Journal of Engineering and Technology for Industrial Applications

ISSN ONLINE: 2447-0228

ITEGAM-JETIA

Manaus, v.11 n.56, p.136-147. November/December., 2025.
DOI: https://doi.org/10.5935/jetia. v11i56.2725

RESEARCH ARTICLEOPEN ACCESS

ENHANCING SMART GRID MANAGEMENT THROUGH MACHINE LEARNING-BASED HOURLY ENERGY FORECASTING

Atmane Hadji[1], Farid Boumaza[2] and Fatah HADJI[3]

[1]LISI Laboratory, Computer Science Department, University Center A. Boussouf Mila, 43000 Mila, Algeria .
[2]Computer Science Department, University of Mohamed El Bachir El Ibrahimi, Bordj Bou Arreridj 34030, Algeria.
[3]Fundamental Department of Science and Technology , University of Jijel, 18000 Jijel, Algeria .

[1]https://orcid.org/0000-0001-6706-6360, [2]http://orcid.org/0000-0002-9785-420X, [3]https://orcid.org/0009-0007-9343-8341

Email: a.hadji@centre-univ-mila.dz, farid.pgia@gmail.com, fatah.hadji@univ-jijel.dz

ARTICLE INFO

Article History
Received: September 22, 2025.
Revised: October 20, 2025.
Accepted: November 1, 2025.
Published: November 30, 2025.

Keywords:
Accurate forecasting ,
Electricity consumption,
Machine Learning,
Hourly energy consumption.

ABSTRACT

Accurate forecasting of electricity consumption with high temporal resolution is essential for smart grid management, especially in the context of increasing integration of intermittent renewable energy sources such as solar and wind power. This transition poses major challenges for grid operators, who must continuously balance supply and demand under conditions of high variability. Traditional forecasting approaches often fail to capture the nonlinear and dynamic patterns of electricity demand, leading to limited predictive reliability. To address these challenges, this study presents a comparative evaluation of six machine learning models for hourly electricity consumption forecasting: Linear Regression, Random Forest, XGBoost, CatBoost, LightGBM, and Multilayer Perceptron (MLP). Model performance was assessed using RMSE, MAE, R^2, Precision, Recall, and F1-score. The results highlight the superiority of ensemble-based models. Random Forest achieved the best performance (RMSE = 396.56, MAE = 161.35, R^2 = 0.9990, F1 = 0.993), followed by LightGBM (RMSE = 544.88, MAE = 309.30, R^2 = 0.9981, F1 = 0.982) and XGBoost (RMSE = 671.42, MAE = 365.31, R^2 = 0.9972, F1 = 0.979). In contrast, Linear Regression and MLP produced significantly weaker results. Overall, this study demonstrates the potential of advanced machine learning, particularly ensemble methods, to substantially improve the accuracy of energy demand forecasting, thereby enabling more reliable smart grid operation and supporting the global transition toward sustainable energy systems.

I. INTRODUCTION

Electricity consumption is a key indicator for assessing global energy demand, as it directly affects the management of electricity grids and energy planning [1]. With the ongoing shift towards more sustainable energy sources, such as renewable , managing electricity consumption has become increasingly challenging [2]. The factors influencing consumption are numerous and interrelated, including climatic conditions, economic dynamics, technological developments, and social behaviors. For example, electricity demand often peaks during the winter and summer months due to the intensive use of heating or air conditioning systems. Other factors such as economic growth, urban expansion, and the proliferation of networked devices also contribute to increased volatility in energy demand [3].

Accurate forecasting of electricity consumption is critical for several reasons. It ensures a balance between supply and demand, helping to avoid overloads or power outages and plays a pivotal role in integrating unstable renewable energy sources into the grid while maintaining its stability. Effective forecasting also contributes to improving supply strategies and reducing operating costs, as well as reducing carbon emissions by prioritizing renewable energies over fossil fuels [4]. However, forecasting energy consumption faces significant challenges, most notably the complexity of the data, the nonlinear relationships between influencing factors, and the variability in the characteristics of electrical grids between regions. Although statistical models and traditional approaches have been used in the past, they often fail to represent the complex temporal correlations and dynamic patterns within electricity consumption data. This shortcoming has led to growing interest in leveraging machine learning techniques to overcome these challenges and improve prediction accuracy [5]. In this context, this work aims to explore the potential of machine learning models in enhancing the accuracy of energy consumption forecasting. We conduct a comprehensive comparative analysis of several machine learning models, including: linear regression, Random Forest, XGBoost, CatBoost, LightGBM, and multilayer perceptron (MLP), based on hourly electricity consumption data. By evaluating these models in terms of accuracy and generalization ability, we seek to provide a deeper understanding of their strengths and weaknesses when used for high-time-resolution energy forecasting. This study emphasizes the pivotal role of machine learning in addressing modern energy management challenges and highlights its contribution to improving the efficiency of electrical networks and supporting the transition to more sustainable energy systems.?The main objectives of this study are as follows:

1. To enhance forecasting accuracy of hourly electricity consumption by applying advanced machine learning techniques capable of capturing nonlinear and dynamic demand patterns.
2. To conduct a comprehensive comparative evaluation of different machine learning models (Linear Regression, Random Forest, XGBoost, CatBoost, LightGBM, and MLP) in terms of prediction accuracy and reliability.
3. To identify the most effective models for high-resolution forecasting using key performance indicators such as RMSE, MAE, R^2, Precision, Recall, and F1-score.
4. To assess the potential of ensemble methods (Random Forest, XGBoost, LightGBM, CatBoost) in outperforming traditional approaches and single neural models.
5. To provide insights for smart grid management, highlighting how accurate predictions can improve energy demand balancing, operational efficiency, and the integration of renewable energy sources.
6. To contribute to the transition toward sustainable energy systems by demonstrating the role of machine learning in supporting reliable, data-driven energy management strategies.

The remainder of this paper is organized as follows: Section II reviews the related works in the field. Section III outlines the methodologies adopted in this study. Section IV details the experimental methodology employed to implement and evaluate the models. Section V presents and discusses the results along with their evaluation. Finally, Section VI concludes the paper and highlights perspectives for future research.

II. RELATED WORKS

For several decades, electricity consumption forecasting has been a key area of research aimed at improving the efficiency and stability of smart grids (SG). Depending on the time horizon, research is generally classified into four categories: (i) very short-term forecasting (minutes to hours) useful for immediate operational decisions [6], (ii) short-term forecasting (days to a week) applied to daily planning [7], (iii) medium-term forecasting (one week to one year) for resource allocation [8], and (iv) long-term forecasting (beyond one year) for strategic planning and capacity development [9]. Initial approaches exploited artificial neural networks (ANNs) due to their ability to model non-linear behaviour. For example, Kohonen's self-organising maps were applied to forecasting daily consumption in Spain, with an error rate of less than 2.5% [10]. Similarly, polynomial networks combined with ordinary differential equations have shown competitive performance in short-term load forecasting [11]. Subsequent work has focused on integrating meteorological factors and subnetwork decomposition to improve accuracy, combining ANN, ARIMA and grey models (GM) [12]. To further enhance this accuracy, hybrid models incorporating feature selection techniques and meta-heuristic algorithms have been proposed [13].

