SN Bus Econ (2025) 5:36
https://doi.org/10.1007/s43546-025-00807-w

ORIGINAL ARTICLE

Estimating future thresholds for the 15% eligibility criteria of the EU taxonomy with limited data availability

Christian Buschbeck[1] · **Jens Teubler**[1]

Received: 18 June 2024 / Accepted: 13 March 2025 / Published online: 25 March 2025

Abstract

The EU Taxonomy Regulation requires, for the acquisition and ownership of buildings, to demonstrate that the asset's primary energy demand (of buildings constructed before 2021) is within the top 15% of the national or regional building stock. Determining the top 15% energy performance of a building stock is challenging because data availability is generally poor. Furthermore, the threshold for this top range will shift with upcoming refurbishment efforts and higher energy efficiency standards. We tackle these issues by proposing a methodology for estimating 1) current thresholds based on more widely available data on energy performance certificates and 2) using existing scenarios to estimate future threshold values. Estimation of current thresholds for residential buildings yields a moderate fit and a threshold value for final energy demand of 74 kWh per square meter and year (or a conservative threshold of 70 kWh for primary energy demand), which is very close to the results reported by other scholars. Estimated future thresholds show a linear decline in final energy demand down to 20 to 45 kWh per square meter and year in 2045, depending on the applied scenario.

Keywords EU taxonomy · Eligibility · 15%-criteria · Primary Energy demand · Building Stock

Acronyms

EPC	Energy performance certificates
EUT	EU Taxonomy regulation on sustainable activities
FE	Final energy
FED	Final energy demand
GHG	Green house gas

✉ Christian Buschbeck
christian.buschbeck@wupperinst.org

Jens Teubler
jens.teubler@wupperinst.org

[1] Research Division Sustainable Production and Consumption, Wuppertal Institute for Climate, Environment and Energy, Döppersberg 3, 42103 Wuppertal, Germany

NZEB Nearly zero energy building
PED Primary energy demand
PEF Primary energy factor
TSC Technical screening criteria

Introduction

The EU Taxonomy Regulation on Sustainable Activities (EUT) is part of the European Commission's action plan on financing sustainable growth (European Parliament and of the Council 2020). It requires the reporting of eligible economic activities by financial intermediaries, as well as a substantiation of whether such activities are taxonomy-aligned. This alignment is assessed with the help of so-called Technical Screening Criteria (TSC). For the activity 'Acquisition and Ownership of Buildings' different criteria can apply depending on the year of construction. For buildings built before 31 December 2020, it must be demonstrated that taxonomy-aligned buildings are within the top 15% of the national or regional building stock regarding their 'operational' primary energy demand (PED).[1]

Institutions that must comply with this regulation therefore need information on the PED of the buildings in their portfolio, but also reliable thresholds to determine the top 15% of PED in the building stock of the respective countries at the time of reporting (for both residential and non-residential building types). So far, such thresholds have not been provided by European or national entities such as the EU building observatory. Moreover, many institutions do not have data on the specific PED of all of their financed buildings, but rather rely on other information such as the nationally regulated energy performance certificates (EPC) of buildings.

The uncertainty of EPC data, but also their frequent use for data analysis in a wide array of research questions is well-discussed in the literature. A meta-study from 2019, for example, found that EPC data now have a wide range of applications that were probably not originally envisioned by policy makers (Pasichnyi et al. 2019). Frequent and often overlapping uses for this type of data are (ibid.) for example the mapping of energy performance in a given context (e.g., for urban planning by Gupta and Gregg (2018)), predicting future energy use (e.g. with the help of machine learning by Paterson et al. (2017)) or addressing the so-called performance gap (e.g. accounting for differences between actual and calculated energy consumption by Balaras et al. (2016)). However, even with an increased availability of EPC data, and even with additional auxiliary variables (or so-called EPC indicators), the performance gap (Shi et al. 2019), the dependency of EPC values on the specific assessor (Jenkins et al. 2017; Hardy and Glew 2019) and the low cross-national harmonization of EPCs pose a serious problem for any type of energy performance

[1] The 15% criterion discussed in the study at hand should not be confused with the 15% proposal by the European Commission in COM/2021/802 (e.g. investigated by Ferrantelli and Kurnitski (2022)). The criterion here refers to the best-performing 15% (primary energy demand) in the building stock, whereas the latter refers to a future threshold for the worst-performing buildings (future category G in EPCs).

mapping, energy or green house gas (GHG) emission predictions, and bench-marking (see (Sesana et al. 2024) for a discussion of the differences between European countries in this regard). The latter seems particular problematic in the context of EUT reporting obligations, since the respective TSC for climate change mitigation allows e.g. real estate companies to adhere either to EPC class A buildings in their portfolio (resulting in energy consumption thresholds that vary strongly between countries) or to the top 15% PED criterion (relying on different methods to derive this value across different actors but usually in reference to EPC classes or values). Moreover, different approaches are chosen in the European context to calculate EPC values in the first place, which further impedes uniform modelling of building stocks (Ferrantelli and Kurnitski 2022).

The introduction of the 15%-criterion into the EU taxonomy (EUT) therefore poses a three-fold problem for benchmarking. Firstly, the EUT-TSC highlight the importance of EPC classifications for taxonomy-alignment but also allow to demonstrate with 'adequate evidence' that a building in question is in the top 15% of the'operational' PED of a regional building stock (the TSC can be found in section 7.7 in Annex 1 of the EUT). Secondly, most companies obliged to report under the EUT only have access to the EPC certificates of their buildings and thus replicate the known performance gap. And thirdly, these EPC classes (or their underlying values) are time-sensitive to a 15% threshold that will likely change over time[2] as well as time-sensitive to the EPC approach selected, regulations in place when they were reported, and the changes in values due to renewals of EPCs(e.g. as demonstrated by Yuan and Choudhary (2023) for buildings in the UK).

This article is based on a previous research project that estimated the 15% thresholds for residential and non-residential buildings in Germany as well as non-residential buildings in France, Spain and the Netherlands. It addresses the challenge of mapping EPC data and classes by EUT reporting companies to the 15% criterion and thus to provide a pragmatic solution to select 'potential taxonomy-aligned'[3] buildings from a given data set for fulfilling a reporting obligation. We close this information gap by providing a methodology for quantifying theses thresholds as well as positioning these thresholds in the EPCs of European Countries in general, and more specifically for both current and future building stocks in Germany.

In theory, determining the required country-specific thresholds is merely a question of data sampling. If a sufficient number and sufficient range of buildings and their respective PEDs is known for each country, the 15th percentile could be calculated and compared to ranges in EPCs. However, such building data sets are often sparse, outdated, and inconsistent between countries (Ali et al. 2020). Such data is also especially difficult to obtain for non-residential buildings. Moreover, national EPCs impede this task not only because they are currently non-comprehensive

[2] We argue that this is also the case under a more flexible interpretation of the TSC that merely requires to demonstrate a threshold which 'at least' represents the stock of buildings built before 2021, since even under this view buildings would phase out of the stock or be renovated.

