Energy Efficiency (2026) 19:5
https://doi.org/10.1007/s12053-025-10405-2

RESEARCH

Beyond energy labels: estimating housing energy efficiency and energy underconsumption using administrative microdata

Lydia Geijtenbeek · Peter Mulder · Robin Niessink

Received: 7 April 2025 / Accepted: 6 December 2025

Abstract A reliable measure of the energy efficiency of housing is essential, both for evaluating the effectiveness of energy efficiency policies and for assessing energy poverty. Across the EU, Energy Performance Certificates (EPCs) or energy labels are commonly used for this purpose. However, these data are often outdated or incomplete, and only weakly correlated with actual energy consumption —a discrepancy known as the performance gap. As a result, EPCs are poorly suited for evaluating energetic housing quality or measuring energy poverty. We address these limitations by developing and implementing a model that estimates housing energy efficiency by combining EPC data with additional administrative resources. Our approach improves upon previous studies through richer data integration and more precise model calibration. The resulting model explains 51% of the variation in energy expenditure based on housing characteristics, compared to 40% when using EPCs alone. We demonstrate the model's application in assessing energy poverty through the LILEQ indicator (i.e. Low Income, Low Energy Quality), showing that it correlates more strongly with survey-based consensual indicators of energy poverty (e.g. EU-SILC), than commonly used indicators based on actual expenditure (e.g. share of energy expenditure). Finally, we illustrate how the model can be used to detect energy underconsumption and monitoring housing quality over time. In summary, we present a calibrated, data-driven model of housing energy efficiency that outperforms EPCs and enables the development of higher-quality, policy-relevant measures of energy poverty and housing conditions.

Keywords Housing energy efficiency · Energy performance certificates (EPCs) · Energy labels · Energy poverty · Microdata · Energy policy evaluation

Introduction

Background and literature

In response to ambitious climate targets and persistently high energy prices, both governments and individual citizens across the European Union (EU)

L. Geijtenbeek · P. Mulder (✉)
Centraal Bureau voor de Statistiek, The Hague, Netherlands
e-mail: p.mulder@tno.nl

L. Geijtenbeek
e-mail: l.geijtenbeek@cbs.nl

P. Mulder · R. Niessink
Energy and Materials Transition, Netherlands Organisation for Applied Scientific Research, Amsterdam, Netherlands
e-mail: robin.niessink@tno.nl

P. Mulder
Energy and Materials Transition, Copernicus Institute for Sustainable Development, Utrecht University, Utrecht, Netherlands

have intensified efforts to improve energy efficiency within the residential sector. As part of this strategy, the revised Energy Performance of Buildings Directive (EU/2024/1275) mandates a 16% improvement in the energy performance of residential buildings by 2030. This requirement reflects the dual importance of energy efficiency: reducing greenhouse gas emissions and mitigating the socio-economic consequences of high energy costs. Residential energy use significantly contributes to both global emissions and household expenditure, thereby affecting income distribution, particularly in the context of rising energy prices. Consequently, energy efficiency has become a focal point for policies at European, national, and regional levels. Accurate, granular, and timely data on energy efficiency, energy use, and energy poverty are essential for designing and evaluating such policies.

Energy performance certificates (EPCs) or energy labels are widely used to assess housing energy efficiency. However, in practice, these labels often suffer from incompleteness and inaccuracy. For instance, in the Netherlands, approximately 39% of dwellings lacked a valid energy label in 2024 (CLO, 2025), while existing EPCs are not always updated following renovations, and are thus often outdated or inaccurate (Van den Wijngaart & van Polen, 2020). Furthermore, a significant body of empirical research has demonstrated that EPCs often do not reflect actual residential energy use. This mismatch, known as the energy performance gap, undermines the utility of EPCs for policymaking, household decision-making, and energy poverty measurement.

Majcen (2016) conducted an early but influential study using Dutch microdata, showing that theoretical EPC-based energy use projections systematically overestimated actual household consumption, particularly in low-efficiency dwellings. Although somewhat dated, her work remains foundational in the Dutch context. Subsequent studies have consistently reinforced these conclusions. Meles et al. (2023) assessed EPC reliability across the housing stock in Ireland and found that EPCs explain only a modest portion of the variance in measured heat loss. Few et al. (2023), in a comprehensive study of the EPC-consumption gap in Great Britain using data from 1,300 homes with both energy label information and actual consumption records, found that EPCs systematically overpredict energy use for homes with lower EPC bands, with the largest discrepancies in the lowest-rated homes. Notably, the performance gap persisted even in homes with behavioral conditions aligned with EPC model assumptions—such as standard occupancy and thermostat use—suggesting that technical or structural assumptions, rather than behavior alone, drive much of the mismatch. This finding is echoed in Van Hove et al. (2022), who analyzed over 100,000 homes in Flanders and found substantial deviations between modeled and actual energy use. Their study also emphasizes multicollinearity and model overspecification as important contributors to EPC inaccuracies. While the former studies address the static performance gap, there is also a dynamic performance gap related to retrofitting. In this context, Peñasco and Díaz Anadón (2023), Allcott and Greenstone (2017), and Fowlie et al. (2018) highlight the systematic overestimation of energy savings following retrofitting measures based on EPC-derived predictions.

In response to the limitations of EPCs, researchers have turned to data-driven models that estimate actual energy consumption using housing and household characteristics. These approaches—ranging from classical statistical models to advanced machine learning algorithms—have demonstrated superior predictive performance and practical scalability. Amasyali and El-Gohary (2018) provide a review of building-level energy prediction models. They distinguish between physics-based ‘forward models’ that simulate energy use using detailed physical and environmental data, and data-driven models that rely on historical consumption patterns and statistical learning. While the former require extensive, often inaccessible inputs, data-driven models can leverage administrative microdata to make reliable predictions across large housing stocks. Recent advances in machine learning further support this shift. Cui et al. (2024) find that models incorporating structural characteristics and household composition yield high predictive accuracy across residential building types. Similarly, Harputlugil and de Wilde (2021) emphasize that human-building interaction is essential for understanding residential energy consumption, which models using occupant and socio-economic data can incorporate more easily than EPCs. These models often use features such as dwelling type, floor area, insulation presence, heating system, construction year, and renewable energy installations (e.g., solar panels). For example, research by Van den Wijngaart and Van Polen (2020), followed up by van Beijnum et al. (2023) use a regression model based on publicly available dwelling features to predict regional residential heat demand. Their approach served as an inspiration to our approach in this paper.

Energy poverty, broadly defined as the inability to attain adequate energy services at affordable cost, is traditionally measured by expenditure thresholds. Common instances are a share of energy expenditure above 10% or twice the median, as used by the European Commission (Rademaekers et al., 2016). Yet these indicators receive criticism for not excluding households with high incomes and strong preferences for energy services above socially accepted standards (Al Kez et al., 2024). Furthermore, they do not account for housing efficiency and therefore fail to distinguish between low energy consumption due to either effective insulation or financial hardship (Mulder et al., 2023). Several alternative frameworks aim to integrate housing quality in energy poverty measurements. The UK's Low Income Low Energy Efficiency (LILEE) measure, for example, combines income with EPC bands A–C as a threshold (Department for Business & Energy and Industrial Strategy, 2022). Yet, as Semple et al. (2024) show, this approach still has limitations: in low-income neighborhoods in London, 27% of households in homes rated EPC A–C reported experiencing energy insecurity, indicating underestimation of energy poverty. Karpinska and Smiech identify 'hidden energy poverty' as a combination of low energy efficiency and budgetary constraints (2020), and construct empirical typologies of energy-poor households using actual energy use data (2023). They argue that reliable poverty assessment requires linking income, housing costs, and energy use, thereby emphasizing the need for robust energy efficiency estimation; something EPCs fail to provide consistently.

