An Analysis of Errors in the Energy Performance Certificate Database

A. Hardy[a,*], D. Glew[a]

[a] *Leeds Beckett University, Leeds, LS2 9EN*

Abstract

Energy Performance Certificates (EPCs) are the adopted method by which the UK government tracks the progress of its domestic energy efficiency policies. Over 15 million EPCs have been lodged, representing a valuable resource for research into the UK building stock. However, the EPC record has a reputation of containing multiple errors. In this work, we identify many such errors and quantify how common they are. We find that 27% of EPCs in the open EPC record display at least one flag to suggests it is incorrect and estimate the true error rate of the EPC record to be between 36-62%. Many of these errors are caused by EPC assessors disagreeing on building parameters such as floor type, wall type and built form. Additionally, flats and maisonettes appear to cause more issues than other property types. This may be due to difficulties in assessing their location in the building and the nature of the surrounding space. We also suggest potential new methods of quality assurance which rely on machine learning and which could allow such errors to be avoided in the future.

Keywords: Energy Performance Directive, Energy Performance Certificates, Retrofit, England, Wales, Dwellings

*Corresponding author
Email address: A.L.Hardy@leedsbeckett.ac.uk (A. Hardy)

Accepted by Energy Policy *April 17, 2019*

1. Introduction

Energy performance certificates (EPCs) are estimated records of a building's energy efficiency and were introduced in response to the EU Energy Performance in Building Directive (EPBD) [1]. The majority of homes in the UK now have EPCs, and they must be produced whenever a home is sold, rented or assessed for certain funding schemes. To generate an EPC, information is gathered about a building such as wall type and levels of insulation, and these characteristics are fed into the Government's Standard Assessment Procedure (SAP) algorithm. The SAP algorithm then makes assumptions about the thermal properties of the building fabric and occupancy to calculate the theoretical heat loss of the property. For new build homes, a full SAP calculation is conducted, whereas pre-existing homes are assessed with a reduced Standard Assessment Procedure (rdSAP). In either case, the SAP algorithm assigns an efficiency rating (EER) out of 100, and also places the property in an EPC band from A to G (see Table 1), where A rated properties are the most energy efficient.

Table 1: Selection of parameters in EPC dataset

Energy Efficiency Rating	EPC Band
≥ 92	A
81 - 91	B
69 - 80	C
55 - 68	D
39 - 54	E
21 - 38	F
1 - 20	G

EPCs are a commonly used resource for UK policy making. For example, EPCs are integrated as a metric for energy efficiency under the Minimum Energy Efficiency Standards (MEES) regulations for the Private Rented Sector (PRS) [2]. EPCs are also the means by which carbon savings are estimated and remunerated in the UK's Energy Company Obligations (ECO) retrofit policy

[3] and the Renewable Heat Incentive [4]. EPCs are not unique to the UK, however. Across Europe, certification processes similar to EPCs are used to describe the efficiency of a nation's housing stock, inform on fuel poverty of occupants and predict the impact of retrofits schemes [5, 6]. The processes used in other countries to generate energy certificates is similar to the process in the UK, in that an algorithm is used to place properties in bands A - G. Some countries, such as France, also add provision to use measured information from fuel bills in the calculation of EPCs [7].

EPCs may also be having impact beyond policy evaluation. Across the EU and the USA, there is some evidence of a price premium for more energy efficient homes [8, 9, 10, 11]. Evidence of the Welsh housing market suggests that A-rated houses could achieve a 12% premium relative to a D-rated home, where as F-rated houses may incur a 6% penalty [12]. The UK government commissioned research to investigate this effect and concluded that A-rated homes could sell for 14% more than equivalent G-rated dwellings [13].

To deliver the millions of EPC assessment surveys that have taken place in the UK, thousands of assessors needed training over a short period of time. A "Domestic Energy Assessor" (DEA) qualification can therefore be achieved relatively quickly without prior surveyor experience. A quality assurance procedure was implemented alongside the introduction of EPC assessments which requires that a sample of the EPCs created by each DEA undergoes a desk based audit. The auditor creates their own version of the EPC based on the data available, and 95% of audit EPCs must be within ± 5 points of the original value [14]. Although EPCs are integrated in to reporting against policy objectives, several studies have highlighted that EPCs may be less accurate than demanded in the auditing procedure. For example, an EPC "Mystery Shopper" study was conducted, in which multiple DEAs assessed the same properties [15]. The average range in EER of all properties was 11 points, and one property reported a range in excess of 30 points. These large ranges were mainly due to differences in assessors interpretation of certain values such as the age band and building form types[15]. Other research supports this finding, and suggests that the ability to

compare individual EPCs with each other may be problematic due to a lack of consistency in assessor quality [16]. Lack of reliability and accuracy in EPCs may not be only a UK phenomenon; The EPCs for Irish properties suggest a discontinuous distribution of EERs, with bunching around the beginning of a EPC band being observed. This bunching may indicate home-owners undertaking retrofits to get them just over the lines or, alternatively, assessors manipulating EPC inputs to ensure the client receives a more positive result [17].

Given the use of EPCs in influencing and assessing policy, the low data quality may be of concern. Additionally, it has been suggested that the impact EPC ratings have on house price in the future could be more substantial if it was seen to be a more accurate reflection of energy efficiency and given more prominence on the transaction process [18]. In this work, we identify EPCs which appear to contain errors and could constitute poor quality data, and assess how common this low-quality data is within the EPC record. Removing the EPCs identified as errors will likely improve the validity of any future analysis. Additionally, the results of this work could direct future quality assurance procedure to ensure a higher standard of data quality is achieved as new EPCs are lodged.

2. Data Analysis

To generate an EPC, a qualified DEA must visit the property and note down key building characteristics. These characteristics are uploaded into SAP software which generates the EER, EPC band and makes certain recommendations for how energy efficiency could be improved. These data are then lodged with a server, which consolidates EPC data from all assessors across the UK. A database of EPC EERs and building characteristics was published in April 2017 as an open access record. A sample of the building parameters present in the public EPC record is displayed in Table 2.

To determine the methods by which errors can be identified in this raw data, a consultation was conducted with professionals responsible for the training and certification of EPC assessors. From these discussions several patterns were

identified in the data as being anomalous. These patterns were consolidated into five different groups; A) Anomalies which appear to be based on a lodgement error, B) Anomalies based on an analysis of building elements such as walls and floors, C) Anomalies based on an analysis of building design such as built form, D) Anomalies based on variables relating specifically to flats, such as their location in the building and E) Anomalies in which installed energy efficiency products have apparently been removed. A final group F) was also created to experiment on if machine learning could be used to identify anomalies. When these error groups were defined, a script was written to identify potential errors within them, as described in detail below.

Table 2: Selection of parameters in EPC dataset

Variable	Description
Energy Efficiency Rating	Output from SAP which describes efficiency of property (out of 100)
Inspection Date	Quoted date of EPC inspection
Lodgement Date	Date on which the EPC was lodged
Property Type	Describes if the property is a house, flat, etc.
Built Form	Describes if the building is detached, semi-detached etc.
Floor Description	Describes if the floor is solid or suspended, along with any insulation
Walls Description	Describes the construction of the walls, including any insulation.
Roof Description	Describes if there is a flat or pitched roof, along with any insulation
Total Floor Area	Floor area of the property in m^2

2.1. Error Group A) Possible Lodgement Errors

When two independent EPCs are conducted for a property, an optimal result would be for the EPCs to agree. However, similarity in certain variables suggests an error has occurred, as explained below:

Error Code 1 - Identical duplicates

The EPC record contain many duplicate entries, where a duplicate entry was defined as an EPC where all parameters are equal, with the exception

of date information and unique certificate identification variables. Given the large number of inputs into a SAP assessment, exact agreement on all input variables, particularly continuous variables such as floor area, is highly unlikely. These duplicate entries may therefore exist as a result of a software or network fault which erroneously lodges the same EPC twice. Alternatively, a DEA may unintentionally submit an EPC twice. These errors were identified by hashing all relevant variables together then searching for duplicate hashes. If any duplicate hashes were found, one EPC was assumed to be genuine and all others were flagged as error code 1.

