

Household Energy Consumption Prediction Using Ensemble Machine Learning Models: A Comparative Analysis

Edris Shamsulhaq*, Fikri Arif Wicaksana

Electrical Engineering, Faculty of Engineering, Computer, and Design, Nusa Putra University
Jl. Raya Cibolang Cisaat - Sukabumi No.21, Sukabumi, 43152, Indonesia
E-mail: edris.shamsulhaq_te23@nusaputra.ac.id

Naskah Masuk: 13 Agustus 2026; Diterima: 15 September 2026; Terbit: 30 September 2026

ABSTRAK

Abstrak - Prediksi konsumsi energi rumah tangga yang akurat memainkan peran penting dalam manajemen jaringan cerdas (smart grid), respons sisi permintaan (demand-side response), dan perencanaan efisiensi energi. Penelitian ini menyajikan analisis komparatif dari empat model pembelajaran mesin ansambel—Random Forest, XGBoost, LightGBM, dan CatBoost—untuk memprediksi konsumsi daya aktif rumah tangga. Dataset Konsumsi Daya Listrik Rumah Tangga yang tersedia untuk umum dari repositori UCI, yang terdiri dari lebih dari dua juta pengukuran, digunakan untuk eksperimen. Tiga skenario fitur dirancang untuk mengevaluasi kinerja model di bawah kondisi input yang berbeda: menggunakan semua fitur, mengecualikan intensitas arus, serta mengecualikan intensitas arus dan daya reaktif. Hasil penelitian menunjukkan bahwa Random Forest mencapai akurasi prediksi tertinggi dengan nilai R^2 sebesar 0,913 pada skenario pengurangan fitur, mengungguli alternatif gradient boosting lainnya. Validasi statistik menggunakan uji peringkat bertanda Wilcoxon digunakan untuk menilai efek dari pengecualian fitur di ketiga skenario tersebut. Penelitian ini membuktikan bahwa prediksi energi yang andal tetap dapat dicapai dengan menggunakan fitur sub-metering, tegangan, dan temporal, sehingga menawarkan standar acuan (baseline) yang praktis untuk aplikasi meteran cerdas (smart meter).

Kata kunci: Prediksi Energi Rumah Tangga, Pembelajaran Ansambel, Random Forest, Kepentingan Fitur, Jaringan Cerdas

ABSTRACT

Abstract - Accurate prediction of household energy consumption plays a vital role in smart grid management, demand-side response, and energy efficiency planning. This study presents a comparative analysis of four ensemble machine learning models—Random Forest, XGBoost, LightGBM, and CatBoost—for predicting household active power consumption. The publicly available Household Electric Power Consumption dataset from the UCI repository, comprising over two million measurements, was used for experimentation. Three feature scenarios were designed to evaluate model performance under different input conditions: using all features, excluding current intensity, and excluding both intensity and reactive power. Results show that Random Forest achieved the highest predictive accuracy with an R^2 of 0.913 in the reduced-feature scenario, outperforming gradient boosting alternatives. Statistical validation using the Wilcoxon signed-rank test was used to assess the effect of feature exclusion across the three scenarios. The study demonstrates that reliable energy prediction remains achievable using sub-metering, voltage, and temporal features, offering a practical baseline for smart meter applications.

Keywords: household energy prediction, ensemble learning, Random Forest, feature importance, smart grid.

Copyright © 2026 Jurnal Teknik Elektro dan Komputasi (ELKOM)

1. INTRODUCTION

Household energy consumption accounts for a substantial portion of total electricity demand worldwide [1]. Understanding and predicting residential energy usage patterns has become increasingly important for utility companies, policymakers, and consumers alike [2]. Accurate short-term load forecasting enables better grid stability, efficient energy distribution, and informed decision-making regarding energy conservation measures [3][4].

The widespread deployment of smart meters in recent years has generated vast amounts of granular energy consumption data [5]. This data availability has opened new opportunities for applying machine

learning techniques to energy prediction problems [6]. Unlike traditional statistical methods, machine learning models can capture complex non-linear relationships between various factors affecting energy usage, such as time of day, seasonal patterns, and appliance-level consumption [7][8].

Several studies have explored the application of machine learning to household energy prediction. Ensemble methods, in particular, have demonstrated strong performance due to their ability to reduce overfitting and capture diverse patterns in the data [9][10]. Random Forest, a bagging-based ensemble, has been widely adopted for its robustness and interpretability. Gradient boosting variants such as XGBoost, LightGBM, and CatBoost have also gained popularity for their competitive accuracy and computational efficiency [11][12].

However, a common methodological concern in energy prediction research relates to feature scenario analysis. Certain electrical variables, such as current intensity, are physically derived from the target variable through fundamental electrical relationships. Including such features can artificially inflate model performance, creating an illusion of accuracy that would not hold in real forecasting scenarios where these variables are unknown. This issue requires careful experimental design to ensure that reported results reflect genuine predictive capability.