[3] We use the term 'potential' to indicate that more than one criterion has to be met (e.g. such as achieving a maximum water-use) for taxonomy-alignment.

among countries and often a non-reliable sources of information, but also because the ranged approach of EPCs means that the top 15% of buildings could occupy more than one class at once. Looking at both the literature and the development of EPC-related regulations, it has also become clear, that the existing performance gap as well as the inconsistency of EPC data currently do and will continue to indirectly also affect the green asset ratio of companies with no fault of their own (as indicated by a small sample of interviews with real estate companies by Backenroth and Lindquist (2021)). The identified, and addressed, research gap in the study at hand thus focuses on determining the 15% threshold for such entities, but a cross-national application of our methodology would likely replicate some of these issues.

First attempts have been made to determine thresholds for the 15%-criterion using both classification schemes and absolute values (Tschätsch 2023; Jakob et al. 2022). However, this threshold will shift with increasing refurbishment activities due to climate change mitigation measures such as new building standards or renovation efforts. This leads to uncertainty for stakeholders and financial institutions regarding planning security for their assets and the operationalization of their reporting obligations. We aim to address this by proposing a methodology for estimating future thresholds of the 15% criterion. It is based on parametric statistics and optimization algorithms and requires a manageable amount of input data.

The goal of the study is to (i) show how such specific numerical thresholds can be determined for current as well as future buildings, (ii) how such data can be aligned with national EPC classes and values in a pragmatic manner, and (iii) to provide a conservative threshold that can be directly used by institutions obliged to report under this part of the EU taxonomy regulation.

The scope for this study are current and future residential building stocks in Germany.[4]

Data and methods

Terminology for energy metrics

The results derived in this study refer to the final energy demand (FED) of buildings that are above or below a particular EPC treshold. It is sometimes unclear though whether the 'performance' in EPC metrics refers to energy consumption, energy need, energy use or energy demand. In fact, this distinction is often not made at all. We therefore consider the definition in the current EU directive on the energy performance of buildings (Art.2(8)) more relevant in this context than the physical definition of these metrics (European-Parliament 2024):

[4] The original research project derived thresholds for other building types and other countries. However, we found that sticking to one example sharpens the focus on the methodology and decided to exclude these cases to evade confusion.

> '[E]nergy performance of a building' means the calculated or metered amount of energy needed to meet the energy demand associated with a typical use of the building, which includes energy used for heating, cooling, ventilation, domestic hot water and lighting.

Energy **demand** (whether final or primary) can therefore refer to the 'metered' or 'calculated' energy consumption of a building which is, in this regulatory definition, equated with energy need or energy demand. A more accurate terminology then further differentiates between the energy performance of a building based on its 'operational rating' (metered energy consumption depending on its residents) or 'asset rating' (calculated energy consumption depending on building type characteristics) (BPIE 2014).

Some countries only use one of these two options to derive the energy performance of buildings and attribute an EPC, but some, such as Germany, allow for both options. The regulatory EPC threshold values for Germany are the same, though, regardless of the type of data used to identify the energy performance class of a building. This has consequences for EPC data and EPC data interpretations, as shown in the following section.

Data

For this study, two datasets are used. The first is published by co2online GmbH and contains a relative frequency histogram (not actual values) showing FED of German residential buildings with a bar width of 10kWh/($m^{2*}a$). The data was originally compiled from 1 949 348 data points (co2online 2022) and before publishing, the dataset was cleaned and processed to properly represent the German building stock (Metzger et al. 2019). This dataset is used to find the right distribution for energy consumption in German residential buildings and to test our methodology against.

The second dataset is published on Statista and contains a relative frequency histogram (not actual values) for EPC classes among German residential buildings. It was compiled by McMakler, a real estate company that issues EPCs, and represents 1 681 buildings with the construction year ranging from 1920 to 2020 (Statista 2021). There is, though, no information regarding the type of data used to determine the EPC classes. It is thus assumed, in accordance with the situation in Germany, that both asset and operation ratings constitute the basis of the EPC frequencies in the sample. Because the actual calculated and metered energy values are not known, it follows, that the share of each sub-set in the entire sample influenced the results to an unknown degree. It is thus very likely that the stated relative frequency would look different if only one method was used to determine the EPC of each building. However, since 'operational ratings' are prone to user behaviour, and 'asset ratings' to the difference between generalized and specific building attributes,[5] this

[5] Both also rely on the spatial and temporal weather conditions, with the 'operational rating' being prone to overestimating outliers in this regard and the 'asset rating' underestimating changes in climate.

Table 1 German EPC classification scheme and corresponding FE ranges (energy demand and energy consumption possible)

Class	FE $[kwh/(m^2 * a)]$
A +	FE ≤ 30
A	30 < FE ≤ 50
B	50 < FE ≤ 75
C	75 < FE ≤ 100
D	100 < FE ≤ 130
E	130 < FE ≤ 160
F	160 < FE≤ 200
G	200 < FE≤ 250
H	250 < FE

Table 2 Shows AIC and BIC values for the tested distributions (rounded to two digits, the value of both measures is the same). All values are shown in million

	Lognormal	Gamma	Exponential	Weibull
AIC and BIC	1.14	1.13	1.21	1.12

fact alone does not allow a prediction about the direction of this shift, even if these shares were known.

The German EPC classification scheme itself consists of 9 classes (from A+ to H), each covering a different range of final energy (FE) (Table 1) (Bundesamt für Justiz 2020).

The second dataset combined with the classification scheme serves as imitation of a situation with poor data availability and is used for calculating the 15%-threshold. The density function fitted in the process is subsequently used to estimate future thresholds.

Finding the right distribution

The methodology described in this work relies on parametric statistics and optimization based parameter estimation. More precisely, a distribution which is believed to describe the data best is fitted to observed values to determine the 15% quantile. In a next step, this distribution can then be updated to reflect the distribution of future energy demand. So the first thing to do was to analyze the structure of existing building stock and decide which distribution is best suited to represent these data. Because energy demands cannot be negative, but every positive value is possible, a positive continuous distribution is needed. Hence, several positive continuous probability distributions were fitted to the data (using the R package 'fitdistrplus') and their fit was compared (Figs. 1 and 2, Table 2). For building stocks in Estonia and Korea a similar approach showed best fits for the lognormal distribution (Ferrantelli

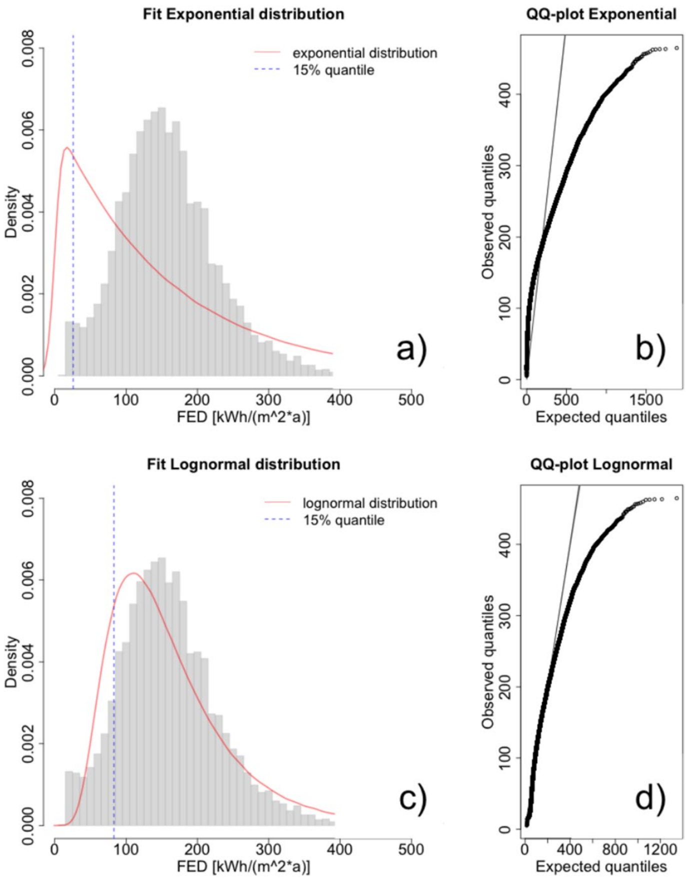

Fig. 1 Histograms showing the distribution of FED values from co2online with fitted density functions and the 15% criterion threshold value for Exponential (**a**) and Lognormal (**c**) distribution with respective QQ-plots for analyzing the quality of the fit (**b** and **d**)

