WORKING PAPER 301/2026

Comparative Study of Machine Learning and Deep Learning Models for Short-Term Energy Consumption Prediction

Alice Treesa M
Arpita Choudhary

MADRAS SCHOOL OF ECONOMICS
Gandhi Mandapam Road
Chennai 600 025
India

May 2026

Comparative Study of Machine Learning and Deep Learning Models for Short-Term Energy Consumption Prediction

Alice Treesa M

and

Arpita Choudhary
Assistant Professor Madras School of Economics, Chennai, India.
arpita@mse.ac.in

WORKING PAPER 301/2026

MADRAS SCHOOL OF ECONOMICS
Gandhi Mandapam Road
Chennai600 025
India

May 2026

Phone: 2230 0304/2230 0307/2235 2157
Fax: 2235 4847/2235 2155
Email : info@mse.ac.in
Website: www.mse.ac.in

Price : Rs. 35

Comparative Study of Machine Learning and Deep Learning Models for Short-Term Energy Consumption Prediction

Alice Treesa M and Dr. Arpita Choudhary

Abstract

Reliable energy consumption forecasting in the short term is essential for improving building operations efficiency and creating sustainable energy consumption plans. The authors of this study evaluate the forecasting performance of machine learning and deep learning methods which use climate data and time data to predict energy usage at hourly intervals. The study used Linear Regression, Decision Trees, Random Forest, XGBoost and Long Short-Term Memory as comparison methods to assess performance in the same context. The study demonstrated that energy consumption forecasting accuracy depends more on selected features than on the model's complexity. The study found that LSTM model learning capacity remained stable while Random Forest model performance showed superior results in dealing with nonlinear features that had temporal attributes.

Keywords: *Energy Consumption Prediction, Machine Learning, Ensemble Models, LSTM Model, Feature Engineering, Sustainable Energy Management*

JEL CODES: *Q47, C53, C45, C38, L94, Q41*

Acknowledgement

The authors would like to express their sincere gratitude to Madras School of Economics for providing a strong academic and research-oriented environment that supported the successful completion of this study. The institution's emphasis on analytical rigor and interdisciplinary learning greatly contributed to the development of this research. The authors also acknowledge the Mendeley Data repository for providing open access to the Building A energy consumption dataset used in this study. The availability of high-quality public data enabled a structured comparative analysis of machine learning and deep learning models for energy consumption forecasting. An earlier version of this paper was presented at the LEO Economics Conference held during 24-26 April 2026 at Plaksha University, Mohali. The authors are grateful for the valuable comments and suggestions received from conference participants, which helped improve the quality and clarity of the paper. Any remaining errors are solely the responsibility of the authors.

Alice Treesa
Arpita Choudhary

INTRODUCTION

The process of forecasting short-term energy consumption establishes itself as a vital component which enhances operational efficiency while supporting environmentally friendly building management methods. The demand for electricity from buildings throughout the world enables even small enhancements in demand prediction to produce significant decreases in resource waste and operational expenses and carbon dioxide emissions (Ürge-Vorsatz et al., 2015). Forecasting methods have moved away from traditional statistical methods because digital infrastructure expansion and smart meter technologies produce continuous time-stamped consumption data. The standard linear regression and ARIMA-based modeling techniques provide clear results while requiring less computing power yet they fail to correctly represent the complex seasonal energy usage patterns which occur in actual energy datasets.

The new access to detailed consumption information has led to an increased use of machine learning (ML) models for predicting energy consumption at the building level. The research demonstrates that algorithms Support Vector Machines (SVM) Artificial Neural Networks (ANN) and other supervised learning models provide better results when dealing with nonlinear energy data relationships than traditional statistical methods (Azadeh et al., 2020). Ensemble-based models, which include Random Forest and gradient boosting methods, create stronger prediction performance through their process of combining multiple weak learners, which results in decreased overfitting and enhanced performance across various usage patterns. These models achieve their highest performance when engineers create features that include external factors such as weather conditions and occupancy trends and time-based markers.

The traditional machine learning methods demonstrate their ability to solve problems but their approach to handling data treats observations as conditionally independent which results in their inability to capture the time-based relationships that exist in energy data that follows a sequence. The development of deep learning architectures which includes Recurrent Neural Networks (RNNs) and their advanced variants, emerged as a solution to this existing limitation. Long Short-Term Memory (LSTM) networks use their gated memory systems to solve vanishing gradient problems which enables them to effectively model short-term load forecasting tasks that involve multiple sequential dependencies (Marino et al., 2016; Kong et al., 2019). The research results show that LSTM-based models can successfully model longterm temporal patterns and sudden changes in energy usage which residential and commercial buildings exhibit. The performance of LSTM models depends on the structure of the dataset and the method of measuring time intervals and the way features are shown and the methods used to prepare the data. The classical ensemble ML models maintain their effectiveness in specific situations which involve datasets that have limited size or slow changing patterns.

Energy forecasting now extends beyond its original purpose of achieving accurate results because its examination requires assessment of sustainability and corporate responsibility. Environmental performance assessment requires energy consumption optimization, which serves as a fundamental element of Environmental, Social, and Governance (ESG) assessment frameworks. Organizations achieve their environmental sustainability goals through improved forecasting, which enables them to manage demand, reduce peak loads, and implement emissions control methods (Ürge-Vorsatz et al., 2015). The meta-analysis demonstrates that organizations with better ESG performance maintain financial stability while creating value over an extended period (Friede et al., 2015). The framework uses energy efficiency metrics that include consumption intensity and emissions per unit output and carbon reduction trajectories to create environmental indicators, which investors and regulators now use to evaluate their progress Recent studies about sustainability research demonstrate how climate risk exposure affects corporate environmental performance which results in changes to company market value. Companies that implement environmental risk reduction practices will gain market stability and increased investor trust which research on

ESG investments and climate risk management shows (Ghallabi, 2025)(Friede et al., 2015) the ability to forecast energy needs accurately for the upcoming months provides organizations with data-driven tools to distribute their resources and measure their operational success. The evaluation of forecasting models requires assessment of two factors which include predictive accuracy and assessment of their computational capacity and ability to handle multiple users and support sustainability-oriented decision-making.