Error Code 2 - EPCs with the same inspection date

Some EPCs are lodged on separate dates and have values which suggest two separate inspections. However, the inspection date for these two EPCs is sometimes identical. The inspection date can be set by the inspector, and identical inspection dates may therefore be caused by human error. An alternative possibility is that such EPCs may be the result of a failed audit. Regulations stipulate that a minimum of 2% of EPCs lodged by an assessor undergo a desk-based audit. If an EPC fails this audit, the EPC assessor is required to re-lodge an EPC after correcting the mistakes identified. Procedure dictates that the DEA should then request to have the failed EPC removed, but it is feasible that DEAs do not always carry out this step. This would cause both EPCs to remain in the data. These errors were identified by hashing the inspection date and building reference number of all EPCs. If duplicates were identified, the more recent EPC (in terms of lodgement date) was assumed to be correct, and others were marked as error code 2.

Error Code 3 - EPCs lodged within a week

Any EPCs which were replaced within a week by a new EPC were treated as potential errors and given error code 3. It is not expected that EPC assessments for a property will frequently occur with less than a week between inspections. Instead, these errors may result from a DEA being made aware of a mistake in

the initial EPC and lodging a new EPC to correct the error. The latest EPC is assumed correct and not marked with any error code.

2.2. Error Group B) Building Structure Discrepancies

Some building characteristics may change over the course of their life, such as the heating system or amount of low-energy lighting. Most structural characteristics, however, are not likely to change. If two EPCs were conducted on the same property, discrepancies in these characteristics were used to identify potential errors.

Error Code 4 - Differing Floor Type

The EPC dataset contains data on the property floor type which takes a general value of "Solid" or "Suspended Timber". In some cases, EPC assessors are able to add more information to these general classifications, detailing the presence of insulation, or assigning a U-value directly to the floor. Some EPCs will also list two floor types for the property (e.g. part solid and part suspended timber floor) which could, for example, occur in properties where a cellar exists under one room.

For a particular property, it is unlikely that the general structure of a floor will change over time - a suspended timber floor property is unlikely to turn into a solid floor property. Similarly if a property has a cellar it is unlikely for this to change. If two EPCs for a property suggest two very different floor constructions, it is therefore likely that one of these EPCs is incorrect. Regular expressions were used to search for the terms "Suspended" and "Solid" within the floor type variable. If a disagreement on the general floor construction was found, all EPCs for that property were marked as a potential error. Errors of this type were given error code 4.

Error Code 5 - Differing Wall Type

Wall types in the UK are typically either of a solid brick or a cavity wall construction, though properties can also have a wall type of timber, stone, and

system build. As with floor types, properties can show two different wall types, particularly if a modern extension has been built. In general however, the main wall type of a property is not likely to change over time and any disagreement on wall type is suggestive of an error. Regular expressions were used to extract the wall type of all EPCs for a property. Any EPCs which disagree on wall type were flagged with error code 5.

Error Code 6 - Disappearing pitched roof

The conversion of a flat roof to a pitched roof is a common occurrence, as a pitched roof affords better insulation, additional space and is more resistant to the elements. However, the conversion of a pitched roof to a flat roof is not likely, and any EPCs which displayed this behaviour were therefore flagged with error code 6.

2.3. Error Group C) Building Design Discrepancies

The general design of a property and its relation to neighbouring properties is recorded during an EPC assessment. In general, these values are unlikely to change over time. Discrepancies in these values for properties with more than one EPC were therefore used to identify potential errors.

Error Code 7 - Differing Property Type

The property type for an EPC can take a value of either House, Flat, Bungalow, Maisonette or Park Home. A building may change is property type by, for example, conversion from house to flat. However, the rate of such conversions in low in the UK, amounting to approximately 5,000 per year [19] (0.5% of all EPCs lodged in a typical year). The property type is therefore not likely to change for a given property, and any EPCs displaying discrepancies on property type were assigned error code 7. It should be noted that some ambiguity surrounds the definition of what constitutes a Maisonette compared to a Flat. Some disagreement between these two property types was therefore expected.

Error Code 8 - Differing Built-Form

The built form of a property describes if it is detached, semi-detached, or terraced. For terrace properties, this parameter further describes the properties location in the terrace (i.e. mid-terrace, end-terrace etc.). It is possible that the built form of a property will change over time, as a detached property could change into two semi- detached properties, but the rate of this change is likely to be low. Any properties whose EPC show a disagreement on built form were therefore marked with error code 8.

2.4. Error Group D) Discrepancies in Flat Parameters

If the EPC assessment is being conducted on a flat, several additional variables are recorded. These additional variables describe the flats location in the building and the nature of the surrounding space. Several of these values are not expected to change over time and can be used to identify potential errors, as described below.

Error Code 9 - Differing Flat Floor Level

For EPCs conducted on flats and maisonettes, the "Floor Level" parameter describes the flat's location in the building. This parameter is important in SAP as, for example, ground floor flats are assumed to lose heat through the ground while other flats will not. Any flats which display EPCs that disagree on their floor level are potential errors, and were marked with error code 9.

Error Code 10 - Differing Top Story Flag

The EPC data for flats and maisonettes contains an additional parameter which flags if it is a top-story property. Again, this is a significant parameter as top-story properties can lose additional heat through their roof. Any EPCs which show a discrepancy in this variable were given an error code of 10.

2.5. Error Group E) Reduced energy efficiency products

The presence of energy efficiency products can change over time. Indeed, it has been an aim of the UK government to increase the presence of energy efficiency products through several incentives such as the Energy Company Obligation[3]. While such products can be installed, many products cannot be easily removed and there is little incentive to do so. A noted reduction in energy efficiency measures may therefore indicate an error has occurred.

Error Code 11 - Reduced Wall Insulation

Wall insulation can be present either as solid-wall insulation or cavity wall insulation. Once wall insulation has been installed it is unlikely that it will be removed, particularly in the case of cavity wall insulation. If one EPC for a property suggests the existence of wall insulation and a later EPC suggests there is no insulation, one of these EPCs is likely incorrect. As it cannot be known from the data which EPC is incorrect, any properties which display this behaviour were marked with error code 11.

Error Code 12 - Reduced Glazing Performance

The "Glazed Type" parameter in the EPC dataset describes if the property possesses single, double or triple glazing. Again, improvement from single to either double or triple glazing is possible and encouraged due to the increased energy efficiency this entails. Conversion from double glazing to single glazing is considerably less likely. In fact, modern building regulations define a maximum allowed value for the heat loss of any new windows installed and, in the vast majority of cases, single glazed windows will not meet this criteria [20]. Conversion from double to single glazing is therefore not likely. Any EPCs which suggest a house has decreased the amount of glazing it possessed to single glazing were marked with error code 12.

Error Code 13 - Decreasing loft insulation

Loft insulation is a significant factor in the calculation of EPCs as a great deal of heat can be lost through a loft. Loft insulation is likely to increase during

home improvements, but decreases in loft insulation are less likely. Any EPCs which display this behaviour were therefore flagged with error code 13.

Error Code 14 - Decreasing Energy Efficiency rating

EPC energy efficiency ratings may increase over time, but should not typically decrease. However, a certain degree of variation in an EPC energy rating is expected even if no changes are made to a property. This variation can occur as assessors may disagree on variables such as floor area or boiler efficiency. The quality assurance procedure for EPCs stipulates that when an EPC is conducted on an identical property, the score should be within ± 5 points 95% of the time, or within ± 10 points 99.99% of the time. If the EPC rating for a property decreases by more than 10, it is therefore potentially an error. However, a legitimate decrease in EPC could occur if, for example, extensions were built onto a property. The discrepancies identified in this manner should therefore be treated with some caution, as the lower energy efficiency may be legitimate. Any EPCs which display this behaviour were flagged with error code 14.