et al. 2022; Park et al. 2016). A common measure to quantify how good a model fits observed data is the log-likelihood. However, log-likelihood does not take model complexity (i.e. the number of parameters) into account, which can lead to fitting a model that performs very good on observed data but poorly estimates unknown data (overfitting). To address this issue, measures that penalize model complexity like the Akaike information criterion (AIC) or the Bayesian information criterion (BIC) were introduced. Unlike in log-likelihood maximization, values of AIC and BIC indicate a better fit, the smaller they are. Because the analyzed distributions differ in their number of parameters, both AIC and BIC where calculated. Additionally, QQ-plots were created to analyze the quality of the fit. In QQ-plots, quantiles of the observed data are plotted against quantiles of the fitted distribution. For an ideal fit, the values would lay on a straight line. It was found that the weibull distribution (Rinne 2008) yields lowest AIC and BIC values (Table 2) and performs best

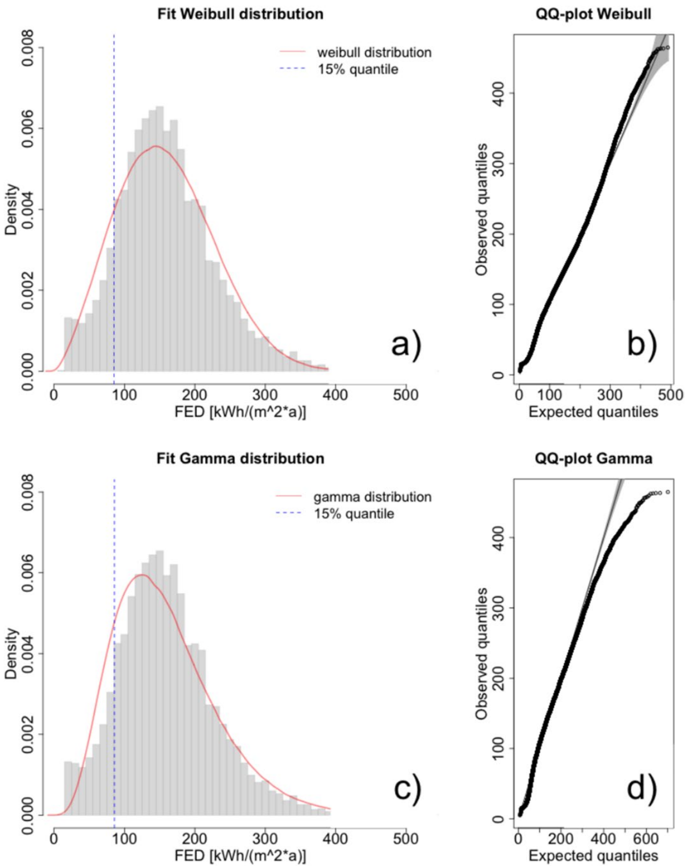

Fig. 2 Histograms showing the distribution of FED values from co2online with fitted density functions and the 15% criterion threshold value for Weibull (**a**) and Gamma (**c**) distribution with respective QQ-plots for analyzing the quality of the fit (**b** and **d**)

in the QQ-plots (Fig. 2). Its shape (right-skewed) seems to reflect the structure of the building stock quite well. Especially for future distributions, this is an important feature, as high energy consumption values will decline and lower values will increase due to refurbishment (BMWi 2020). Based on these insights, we regarded the weibull distribution as a suitable candidate for estimating the thresholds of the 15% criterion.

Estimating thresholds with limited data availability

In general, the availability of monitored data regarding the distribution of actual PED or FED of national building stocks is poor (Do and Kristen 2018). Data regarding the distribution of building energy classes is given more frequently and for larger samples (Li et al. 2019), but suffers from the so-called 'energy performance gap'. The latter usually refers to the difference between energy performance estimations relying on ’static’ characteristics such as regional climate and building types in EPCs and the actual energy use of buildings further determined by additional parameters such as actual weather or user behaviour (Shi et al. 2019). The EPC classes and their underlying values themselves can also differ for the same building as shown for EPC results in the UK by e.g. Jenkins et al. (2017) and Hardy and Glew (2019). Moreover, these EPC results could even differ based on the country and its framework for determining the EPC relevant energy metric (as using metered energy consumption instead of calculated energy need is likely to lead to different results for the same building in similar climate conditions).

But even if EPC data was fully reliable, the information on how many buildings fall within a certain class would not be sufficient to estimate threshold values because the distribution of energy demands within an efficiency class is not known. For example, if both the best- and second-best class each comprise 10% of the observed values, it is not known where exactly the threshold for the 15% criterion lies. This problem gets more severe the larger the range of an EPC is defined. Another information which is often publicly available is the overall average for FED, PED or similar statistical metrics. The methodology we are proposing builds upon optimization based parameter optimization and the method of least squares. We aim to fit a distribution, whose properties allow reflecting energy demands (here, weibull), to both the share of values within each class and another given statistical metric (e.g. the overall average). For estimating the 15 % criterion, the fit to cumulative shares of each class is more meaningful than the fit to each individual class, as quantiles can also be regarded as cumulative shares. Therefore, the fitting is done by minimizing the sum of squares of differences between observed and modelled values, with the cumulative shares of EPCs and the statistical metric being the independent variables. Because two different units are compared (e.g. percentage points and $kWh/(m^2 * a)$) the difference between modelled and source data is calculated in percent. Both the fit to the shares of energy classes and the fit to the statistical metric are regarded equally important. Therefore, they should weigh the same. This is achieved by taking the difference between observed and modelled values for the statistical metric into account once for each EPC. The optimization is performed in R using the ’optim’ function and ’SANN’ algorithm.

Selecting thresholds from distributions

The energy benchmarks used for this study refer to the final energy demand (FED) of buildings based on their EPC class rather than the required primary energy demand. Whereas the first metric relates to either calculated or metered energy consumption in buildings (at least in regard to the EPC scheme in Germany), the latter also takes the regional energy provision and conversion into account (e.g. from power plants combusting coal).

We argue that deriving the 15% threshold should be based on the best-available data (in our case FED values from a EPC frequencies fitted to actual energy consumption distributions), but that the goal of the study is to match these thresholds with the metrics available to reporting agencies (which is usually based on EPC classes and can either or both be expressed in FED or PED values). This requires to translate the derived FED values into both PED values and EPC classes, with the latter representing the less accurate criterion.

