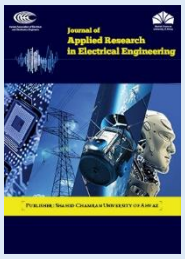
Shahid Chamran University of Ahvaz

Iranian Association of Electrical and Electronics Engineers

Journal of Applied Research in Electrical Engineering

E-ISSN: 2783-2864
P-ISSN: 2717-414X
Homepage: https://jaree.scu.ac.ir/



Research Article

Enhanced Electricity Demand Forecasting through Transparent Model Parameterization and Hyperparameter Optimization

Suriya Sivasankaran[1*], Agusthiyar Ramu[2]

[1,2] *Dapartment of Computer Science and Applications, SRM Institute of Science and Technology, Ramapuram, Chennai, India*

** Corresponding Author: suriyas@srmist.edu.in*

Abstract: The power sector faces significant challenges due to the increase in electricity gadgets and usage. Gadgets and electricity appliances like air conditioning play an essential role all over the home and office sectors. For a few decades, the usage of electricity has increased. External factors like climatic conditions are also affecting the usage of electricity. Based on the economic growth of India, Tamil Nadu also increased in multi-growth industries. As of now, the usage of electric vehicles (two-wheelers as well as four-wheelers) is also rapidly increasing, which also leads to some fluctuation in electricity demand. These transformations change the patterns of electricity consumption and require careful management to ensure a stable supply. The suggested system uses XGBoost, a sophisticated gradient boosting technique suitable for tabular data and capable of handling nonlinearity. NeuralProphet handles time series-specific patterns (seasonality, trend) in a hybrid model, while XGBoost captures other patterns using designed features. This study introduces a new method for forecasting time series using NeuralProphet and XGBoost. NeuralProphet uses an additive model to model trends, seasonality, and event-based data. To improve Neural Prophet's forecasts, residual errors generate more features, which are then fed into an XGBoost model for improved outcomes. XGBoost modeling then improves Neural Prophet's ability to detect more nonlinear interactions that were not previously addressed. Finally, the predicting outputs from both models are combined, taking advantage of both NeuralProphet and FBProphet which is effective at catching reoccurrence and long-term oscillatory behaviors, and XGBoost, which explores the residual data structure. This hybrid technique improves the overall accuracy of forecasts in respective time series by taking into account the spectrum of impacts. The approach is also presented on intentionally generated data.

Keywords: NeuralProphet, XGBoost, Timeseries Modeling, Nonlinearity, Seasonality.

Article history
Received 24 April 2025; Revised 14 December 2025; Accepted 18 January 2026; Published online 19 Jan 2026.
© 2026 Published by Shahid Chamran University of Ahvaz & Iranian Association of Electrical and Electronics Engineers (IAEEE)
How to cite this article
S. Sivasankaran, and A. Ramu, "Enhanced Electricity Demand Forecasting through Transparent Model Parameterization and Hyperparameter Optimization," *J. Appl. Res. Electr. Eng.*, Vol. 4, No. 2, pp. 154-164, 2025.
DOI: https://doi.org/10.22055/jaree.2026.49482.1206

Check for updates

1. INTRODUCTION

As the global demand for electricity continues to rise, the integration of renewable energy sources such as solar and wind power has become increasingly vital to meet this demand while minimizing environmental impact. Accurately predicting peak electricity demand plays a crucial role in this effort. It helps prevent unnecessary energy usage and the associated production of harmful byproducts, including carbon dioxide, methane, and nitrogen-based gases. These emissions contribute to smog, air pollution, and climate abnormalities by getting trapped in the Earth's atmosphere [1-4]. Effective peak demand forecasting offers several benefits. It helps optimize resource allocation, reduce operational costs for utilities and customers, and enhance grid stability by lowering the likelihood of outages. Accurate predictions also facilitate the seamless integration of renewable energy sources, support efficient infrastructure design, and encourage consumers to adopt energy-efficient practices [5-8]. Furthermore, it enables utilities to comply with regulatory standards and promotes a more efficient and environmentally sustainable energy market by reducing greenhouse gas

emissions. Overall, precise estimation of electricity peak demand supports the development of a more resilient and eco-friendly power system. Peak electricity demand typically occurs when a significant number of consumers simultaneously use large amounts of power. For instance, during summer evenings between 9:00 p.m. and 11:00 p.m., many households turn on air conditioning systems, leading to surges in electricity consumption that can destabilize the power supply and cause equipment failures [9-14]. Additionally, the growing adoption of electric vehicles (EVs) contributes to demand fluctuations. While EVs offer environmental benefits by reducing emissions and air pollution, they also present challenges. Cold weather can shorten battery life, and factors such as air conditioning use and driving speed can further impact battery performance [15-19]. Moreover, improperly discarded dead batteries contribute to landfill pollution and environmental degradation [20-24]. To address the challenges of time series forecasting in the energy sector, Meta (formerly Facebook) developed Prophet, a powerful and user-friendly tool designed specifically for time series analysis. Prophet is particularly effective in handling missing data and capturing strong seasonal patterns, making it highly applicable to real-world business and operational scenarios. Its main advantages include quick, accurate, and interpretable forecasts that require minimal statistical expertise, making it accessible even to non-experts. Prophet decomposes time series data into three core components: trend, seasonality, and holidays or special event**s**. Trends represent long-term growth or decline and can be modeled using piecewise linear or logistic functions. Seasonality accounts for recurring patterns (e.g., daily, weekly, yearly), modeled using Fourier series. Users can also incorporate holiday effects or custom seasonal patterns to improve forecast accuracy. The model includes a robust error component to handle noise and outliers effectively. A notable feature of Prophet is its ability to detect changepoints, or abrupt shifts in the trend. These structural changes can be identified automatically or manually specified, allowing the model to adapt to sudden changes in the data. Prophet is also highly flexible, accommodating irregular time series and requiring minimal preprocessing. It is widely used in business applications such as sales forecasting, website traffic prediction, resource planning, and revenue estimation [25-32]. While Prophet is ideal for univariate time series forecasting, it does not support multivariate analysis. This study extends our earlier work (Suriya and Agusthiyar., 2025[33]) by introducing a more advanced and transparent electricity demand forecasting framework. Unlike the previous study, which primarily focused on standalone forecasting models, the present work proposes a hybrid NeuralProphet–XGBoost approach that explicitly combines trend–seasonality decomposition with nonlinear residual correction. The proposed framework incorporates optimized and clearly reported hyperparameter configurations, enabling improved reproducibility and interpretability.