2.6. Error Group F) Random Forest (RF) Errors

If multiple EPCs disagree on a building characteristic then one is likely erroneous. The previous error identification methods could not discriminate between the correct and incorrect EPCs and all were therefore marked as potential errors. A more sophisticated approach is to isolate only the incorrect EPC based upon how different it is from other EPCs. For example, if one EPC lists a house as having a pitched roof and four other EPCs say is has a flat roof, the pitched-roof EPC is likely incorrect. However, the choice of how different an EPC must be to be classed as an error is not immediately obvious. For example, does an EPC have to disagree with 90% of other EPCs before it can robustly be classified as an error, or is 80% disagreement sufficient. To overcome this, we trained a random forest to identify errors automatically. Random forests are built from decision trees, and decision trees apply rules such as "If an EPC's roof is different to 80 % of the other EPCS, it is an error". These rules will be

chosen by decision trees in such a way that the most EPCs can be correctly classified as errors. However, single decision trees are prone to overfitting and a random forest mitigates this by using thousands of different decision trees. Each tree can then individually vote on which EPCs they think are errors and any instances of overfitting will be compensated for by those decision trees that did not overfit that particular parameter. All of the decision tree votes are then tallied to determine the final classification for the EPC.

Two random forests (RFs) were created, each exploiting different patterns that are expected in EPC data. These random forests are explained below.

Error Code 15 - RF on Multiple EPCs for a single house

This RF was trained to search for EPCs which appear erroneous when compared to other EPCs from the same property. To achieve this, several additional variables were calculated to describe how similar an EPC is to others. For example, the quoted floor area of an EPC was divided by the average floor area of all EPCs for the property. If all EPCs agree on floor areas, this variable would therefore take a value of 1. A training dataset for this RF was then created by taking real EPCs for a property and substituting in values from a different property. This substitution was known to be an "error", allowing the training and testing of the random forest. However, one limitation with this method is that the input data may naturally contain errors that were not classified as such in the training data. To mitigate this, we remove any EPCs which were identified by the previous error methods before training the RF. The RF identified 80% of the errors present in the training data. It misclassified 7% of the "correct" data as errors. The full list of variables included in the model is displayed in section 5. Any errors identified by this random forest where given the error code 15.

Error Code 16 - RF on EPCs of neighbouring houses

In the UK, it is common for an entire street to be built at the same time and to similar specifications. As a result, it is common for properties to be similar to their neighbours, such as a a terraced property with solid walls being surrounded

by similar terraced houses. Several additional variables were therefore calculated for each EPC to describe the similarity of a house compared to houses within the same postcode. A training dataset was created by taking real EPCs for a street and substituting in a random EPC from a different location. Again, the input data may naturally contain errors that were not classified as such in the training data and we mitigate this by removing any EPCs which were identified by the previous error methods before training the RF. The RF identified 62% of the errors present in the training data. It misclassified 7% of the "correct" data as an error. Further details on this random forest can be found in the appendix. Dwellings identified by this RF were classed as error code 16.

3. Discussion

In this section, an analysis of the identified errors is conducted to determine if errors occur more frequently in certain dwelling types or areas of the country. Because many of the errors are identified by comparison with a second EPC for the same property, any dwellings which only have one EPC are less likely to be flagged as a potential error. To mitigate this effect, a sub-sample was extracted from the data which removed any properties for which only 1 EPC was available. Furthermore, a property with more EPCs is more likely to have a potential error identified. For example - a property with 10 EPCs is more likely to have an error identified than a property with only 2 EPCs. If certain properties types, such as flats, have more EPCs than other property types then more errors would be identified for flats simply by virtue of them having more EPCs. To account for this effect and allow comparison between different property variables, we reduce the sub-sample further to only include properties for which exactly 2 EPCs were available.

3.1. Total Error rates

Figure 1 shows the incidence of each error code in the EPC record as a whole. In total, 27% of all EPCs display at least one flag to indicate it has a potential

error. The error percentages for the sub-sample is displayed in Figure 2. As expected, the total error rate in this sample is higher as these EPCs have more opportunity to be identified as an error. Error type 15 does not occur in the sub-sample however, as the RF that identified these errors requires more than 2 EPCs for a property.

Within the sub sample, 62% of the EPCs displayed at least one error code. This percentage reflects the amount of EPCs which would have to be removed from the sub-sample to ensure all known errors are accounted for. However, the real percentage of errors in the sub sample is likely to be lower. This is because, if a disagreement exists between building parameters of two EPCs, one EPC of them is potentially correct. The algorithm used in this work typically marked both EPCs as an error as distinguishing between the correct and incorrect EPC is not possible with certainty. Only the RF methods were able to make a choice between which EPC is likely to be correct. The value of 62% should therefore be considered an upper limit on the total errors present in the sample. In the best-case scenario that one EPC is always correct, the percentage of errors in the sub-sample would be reduced to 36% . In reality, the percentage of errors present in the EPC data as a whole likely lies in-between these values.

Both figures suggest that different error codes occur in different numbers, with particularly common errors being codes being 4, 5, 8 and 13; These correspond to disagreements on floor type, wall type, built form and decreasing loft insulation. The high incidence of EPCs which disagree on floor type and wall type is perhaps surprising, as these parameters should be relatively simple to assess. Properties which show two types of floor or wall types due, for example, to the presence of a cellar or extension may cause some of these discrepancies. In such properties the required procedure is to split the assessment into "main" and "extension", allowing both constructions to be reflected [21]. The incidence of this error code is perhaps indicative of this procedure being inconsistently applied or interpreted differently between assessors.

The high incidence of EPCs which disagree on built form is also surprising, as built form (i.e. whether the property is detached, semi-detached, mid terraced,

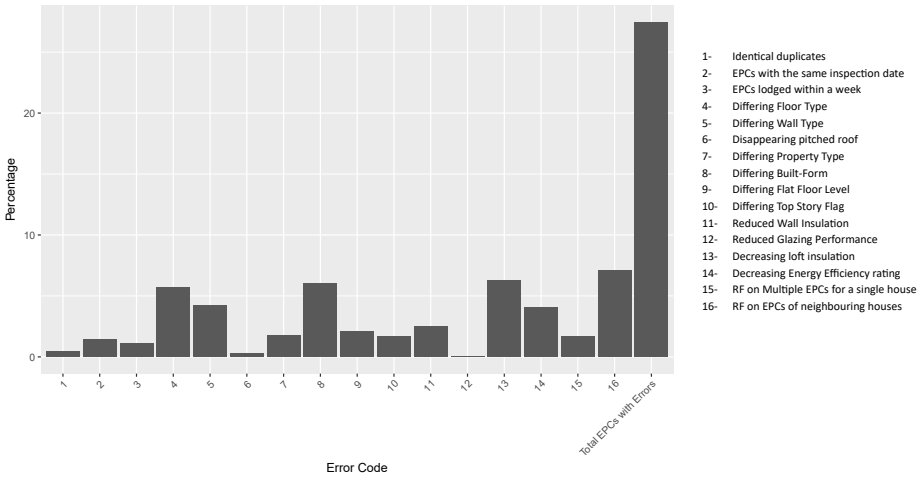

Figure 1: Total error percentages in the whole EPC dataset. Each number corresponds to an individual error type identified. In total, 27% of all EPCs analysed show at least one error. Uncertainties on these bars were estimated using an exact binomial test and were found to be negligible.