The present study addresses this concern by systematically evaluating model performance across three feature scenarios with varying levels of physical coupling to the target variable. The primary objectives of this research are to compare the predictive performance of four prominent ensemble learning algorithms on household energy consumption data, investigate the impact of removing physically derived electrical features on model accuracy, identify the most important predictors of household energy usage through multiple feature importance techniques, and provide a statistically validated baseline for future research in household energy prediction.

The remainder of this paper is organized as follows. Section 2 reviews related work in energy consumption prediction. Section 3 describes the dataset, preprocessing steps, and experimental methodology. Section 4 presents and discusses the results, including exploratory analysis, model comparisons, and statistical validation. Section 5 concludes the study and outlines directions for future work.

2. LITERATURE REVIEW

Accurate prediction of household energy consumption is essential for smart grid operations and demand-side management [13]. Ensemble learning methods have been widely adopted due to their ability to reduce overfitting and capture complex patterns in energy data. For instance, Random Forest and gradient boosting variants such as XGBoost, LightGBM, and CatBoost have been compared for household power prediction, with Random Forest showing strong robustness in realistic scenarios [14]. Other studies have proposed stacking, voting, and averaging ensembles to improve residential load forecasting, often achieving R^2 values above 0.98[15].

Deep learning approaches, including Long ShortTerm Memory (LSTM) and residual neural networks, have also been applied to day ahead household forecasting, demonstrating competitive accuracy when integrating weather and calendar features [16]. However, these models often require large datasets and offer limited interpretability [17].

Feature scenario analysis and data quality significantly affect prediction performance. Removing physically derived variables, such as current intensity, can prevent artificially inflated accuracy and reveal the true contribution of temporal and sub metering features [14]. Historical consumption lags and average past usage are consistently found to be more influential than environmental factors like temperature and humidity [17].

The COVID 19 pandemic introduced abrupt changes in consumption patterns, prompting the use of sliding windows and ensemble stacking to maintain prediction accuracy under disruptive conditions [18]. Additionally, explainable AI techniques, particularly SHAP, have been integrated with ensemble models to identify key drivers such as peak hour usage and household size, enhancing both performance and transparency [19].

However, relatively limited attention has been given to evaluating ensemble models under systematically reduced feature settings, particularly when physically coupled electrical variables are excluded. This study addresses this gap by comparing four ensemble models across three feature scenarios and examining the resulting changes in predictive performance and feature importance.

3. RESEARCH METHODOLOGY

3.1. Dataset Description

This study utilizes the publicly available Household Electric Power Consumption dataset from the UCI Machine Learning Repository [20]. The dataset contains 2,049,280 measurements collected from

a single household in Sceaux, France, between December 2006 and November 2010. Each record includes seven electrical features: global active power (kW), global reactive power (kW), voltage (V), global current intensity (A), and three sub-metering values (Wh) corresponding to kitchen, laundry, and electric water heater/air conditioner usage. The target variable for prediction is `Global_active_power`.

3.2. Data Preprocessing

The dataset was provided as a pre-cleaned CSV file with no missing values. Two time-based features, Hour (0–23) and Month (1–12), were extracted from the `Datetime` column to capture temporal consumption patterns. All numerical features were cast to 32-bit floating point or integer types to reduce memory usage. No further scaling or normalization was applied, as all employed models are tree-based and insensitive to feature scaling.

3.3. Exploratory Data Analysis

A comprehensive exploratory data analysis was conducted on the full dataset (2,049,280 records) prior to sampling. Descriptive statistics were computed for each variable. A histogram of `Global_active_power` was plotted to visualize its distribution. Average hourly and monthly consumption patterns were examined to identify temporal trends. A Pearson correlation heatmap was generated from a representative random subset of 100,000 records.

3.4. Sampling Strategy

To make model training computationally feasible, a random sample of 300,000 records was drawn without replacement (random state = 42). All subsequent modeling and feature scenario analysis steps were performed exclusively on this sample. A sensitivity check confirmed that the sample accurately represents the full data distribution.

3.5. Feature Importance Analysis

Feature importance was assessed using three complementary techniques, all applied strictly to the training data:

- Mutual Information (MI): captures both linear and non-linear dependencies between each feature and the target.
- Random Forest Built-in Importance: mean decrease in impurity computed during construction of a Random Forest with 100 trees.
- Permutation Importance: drop in R^2 score when values of a single feature are randomly shuffled.

3.6. Machine Learning Models

Four ensemble learning algorithms were selected:

- Random Forest (RF): ensemble of decision trees trained on bootstrap samples, configured with 100 trees.
- XGBoost (XGB): gradient-boosted tree algorithm using histogram-based split finder.
- LightGBM (LGBM): gradient boosting framework using leaf-wise tree growth.
- CatBoost (CAT): gradient boosting with native categorical feature support.
- All models were implemented with a fixed random seed (42) for reproducibility.