The PED is equal to the FED of a building multiplied with the primary energy factor (PEF). These PEFs are usually defined to have a value of 1 or higher for any type of fossil fuels involved, since there are always losses from energy provision, conversion and transport. Although some sub-metrics of PED distinguish between energy from fossil fuels and energy from renewables (with corresponding PEF_{fossil} then being lower than 1), one can therefore safely assume that achieving a FED threshold also implies achieving at least a corresponding PED threshold. This is why we recommend, and utilize, the following prioritisation procedure for selecting data from a reporting entity's portfolio:

1. If the FED of building is known, this value should be lower than the FED 15% threshold including a small error margin (e.g. by rounding down the threshold value).
2. If the FED of a building is not known, but the PED is known, this value should be lower than the calculated FED 15% threshold rounded down to the nearest multiple of 5.
3. If neither the FED nor PED is known, but the EPC class is available, only building classes are eligible for selection that are fully entailed by the 15% threshold.

This prioritization thus ensures that reporting entities err on the side of caution by only including assets that are extremely likely to adhere to the criteria set out in the EU taxonomy. Or in other words: the actual green asset ratio might be higher, but the reporting entity lacks the evidence to support this claim. It also ensures that at least a portion of the uncertainty in the data itself (here relative frequencies of EPC classes in a country that allows for both 'operational ratings' and 'asset ratings') will likely be mitigated in a conservative manner as well.

Estimating future thresholds

For future distributions, estimating threshold values is challenging because it is unknown how the structure of the building stock will evolve (Dascalaki et al. 2016; IWU 2016). In contrast, there are **scenarios** for the future evolution of the **average** FED for e.g. Germany (Bürger et al. 2016; Glaichen et al. 2021; Mellwig et al. 2018). The method introduced in this study is based on this fact and inspired by techniques like the method of moments, where statistical metrics (moments) are used to derive distributions that exhibit specific properties (Ashkar et al. 1988). This idea is taken up as we aim to optimize the parameters of the current distribution, so that its properties reflect future conditions as good as possible.

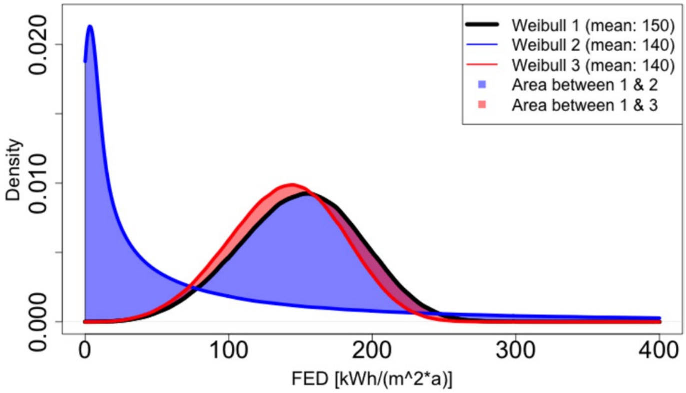

Fig. 3 Shows the probability density functions of three Weibull distributions. The first (black) reflects current, the second and third (blue and red) reflect future probability density functions. The area between current and future probability density functions is colored according to the respective future distribution

Desired properties of future distributions

The distribution of future FEDs can only be a plausible reflection of the future building stock if it exhibits certain properties. These properties are defined in the following.

First, it is very likely that many buildings will be refurbished to achieve lower energy consumption and GHG emissions. As a consequence, the average value for FED will probably decline. To reflect this, the distributions' mean value should equal estimated future mean values derived from scenario analysis.

Second, the future building stock evolves from the current one. Even if refurbishment is performed at high rates, many buildings that are used today will still be in use for several decades. Some of them will show an increased energy efficiency from renovation, whereas others will not. So evidently, the building stock is a system that only changes gradually. To reflect this, the distribution for future FEDs should resemble the current distribution as much as possible. For quantifying this 'similarity' between distributions, we propose to calculate the area between probability density functions of future and current distributions. Figure 3 exemplarily shows the relationship between the shape of probability density functions and the area between them.

The desired properties for future distributions can be summarized to:

a) The mean value of the future distribution should equal the estimated future mean value derived from scenarios.
b) The area between the current and future probability density functions should be minimized.

This yields an optimization problem with one constraint (a) and one objective (b). The number of variables of the optimization problem depends on the parameters of the distribution. The weibull distribution is defined by two parameters (shape and scale). The optimization can be simplified due to another feature of the weibull distribution: The mean value μ depends directly on the parameters shape (λ) and scale (k).

$$(\mu) = \lambda\Gamma\left(1 + \frac{1}{k}\right) \text{ with } \Gamma \text{ being the gamma function}$$

Due to this relationship, only one of the parameters needs to be altered during the optimization process, and the other can be determined using the first parameter and the targeted mean value. This leads to a simplification of the optimization problem (1 instead of 2 parameters) and ensures that the resulting distribution has the wanted mean value. The optimization is performed in R using the 'optim' function and 'BRENT' algorithm.

Scenarios for future mean values

The future mean FED is estimated based on scenarios by the German Federal Environment Agency (UBA) (Bürger et al. 2016) and by Agora Energiewende (Glaichen et al. 2021; Mellwig et al. 2018). The scenarios from UBA, referred to as Target 40, Target 55 and Target 70, represent different pathways to mitigate impacts in the building sector and translate into a reduction in FED of 40, 55 and 65% respectively (compared to consumption levels in 2008). These scenarios differ in the way in which the overarching goal, a reduction of 80% primary energy demand until 2050, is to be achieved: In Target 70 the focus lies on refurbishment while efforts in Target 40 concentrate more on changing the energy source. Accordingly, Target 55 is moderate in both aspects. Agora Energiewende conducted a study which estimates a reduction of 44% FED until 2050, compared to consumption levels of 2011 (Mellwig et al. 2018). After the national goal of climate neutrality for Germany was set to 2045 instead of 2050, the scenario was refined to a reduction of 32% FED until 2050, compared to consumption levels of 2018.

Because no further information for these decarbonization pathways are given, a linear interpolation between current and future consumption levels was applied to fill the gaps. Taking these five scenarios into account, it can be shown, that the UBA scenarios Target 40 and Target 70 reflect extreme cases and can serve as guardrails that mark out the possibility space (Fig. 4). Therefore, these two are used for further analysis.