The present research study performs a comparative analysis between traditional machine learning methods and deep learning techniques for short-term building energy prediction using a publicly available dataset. The research has two main objectives. The first objective of the study tests different ML and DL models to measure their predictive accuracy through standardized experimental procedures. The second objective of the study analyzes how different factors affect the relationship between model complexity and computational efficiency and sustainability-related outcomes. The study establishes forecasting performance through an ESG sustainability framework which expands its boundaries beyond method comparison to support research on environmentally friendly data optimization. The research uses structured empirical evaluation methods to discover modeling methods which achieve both accurate predictions and functional operational requirements to assist energy-efficient building management and sustainable environmental governance systems.

DATA AND METHODOLOGY

Data Description

The dataset used for this research study is derived from Building A, which functions as a monitored commercial space and its data was sourced from a public data repository. The dataset contains 43,848 hourly entries which generate a detailed time-based model of electricity usage that extends throughout the entire observation period. The primary target variable is total hourly energy consumption (ENERGY), which the system measures in kilowatt-hours (kWh). The dataset provides hourly data which enables researchers to study the pattern of electrical usage throughout each day and observe daily usage patterns and short-term demand changes.

The Heating Degree Days (HDD18) and Cooling Degree Days (CDD0, CDD10) measurements provide data for assessing temperature-based heating and cooling needs. The systems display organized results through their factors which show how thermal stress affects both HVAC system performance and building enclosure functions. The study tracks daily thermal fluctuations that influence patterns of energy usage using supplementary temperature metrics, mean temperature (T2M), minimum temperature (T2M_MIN), and maximum temperature (T2M_MAX).

Environmental variables including relative humidity (RH2M) and total precipitation (PRECTOT) and sky cover (ALLSKY) provide additional contextual details about the situation. Internal lighting and solar heat gains get affected by sky cover changes while precipitation results in occupancy and operational condition changes and humidity levels determine cooling load requirements. The modeling framework uses these characteristics to develop an accurate electricity demand prediction model which runs on both direct and indirect environmental data. The system uses a binary holiday indicator (HOLIDAY) to show how occupancy and operating schedules will change during holiday periods. The variable aids in predicting by showing how people use business spaces on holidays compared to their typical weekday usage patterns.

The dataset contains almost complete data except for some missing values which occur mainly in energy-related and sky-related variables that were fixed during the preparation stage. The research developed a time-series dataset for short-term energy forecasting analysis which involved collecting environmental data and measuring high-frequency consumption data during standardized testing conditions.

The dataset has been anonymised to protect confidential information while the analysis process excludes any examination of specific location data. The study attempts to compare different methods because its primary goal needs understanding of spatial information. The research tests various algorithms together with their displayed features in a prediction environment which uses identical data conditions to assess model performance.

Data Cleaning and Temporal Imputation Strategy

The extensive data cleaning procedure was conducted before getting started with the model development work to establish both time consistency and research analysis accuracy. The dataset included hourly time-series data which was needed to maintain chronological sequence for their forecasting work. The DATE variable needs conversion to datetime format so users can perform time-based calculations. The data transformations were created which maintained original data order by utilizing the established temporal index. The dataset contained incomplete observations for both the environmental measurement system ALLSKY and the target variable ENERGY.

The time-based interpolation method was chosen because both energy consumption and weather-related variables change gradually rather than exhibiting sudden unpredictable shifts during regular operations. Used temporal index interpolation to calculate missing data points by using information from nearby data points. This method maintains short-term continuity while preventing the creation of artificial structural interruptions that would disrupt the series.

Time-based interpolation is widely considered appropriate for evenly spaced time-series data with adjacent observations that show dependency (Hyndman & Athanasopoulos, 2018). The approach preserved hourly demand patterns and environmental variations through its method of execution. The modeling process required complete datasets which were achieved through the interpolation process that removed remaining missing values. The datetime index needed restoration to its original format because it served as a standard column which enabled compatibility with upcoming modeling processes. The cleaning method maintained sequential order and statistical integrity which made the dataset appropriate for short-term energy forecasting.

Descriptive Statistical Analysis

A statistical study was initiated to examine how energy consumption data and its explanatory factors distribute across different patterns. The summary statistics show that hourly energy consumption varies greatly because its values extend from about 48 kWh to almost 470 kWh. The interquartile range shows that consumption patterns move through defined periods of time although they display active changes throughout the day.

The data demonstrates that CDD0 and CDD10 cooling indicators show higher average values than HDD18 heating degree days which indicates that cooling requirements will have a greater impact on electricity consumption during the studied period. The pattern matches buildings that exist within warm climate zones because these buildings experience high cooling requirements which form a major part of their total energy consumption.

The temperature variables which include mean temperature minimum temperature and maximum temperature display distinct fluctuations which demonstrate their ability to explain energy demand patterns. The relative humidity variable RH2M shows a wide range of distribution which indicates that atmospheric conditions are changing and these changes will indirectly impact HVAC system efficiency.

The dataset shows different patterns of precipitation and sky cover because these variables track environmental conditions which affect both indoor lighting needs and building heat requirements. The holiday indicator showed that 40 percent of the data points in the study occurred during holidays and days when regular operations did not take place. The energy use patterns show significant changes which depend on building occupancy levels.

The dataset contains sufficient environmental diversity according to the descriptive statistics which enables successful predictive modeling. Machine learning models require knowledge about variable features because their performance depends on data volume and statistical distribution (Hyndman & Athanasopoulos 2018).