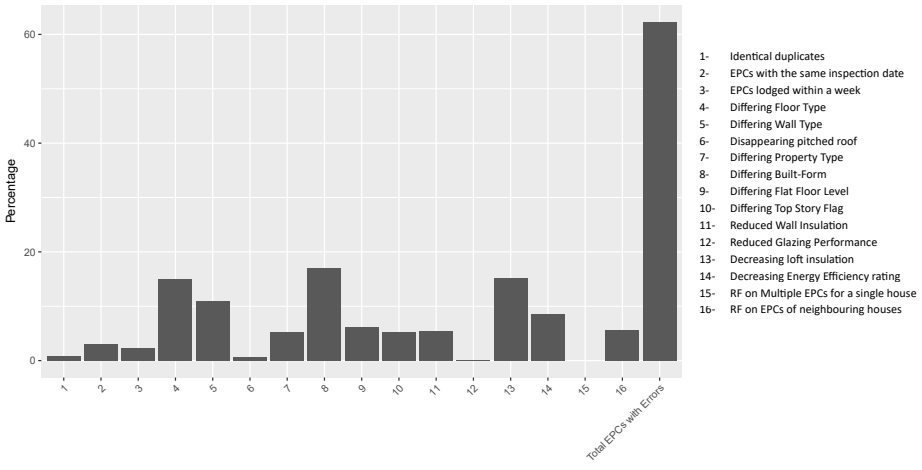

Figure 2: Total error percentages on EPCs which have been repeated for a property. Each number corresponds to an individual error type identified. In total, 62% of EPCs analysed within this sub sample show at least one error. Uncertainties on these bars were estimated using an exact binomial test and were found to be negligible.

etc.) is one of the more obvious building characteristics. An analysis of the property types for built form errors reveals that flats cause the majority of the errors. The SAP methodology does not use the built form of flats and maisonettes in its calculations, and it is possible that DEAs therefore do not pay much attention to this variable. It is noted in training documentation that DEAs should select a built form which best represents the flat's character, but this is open to interpretation and the high number of discrepancies may therefore reflect the difficulty in assigning built form for a flat. The built form of a house, meanwhile, does effect the SAP calculations as it describes how many walls are open to the exterior environment. However, semi-detached and end-terrace built forms are treated in the same manner in SAP, as they have an equal number of exposed walls. 45% of the built form discrepancies for houses consist of disagreements on whether a property is semi-detached or end-terrace. These particular discrepancies will not cause differences in the outputs of SAP, but will cause erroneous results if the EPC dataset is used to calculate statistics on the built form of the UK housing stock. The remaining 55% of built form discrepancies for houses are likely to be causing incorrect EERs.

Many EPCs suggest that loft insulation has decreased. Some of these may be genuine, for example if the dwelling has moved loft insulation to make way for storage. However, in many cases these errors likely result from at least one EPC assessor not accurately measuring the loft insulation, or not accessing the loft and instead using an assumed value. Indeed, assessors are only permitted to enter the loft space to measure the insulation levels where it is deemed safe to do so. In either case, previous research has also found that DEAs often record different levels of loft insulation for the same property [15].

3.2. Errors for different property types

An analysis of the error rates for the different property types was calculated to determine if some property types appear to cause more issues than others. This analysis was conducted on the sub-sample containing properties for which 2 EPCs were available. Figure 3 displays the percentages of EPCs lodged for each

property type which are erroneous. Flats and Maisonettes display a significantly higher total percentage than other property types. This is partly due to the effect of error group D, which relates only to flats and maisonettes. Additionally, error group C which relations to discrepancies in property type and built form is more common for these two property types. As mentioned previously, the built form of flats and maisonettes is not used in SAP calculations and this may cause DEAs to give little attention to this variable. There is also some ambiguity around what constitutes a flat or maisonette, and this may also contribute to the increase in error percentages for these types.

Park homes and, to a lesser extent, maisonettes were identified by the random forests (RF) more often than other property types. Many of these errors are likely to be a result of how the RF operates: The RF compares the parameters of one house to its neighbours in the same postcode. If maisonettes and park homes exist in small numbers alongside other property types, then the random forest may be misclassifying some of these park homes as errors. Correcting any misclassified park homes would would likely require a manual inspection.

3.3. Errors for different UK locations

There are 348 local authorities in England and Wales. The error rate in each of these local authorities was calculated using the sub-sample which containing properties with 2 EPCs. The results were then plotted on a map, with the results presented in Figure 4. The highest error rate was found for the local authority of Westminster, London at (81.8±0.5)%. The lowest error percentage was found in Gedling, Nottinghamshire at (39.2±1)%. In general, higher error percentages occur within the London area, as is apparent from the map showing only the London local authorities (figure 5). The higher error percentages found in London may be due to the increased number of flats compared to the rest of the country. As suggested in section 3.2, flats cause a larger number of issues than other property types. Data from the Valuation Office Agency in 2014 [22] suggests that 53% of London's accommodation is comprised of flats,

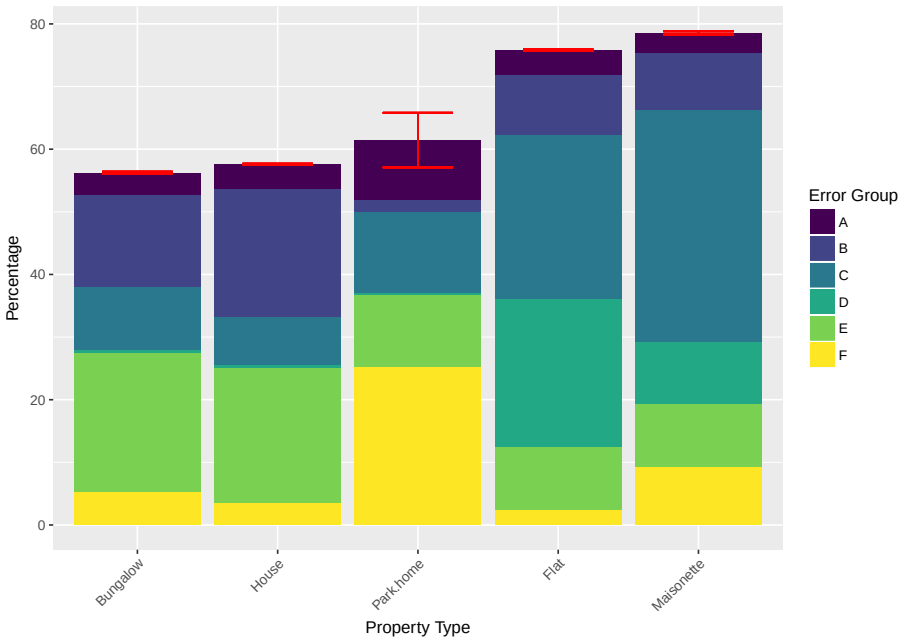

Figure 3: Error percentages for different property types. Uncertainties on these bars were estimated using an exact binomial test. The colour scheme reflects the error group, as described in section 2

