is clearly significant that on average there is a far smaller difference in energy use between homes rated in different bands than the EPC predicts.

While the exact cause of the discrepancy between the metered and EPC-modelled PEUI is unclear, in the previous section we identified several factors which plausibly contribute to the discrepancy. The cause is likely to be a combination of factors whose influence varies between homes. Nonetheless, research has consistently shown systematically lower metered than modelled energy use in homes which are inefficient according to the EPC, with the difference increasing as EPC efficiency rating decreases. Given the heavy reliance on EPCs in a policy context, and the demonstration in this work that some of the commonly suggested reasons for the difference are unlikely to be correct, it is imperative that the causes of the discrepancy are identified so that future EPCs are fit for purpose.

Data availability

Data is available via the Smart Energy Research Lab (SERL), for information on accessing the data please visit: serl.ac.uk, research code will be made available via: github.com/smartEnergyResearchLab

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was funded by EPSRC (Engineering and Physical Sciences Research Council) grant number EP/P032761/1. Oreszczyn's input has been additionally supported by EPSRC grant EP/R035288/1.

There are over 30 individuals across 8 organisations in the SERL Consortium (University College London, the University of Essex (UK Data Archive), University of Edinburgh, Cardiff University, Loughborough University, Leeds Beckett University, the University of Southampton and the Energy Saving Trust) who have contributed to the development of SERL and thus the content of this paper. Particular thanks go to the SERL technical team at the UK

Data Archive: Darren Bell, Deirdre Lungley, Martin Randall and Jacob Joy and to SERL Consortium Manager James O'Toole at UCL. We acknowledge support from the SERL Independent Advisory Board, Data Governance Board and Research Programme Board which played critical role in the establishment and ethical operation of SERL. We are also grateful to the 13,000 + participants who have provided their consent to participate in the SERL Observatory.

Please note that neither the European Commission nor the European Centre for Medium-Range Weather Forecasts is responsible for any use that may be made of the Copernicus information or data it contains.

Appendix: Fisher's exact test result.

Table 6.

Table 6
Results of Fisher's exact test of the association between whether a home has an EPC rating of A or B and other relevant contextual variables. Variables where the p-value is less than 0.05 are considered to show a statistically significant association with homes that are rated A and B are shown in bold font. The source of the variable is noted, and for SERL survey variables the question number is shown.

Variable source	Variable name	p-value	Variable source	Variable name	p-value
EPC	**energyTariff**	**<0.01**	EPC	numberOpenFireplaces	0.30
EPC	**fixedLightingOutletsCount**	**<0.01**	EPC	propertyType	0.34
EPC	**floorDescription**	**<0.01**	EPC	heating_control_smart	0.37
EPC	**lightingEnergyEff**	**<0.01**	EPC	heating_control_manual	0.37
EPC	**lowEnergyFixedLightCount**	**<0.01**	EPC	lightingDescription	0.38
EPC	**mainheatcEnergyEff**	**<0.01**	SERL survey	B10_sum, number of appliance options selected from a list	0.40
EPC	**mainheatcontDescription**	**<0.01**	EPC	glazedType	0.48
EPC	**mainheatDescription**	**<0.01**	EPC	extensionCount	0.51
EPC	**roofDescription**	**<0.01**	EPC	numberHabitableRooms	0.52
EPC	**roofEnergyEff**	**<0.01**	SERL survey	A13_01, how often are lights switched off when leaving a room	0.52
EPC	**secondheatDescription**	**<0.01**	EPC	numberHeatedRooms	0.55
EPC	**transactionType**	**<0.01**	SERL survey	A11, are there unheated living spaces in winter	0.58
EPC	**wallsDescription**	**<0.01**	SERL survey	A14, how much effort to limit energy use	0.65
EPC	**wallsEnergyEff**	**<0.01**	SERL survey	A1502, how often are windows opened on a warm day	0.67
EPC	**windowsDescription**	**<0.01**	EPC	builtForm	0.69
EPC	**windowsEnergyEff**	**<0.01**	EPC	flatTopStorey	0.71
EPC	**building_age**	**<0.01**	SERL survey	D4, how do you manage financially	0.73
EPC	**hotwaterDescription**	**<0.01**	SERL survey	A13_02, how often are more clothes worn when cold	0.84
SERL survey	**A5_degC, reported thermostat set point in bands of 2°C**	**<0.01**	SERL survey	A1501, how often are windows open on a cold day	0.87
SERL participant summary	**Region**	**<0.01**	SERL survey	A10, how often are heating controls adjusted if the home is unoccupied	0.93
EPC	**totalFloorArea in bands of 50 m^2**	**0.01**	EPC	lowEnergyLighting	0.96
EPC	**heating_control_temp_setting**	**0.02**	EPC	glazedArea	1
EPC	localAuthority	0.06	EPC	mainheatEnergyEff	1
SERL survey	C1_new, number of occupants	0.14	EPC	mechanicalVentilation	1
EPC	hotWaterEnergyEff	0.17	EPC	multiGlazeProportion	1
SERL participant summary	IMD_quintile	0.20	EPC	tenure	1
EPC	lodgement_year	0.24			