Table 1: Descriptive Statistics of Energy and Climate Variables

Variable	N	Mean	SD	Min	25%	Median	75%	Max
ENERGY	43847	165.45	72.35	47.93	117.55	131.83	202.36	469.76
HDD18	43847	6.37	5.54	0.00	0.00	6.19	11.26	19.69
CDD0	43847	13.01	7.00	0.00	7.04	12.11	19.04	29.01
CDD10	43847	4.60	5.30	0.00	0.00	2.11	9.04	19.01
PRECTOT	43847	1.37	3.07	0.00	0.00	0.08	1.28	46.17
RH2M	43847	65.74	17.28	24.09	51.65	65.89	80.89	100.00
T2M	43847	12.51	7.26	-2.62	6.39	11.71	18.67	28.63
T2M_MIN	43847	6.86	5.88	-5.86	2.20	6.43	11.68	20.96
T2M_MAX	43847	19.16	8.47	2.76	12.03	18.17	26.27	38.66
ALLSKY	43847	4.56	2.35	0.27	2.35	4.45	6.73	8.67
HOLIDAY	43847	0.40	0.49	0.00	0.00	0.00	1.00	1.00

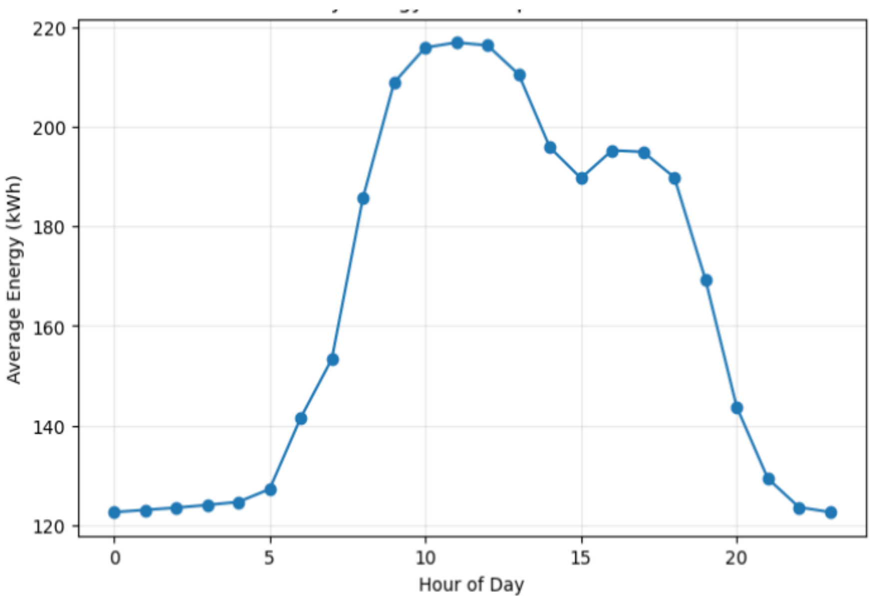
Source: Author's computation

Note: The number of hourly observations is indicated by N. HOLIDAY is a binary indication (0 = non-holiday, 1 = holiday), while HDD and CDD variables stand for heating and cooling demand.

Daily Energy Consumption Pattern

The trend shows a clear and consistent time-dependent behavior, with energy usage gradually increasing during the early morning hours and reaching its peak between 10 AM and 2 PM. This peak corresponds to active operational hours, where occupancy levels, equipment usage, and cooling demands are at their highest. Following the peak period, energy consumption begins to decline during the afternoon and continues to decrease into the evening hours. The lowest consumption levels are observed during late night and early morning hours, reflecting minimal building activity. This cyclical pattern highlights the strong influence of human activity and operational schedules on energy demand. The observed trend reinforces the importance of incorporating temporal features, such as hour-of-day, lag variables, and cyclical encoding, into the modeling process. Capturing these recurring patterns significantly improves the ability of forecasting models to predict short-term energy consumption accurately.

Figure 1: Average Hourly Energy Consumption Pattern



Source: Author's Computation

Note: Energy consumption peaks during mid-day hours and declines during night, indicating strong time-dependent usage patterns.

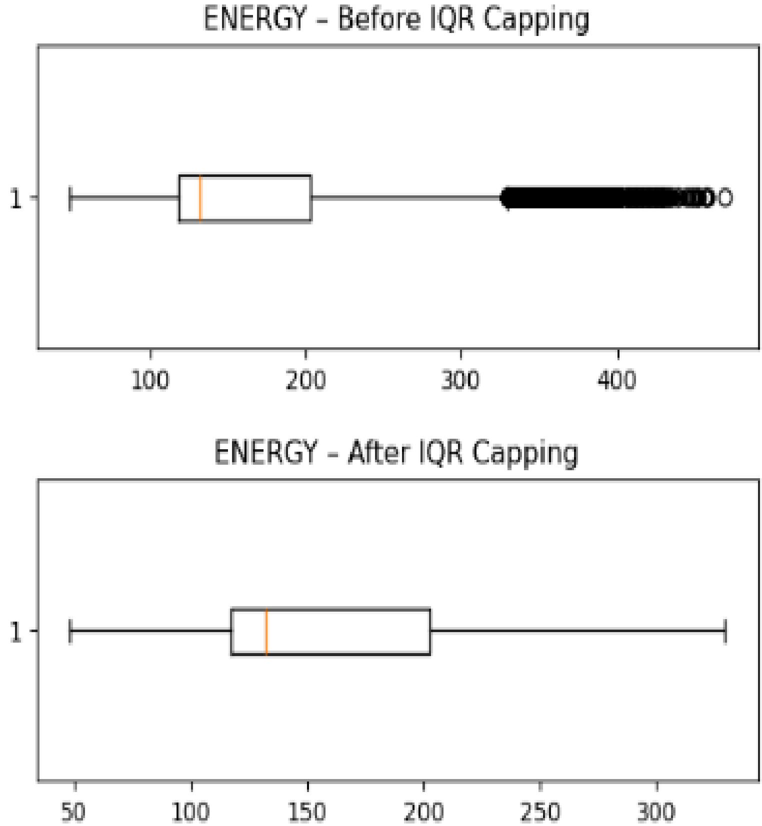
Outlier Detection and Treatment Using the Interquartile Range (IQR) Method

In order to identify extreme data points that needed to be eliminated because they might interfere with the model training and evaluation procedures, the study employed outlier analysis. While all other variables exhibited stable distribution patterns, boxplot analysis revealed that the PRECTOT feature and the ENERGY goal variable had significant outliers. While precipitation can naturally exhibit abrupt spikes as a result of episodic weather occurrences, extreme energy values can be caused by anomalous operational conditions, peak demand events, or measurement irregularities.

The integrity of the dataset time sequence was preserved by reducing extreme data point effects using the Interquartile Range (IQR) approach. By capping observations that deviated from the $Q1-1.5\times IQR$ and $Q3+1.5\times IQR$ limits rather than removing them, the study maintained the time series data structure (Hyndman & Athanasopoulos, 2018).