The use of deep learning has also proven effective, particularly for capturing residential consumption sequences [14], although this involves high computational costs. In the same vein, extending the ARMAX model to Hilbert space has improved the forecasting of functional time series such as electricity prices [15]. From 2018 onwards, the focus shifted to combining optimisation methods and deep learning. For example, approaches combining deep neural networks and optimisation heuristics have achieved significant gains in accuracy [16]. The use of optimised Elman-type networks has also confirmed their effectiveness in forecasting loads in smart grids [17], while online SVM-based methods have been explored for real-time forecasting [18]. At the same time, studies applied to microgrids have shown the impact of forecasting errors on the balance between supply and demand. Comparisons between neural networks, multilayer perceptrons, LSTMs and ensemble prediction networks (EPNs) have shown that the latter significantly reduce forecasting errors and operating costs [19]. New algorithm variants, such as the Extreme Learning Machine (ELM) and its improved kernel-based version (KELM), have demonstrated a greater ability to capture the complexity of electricity prices [20].

More recently, the integration of reinforcement learning for multi-period scheduling of microgrids [21] has reduced costs while speeding up computation. New combinations such as LS-SVM combined with the BCC algorithm [22] have also proven effective in terms of speed and accuracy. For renewable production, ANN models applied to solar forecasting [23] and improved regression methods [24] have increased the reliability of forecasts, while the use of K-means for irradiance has been criticised for its lack of robustness [25]. Finally, several recent approaches explore hybrid and ensemble models. Models combining neural networks, wavelet transformation and simulated annealing (FFANN-WT-SA) have improved one-day forecasting [26], while ensemble techniques such as Gradient Boosted Trees have shown superiority over other methods [27]. The evaluation of multiple algorithms (ARIMA, SARIMA, SVM, XGBoost, RNN, LSTM) confirmed that sequential models (RNN, LSTM, RNN-LSTM) provide the best performance for university campus microgrids [28].In summary, the evolution of the work shows progress: from classical neural models to hybrid methods integrating optimization , deep learning and IoT. Recent approaches highlight a trade-off between increased accuracy and computational costs, while paving the way for integrated solutions for smart grids and microgrids.

III. METHODOLOGIES

In this study, we present an in-depth analysis of the application of machine learning techniques in predicting electricity consumption based on hourly time-accurate data. The analysis included a comparison of several models [20], ranging from traditional methods such as linear regression to more advanced techniques such as Random Forest, XGBoost, CatBoost, LightGBM, and multi-layer perceptron (MLP) neural networks from the field of deep learning [29]. Each algorithm has distinct characteristics and specific advantages that affect its ability to capture complex temporal patterns and improve prediction accuracy. This work highlights the importance of combining traditional and modern approaches to address the challenges associated with the complexity of energy data, while offering a new perspective on developing solutions tailored to the needs of contemporary energy systems, highlighting the strengths and weaknesses of each approach [30].

III.1. MODELS OVERVIEW

This section presents the machine learning algorithms used in the proposed approach [31].

III.1.1 Linear Regression (LR)

Linear regression is a statistical method used to model the relationship between a dependent variable and independent variables on a linear basis. It relies on minimizing the mean square error between actual and predicted values. Despite its simplicity, it is widely used as a baseline in forecasting studies. In the context of energy consumption, this model allows for capturing simple relationships between consumption and factors such as temperature or time of day. Although it is limited in dealing with non-linear relationships, it remains an important option due to its transparency and interpretability.

III.1.2 Random Forest (RF)

Random Forest is an ensemble algorithm that builds multiple decision trees on random samples of data and then aggregates their results to improve robustness and reduce overfitting . In energy consumption forecasting, its strength lies in handling complex relationships and identifying the variables that most influence consumption, which helps to understand the drivers of demand. Its robustness to noise and high-dimensional data also makes it an effective choice.

III.1.3XGBoost

XGBoost is an improved version of Gradient Boosting, where models are built sequentially to correct the errors of previous models. It incorporates advanced techniques such as regularization and parallel processing, giving it high performance with good computational efficiency. Its ability to handle complex relationships and time series data makes it particularly relevant for forecasting hourly consumption, which is characterized by high volatility. Its internal regularization prevents overfitting and enhances the reliability of forecasts.

III.1.4 CatBoost

Designed to handle categorical variables efficiently, CatBoost is a stepwise boosting algorithm that provides advanced regularization techniques to reduce overfitting and produces accurate and fast forecasts. For hourly data that includes categorical variables (such as days of the week), CatBoost provides accurate modeling with minimal preprocessing, which is an important advantage in this domain.

III.1.5 LightGBM

LightGBM is a fast and efficient gradient boosting model that uses a leaf-wise growth strategy to improve accuracy when dealing with large and complex datasets. This model is particularly effective with large-scale time series data, combining speed and high accuracy with a good ability to handle unbalanced data.

III.1.6 Multi-Layer Perceptron (MLP)

MLP is a type of artificial neural network consisting of several fully interconnected layers. Thanks to the use of nonlinear activation functions, it can capture complex relationships in data, making it suitable for multiple tasks, including time series forecasting. In the context of energy, it can model complex patterns such as consumption peaks. Its flexibility and ability to learn deep relationships make it a promising choice for dealing with dynamic and complex data.

III.1.7 Model selection criteria

The selection of these models was based on rigorous scientific criteria [32]. , the most important of which are :

1. Handling linear and nonlinear relationships: Combining linear (e.g., linear regression) and nonlinear (e.g., RF and XGBoost) models to achieve comprehensive analysis.
2. Performance with time series: Models such as XGBoost and LightGBM have shown superiority in handling complex hourly data.
3. Robustness and efficiency: Ensemble models such as RF, CatBoost, and LightGBM offer high accuracy while reducing the risk of overfitting.

4. Flexibility in data processing: CatBoost excels with categorical data, while MLP offers the ability to capture complex dynamic relationships.
5. This methodological diversity provides a solid basis for comprehensive comparison and ensures interpretable and reliable results, contributing to a more accurate understanding of changes in energy consumption and more effective forecasting.

III.2 PERFORMANCE EVALUATION MEASURES

Evaluating the performance of machine learning models is crucial for assessing their effectiveness and ability to make accurate predictions. Several metrics are commonly used to evaluate prediction accuracy and model quality. Among these are **RMSE** (Root Mean Square Error), **MAE** (Mean Absolute Error), and **MAPE** (Mean Absolute Percentage Error).

IV.2.1 RMSE (Root Mean Square Error)

RMSE is a widely used metric for measuring the difference between predicted values and actual values. It penalizes larger prediction errors more heavily due to its squared error approach. The RMSE measures the average squared difference between predicted and actual values

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(yi-\hat{y}i)}^{2} \tag{1}$$

- y_i: Actual values.
- $\hat{y}_i$: Predicted values.
- n: Total number of observations.