Goal and scope

This paper addresses the limitations discussed above of both EPCs and energy poverty indicators, by proposing a model that estimates housing energy efficiency more accurately than EPCs, and to apply this to the assessment of energy poverty and underconsumption.

For this model, we integrate EPCs with a broad range of administrative data, thus building upon previous models that use Dutch microdata to estimate housing quality above or in the absence of EPCs (e.g., Van den Wijngaart & Van Polen, 2020, van Beijnum et al., 2023, Mulder et al., 2023). Similarly to these models, we aim to generate robust estimates of energy efficiency that remain unaffected by occupant behavior, while ensuring that the model can be applied to most of the Dutch housing stock. To achieve this, we use data available consistently across all Dutch homes from 2019 at Statistics Netherlands (CBS).

Our model differs in a few respects from similar models. First, since our main context is energy poverty, our dependent variable is total net energy expenditure, and not energy consumption or energy use for space heating. This means that our measure of housing quality does not only depend on insulation levels, but also on energy generation (e.g., by solar panels), and the share of electricity in energy consumption. Second, although we base our estimates primarily on characteristics of a dwelling, we also control for social factors of the inhabitants. Third, we use additional variables such as the presence of solar panels and main heat source. And finally, we use a sparser and more finely tuned specification with fewer degrees of freedom.

The model presented in this paper builds on ongoing collaboration between CBS and TNO to monitor and analyze energy poverty in the Netherlands. Central to this work is the development of the LEQ (Low Energy efficiency Quality) indicator, designed to identify households living in poorly insulated dwellings—a key structural driver of energy poverty. Introduced as a building block in a broader monitoring framework established in 2020, LEQ supports more accurate diagnosis and targeting of vulnerable households. This framework has informed national policy through an annual regionally detailed assessment of energy poverty across the country and is detailed in a series of publications, including Statistics Netherlands (2024), Batenburg et al. (2024) and Mulder et al. (2023).

The remainder of this paper is structured as follows: Sect. 2 describes the data sources and model design; Sect. 3 demonstrates applications to energy efficiency and poverty analysis; Sect. 4 discusses results, limitations and implications; while Sect. 5 concludes.

Methodology

Data and variable selection

This research uses the rich administrative data set that is available to researchers at Statistics Netherlands (CBS). The development and estimation of our model was based on a population of 7.8 million houses in 2019, for which many attributes are available.

The dependent variable in all of our estimations is the *net energy expenditure* at an address. Energy expenditure is calculated by using the net supply of gas[1] and electricity to any home in the Netherlands, and combining these with average gas and electricity prices (including fixed and variable costs and taxes). This means that energy expenditure is not restricted to energy consumption for heating, but includes all energy costs and revenues attributable to a certain home, including the energy costs of hot tap water or electrical appliances, the financial benefits of electricity generations by solar panels, and cost differences in gas versus electrical heating. Note that the Dutch EPC-label system also combines the level of insulation with electricity generation and gas-independence in determining the label. The reason that we use net energy *expenditure* as our metric is that expenditures are the relevant concept in the context of energy poverty.

Out of the vast available data at CBS, we considered both physical characteristics of dwellings (e.g., floor area, construction period) and socio-economic characteristics of households (e.g., disposable income, household size) that were found to be related to energy consumption in other existing research, or were otherwise deemed relevant. From other studies (e.g., van Beijnum et al., 2023) we have learned that it is important to keep the number of distinct variables low, so that relevant interactions can be included without adding too many degrees of freedom to the model.

We used decision trees for our variable selection (or, more specifically, regression trees, since our dependent variable *energy expenditure* is continuous). In a decision tree, all variables are assessed for their ability to reduce the variation in the outcome variable by recursively splitting the data points in two groups that have a lower variability in the outcome variable than the original group. For the variable selection we use two metrics that are derived from the regression tree algorithm. The first is the *importance* metric (Ishwaran, 2007), which ranks the variables in their ability to distinguish high from low energy expenditure dwellings. A drawback of the importance metric is that correlated variables typically have similar importance, even if one variable always performs better in reducing variation of the outcome variable. Therefore we use a second metric, the *weighed node count*, i.e. the number of times that a variable is selected in the decision tree, weighed by the share of cases in that particular node of the tree, which tends to exclude variables that are correlated to a variable with a stronger distinguishing power.

Table 1 shows the initial variable list. Next are the two outcomes for two different regression trees (tree 1 and tree 2). For the first tree, which has energy expenditure as its dependent variable (tree 1), we find that floor area has the largest importance, followed by building type, property value, income, and number of people. The largest node count is found for area, building type, people and solar, while the node count is zero for property value and income. This suggests that these latter two variables do not directly influence energy expenditure, but rather have an indirect effect through area. To test this assumption, we did a second regression tree analysis (tree 2), which has energy expenditure corrected for area as its dependent variable.[2] There we find that the importance of income increases, yet that of property value decreases significantly.

Our final selection are the bold face variables in Table 1. Of these, area is included because of its high scores in tree 1, and other variables are selected if their importance or node count in tree 2 is above the

[1] About 6–7% of houses use district heating, instead of gas or electricity. For these dwellings we have no data on energy consumption nor prices. We therefore use estimates for the energy use combined with gas prices.

[2] For the correction we used the piece wise linear specification as described below, and mentioned in Table 2.

Table 1 Variables that were considered and the values of their selection metrics

Variable	Description of variable	Tree 1, dependent = energy expenditure		Tree 2, dependent = energy expenditure, corrected for area	
		Importance	Node count	Importance	Node count
Area	Floor area of the house (in m^2)	**1.00**	**1.57**	0.14	**0.41**
People	Number of inhabitants	0.35	1.19	**0.82**	**1.46**
Solar	Presence of solar panels (yes/no)	0.23	0.66	**1.00**	**1.00**
BuildType	One of: detached house, semi-detached house, corner house, terraced house, apartment	0.79	1.25	**0.50**	**0.99**
BuildYear	Year of construction of the dwelling	0.14	0.39	**0.57**	**1.17**
Label	Certified energy label (EPC): A (incl., A +, A + +, …), B, C, D, E, F, or No label	0.10	0.27	0.27	**0.36**
HeatSource	Main heat source. One of: individual gas heating, block heating, district heating, individual electric heating	0.13	0.18	0.23	**0.33**
Income	Disposable household income (at address), in percentiles	0.38	0.00	**0.44**	0.00
PropertyValue	Taxable property value of the house	0.49	0.00	0.21	0.11
HasGas	House had a connection to the gas grid; either for heating or (also) warm water, and/or cooking	0.04	0.00	0.09	0.00
InhabitYears	Number of years that the current residents live in the house	0.04	0.00	0.05	0.00
Ownership	Ownership type of the house	0.13	0.00	0.03	0.00
LabelYear	Year that the label (EPC) was issued	0.10	0.00	0.02	0.00
IncomeSource	Main income source of the household (e.g., wage, profits, benefits for unemployment, social security, or sickness, pension, etc.)	0.02	0.00	0.02	0.00
Age	Age of main resident of the household	0.03	0.00	0.02	0.00