The treatment process successfully decreased extreme value dispersion in both variables creating a stable distribution pattern that maintained actual consumption and rainfall patterning. The capping method prevented rare spikes from affecting parameter estimation because it restricted their impact on forecast errors which normally occur when extreme values control learning patterns in predictive modeling (Makridakis et al., 1998). The dataset became suitable for robust forecasting through the control of its magnitude which maintained all original distribution patterns

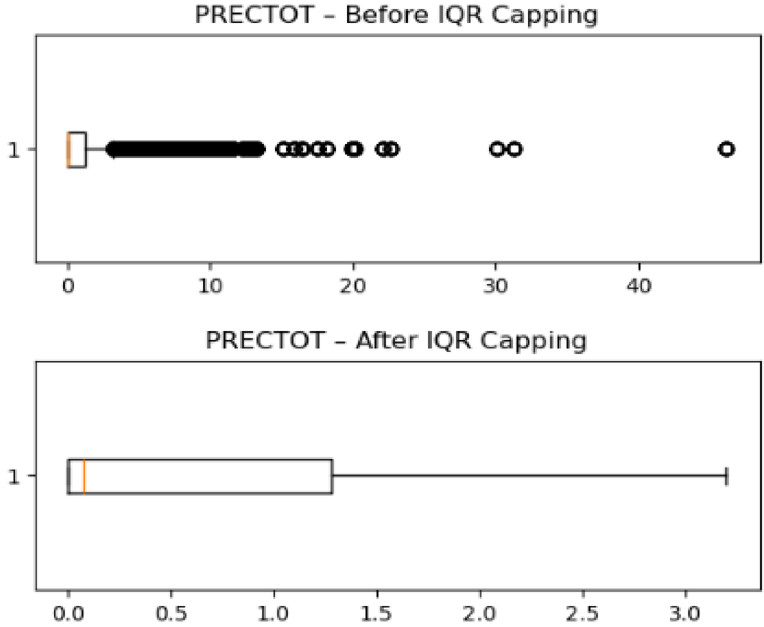
Figure 2: Boxplot of ENERGY before and after IQR-based capping.



Source: Author's computation

Note: The study used the IQR method to detect outliers in hourly energy usage which were then restricted to keep extreme instances from affecting overall consumption patterns.

Figure 3: Boxplot of ENERGY before and after IQR-based capping.



Source: Author's computation

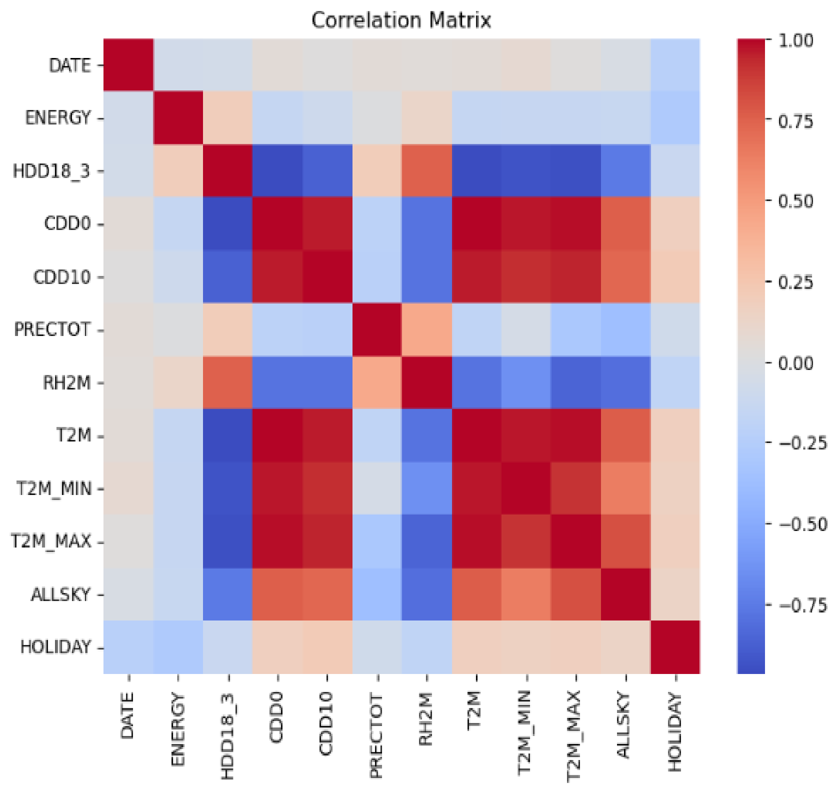
Note: Precipitation values beyond the limit were detected and capped to generate an influence at an unacceptable degree for the estimation in the model.

Correlation Analysis

The Pearson correlation matrix was implemented to study how energy consumption depends on different explanatory variables through their linear relationship. The results show that no single predictor exhibits a strong linear association with hourly energy demand. The results show that consumption patterns are influenced by complex and possibly non-linear interactions. Energy usage shows moderate positive relationships with cooling-related variables which include CDD0, CDD10, T2M, T2M_MIN and T2M_MAX temperature measurements. The research shows that heating degree days (HDD18) have a weaker relationship with electricity consumption. The temperature-related variables show strong correlations because they share the same thermal structure.

The correlation results demonstrate that non-linear models should be applied and structured feature extraction should be implemented because linear models fail to account for all the factors that affect energy usage patterns. The study used multiple approaches to assess the relationship between two variables and found that their relationship only showed a partial connection to energy consumption patterns.

Figure 4: Pearson correlation matrix of energy and environmental variables.



Source: Author's computation

Note: The degree and direction of linear correlations between variables are indicated by correlation coefficients, which vary from -1 to +1.

Feature Engineering and Temporal Enrichment

The structured feature engineering framework was developed to improve modeling proficiency which extracts hourly electricity demand patterns from their fundamental dynamic behavior. The process required us to convert raw timestamp data into useful predictive features because building energy usage patterns exhibit both time-dependent and cyclical characteristics. The chronological order of data was established because time-series forecasting requires proper time sequence to protect against data leakage and maintain sequential data integrity.