IV.2.2. MAE (Mean Absolute Error)

MAE measures the average of the absolute errors between predicted values and actual values. Unlike RMSE, it does not disproportionately penalize large errors, making it more robust against outliers. MAE provides a simpler and less outlier-sensitive estimate of error, which is useful when it is important to treat all errors equally. The MAE calculates the average absolute differences between predicted and actual values.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | \, yi - \hat{y}i \, | \tag{2}$$

- $|y_i-\hat{y}_i|$: Absolute difference between actual and predicted values.

IV.2.3. Coefficient of Determination (R²)

The **Coefficient of Determination (R^2)** is a statistical metric used to assess how well a regression model explains the variance in the dependent variable (target). It indicates the proportion of the total variance in the actual data that is captured by the model's predictions. The R^2 indicates the proportion of variance in the data explained by the model.

$$R^2 = \frac{\frac{1}{n}\sum_{i=1}^{n}(yi-\hat{y}i)^2}{\frac{1}{n}\sum_{i=1}^{n}(yi-\bar{y})^2} \tag{3}$$

- y_i: Actual values.
- $\hat{y}_i$:Predicted values.
- $\bar{y}$: Mean of actual values.
- R^2: ranges from 0 (no variance explained) to 1 (all variance explained).

To assess and compare the performance of our models, we employ the standard evaluation metrics **precision**, **recall**, and the **F1-score**. Precision reflects the proportion of correctly predicted positive instances among all instances labeled as positive by the model, thus measuring the reliability of predictions. Recall quantifies the proportion of correctly predicted positives relative to all actual positives, highlighting the model's capacity to capture relevant cases. The F1-score, defined as the harmonic mean of precision and recall, provides a balanced indicator that accounts for both accuracy and completeness [33]. As emphasized the precision corresponds to the ratio of true positives to the total number of predicted positives.

$$\text{Precision} = \frac{\text{Correct}}{\text{Correct} + \text{Spurious}} \tag{4}$$

Recall is the fraction of the valid annotations over the total amount of annotations. It is formally defined as:

$$\text{Recall} = \frac{\text{Correct}}{\text{Correct} + \text{Missing}} \tag{5}$$

F-measure is defined as the harmonic mean of two factors, precision and recall. It is formally as:

$$F1 = \frac{2 * \text{Precision} * \text{Recall}}{(\text{Precision} + \text{Recall})} \quad (6)$$

IV. EXPERIMENTAL METHODOLOGY

IV.1 DATASET DESCRIPTION

The dataset employed in this study originates from the ENTSO-E[1] (European Network of Transmission System Operators for Electricity) platform, which serves as a key reference source for collecting, aggregating, and disseminating data related to electricity generation and consumption across Europe. Since 2019, the statistics published by ENTSO-E have been derived from the Transparency Platform, in compliance with the Transparency Regulation, ensuring standardized formats and near real-time availability of information. Although the primary purpose of these data is not statistical reporting, their reliability, temporal granularity, and wide geographical coverage make them an indispensable resource for modeling and analyzing energy demand. For the purpose of this study, we selected the dataset corresponding to the year 2024, specifically the Monthly Hourly Load Values, which provide hourly aggregated electricity load values by European country.

Each record in the dataset corresponds to a consumption measurement for a given hour, allowing for a detailed analysis of temporal dynamics in electricity demand, including daily, weekly, and seasonal variations. The dataset is structured into several well-defined fields (Table 1). This homogeneous structure facilitates straightforward processing and comparative analyses. The strength of this dataset lies in its hourly granularity combined with its European scope, which enables researchers to examine short-term variability in electricity demand while also capturing structural trends over longer periods. Moreover, by aggregating loads at the national level, ENTSO-E provides a consolidated basis for cross-country comparisons of energy consumption patterns, the influence of climatic conditions, economic cycles, and the effects of energy policies.

These characteristics make the Monthly Hourly Load Values 2024 dataset a central resource for studies focused on electricity demand forecasting, smart grid optimization, and resilience assessment of energy systems at the continental scale. This study utilized a dataset of approximately 25,000 hourly records of electricity consumption, providing a detailed representation of load fluctuations and enabling the analysis of consumption patterns over time. The dataset is particularly valuable due to several characteristics: its relatively large size offers a robust foundation for experimental evaluation; its fine-grained temporal resolution (hourly) makes it highly suitable for short-term forecasting; its broad geographical coverage allows for cross-regional comparisons of consumption behaviors; and the inclusion of timestamp values contributes to model stability and enhances forecasting accuracy. The dataset consists of several key attributes, as follows:

Table 1: Dataset atributes.

Column Name	Example	Meaning	Use in Modeling
MeasureItem	Monthly Hourly Load Values	Type of measurement (here: hourly load values aggregated monthly)	Metadata, useful for filtering measurement types
DateUTC	01-01-2024 00:00	Exact date and time of the measurement (UTC)	Used to create time-related features (hour, day, month, season, weekday)
DateShort	01-01-2024	Simplified date without time	Enables daily analysis or aggregation
TimeFrom	00:00	Start time of the hourly interval	Defines the measurement window
TimeTo	01:00	End time of the hourly interval	Defines the measurement window
CountryCode	AL	ISO country code (e.g., Albania)	Categorical variable (requires encoding if multiple countries are included)
Cov_ratio	100	Data coverage ratio (%)	Useful for assessing data quality and filtering incomplete records
Value	731	Recorded electrical load in MW	Main target variable
Value_ScaleTo100	731	Scaled value (0–100) or same as Value	Optional, useful for normalization and stable learning
CreateDate	06-12-2024 14:36:48	Record creation timestamp	Metadata (traceability), usually not used in modeling
UpdateDate	06-12-2024 14:36:48	Last update timestamp	Metadata (data freshness), usually not used in modeling

Source: Authors, (2025).

IV.2 DATASET PREPARATION AND PREPROCESSING

In this study, the dataset (year 2024, approximately 2,500 records) was carefully preprocessed to ensure data quality, consistency, and suitability for machine learning applications. The preparation process involved multiple stages:

1. Temporal conversion and feature extraction: The DateUTC column was transformed into a standard datetime format, from which temporal attributes were derived, namely hour, day, month, and weekday. These variables capture short- and medium-term temporal dynamics of electricity consumption.

[1] https://www.entsoe.eu/data/power-stats/

2. Categorical encoding: The variable CountryCode, representing geographical regions, was encoded into numerical values using categorical encoding techniques. This transformation allowed the integration of categorical information into models that require numeric inputs.
3. Definition of target and predictors: The target variable is energy consumption (Value), while the selected predictors are [hour, day, month, weekday, CountryCode]. This feature set balances temporal resolution with geographical diversity, ensuring interpretability and predictive strength.
4. Train/test split: To enable robust evaluation, the dataset was partitioned into 70% training and 30% testing sets, maintaining temporal consistency to avoid information leakage.
5. Normalization and scaling: Specific preprocessing was tailored to the requirements of the algorithms:For distance- and gradient-based models (e.g., Linear Regression, Multi-Layer Perceptron (MLP), Long Short-Term Memory (LSTM)), normalization or standardization of input features was applied to improve numerical stability, convergence speed, and overall model accuracy. For tree-based models (Random Forest, XGBoost, CatBoost, LightGBM), feature scaling was not strictly necessary, as these methods rely on decision rules and are invariant to monotonic transformations of the input space.