The first two columns contain variable name and description. Columns 3–6 give importance and node count for two different regression trees: tree 1 has energy expenditure as dependent variable, and tree 2 has energy expenditure corrected for floor area as dependent variable. Bold face variables were selected, and bold face values indicate that they were used as selection criterium

Table 2 Variables that were selected and a description of their functional form

Variable name	Specification and form of the variable
Area	Linear, capped from above at $302m^2$; for use in interactions with other variables
AreaPWL	Piece wise linear (PWL), in 12 pieces with selection of split points based on significance in a separate regression model
BuildPeriod	Categorical, in 9 periods (i.e., oldest – 1924–1947–1963–1995–2002–2015–2017 – most recent), determined by a separate regression model, which reflect changes in historical building regulations
BuildType	Categories: detached house, semi-detached house, corner house, terraced house, apartment, unknown
Label	Categories: A (incl., A +, A + +, …), B, C, D, E, F, or No label
Solar	Boolean (true/false)
HeatSource	Categories: individual gas heating, block heating, district heating, individual electric heating, unknown
People	Categories 0, 1, 2, 3, 4, and 5 or more
Income	Categorical, in 9 income brackets based on percentiles of disposable household income (i.e., unknown or 1 st percentile, percentiles 2–4, 5–19, 20–53, 54–68, 69–89, 90–96, 97–99, and 100th percentile)

Readers that want to know the exact shape (including cut-offs) will receive this information on request by the first author

Table 3 Comparison of different versions of the model, which differ in the number of two-way interactions that were included

Description of the model	Degrees of freedom	Model quality (adjusted R2)
No interactions, only separate variables	50	0.566
All interactions between two variables	649	0.586
Selected set of interactions (see Eq. 1 and Table 4)	137	0.579

(somewhat arbitrary) threshold of 0.30. Note that type of ownership was excluded due to its low scores on both metrics for both trees, although it was included in previous models (Mulder et al. (2023); van Beijnum et al., 2023).

Model

For the selected variables, we evaluated the shape of the relationship with energy expenditure, and found strong non-linearities for most numerical variables. Therefore, most variables are included in the model as categorical variables, with categories that mirror the observed non-linearities. Only area appears roughly linear, yet with an upper bound at about 300m^2. Since area is such a strong predictor, we assessed other specifications than purely linear, such as loglinear, polynomial and piece wise linear, and found that a higher degree polynomial or piece-wise linear function best captures the slight S-shape of the relation between area and energy expenditure. In the final model we use a piece-wise linear form for area combined with capped linear form to be used in interactions with other variables. Table 2 lists the variables that are included in the final model, along with a description of their selected form.

Both the regression tree and existing research indicate that interactions of variables are potentially relevant for the prediction of energy consumption. Van den Wijngaart and Van Polen (2020) used a full 4-way interaction of all their variables (i.e. building period, EPC label, house type, floor area), while van Beijnum et al. (2023) used a 5-way interaction of their set of variables that adds type of ownership, while Mulder et al (2023) included a 3-way interaction of building period, house type and floor area. This resulted in large models that are computationally hard to estimate, and are potentially prone to overspecification.

For the current model we only considered two-way interactions. We evaluated several instances, ranging from no interactions at all to including all two-way interactions, on data from 2020. The main results are summarized in Table 3. The final model (see also Eq. 1) attempts to strike a balance between accuracy and sparsity, by only selecting the interactions with the largest F statistic. This final model can explain more than half of the variation in energy expenditure ($R^2=0.584$) in our base year 2019, and is thus reasonably effective.

The resulting model includes both main effects and a selection of two-way interactions that were found to significantly improve explanatory power. The full specification of the model is given in Eq. 1 below. Note that most of the terms in this equation relate not a single variable, but sets of dummy variables. For instance, BuildPeriod consists of 8 separate dummy variables for building periods (excluding the last, as reference category).

Equation 1 Formal specification of our model, with energy costs as a function of characteristics of both the house and its inhabitants

$$
\begin{aligned}
EnergyCost = {} & \alpha + \beta_1 Area + \beta_2 AreaPWL + \beta_3 Solar \\
& + \beta_4 BuildPeriod + \beta_5 Label + \beta_6 HeatSource \\
& + \beta_7 BuildType + \beta_8 Income + \beta_9 People \\
& + \beta_{10} Area * Solar + \beta_{11} Area * Income + \beta_{12} Area \\
& * HeatSource + \beta_{13} Area * People + \beta_{14} Area * Label \\
& + \beta_{15} Area * BuildPeriod + \beta_{16} Solar * BuildType \\
& + \beta_{17} Solar * HeatSource + \beta_{18} Solar * People \\
& + \beta_{19} Solar * Label + \beta_{20} HeatSource * People \\
& + \beta_{21} BuildYear * Label
\end{aligned} \tag{1}
$$

Table 4 shows some summary statistics for the terms in Eq. 1. First, we list the degrees of freedom for each term, or the number of underlying dummy variables that are used to model a term (e.g., BuildPeriod consists of 8 dummies for different building periods, BuildType had 4 dummies for different building types). The degrees of freedom are especially large for the interactions (e.g., 25 for the interaction

Table 4 Combined significance, F-score and degrees of freedom for (sets of) variables

Variable, set of variables, or interaction between two (sets of) variables	Degrees of freedom (i.e., number of dummy variables)	F-score	Combined significance
Intercept	1	4,306.561	0.000
Solar	1	4,624.407	0.000
Area (linear, incl. dummy for unknown)	2	3.799	0.024
AreaPWL (piece wise linear)	12	70.925	0.000
BuildPeriod	8	4,234.216	0.000
Label	7	856.989	0.000
BuildType	5	101,862.046	0.000
Income	8	561.676	0.000
People	4	9,115.203	0.000
HeatSource	4	858.785	0.000
Solar * Area	1	6,442.281	0.000
Income * Area	8	3,197.782	0.000
HeatSource * Area	4	546.129	0.000
People * Area	4	1,507.040	0.000
Label * Area	5	481.251	0.000
BuildPeriod * Area	5	2,374.237	0.000
BuildType * Solar	5	3,456.916	0.000
HeatSource * Solar	4	1,348.715	0.000
People * Solar	4	1,240.410	0.000
Label * Solar	5	1,448.728	0.000
People * HeatSource	16	550.522	0.000
BuildPeriod * Label	25	921.223	0.000
Total regression	138	66,866.826	0.000

of building period and EPC label), as these require a separate dummy variable for each combination of variables (e.g., each combination of building period and EPC label). The bottom row shows the degrees of freedom for the entire model, i.e. the number of coefficients, which is 138. Note that due to this substantial number of coefficients, we do not include a full table of estimated parameters.[3]

Next provide information on the significance of the terms by means of their F-statistic and its p-value. We find that for each of the terms the combined effect of the underlying dummy variables was highly significant (at $p < 0.001$). Since the p-values are all very low, we use F-scores to determine which variables are most influential in the regression. Note that variables that are also included in one or more interactions typically have a lower F-score, since these interactions already account for some of the explanatory power of the single variable. For instance, building type has the largest F-score while only being included in a single interaction, while area is included in six interactions and has the lowest F-score. Among the most influential predictors are the type of dwelling, number of residents, presence of solar panels in interaction with floor area, and construction period. Notably, interaction effects between solar presence and housing characteristics (e.g., area, label, build type) suggest non-uniform impacts of energy improvements across the housing stock. For (semi-) detached houses and larger houses the effect of solar panels on energy expenditure is larger; possibly because these houses typically have larger roofs that can in turn support more solar

[3] Also note that due to the large number of interactions (i.e. 12 of 22 terms), it is difficult to interpret the estimated coefficients separately. For instance, some interaction with area could have a negative parameter estimate, but as long as the estimate for Area is larger (in absolute terms), the total effect would still be positive.