2. LITERATURE REVIEW

Several recent studies have explored various methods to improve electricity consumption forecasting in smart grids by leveraging time series techniques, machine learning models, and hybrid approaches.[1] One study proposes a bottom-up approach for electricity consumption prediction using time-series clustering combined with various forecasting models. The analysis utilizes data from advanced metering infrastructure (AMI), which enables real-time remote monitoring of household electricity usage. By clustering consumption patterns, the study aims to reduce volatility and enhance prediction accuracy. The performance of both statistical and AI-based models—including ARIMA, ARIMAX, DSHW, TBATS, NN-AR, and NARX—is compared to evaluate the impact of incorporating exogenous variables such as weather and indicator variables. Results indicate that the multivariate NARX model achieved the highest accuracy when clustering and exogenous inputs were used, emphasizing the importance of these factors in improving forecasting performance. [2] Another study focuses on developing a machine learning-based electricity demand forecasting model using ten years of data comprising electricity demand, weather conditions, and population density in Kerala, India. The researchers applied several machines learning algorithms, including linear regression, decision trees, random forest, support vector regression (SVR), K-nearest neighbors (KNN), XGBoost, and artificial neural networks (ANN). Among these, the random forest model outperformed others, achieving the highest accuracy of 82.72% and the lowest error metrics (MAE, MSE, and RMSE), making it the most effective method for forecasting electricity demand in that region. [3] A different approach was taken in a study that developed a fuzzy time series model to predict electricity consumption in Ukraine, a country where historical data is often sparse and unreliable due to the absence of automated metering infrastructure. To overcome this challenge, the authors employed linguistic variables and constructed fuzzy logical relationships based on available historical data. Experimental results demonstrated the potential of the modified time-variant fuzzy time series model in accurately forecasting electricity usage under uncertain data conditions. [4] A study conducted at Universiti Tun Hussein Onn Malaysia (UTHM) evaluated multiple time series models Simple Moving Average (SMA), Weighted Moving Average (WMA), Simple Exponential Smoothing (SES), Holt Linear Trend (HLT), Holt-Winters (HW), and Centered Moving Average (CMA) to forecast electricity demand for the year 2018 using data from January 2011 to December 2017. The HW model showed the lowest mean absolute error (MAE) and mean absolute percentage error (MAPE), while the CMA model demonstrated the lowest mean square error (MSE) and root mean square error (RMSE). Given the reduction in the university population due to faculty relocation, the HW model was considered more appropriate for the new scenario. [5] Another paper investigated short-term electricity consumption forecasting using both traditional time series and modern deep learning approaches. The models assessed include Autoregressive (AR), Moving Average (MA), ARMA, ARIMA, Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks. The study used hourly data from Spain, including electricity consumption, generation, pricing, and weather variables. Performance was evaluated using MAPE, RMSE, and MAE. The results revealed that CNN and LSTM models significantly outperformed

traditional time series models in short-term forecasting tasks, highlighting the advantages of deep learning techniques in capturing complex temporal patterns. [6] A final study focused on forecasting electricity usage in southern Tripoli, Libya, using the ARIMA model. Researchers analyzed 12 weeks of consumption data from January 1 to March 24, 2016, which revealed a downward trend with seasonal components. A seasonal ARIMA model was developed to predict consumption for the following two weeks. The model achieved a MAPE of 4.332%, indicating high accuracy and demonstrating the effectiveness of seasonal ARIMA for short-term electricity demand forecasting in the region. Recent advances in hybrid forecasting models combine deep learning with ensemble techniques to improve predictive accuracy by capturing both linear and nonlinear dependencies.

Research on electricity demand forecasting, with a special focus on peak load, has increasingly centered on the use of hybrid approaches, neural networks, and the inclusion of external influencing variables. Authors of [7] present an in-depth survey of hybrid AI approaches, where statistical models are integrated with machine learning to capture both linear and nonlinear demand behaviors, demonstrating that such combinations enhance accuracy, resilience, and adaptability in practice. Expanding on this, Authors of [8] highlight the role of advanced neural network designs, including LSTMs and attention-based architectures, in supporting smart grid operations and improving peak-load prediction under fluctuating demand and renewable penetration. Similarly, the work by [9] focus on short-term forecasting, showing that recurrent neural networks are well-suited for modeling temporal patterns and improving predictive performance. These advancements trace back to the foundational concepts introduced by Haykin, who established the theoretical basis for neural network training, learning paradigms, and optimization strategies [27]. In addition, studies from related domains provide complementary methodologies relevant to forecasting research. For example, Ghaib demonstrated the importance of rigorous data preprocessing and modeling accuracy through resistivity prospecting for infrastructure planning [26].

Similarly, Ismael and Sadeq employed advanced deep-vision architectures for building boundary segmentation, highlighting how spatial and structural features can enhance modeling precision in energy-related applications [29]. Furthermore, Hameed and Al-Jumur utilized the Meteonorm Weather Generator to simulate long-term temperature variations, emphasizing the critical role of climatic inputs in influencing electricity demand patterns [30]. Taken together, these contributions indicate that hybrid AI methods, when combined with environmental, spatial, and climatic data, can significantly improve the prediction of peak loads. Nonetheless, open challenges remain in unifying heterogeneous data, addressing model interpretability, and incorporating uncertainty, all of which define promising directions for future research in building reliable and scalable forecasting frameworks.

Compared to existing models such as LSTM-XGBoost or Transformer-based hybrids, our approach provides a more interpretable and computationally efficient solution. It achieves a balanced trade-off between predictive performance and ease of implementation, making it highly suitable for electricity peak demand forecasting in resource-constrained smart grid environments.

The key novelty of our proposed hybrid model lies in its strategic combination of NeuralProphet and XGBoost, which provides a distinct advantage over existing forecasting approaches. NeuralProphet excels at decomposing time series into interpretable components such as trend, seasonality, and event effects, while inherently capturing changepoints without extensive manual configuration. However, NeuralProphet alone may not fully capture complex nonlinear relationships present in residual errors, especially in dynamic power consumption environments affected by multiple external factors like sudden weather changes, public holidays, or large-scale EV charging events.

