AU-26-C034 https://doi.org/10.63044/s26evr34

Mitigating Physically Implausible Energy Performance Forecasting in Residential Buildings Using Logarithmic Transformation-Based Machine Learning Analysis: A Case Study Using RECS and ResStock Datasets

Aditya Ramnarayan
Student Member ASHRAE

Fatih Evren, PhD
Full Member ASHRAE

Patricia Gunderson, PE
Full Member ASHRAE

ABSTRACT

Data-driven analysis of residential building energy consumption is essential for understanding how homes perform, improving model accuracy, and guiding large-scale energy efficiency decisions. In this study, we apply several machine learning methods, including CatBoost, LightGBM, Random Forest, XGBoost, and a Neural Network, to predict total energy consumption as well as space heating and cooling loads across two diverse datasets – the Energy Information Administration's Residential Energy Consumption Survey (RECS) of nearly 18,500 real households, and the Department of Energy's Residential Stock (ResStock) database, a compendium of nearly 550,000 statistically-developed building simulations. To address the significant skewness often observed in energy consumption distributions, we applied a logarithmic transformation to all energy performance metrics, including total energy consumption and space heating and cooling energy use. In its raw form, residential energy performance (whether total usage, space heating, or space cooling) often exhibits a strongly right-skewed distribution: a small fraction of homes accounts for disproportionately large energy demands. Such heavy-tailed data can bias statistical models, destabilize variance estimates, and undermine predictive accuracy. By applying a natural logarithm transformation (with an additive constant, e, to handle low or zero values), we compress extremely high values and stretch values near zero, yielding a distribution that is more symmetric and statistically tractable. After fitting our models in the transformed space, we back-transform predictions to the original energy scale for interpretation. This log-transformation approach enhances the statistical robustness of our models while preserving their physical interpretability, effectively mitigating physically implausible negative forecasts and improving predictive consistency across highly skewed energy variables. These data-driven models serve as fast, scalable predictors of residential energy performance for individual homes. By basing predictions on actual (RECS) and/or modeled consumption (ResStock), they improve the accuracy of national energy demand estimates and enable identification of potential energy efficiency upgrades. This approach enhances traditional tools with greater empirical predictive accuracy, particularly for skewed end uses such as cooling. Using the logarithmic transformation approach resulted in significant model performance improvement for cooling energy consumption forecasts, with R^2 values rising from 0.867 to 0.928 and decreases in MAE and RMSE values

Aditya Ramnarayan, is a graduate intern at Pacific Northwest National Laboratory and a PhD Student at the University of Maryland, **Fatih Evren** is a senior researcher at Pacific Northwest National Laboratory, WA, **Patricia Gunderson** is a senior researcher at Pacific Northwest National Laboratory, WA.

from 5943 to 4517 and 9441 to 6911, respectively, thereby underscoring the value of this approach in eliminating physically implausible values in energy performance forecasting.

INTRODUCTION

In 2024, the residential sector in the U.S. consumed approximately 18.4 quadrillion British thermal units (Btu) of energy, representing approximately 19.6% of the total primary energy consumption in the United States (*U.S. Energy Information Administration (EIA)*, n.d.). The residential sector is the largest consumer of electricity, consuming 1.55 trillion kWh and accounting for 38% of the nation's overall electricity consumption (*U.S. Energy Information Administration (EIA)*, n.d.). The number of housing units in the U.S. continues to increase, with the most recent count indicating approximately 144 million housing units (Bawaneh et al., 2024). As the housing stock expands, demand for residential energy and the associated costs are expected to rise, underscoring the need for more precise forecasting tools. Furthermore, in 2024, the average price of electricity for the residential sector in the US was 16.48 cents/kWh, rising faster since 2020 than in the previous decade (*U.S. Electricity Retail Price by Sector 2024*). As prices continue to rise and demand for energy grows, accurate and inexpensive electricity usage prediction supports proactive load management by utilities and empowers building owners to invest in energy upgrades and take control of their utility bills, promoting economic stability for all Americans.