This preprocessing pipeline ensured that the dataset was both statistically consistent and adapted to the specific conditions of each algorithm, thereby improving the reliability of the subsequent experimental results.

IV.3 IMPLIMENTATION

The proposed application follows a systematic pipeline designed to forecast electricity consumption based on hourly load values. The process (see Figure 1) begins with data preprocessing, which ensures the reliability and consistency of the dataset through the handling of missing values, removal of invalid entries, feature engineering, standardization, and splitting into training and testing subsets. Subsequently, different families of models are trained, encompassing traditional machine learning approaches (Random Forest, Linear Regression), gradient boosting techniques (CatBoost, XGBoost), and a deep learning model (MLP Classifier). This diversified modeling strategy enables a comprehensive evaluation of predictive capabilities across distinct paradigms of learning. Finally, the trained models are assessed using a set of well-established evaluation metrics, including Precision, Recall, F1-Score, RMSE, MAE, and R^2, complemented by confusion matrices and a comparative performance analysis. Such a methodology not only ensures rigorous validation of the models but also provides valuable insights into their respective strengths and weaknesses, thereby contributing to a robust framework for accurate short-term electricity demand forecasting.

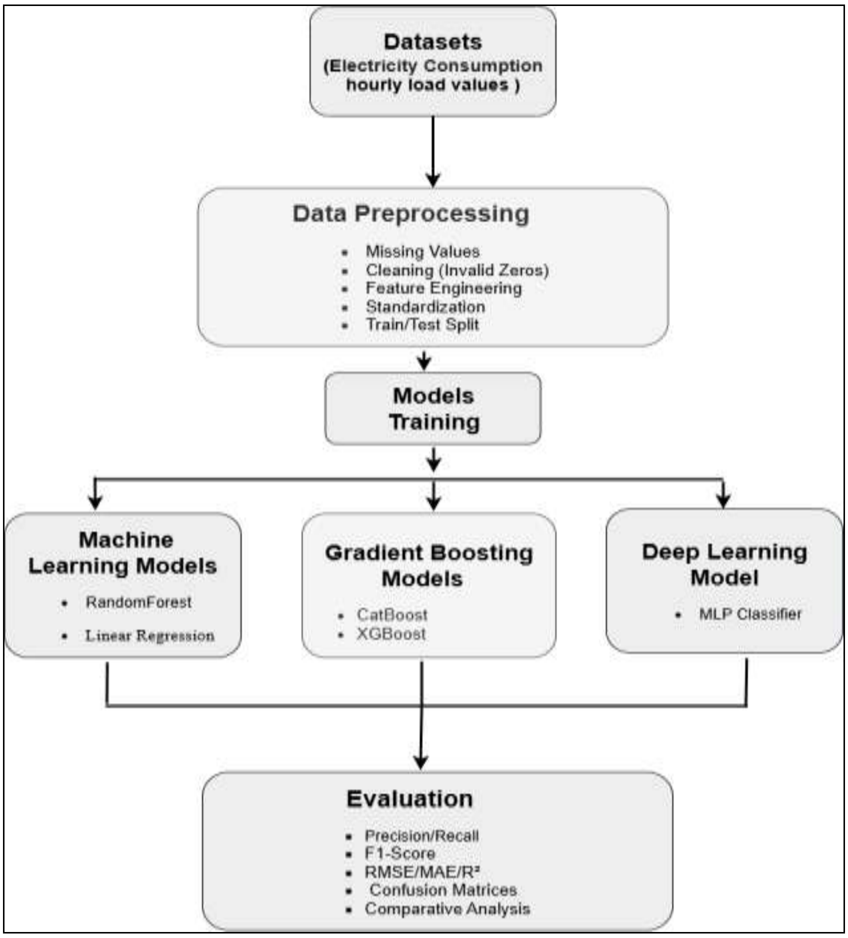

Figure 1: Comparison of Predicted and Actual Values with Linear Regression.
Source: Authors, (2025).

Table 2: Algorithm Conditions, Preprocessing, and Key Parameters.

Algorithm	Conditions	Preprocessing Required	Key Parameters (from code)	Remarks
Linear Regression	Assumes linear relationship between features and target. Sensitive to outliers and collinearity.	Standardization recommended. Outlier handling desirable.	Default sklearn parameters (no regularization applied).	Interpretable baseline, but limited for nonlinear data.
Random Forest	Ensemble of decision trees, robust to outliers and scaling.	No normalization required. Only categorical encoding.	n_estimators=100, random_state=42.	Captures nonlinearities, provides feature importance.
XGBoost	Gradient boosting with regularization to prevent overfitting.	No scaling needed. Encoded categorical features accepted.	n_estimators=100, objective=reg:squarederror, random_state=42.	Strong performance on tabular data, needs tuning (learning rate, depth).
CatBoost	Gradient boosting optimized for categorical features. Reduces overfitting with ordered boosting.	Manual encoding used here (CatBoost can handle categories natively). Scaling unnecessary.	iterations=500, learning_rate=0.1, depth=6, random_state=42.	Handles categorical data efficiently, less preprocessing effort.
LightGBM	Gradient boosting framework optimized for large-scale data.	No scaling needed. Works with categorical and numerical features.	n_estimators=500, learning_rate=0.1, random_state=42.	Faster than XGBoost, but can be sensitive on small datasets.
MLP (Neural Network)	Sensitive to feature scaling. Requires dense input without missing values.	Normalization mandatory (e.g., StandardScaler).	hidden_layer_sizes=(100, 50) → 2 hidden layers, activation='relu', solver='adam', max_iter=500.	Captures complex nonlinearities; longer training; risk of overfitting on small data.

Source: Authors, (2025).

V. RESULTS AND EVALUATION

V.1 RESULTS

To assess the predictive performance of the different machine learning algorithms applied to our dataset, we relied on three commonly used evaluation metrics: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the Coefficient of Determination (R^2). These indicators respectively measure the average deviation between predicted and actual values, the average absolute prediction error, and the proportion of variance explained by the model. The results obtained for each of the six algorithms considered (Linear Regression, Random Forest, XGBoost, CatBoost, LightGBM, and Multi-Layer Perceptron) are summarized in Table 3, highlighting the performance differences across the approaches.

Table 3: Results of perfermance models.

Models	RMSE	MAE	R^2
Linear Regression	12504.01	8575.82	0.0217
Random Forest	396.56	161.35	0.9990
XGBoost	671.42	365.31	0.9972
CatBoost	911.29	503.25	0.9948
LightGBM	544.88	309.30	0.9981
MLP	6014.80	4103.33	0.7736

Source: Authors, (2025).

In this study we applied ten-fold cross-validation (see Table 4) to ensure a robust assessment of model performance and generalization. Unlike a simple train/test split, this method reduces the influence of random data partitioning and provides a more reliable estimate of predictive accuracy. The results show strong consistency across folds, which highlights the quality and homogeneity of the dataset. Such stability indicates that the data are representative and well-structured, allowing for reliable evaluation without major imbalances or extreme fluctuations. Therefore, cross-validation not only strengthens methodological rigor but also serves as an indirect indicator of data quality. By confirming that the reported results are stable and reproducible, it ensures that the findings are genuinely generalizable and scientifically valid.Table 2: Result of model performance.