3.7. Experimental Design (Scenarios A–C)

Three feature subsets were defined:

- Scenario A (All Features): all ten features including Voltage, `Global_intensity`, `Global_reactive_power`, `Sub_metering_1–3`, Hour, Day, Month, and Weekday.
- Scenario B (Without `Global_intensity`): removes `Global_intensity`, which is strongly correlated with the target.
- Scenario C (Without `Global_intensity` and `Global_reactive_power`): additionally drops `Global_reactive_power`.

For each scenario, the sample was split into training (80%, 240k records) and test (20%, 60k records) sets with shuffle enabled and fixed random state.

3.8. Hyperparameter Optimization

Random Forest was selected for hyperparameter tuning using Randomized Search with 5-fold cross-validation. Twenty random combinations of parameters were evaluated using R^2 as the scoring metric.

The baseline Random Forest employed 100 trees, while the tuned model used the optimal configuration identified through search.

3.9. Performance Evaluation Metrics

Four standard regression metrics were used:

- Mean Absolute Error (MAE)
- Root Mean Squared Error (RMSE)
- Mean Absolute Percentage Error (MAPE)
- Coefficient of Determination (R^2)

4. Results and Discussion

4.1. Exploratory Data Analysis Results

Table 1. Descriptive Statistics of the Full Dataset

Feature	Mean	Std	Min	25%	50%	75%	Max
Global_active_power	1.09	1.06	0.08	0.31	0.60	1.53	11.12
Global_reactive_power	0.12	0.11	0.00	0.05	0.10	0.19	1.39
Voltage	240.84	3.24	223.20	239.00	241.00	243.00	254.20
Global_intensity	4.63	4.45	0.20	1.40	2.60	6.40	48.40
Sub_metering_1	1.12	3.69	0.00	0.00	0.00	0.00	88.00
Sub_metering_2	1.30	4.79	0.00	0.00	0.00	0.00	80.00
Sub_metering_3	6.46	12.62	0.00	0.00	0.00	3.00	84.00

The target variable `Global_active_power` exhibits a right-skewed distribution (Figure 1), with a mean of 1.09 kW and a large proportion of low consumption values representing idle periods.

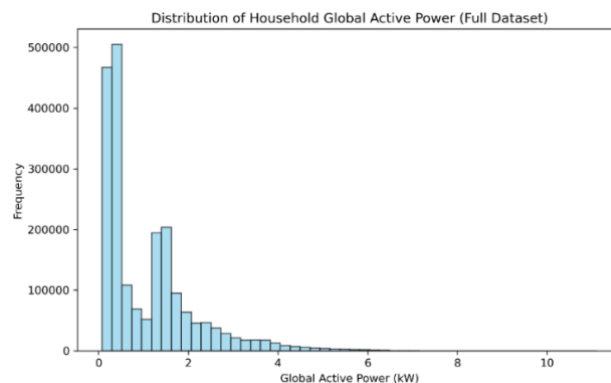


Figure 1. Distribution of household global active power across the full dataset, showing strong right-skewness with most values concentrated near zero.

The average hourly consumption (Figure 2) displays a distinct evening peak between 18:00 and 20:00. Monthly patterns reveal higher demand during winter months (December–February) and lower in summer.

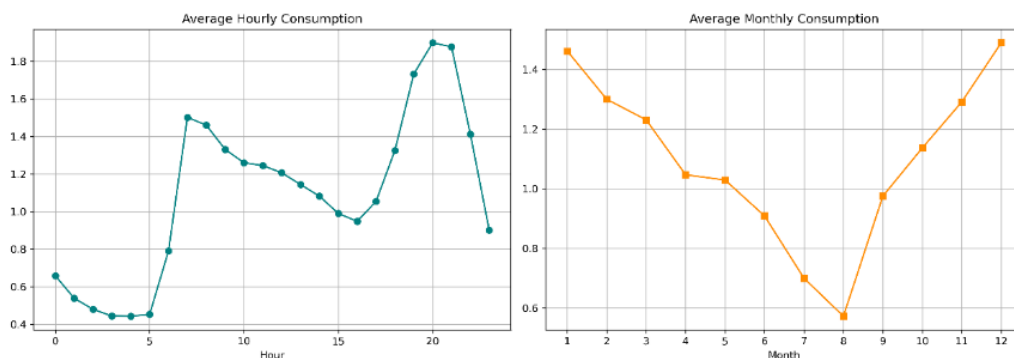


Figure 2. Average household energy consumption: (left) hourly pattern with evening peak at 18:00–20:00; (right) monthly pattern with higher winter demand.

The correlation heatmap (Figure 3) shows Global_intensity almost perfectly correlated with Global_active_power ($r \approx 0.99$). Global_reactive_power and Sub_metering_3 also show moderate to high correlations, confirming physical linkage and motivating Scenarios B and C.