Existing methods for estimating building energy consumption primarily involve simulation (Higgins et al., 2024; Liu et al., 2022; Sun et al., 2022) and data-driven models (Amasyali & El-Gohary, 2018; Ramnarayan et al., 2025; Zargarzadeh et al., 2024). Building energy simulation software calculates the energy required to meet specific performance standards, accounting for factors such as building construction, weather, occupancy, and air leakage (Coakley et al., 2014). Although building energy simulations produce highly accurate results, their widespread adoption faces several challenges. These challenges primarily stem from uncertainties in the input data, substantial time and cost requirements, and the potential for significant prediction errors resulting from model complexity and insufficient reliable data (Gerrish et al., 2017). In contrast, data-driven (black-box) models use monitored data from building metadata to create mathematical or statistical models capable of accurately predicting system behavior (Di Stefano et al., 2023). Data-driven building energy consumption prediction modeling relies on historical data rather than detailed information about the simulated building (Esen et al., 2017). This study uses the Residential Energy Consumption Survey (RECS) and the ResStock datasets to build and evaluate machine learning models that forecast total energy consumption, space heating, and space cooling. The RECS dataset comprises approximately 18,500 households, which are statistically representative of the 123.5 million residences in the United States. On the other hand, ResStock, developed by the National Lab of the Rockies (NLR), models and simulates energy consumption and savings for residential buildings nationwide. It creates a dataset comprising of ~550,000 data points statistically representative of the country's housing stock.

Owing to the large volume of data across both datasets, machine learning models are well-suited to this task owing to their innate capabilities in handling complex and nuanced data as well as their predictive capabilities. This study compares the performance of the two datasets as benchmarking approaches to develop the best-performing methods. The motivation behind this study and methodology is to develop generalizable, robust, and accurate ML models capable of reliable energy performance forecasting. In modeling residential energy consumption and heating/cooling loads, the target variable often spans several orders of magnitude. Using simple regression on the raw scale may yield negative predictions, even though these quantities cannot be negative. To address the constraint of a lower bound (zero) and the common right-skewed distribution of energy usage, we apply a logarithmic transformation to the target variable (F. Zhang et al., 2016; Y. Zhang et al., 2015a). In this study, we examine the improvements in regression forecasts with the scope of achieving physically plausible forecasts.

METHODOLOGY

This study leverages two nationally representative datasets: the Residential Energy Consumption Survey (RECS) and the ResStock dataset, aiming to develop machine learning models for forecasting energy performance. The methodology encompasses a comprehensive framework that integrates data preprocessing, manual feature selection, model training and testing, hyperparameter tuning, and meta-stacking. To thoroughly evaluate the performance of the proposed model, the data is split into two subsets: 80% for training and 20% for testing. During the model training phase, hyperparameter tuning and model validation were conducted using k-fold cross-validation strictly on the 80% training dataset. The remaining 20% hold-out test set was kept entirely unseen during tuning and was used solely for final model evaluation. Five machine learning models, namely, CatBoost, LightGBM, Random Forest, XGBoost, and a Neural Network are trained on these datasets, and their performance is assessed using metrics such as R-squared (R^2), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE), to identify the best-performing model. These five machine learning algorithms were selected for their effectiveness in managing large-scale tabular data. Tree-based ensemble methods—Random Forest, XGBoost, LightGBM, and CatBoost—show strong resilience to non-linear relationships, outliers, and mixed data types without requiring extensive feature scaling. A Neural Network is also included to serve as a robust deep learning baseline, effectively capturing complex interactions between building characteristics and energy performance. The top-performing model is fine-tuned to enhance its performance. After this, a meta-stacking approach is implemented to boost model efficacy (Naz et al., 2025). This strategy aims to create a robust and interpretable forecasting pipeline, thereby enhancing the reliability and transparency of the predictive outcomes.

Following model testing, the regression forecasts are scrutinized by visualizing an Actual vs. Predicted scatter plot. Next, the number of negative forecasts is computed for the three forecasting scenarios – total energy consumption, space heating, and space cooling. The logarithmic transformation methodology is then implemented to mitigate the presence of negative forecasts. To address both the lower-bound constraint (zero) and the right-skewed distribution typical of energy use, we apply a logarithmic transformation to the target variable. Specifically, we define $y^{*} = \ln(y + c)$ where y^* represents the original consumption metric (in kWh or equivalent), and c is a small positive constant added to shift zero-valued observations into the positive domain (since *ln (0)* is undefined). This transformation offers three key advantages (F. Zhang et al., 2016): (1) it ensures that the exponentiated model output, $\hat{y} = \exp(\hat{y}^{*}) - c$, remains strictly non-negative, thus preventing negative forecasts; (2) it compresses the right tail of the distribution, which helps mitigate heteroscedasticity and reduces the impact of extreme high-consumption outliers on the loss function; and (3) it linearizes multiplicative relationships between predictors and consumption, enhancing the fit of linear (or tree-based) algorithms that operate under additive error assumptions. Caution is necessary when adding a constant, as it modifies the effective scale of the data. This adjustment requires careful back-transformation to mitigate biases from Jensen's inequality and retransformation errors (Kolter & Ferreira, 2011). Additionally, the choice of this constant can subtly impact predictions, especially at the lower end of the distribution. We trained all models using this log-transformed data and evaluated the outputs on the original scale by exponentiating and subtracting the constant. This methodology offers a systematic approach to prevent negative forecasts while preserving interpretability and enhancing model stability in the context of residential energy prediction.