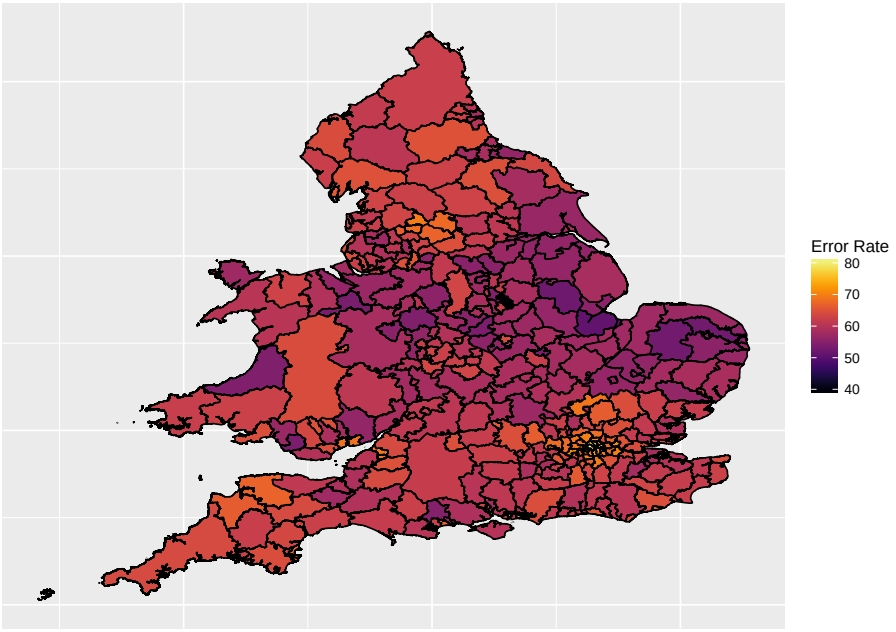

Figure 4: Map of the 348 local authorities in England and Wales with a color scheme describing the error percentage in each local authority.

compared to 22% in England and Wales as a whole. A comparison of the percentages of flats in in each local authority to the Error Percentage further reveals this relationship. However, the relationship between flats and errors does not fully explain the data, as the local authority of Gedling stands out as having a relatively low error percentage despite the presence of flats.

3.4. Errors by Property Floor Area

The error percentage for each property floor area in the sub-sample is displayed in Figure 7, where property areas are combined into bins of width $10\,\mathrm{m}^2$. Error group A is fairly consistent as a function of property area, suggesting that this type of error will happen independently of how big the property is. The remaining error groups show noticeable variation, however, and many of these differences can be explained by the fact that larger properties are more likely to be houses instead of flats. For example, error group B includes errors on

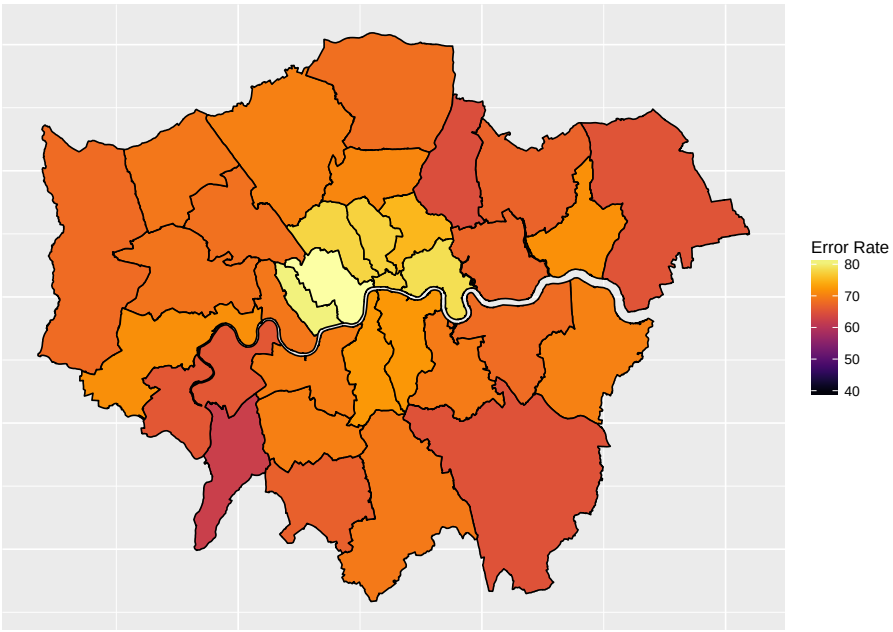

Figure 5: Map of the 33 London local authorities with a color scheme describing the error percentage in each local authority.

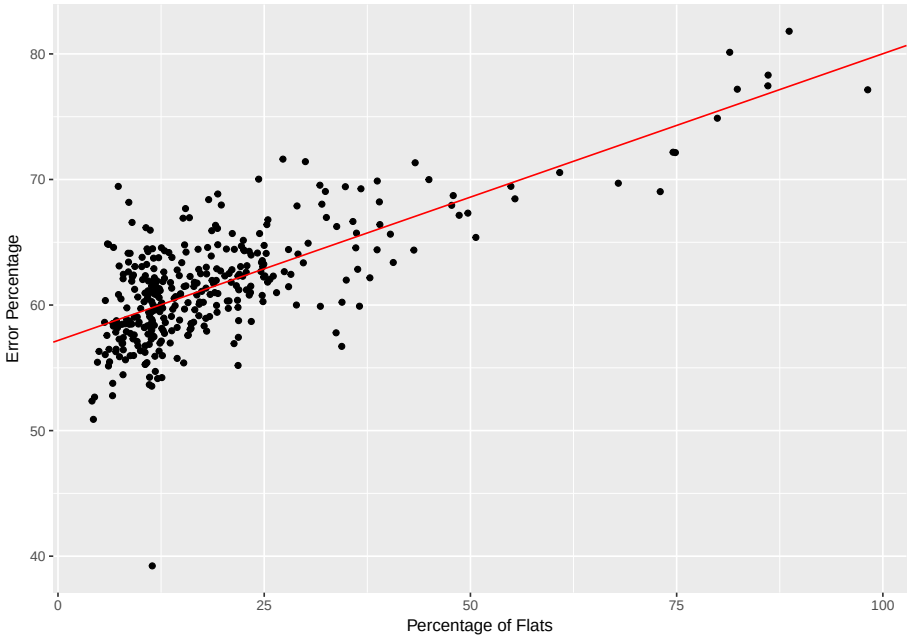

Figure 6: Scatter plot of the percentage of properties which are flats/maisonettes in each local authority with the error percentage in that local authority. The red line showing the result of a linear regression in the data to highlight the trend.

floor type which are more applicable to houses (see Figure 3, thus explaining the increase at larger property sizes. Error group C includes errors of property type which are more common for flats. Error group D relates to errors found in flats only, so it would be expected that these decrease in larger properties. The increase in error group E with property size is dominated by reductions in loft insulation and houses are more likely to have loft insulation. Finally, error group F include errors identified by the RFs. These may become more common with increasing property size as total floor area was a parameter used by the RF to determine errors. A property of $300\,\mathrm{m}^2$ in a street of $70\,\mathrm{m}^2$ houses would likely be flagged as an error.

UK minimum space standards require that a new one bedroom flat has a floor area of at least $37\,\mathrm{m}^2$ [23]. It is therefore likely that all properties in the 0-$10\,\mathrm{m}^2$ bin are in fact incorrect. Such small values for floor area could be caused by a low default value being unchanged during the EPC assessment. The only error identification method used in this work able to flag these small properties as errors is the RFs and, although they did ID some properties in this bin as errors, future work could train the RFs to identify extremely small properties as incorrect. It is also surprising to find properties in the 10-$20\,\mathrm{m}^2$ bin. However only $\sim$6000 EPCs exist in the 10-$20\,\mathrm{m}^2$ bin compared to an average of $\sim$120,000 in all other bins. Some of these floor areas may therefore be genuine and reflective of a low number of very small dwellings.

Given that the minimum space standards were formailised in 2015 [23], it may be expected that new-dwelling EPCs with a floor area less that $37\,\mathrm{m}^2$ would dramatically decrease after 2015. However, the current data does not show this trend occurring. This suggests either that many new dwellings are not adhering to minimum space standards, or that many errors on the size of these properties persist. If the latter case could be confirmed to be correct, then a floor area less than $37\,\mathrm{m}^2$ could be used in future to check to additional errors.