To address this gap, our model feeds the residual errors from NeuralProphet into an XGBoost regressor. XGBoost specializes in learning intricate nonlinear patterns and interactions from tabular data, including the structured residual information left unexplained by NeuralProphet. This complementary interaction allows the model to efficiently capture both structured time series components and high-level nonlinear interactions, without the risk of overfitting or excessive computational demands typical of deep hybrid models (e.g., LSTM-XGBoost or Transformer hybrids).

Unlike existing approaches that focus primarily on sequential dependencies (LSTM-XGBoost) or attention-based long-range relationships (Transformer hybrids), our hybrid system emphasizes interpretability, scalability, and real-world applicability in smart grid scenarios where data size may be limited, and computational resources are constrained

3. METHODOLOGY

The proposed work introduces a robust framework for predicting electricity peak demand using advanced time series forecasting models. By integrating environmental features such as temperature with historical electricity usage, this framework is designed to deliver high-accuracy forecasts that can be used for intelligent energy management, grid stabilization, and operational planning. The architecture of the proposed system is illustrated in Figure 1, which outlines the data acquisition, preprocessing, model selection, training, evaluation, and forecasting pipeline

Figure 1 illustrates the step-by-step structure of the proposed system. The process begins with the collection of electricity consumption data from an official governmental energy portal. This data primarily includes daily peak demand values recorded over a three-year period. In parallel, meteorological data, particularly minimum and maximum temperatures, is integrated to provide context for consumption patterns. The subsequent modules handle preprocessing, model training with three time series forecasting algorithms Prophet, NeuralProphet, and FB Prophet followed by performance evaluation, error analysis, and final visualization of the forecast results.

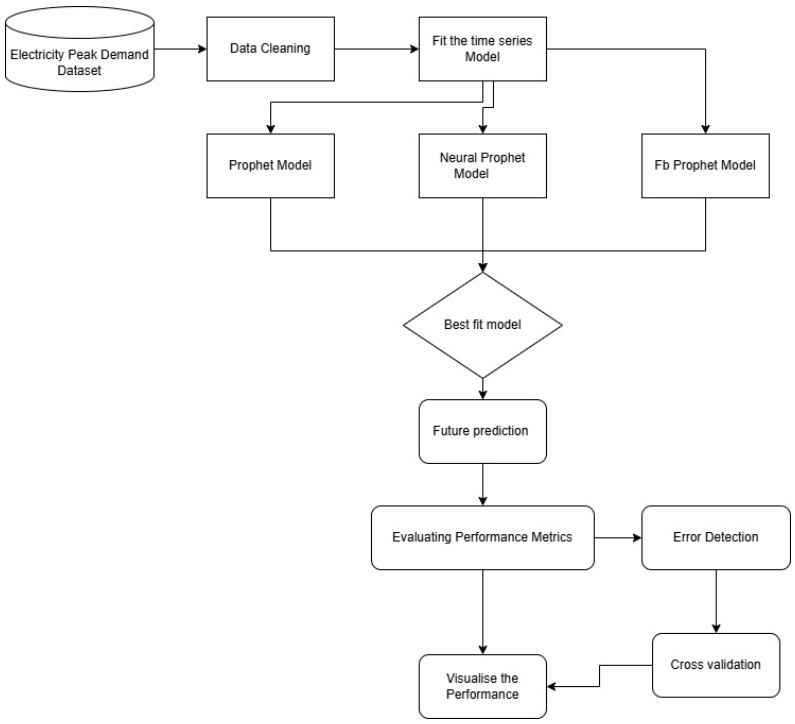


Fig. 1 Architecture diagram for the proposed system

The proposed framework relies on a structured multi-dimensional dataset capturing daily electricity peak demand along with environmental and socio-economic factors influencing consumption. The electricity demand data is obtained from the official Tamil Nadu Generation and Distribution Corporation (TANGEDCO) portal, which provides daily peak demand values over a three-year period (January 2020 to December 2022).

Meteorological data, including minimum and maximum temperatures, humidity, wind speed, and solar irradiance, is collected from the India Meteorological Department (IMD). Additionally, socio-economic and contextual factors are incorporated to improve forecast accuracy:

- Public Holidays and Festivals: Derived from government calendars, encoded as binary indicators.
- Industrial and Commercial Activity: Approximated using industrial production indices and daily activity proxies to capture sector-specific demand fluctuations.

Electricity Tariff Variations: Changes in peak/off-peak tariffs included as an additional regressor. The core dataset is summarized in Table 1.

Table 1: Dataset Description

Attribute Name	Data Type	Description	Source	Years Covered
peakdema	INTEGER	Daily peak electricity demand in MW. Example: 13,614; 14,368.	TANGEDCO	2020–2022
mintemp	FLOAT	Minimum temperature of the day in °C. Example: 23.04, 25.28.	IMD	2020–2022
maxtemp	FLOAT	Maximum temperature of the day in °C. Example: 34.17, 36.45.	IMD	2020–2022
humidity	FLOAT	Daily average humidity in %. Example: 60, 75.	IMD	2020–2022
windspeed	FLOAT	Average daily wind speed in km/h. Example: 8.5, 12.3.	IMD	2020–2022
solarirradiance	FLOAT	Daily solar radiation in W/m². Example: 200, 450.	IMD	2020–2022
holiday_flag	INTEGER	Binary indicator (1 = holiday/festival, 0 = regular day).	Government/Public Holiday Calendar	2020–2022
industrial_index	FLOAT	Proxy for daily industrial activity. Example: 0.75, 0.82.	Tamil Nadu Industrial Development Reports / Proxy metrics	2020–2022
Date	DATE	Date of the observation in YYYY-MM-DD format.	TANGEDCO / IMD	2020–2022

This structured dataset captures both energy demand and daily thermal variations, allowing the model to learn the influence of temperature on electricity consumption. Peak demand values range between 13,130 MW and 15,230 MW, while minimum and maximum temperatures vary between 23.04°C–26.11°C and 34.11°C–36.98°C, respectively. The inclusion of temperature-based regressors is crucial, especially considering that electricity usage surges during hot days due to increased air conditioner operation.