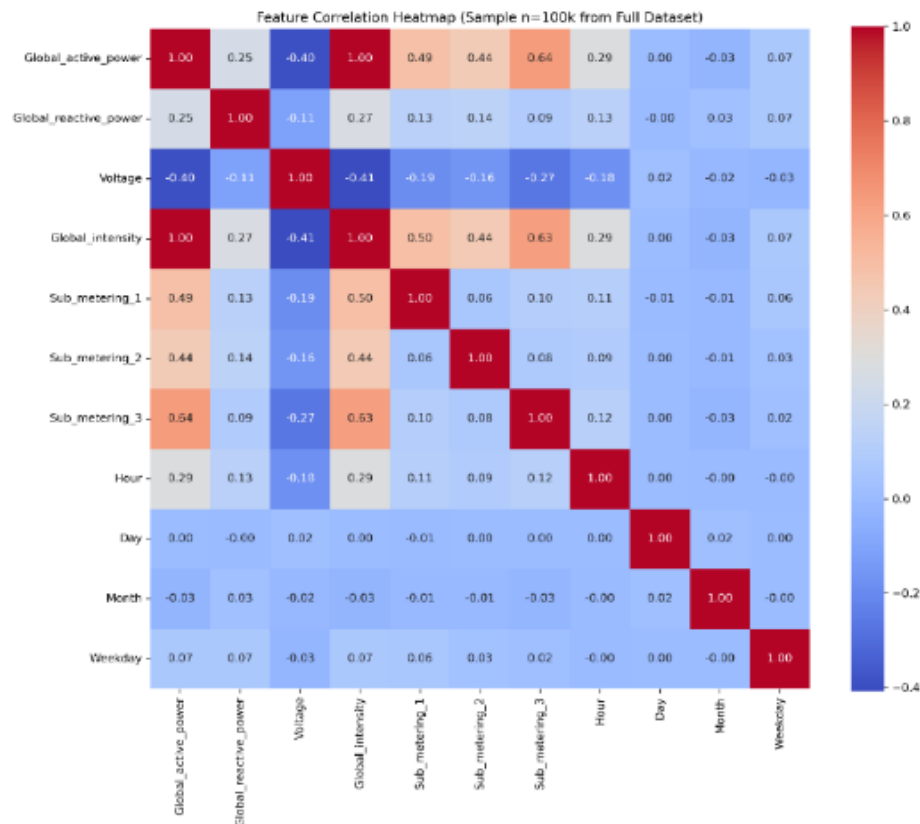


Figure 3. Pearson correlation heatmap showing Global_intensity nearly perfectly correlated with Global_active_power ($r \approx 0.99$).

4.2. Feature Importance Analysis

Figure 4 presents the top 10 features across all three importance methods. Sub_metering_3 consistently ranks among the most influential predictors across all methods. Temporal features, particularly Hour and Month, demonstrate substantial importance, confirming the strong role of time-based consumption patterns. Features like Day and Weekday provide less but still meaningful information.

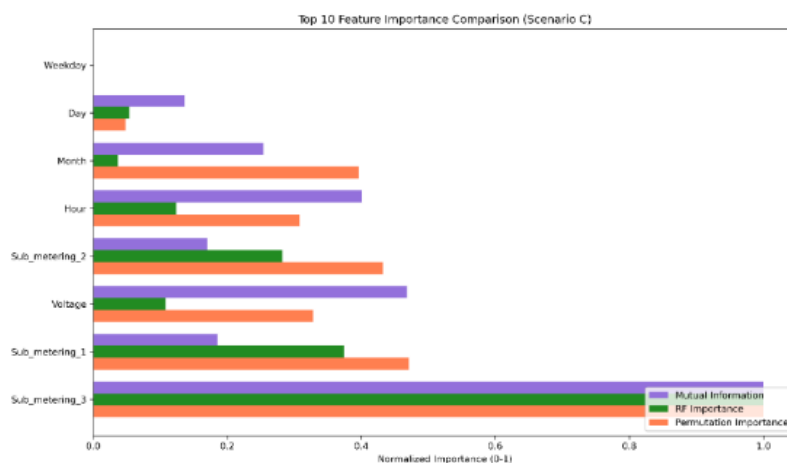


Figure 4. Top 10 features ranked by Mutual Information, Random Forest importance, and Permutation Importance. Sub_metering_3 and Hour consistently rank highest

4.3. Scenario A Results

Table 2. Model Performance in Scenario A (All Features)

Model	MAE	RMSE	MAPE	R ²	Training Time (s)
CatBoost	0.01730	0.02936	0.02981	0.999242	56.81
XGBoost	0.01748	0.02959	0.02959	0.999230	2.70
Random Forest	0.01572	0.03075	0.02163	0.999168	245.59
LightGBM	0.01802	0.03095	0.03176	0.999157	2.55

All algorithms achieve R² of approximately 0.9992. CatBoost marginally leads with R² of 0.999242. These results reflect the inclusion of Global_intensity, which provides highly informative information about the target, and serve primarily as a reference benchmark.