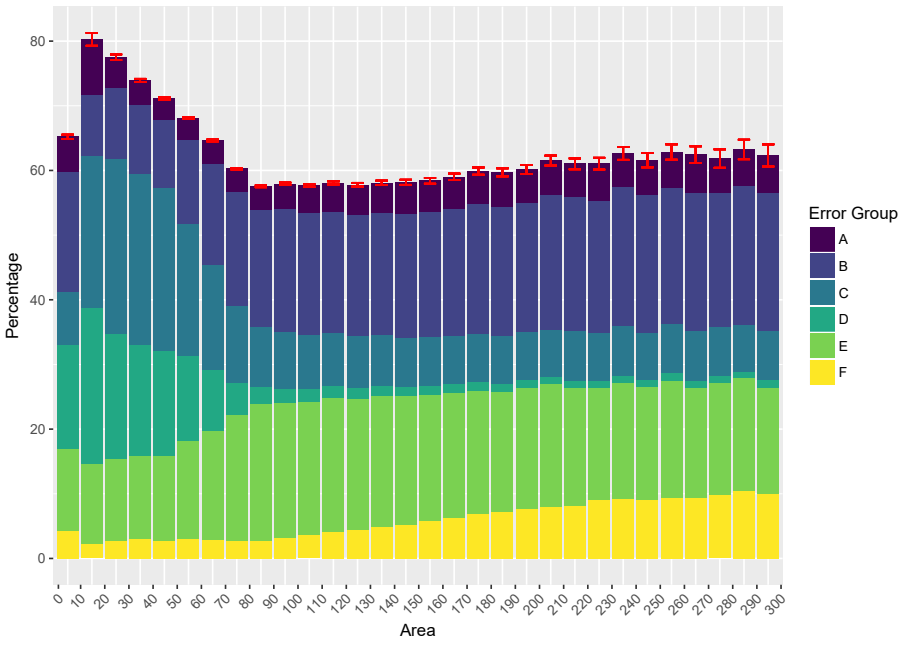

Figure 7: Error percentages for different property floor areas. Uncertainties on these bars were estimated using an exact binomial test. The colour scheme reflects the error group, as described in section 2

3.5. Errors by reason for the EPC

When an EPC is created, one of the parameters recorded is the reason for the EPC being lodged. Having a valid EPC is a requirement when properties are sold or rented, and many EPCs therefore exist for these reasons. In addition, many funding schemes have required EPCs as part of the assessment process. The error percentage for each EPC reason on the sub-sample is displayed in Figure 8, and ordered form lowest to highest.

EPCs lodged following a green deal have the lowest error incidence. The green deal was a government-led scheme designed to finance retrofits and included an EPC assessment both before and after the retrofit measures were installed. These assessments could be completed by the same DEA who conducted the original and, if this is the case, the DEA would not need to re-visit the property. A DEA could instead lodge the new EPC using information from their previous assessment, plus information on what retrofit measures were installed. EPCs lodged under this scenario would be more consistent and this could cause the lower error percentage in this group.

3.6. Effect of potential errors on energy efficiency rating

The EPC process results in am energy efficiency rating (EER) out of 100, from which the property is rated on a scale of A to G (see Table 1). Although the analysis presented in the previous sections suggest that errors are common in the EPC database, we have not yet quantified the effect these errors may be having on EERs. Ideally, this would be achieved by re-creating the EPC and altering the erroneous variable to see its effect. Unfortunately, the public EPC record does not contain all the information required to re-create the original EPC and we instead employ another method.

For errors that are a result of unlikely building changes (error codes 4-16), we extract two EPCs which contain the only the error of interest and take their difference in EER. For example, if we are interested in error code 4 (a disagreement in floor type), we find EPCs for a property whose building characteristics

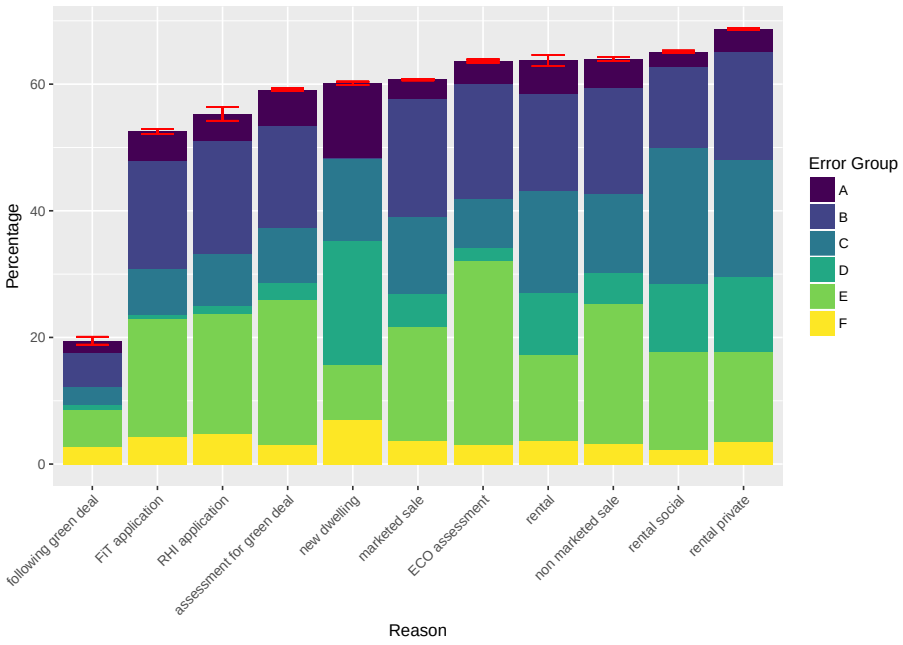

Figure 8: Error percentages for different EPC reasons. Uncertainties on these bars were estimated using an exact binomial test. The colour scheme reflects the error group, as described in section 2

differ only in the stated floor type, and use these to calculate the typical difference in EER. This way, we isolate the effect of the potential error as much as possible. A potential error could either serve to overestimate or underestimate the true EER. Unfortunately, it is not possible to tell from the data alone whether overestimation of underestimation is more prevalent as it is not known what the true values for the building parameters are. We therefore analyse only absolute values of EER difference. Error code 12 (loss of double glazing) was not included in this analysis as only 12 properties meet the criteria required above.

A box plot showing the differences in EER for each error code in displayed in figure 9. Outliers, classed as any values outside 1.5 times the interquartile range above the upper quartile and below the lower quartile, were not shown in this figure for clarity.

Figure 9 suggests that making an individual mistake on a EPC assessment is unlikely to make a considerable difference in the EER. Indeed, error codes 7, 9 and 11 have, on average, no effect on the final EER. Error code 5 (differing wall types) has the most effect of any individual error on EER.

Error codes 2 and 3 are identified by anomalies in the dates of EPCs and not by unlikely change in the building fabric. The typical effect of these errors on EER was therefore calculated separately. Two EPCs from a property which displays the error in question were taken and the the difference in EER between them was calculated. The building parameters in these analyses were allowed to change and, as described in section 2, the reason for these errors may be DEAs correcting mistakes in an earlier assessment. The difference in EER for these error types may therefore indicate a typical change when multiple mistakes are made in the building parameters. A box plot showing the differences in EER for each error code in displayed in figure 10. We do not include error code 1 (identical duplicates) in this analysis as the difference in EER will naturally be zero.