RESULTS

Pre-logarithmic transformation

Once the models have been trained, validated, and tested, model performance is evaluated by inspecting the accuracy metrics results. Table 1 highlights the performance of the models. CatBoost ML provided the most accurate forecasts, shown in bold type. Figure 1 shows the actual vs. predicted regression plots (before logarithmic transformation) for the RECS dataset using the CatBoost model.

Table 1: Performance metrics pre-logarithmic transformation for the RECS dataset

Target Variable	ML Models	R^2	MAE	RMSE
Total Energy Consumption	LightGBM	0.714	17,795	24,294
	Random Forest	0.691	18,506	25,242
	CatBoost	0.726	17,338	23,788
	XGBoost	0.695	18,444	25,090
	Neural Network	0.718	17,750	24,134
Space Heating	LightGBM	0.728	11,115	16,472
	Random Forest	0.707	11,453	17,097
	CatBoost	0.732	11,001	16,348
	XGBoost	0.697	11,736	17,364
	Neural Network	0.728	11,127	16,459
Space Cooling	LightGBM	0.641	2,369	3,512
	Random Forest	0.631	2,387	3,565
	CatBoost	0.653	2,346	3,455
	XGBoost	0.609	2,487	3,669
	Neural Network	0.628	2,481	3,576

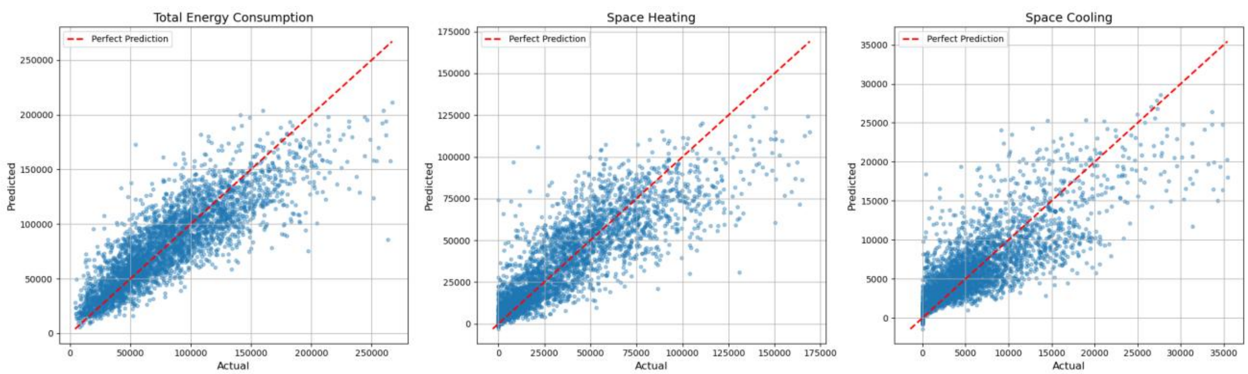

Figure 1: Actual vs Predicted Regression Plots pre-logarithmic transformation for the RECS dataset

While comparing the forecasts for total energy consumption with existing literature (Cui et al., 2024, 2025; Korsavi et al., 2025) for the RECS dataset, Table 1 demonstrates that our accuracy metrics outperform existing studies [Our R^2: 0.726, Cui et al. R^2: 0.690, Korsavi et al. R^2: 0.712]. However, upon scrutinizing the forecasts in Figure 1, we see that those for Space Heating and Cooling are negative. Upon investigation, we calculated the number of zero values in the forecasts, as shown in Table 2.

Table 2: Number of negative forecasts for the CatBoost ML model predictions of the RECS dataset

Parameter	Number of negative counts	% Compared to the test dataset
Total Energy Consumption	0	0
Space Heating	38	1.06
Space Cooling	189	5.30

Similarly, Table 3 highlights the models' performance, and Figure 2 shows the actual vs. predicted regression plots (before logarithmic transformation) for the ResStock dataset using all tested ML models. Again, CatBoost ML provided the most accurate forecasts, shown in bold type

Table 3: Performance metrics pre-logarithmic transformation for the ResStock dataset

Target Variable	ML Models	R^2	MAE	RMSE
Total Energy Consumption	LightGBM	0.814	5,178	7,285
	Random Forest	0.878	4,119	5,903
	CatBoost	0.913	3,547	4,987
	XGBoost	0.892	3,970	5,548
	Neural Network	0.896	3,857	5,437
Space Heating	LightGBM	0.829	11,264	16,698