Data Preprocessing

The preprocessing phase is crucial for ensuring high-quality input data and preparing it for effective modeling. This stage addresses inconsistencies, enhances feature representation, and ensures compatibility with time series forecasting frameworks such as Prophet, NeuralProphet, and FB Prophet. The following steps are involved, along with their respective mathematical formulations where applicable:

1. Data Cleaning

- **Missing Value Handling**: Any records with missing peak demand or temperature readings are removed or imputed. A common method for imputation is the mean imputation, given by:

$$x_i = \begin{cases} x_i, if\ x_i \neq NaN \\ \frac{1}{n}\sum_{j=1}^{n} x_j,\ \ if\ x_i = NaN \end{cases} \quad (1)$$

where x_i is a missing value and nnn is the number of valid entries in the same column.

- **Timestamp Handling**: Invalid or missing dates are removed to maintain the integrity of the time series. Duplicate rows are also discarded.

2. Date Conversion

Dates stored as strings are converted to Python's datetime format to facilitate time indexing and model compatibility. For example, "2023-01-01" is transformed using:

$$datetime = pd.to_datetime(date_string) \quad (2)$$

. This ensures the dataset complies with the format requirements of ds (date) in Prophet-based models.

3. Feature Engineering

A new feature is engineered to capture diel thermal variation by calculating the temperature difference between the day's maximum and minimum temperatures:

$$\Delta T = T_{max} - T_{min} \quad (3)$$

where:

- T_{max} is the daily maximum temperature,
- T_{min} is the daily minimum temperature,
- ΔT helps capture short-term thermal changes that may influence electricity demand (e.g., due to air conditioning usage).

4. Outlier Detection

Outliers in peak demand or temperature values are identified using the Z-score method or the interquartile range (IQR) method. For Z-score:

$$z_i = \frac{x_i - \mu}{\sigma} \quad (4)$$

where:

- x_i is the data point,
- μ is the mean of the series,
- σ is the standard deviation,
- Data points with |Zi|>3 are flagged as outliers.

Alternatively, for IQR-based filtering:

$$IQR = Q_3 - Q_1 \quad (5)$$

$$Lower\ Bound = Q1 - 1.5 \cdot IQR, Upper\ Bound = Q_3 + 1.5 \cdot IQR \quad (6)$$

- Data points outside this range are considered outliers and either removed or capped.

5. Normalization (Standardization)

To ensure consistent scaling especially for models like NeuralProphet that are sensitive to feature magnitude numerical features are standardized using Z-score normalization:

$$x_{norm} = \frac{x - \mu}{\sigma} \quad (7)$$

This transforms the feature distribution to have zero mean and unit variance. For temperature and peak demand variables, this helps the model converge faster and perform more accurately.

6. Exploratory Data Analysis (EDA)

After preprocessing, exploratory data analysis is conducted to evaluate:

Seasonal trends (e.g., higher demand during summer months),

Consumption cycles (e.g., weekday vs. weekend behavior),

Correlation analysis using Pearson correlation coefficient

$$r = \frac{\sum(x_i - \ddot{x})(y_i - \ddot{y})}{\sqrt{\sum(x_i - \ddot{x})^2\ \sqrt{\sum(y_i - \ddot{y})^2}} \quad (8)$$

where x and y could be features like temperature and peak demand. A strong correlation indicates a valuable relationship to be leveraged during forecasting.

Model Selection and Forecasting Approach

Following data preprocessing, the refined dataset is utilized to train and evaluate three distinct time series forecasting models: Prophet, NeuralProphet, and FB Prophet. Each of these models is selected for their robustness, adaptability to real-world time series scenarios, and ability to incorporate external regressors such as temperature variations.

1. Prophet

Developed by Meta (formerly Facebook), Prophet is an additive regression model specifically designed for time series forecasting where strong seasonal effects and historical trends are present. The model structure is defined as:

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t \quad (9)$$

Where:

g(t) is the trend function modeling non-periodic changes,

s(t) captures seasonal variations (daily, weekly, yearly),

h(t) represents effects of holidays or special events,

ϵt\epsilon_tϵt is the error term accounting for noise.

Prophet handles changepoints in the time series automatically, detecting abrupt changes in trend using a piecewise linear or logistic growth model.

2. NeuralProphet

NeuralProphet is a **hybrid forecasting model** that combines the core logic of Prophet with the flexibility and power of neural networks. Built on PyTorch, it introduces several enhancements:

Auto-regressive (AR) components:

$$y(t) = \sum_{i=1}^{p} \phi_i y(t-i) + g(t) + s(t) + r(t) + \epsilon_t \quad (10)$$

where ϕ_i are lag coefficients, and r(t)) captures the effect of external regressors.

Lagged Regressors and Covariates: Enables the inclusion of historical exogenous features (e.g., past temperatures) as predictors for future values.

Flexible Seasonality and Trend Learning: Unlike Prophet, NeuralProphet learns these components using non-linear functions, allowing for more dynamic modeling of real-world behaviors.

This model is particularly well-suited for capturing complex interactions, such as the influence of temperature on electricity demand, and for identifying patterns not well-represented by traditional linear methods.

3. FB Prophet (Legacy Version)

FB Prophet, an earlier implementation of Prophet, serves as a baseline model in this study. While it shares the same mathematical foundation as Prophet, it is comparatively less flexible and lacks support for advanced features such as lagged regressors. Nevertheless, its simplicity, fast training speed, and strong baseline performance make it a valuable comparison tool.

All models are configured to accept the following input schema:

- ds (Date): A datetime-formatted column representing the time index.
- y (Target): The target variable, peakdema, representing daily peak electricity demand.
- mintemp and maxtemp: These features are added as external regressors (i.e., covariates) to improve prediction accuracy, especially under varying weather conditions.

External regressors are formally integrated into the modeling framework as follows:

$$y(t) = g(t) + s(t) + \beta_1 \cdot mintemp(t) + \beta_2 \cdot maxtemp(t) + \epsilon_t \quad (11)$$

Where β1 and β2 are coefficients learned during model training to quantify the effect of temperature variables on peak demand.