The distributions for error codes 2 and 3 are very similar, with both suggesting a median difference in EER or 4. If an error of 4 were applied to all

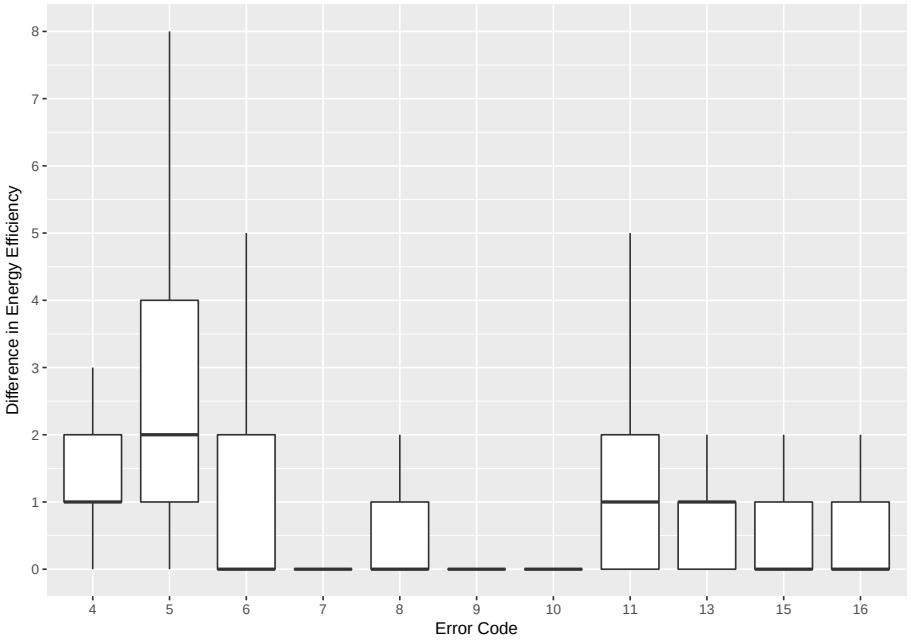

Figure 9: Box plot displaying the effect that each error has the Energy Efficiency Rating. For clarity, this figure does not include outliers. Of the errors studied in this work, discrepancies in floor type have the largest effect on average.

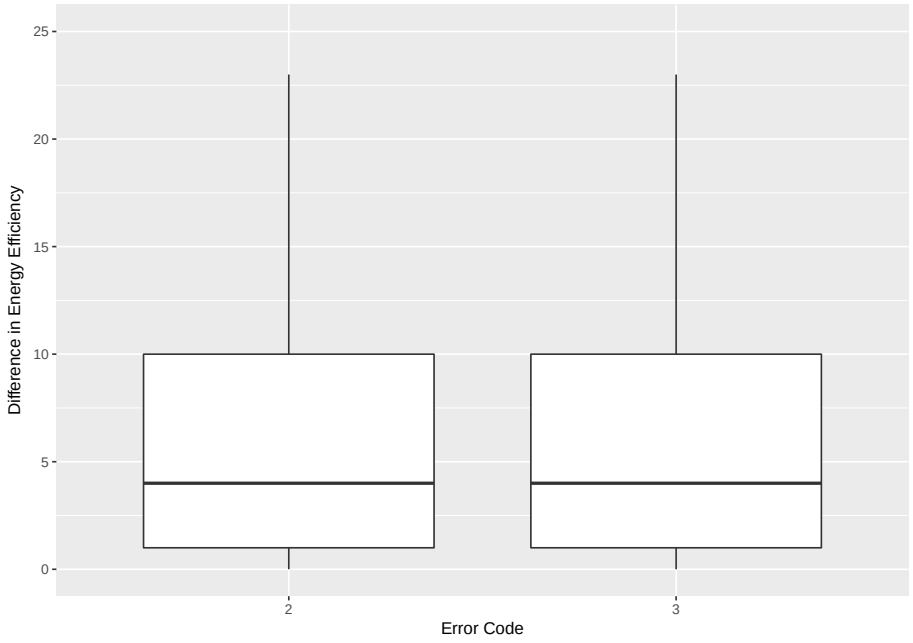

Figure 10: Box plot displaying the difference in Energy Efficiency between two EPCs that show a date error. For clarity, this figure does not include outliers. These errors are likely caused by EPC assessors correcting previous mistakes, and may therefore be reflective of typical typical errors in the EER if multiple mistakes go un-noticed.

EERs in the dataset, the EPC would shift into the wrong band 30% of the time. This could have implications for the EPC price premiums noted in several works [13, 9, 12] as it may mean that up to 30% of property transactions are based on incorrect information. Price premiums are also a noted phenomenon in other European countries [8, 11], and it could therefore be beneficial to conduct a similar analysis of errors within other European EPC systems. However some uncertainty exists around the ability to isolate the effect of price premiums to the EPC alone [24]. Any price premium may in fact be because homes with higher EPC ratings also tend to have traits which are generally more desirable, but which cannot be accounted for in typical hedonic pricing models[25]. Further study of EPC errors may actually be able to help resolve this issue. If a sample of "A" grade houses which have been misclassified as lower grades is extracted, their price may reveal if premiums are the result of the energy efficiency or something more intrinsic.

Unlike other error codes, error code 3 may allow the typical direction of changes in the EER to be understood. For these errors, we take the EER from the most recent EPC and subtract the older EER from this value. The most recent EPC is likely the one that the DEA found to be more acceptable. Any DEAs trying to bump houses into a higher EPC band may try to cause a positive difference in the EER. DEAs trying to have a property qualify for an assistance program may attempt to lower the EPC band, causing a negative difference in the EER. In fact, the distribution of these differences in EER is symmetrical around zero. This suggests that if any DEAs are biasing their assessments in one direction, an equal amount of DEAs are biasing their assessments in the other direction. Alternatively, the DEAs performing these changes are doing so only to correct errors and are not trying to bias the assessment in any way. This would likely also result in the symmetrical distribution witnessed.

4. Conclusions and Policy Implications

At least 27% of all EPCs lodged between 2008 and 2016 have a discrepancy which indicates an error has been made. Many errors were identified via comparison of two EPCs for the same property and if only one EPC is available for a property it is therefore more challenging to find any errors. Correcting for this suggests that the true error percentage for the EPC record is in the range 36-62%. The errors identified cause an approximate difference in energy efficiency rating of 4 points, and an error of 4 points on each EPC in the sample would result in 30% of homes being placed in the wrong EPC band. The volume of errors present in the data suggests much greater care should be taken when using EPC data, whether it is assessing individual homes for retrofits, or calculating bulk statistics from the dataset as a whole. Removing the EPCs identified as potential errors would go some way to increasing the accuracy of any work involving EPCs, however.

Error rates are not consistent across local authorities, or across EPC reasons. The cause of the variation may be explained by differing percentages of flats/maisonettes in these groups, as flats/maisonettes appear to cause more issues than other property types. The difficulties associated with flats/maisonettes may be due to ambiguity around their built form, property type, and the additional complexity associated with determining their location in the larger building. As flats/maisonettes make up 20% of the residential dwellings in England and Wales (and nearly 100% of the residential dwellings in some local authorities), reducing the challenges associated with flats should be a priority for future EPC schemes. More training and strict rules for what is expected for flats/maisonettes could help to remove the ambiguity and around the assessment of flats.

The quality assurance for EPCs assessments require that they be within ± 5 points of the audit value 95% of the time. Since the errors observed in this paper resulted in an average change in the energy efficiency rating of 4 points, the errors are not likely to cause an audit to fail. However, given that many of these

errors are caused by simple-to-assess features such as floor types, wall types and built form, it would be regrettable to ignore these errors. Furthermore, EPCs are delivered alongside suggested energy efficiency improvements and incorrect building characteristics may lead to unsuitable products being recommended for a house.

A campaign of more targeted auditing could be effective at reducing the number of discrepancies found. The method of auditing has previously selected most EPCs for audit randomly, but a targeted procedure could audit those which have been identified as having a discrepancy. There is already a move into "smart auditing" occurring within the EPC quality assurance procedure. Smart auditing will involve selecting EPCs for audit if they meet certain criteria such as having no main heating systems declared, but a gas supply present. The smart auditing criteria currently being considered does not offer comparison with any other EPCs in the record. The work presented in this paper suggests that the criteria for smart auditing could be expanded considerably and include many more rules that would trigger an audit. However, if expanded rules were implemented, the volume of EPCs chosen for smart audit may be unmanageable for the current number of EPC auditors.