Table 5 Comparing prediction of energy consumption by labels versus estimated efficiency

Energy efficiency indicator	Model specification	Degrees of freedom	R^2
EPC label, in 8 categories (A-G, and 'No label')	EnergyCost = Label + Label * Area	15	0.403
Estimated energy efficiency score	EnergyCost = EEScore + EEScore * Area	2	0.515
Estimated energy efficiency score, in eight categories of equal share	EnergyCost = EEScore_cat + EEScore_cat * Area	15	0.480

panels. For EPC-labels, on the other hand, we find that the positive impact of solar panels is smaller for better labels (e.g., A, B), which is probably because solar panels are also incorporated in the label.

In its raw form, this model can be used to estimate annual energy expenditure of a certain combination of characteristics for a house and its inhabitants. This is relevant for predictions or scenario studies. Yet in this paper we use the model for two other types of estimates, as will be described in the next section.

Results and applications

The results of the previous section lay the empirical foundation for several policy-relevant applications, which we explore in this section: estimating energy efficiency, identifying energy poverty, and detecting over- and underconsumption. Together, these three applications illustrate the versatility and policy relevance of such a model, providing new insights into both structural inefficiencies and behavioral adaptations within the residential energy landscape.

Estimating energy efficiency

In this context, energy efficiency relates to the energy costs that are inherent to a building. For this, we use our model while abstracting from household characteristics and floor area to obtain an estimate for the energy efficiency of a dwelling, which we define as the energy expenditure that is inherent to a building. To isolate this intrinsic property from occupant behavior and household characteristics, we simulate each dwelling's energy cost using standardized values for floor area ($100m^2$), and number of residents (2 people), and income (median). This procedure results in an estimated yearly energy cost for the house, which is independent of its inhabitants or floor area. We will refer to this value as Energy Efficiency Score (or EEScore).

In Table 5 we compare our model-based efficiency score with EPC labels. To this end, we model the actual energy expenditure as a linear combination of efficiency and floor area, and measure the predictive power of the model in the form of its R^2. We evaluated three models: one for energy labels (EPCs) in 8 categories, including a category 'No label', a second which uses our continuous EEScore, and a third which uses a discrete version of our measure. We added the third option, because being continuous may give the EEScore an unfair advantage with respect to the discrete EPC labels. Thus, for fair comparison, we also created a discrete version of the EEScore in 8 levels (like the 8 categories of A-G with an additional category 'No label').

Table 5 demonstrates that our efficiency score explains a greater proportion of variation in energy use ($R2 = 0.515$) than EPC-labels ($R2 = 0.403$). Furthermore, this difference is only for a smaller part due to its continuous nature, since two thirds of the increase in predictive power is already present for the discrete version of the EEScore ($R2 = 0.480$). In other words: our modelled version of housing energy efficiency (EEScore) explains 51% of the variance in actual energy cost, versus 40% for using labels and floor area alone.

Table 6 indicates how EPC labels align with our model-based efficiency score (EEScore), or energy quality. To this end we use two binary indicators that were derived from the efficiency score: Low Energy Quality (LEQ) coincides with the 50 percent least efficient homes, and Very Low Energy Quality (VLEQ) with the 15 percent least efficient homes, both with respect to the Dutch housing population in 2019. Table 6 shows how EPC labels and model-based efficiency quality are relatively well-aligned, with labels A-C being most prevalent in non-LEQ

Table 6 Relation between our estimate of energy efficiency and energy labels (EPC) in 2019

Category of model-based efficiency score	Total	Share of total (%)	Share of houses in the row with a certain EPC energy label (%)							
			A	B	C	D	E	F	G	No label
Total housing stock	7,814,910	100	8	8	14	8	5	3	2	53
Non-Low Energy Quality (non-LEQ)	3,905,890	50	15	13	22	8	3	1	1	36
Low Energy Quality (LEQ)	3,909,030	50	1	2	6	9	6	4	3	69
Very Low Energy Quality (VLEQ)	1,196,680	15	0	1	2	2	2	3	4	87

homes, D-F labels in LEQ homes, and G labels in VLEQ homes. Yet, some misalignment occurs, which is difficult to explain. For instance, of the homes classified by the model as inefficient, 1% have an EPC rating of A, which could be due to either misclassification of energy labels, outliers in energy consumption, or structural overconsumption for certain combinations of model variables.

Next, and most importantly, while 53% of all homes do not have a valid EPC label in 2019, among the homes with the largest predicted energy costs (VLEQ) a majority of 87% does not have a valid EPC label. This means that identifying low energy efficient homes based on labels alone would exclude most of such homes.

Monitoring energy efficiency

One key application of our model is tracking the evolution of energy quality in the housing stock over time. Using the calculation method and indicators as described above, we find that the number of homes with low energy quality (LEQ) in the Netherlands decreased by a large share,[4] from 49.8% in 2019 to 22.9% 2023, while the number of very low energy quality homes (VLEQ) decreased from 15.3% in 2019 to 9.0% in 2023. See also Fig. 1, which illustrates the share of dwellings classified as low energy quality (LEQ) or very low energy quality (VLEQ) in municipalities in the Netherlands in 2019 and 2023. The figure shows a discernible trend to better quality; whereas in 2019 municipality averages were mostly between 50–60% for LEQ and 30–40% for VLEQ, in 2023 the most common ranges are 20–30% for LEQ and 10–20% for VLEQ.

Further note that EPC labels are not a reliable tool for tracking the share of low-quality homes, because most low quality homes have no label. For instance, in the Netherlands 5.0% of the homes changed from no label to a low quality label (D-G) between 2019 and 2023, while in the same period 4.7% of homes transitioned from a low quality to high quality (A-C) label. As a result, the net effect is that the share of low quality labels increased by 0.3%, thus giving the false impression that the average quality of the Dutch housing stock deteriorated. This means that in the face of homes without a label, EPC-labels are not a useful tool for monitoring the energy quality of the existing housing stock. Thus, our new measure fills a problematic gap in the current data.

There is one caveat in using LEQ for assessing changes in housing energy efficiency, which is that not all improvements can be detected by this measure. We can only detect changes that reflect in the administrative data. These are attaining an EPC label (when it was previously missing), getting a new different EPC label, newly installed solar panels, or a change in main heat source.[5] Among the 1.4 million homes that were LEQ in 2019 and not in 2023, we saw the following changes: 73% had newly installed solar panels, 49% had a different label (including changes between an actual label and the category 'No label'), and 12% had a different main heat source, e.g. district heating, or electric heating. Of these houses 67% had only one change: 44% only had new solar

[4] Note that we only look at the homes that were present in both years, thus excluding demolished or newly built homes from these percentages. Which also explains why the percentages LEQ and VLEQ of 2019 are not exactly 50.0% and 15.0%.

[5] There is also a small number of homes with a different LEQ-status due to administrative changes in housing type (14,000) or construction year (28,000).