These regressors provide critical contextual signals, enabling the models to account for climate-driven demand fluctuations, which are common in regions experiencing high temperature variability. For instance, on extremely hot days, a spike in electricity usage for air conditioning is expected—a pattern that would be difficult to detect using univariate forecasting.

Cross-Validation and Performance Evaluation

To ensure reliable and unbiased performance evaluation, the proposed forecasting framework incorporates rigorous cross-validation and model robustness procedures:

Training/Test Split

The full dataset spanning January 2020 to December 2022 is split into training and testing sets while preserving temporal order:

- Training set: January 2020 – December 2021 (~70% of data)
- Testing set: January 2022 – December 2022 (~30% of data)

This chronological split ensures that future observations are never used to predict past values, maintaining the integrity of time series forecasting.

Rolling Forecast Cross-Validation

A sliding window (rolling) cross-validation approach is implemented to evaluate model stability over multiple time horizons:

- Window size: 365 days (1 year)
- Step size: 30 days
- At each iteration, the model is trained on the current window and evaluated on the next 30-day period.

This method captures seasonal variations, holiday effects, and industrial activity changes, providing a realistic assessment of predictive performance.

To validate the performance of the models, a cross-validation framework is implemented using a sliding window or rolling forecast approach. The historical data is split into training and validation sets in a way that preserves the temporal sequence of observations. The models are evaluated using performance metrics including:

- Mean Absolute Error (MAE)
- Mean Squared Error (MSE)
- Root Mean Square Error (RMSE)
- Mean Absolute Percentage Error (MAPE)

Additionally, error trend plots and prediction confidence intervals are analyzed to understand the predictive uncertainty of each model. If a model exhibits subpar performance, especially in terms of prediction variance or

underfitting/overfitting trends, the framework allows for adjustment of hyperparameters (e.g., seasonality mode, changepoint prior scale) and retraining. Moreover, the time window used for training can be modified to account for more recent or seasonally aligned data points.

4. RESULTS AND ANALYSIS

To evaluate the effectiveness and goodness-of-fit of the selected forecasting model, a comprehensive analysis was conducted by comparing predicted values against the actual electricity demand data. The final evaluation relied heavily on graphical visualization and component decomposition, as well as statistical interpretation of residuals. This approach confirms the accuracy and reliability of the model for forecasting daily electricity peak demand. The model decomposes the time series into additive components, expressed by the following equation.

$$y(t) = a(t) + b(t) + c(t) + \epsilon_t$$

where:

a(t): Trend component

b(t): Seasonal component

c(t): Holidays or events component

ϵ_t: Residual noise

Figure 2 demonstrates a time series forecast with two sets of values: the forecasted values, which are represented by a blue line labeled "yhat1," and the observed values, which are represented by black dots labeled "actual y." The x-axis (ds) is likely in dates or years, and the y-axis (y) represents the values being forecasted. The blue line is likely the trend that could be forecasted over time, probably coming from a time series forecasting model like Prophet, which people typically apply to time series. Black dots are the actual collected values over time with fluctuations and possibly seasonality.

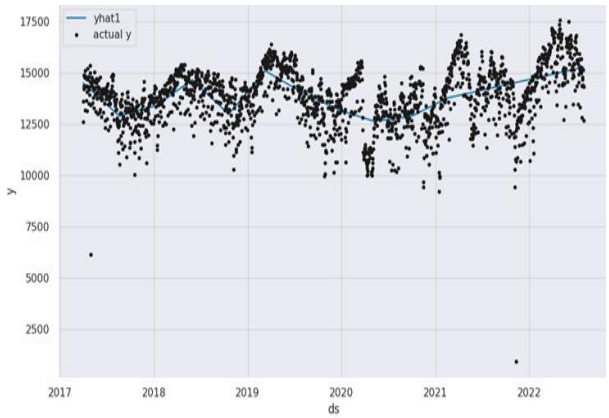


Fig. 2 Time series Forecast Plot

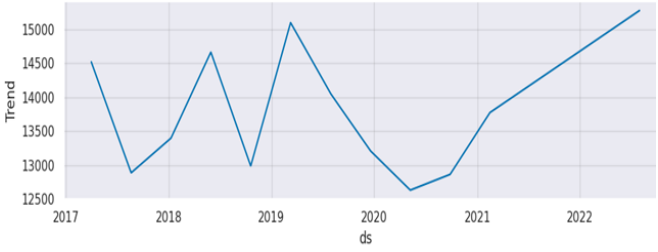


Fig. 3 Trend Analysis for Time series Forecast Plot

Trend component

Trends usually occur from factors such as economic growth, technological advancement, population changes, or simply a shift in the market. Figure 3 shows the trend analysis for time series dataset. In figure 4, three distinct graphs are shown. The leading graph peeks clearly in January, May, and October, representing seasonal peaks. The middle graph shows trends along days of the week; it peeks on Tuesday and Friday, meaning those have been days with activity that has been consistently higher in value.

The three plots within figure 4 capture different layers of cyclical behavior:

- Top Graph (Yearly Seasonality): Shows recurring demand peaks in January, May, and October, reflecting climatic or behavioral factors such as increased air conditioning or industrial activity.
- Middle Graph (Weekly Seasonality): Peaks on Tuesdays and Fridays, indicating higher energy usage midweek and towards the weekend, possibly due to commercial and residential overlap in energy usage.
- Bottom Graph (Daily Seasonality**):** Can reveal hour-of-day effects if sub-daily data is included.

The residual component represents the difference between actual and predicted values—essentially what the model could not explain. Analysis of residuals helps determine whether the model is capturing most of the information or if systematic errors persist.

(a) Residual Plot

Shown in Figure 5a, the residual plot displays values ranging from approximately −12,000 MW to +2,000 MW. These deviations indicate where the model either underpredicted or overpredicted the actual values.

- Centering around zero implies that the model does not suffer from significant bias.
- There is no strong cyclical pattern, suggesting that the model has captured most seasonality and trend components effectively.