Table 4: Results of perfermance for cross-validation (10 folds)

Models	RMSE	MAE	R^2
Linear Regression	12504.01	8575.81	0.0224
Random Forest	396.55	161.35	0.9992
XGBoost	671.41	365.30	0.9971
CatBoost	911.29	503.25	0.9949
LightGBM	544.88	309.29	0.9982
MLP	6014.80	4103.33	0.7736

Source: Authors, (2025).

Table 5: Results of perfermance models Precision, Recall and F1.

Models	Precision	Recall	F1
Linear Regression	0.502	0.999	0.669
Random Forest	0.993	0.994	0.993
XGBoost	0.977	0.982	0.979
CatBoost	0.964	0.980	0.972
LightGBM	0.981	0.982	0.982
MLP	0.687	0.919	0.786

Source: Authors, (2025).

The Table 5 summarizes the results of the six evaluated machine learning algorithms in terms of Precision, Recall, and F1-score. These metrics allow us to assess both the correctness and completeness of the predictions, as well as the overall balance between them. The correlation matrix presented in Figure 2 provides an overview of the relationships between temporal features (hour, day, month, and weekday), the country code, and electricity consumption (Value). Overall, the results indicate very weak correlations between the consumption variable and most temporal features, suggesting that electricity demand patterns are highly nonlinear and cannot be explained by simple linear dependencies.

The negative correlation observed between consumption and the country code (-0.14) highlights some regional variability, but its low magnitude confirms that additional contextual and external variables (e.g., weather conditions, socio-economic factors, or seasonal events) are necessary to capture meaningful trends. Moreover, the absence of strong correlations among temporal variables themselves indicates that these predictors carry complementary rather than redundant information, justifying their inclusion in machine learning models. This analysis underscores the complexity of electricity consumption dynamics and reinforces the relevance of advanced predictive approaches, such as ensemble methods and deep learning, which are better suited to capture hidden nonlinearities and interactions between influencing factors.

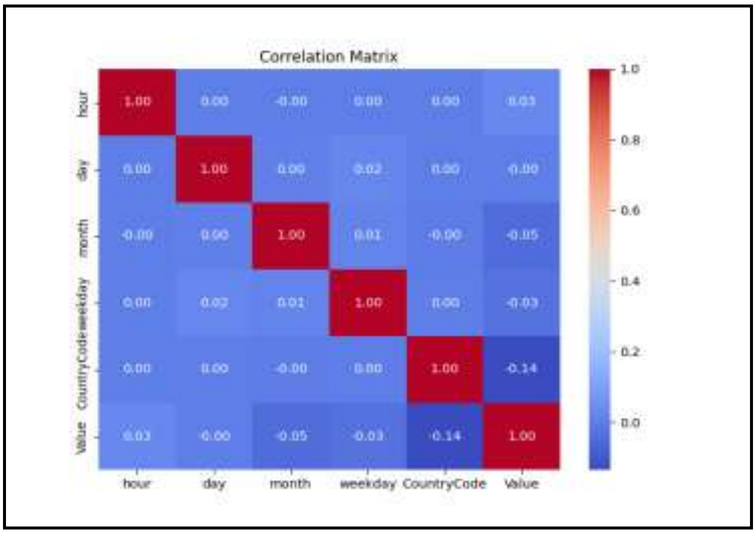

Figure 2: Correlation matrix.
Source: Authors, (2025)

V.2 ANALYSIS AND DISCUSSION

The study on Machine Learning for Energy Consumption Forecasting using hourly data provided a comparative analysis of six algorithms: Linear Regression, Random Forest, XGBoost, CatBoost, LightGBM, and Multi-Layer Perceptron (MLP). Below is an in-depth analysis and discussion of the results based on the evaluation metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2).

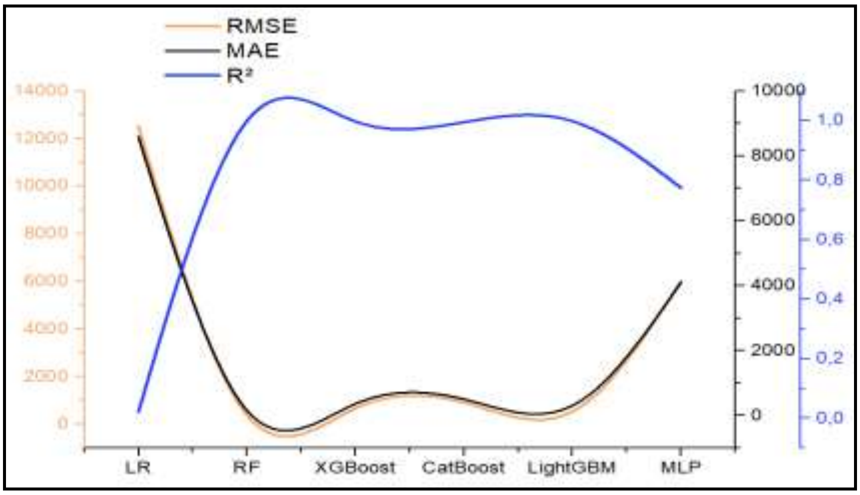

Figure 3: Evaluation of RMSE, MAE and R2 metrics of Models.
Source: Authors, (2 025).

This study provides a comprehensive comparative analysis of six machine learning models—Linear Regression, Random Forest, XGBoost, CatBoost, LightGBM, and Multi-Layer Perceptron (MLP)—for hourly electricity consumption forecasting, evaluated using RMSE, MAE, and R^2. Results indicate that Linear Regression, despite its simplicity and interpretability, performed the worst, failing to capture the nonlinear and dynamic patterns in the data. Among all models, Random Forest achieved the highest accuracy, with the lowest RMSE and MAE and an R^2 close to 1, confirming its robustness and strong capability to model complex temporal dependencies.

LightGBM also demonstrated excellent performance, outperforming XGBoost and CatBoost thanks to its efficient leaf-wise growth strategy, which balances speed, precision, and scalability, making it well-suited for real-time applications. XGBoost, while slightly less accurate than Random Forest, remained a strong candidate due to its computational efficiency and regularization features, particularly for large-scale tasks. CatBoost, although slightly less precise than the other ensemble methods, showed competitive results and stands out for its ability to natively handle categorical data, reducing preprocessing complexity. On the other hand, the MLP achieved only moderate performance, with higher RMSE and MAE and a lower R^2 (0.77), suggesting difficulties in capturing nonlinear temporal dependencies, likely due to limited tuning or insufficient feature engineering. Overall, ensemble methods clearly outperformed both linear and neural approaches, highlighting their suitability for structured, tabular datasets such as energy consumption. In practice, Random Forest is recommended for applications requiring interpretability and robustness (e.g., grid optimization), LightGBM for fast and scalable real-time forecasting, XGBoost for large-scale tasks where efficiency is critical, and CatBoost for datasets involving categorical variables.