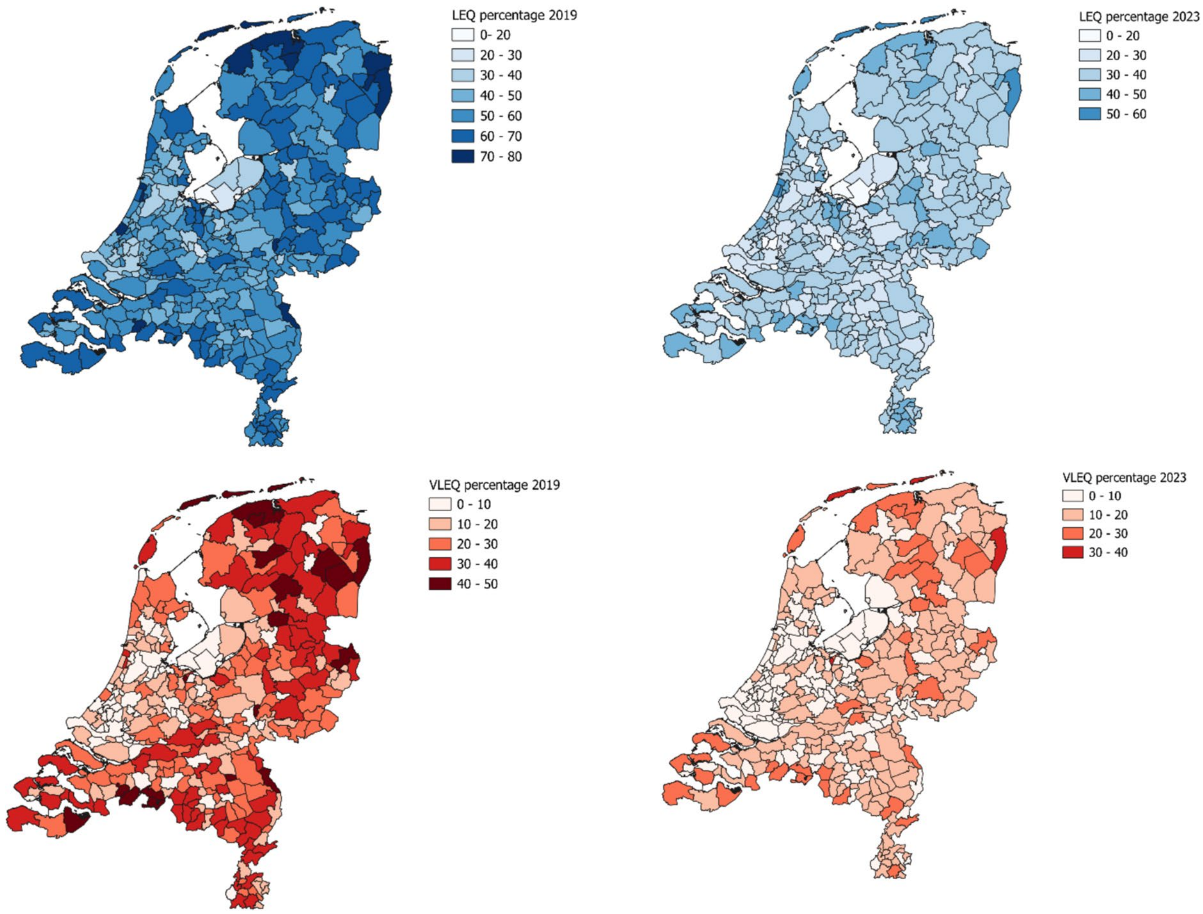

Fig. 1 Share of low energy quality (LEQ), or very low energy quality (VLEQ), houses in the Netherlands, 2019–2023

panels, 22% only had an improved EPC label, and 1% only a different heat source. Note that although solar panels do not improve home insulation, they do typically decrease the energy expenditure of a home, and are therefore incorporated in our notion of housing energy efficiency.

Even though improvements in energy efficiency as measured by LEQ are based on administrative data, they still correlate strongly with energy expenditure. When we compare energy expenditure of homes between 2019 and 2023, we find that on average the energy expenditure decreases by 380 euros per year (from 1,790 euros in 2019 to 1,410 euros in 2023, in prices of 2019). Note that homes with an A label in 2019 without any observed changes still decreased their energy expenditure by 130 euros on average, which suggests that this decrease appears to be at least partly caused by changes in behavior, such as setting the thermostat lower or showering shorter. Yet for homes that went from LEQ to not-LEQ this increase was significantly larger with 780 euros, as opposed to 340 euros for homes that remained LEQ. For homes that transitioned from VLEQ tot not-LEQ the decrease was even larger, with a decrease of 1,170 euro per year, compared to a decrease of 470 for homes that remained VLEQ. Thus, one can easily image that improving a home to non-LEQ could reduce energy poverty.

Estimating energy poverty

Most expenditure based energy poverty indicators rely on observed energy expenditure in relation to income. However, this approach risks overlooking households who under-consume energy due to financial constraints, since low energy use can be misinterpreted

Table 7 Overview of Dutch energy poverty indicators and the share of households that are considered energy poor according to each indicator, 2019

Dutch energy poverty indicators	Count[a]	Percentage
Total households for which energy poverty was measured	6,963,900	100
Households with high energy quote (HEQ)	513,500	7.4
Households with low income and high energy expenditure (LIHE)	436,800	6.3
Households with low income and low energy quality (LILEQ)	392,100	5.6
Households with low income and high energy expenditure *or* low energy quality (LIHELEQ)	601,700	8.6

[a]Note that the monitoring excludes households, with the largest groups being households with (partially) unknown income, student households, and households that share a common address. This leaves out roughly 1 million Dutch households (Statistics Netherlands, 2024). Statistics Netherlands is working on filling some of these gaps in the next update

as energy efficiency. By interpreting our model-based estimates as 'required energy expenditure', we can instead identify households whose structural energy needs are high but whose income is insufficient to meet those needs, irrespective of whether they consume much energy or not. This enables more accurate targeting of households experiencing 'hidden' or structural energy poverty.

Table 7 presents a representative set of energy poverty indicators, including two that incorporate our Low Energy Quality (LEQ) measure, as described in the previous section. TNO and CBS have developed a national framework for monitoring energy poverty, introduced in 2020 and updated annually (Statistics Netherlands 2025, Batenburg et al., 2024). This framework comprises multiple indicators, including HEQ (High Energy Quote), LIHE (Low Income High Expenditure), and the LEQ-based indicator LILEQ (Low Income Low Energy Quality). A brief explanation follows:

– HEQ is defined as having energy expenditure exceeding 10% of disposable household income. Since this would otherwise include many households with substantial wealth yet low income (such as entrepreneur-retirees), the indicator explicitly excludes households in the top 10% of the wealth distribution. This indicator closely aligns with widely used expenditure-share indicators, such as the 10% or 2 M indicators (Al Kez et al., 2024). Disadvantages were discussed in Sect. 1.
– LIHE identifies households with an income below 130% of the poverty line and energy costs exceeding r the average expenditure of a label C dwelling in 2019. This measure is conceptually similar to the Low Income High Cost (LIHC) measure.[6]
– LILEQ classifies households with an income below 130% of the poverty line who live in dwellings among the 50% least energy-efficient in 2019. Note that LEQ was specifically designed to capture housing quality as a structural factor in energy poverty. The LILEQ indicator is closely related to the Low Income Low Energy Efficiency (LILEE) indicator used in the UK8.
– LIHELEQ combines the previous two, identifying all households with an income below 130% of the poverty line who either face high energy costs (LIHE), live in an inefficient home, or both. CBS and TNO use LIHELEQ as the main indicator for 'energy poverty'.