	Random Forest	0.892	8,306	13,309
	CatBoost	0.927	6,952	10,960
	XGBoost	0.908	7,970	12,244
	Neural Network	0.901	8,241	12,701
Space Cooling	LightGBM	0.763	8,452	12,582
	Random Forest	0.84	6,492	10,357
	CatBoost	0.867	5,944	9,441
	XGBoost	0.842	6,668	10,278
	Neural Network	0.854	6,236	9,865

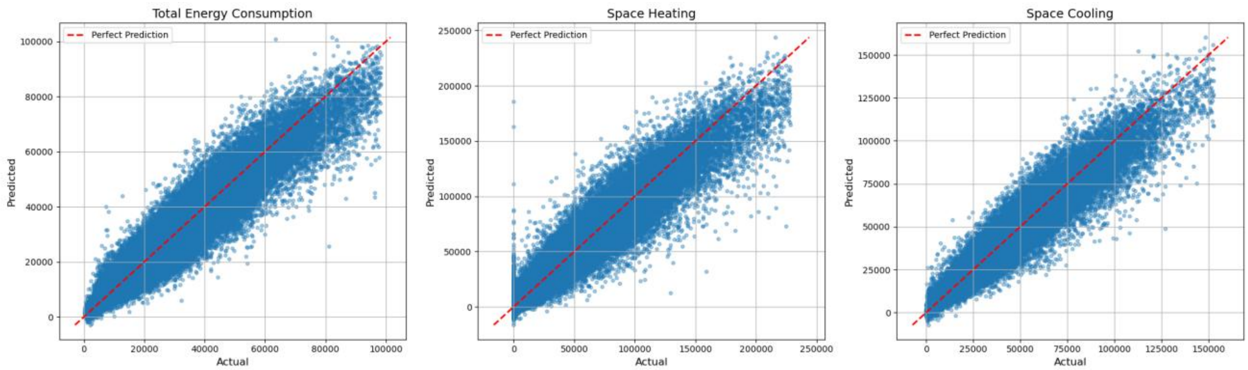

Figure 2 Actual vs Predicted Regression Plots pre-logarithmic transformation for the ResStock dataset

Similar to the RECS forecasts, we found that the developed ML models for the ResStock dataset outperformed existing literature (Asamoah & Shittu, 2025), both for energy consumption as well as space heating and cooling. However, the same problem with negative forecasts persisted with the forecasts for the ResStock dataset as well. Table 4 quantifies the number and percentage of negative forecasts.

Table 4: Number of negative forecasts for the CatBoost ML model predictions of the ResStock dataset

Parameter	Number of negative counts	% of the test dataset
Total Energy Consumption	41	0.04
Space Heating	4877	4.59
Space Cooling	191	0.23

Although the percentages of negative predictions are relatively low, it still highlights an important concern in adopting black-box models for energy modeling and forecasting purposes – lack of interpretability and inherent constraints can lead to implausible predictions, such as physically impossible negative values or results that violate fundamental laws of physics, undermining trust and practical application (Rudin et al., 2022).

Post-logarithmic transformation

By applying logarithmic modeling to the energy data and then back-transforming the predictions with the exponential function, it is mathematically ensured that all resulting forecasts will be non-negative. This approach effectively mitigates the risk of negative value predictions and enhances the robustness of the forecasting process (Y. Zhang et al., 2015b). To improve model stability and address the positive skew in the target variable in the RECS and ResStock datasets, we applied a logarithmic transformation before regression analysis. We added a small constant ($c = 1 \times 10^{-3}$) to all target values to handle the logarithm's undefined nature at zero. This transformation reduced heteroscedasticity, lessened the effect of extreme values, and emphasized multiplicative relationships. After modeling, we exponentiated the predicted values to revert them to their original scale. This reduced the number of negative values in the predictions of total energy consumption and space heating and cooling to zero across both datasets. There were no negative values in the CatBoost forecasts for total energy consumption for RECS (Table 2), so this target variable didn't undergo log transformation. Using the ResStock dataset CatBoost initially produced negative values for all three target variables (Table 4). Tables 5 and 6 show the performance metrics post-logarithmic transformation of the RECS and ResStock datasets, respectively. As with the pre-logarithmic transformation, CatBoost (bold type) was the most accurate ML for the post-logarithmic transformation.