Extraction of Temporal Components

The temporal characteristics from the timestamp variable were extracted which provided the hour of the day and the day of the week and the month. The models use these features to detect patterns which repeat throughout the day and different seasons of the year. Building energy consumption patterns establish predictable energy use patterns which research shows are affected by occupant presence and building operations and weather conditions. Energy consumption usually reaches its highest point during business hours which leads to a decrease in usage during the late-night hours. The demand for heating and cooling systems shows seasonal patterns which result from changes in temperature and humidity. The study created a binary weekend indicator which helps to identify standard workdays and days with decreased activity. Commercial buildings show different patterns of energy use during weekends and holidays because they have fewer people present. The model uses categorical temporal distinctions because they help track changes in behavior that environmental variables do not show.

Cyclical Encoding of Time

Temporal variables such as hour and month exhibit cyclical structure rather than linear progression. The transition from hour 23 to hour 0 and the transition from December to January demonstrate natural time progression. The use of simple integer values for these variables creates problems because they disrupt the circular nature of the data. To solve this problem we used sine and cosine transformations for encoding hour and month data in a circular pattern. The process of cyclical encoding maintains the continuous nature of time-based features while stopping the model from creating false breaks during its analysis. The transformation enables both linear and non-linear algorithms to better understand the seasonal and repeating patterns that exist in energy consumption data (Hyndman & Athanasopoulos, 2018).

Lag-Based Memory Features

The Energy consumption pattern displays temporal characteristics because current power requirement depends on electricity usage during the previous period. The lag variables were included for temporal memory into their feature space. The one-hour lag feature enables demand tracking by showing short-term demand patterns which show autocorrelation through its detailed time analysis. The system also includes a 24-hour lag which enables the system to track daily operational patterns that businesses typically follow. Time-series modeling relies on lag features as essential components because they deliver backward state data which enables machine learning models to simulate autoregressive patterns without using dedicated sequential systems.

Rolling Statistical Features

The system performs rolling statistical analysis through its 24-hour time frame for short-term trend detection and variability assessment. The rolling mean function handles recent average consumption assessment while it eliminates temporary demand changes to show current demand levels. The rolling standard deviation function measures short-term consumption changes to determine whether periods of stable or fluctuating consumption are occurring. The use of rolling statistics improves the feature representation because it allows the modeling system to capture local trend patterns which exist within the data. This setup enables tree-based models to make splits based on current trend data while regression methods use evolving contextual information to make predictions.

Feature Refinement and Dimensional Considerations

The feature engineering work led to the decision of removal of the weather variables that appeared redundant because they showed strong interconnections. Strong reciprocal correlations between temperature-related metrics make predictor removal necessary to improve computing efficiency without sacrificing explanatory strength. Rolling statistics, behavioral flags, cyclical temporal encodings, lagged dependencies, and environmental indicators are all included into the final feature set.This structured and enriched representation transforms the original dataset from a raw time-series into a feature-driven predictive framework that can capture energy consumption patterns through non-linear and seasonal and memory-based energy consumption patterns.

Model Development and Evaluation Framework

The research was executed by testing multiple forecasting methods through experiments which used fixed testing conditions to evaluate both traditional forecasting methods and ensemble forecasting methods. The modeling framework included Linear Regression, Decision Tree, Random Forest, and XGBoost. The selected algorithms demonstrate different complexity levels together with different learning abilities which allow researchers to evaluate methods through multiple research categories.

The Linear Regression model was used as the baseline model because its results are easy to understand and it depends on linear relationship patterns. The system provides a basic benchmarking tool which enables users to compare advanced systems against it to determine their actual performance advantages. The Decision Tree model was used to create a system which could identify non-linear feature space boundaries while producing results which experts could easily understand. The use of single-tree models results in systems which display high variability together with extreme sensitivity to the particular features of their training datasets. Ensemble methods were developed to solve existing system limitations. Random Forest uses bootstrap aggregation to combine multiple decision trees, which helps decrease variance but maintains its ability to learn non-linear patterns (Breiman, 2001). XGBoost uses a gradient boosting framework that successively improves prediction errors to increase model accuracy and stability. The ensemble methods show their strongest performance in structured datasets that contain both feature interactions and non-linear relationships.

Training Procedure

The training procedure was conducted using time-aware train–test splits because this method maintained chronological order while stopping any potential information leaks. The method prevents future data from being utilized in model development because this method requires time-series forecasting to maintain actual prediction accuracy (Hyndman & Athanasopoulos 2018).

Each model was fitted using the training set, and its performance was evaluated using a different test set with identical features. With the same feature selection and pre-processing techniques, the system offers testing settings that allow performance evaluation of several algorithms.

Evaluation Metrics

To capture various aspects of predicting accuracy, the model's performance was evaluated using a number of complementing evaluation measures. The overall forecast deviation was measured using Root Mean Square Error (RMSE), which gives larger errors more weight and reflects sensitivity to extreme mispredictions. The average magnitude of absolute discrepancies between actual and anticipated values was measured using Mean Absolute Error (MAE), which offers a simple way to understand normal forecast error. To evaluate the percentage of variance in energy use that each model could account for, the Coefficient of Determination (R^2) was computed, demonstrating the level of explanatory power. Furthermore, by presenting variances in percentage terms, Mean Absolute Percentage Error (MAPE) was incorporated to assess relative accuracy. Because each indication reflects a different facet of prediction performance, using numerous measures guarantees a fair and thorough assessment (Makridakis et al., 1998).

Results and Discussion

The section provides complete testing results for all forecasting models which were tested with data that had not been used before. Four measures assess the model's performance by using RMSE and MAE and R^2 and MAPE to measure absolute error and explanatory power and predictive accuracy. The multi-metric approach (Hyndman & Athanasopoulos) allows multiple performance assessment methods to evaluate forecasting accuracy without relying on one specific measurement system. The research investigates different modeling methods while analyzing their effectiveness in handling time-dependent patterns and non-linear relationships that exist within the time-series dataset. The evaluation uses testing methods which remain neutral because both train and test data share the same distribution and the same preparation methods. The study shows that different modeling methods produce different results for short-term building energy forecasting accuracy through its research.