As outlined by Mulder et al. (2023), this monitoring system enables policymakers to identify energy-poor households according to different criteria, tuned for different policy perspectives. For instance, HEQ targets households that are sensitive to increases of energy prices, LIHE detects households that are at risk of falling below the poverty line when energy prices rise, while the use of LILEQ allows detection of households with low incomes who live in poorly insulated homes.

Next to expenditure based indicators on energy poverty, policy and research also use consensual indicators, based on survey data (Al Kez et al., 2024). A commonly used set of such indicators is included in

[6] Both LILEE and LIHC were coined by the Department for Business & Energy and Industrial Strategy (2022).

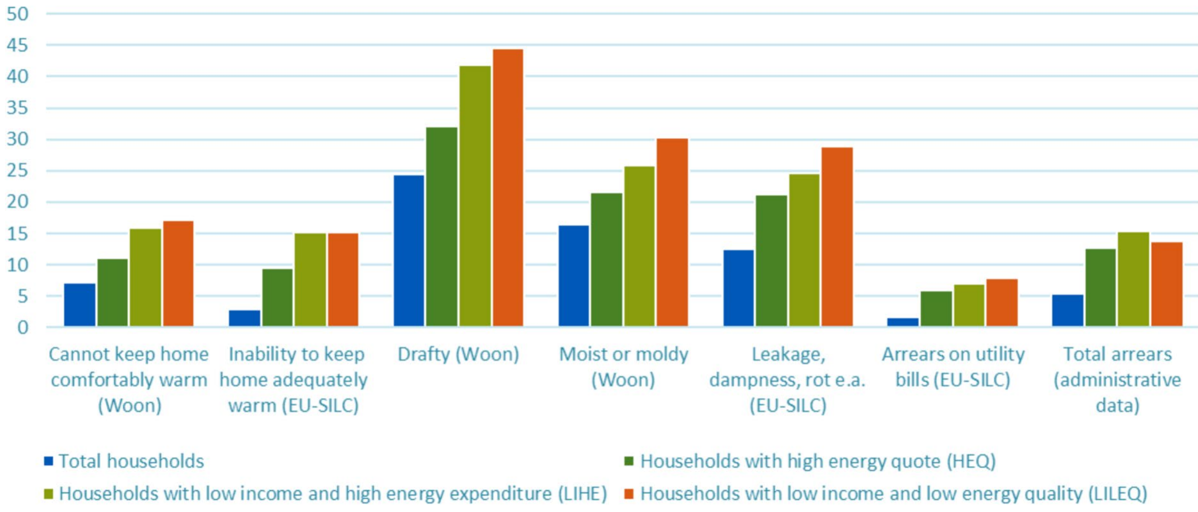

Fig. 2 Households that respond positively to survey questions related to energy poverty, as a share of all households or households that are energy poor according to one of three typical indicators of energy poverty

the European Union Statistics on Income and Living conditions (EU-SILC, 2019). These consist of (1) Living in drafty, damp, or mold-prone homes; (2) Difficulty in maintaining an adequately warm indoor environment during winter; (3) Arrears on energy or utility bills. These indicators may serve as external validation points to assess which administrative measures best identify households experiencing actual energy hardship.

Figure 2 shows how these consensual indicators relate to the three main types of expenditure-based indicators of Table 7: LEQ, LIHE and LILEQ. We use three different sources for these indicators. Survey data from the Dutch housing survey (Woon, 2018), and the European Union Statistics on Income and Living conditions (EU-SILC, 2019), complemented by administrative data on general arrears. The figure shows the percentage of positive answers to the questions as a percentage of all responses, aggregated at the address level.

For these indicators the share of positive answers in the general population (left bar of each group) is relatively low, ranging from 1% with arrears on utility bills, to 25% with a drafty home. Yet, for each of the groups with energy poverty, the responses are markedly higher. This indicates that indeed groups that are labeled as energy poor based on administrative data are more likely to actually undergo energy hardship. Furthermore, the results show that these effects are highest for LILEQ for 6 out of 7 variables, except for 'Total arrears', which is most common among energy poor according to LIHE (since higher bills are more likely to result in arrears). Thus, LILEQ appears to capture energy poverty relatively well. Meanwhile, households with HEQ, though commonly used, are much less likely to suffer from energy poverty related problems as mentioned in the survey-based indicators, which raises questions about its diagnostic value. These findings highlight the advantage of using model-based structural indicators like LEQ/LILEQ to capture the multidimensional nature of energy poverty more accurately.

Estimating over- or underconsumption of energy

A final application of the model involves comparing estimated versus actual energy expenditure to infer behavioral effects. Underconsumption, the case where actual use falls significantly below expected levels, may indicate coping strategies in response to energy poverty (e.g., keeping the home cold). Overconsumption, by contrast, may signal inefficient behavior, poorly adjusted heating systems, or misaligned comfort preferences.

We infer over- and underconsumption by comparing actual energy expenditure with 'expected' energy consumption. We estimate expected consumption by using our model to estimate energy expenditure based

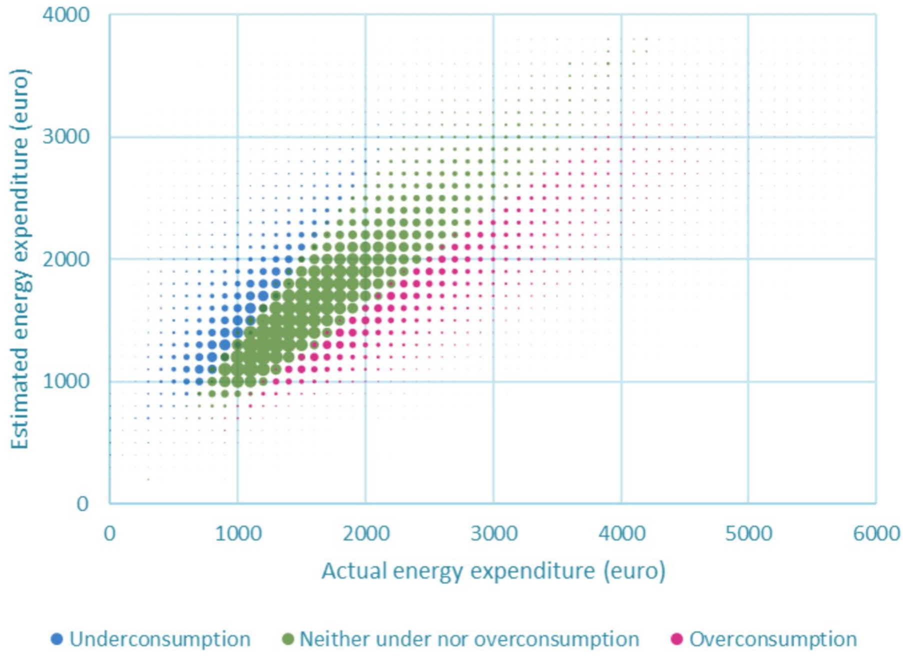

Fig. 3 Relation between actual and expected energy expenditure in relation with our estimates of over- and underconsumption

on both dwelling and household characteristics. With one alteration, which is that we abstract from income, by replacing actual household income by the median income. This is important because income and energy consumption are correlated, and including the income effect could set the bar for energy expenditure too low for low-income households.