Table 5: Performance metrics post-logarithmic transformation for the RECS dataset

Target Variable	ML Models	R^2	MAE	RMSE
Total Energy Consumption	LightGBM	0.714	17,795	24,294
	Random Forest	0.691	18,506	25,242
	CatBoost	0.726	17,338	23,788
	XGBoost	0.695	18,444	25,090
	Neural Network	0.718	17,750	24,133
Space Heating	LightGBM	0.69	11,711	17,547
	Random Forest	0.683	11,877	17,758
	CatBoost	0.706	11,420	17,100
	XGBoost	0.669	12,056	18,134
	Neural Network	0.68	11,855	17,839
Space Cooling	LightGBM	0.554	2,684	3,913
	Random Forest	0.517	2,777	4,073
	CatBoost	0.562	2,646	3,877
	XGBoost	0.521	2,834	4,055
	Neural Network	0.541	2,700	3,969

For the RECS dataset, the log transformation reduced the influence of extreme percentage-based errors but, in some cases, degraded performance metrics evaluated on the original scale (R^2, MAE, RMSE). After exponentiating predictions back to the original scale, small deviations in log space can become amplified for larger target values, leading to increased absolute errors, which for space cooling ranged from approximately 10% to 15% for all metrics and all MLMs, though slightly less for Neural Network. For space heating the reduction in accuracy following log transformation was typically less than 8% for all combinations. The results suggest that the original model exhibited instability driven by very small true values, and the logarithmic transformation smoothed the relative scale of the data. Although the transformation improved the stability of relative error behavior, reconversion to the original metric space introduced some distortion, resulting in a slight decline in absolute performance metrics. Forecasts for the ResStock dataset were broadly similar pre- and post-transformation for total energy consumption and space heating. Reduced accuracy for total energy was less than 5.5% for all metrics and for all MLMs except Neural Network, for which the transformation produced slightly better accuracy for all three metrics. For space heating changes to accuracy metrics post-transformation were mixed but typically less than 3% in either direction (Table 6), while all three metrics for space cooling improved substantially following transformation. For total energy, the log transformation resulted in only minor changes to absolute metrics, indicating comparable predictive accuracy on the original scale. The effect was more pronounced for heating and cooling, where extreme error behavior associated with zero or near-zero target values was substantially mitigated after transformation. Overall, the logarithmic transformation regularized model behavior by reducing small-denominator instability and yielding more coherent and consistent error patterns across all three forecasting tasks. Figures 3 and 4 present the actual versus predicted regression plots for CatBoost after logarithmic transformation for total energy consumption, space heating, and space cooling for the RECS and ResStock datasets, respectively.

Table 6: Performance metrics post-logarithmic transformation for the ResStock dataset

Target Variable	ML Models	R^2	MAE	RMSE
Total Energy Consumption	LightGBM	0.795	5,397	7462
	Random Forest	0.865	4,346	6052
	CatBoost	0.903	3,672	5133
	XGBoost	0.884	4,096	5620
	Neural Network	0.898	3,815	5252
Space Heating	LightGBM	0.832	11,477	16837

	Random Forest	0.895	8,566	13326
	CatBoost	0.926	7,037	11160
	XGBoost	0.916	7,961	11905
	Neural Network	0.922	7,242	11485
Space Cooling	LightGBM	0.845	7,088	10160
	Random Forest	0.899	5,455	8212
	CatBoost	0.928	4,517	6911
	XGBoost	0.911	5,236	7675
	Neural Network	0.921	4,843	7253

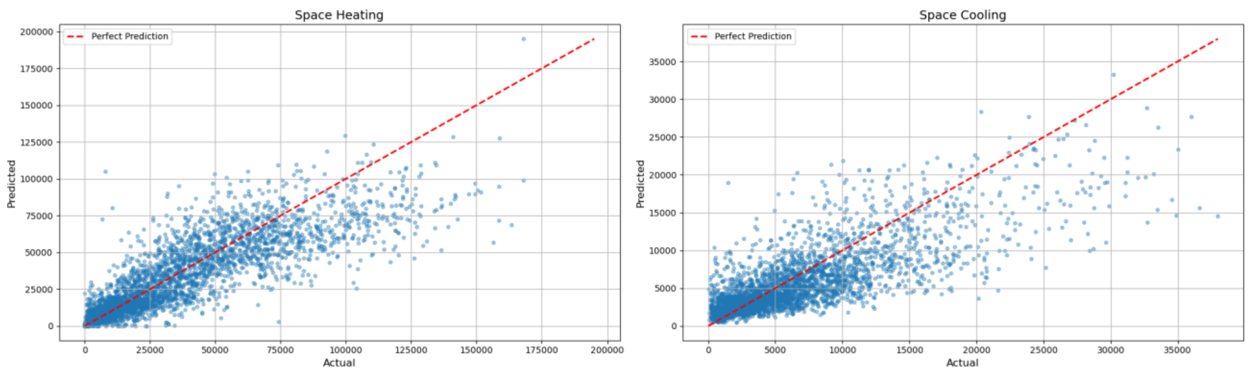

Figure 3 Actual vs Predicted Regression Plots post-logarithmic transformation for the RECS dataset

As shown in Figure 3, the negative predictions no longer occur, and all forecast values are physically plausible owing to the logarithmic transformation methodology. Similarly, for the ResStock dataset, negative values no longer appear in Figure 4, thereby improving accuracy in space cooling forecasting and increasing confidence in this approach.