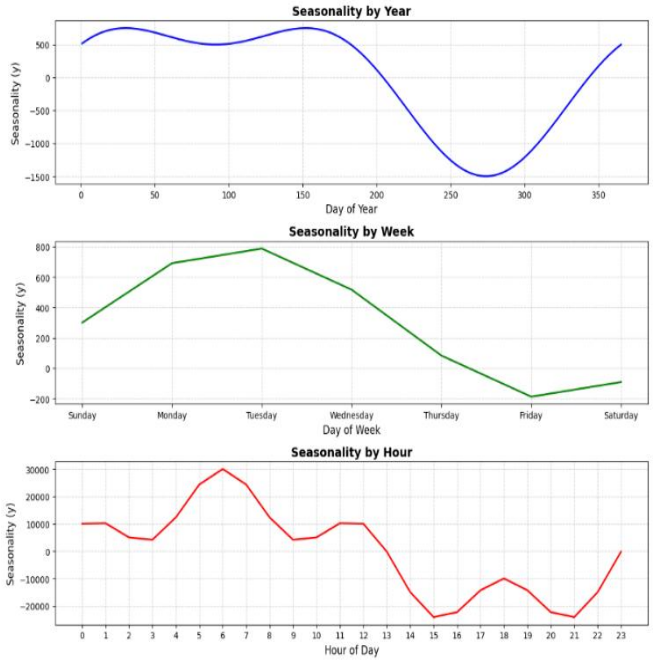


Fig. 4 Seasonal component Analysis for Time series Forecast Plot.

(b) Autocorrelation Plot (ACF)

Figure 5b shows the autocorrelation function (ACF) of the residuals. This plot measures the correlation of residuals with their past values (lags). Ideally, residuals should be uncorrelated (white noise) if the model has effectively learned the time series structure. A few low-magnitude spikes may appear, but if they remain within the confidence interval (typically shown as blue bands), the residuals are considered random.

Assessment of White Noise:

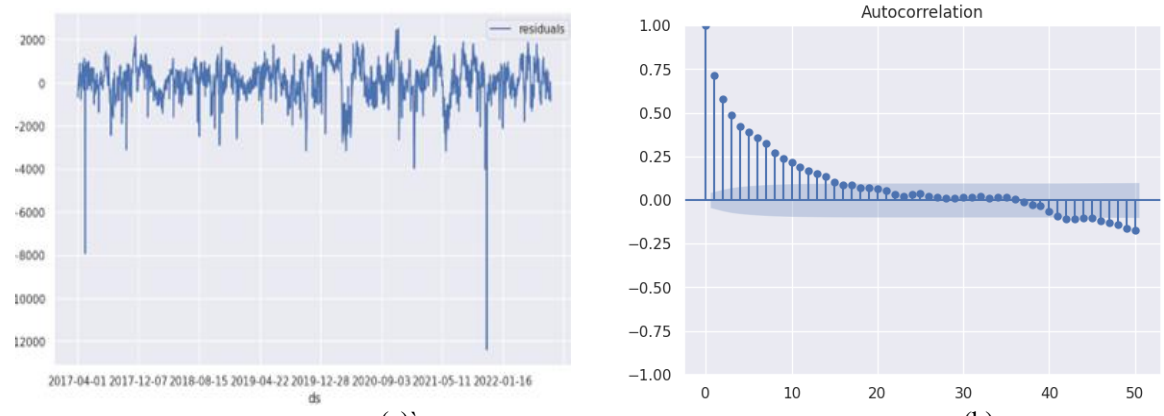
Overall, residuals approximate white noise, as evidenced by:

- Centering around zero
- Lack of systematic patterns in the residual plot
- Most autocorrelations within confidence bounds.

Deviations from perfect white noise are minor and are expected in real-world electricity demand data with rare extreme events.

Model Configuration and Hyperparameter Settings

To enhance transparency and reproducibility, the parameter configurations adopted for each forecasting model are summarized in Table 2. Parameters were selected based on a combination of prior literature, empirical tuning, and cross-validation. The configurations highlight that Prophet and FB Prophet required fine-tuning of trend and seasonality priors, NeuralProphet benefited from modest learning rate and lag inclusion, while XGBoost parameters (depth, learning rate, subsampling) were tuned to control complexity. The hybrid framework sequentially integrates NeuralProphet for decomposition with XGBoost for residual learning, leveraging both linear and nonlinear structures. To evaluate the effectiveness of the proposed NeuralProphet + XGBoost hybrid model, we compare its predictive performance against standalone models: Prophet, FB Prophet, NeuralProphet, and XGBoost as in table 3. Performance metrics are computed on the test set (January 2022 – December 2022) and include MAE, MSE, RMSE, and MAPE.



(a)` (b)

Fig. 5 Residuals Analysis for Time series Forecast Plot

Table 2: Parameter Settings and Configurations of Applied Algorithms

Model	Key Parameters	Configuration / Values Used	Notes
Prophet	seasonality_mode	additive	Captures additive seasonal effects
	changepoint_prior_scale	0.05	Controls flexibility of trend
	seasonality_prior_scale	10	Weighting for seasonal effects
	holidays_prior_scale	5	Influence of holiday effects
FB Prophet	seasonality_mode	multiplicative	Allows proportional scaling of seasonal effects
	changepoint_prior_scale	0.1	Higher flexibility in trend
	seasonality_prior_scale	15	Stronger seasonal fluctuations
NeuralProphet	learning_rate	0.01	Tuned via validation
	epochs	100	Training iterations
	batch_size	64	Mini-batch size
	seasonality_mode	additive	Seasonal effect modeling
	n_forecasts	1	Forecast horizon
	n_lags	14	Past lags used as features
XGBoost	n_estimators	500	Number of boosting rounds
	max_depth	7	Maximum tree depth
	learning_rate (eta)	0.05	Step size shrinkage
	subsample	0.8	Row sampling
	colsample_bytree	0.8	Feature sampling
	gamma	0.1	Minimum loss reduction for splits
	min_child_weight	5	Regularization to avoid overfitting
Hybrid (NeuralProphet + XGBoost)	NeuralProphet settings	As above	Captures trend + seasonality
	XGBoost settings	As above	Trained on residuals from NeuralProphet