We define underconsumption as having less than 75% of expected energy expenditure, and overconsumption as exceeding 125%. This method leads to a 17% share of households for both under- and overconsumption for the Netherlands in 2019. The categories of over- and underconsumption reveal behavioral adaptations, particularly among low-income groups that are more likely to under consume energy due to budget constraints, thus risking a severe form of energy poverty.

Figure 3 shows the relation between actual energy consumption of a household (x-axis) and 'expected' energy consumption at median income levels (y-axis). Expenditures are rounded to the nearest 100-fold. Each combination of rounded expenditures contains a dot which size is proportional to the number of households; the slight overlap between groups is due to rounding.

Underconsumption can have several causes, from frequently being away from home, an environmentally friendly lifestyle, or freeriding on the heating efforts of neighbors in an apartment block. Yet underconsumption can also indicate energy poverty, when people turn down the thermostat to be able to pay the bills. We are especially interested in underconsumption in relation to energy poverty, therefore we focus on the combination of underconsumption and energy poverty, under the assumption that households with a low income and low quality housing are more likely to suffer from underconsumption through energy poverty. We would like to establish this link further in future research, but for now our applications towards underconsumption are still very preliminary and work in progress.

Table 8 shows the proportion of total or energy poor households with over- and underconsumption. While the total population is symmetric, with 17% for both over- and under consumption, we find that overconsumption is very common (46%) among households with a high energy quote (HEQ), still quite large (37%) for low income households with a high energy expenditure, while being smaller than average (14%) for LILEQ. For underconsumption we find that among LILEQ underconsumption is just as common as in the general population. Batenburg et al. (2024) use a similar method to derive under- and overconsumption, and find a slightly larger share of underconsumption among the energy poor (LILEQ) than among all households.

Table 8 The share of over- and underconsumption for all and energy poor households

Dutch energy poverty indicators	Underconsumption	Neither under- nor overconsumption	Overconsumption
Total households for which energy poverty was measured	17	67	17
Households with high energy quote (HEQ)	1	52	46
Households with low income and high energy expenditure (LIHE)	1	62	37
Households with low income and low energy quality (LILEQ)	17	69	14
Households with low income and high energy expenditure and low energy quality (LIHELEQ)	11	62	27

Discussion

Discussion of results

This research was initiated because of two weaknesses found in common indicators in the field of housing and energy. First, the standard indicator for energy efficiency (i.e. EPC labels) fails to adequately identify the homes that are associated with the highest (independent of inhabitants) energy expenditure, and thus the households most at risk (Few et al., 2023; Majcen, 2016; Majcen et al., 2013; Meles et al., 2023; Van Hove et al., 2022). Second, energy poverty indicators that are based on energy expenditure cannot make meaningful assessments of over- and underconsumption (Batenburg et al., 2024). Yet underconsumption and its related health risks are potentially a large problem associated with energy poverty.

Our model addresses both issues. The quality of monitoring energy (in)efficient houses improves dramatically, since houses without an EPC-label can still be accurately classified. This helps both municipalities and the national government in monitoring the effects of their retrofit policies, and adjust them accordingly. Next, our LILEQ indicator improves upon existing measures of energy poverty, and makes it possible to monitor households with underconsumption, who arguably experience the largest negative effects of energy poverty.

Although this research is tailored to Dutch data and policy frameworks, its methodology is generalizable. With appropriate adjustments for local data availability, similar models could be developed in other (EU) countries to support national energy efficiency strategies and harmonized poverty diagnostics. The quality of the indicator depends on the availability and quality of administrative data sets. But even a simple combination of EPC-label with floor area and build year, for which data are widely available throughout countries, would already make a good indicator.

Limitations & further improvements

Our model has two main limitations, which we discuss below, along with potential directions for improvement. First, the model itself is open to improvement: predictions are good but could potentially be better, and its large number of parameters make the model tedious to calculate with a risk of overfitting. Herein lies a fine balance, between adding new variables that increase the prediction quality on the one hand, and on the other hand making the model more compact and efficient by deleting parameters that do not add sufficient predictive power. We suggest these improvements:

– *Incorporation of new administrative variables.* With recent expansions in administrative datasets, new variables have become available. For instance: the capacity of solar panels, ownership of electric vehicles, use of retrofitting subsidies. Integrating these into our model could improve accuracy and provide further insights into household energy behavior.
– *Model simplification.* While the current model includes piecewise linear terms for living area to capture non-linearities, tentative tests suggest that a simpler linear specification may perform equally well when combined with other variables.

Next, the interactions between variables add a lot of dimensions to the model, while the increase in predictive power (R^2) appears limited. Streamlining the model could improve computational efficiency and interpretability.

Second is a more fundamental issue. Our model depends entirely on administrative data. As a result, improvements that are not formally registered—such as insulation upgrades without a new EPC or behavioral changes (e.g., reduced heating use)—are not captured. For instance, after the soaring energy prices in 2022, many households changed their standard thermostat settings or reduced heating in their bedrooms. If sustained**, t**hese behavioral changes would lower the inherent energy expenditure for most dwellings. Yet they are not incorporated in our present model and indicator. Accordingly, we suggest the following avenues for further research and refinement:

- Add survey or smart-meter data to incorporate behavioral aspects of energy use. If we have reliable information about behavioral changes, we can use these to distinguish non-observed retrofits from behavior change. This could be done through a panel survey, or by using smart meter data. CBS is currently evaluating a panel survey for heating and cooling behavior, which could supply the necessary information. Smart meter data could be employed to distinguish between different kinds of energy use (e.g. heating, warm water, EV's) and thus help to better distinguish energy use that is relation to the building form energy use that is more related to its inhabitants.
- Remote sensing of inside temperatures would, when combined with information on the energy use for heating, make it possible to make fine calculations on the isolation of a building. Yet cost of data collection and sensitivity to privacy may make this difficult.
- We can incorporate structural and widespread changes in energy behavior (such as a lower default thermostat level as an aftermath of the energy crises of 2022), by re-fitting the model with energy expenditure data in a new year. Care must be taken though, to correctly separate aggregate or average behavioral effects from housing quality. CBS will implement a shift of base year from 2019 to 2024 in the first half of 2025.

Conclusion

This paper sets out to develop and validate a model for estimating housing energy efficiency that overcomes the limitations of conventional Energy Performance Certificates (EPCs). Our analyses demonstrate that our model captures the underlying energy performance of homes more reliably than EPCs, which are often outdated, missing, or only weakly correlated with actual energy use. Through the integration of structural dwelling characteristics and socio-demographic information, the model enables an estimation of energy needs which abstracts from individual behavior. Thus, our model allows estimation of the inherent energy cost of housing, or housing energy efficiency, more accurately than EPC-labels, while also including dwellings without an EPC label.

A key contribution of this work is the application of our model to energy poverty analysis. We offer an alternative to commonly used indicators for identifying energy poverty that are based on energy expenditure. The Low Income Low Energy Quality (LILEQ) indicator, derived from this model, has been incorporated into the national energy poverty monitoring framework established by CBS and TNO. It complements other indicators such as LIHE (Low Income High Expenditure) and HEQ (High Energy Quote), and has proven especially useful for capturing structural vulnerabilities in the housing stock. As shown in our validation, which uses survey-based indicators (Sect. 4.2, Fig. 1), the LEQ-based indicator (LILEQ) aligns most closely with self-reported problems related to housing quality and thermal comfort. This reinforces its value as a diagnostic tool in social and housing policy.