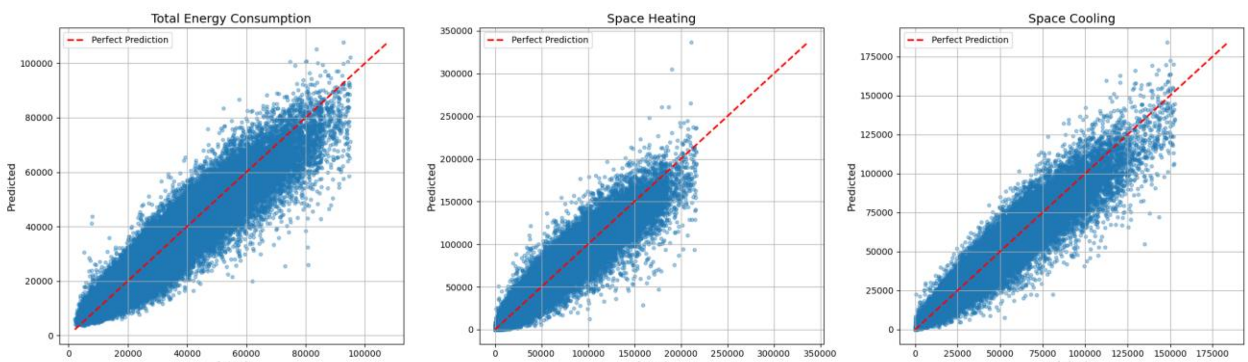

Figure 4 Actual vs Predicted Regression Plots post-logarithmic transformation for the ResStock dataset

DISCUSSION

Standard machine learning (ML) energy forecasting models can produce physically implausible values, such as negative energy consumption or negative heating and cooling loads, due to inherent limitations which do not follow thermodynamic principles. Although these models excel at identifying intricate statistical patterns in data, their reliance on purely mathematical data-driven approaches can lead to predictions that contradict the laws of physics (Quarteroni et al., 2025). This study applied a transformation by using a log function on the target variable, like energy consumption or heating/cooling loads, before training various machine learning models. Logarithmic transformation stabilizes variance, helps correct skewed data distributions often present in energy datasets, eliminates nonsensical negative values, and brings them onto a comparable scale. Sensible heat transfer follows a multiplicative relationship; the product of the heat transfer coefficient (U), area (A), and temperature differential (ΔT). Standard machine learning algorithms use additive operations, which can limit their effectiveness. A logarithmic transformation concerts this to an additive relationship: (ln (U x A x ΔT) becomes ln(U) + ln(A) + ln(ΔT), allowing tree-based and linear models to better capture building energy mechanics, and prevents models from yielding unrealistic negative energy values. Comparing both datasets and all three target variables, CatBoost emerged as the superior mode, possibly due to its inherent capability to process data natively, reducing the need for extensive manual preprocessing, typically required by conventional gradient boosting frameworks (Hancock & Khoshgoftaar, 2020).

In the initial applications of this technique, a constant value of 10^{-3} was selected to eliminate the negative values associated with energy performance forecasting. When this constant was applied to the RECS dataset, the accuracy metrics for total energy consumption remained unchanged because the forecasts contained no negative values. However, for space heating and cooling, the R^2 values decreased, whereas the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) remained unchanged. For high-variance variables like space heating and space cooling, the log-transformation introduced a bias-variance trade-off that was detrimental to the model's explanatory power. The decrease in R^2 suggests that the model had difficulty capturing the data's dynamic range, likely because the peak values were compressed on the logarithmic scale. This clearly illustrates that the method is sensitive to the target distribution and may not be appropriate for intermittent loads compared to the baseline. For the ResStock dataset, the R^2 values for total energy consumption and space heating were nearly unchanged, whereas those for space cooling increased. The MAE and RMSE values showed only slight improvement in forecasting total energy consumption and space heating, while space cooling showed a substantial decrease (over 1,000 kWh). Furthermore, the logarithmic transformation methodology helped eliminate 227 negative forecasts (1.23% of the dataset) from the RECS dataset and 5,109 (0.93% of the dataset) from the ResStock dataset. The value of models that provide better predictive metrics but also blindly produce impossible predictive results is somewhat illusory.