4.4. Scenario B Results

Table 3. Model Performance in Scenario B (Without Global_intensity)

Model	MAE	RMSE	MAPE	R ²	Training Time (s)
Random Forest	0.16344	0.31053	0.19376	0.915159	250.97
CatBoost	0.21417	0.35494	0.29213	0.889157	50.01
XGBoost	0.21511	0.35720	0.29086	0.887737	2.37
LightGBM	0.22939	0.37886	0.30911	0.873713	2.19

With Global_intensity removed, R² values fall to 0.87–0.92. Random Forest achieves the highest R² of 0.915. The substantial decline confirms that Global_intensity provided the dominant share of predictive information, yet remaining features still capture meaningful consumption patterns.

4.5. Scenario C Results

Table 4. Model Performance in Scenario C (Without Global_intensity and Global_reactive_power)

Model	MAE	RMSE	MAPE	R ²	Training Time (s)
Random Forest	0.17310	0.31462	0.23821	0.912910	258.74
CatBoost	0.23026	0.37264	0.34046	0.877826	46.60
XGBoost	0.23042	0.37451	0.33765	0.876598	1.72
LightGBM	0.24533	0.39606	0.36033	0.861987	1.56

Scenario C performance decreases only marginally compared to Scenario B. Random Forest achieves R² of 0.913, a drop of merely 0.002. This suggests that Global_reactive_power provides limited additional predictive information once Global_intensity is excluded.

To assess the stability of the reported results, a 10-fold shuffled cross-validation was additionally performed as a robustness check. The average R² scores obtained for Scenario C were 0.9094 ± 0.0027 for Random Forest, 0.8740 ± 0.0031 for XGBoost, 0.8593 ± 0.0035 for LightGBM, and 0.8752 ± 0.0029 for CatBoost. These values are consistent with the hold-out test results presented in Table 4, confirming that the observed performance rankings are not highly sensitive to the specific train-test split.

4.6. Hyperparameter Optimization Results

Table 5. Hyperparameter Tuning Results (Random Forest, Scenario C)

Model Configuration	MAE	RMSE	MAPE	R ²
Random Forest (Default, 100 trees)	0.17310	0.31462	0.23821	0.912910
Random Forest (Tuned, 150 trees)	0.17528	0.31285	0.24530	0.913883

The tuned model achieved R² of 0.9139 with marginal improvement ($\Delta R^2 \approx 0.001$), suggesting that default Random Forest configuration already provides near-optimal performance.

Table 6. Summary of Best-Performing Model in Each Configuration

Scenario / Configuration	Best Model	R ²
Scenario A (All features)	CatBoost	0.99924
Scenario B (w/o intensity)	Random Forest	0.91516
Scenario C (w/o intensity & react)	Random Forest	0.91291
Scenario C – Tuned RF (150 trees)	Random Forest	0.91388

Note: The tuned configuration is not an additional feature scenario; it represents hyperparameter optimization applied to Scenario C.

Figure 5 presents the grouped bar chart comparing R^2 scores across all scenarios and models. The visual contrast between Scenario A and realistic scenarios (B and C) underscores the importance of proper feature exclusion.

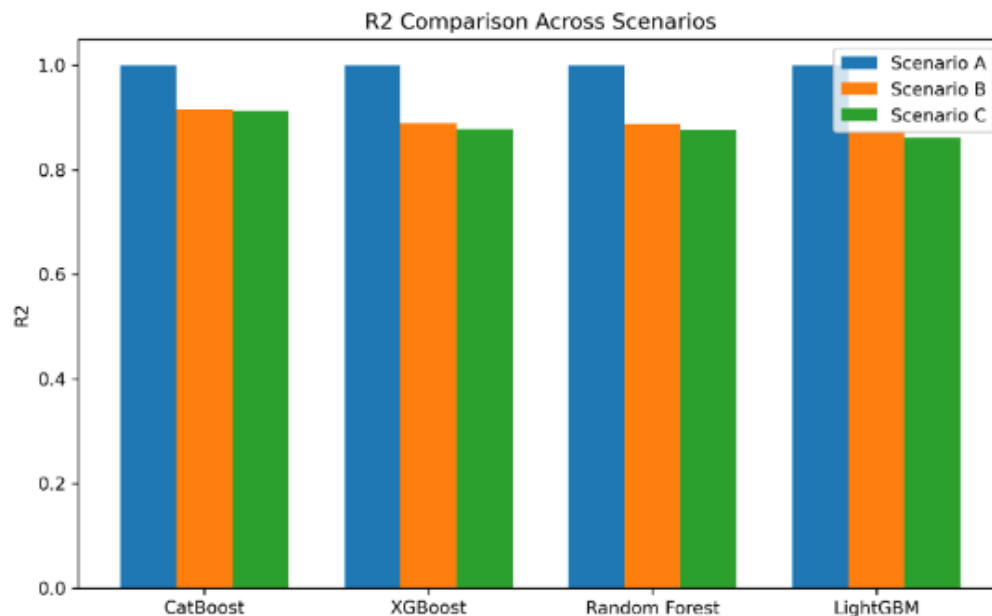


Figure 5. Random Forest achieves the highest R^2 among the evaluated models in Scenarios B and C