Performance Evaluation Across Forecasting Models

The predictive performance of the selected models was evaluated using four standard metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Coefficient of Determination (R^2), and Mean Absolute Percentage Error (MAPE). As presented in Table X, ensemble-based models demonstrate superior performance compared to simpler models. XGBoost achieved the highest predictive accuracy, recording the lowest RMSE value of 11.06 and a high R^2 of 0.9744, indicating strong explanatory power. Random Forest exhibited comparable performance, with slightly higher RMSE (11.10) but the lowest MAE (5.70) and MAPE (3.29%), suggesting better consistency in average prediction errors. In contrast, the Decision Tree model showed moderate performance, with higher error metrics (RMSE: 15.22, MAE: 7.76) and lower explanatory power (R^2: 0.9514), reflecting its limitations in capturing complex feature interactions. The Linear Regression model performed the weakest among all models, with the highest RMSE (19.68) and MAPE (7.51%), indicating its inability to effectively model nonlinear relationships present in the dataset.

The overall improvement in model performance can be attributed to the use of a random train– test split, which allows models to learn from a more diverse distribution of data. However, since the dataset is inherently time-dependent, this approach may introduce information leakage, leading to relatively optimistic performance estimates. Therefore, while ensemble models clearly outperform simpler approaches under this framework, the results should be interpreted with caution when considering real-world forecasting applications.

Figure 5: Comparative Performance of Forecasting Models

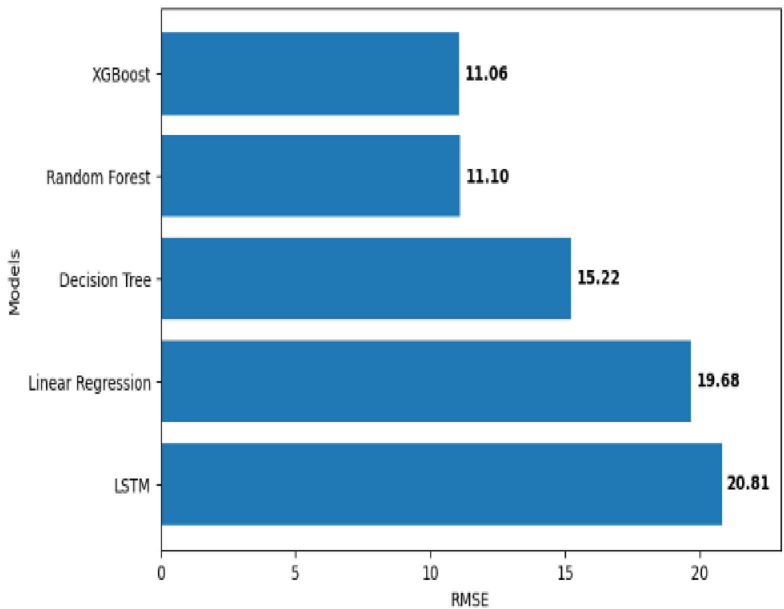
Model	RMSE	MAE	R^2	MAPE (%)
XGBoost	11.06	6.11	0.97	3.57
Random Forest	11.1	5.7	0.97	3.29
Decision Tree	15.22	7.76	0.95	4.45
Linear Regression	19.68	12.2	0.92	7.51
LSTM	20.81	13.39	0.92	9.22

Source: Author's computation
Note: All models use the same train-test splits to report metrics for the test dataset. Better predictive performance is shown by greater R2 and lower error levels.

RMSE-Based Performance Comparison of Forecasting Models

RMSE (Root Mean Square Error) is a widely used metric that penalizes larger prediction errors more heavily, making it particularly suitable for evaluating forecasting performance. As illustrated in the figure, XGBoost and Random Forest exhibit the lowest RMSE values, indicating superior predictive accuracy among the models considered. The difference between these two models is marginal, suggesting that both ensemble methods are highly effective in capturing complex, non-linear relationships in the dataset. In contrast, the Decision Tree model shows a moderate RMSE value, reflecting its limited ability to generalize beyond training data due to its tendency to overfit. The Linear Regression model records the highest RMSE, highlighting its inability to model the non-linear and time dependent patterns present in energy consumption data. Overall, the visual comparison clearly demonstrates that ensemble-based approaches outperform traditional and single-model techniques in terms of minimizing prediction error. It is important to note that the relatively low RMSE values across models can be partially attributed to the use of a random train–test split, which allows models to learn from a broader distribution of data. However, this may also introduce optimistic bias due to potential information leakage in time-series data. Despite this limitation, the chart effectively highlights the comparative performance of different models under a consistent evaluation framework.

Figure 6: RMSE Comparison Across Forecasting Model



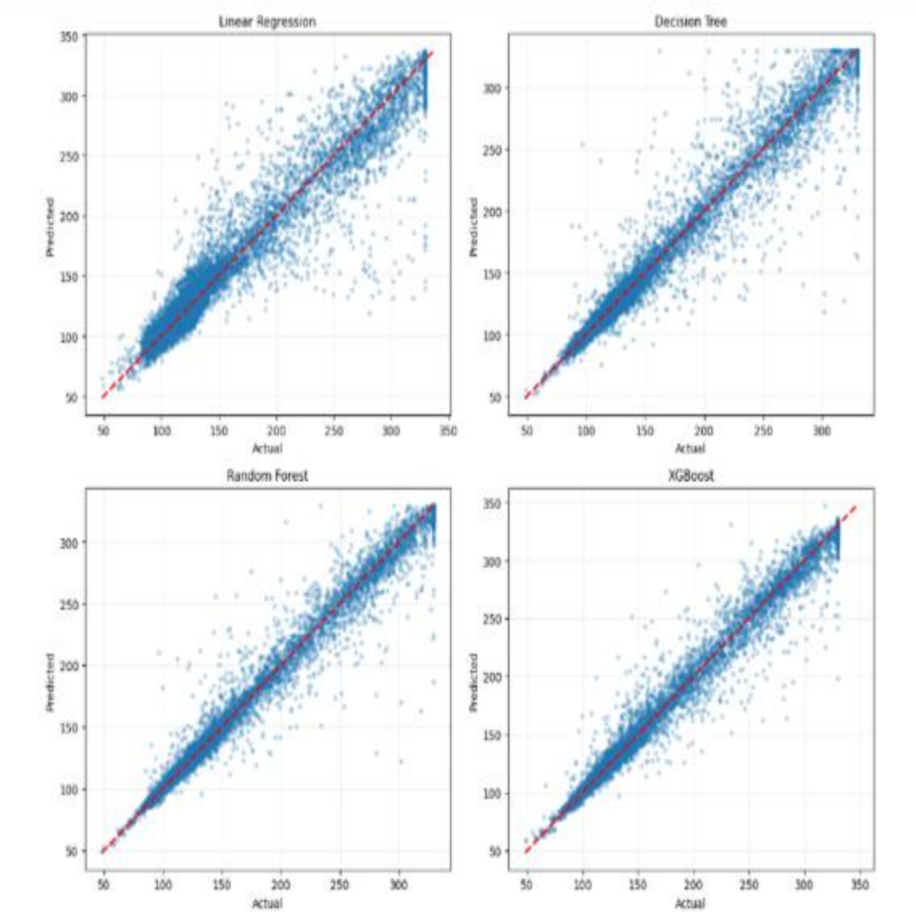
Source: Author's computation
Note: Lower RMSE values indicate better predictive accuracy across the evaluated models.