Another contribution lies in the detection of energy over- or underconsumption. Using model-estimated energy costs as a reference, we identify households with a structural mismatch between income and energy needs; a group that traditional, expenditure-based indicators, often overlook. This approach enables the detection of underconsumption, a condition in which low-income households use less energy than required for adequate comfort due to financial constraints. Such households may not appear energy-poor under conventional metrics but face real material deprivation. By providing a way to identify this condition, our model addresses an important blind spot in current monitoring practices.

In conclusion, this paper provides both a methodological contribution and a practical tool for advancing energy transition goals in the residential sector. By leveraging administrative data to estimate housing energy efficiency more accurately, we offer a foundation for more targeted, equitable, and evidence-based energy policy. As the energy transition accelerates and policy interventions become increasingly urgent, tools like the LEQ model will play a critical role in ensuring that no household is left behind.

Acknowledgements We like to thank Manon Middelkoop, Ruud van den Wijngaart, Anika Batenburg, Charlotte Brand, Francesco Dalla Longa, Krista Keller, Reinder Lok, Lindsey van der Meer and Koen Straver from CBS and TNO at the time of development for their respective contributions to building a framework and comprehensive dataset on measuring energy poverty in The Netherlands over several years. Also thanks to Robert Harmsen for his practical assistance and useful comments, and to an anonymous reviewer for her/his sharp observations and useful advise.

Author contribution L.G. and P.M. revised the final version of the manuscript. R.N. provided the research, text and graphics for Fig. 1 and reviewed previous versions the manuscript. P.M. initiated early versions of the energy poverty incators described in the paper, co-authored title, abstract, literature, discussion and conclusion, while being the main author of the introduction. L.G. co-authored most sections of the paper and was the main author for the methodology and result sections.

Funding No specific funding for this paper.

Data availability We used data hosted by Statistics Netherlands. These data can be accesed by certified research institutions. See: https://www.cbs.nl/en-gb/our-services/customised-services-microdata/microdata-conducting-your-own-research.

Declarations

Clinical trial number Not applicable.

Competing interests The authors declare no competing interests.

References

Al Kez, D., Foley, A., Lowans, C., & Del Rio, D. F. (2024). Energy poverty assessment: Indicators and implications for developing and developed countries. *Energy Conversion and Management, 307*, Article 118324.

Allcott, H., & Greenstone, M. (2017). Measuring the welfare effects of residential energy efficiency programs. *NBER Working Paper No. 23386*.

Amasyali, K., & El-Gohary, N. M. (2018). A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews, 81*, 1192–1205.

Batenburg, A., van der Meer, L., Hopman, B., Dalla Longa, F., Geijtenbeek, L., van Middelkoop, M., Mulder, P., & Wijlhuizen, E. (2024). Energiearmoede in Nederland 2019–2023: Een overzicht van 2019 tot en met 2023 en een verdieping op onderconsumptie (TNO 2024 R10801). TNO. https://publications.tno.nl/publication/34642676/7qBbU9FJ/TNO-2024-R10801.pdf

Beijnum, B. van, van den Wijngaart, R., Luteijn, G., van der Molen, F., Geijtenbeek, L., Rovers, V., & Tigchelaar, C. (2023). Referentieverbruik warmte woningen. Achtergrondrapport, Den Haag: Planbureau voor de Leefomgeving. https://www.pbl.nl/uploads/default/downloads/pbl-2023-referentieverbruik-warmte-woningen-achtergrondrapport_5168.pdf

CLO. (2025). Energielabels van woningen, 2010 t/m 2024 (indicator 0556, versie 11, 18 maart 2025), www.clo.nl. Centraal Bureau voor de Statistiek (CBS), Den Haag; PBL Planbureau voor de Leefomgeving, Den Haag; RIVM Rijksinstituut voor Volksgezondheid en Milieu, Bilthoven; en Wageningen University and Research, Wageningen.

Cui, X., Lee, M., Koo, C., & Hong, T. (2024). Energy consumption prediction and household feature analysis for different residential building types using machine learning and SHAP. *Energy and Buildings, 309*, Article 113997.

Department for Business & Energy and Industrial Strategy. Fuel poverty methodology handbook: Low Income Low Energy Efficiency (LILEE). London, United Kingdom; 2022. <https://www.gov.uk/government/publications/fuel-poverty-statistics-methodology-handbook>.

European Commission. (2025). "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)". European Data Portalaccess-date=26 May 2025. 24 April 2024.

Few, J., Manouseli, D., McKenna, E., Pullinger, M., Zapata-Webborn, E., & Oreszczyn, T. (2023). The over-prediction of energy use by EPCs in Great Britain. *Energy and Buildings, 288*, Article 113024.

Fowlie, M., Greenstone, M., & Wolfram, C. (2018). Do energy efficiency investments deliver? *The Quarterly Journal of Economics, 133*(3), 1597–1644.

Harputlugil, T., & de Wilde, P. (2021). The interaction between humans and buildings for energy efficiency. *Energy Research & Social Science, 71*, Article 101828.

Ishwaran, H. (2007). Variable importance in binary regression trees and forests.

Karpinska, L., & Smiech, S. (2020). Conceptualising housing costs: The hidden face of energy poverty in Poland. *Energy Policy, 147*, Article 111819.

Karpinska, L., & Smiech, S. (2023). Multiple faces of poverty: Exploring housing-costs-induced energy poverty in Central and Eastern Europe. *Energy Research & Social Science, 105*, 103273.A.

Majcen, D. (2016). Predicting energy consumption and savings in the housing stock: A performance gap analysis in the Netherlands. *A+be | Architecture and the Built Environment, 6*(4), 1–224.

Majcen, D., Itard, L., & Visscher, H. (2013). Theoretical vs. actual energy consumption of labelled dwellings in the Netherlands: Discrepancies and policy implications. *Energy Policy, 54*(2013), 125–136.

Meles, T. H., Farrell, N., & Curtis, J. (2023). How well do building energy performance certificates predict heat loss? *Energy Efficiency, 16*, 74. https://doi.org/10.1007/s12053-023-10146-0

Mulder, P., Dalla Longa, F., & Straver, K. (2023). Energy poverty in the Netherlands at the national and local level: A multi-dimensional spatial analysis. *Energy Research & Social Science, 96*, Article 102892. https://doi.org/10.1016/j.erss.2022.102892

Peñasco, C., & Díaz Anadón, L. (2023). Assessing the effectiveness of energy efficiency measures in the residential sector. *Energy Economics, 117*, Article 106435.

Rademaekers, K., Yearwood, J., Ferreira, A., Pye, S. T., Hamilton, I., Agnolucci, P., Grover, D., Karásek, J. & Anisimova, N. (2016). *Selecting Indicators to Measure Energy Poverty*; European Commission, DG Energy: Brussels, Belgium.

Semple, T., Rodrigues, L., Harvey, J., et al. (2024). An empirical critique of the low income low energy efficiency approach. *Energy Policy, 186*, Article 114014.

Statistics Netherlands (2024). Methodedocument monitor energiearmoede 2019–2022. https://www.cbs.nl/nl-nl/longread/aanvullende-statistische-diensten/2024/methodedocument-monitor-energiearmoede-2019-2022

Van Hove, M. Y. C., et al. (2022). Large-scale statistical analysis and modelling of real and regulatory energy use. *Building Research & Information, 51*(2), 203–222.

Van den Wijngaart, R., & Van Polen, S. (2020). Bepaling energiebesparing door isolatie van woningen in de Startanalyse 2020, PBL-publicatienummer 4284.