Results

Estimating current threshold values for the 15%-criterion based on EPCs

The proposed procedure was applied to EPC class data and tested against the more granulate consumption data. The fit can be evaluated based on the shares of each EPC class as well as their cumulative shares (Fig. 5). Except for the highest class (H), where a huge difference can be seen, the share of each EPC deviates moderately

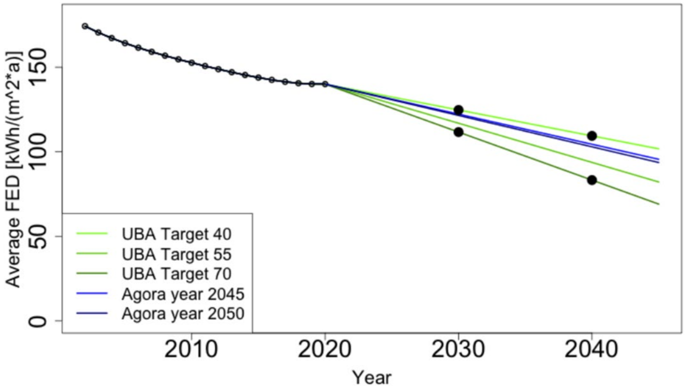

Fig. 4 Shows the trend in average FED according to five scenarios from the German Federal Environment Agency and Agora Energiewende. The trend was calculated with a linear interpolation

between fitted and observed values. The fitted average is very close to the observed value. In general, the fit appears to be better for lower classes (A+ to D) than for higher ones (E to H). This is also visible in cumulative shares where the two curves are very close together for lower EPCs but deviate for higher ones (Fig. 5b). This might stem from the percentage deviation being used in the optimization and in consequence a disproportional fitting on classes with lower shares. For the estimation of threshold values, the poor fit of higher EPCs is not a problem because only lower ones are of concern for estimating the 15% criterion. But for the estimation of future threshold values, the shape of the whole distribution influences the result. Therefore, this shortcoming will propagate, which is also discussed below. Interestingly, when compared to the FED values from co2online (2022), it is vice versa: The distribution seems to better fit with higher than lower FED values (Fig. 6). This discrepancy can only be explained by the different data sources.

The obtained threshold value is 74 $kWh/(m^2 * a)$ for the FED of residential buildings, which is higher than the 70 $kWh/(m^2 * a)$ reported in Tschätsch (2023). Yet, it has to be noted that the value reported by Tschätsch (2023) was derived by taking data provided by co2online (2022), for which the cumulative share indicates that the threshold must lay between 70 and 80 $kWh/(m^2 * a)$, but the threshold was set to the lower boundary as a conservative assumption. This is a pragmatic and feasible approach and reveals that the two results are actually very close. Despite the stochastic nature of the optimization algorithm, this result can be regarded as robust, because the standard deviation for this value is 0.7 for 100 optimization runs.

Estimating future threshold values for the 15%-criterion

For the two most extreme scenarios, the original weibull distribution was updated with future average FED values for the years 2030 and 2040, and the thresholds for the 15%-criterion were calculated. The results are shown in Fig. 7 and Table 3. While probability density functions usually shift to the left when average values are

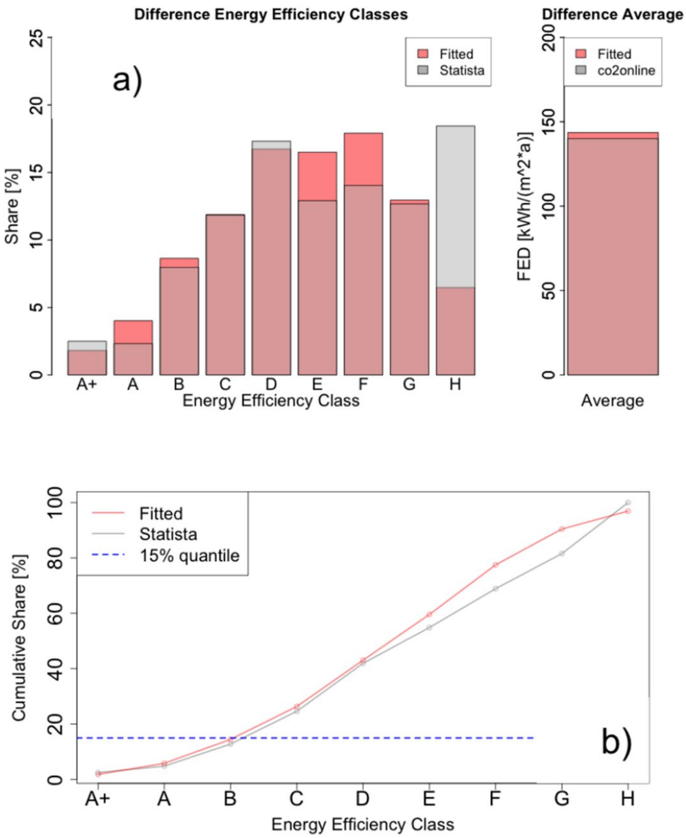

Fig. 5 Evaluation of the fitted distribution based on EPC classes in German residential building stock. a) Difference between fitted shares of EPCs and reported values from Statista. The difference between fitted and reported average is given as well. b) Comparison of fitted and reported cumulative shares of EPCs

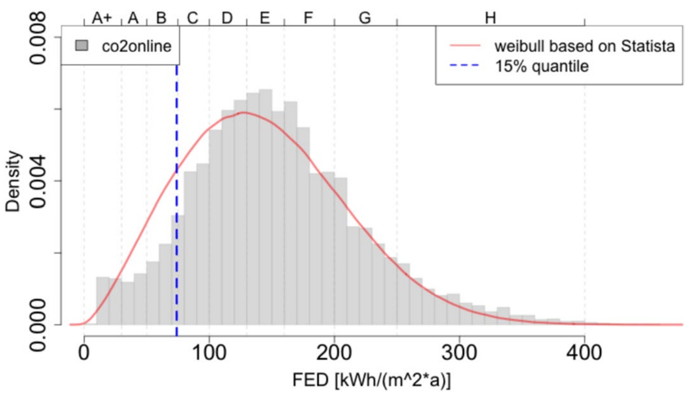

Fig. 6 Evaluation of the fitted distribution based on EPCs (red) and more detailed data provided by co2online (grey)

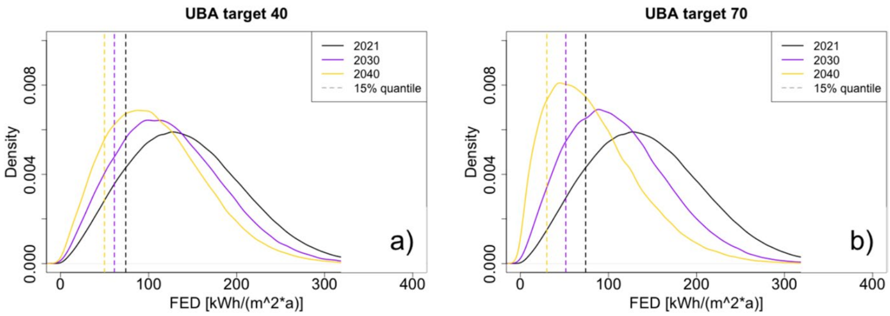

Fig. 7 Shift of the fitted Weibull distributions with future average FED values. The black line represents the current fit, the colored lines represent future fits. The dotted lines show the respective threshold for the 15% criterion

Table 3 Calculated threshold values and conservative estimates for the 15%-criterion in the years 2030 and 2040 for the two UBA scenarios (in $kWh/(m^2 * a)$)

	Target 40		Target 70	
	Calculated	Conservative Estimate	Calculated	Con-servative Estimate
2021	74	70	74	70
2030	63	60	52	50
2040	55	55	31	30

reduced, a continuous positive distribution (as the weibull distribution) cannot take on negative values. As a consequence, decreasing the average value while staying as close to the original density function as possible, leads to blowing up the left part (lower values) and lowering the tail (higher values). This seems reasonable for the evolution of the building stock, because with increasing refurbishment and the construction of more energy efficient buildings, the share of energy efficient buildings rises, while the share of energy intensive buildings decreases. With the shift of probability density functions, the thresholds for the 15%-criterion shift towards lower values as well. Obviously, the thresholds for the scenario with more ambitious refurbishment (Target 70) are lower than for the other one.