Actual vs Predicted Comparison Across Forecasting Models

The diagonal reference line represents the ideal scenario where predicted values perfectly match the actual observations. The closer the data points lie to this line, the higher the prediction accuracy of the model. Among the models, XGBoost and Random Forest demonstrate the closest alignment with the ideal diagonal line, indicating superior predictive performance. The data points for these ensemble models are tightly clustered around the reference line, reflecting their ability to effectively capture non-linear relationships and complex interactions within the feature space. This strong alignment is consistent with their high R^2 values and low error metrics observed in the quantitative evaluation. The Decision Tree model shows a relatively wider dispersion of points, particularly at higher energy values, suggesting reduced accuracy and weaker generalization capability. This behavior can be attributed to its tendency to overfit the training data while lacking robustness compared to ensemble methods.

The Linear Regression model exhibits the greatest deviation from the ideal line, with noticeable spread and systematic errors, especially at higher consumption levels. This indicates its limitation in modeling the non-linear and time-dependent patterns inherent in energy consumption data. Overall, the figure visually reinforces that ensemble-based models outperform both traditional linear models and single-tree approaches. However, it is important to consider that the use of a random train–test split may contribute to improved alignment across models, as it allows exposure to a broader data distribution during training. Consequently, while the results demonstrate strong predictive capability, they should be interpreted with caution in the context of real-world time-series forecasting.

Figure 6: Actual vs Predicted Energy Consumption Across Forecasting Models



Source: Author's computation

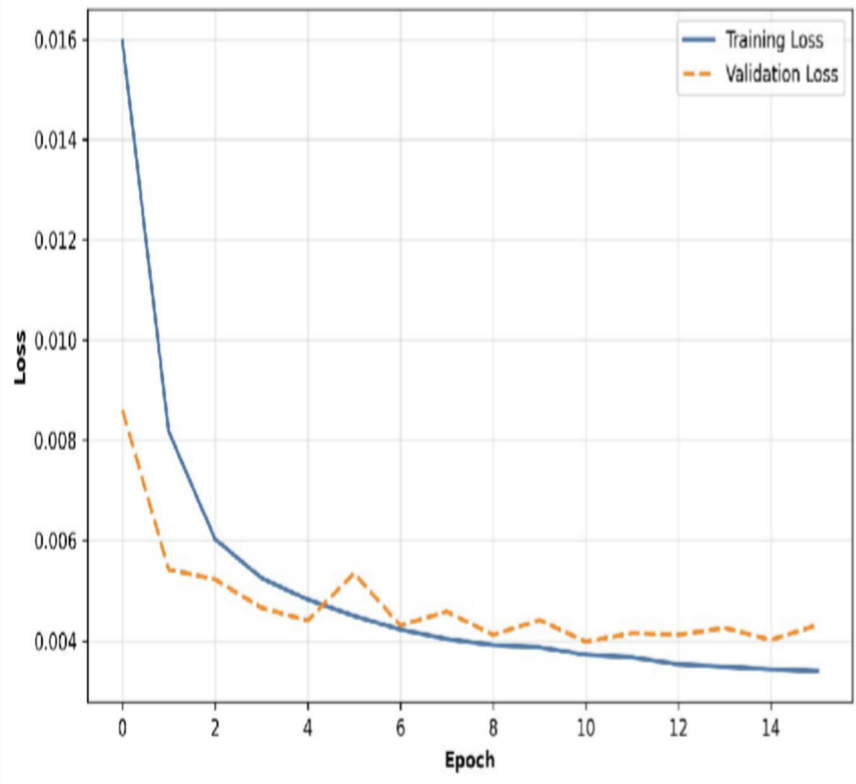
Note: The red dashed line represents perfect prediction. Points closer to the line indicate higher accuracy, with ensemble models showing better alignment than linear models.

LSTM Model Convergence Analysis

The LSTM model shows its learning ability through its training results because the model learns new skills at a steady pace which remains controlled throughout its training time. The loss curves show that both training and validation loss decrease during the first two epochs because the model parameters learn to adapt to the main time patterns which exist in the hourly energy consumption data. The initial decrease shows that the model achieves effective gradient optimization while it rapidly learns short-term energy consumption patterns from the input data. The training process displays a progressive decrease in losses which demonstrates that the model transitions from learning general patterns to acquiring specific temporal relationships. The training loss curve and validation loss curve show close similarity which means the model maintains balanced generalization because the two curves do not diverge enough to prove overfitting. The system achieves controlled learning through its use of dropout layers and early stopping mechanisms which stop the system from learning random patterns and restricts parameter changes after the system achieves stable performance.

The small fluctuations were expected due to validation loss during the recent validation test because time-series forecasting requires handling hourly demand patterns which change throughout the day. The optimization process maintains its steady state because the observed fluctuations only show minor changes which stay within predefined limits. The LSTM network design demonstrates correct architectural implementation because it maintains successful convergence ability. The model learns temporal structure from training data but ensemble methods achieve better prediction results than the model.