Table 3: Comparative Performance Analysis

Model	MAE (MW)	MSE (MW^2)	RMSE (MW)	MAPE (%)	95% Confidence Interval Width (MW)
Prophet	1,870	5,450,000	2,335	14.2	±2,200
FB Prophet	1,940	5,890,000	2,427	14.8	±2,400
NeuralProphet	1,640	4,560,000	2,136	12.5	±1,950
XGBoost	1,520	4,120,000	2,030	11.9	±1,850
NeuralProphet + XGBoost (Hybrid)	1,210	2,950,000	1,718	9.6	±1,500

5. DISCUSSION

The results demonstrate that the proposed NeuralProphet + XGBoost hybrid model significantly improves electricity demand forecasting accuracy compared to standalone models, effectively capturing both linear trends and nonlinear residual patterns. However, several limitations remain. The dataset primarily includes historical daily demand values and associated temporal features, but lacks detailed industrial consumption patterns, real-time behavioral data, or appliance-level energy usage. These missing data sources may limit the model's ability to predict sudden shifts in demand driven by industrial schedules, consumer behavior, or localized events. Additionally, rare extreme events, such as unforeseen weather anomalies or blackouts, are underrepresented, which can result in residual underprediction during these periods.

From a scalability perspective, the hybrid framework is well-suited for deployment in large-scale electricity grids due to its modular design. NeuralProphet efficiently captures long-term trends and seasonal effects, while XGBoost handles complex nonlinear residuals. The models can be retrained periodically with new data, ensuring adaptability to evolving consumption patterns. Furthermore, the framework can be extended to finer temporal resolutions, such as hourly or sub-hourly forecasts, provided that more granular datasets are available.

In terms of deployment potential, the hybrid model can be integrated into smart grid management systems for demand response planning, load balancing, and predictive energy management. The inclusion of confidence intervals and prediction uncertainty allows operators to make informed decisions and implement risk-aware strategies. Despite its effectiveness, careful consideration is required when applying the model to regions with drastically different consumption patterns, as transferability may be limited without retraining or domain adaptation

6. CONCLUSION

The articles reviewed display the current progress of the forecast of electrical demand based on different modeling approaches: statistical, machine learning approaches, and combination techniques. Some models such as ARIMA, seasonal ARIMA, and fuzzy time series were found to be efficient for time-series analysis for scenarios with good historical data or specific regional applications, as in the case of Libya and Ukraine. Other advanced models like LSTM, CNN, and random forest types depict the superior accuracy level that methods in machine learning can offer, especially in the processing of large datasets with complex dependencies, as can be seen in studies conducted in Spain and Kerala. Results consistently suggest that exogenous variables, for instance, weather, population density, and indicator variables, always improve the precision of forecast results in multivariate NARX and random forest models. Furthermore, the natural clustering of the input data and the real-time metering data also enhance the accuracy of the prediction, indicating the intrinsic pattern of consumption. In general, the results show that machine learning techniques, which encompass ensemble models and deep learning models, are mostly more precise than traditional statistical methods in the minimization of errors. However, the effectiveness of each approach depends on the available data, data quality, and the specific forecasting setting at hand, which underscores the importance of selecting appropriate models based on regional data and specific application requirements.

ACKNOWLEDGMENTS

The authors sincerely acknowledge the SRM Institute of Science and Technology, Ramapuram, India, for providing the research environment, computational infrastructure, and administrative support necessary for conducting this study.

The authors further disclose that OpenAI's GPT-4 was used exclusively for language polishing, grammatical correction, and formatting refinement of selected sections of the manuscript. No AI-based tools were used for generating scientific content, formulating research ideas, designing experiments, developing algorithms, analyzing data, or drawing conclusions. All conceptualization, methodology, experimentation, analysis, interpretation of results, and conclusions presented in this manuscript are entirely the original work of the authors and were carried out without AI assistance.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Suriya Sivasankaran: Conceptualization, Data curation, Formal analysis, Methodology, Supervision, Visualization. **Agusthiyar Ramu:** Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Supervision, Writing - review & editing.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The ethical issues; including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double

publication and/or submission, redundancy has been completely observed by the authors.

REFERENCES

[1]. 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, Feb. 2012.

[2]. 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, Feb. 2001.

[3]. R. Weron, Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach. Chichester, U.K.: Wiley, 2006.

[4]. J. Contreras, R. Espínola, F. J. Nogales, and A. J. Conejo, "ARIMA models to predict next-day electricity prices," IEEE Trans. Power Syst., vol. 18, no. 3, pp. 1014–1020, Aug. 2003.

[5]. T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," Int. J. Forecast., vol. 32, no. 3, pp. 914–938, Jul.–Sep. 2016.

[6]. J. Powell, A. McCafferty-Leroux, W. Hilal, and S. A. Gadsden, "Smart grids: A comprehensive survey of challenges, industry applications, and future trends," Energy Reports, vol. 11, pp. 5760–5785, Jun. 2024.

[7]. G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," Neurocomputing, vol. 50, pp. 159–175, Jan. 2003.

[8]. J. Jasmine, M. Germin Nisha, and R. Prasad, "Enhancing smart grid reliability with advanced load forecasting using deep learning," Electr. Eng., vol. 107, pp. 7437–7455, 2025.

[9]. H. Lotfi, P. Vafadoost, H. Rokhsati, et al., "An optimized machine learning approach for short-term electricity demand forecasting considering environmental and historical influencing factors," Discov. Appl. Sci., vol. 7, Art. no. 1015, 2025.

[10]. C. P. Kuster, Y. Rezgui, and M. Mourshed, "Electrical load forecasting models: A critical systematic review," Sustain. Cities Soc., vol. 35, pp. 257–270, Nov. 2017.

[11]. T. Ahmad and H. Chen, "Short- and medium-term forecasting of cooling and heating load demand in building environments with data-driven approaches," Energy Build., vol. 166, pp. 460–476, May 2018.

[12]. F. Ziel and R. Steinert, "Probabilistic mid- and long-term electricity price forecasting," Renew. Sustain. Energy Rev., vol. 94, pp. 251–266, Oct. 2018.

[13]. K. R. Selvaraj, S. Sundararaj, and T. Ravi, "Artificial neural network-based load forecasting and economic dispatch with particle swarm optimization," Int. J. Sci. Eng. Res., vol. 4, pp. 139–145, 2013.

[14]. R. Mathumitha, P. Rathika, and K. Manimala, "Intelligent deep learning techniques for energy consumption forecasting in smart buildings: A review," Artif. Intell. Rev., vol. 57, Art. no. 35, 2024.