Evolution of threshold values

The evolution of the threshold values for the 15%-criterion shows an almost linear decline (Fig. 8). Still, the two curves have different slopes, which leads to small differences in the beginning and a huge discrepancy until 2045: The Target 40 scenario shows a threshold value of 45.6 while the Target 70 scenario shows a threshold value of 20.3. Due to the stochastic nature of the applied method, all the curves are not completely smooth. If eligibility is determined based on EPCs, then the threshold must lay above that class for the class to be valid. From Fig. 8 can be concluded

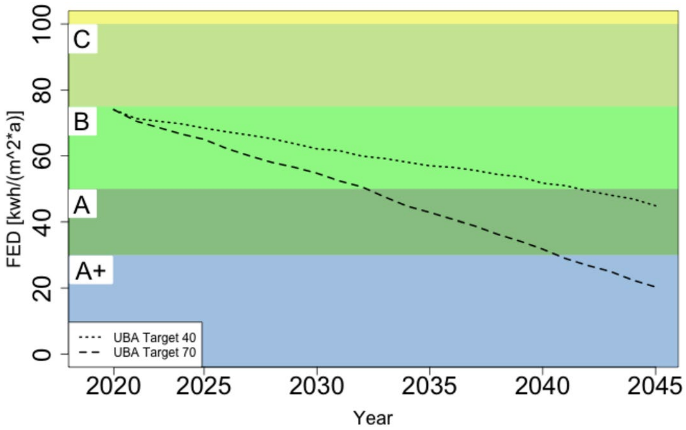

Fig. 8 Evolution of threshold values for the 15%-criterion for two extreme scenarios, depicting refurbishment in German residential buildings (dotted lines). The background colors indicate the ranges of EPCs in Germany. The values can be found in the supplementary material

that based on these scenarios, the necessary EPC for adhering to the 15%-criterion will change from A to A+ between 2033 and 2041 for residential buildings.

Discussion

Limitations

We identify the following limitations.

Although we consider it a strength of the method, that it only requires data on the relative frequencies of EPC classes in a given context, the data that was used in this case is not ideal. Since it is unknown to which extent these EPC classes were attributed based on metered compared to calculated energy consumption (both of which is allowed in Germany), it is also unknown how different shares of the two distinct types of energy metrics would have influenced the histogram that constitutes the basis for the results. Given the fact that the status quo results in this paper are well within the range of the results from a comparable study (purely based on consumption data), we do not think that this affected the overall findings to a large degree. However, this is not necessarily the case. One had just to imagine a behaviour-independent case in which a surprisingly mild winter would disproportionally increase the gap between metered and calculated final energy use in only one direction.

We also assume that the distribution family (weibull) we are fitting to current FED values for German residential buildings also applies to building stocks in other countries and in the future. Of course, this is not a given and due to that, the actual threshold might deviate from the estimates obtained with this methodology. This deviation is likely higher when looking at building types rather than the total set of buildings. Especially non-residential buildings might exhibit larger differences between buildings types (e.g. when comparing a building for commercial use to an office building) than the difference between a single building and the mean average of energy demand. More uncertainty is added because stochastic algorithms are used in the optimization process. This uncertainty might be alleviated, if the optimization is performed multiple times and threshold values are determined based on an average (Monte Carlo Simulation).

An alternative to the approach shown here would be to not only account for the current distribution of energy demand in buildings, but also to look at developments in the past for establishing a base-line of future distributions. However, such a scenario approach would come with caveats as well since the future activities in question, constructing new buildings and removing or renovating existing buildings, are strongly influenced by policy choices as well as the capital costs. One also has to keep in mind that the scenario choice strongly influenced the outcome. In our example, we tried to address this by looking at two scenarios, which mark out the possibility space.

The fitting based on EPCs showed a good fit for higher energy efficiency but deviated substantially for lower energy efficiency. This is not problematic for estimating current threshold values, but changes the shape of the density function of this distribution, which in turn influences the outcome of the estimation for future threshold values. Here, we argue that the tail of the weibull distribution apparently contributes little to the difference in area between density functions, when Fig. 3 is regarded. Therefore, it is expected that the added uncertainty for the estimation of future threshold values is limited.

Overall, there are various sources of uncertainties for the presented method. Nonetheless, we think it can be a valuable tool for estimating threshold values from a wider variety of available data (especially since large datasets with actual or estimated PED or FED values are rare). After all, the reason to come up with approximations like this stems from the fact that there is currently not sufficient data nor guidance to demonstrate if a building is in line with a technical screening criteria in the EU taxonomy regulation.

Application of thresholds

Only one other study reports on the current FED and PED thresholds in the residential building stock in Germany that is comparable to the study at hand (Tschäsch et al. 2022). It sets this threshold at 70 $kWh/(m^2 * a)$ for FED and 74 $kWh/(m^2 * a)$ for PED. This is in line with our proposed threshold of 74 $kWh/(m^2 * a)$ FED in Germany, which, based on the fit to the available data, is robust enough to be used as eligibility criteria to determine which buildings in a portfolio adhere to the 15% criterion. The future thresholds also indicate at which point in time some portion of this portfolio no longer belongs to this set of buildings, or at which point in time a higher EPC has to be used to identify eligible buildings.

In another context, the methodology was also applied to a dataset showing relative frequencies of energy consumption for French office buildings (Observatoire de l'Immobilier Durable 2023). The same source also published a value for the 15%-threshold, so that another test of the proposed methodology was possible: The reported value was 161 $kWh/(m^2 * a)$ PED and the value derived with our method was 158 $kWh/(m^2 * a)$ PED, which underlines the functionality of the method.

However, neither households nor commercial tenants are required to provide information on their FED to their banks. It is therefore not surprising that building datasets would exhibit gaps regarding this as well as other information. We suggest to look for other potential thresholds in these cases such as PED, the EPC label, or the year of construction in light of national building regulations.

The PED can be derived from the FED and the national primary energy factor (PEF) for the energy carriers used in a building. This unit-less factor (e.g. kWh/kWh) depends on the ratio of energy transformation to final demand in a local energy system. Since this value is always higher than 1 for the worst-case of energy provision via fossil energy carriers (Hitchin et al. 2018), any PED threshold that is equal to a FED threshold therefore also must comply with the 15%-criterion (see also section 2.5 for a selection hierarchy). In the case at hand, residential buildings in Germany, slighter lower conservative estimates of 70 $kWh/(m^2 * a)$ PED seem appropriate to ensure eligibility with EU taxonomy. As for EPC, the current technical screening criteria already account for class A to be appropriate for residential buildings, which is confirmed by our data. An additional criterion are national building standards in line with nearly zero energy building (NZEB) regulations. In the case of residential buildings in Germany, current as well as earlier regulations can be compared to the FED and PED thresholds. The German EnEV regulation from 2009 for example suggests that any residential buildings build after 2009 would have a FED of 60 to 70 $kWh/(m^2 * a)$ (Bigalke et al. 2016) and would therefore be in line with the thresholds suggested here.

Future regulations might also require or include thresholds on the direct GHG emissions of buildings. If such data is already available, it can thus also be used to derive a threshold based on the GHG intensity. From the perspective of a worst-case scenario (heating with oil), an FED of 75 $kWh/(m^2 * a)$ could thus correspond to a threshold of 22.5 $CO2 - equivalents/(m^2 * a)$(with 300 g CO2e/kWh according to Braune et al. (2020)).