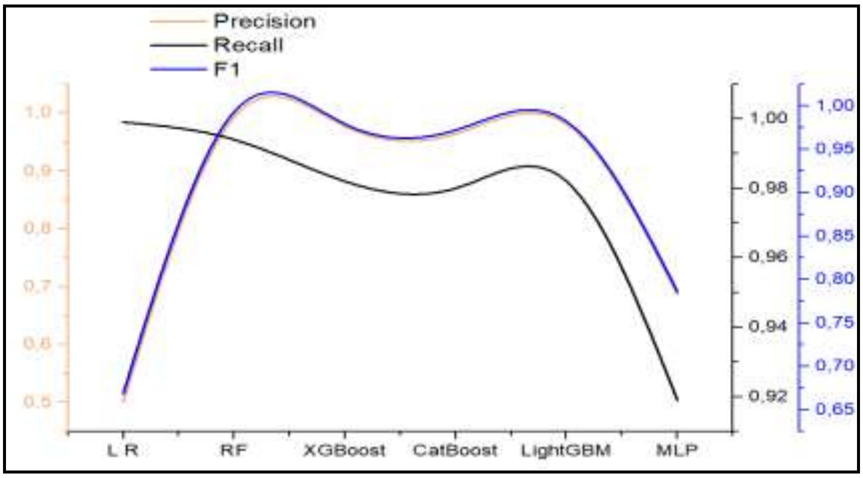

Figure 4: Evaluation of Precision, Recall and F1 mesures of Models.
Source: Authors, (2025).

The evaluation of the six machine learning models (Figure 4) using precision, recall, and F1-score highlights significant differences in their ability to balance accuracy and completeness. Linear Regression, although achieving an almost perfect recall (0.999), suffers from very low precision (0.502), indicating a high proportion of false positives and resulting in a modest F1-score (0.669). This confirms its inadequacy for this type of task due to its simplicity and inability to capture complex relationships. In contrast, Random Forest demonstrates remarkably balanced results (Precision = 0.993, Recall = 0.994, F1 = 0.993), reflecting the robustness of ensemble methods and their capacity to reduce variance through the aggregation of decision trees. Boosting algorithms, namely XGBoost (F1 = 0.979) and LightGBM (F1 = 0.982), also confirm their effectiveness by combining high precision and recall thanks to advanced optimization mechanisms and efficient handling of imbalanced data, with LightGBM further standing out for its computational efficiency.

CatBoost, with a competitive F1-score (0.972), occupies an intermediate po sition, offering good coverage but slightly lower precision, which reflects a relatively higher number of false positives despite its strength in managing categorical variables. Finally, the Multi-Layer Perceptron achieves more modest results (F1 = 0.786), with satisfactory recall (0.919) but limited precision (0.687), indicating a tendency to overestimate positive cases and a higher dependency on fine-tuning hyperparameters and data quality. Overall, this analysis confirms the superiority of ensemble-based methods relying on bagging and boosting, which provide an optimal trade-off between precision and recall, resulting in F1-scores close to 1. These approaches thus emerge as the most robust and effective for this specific problem, with Random Forest, XGBoost, and LightGBM standing out as particularly relevant choices that combine stability, efficiency, and strong generalization ability.

Each graph corresponding to a model illustrates the comparison between actual data and predicted values.

- **X-axis (Index)**: represents the data points, such as hours or time steps. In this case, approximately 100 samples are shown.
- **Y-axis (Energy Consumption)**: indicates the energy consumption, expressed in either actual or predicted values.
- **Blue curve (Actual Values)**: shows the real energy consumption, characterized by high variability and the presence of significant peaks, sometimes exceeding 50,000 units.
- **Orange curve (Predicted Values)**: represents the values estimated by the model. These predictions appear much smoother and fail to closely follow the peaks and abrupt variations in the actual data, highlighting the model's limitations in capturing sudden fluctuations in energy demand.

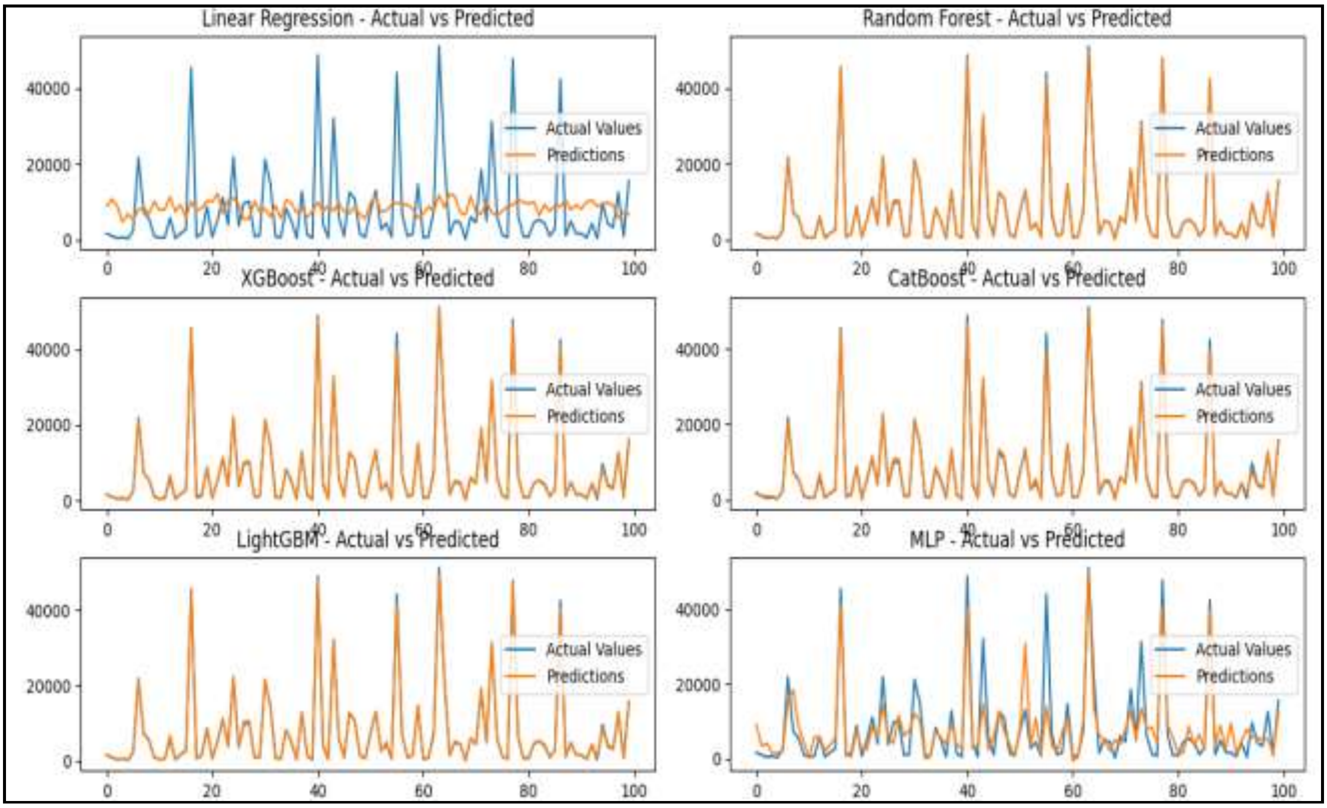

Figure 5: Comparison of Predicted and Actual Values with Six Models.
Source: Authors, (2025).