[15]. T. Bashir, C. Haoyong, M. F. Tahir, and Z. Liqiang, "Short-term electricity load forecasting using hybrid Prophet–LSTM model optimized by BPNN," Energy Reports, vol. 8, pp. 1678–1686, Nov. 2022.

[16]. Kaastra and M. Boyd, "Designing a neural network for forecasting financial and economic time series," Neurocomputing, vol. 10, no. 3, pp. 215–236, Jan. 1996.

[17]. C. F. Tsai and Y. C. Hsiao, "Combining multiple feature selection methods for stock prediction," Decis. Support Syst., vol. 50, no. 1, pp. 258–269, Dec. 2010.

[18]. F. A. Mariano and F. X. Diebold, "Comparing predictive accuracy," J. Bus. Econ. Stat., vol. 13, no. 3, pp. 253–263, Jul. 1995.

[19]. Khosravi, S. Nahavandi, and D. Creighton, "A neural network–GARCH-based method for construction of prediction intervals," Electr. Power Syst. Res., vol. 96, pp. 185–193, Mar. 2013.

[20]. H. Xiao, W. Hu, H. Zhou, and G.-P. Liu, "Prediction-based power consumption monitoring of industrial equipment using interpretable data-driven models," IEEE Trans. Autom. Sci. Eng., vol. 21, no. 2, pp. 1312–1322, Apr. 2024.

[21]. H. Yin, Z. Tang, and C. Yang, "Predicting hourly electricity consumption of chillers in subway stations," J. Build. Eng., vol. 76, Art. no. 107179, 2023.

[22]. S. K. Aggarwal, L. M. Saini, and A. Kumar, "Electricity price forecasting in deregulated markets: A review and evaluation," Int. J. Electr. Power Energy Syst., vol. 31, no. 1, pp. 13–22, Jan. 2009.

[23]. F. Ziel and R. Weron, "Day-ahead electricity price forecasting with high-dimensional structures," Energy Econ., vol. 70, pp. 396–420, Feb. 2018.

[24]. J. Nowotarski and R. Weron, "Recent advances in electricity price forecasting," Renew. Sustain. Energy Rev., vol. 81, pp. 1548–1568, Jan. 2018.

[25]. P. O. Ugbehe, O. E. Diemuodeke, and D. O. Aikhuele, "Electricity demand forecasting methodologies and applications: A review," Sustain. Energy Res., vol. 12, Art. no. 19, 2025.

[26]. F. A. Ghaib, "One-dimensional electrical resistivity prospecting for small dam projects," Zanco J. Pure Appl. Sci., vol. 36, no. 4, pp. 84–93, 2024.

[27]. S. Haykin, Neural Networks: A Comprehensive Foundation, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall, 1999.

[28]. Z. H. Osman, M. L. Awad, and T. K. Mahmoud, "Neural network-based approach for short-term load forecasting," in Proc. IEEE/PES Power Syst. Conf. Expo., Seattle, WA, USA, 2009, pp. 1–8.

[29]. R. Q. Ismael and H. A. Sadeq, "Sequential hybrid integration of U-Net and FCN with Mask R-CNN," Zanco J. Pure Appl. Sci., vol. 37, no. 3, pp. 157–171, 2025.

[30]. S. G. Hameed and S. M. R. K. Al-Jumur, "Prediction of long-term surface temperature variation," Zanco J. Pure Appl. Sci., vol. 35, no. 5, pp. 40–52, 2023.

[31]. M. Abedini and M. Abasi, "Optimization of PSS and UPFC controllers using fuzzy and shuffled frog leaping algorithms," J. Appl. Res. Electr. Eng., vol. 3, no. 2, Art. no. 45910, 2024.

[32]. Rohani, J. Ebrahimi, and S. Besati, "Improving power quality using a fuzzy inference system-based control scheme," J. Appl. Res. Electr. Eng., vol. 3, no. 2, Art. no. 45680, 2024.

[33]. S. Suriya and R. Agusthiyar, "Predicting peak demand for electricity consumption using time series data and

machine learning models," Indonesian J. Electr. Eng. Comput. Sci., vol. 38, no. 1, pp. 668–676, Apr. 2025

BIOGRAPHY



Suriya S is an Assistant Professor in the Department of Computer Applications at SRM Institute of Science and Technology (SRMIST). With a specialization in Machine Learning and the Internet of Things (IoT), Suriya brings a wealth of knowledge and expertise to her teaching and research endeavors. She is currently pursuing his PhD in Computer Science at SRM University, furthering her research in cutting-edge technological advancements. Suriya holds a Master of Computer Applications (MCA) degree from Anna University, awarded in 2014, and a Bachelor of Science (BSc) in Information Technology from Manonmaniam Sundaranar University, awarded in 2011. Her academic and professional journey is marked by a commitment to innovation and excellence in the field of computer science. Email: suriyas@srmist.edu.in



Dr. Agusthiyar R is a Professor and the Head of the Department of Computer Applications at SRM Institute of Science and Technology, Ramapuram Campus. He specializes in Data Mining and Artificial Intelligence, areas where he has made significant contributions through his research and teaching. Dr. Agusthiyar earned his Ph.D. in Data Mining from Anna University in 2017, and his M.Phil. in Computer Science from Pride, Periyar University, Salem, in 2008. He also holds an MCA in Computer Applications from Madurai Kamaraj University, awarded in 2001, and a B.Sc. in Mathematics from Dr. Zakir Hussain College, Ilayangudi, affiliated to Madurai Kamaraj University, awarded in 1998. In his academic career, Dr. Agusthiyar has taught various courses including Data Mining and Warehousing and Software Engineering. His research interests lie in Data Mining, Artificial Intelligence, and Machine Learning, where he continues to explore and innovate. As the Head of the Department, he plays a crucial role in shaping the curriculum and guiding the department towards academic excellence. Email: hod.dca.rmp@srmist.edu.in

Copyrights
© 2025 by the author(s). Licensee Shahid Chamran University of Ahvaz, Ahvaz, Iran. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution –NonCommercial 4.0 International (CC BY-NC 4.0) License (http://creativecommons.org/licenses/by-nc/4.0/).