Conclusion

We proposed a methodology which does not require many resources and gives an idea, how energy efficiency might be structured in future building stocks. This can be a helpful tool for financial service providers and stakeholders that aim for planning security when it comes to the 15% eligibility criterion. Because the results obtained (shown in Fig. 8) are based on scenarios and more uncertainty is added due to stochasticity within the proposed method, they should always be interpreted with care. When in doubt, we suggest to work with conservative values so that adhering to the thresholds (or translations of the thresholds into other variables as suggested in section 4.2) ensures that each selected building aligns with the substantial contribution criteria. Examples for this approach are rounding down rather than taking the concrete values for FEDs or working with building regulations that require buildings to be unequivocally more energy-efficient than the threshold values.

With this in mind, we strongly advise taking several scenarios into account, when this method is applied. For the analyzed case study, the residential building stock in Germany, the results show that the threshold value for the 15%-criterion will decline linearly and ultimately lay between roughly 20 and 45 kWh per square meter and year in 2045. This relatively wide range reflects the most extreme scenario assumptions that lay out the possibility space. At the same time, it highlights how sensitive scenario assumptions are for the outcome. Hence, improving the quality of scenarios, e.g. by elaborating a reduction pathway instead of assuming a linear decline, could increase the robustness of future threshold values. The problem of estimating current threshold values could also be tackled by further methodological improvements. However, a more straightforward approach would be to improve data availability and quality regarding the building stock. After all, in contrast to future building stocks, necessary information could be obtained simply by collecting data. Of course this would imply a great effort, but making building stock data available on a large scale would not only benefit determining eligibility, but is crucial for managing and transforming building stocks towards higher energy efficiency and less emission intensity.

This study, as well as similar studies (Bene et al. 2023; Jakob et al. 2022; Tschätsch 2023), highlighted the challenges of selecting buildings (or financing thereof) in a portfolio that adhere to the 15% criterion in the EU taxonomy. There are currently large variations in both the national definitions of EPCs for residential and non-residential buildings, but also regarding data availability on the building stock in each country and for each category of buildings. However, and more importantly, there are also different national standards regarding the definition of energy-efficient and NZEB buildings in each country (Economidou et al. 2020). Some countries, like Germany, apply a relative reduction of PED in relation to a reference building, and others, like Austria, require adherence to concrete absolute values for the FED or PED of a building. Future research, but also future policies that integrate the EU taxonomy regulation, should take this into account when this and other disclosure regulations are about to be revised. One solution might be to match the national buildings regulations and standards to the 15% criterion in such a way, that the year of construction or modernization can be used to select buildings in a portfolio, since this information is almost always available to the reporting entities. This could be achieved by, for example, investigating time trends in clustered EPC data (similar to the approach for determining a ZEB year in Ferrantelli et al. (2022)).

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s43546-025-00807-w.

Author Contributions CB compiled the relevant data and developed the methodology together with JT. The calculation and presentation of the result was carried out by CB. The contextualization of the study (i.e., Introduction and Discussion) was mainly done by JT. Both authors read and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. The research that led to this publication (rather than the study itself) was funded by the Münchener Hypothekenbank eG.

Data Availability All data generated or analyzed during this study are included in this published article. The code for statistical analysis is released under a permissive license on https://github.com/christianbuschbeck/15perc.git

Declarations

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

Research involving human and animal participants The research did not involve human participants or animals.

References

Ali U, Shamsi MH, Bohacek M, Hoare C, Purcell K, Mangina E, O'Donnell J (2020) A data-driven approach to optimize urban scale energy retrofit decisions for residential buildings. Appl Energy 267:114861. https://doi.org/10.1016/j.apenergy.2020.114861

Ashkar F, Bobée B, Leroux D, Morisette D (1988) The generalized method of moments as applied to the generalized gamma distribution. Stoch Hydrol Hydraul 2(3):161–174. https://doi.org/10.1007/BF01550839

Backenroth S, Lindquist C (2021) *A Green New World - How the Real Estate Sector is Working with Sustainable Financing in Regard to the EU Taxonomy*(Degree Project). Stockholm:KTH Royal Institute of Technology. Retrieved 2024-10-18, from https://www.diva-portal.org/smash/get/diva2:1572516/FULLTEXT01.pdf

Balaras CA, Dascalaki EG, Droutsa KG, Kontoyiannidis S (2016) Empirical assessment of calculated and actual heating energy use in Hellenic residential buildings. Applied Energy 164:115–132. https://doi.org/10.1016/j.apenergy.2015.11.027 Retrieved 2024-10-17, from https://linkinghub.elsevier.com/retrieve/pii/S0306261915014774

Bene M, Ertl A, Horvath A, Mónus G, Székely J (2023) Estimating the energy demand of the residential real estate stock in Hungary based on energy performance certificate data. Financ Econ Rev 22(3):123–151. https://doi.org/10.33893/FER.22.3.123

Bigalke U, Armbruster A, Lukas F, Krieger O, Schuch C, Kunde J (2016) *dena GEBÄUDEREPORT Statistiken und Analysen zur Energieeffizienz im Gebäudebestand* (Tech. Rep.). Berlin

BMWi (2020) *Langfristige renovierungsstrategie der Bundesregierung* (Tech. Rep.). Bundesministerium für Wirtschaft und Energie. Retrieved from https://www.bmwk.de/Redaktion/DE/Publikationen/Energie/langfristige-renovierungsstrategie-der-bundesregierung.pdf?__blob=publicationFile&v=1

BPIE (2014) *Energy Performance Certificates across the EU* (Tech. Rep.). Brussels Buildings Performance Institute Europe (BPIE. Retrieved 2024-12-20, from https://www.bpie.eu/wp-content/uploads/2015/10/Energy-Performance-Certificates-EPC-across-the-EU.-A-mapping-of-national-approaches-2014.pdf

Braune A, Lemaitre C, Geiselmann D, Kreißig J, von Gemmingen U, Klemm S (2020) *Frameworkfor carbon neutral buildings and sites.* Retrieved 2021-11-24, from https://www.dgnb.de/?eID=dumpFile&t=f &download=1 &f=7140 &token=4639f8880b28beb12bf771830b49e105432e310e

Bundesamt für Justiz (2020) *GEG Annex 10.* Retrieved 2024-10-29, from https://www.gesetze-im-internet.de/geg/anlage_10.html

Bürger V, Hesse T, Quack D, Palzer A, Köhler B, Herkel S, Engelmann P (2016) *Klimaneutraler Gebäudebestand 2050* (Tech. Rep.). German Federal Environmental Agency. Retrieved from https://www.umweltbundesamt.de/sites/default/files/medien/378/publikationen/climate_change_06_2016_klimaneutraler_gebaeudebestand_2050.pdf

co2online (2022) *Verteilung Heizenergieverbrauch.* Retrieved from https://www.wohngebaeude.info/daten/#/heizen/bundesweit

Dascalaki EG, Balaras CA, Kontoyiannidis S, Droutsa KG (2016) Modeling energy refurbishment scenarios for the Hellenic residential building stock towards the 2020 & 2030 targets. Energy Build 132:74–90. https://doi.org/10.1016/j.enbuild.2016.06.003