References

[1] Agresti, A. (2002) *Categorical Data Analysis*. 2nd edn. John Wiley & Sons. 10.1142/9789811217524_0012.
[2] A.S. Anđelković et al., Building energy performance certificate—a relevant indicator of actual energy consumption and savings?, Energies 14 (12) (2021), https://doiorg/10.3390/en14123455.
[3] BEIS, Proposed Changes To Government's Standard Assessment Procedure Proposed Changes To Government's Standard Assessment Procedure (SAP) Available at: https://www.gov.uk/government/consultations/public-consultation-on-proposals-to-amend-the-standard-assessment-procedure-sap, 2016.
[4] BEIS, Energy Follow Up Survey: Heating patterns and occupancy Final report Available at: gov.uk/government/publications/energy-follow-up-survey-efus-2017-reports, 2021.
[5] BEIS, Energy Follow Up Survey: thermal comfort, damp and ventilation Final report Available at: gov.uk/government/publications/energy-follow-up-survey-efus-2017-reports, 2021.
[6] BEIS, Fuel Poverty Methodology Handbook (Low Income Low Energy Efficiency) Available at: gov.uk/government/publications/fuel-poverty-statistics-methodology-handbook, 2022.
[7] BRE, SAP 2012: The Government's Standard Assessment Procedure for Energy Rating of Dwellings, BRE, Watford, UK, 2014, Available at: https://files.bregroup.com/SAP/SAP-2012_9-92.pdf.
[8] BRE (2022) *The Government's Standard Assessment Procedure for Energy Rating of Dwellings Version 10.2*. Watford, UK. Available at: https://bregroup.com/sap/sap10/
[9] 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, https://doi.org/10.1080/09613218.2017.1312897.
[10] CIBSE (2015) *CIBSE Guide A: Environmental Design*. 8th edn. London.
[11] B. Coyne, E. Denny, Mind the Energy Performance Gap: testing the accuracy of building Energy Performance Certificates in Ireland, Energy Efficiency 14 (6) (2021), https://doi.org/10.1007/s12053-021-09960-1.
[12] S. Cozza et al., 'Do energy performance certificates allow reliable predictions of actual energy consumption and savings?, Learning from the Swiss national database', *Energy and Buildings* 224 (2020), https://doiorg/10.1016/j.enbuild.2020.110235.
[13] J. Crawley et al., Quantifying the Measurement Error on England and Wales EPC Ratings, Energies 12 (3523) (2019), https://doi.org/10.3390/en12183523.
[14] DLUCH, Energy Performance of Buildings Certificates Statistical Release October to December 2021 England and Wales Available at: https://www.gov.uk/government/statistics/energy-performance-of-building-certificates-in-england-and-wales-october-to-december-2021, 2022.
[15] Elam, S. *et al.* (2022) *Smart Energy Research Lab Observatory Data, 2019-2022: Secure Access. (Data collection)*. 6th edn. UK Data Service. SN:8666. 10.5255/UKDA-SN-8666-6.
[16] Etude (2021) *Making SAP and RdSAP 11 fit for Net Zero*. Available at: https://etude.co.uk/how-we-work/the-future-of-sap-calculations/.
[17] European Parliament (2010) *Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings (recast)*.
[18] J. Few et al., Smart Energy Research Lab: Energy use in GB domestic buildings 2021, Smart Energy Research Lab (Smart Energy Research Lab (SERL) Statistical Reports), London, UK, 2022.
[19] T. Gledhill, A study into the variability of UK domestic energy assessments, University of Salford, 2022.

[20] B. Hårsman, Z. Daghbashyan, P. Chaudhary, On the quality and impact of residential energy performance certificates, Energy and Buildings 133 (2016) 711–723, https://doi.org/10.1016/j.enbuild.2016.10.033.

[21] HMG (2015) *The Energy Efficiency (Private Rented Property) (England and Wales) Regulations 2015*. UK.

[22] M.Y.C. Van Hove et al., Large-scale statistical analysis and modelling of real and regulatory total energy use in existing single-family houses in Flanders, Building Research and Information (2022), https://doi.org/10.1080/09613218.2022.2113023.