Traditional regression models typically aim to minimize the Mean Squared Error (MSE), which leads to metrics such as R^2 and RMSE becoming key targets for optimization (Dumre et al., 2024). However, these metrics focus primarily on overall statistical fit and penalize significant deviations without considering whether the predicted values adhere to fundamental physical laws. In the context of energy performance forecasting, employing a logarithmic transformation involves a conscious trade-off: conventional accuracy metrics such as R^2, MAE, and RMSE may be compromised in favor of ensuring the plausibility of the results. This strategic choice emphasizes physically plausible outcomes, such as non-negative energy consumption, space heating, and space cooling, rather than solely striving for the lowest statistical error on the original scale. This is an important aspect and step to consider when developing ML models, since these forecasts have implications for energy-efficiency retrofit decisions, which are driven by financial incentives, shareholder investment, and utility allowances for income-qualified housing.

CONCLUSIONS

Machine learning models CatBoost, LightGBM, Random Forest, XGBoost, and a Neural Network were developed for two open-source, nationally representative datasets, RECs and ResStock. These models aimed to forecast energy consumption, space heating, and space cooling by utilizing a combination of data-driven and statistical modeling techniques. The methodology included data preprocessing, outlier removal, hyperparameter tuning, and meta-stacking. During the forecasting process, several physically implausible negative forecasts were identified in both datasets. To rectify these anomalies, a logarithmic transformation method was employed. This involved adding a very small constant value to eliminate negative forecasts when converting the variables back from the logarithmic scale. This approach effectively eliminates physically implausible negative consumption values that purely data-driven models might produce when attempting to minimize RMSE on skewed datasets, with only minor changes in the accuracy of R^2 and MAE metrics. Both before and after the logarithmic transformation, CatBoost ML provided the most accurate results.

ACKNOWLEDGMENTS

Author Contributions: AR conducted the data analysis and drafted the original version of the manuscript. FE supervised the study, and both FE and PG provided critical revisions on model development and contributed to the interpretation of the results. All authors reviewed and approved the final version of the manuscript. The authors would like to thank Sam Rosenberg for his valuable input during the model development.

NOMENCLATURE

y= Original energy consumption (kWh or BTU)

y^*= Log-transformed energy consumption

$\hat{y}$= Predicted energy consumption (original scale)

$\hat{y}^*$= Predicted energy consumption (log scale)

c= Small positive constant added prior to logarithmic transformation

R^2= Coefficient of determination

MAE= Mean Absolute Error

$RMSE$= Root Mean Squared Error

MSE= Mean Squared Error

n= Number of observations

REFERENCES

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. https://doi.org/10.1016/j.rser.2017.04.095

Asamoah, P. B., & Shittu, E. (2025). Evaluating the performance of machine learning models for energy load prediction in residential HVAC systems. *Energy and Buildings*, *334*, 115517. https://doi.org/10.1016/j.enbuild.2025.115517

Bawaneh, K., Das, S., & Rasheduzzaman, Md. (2024). Energy Consumption Analysis and Characterization of the Residential Sector in the US towards Sustainable Development. *Energies*, *17*(11), 2789. https://doi.org/10.3390/en17112789

Coakley, D., Raftery, P., & Keane, M. (2014). A review of methods to match building energy simulation models to measured data. *Renewable and Sustainable Energy Reviews*, *37*, 123–141. https://doi.org/10.1016/j.rser.2014.05.007

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: Toward energy-efficient buildings. *Energy and Buildings*, *309*, 113997. https://doi.org/10.1016/j.enbuild.2024.113997

Cui, X., Lee, M., Uddin, M. N., Zhang, X., & Zakka, V. G. (2025). Analyzing different household energy use patterns using clustering and machine learning. *Renewable and Sustainable Energy Reviews*, *212*, 115335. https://doi.org/10.1016/j.rser.2025.115335

Di Stefano, A. G., Ruta, M., & Masera, G. (2023). Advanced Digital Tools for Data-Informed and Performance-Driven Design: A Review of Building Energy Consumption Forecasting Models Based on Machine Learning. *Applied Sciences*, *13*(24), 12981. https://doi.org/10.3390/app132412981

Dumre, P., Bhattarai, S., & Shashikala, H. K. (2024). Optimizing Linear Regression Models: A Comparative Study of Error Metrics. *2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS)*, 1856–1861. https://doi.org/10.1109/ICTACS62700.2024.10840719

Esen, H., Esen, M., & Ozsolak, O. (2017). Modelling and experimental performance analysis of solar-assisted ground source heat pump system. *Journal of Experimental & Theoretical Artificial Intelligence*, *29*(1), 1–17. https://doi.org/10.1080/0952813X.2015.1056242