4.7. Statistical Analysis

Since the distributional assumptions of the cross-validation performance differences were not assumed to be normal, the non parametric Wilcoxon signed rank test was employed to statistically validate the observed performance differences. The test was applied to compare the cross-validation R^2 scores obtained under the three feature scenarios for each model individually, rather than directly comparing different algorithms against each other. This approach ensures that the analysis focuses on the impact of feature exclusion on model stability. A significance level of $\alpha = 0.05$ was adopted for all comparisons.

Table 7 summarizes the p values obtained for scenario comparisons per model. For Random Forest, the comparison between Scenario B and C yielded a p value of 0.0645, which is above the 0.05 threshold. This indicates that removing Global_reactive_power (Scenario C) does not cause a statistically significant degradation in Random Forest's performance compared to the setting that still retains it (Scenario B). In contrast, the same B vs C comparison produced p values of 0.0020 for XGBoost, LightGBM, and CatBoost, confirming that boosting algorithms are significantly more sensitive to the exclusion of reactive power. For all models, the comparisons between Scenario A and B, as well as Scenario A and C, were highly significant ($p = 0.0020$), confirming that the removal of Global_intensity introduces a substantial and statistically meaningful decline in predictive accuracy.

Table 7. Wilcoxon Signed Rank Test – Scenario Comparisons per Model

Model	A vs B	B vs C	A vs C
Random Forest	0.0020 *	0.0645	0.0020 *
XGBoost	0.0020 *	0.0020 *	0.0020 *
LightGBM	0.0020 *	0.0020 *	0.0020 *
CatBoost	0.0020 *	0.0020 *	0.0020 *

* Significant at $\alpha = 0.05$.

4.8. Comparative Discussion

Key findings from the results include:

1-Impact of physically derived features: Scenario A's near-perfect R^2 (≈ 0.999) illustrates how electrically coupled variables can dominate learning. Removing Global_intensity drops performance

to levels comparable to practically constrained input settings, yet remaining features still explain over 91% of variance.

2-Model robustness: Random Forest demonstrated highest robustness after removing physically related variables, outperforming gradient boosting by 2–3 percentage points in R^2 across Scenarios B and C.

3-Redundancy of reactive power: Dropping `Global_reactive_power` causes negligible deterioration, suggesting that `Global_reactive_power` provides limited additional predictive information, particularly when using Random Forest under the current feature setting.

4-Tuning impact: Marginal gain from hyperparameter optimization ($\Delta R^2 \approx 0.001$) indicates default Random Forest is already well-suited to this problem.

Figure 6 shows the Actual vs. Predicted scatter plot for the best model. Points cluster tightly along the diagonal, indicating good agreement. Figure 7 presents the residual plot, with errors symmetrically distributed around zero and no strong patterns, suggesting minimal systematic bias.

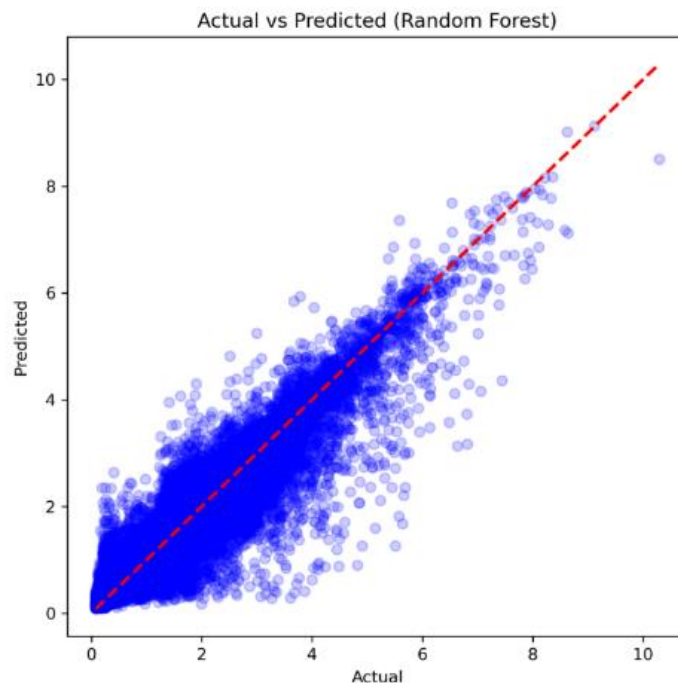


Figure 6. Actual vs. predicted values for Random Forest (Scenario C). Points cluster along the diagonal, indicating strong predictive agreement