Do H, Kristen SC (2018) Residential Building Energy Consumption: a Review of Energy Data Availability, Characteristics, and Energy Performance Prediction Methods. Curr Sustain/Renew Energy Rep 5:76–85. https://doi.org/10.1007/s40518-018-0099-3

Economidou M, Todeschi V, Bertoldi P, D'Agostino D, Zangheri P, Castellazzi L (2020) Review of 50 years of EU energy efficiency policies for buildings. Energy Build 225:110322. https://doi.org/10.1016/j.enbuild.2020.110322

European Parliament and of the Council (2020) *Regulation (EU) 2020/852 of the European Parliament and of the Council of 18 June 2020 on the establishment of a framework to facilitate sustainable investment, and amending Regulation (EU) 2019/2088 (Text with EEA relevance).* Retrieved 2024-01-18, from https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX%3A32020R0852 (Usr_lan: EN)

European-Parliament (2024) *Directive (EU) 2024/1275 of the European Parliament and of the Council of 24 April 2024 on the energy performance of buildings (recast) (Text with EEA relevance).* Retrieved 2024-12-20, from http://data.europa.eu/eli/dir/2024/1275/oj/eng

Ferrantelli A, Kurnitski J (2022) Energy performance certificate classes rating methods tested with data: how does the application of minimum energy performance standards to worst-performing buildings affect renovation rates, costs, emissions, energy consumption? Energies 15(20):7552. https://doi.org/10.3390/en15207552

Ferrantelli A, Belikov J, Petlenkov E, Thalfeldt M, Kurnitski J (2022) Evaluating the energy readiness of national building stocks through benchmarking. IEEE Access 10:45430–45443. https://doi.org/10.1109/ACCESS.2022.3170716

Glaichen P, Deutsch M, Langenheld A, Peter F, Hauser P, Hein F, Weiß M (2021) *Klimaneutrales Deutschland 2045* (Tech. Rep.). Agora Energiewende. Retrieved from https://static.agora-energiewende.de/fileadmin/Projekte/2021/2021_04_KNDE45/A-EW_231_KNDE2045_Langfassung_DE_WEB.pdf

Gupta R, Gregg M (2018) Targeting and modelling urban energy retrofits using a city-scale energy mapping approach. J Clean Prod 174:401–412. https://doi.org/10.1016/j.jclepro.2017.10.262

Hardy A, Glew D (2019) An analysis of errors in the Energy Performance certificate database. Energy Policy 129:1168–1178. https://doi.org/10.1016/j.enpol.2019.03.022

Hitchin R, Thomsen KE, Wittchen KB (2018) *Primary Energy Factors and Members States Energy Regulations.* Retrieved 2021-11-24, from https://epbd-ca.eu/wp-content/uploads/2018/04/05-CCT1-Factsheet-PEF.pdf

IWU (2016) *Monitor Progress towards Climate Targets in European Housing Stocks Main Results of the EPISCOPE Project* (Tech. Rep.). Institut Wohnen und Umwelt GmbH. Retrieved from https://episcope.eu/fileadmin/episcope/public/docs/reports/EPISCOPE_FinalReport.pdf

Jakob M, Ostermeyer Y, Nägeli C, Hofer C (2022) Overcome data gaps to benchmark building stocks against climate targets related to the EU taxonomy and other decarbonisation initiatives. IOP Conf Ser: Earth Environ Sci 1085(1):012042. https://doi.org/10.1088/1755-1315/1085/1/012042

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

Li Y, Kubicki S, Guerriero A, Rezgui Y (2019) Review of building energy performance certification schemes towards future improvement. Renew Sustain Energy Rev 113:109244. https://doi.org/10.1016/j.rser.2019.109244

Mellwig P, Pehnt M, von Oehsen A, Blömer S, Lempik J, Werle M, Linke C (2018) *Building sector Efficiency: A crucial Component of the Energy Transition - Final report on a study conducted by the Institut für Energie- und Umweltforschung Heidelberg (Ifeu), Fraunhofer IEE and Consentec* (Tech.

Rep.). Agora Energiewende. Retrieved from https://static.agora-energiewende.de/fileadmin/Projekte/2017/Heat_System_Benefit/163_Building-Sector-Efficiency_EN_WEB.pdf

Metzger S, Jahnke K, Walikewitz N, Otto M, Grondey A, Fritz S (2019) *Wohnen und Sanieren* (Tech. Rep.). Umweltbundesamt. Retrieved from https://www.umweltbundesamt.de/sites/default/files/medien/1410/publikationen/2019-05-23_cc_22-2019_wohnenundsanieren_hintergrundbericht.pdf

Observatoire de l'Immobilier Durable (2023) *2023 Barometre de la performance energetique et entvironmentale des batiments* (Tech. Rep.). Paris. Retrieved 2024-10-29, from https://conseils.xpair.com/agenda_news/barometre-2023-performance-energetique-environnementale-batiments.htm

Park HS, Lee M, Kang H, Hong T, Jeong J (2016) Development of a new energy benchmark for improving the operational rating system of office buildings using various data-mining techniques. Appl Energy 173:225–237. https://doi.org/10.1016/j.apenergy.2016.04.035

Pasichnyi O, Wallin J, Levihn F, Shahrokni H, Kordas O (2019) Energy performance certificates - new opportunities for data-enabled urban energy policy instruments? Energy Policy 127:486–499. https://doi.org/10.1016/j.enpol.2018.11.051

Paterson G, Mumovic D, Das P, Kimpian J (2017) Energy use predictions with machine learning during architectural concept design. Sci Technol Built Environ 23(6):1036–1048. https://doi.org/10.1080/23744731.2017.1319176

Rinne H (2008) *The Weibull Distribution A Handbook*. Retrieved from https://doi.org/10.1201/9781420087444

Sesana MM, Salvalai G, Della Valle N, Melica G, Bertoldi P (2024) Towards harmonising energy performance certificate indicators in Europe. J Build Eng 95:110323. https://doi.org/10.1016/j.jobe.2024.110323

Shi X, Si B, Zhao J, Tian Z, Wang C, Jin X, Zhou X (2019) Magnitude, causes, and solutions of the performance gap of buildings: a review. Sustainability 11(3):937. https://doi.org/10.3390/su11030937

Statista (2021) *Verteilung der Energieeffizienzklassen bei Immobilien in Deutschland im Jahr 2021*. Retrieved from https://de.statista.com/statistik/daten/studie/1284714/umfrage/verteilung-energieeffizienzklassen-immobilien/

Tschäsch C, Rößler J, Adelhoefer C (2022) *EU Taxonomy: Climate Change Mitigation Real Estate Activities - Eligibility Criteria - Top 15% in Germany.*

Tschätsch C (2023) *EU taxonomy: Aquisition and ownership of buildings -Derivation of top 15% of existing building stock in Germany* (Tech. Rep.). Retrieved from https://www.vdpresearch.de/wp-content/uploads/2023/10/Drees_Sommer_Veroeffentlichung-VDP_top15_summary_update_2023.pdf

Yuan M, Choudhary R (2023) Energy Performance Certificate renewal - An analysis of reliability of simple non-domestic buildings ' EPC ratings and pragmatic improving strategies in the UK. Energy Policy 178:113581. https://doi.org/10.1016/j.enpol.2023.113581