The Figure 5 illustrates the comparison between actual energy consumption values and values predicted by different machine learning models, including Linear Regression, Random Forest, XGBoost, CatBoost, LightGBM, and MLP. Firstly, we can see that linear regression fails to reproduce the variability of the data, with almost constant predictions that ignore consumption peaks, confirming its inability to capture the complex non-linear relationships present in the time series. On the other hand, ensemble models such as Random Forest, XGBoost, CatBoost and LightGBM show significantly better performance: their prediction curves closely follow the actual values, even during extreme fluctuations exceeding 40,000 units. These methods exploit the combination of decision trees and boosting strategies, which allows them to better model non-linearities and reduce prediction errors. The MLP model, based on neural networks, also shows good adaptability to dynamic variations in consumption, but with slightly lower accuracy than gradient boosting models, probably due to the complexity of the hyperparameters to be optimised. Overall, the results confirm the superiority of advanced models, particularly XGBoost, CatBoost and LightGBM, which offer excellent correspondence with actual data and appear to be promising solutions for forecasting hourly energy consumption in smart grids.

Table 5: Comparison of Machine Learning Models for Hourly Energy Consumption Forecasting.

Model	Strengths	Limitations
Linear Regression (LR)	Simple, interpretable, fast to train. Serves as a baseline reference.	Cannot capture nonlinear relationships. Poor performance on complex and volatile data.
Random Forest (RF)	Robust to noisy data, good accuracy, reduces overfitting, identifies key variables.	Less effective than boosting methods on highly dynamic data. Can be slower on large datasets.
XGBoost	High accuracy, efficient handling of nonlinearities, built-in regularization, strong performance on time series.	Sensitive to hyperparameters (complex tuning). Training can be time-consuming.
CatBoost	Handles categorical variables automatically, strong regularization, high accuracy without heavy preprocessing.	Relatively new, less documented than XGBoost/LightGBM. May be sensitive to parameter selection.
LightGBM	Very fast and efficient on large datasets, excellent accuracy, good handling of imbalanced data.	Risk of overfitting if not properly regularized. Less robust than CatBoost for small datasets.
Multi-Layer Perceptron (MLP)	Captures very complex and dynamic relationships, flexible and adaptable to diverse contexts.	Requires careful hyperparameter tuning, prone to overfitting, demands high computational resources.

Source: Authors, (2025).

V. CONCLUSIONS

In this study, we have investigated the application of various machine learning techniques for the accurate prediction of electricity consumption, particularly at a high temporal resolution (hourly data). By comparing models such as Linear Regression, Random Forest, XGBoost, CatBoost, LightGBM, and Multi-Layer Perceptron (MLP), we have demonstrated the significant advantages of machine learning approaches over traditional statistical methods in forecasting energy demand. Our results show that ensemble models like Random Forest, LightGBM, and XGBoost outperform other techniques in terms of accuracy, with Random Forest exhibiting the best performance, achieving near-perfect predictions. The findings highlight the importance of leveraging machine learning for energy forecasting, particularly given the challenges posed by the integration of renewable energy sources, which are inherently variable. The superior predictive power of models like Random Forest and LightGBM ensures that grid operators can better balance supply and demand, leading to more stable and efficient energy systems.

Despite the promising results, it is important to note that challenges remain, including data complexity, model interpretability, and the need for further refinement of predictive algorithms to account for regional and sectoral variations in energy consumption. Future work should explore hybrid models, incorporating external factors such as weather conditions, economic changes, and social behaviors, which can further enhance prediction accuracy. Ultimately, this research underscores the potential of machine learning in advancing the field of energy management, contributing to more sustainable and cost-efficient energy systems. The integration of these advanced forecasting techniques will play a critical role in addressing the growing demand for energy while supporting the global transition to cleaner, renewable energy sources.

VI. AUTHOR'S CONTRIBUTION

Conceptualization: Atmane Hadji, Farid Boumaza and Fatah Hadji.
Methodology: Atmane Hadji, Farid Boumaza.
Investigation: Atmane Hadji, Farid Boumaza.
Discussion of results: Atmane Hadji, Farid Boumaza and Fatah Hadji.
Writing – Original Draft: Atmane Hadji.
Writing – Review and Editing: Atmane Hadji, Farid Boumaza.
Resources: Atmane Hadji .
Supervision: Atmane Hadji.
Approval of the final text: Atmane Hadji, Farid Boumaza and Fatah Hadji.

VIII. REFERENCES

[1] L. Lin, R. Wijayawardana, V. Rao, H. Nguyen, E. W. Gnibga, and A. A. Chien, "Exploding AI power use: An opportunity to rethink grid planning and management," in Proc. 15th ACM Int. Conf. Future and Sustainable Energy Systems, pp. 434–441, Jun. 2024. https://arxiv.org/pdf/2311.11645

[2] K. Misiurek, T. Olkuski, and J. Zyśk, "Review of methods and models for forecasting electricity consumption," *Energies*, vol. 18, no. 15, p. 4032, 2025.https://www.mdpi.com/1996-1073/18/15/4032

[3] Y. Zou, C. Wang, H. Najam, A. A. Nassani, G. Djuraeva, and D. Oscar, "Energy consumption prediction for households in a society with an ageing population," *Energy Strategy Reviews*, vol. 57, p. 101622, 2025. https://www.sciencedirect.com/science/article/pii/S2211467X24003316

[4] X. Dong, Y. Luo, S. Yuan, Z. Tian, L. Zhang, X. Wu, and B. Liu, "Building electricity load forecasting based on spatiotemporal correlation and electricity consumption behavior information," *Applied Energy*, vol. 377, p. 124580, 2025. https://www.sciencedirect.com/science/article/abs/pii/S0306261924019639

[5] A. Hossain, M. H. Ridoy, B. R. Chowdhury, M. N. Hossain, M. N. S. Rabbi, M. A. Ahad, … and M. S. Hasan, "Energy demand forecasting using machine learning: Optimizing smart grid efficiency with time-series analytics," *J. Environ. Agric. Studies*, vol. 5, no. 1, pp. 26–42, 2024. https://al-kindipublishers.org/index.php/jeas/article/view/9279

[6] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: A review and evaluation," *IEEE Trans. Power Syst.*, vol. 16, no. 1, pp. 44–55, 2001. https://ieeexplore.ieee.org/docume

[7] M. Abu-Shikhah, J. Elkarmi, and E. Aloquili, "Medium-term electric load forecasting using multivariate regression and principal component analysis," *Energy*, vol. 36, no. 7, pp. 4259–4266, 2011.