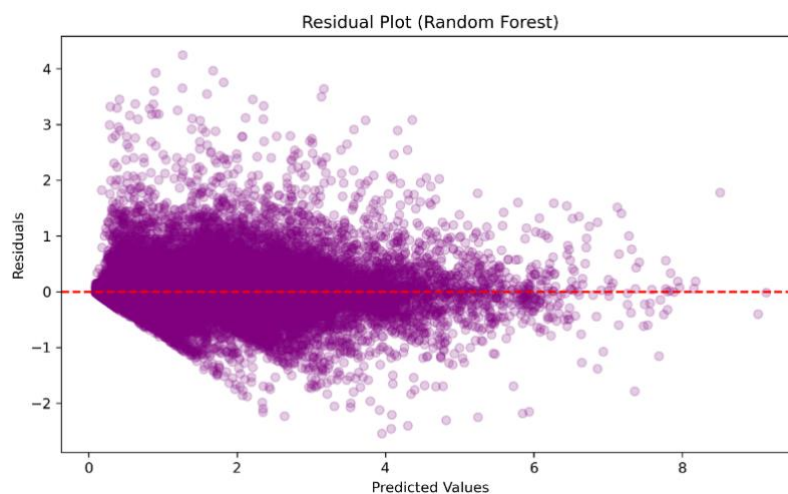


Figure 7. Residual plot for Random Forest (Scenario C). Errors are symmetrically distributed around zero with no systematic patterns

4.9. Practical Implications

- a. Smart meter applicability: Although current intensity can be measured by smart meters, it may not be available for future consumption forecasting. Scenario C approximates a practically constrained feature set that may reflect limited-input conditions in some smart meter applications and still achieves high accuracy.
- b. Model selection: Random Forest achieved the best balance between accuracy and robustness for the investigated dataset.

4.10. Limitations and Future Work

Several limitations warrant acknowledgment. The dataset originates from a single household in France, restricting generalizability. All experiments used a single random train/test split. External validation from an independent time period was not performed.

Additionally, the use of a single random train-test split (with shuffle enabled) limits the ability to claim true out-of-sample forecasting performance, as it does not fully preserve the sequential nature of time-series data.

Future work should adopt chronological splitting to better evaluate real-world predictive applicability, it also should extend to multi-household datasets to assess cross-building generalization. Comparing tree-based ensembles with sequence-aware models such as transformers or recurrent neural networks could reveal whether temporal dynamics offer additional predictive power.

5. CONCLUSION

This study compared four ensemble machine learning models for household energy prediction across three feature scenarios. Among the evaluated models, Random Forest achieved the highest accuracy in the reduced-feature scenarios, reaching an R^2 of 0.913 in Scenario C. Conversely, including Global_intensity artificially inflated performance to near-perfect levels with an R^2 of approximately 0.999. After Global_intensity was excluded, Global_reactive_power provided limited additional predictive information. Statistical testing confirmed that removing Global_intensity significantly affects model performance, while removing Global_reactive_power had no significant effect for Random Forest. Furthermore, hyperparameter optimization yielded only marginal improvements. Ultimately, these findings support adopting Scenario C with Random Forest as a reliable baseline for household energy prediction using commonly available smart meter features.

REFERENCES

- [1] T. M. Olatunde, A. C. Okwandu, D. O. Akande, and E. IJSTRA, "Reviewing the impact of energy-efficient appliances on household consumption," *International Journal of Science and Technology Research Archive*, vol. 6, no. 2, pp. 001–011, Apr. 2024, doi: 10.53771/IJSTRA.2024.6.2.0038.
- [2] A. Bandyopadhyay and A. Bhattacharya, "Residential Appliance Usage Patterns From Overall Energy Consumption Data: A Statistical Machine Learning Approach," *ASME International Mechanical Engineering Congress and Exposition, Proceedings (IMECE)*, vol. 8A-2021, Jan. 2022, doi: 10.1115/IMECE2021-70122.
- [3] T. M. Ghazal, J. I. Janjua, W. Abushiba, M. Ahmad, T. Abbas, and K. Ateeq, "Cutting-Edge Predictive Methods for Enhancing Short-Term Load Prediction," *International Conference on Electrical and Computer Engineering Researches, ICECER 2024*, 2024, doi: 10.1109/ICECER62944.2024.10920448.
- [4] K. Okoli and Y. A. Bekeneva, "Accurate Neural Prophecy for Short-Term Load Forecasting of Optimal Renewable Energy," *Proceedings of 2023 5th International Conference on Control in Technical Systems, CTS 2023*, pp. 196–198, 2023, doi: 10.1109/CTS59431.2023.10288952.
- [5] Y. Wang, Q. Chen, T. Hong, and C. Kang, "Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 3125–3148, May 2019, doi: 10.1109/TSG.2018.2818167.
- [6] S. M. Qaisar, D. A. Bashawyah, F. Alsharif, and A. Subasi, "A comprehensive review on the application of machine learning techniques for analyzing the smart meter data," *Machine Learning for Sustainable Development*, pp. 53–75, Jul. 2021, doi: 10.1515/9783110702514-004/XML.
- [7] A. Wong et al., "Machine Learning Models Application in Daily Forecasting of Hourly Electricity Usage," *SysCon 2022 - 16th Annual IEEE International Systems Conference, Proceedings*, 2022, doi: 10.1109/SYSCON53536.2022.9773835.
- [8] M. Tarani and T. Sowjanya, "Smart Forecasting of Power Consumption with Machine Learning Models," *International Scientific Journal of Engineering and Management*, vol. 05, no. 07, pp. 1–9, Jul. 2026, doi: 10.55041/ISJEM08278.
- [9] P. A. Ujjainkar, T. H. Khan, S. Nirkhi, S. Mudpalliwar, S. Chavan, and K. K. Baraskar, "Energy