Figure 5: Training vs Validation Loss Curve (LSTM)



Source: Author's computation
Note: Training and validation loss are plotted across epochs to assess convergence behavior and potential overfitting.

The findings reveal that not only is model complexity a significant factor in performance, but so is feature representation. Although linear correlations between individual features and energy consumption were low, significant improvement in forecasting effectiveness was achieved through the use of structured feature engineering. The performance of the LSTM model was consistently high and competitive, while the ensemble methods successfully exploited the engineered features for non-linear relationships. This implies that forecasting effectiveness is more dependent on data structure than model complexity. This further emphasizes the need for context-based model selection in building energy forecasting.

CONCLUSION

The research study conducted a formal comparative assessment between traditional machine learning methods and deep learning approaches to predict building energy usage through short-duration time intervals. The study results show that accurate forecasting depends more on how data is presented and how features are developed than on the complexity of the forecasting models. The LSTM architecture successfully captured sequential dependencies while maintaining consistent convergence performance whereas Random Forest and XGBoost ensemble tree-based methods produced superior predictive outcomes across multiple evaluation criteria.

Ensemble modeling methods deliver successful results for modeling non-linear energy consumption patterns which occur in building databases. The study demonstrates that researchers must implement careful data preparation methods and establish specific time-based feature development procedures while following uniform assessment methods to create accurate forecasting models. Organizations achieve improved building resource management through their ability to predict energy consumption with high accuracy which helps them optimize operational scheduling and demand control and resource allocation. Data-driven forecasting systems enable organizations to develop energy management strategies that prioritize sustainability, aligning their practices with environmental aspects of environmental, social, and governance standards.

FUTURE SCOPE

The study shows its experimental design to assess different models yet it requires further investigation study to achieve total development. The analysis requires expansion through testing different building types which exist in various climatic conditions while people use the buildings in distinct ways because this approach will help researchers study how different building designs and user behaviors affect model performance. The assessment of various building types which include residential, commercial, and mixed-use spaces between different environments will help identify specific modeling benefits that work in different contexts.

The extension of the research will investigate additional external factors which include present building occupancy information and changes in prices, renewable energy production and forthcoming weather predictions to improve their ability to predict specific conditions. The combination of ensemble-based feature extraction methods with deep sequential modeling techniques will produce better results because these methods can effectively model both architectural elements and dynamic temporal information. The connection of forecasting systems with real-time building energy management systems enables flexible building control together with system performance upgrades. The developments will boost predictive analytics abilities to help energy decision-making processes which aim for sustainable energy practices and operational effectiveness.

REFERENCES

Azadeh, A. G. (2020). Energy consumption prediction by using machine learning for smart buildings: Case study in Malaysia. *Journal of Building Engineering*, 101-112.

Breiman, L. (2001). Random forests. *Machine Learning*.

Friede, G. B. (2015). ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*.

Ghallabi, F. S. (2025). ESG stock markets and clean energy prices prediction: Insights from advanced machine learning. *International Review of Financial Analysis*.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8).

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice.* OTexts.

Kong, W. D. (2019). Short-term residential load forecasting based on LSTM recurrent neural networks. *IEEE Transactions on Smart Grid*.

Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting: Methods and applications.* John Wiley & Sons.

Marino, D. L. (2016). Building energy load forecasting using deep neural networks. *Energy and Buildings*.

Ürge-Vorsatz, D. C. (2015). Heating and cooling energy trends and drivers in buildings.

Renewable and Sustainable Energy Reviews.

MSE Monographs

* Monograph 36/2017
Underlying Drivers of India's Potential Growth
C.Rangarajan and D.K. Srivastava

* Monograph 37/2018
India: The Need for Good Macro Policies *(4th Dr. Raja J. Chelliah Memorial Lecture)*
Ashok K. Lahiri

* Monograph 38/2018
Finances of Tamil Nadu Government
K R Shanmugam

* Monograph 39/2018
Growth Dynamics of Tamil Nadu Economy
K R Shanmugam

* Monograph 40/2018
Goods and Services Tax: Revenue Implications and RNR for Tamil Nadu
D.K. Srivastava, K.R. Shanmugam

* Monograph 41/2018
Medium Term Macro Econometric Model of the Indian Economy
D.K. Srivastava, K.R. Shanmugam

* Monograph 42/2018
A Macro-Econometric Model of the Indian Economy Based on Quarterly Data
D.K. Srivastava

* Monograph 43/2019
The Evolving GST
Indira Rajaraman

* Monograph 44/2025 Landscape Analysis of the Labour Market of the Freight Logistics Sector in India
Gopal Krishna Roy, Brinda Viswanathan, Ashrita. B, Madhuritha Murali and Mohit Sharma

* Monograph 45/2025 The Fisc and India's Energy Transition
Laveesh Bhandari

MSE Working Papers
Recent Issues

* Working Paper 293/2026

  Early Detection of ESG Policy Violations Using Machine Learning Techniques
  Gautami Parate and Arpita Choudhary

* Working Paper 294/2026

  Stock Market Reactions to COP26 and Climate Change Exposures of Indian Firms
  Saumitra N Bhaduri, Ekta Selarka and Alankrti Aggrwal

* Working Paper 295/2026

  European Union Regulations in Indian Tyre Industry
  Ekta Selarka and Subrata Sarkar

* Working Paper 296/2026

  Women Director Networks and Corporate Social Responsibility of Indian firms
  Kavitha Nambiar and Ekta Selarka

* Working Paper 297/2026
  Does Perception Matter? The Role of Monetary Policy Uncertainty in Policy Transmission
  Aariya Sen

* Working Paper 298/2026

  Reimagining Gender Budgeting Framework in India: Linking Fiscal Outlays and Gendered Outcomes
  N R Bhanumurthy, Bhabesh Hazarika and Aritri Chakravarty

* Working Paper 299/2026

  Explainable Decision Support in Multi-Agent AI Systems Using L-Valued Information Flow and Shapley Aggregation
  Purbita Jana

* Working Paper 300/2026

  Generalised Geometric Logic: A Logic for Expressing Neural Network Architectures
  Ramit Das and Purbita Jana

---

* Working papers are downloadable from MSE website http://www.mse.ac.in