[8] S. Fan and R. J. Hyndman, "Short-term load forecasting based on a semi-parametric additive model," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 134–141, 2012. https://ieeexplore.ieee.org/document/5985500

[9] J. W. Taylor and R. Buizza, "Using weather ensemble predictions in electricity demand forecasting," *Int. J. Forecast.*, vol. 19, no. 1, pp. 57–70, 2003. https://www.sciencedirect.com/science/article/abs/pii/S0169207001001236

[10] C. L. Borges and V. F. Martins, "A self-organizing map for short-term load forecasting using daily profiles," *IEEE Trans. Power Syst.*, vol. 21, no. 1, pp. 109–118, 2006.

[11] L. Zjavka and V. Snášel, "Short-term power load forecasting with ordinary differential equation substitutions of polynomial networks," *Electr. Power Syst. Res.*, vol. 137, pp. 113–123, 2016. https://www.sciencedirect.com/science/article/abs/pii/S0378779616301092

[12] D. Liu, L. Zeng, C. Li, K. Ma, Y. Chen, and Y. Cao, "A distributed short-term load forecasting method based on local weather information," *IEEE Syst. J.*, vol. 12, pp. 208–215, 2018. https://ieeexplore.ieee.org/document/7549109

[13] M. L. Khashei, M. Bijari, and G. A. R. Ardali, "Polynomial neural networks for short-term load forecasting," *Energy*, vol. 35, no. 5, pp. 2096–2103, 2010.

[14] L. Zhang et al., "Hybrid models combining ARIMA, neural networks, and grey models for short-term load forecasting," *Applied Energy*, vol. 88, no. 11, pp. 3684–3690, 2011.

[15] Y. Chen, Y. Xu, and W. Su, "Short-term load forecasting: A review of hybrid models," *Electr. Power Syst. Res.*, vol. 146, pp. 325–335, 2017. https://www.sciencedirect.com/science/article/abs/pii/S036054421831065X

[16] A. Mocanu, P. H. Nguyen, M. Gibescu, and W. L. Kling, "Deep learning for estimating building energy consumption," *Sustain. Energy, Grids and Netw.*, vol. 6, pp. 91–99, 2016. https://www.sciencedirect.com/science/article/abs/pii/S2352467716000163

[17] A. Antoniadis, E. Paparoditis, and T. Sapatinas, "A functional time-series approach for electricity load forecasting," *Int. J. Forecast.*, vol. 22, no. 4, pp. 715–730, 2006. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5067419

[18] J. Lago, F. De Ridder, and B. De Schutter, "Forecasting spot electricity prices: Deep learning approaches and empirical comparison," *Applied Energy*, vol. 221, pp. 386–405, 2018. https://www.sciencedirect.com/science/article/pii/S030626191830196X

[19] H. Shi, M. Xu, and R. Li, "Deep learning for household load forecasting—A novel pooling deep RNN," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 5271–5280, 2018. https://ieeexplore.ieee.org/document/7885096

[20] X. Kong et al., "Online support vector regression for short-term load forecasting in smart grid," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6876–6888, 2018.

[21] M. Z. A. Bhotto, R. Jones, S. Makonin, and I. V. Bajic, "Short-term demand prediction using an ensemble of linearly-constrained estimators," *IEEE Trans. Power Syst.*, vol. 36, pp. 3163–3175, 2021. https://ieeexplore.ieee.org/document/9318533

[22] S. P. Muni and R. Sharma, "Short-term electricity price prediction using kernel-based machine learning techniques," in Proc. 1st Odisha Int. Conf. Electr. Power Eng., Commun. Comput. Technol. (ODICON), Bhubaneswar, India, pp. 1–5, Jan. 2021. https://ieeexplore.ieee.org/document/9428972

[23] Y. Li, R. Wang, and Z. Yang, "Optimal scheduling of isolated microgrids using automated reinforcement learning-based multi-period forecasting," *IEEE Trans. Sustain. Energy*, vol. 13, pp. 159–169, 2022. https://ieeexplore.ieee.org/document/9516934

[24] Z. B. Shi, Y. Li, and T. Yu, "Short-term load forecasting based on LS-SVM optimized by bacterial colony chemotaxis algorithm," in Proc. Int. Conf. Inf. Multimedia Technol., Beijing, China, pp. 306–309, Dec. 2009. https://ieeexplore.ieee.org/document/5381196

[25] R. Sabzehgar, D. Z. Amirhosseini, and M. Rasouli, "Solar power forecast for a residential smart microgrid based on numerical weather predictions using artificial intelligence methods," *J. Build. Eng.*, vol. 32, p. 101629, 2020. https://www.sciencedirect.com/science/article/abs/pii/S2352710219330414

[26] M. Mohamed, A. Chandra, M. A. Abd, and B. Singh, "Application of machine learning for prediction of solar microgrid system," in Proc. IEEE Int. Conf. Power Electron., Drives Energy Syst. (PEDES), Jaipur, India, pp. 1–5, Dec. 2020. https://ieeexplore.ieee.org/document/9379497

[27] E. Scolari, F. Sossan, and M. Paolone, "Irradiance prediction intervals for PV stochastic generation in microgrid applications," *Sol. Energy*, vol. 139, pp. 116–129, 2016. https://www.sciencedirect.com/science/article/abs/pii/S0038092X16304364

[28] Y. J. Ma and M. Y. Zhai, "Day-ahead prediction of microgrid electricity demand using a hybrid artificial intelligence model," *Processes*, vol. 7, p. 320, 2019. https://www.mdpi.com/2227-9717/7/6/320

[29] W. Tian, C. Lei, and M. Tian, "Dynamic prediction of building HVAC energy consumption by ensemble learning approach," in Proc. Int. Conf. Comput. Sci. Comput. Intell. (CSCI), Las Vegas, NV, USA, vol. 8, pp. 254–257, Dec. 2018. https://ieeexplore.ieee.org/abstract/document/8947463

[30] I. Hajjaji, H. E. Alami, M. R. El-Fenni, and H. Dahmouni, "Evaluation of artificial intelligence algorithms for predicting power consumption in university campus microgrid," in Proc. Int. Wireless Commun. Mobile Comput. (IWCMC), Harbin, China, pp. 2121–2126, Jul. 2021.https://ieeexplore.ieee.org/abstract/document/9498891

[31] A. Hadji and M. K. Kholladi, "Disaster information extraction: Evaluation of NLP techniques using JAPE rules, ontologies and machine learning approaches," in Proc. Int. Conf. Comput. Sci. Complex Syst. Appl. (CCSA), Cham: Springer, pp. 294–313, 2024. https://link.springer.com/chapter/10.1007/978-3-031-90758-6_22

[32] A. Hadji, F. Boumaza, and D. S. Bali, "Sentiment analysis of digital currency discussions: A machine learning and ontology approaches," in Proc. Int. Conf. Informatics Appl. Math. (IAM), CEUR Workshop Proc., vol. 2239, pp. 22–35, Dec. 2024. https://ceur-ws.org/Vol-3922/paper3.pdf

[33] A. Hadji, M. K. Kholladi, and N. Borisova, "Enhancing spatial information extraction from Arabic text: A hybrid approach with ontology and rule-based," *Ingénierie des Systèmes d'Information*, vol. 29, no. 4, p. 1261, 2024. https://www.iieta.org/journals/isi/paper/10.18280/isi.290402.