- Consumption Prediction using Ensemble Learning and Streamlit Application,” 2025 International Conference on Artificial Intelligence and Quantum Computation-Based Sensor Application, ICAIQSA 2025 - Proceedings, 2025, doi: 10.1109/ICAIQSA67794.2025.11440429.
- [10] G. Airlangga, “Evaluation of Ensemble and Hybrid Models for Predicting Household Energy Consumption: A Comparative Study of Machine Learning Approaches,” *Indonesian Journal of Artificial Intelligence and Data Mining*, vol. 8, no. 1, pp. 94–101, Dec. 2025, doi: 10.24014/IJAIDM.V8I1.32819.
- [11] I. D. Mienye and Y. Sun, “A Survey of Ensemble Learning: Concepts, Algorithms, Applications, and Prospects,” *IEEE Access*, vol. 10, pp. 99129–99149, 2022, doi: 10.1109/ACCESS.2022.3207287.
- [12] Er. M. Kaur, Er. G. Kaur, Dr. R. Rana, and Mr. S. Sharma, “A Comparative Study of XGBoost and Random Forest for Modern Machine Learning Applications,” *International Journal of Scientific Research in Engineering and Management*, vol. 09, no. 12, pp. 1–9, Dec. 2025, doi: 10.55041/IJSREM55183.
- [13] G. S. Ramnath, R. Harikrishnan, S. M. Muyeen, and K. Kotecha, “Household electricity consumption prediction using database combinations, ensemble and hybrid modeling techniques,” *Sci. Rep.*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-57550-9.
- [14] D. Attipoe, D. K. Moulla, E. Mnkandla, and A. Abran, “Predicting Residential Energy Consumption in South Africa Using Ensemble Models,” *Applied Computational Intelligence and Soft Computing*, vol. 2025, no. 1, 2025, doi: 10.1155/acs/5211419.
- [15] J. Moon, M. Maqsood, D. So, S. W. Baik, S. Rho, and Y. Nam, “Advancing ensemble learning techniques for residential building electricity consumption forecasting: Insight from explainable artificial intelligence,” *PLoS One*, vol. 19, no. 11 November, Nov. 2024, doi: 10.1371/journal.pone.0307654.
- [16] I. Kiprijanovska, S. Stankoski, I. Ilievski, S. Jovanovski, M. Gams, and H. Gjoreski, “HousEEC: Day-ahead household electrical energy consumption forecasting using deep learning,” *Energies (Basel)*, vol. 13, no. 10, May 2020, doi: 10.3390/en13102672.
- [17] E. Fitri, S. N. Nugraha, and M. Ernawati, “ELECTRICITY CONSUMPTION PREDICTION AND INFLUENTIAL FACTORS ANALYSIS USING MACHINE LEARNING REGRESSION,” *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, vol. 11, no. 4, pp. 1197–1211, May 2026, doi: 10.33480/jitk.v11i4.7914.
- [18] P. H. K. Utama et al., “Adaptive Ensemble Learning for Enhancing Building Energy Consumption Prediction: Insights from COVID-19 Pandemic Energy Consumption Dynamics,” *Journal of Engineering and Technological Sciences*, vol. 57, no. 2, pp. 161–174, Apr. 2025, doi: 10.5614/j.eng.technol.sci.2025.57.2.2.
- [19] B. B. Muslim and J. Munawar, “Optimizing Daily Household Energy Consumption Prediction Using Ensemble Machine Learning, Feature Selection, and Explainable Artificial Intelligence,” *Journal of Intelligent Systems Technology and Informatics*, vol. 2, no. 2, pp. 82–90, Jul. 2026, doi: 10.64878/jistics.v2i2.224.
- [20] G. Hebrail and A. Berard, “Individual Household Electric Power Consumption,” 2006, UCI Machine Learning Repository. doi:10.24432/C58K54.