An alternative to smart auditing would be to allow DEAs to correct potential errors before the EPC is lodged. For example, if an EPC is lodged that describes a solid wall property but a historic EPC describes a cavity wall, then the software could ask for confirmation that the DEAs wall type is correct. This example would only apply if an EPC already exists for a property, but a procedure similar to the RF used in this work could also be applied to compare to neighbours properties and create a similar alert if a potential error is found. The DEA could, in fact, be removed from this process entirely and machine learning could be relied on to automatically correct any errors found in the EPC record. This paper therefore shows there is potential for machine learning to improve EPC auditing practices. However, if machine learning alone is used then this procedure would be prone to misclassifying any EPCs which are correct, but appear anomalous. In either case, provision would have to be made to

ensure DEAs do not become overly dependent on any automatic checks which are implemented.

This paper has highlighted that care is needed when performing detailed analysis of EPC records. It has also identified that many of the inaccuracies are avoidable and training schemes perhaps should be adapted in light of this. It is recommended that the lodging and auditing system for EPCs becomes more sophisticated to ensure that more errors are spotted before the certificates are formally lodged. These recommendations could substantially improve the EPC record, making it a more reliable data source and improve confidence in its use for both policy makers and households.

5. Appendix

The random forests (RFs) used in this work exploited two patterns expected to be in the data. The first RF exploited the fact that certain building characteristics were not expected to change for a particular dwelling. The second exploited that fact that properties on the same street were expected to have some similar building characteristics.

5.1. RF on Multiple EPCs for a Single Dwelling

The training dataset for this RF was created by first randomly selecting EPCs for properties who possessed at least 3 EPCs from the record. This resulted in a sample of 8324 EPCs. In approximately one third of these EPCs, values for the building parameters were changed to that random EPC for another house. Where this occurred, the entry was marked as an error to be used in the training. Several variables were then created for each EPC to describe how similar it is to EPCs from the same property. For discrete variables such as wall type, the similarity was described as fraction of the marginal frequencies for each variable. For example, if a property possessed 3 EPCs with wall type of solid, solid, and cavity, the similarity variables would be 0.66, 0.66 and 0.33, respectively. The same regular expression rules described in section 2 were used

when assessing text strings. For continuous variables such as total floor area, the similarity was expressed as a percentage of the average.

A table containing the variables on which this RF was trained is is displayed in Table 3. The confusion matrix resulting from this training is displayed in Table 4.

Table 3: Variables included in Random Forest

Incidence of Roof Type
Incidence of Built Form
Incidence of Property Type
Incidence of Floor Type
Incidence of Mains Gas supply
Percentage of Environment Impact relative to mean
Percentage of Energy Efficiency relative to mean
Percentage of Floor area relative to mean
Incidence of Wall Type

Table 4: Confusion Matrix for Random Forest

	False	True	Class Error
False	2196	559	0.20
True	409	5160	0.07

5.2. RF on EPCs from neighbours dwellings

A training dataset for this RF was created by first randomly selecting postcodes which possessed at least 3 EPCs from the EPC record. This resulted in a sample of 9772 EPCs. For each postcode in this sample, around one third of the EPCs were selected and values for the building parameters were changed to that of EPCs from other postcodes. Where this occurred, the entry was marked as an error to be used for training. Variables were then created for each EPC to describe how similar it is to EPCs from the same postcode. Discrete and continuous variables were handled in the same manner as for the previous RF.

A table containing the variables on which this RF was trained is is displayed in Table 3. The confusion matrix resulting from this training is displayed in Table 4.

Table 5: Variables included in Random Forest

Incidence of Roof Type
Incidence of Built Form
Incidence of Property Type
Incidence of Floor Type
Percentage of Floor area relative to mean
Incidence of Wall Type
Number of houses on street

Table 6: Confusion Matrix for Random Forest

	False	True	Class Error
False	1693	1051	0.38
True	525	6503	0.07

6. Acknowledgements

This work has been supported by the Consumer Data Research Centre, an ESRC Data Investment, under project ID CDRC 88, ES/L011840/1; ES/L011891/1

References

[1] EPBD, Directive 2010/31/eu of the european parliament and of the council of 19 may 2010 on the energy performance of buildings (recast), Official Journal of the European Union 18 (06) (2010) 2010.

[2] United kingdom statutory instrument no. 660. the energy efficiency (private rented property) (england and wales) (amendment) regulations 2016 (2016).

[3] The electricity and gas (energy company obligation) order 2014 (2014).

[4] The renewable heat incentive scheme and domestic renewable heat incentive scheme (amendment) regulations 2018 (2018).

[5] K. G. Droutsa, S. Kontoyiannidis, E. G. Dascalaki, C. A. Balaras, Mapping the energy performance of hellenic residential buildings from epc (energy performance certificate) data, Energy 98 (2016) 284–295.

[6] P. Florio, O. Teissier, Estimation of the energy performance certificate of a housing stock characterised via qualitative variables through a typology-based approach model: a fuel poverty evaluation tool, Energy and Buildings 89 (2015) 39–48.

[7] BPIE, Energy performance certificates across europe: From design to implementation (2010).

[8] D. Brounen, N. Kok, On the economics of energy labels in the housing market, Journal of Environmental Economics and Management 62 (2) (2011) 166–179.

[9] F. Fuerst, P. McAllister, A. Nanda, P. Wyatt, Does energy efficiency matter to home-buyers? an investigation of epc ratings and transaction prices in england, Energy Economics 48 (2015) 145–156.

[10] B. Bloom, M. Nobe, M. Nobe, Valuing green home designs: A study of energy star® homes, Journal of Sustainable Real Estate 3 (1) (2011) 109–126.

[11] M. Hyland, R. C. Lyons, S. Lyons, The value of domestic building energy efficiencyevidence from ireland, Energy Economics 40 (2013) 943–952.

[12] F. Fuerst, P. McAllister, A. Nanda, P. Wyatt, Energy performance ratings and house prices in wales: An empirical study, Energy Policy 92 (2016) 20–33.

[13] F. Fuerst, P. McAllister, A. Nanda, P. Wyatt, An investigation of the effect of epc ratings on house prices.

[14] HM Government, Future administration of the energy performance buildings directive quality assurance (epbd qa) regime (2011).

[15] DECC, Green deal assessment mystery shopping research (2014).

[16] B. Hårsman, Z. Daghbashyan, P. Chaudhary, On the quality and impact of residential energy performance certificates, Energy and Buildings 133 (2016) 711–723.

[17] M. Collins, J. Curtis, Bunching of residential building energy performance certificates at threshold values, Applied Energy 211 (2018) 662–676.

[18] H. Amecke, The impact of energy performance certificates: A survey of german home owners, Energy Policy 46 (2012) 4–14.

[19] DCLG, Housing supply; net additional dwellings, england: 2016-17 (2017).

[20] HM Government, The building reguations 2010 (2010).

[21] BRE, Rdsap 2012 version 9.93 (2017).
URL https://www.bre.co.uk/filelibrary/SAP/2012/RdSAP-9.93/RdSAP_2012_9.93.pdf

[22] Valuation Office Agency, Dwellings by property build period and type, lsoa and msoa (2015).
URL https://files.datapress.com/london/dataset/property-build-period-lsoa/2015-12-30T15:34:13/property-type-2014-lsoa.csv

[23] DCLG, Technical housing standards nationally described space standard (2015).

[24] L. Murphy, The influence of the energy performance certificate: The dutch case, Energy Policy 67 (2014) 664–672.

[25] J. O. Olaussen, A. Oust, J. T. Solstad, Energy performance certificates–informing the informed or the indifferent?, Energy Policy 111 (2017) 246–254.