[23] G.M. Huebner et al., Explaining domestic energy consumption – The comparative contribution of building factors, socio-demographics, behaviours and attitudes, Applied Energy 159 (2015) 589–600, https://doi.org/10.1016/j.apenergy.2015.09.028.

[24] G.M. Huebner et al., Survey study on energy use in UK homes during Covid-19, Buildings and Cities 2 (1) (2021) 952, https://doi.org/10.5334/bc.162.

[25] G. James et al., An introduction to statistical learning with applications in R. (2013), https://doi.org/10.1017/CBO9781107415324.004.

[26] D. Jenkins, S. Simpson, A. Peacock, Investigating the consistency and quality of EPC ratings and assessments, Energy 138 (2017) 480–489, https://doi.org/10.1016/j.energy.2017.07.105.

[27] R.V. Jones, A. Fuertes, K.J. Lomas, The socio-economic, dwelling and appliance related factors affecting electricity consumption in domestic buildings, Renewable and Sustainable Energy Reviews 43 (2015) 901–917, https://doi.org/10.1016/j.rser.2014.11.084.

[28] S. Kelly, D. Crawford-Brown, M.G. Pollitt, Building performance evaluation and certification in the UK: Is SAP fit for purpose?, Renewable and Sustainable Energy Reviews 16 (9) (2012) 6861–6878, https://doiorg/10.1016/j.rser.2012.07.018.

[29] A. Marshall et al., 'Investigating the Impact of Renewing Floor Coverings on the Energy Performance of Dwellings with Suspended Timber Floors, Tested under Controlled Conditions', Sustainable Design of the Built Environment, 2018.

[30] E. McKenna et al., Explaining daily energy demand in British housing using linked smart meter and socio-technical data in a bottom-up statistical model, Energy and Buildings (2022), https://doi.org/10.1016/j.enbuild.2022.111845.

[31] MHCLG, Energy Performance Certificates For Buildings - Action Plan Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/922660/EPC_Action_Plan.pdf, 2020.

[32] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2022, Available at: https://www.r-project.org/.

[33] K. Raushan, C. Ahern, B. Norton, Determining realistic U-values to substitute default U-values in EPC database to make more representative; a case-study in Ireland, Energy and Buildings 274 (2022), https://doi.org/10.1016/j.enbuild.2022.112358.

[34] A. Stone et al., Key factors determining the energy rating of existing English houses, Building Research & Information 42 (6) (2014) 725–738, https://doi.org/10.1080/09613218.2014.905383.

[35] A.J. Summerfield et al., What do empirical findings reveal about modelled energy demand and energy ratings? Comparisons of gas consumption across the English residential sector, Energy Policy 129 (March) (2019) 997–1007, https://doi.org/10.1016/j.enpol.2019.02.033.

[36] M. Sunikka-Blank, R. Galvin, Introducing the prebound effect: The gap between performance and actual energy consumption, Building Research and Information 40 (3) (2012) 260–273, https://doi.org/10.1080/09613218.2012.690952.

[37] M. Szulgowska-Zgrzywa et al., Impact of users' behavior and real weather conditions on the energy consumption of tenement houses in Wroclaw, Poland: Energy performance gap simulation based on a model calibrated by field measurements, Energies 13 (24) (2020), https://doi.org/10.3390/en13246707.

[38] E. Webborn et al., The SERL observatory dataset: Longitudinal smart meter electricity and gas data, survey, EPC and climate data for over 13,000 households in Great Britain, Energies 14 (21) (2021), https://doi.org/10.3390/en14216934.

[39] E. Webborn et al., (2022) 'Increasing response rates and improving research design: Learnings from the Smart Energy Research Lab in the United Kingdom', Energy Research and Social Science 83 (November 2021), https://doi.org/10.1016/j.erss.2021.102312.

[40] H. Wickham, Reshaping data with the reshape package Available at: Journal of Statistical Software 21 (12) (2007). https://www.jstatsoft.org/v21/i12/.

[41] H. Wickham, ggplot2: Elegant Graphics for Data Analysis, Springer-Verlag, New York, 2016, Available at: https://ggplot2.tidyverse.org.

[42] H. Wickham et al., dplyr: A Grammar of Data Manipulation. R package version 1.0.8, Available at (2022). https://cran.r-project.org/package=dplyr.

[43] C.O. Wilke, cowplot: Streamlined Plot Theme and Plot Annotations for "ggplot2", Available at (2020). https://cran.r-project.org/package=cowplot.

[44] J. Young, Towards Zero Energy Buildings: Lessons Learned From the Bedzed Development, University College London, 2015.