Gerrish, T., Ruikar, K., Cook, M., Johnson, M., Phillip, M., & Lowry, C. (2017). BIM application to building energy performance visualisation and management: Challenges and potential. *Energy and Buildings*, *144*, 218–228. https://doi.org/10.1016/j.enbuild.2017.03.032

Hancock, J. T., & Khoshgoftaar, T. M. (2020). CatBoost for big data: An interdisciplinary review. *Journal of Big Data*, *7*(1), 94. https://doi.org/10.1186/s40537-020-00369-8

Higgins, J., Ramnarayan, A., Family, R., & Ohadi, M. (2024). Analysis of Energy Efficiency Opportunities for a Public Transportation Maintenance Facility—A Case Study. *Energies*, *17*(8), 1907. https://doi.org/10.3390/en17081907

Kolter, J., & Ferreira, J. (2011). A Large-Scale Study on Predicting and Contextualizing Building Energy Usage. *Proceedings of the AAAI Conference on Artificial Intelligence*, *25*(1), 1349–1356. https://doi.org/10.1609/aaai.v25i1.7806

Korsavi, S. S., Azari, R., Iulo, L. D., & Mahdavi, M. (2025). Determinants of U.S. residential energy consumption at national and state levels: Policy implications. *Energy Policy*, *202*, 114594. https://doi.org/10.1016/j.enpol.2025.114594

Liu, X., Tang, H., Ding, Y., & Yan, D. (2022). Investigating the performance of machine learning models combined with different feature selection methods to estimate the energy consumption of buildings. *Energy and Buildings*, *273*, 112408. https://doi.org/10.1016/j.enbuild.2022.112408

Naz, M., Khalid, A., Hameed, A., Taj, R., Mumtaz, W., Alotaibi, F. A., & Alnfiai, M. M. (2025). Meta-Ensemble Learning for Heart Disease Prediction: A Stacking-Based Approach With Explainable AI. *IEEE Access*, *13*, 137271–137290. https://doi.org/10.1109/ACCESS.2025.3588683

Quarteroni, A., Gervasio, P., & Regazzoni, F. (2025). Combining physics-based and data-driven models: Advancing the frontiers of research with scientific machine learning. *Mathematical Models and Methods in Applied Sciences*, *35*(04), 905–1071. https://doi.org/10.1142/S0218202525500125

Ramnarayan, A., De Castro, F., Sarmiento, A., & Ohadi, M. (2025). Machine Learning-Based Carbon Compliance Forecasting and Energy Performance Assessment in Commercial Buildings. *Energies*, *18*(15), 3906. https://doi.org/10.3390/en18153906

Rudin, C., Chen, C., Chen, Z., Huang, H., Semenova, L., & Zhong, C. (2022). Interpretable machine learning: Fundamental principles and 10 grand challenges. *Statistics Surveys*, *16*(none). https://doi.org/10.1214/21-SS133

Sun, M., Han, C., Nie, Q., Xu, J., Zhang, F., & Zhao, Q. (2022). Understanding building energy efficiency with administrative and emerging urban big data by deep learning in Glasgow. *Energy and Buildings*, *273*, 112331. https://doi.org/10.1016/j.enbuild.2022.112331

U.S. electricity retail price by sector 2024. (n.d.). Statista. Retrieved December 10, 2025, from https://www.statista.com/statistics/200197/average-retail-price-of-electricity-in-the-us-by-sector-since-1998/

U.S. primary energy production, consumption, and exports increased in 2024—U.S. Energy Information Administration (EIA). (n.d.). Retrieved December 10, 2025, from https://www.eia.gov/todayinenergy/detail.php?id=65524

Zargarzadeh, S., Ramnarayan, A., Castro, F. D., & Ohadi, M. (2024). ML-Enabled Solar PV Electricity Generation Projection for a Large Academic Campus to Reduce Onsite CO2 Emissions. *Energies*, *17*(23), 6188. https://doi.org/10.3390/en17236188

Zhang, F., Mockus, A., Keivanloo, I., & Zou, Y. (2016). Towards building a universal defect prediction model with rank transformed predictors. *Empirical Software Engineering*, *21*(5), 2107–2145. https://doi.org/10.1007/s10664-015-9396-2

Zhang, Y., Wang, J., & Luo, X. (2015a). Probabilistic wind power forecasting based on logarithmic transformation and boundary kernel. *Energy Conversion and Management*, *96*, 440–451. https://doi.org/10.1016/j.enconman.2015.03.012

Zhang, Y., Wang, J., & Luo, X. (2015b). Probabilistic wind power forecasting based on logarithmic transformation and boundary kernel. *Energy Conversion and Management*, *96*, 440–451. https://doi.org/10.1016/j.enconman.2015.03.012