Research Article

Electrical Load Prediction Using Statistical, Deep Learning, and Hybrid Time Series Models

Ali Hussain [1*], Syed Taha Ahmed [2], Zohaib Ahmed [3]

[1]Faculty of Computer Science & Information Technology, The University of Lahore, Lahore, Pakistan
[2] MU Institute for Data Science and Informatics, University of Missouri, Missouri, United States
[3] Faculty of information Technology, Beijing University of Technology, Beijing, China
[*]Corresponding Author: Ali Hussain Author Email: ali.hussain1@cs.uol.edu.pk

Received: May 27, 2025 **Accepted:** December 17, 2025 **Published:** December 31, 2025

Abstract: Accurate electrical load forecasting is essential for the reliable, economical, and secure operation of modern power systems, as it directly supports generation scheduling, grid stability, and long-term planning decisions. However, electricity demand exhibits strong nonlinearity, nonstationary, and pronounced seasonal patterns, which significantly limit the effectiveness of conventional forecasting techniques, particularly over extended prediction horizons. To address these challenges, this study presents a comprehensive and unified benchmarking framework that systematically evaluates classical statistical models, deep learning architectures, and a residual-based hybrid approach for hourly electrical load forecasting. Specifically, Autoregressive Integrated Moving Average (ARIMA), Seasonal ARIMA (SARIMA), Long Short-Term Memory (LSTM), and a hybrid ARIMA–LSTM model are investigated using the publicly available PJM East (PJME) hourly electricity consumption dataset. The proposed hybrid framework decomposes the load series into linear–seasonal and nonlinear components, allowing ARIMA to capture structured temporal patterns while LSTM learns complex nonlinear residual dependencies. A consistent preprocessing pipeline and identical experimental settings are employed to ensure a fair and transparent comparison across short-term, medium-term, and long-term forecasting horizons. Forecasting performance is evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). Experimental results demonstrate that the hybrid ARIMA–LSTM model consistently outperforms standalone statistical and deep learning models across all forecasting horizons, achieving the lowest prediction errors and exhibiting improved robustness against error accumulation. These findings confirm that residual-based hybrid learning effectively combines the interpretability of statistical models with the nonlinear modeling capability of deep learning, offering a practical and reliable solution for real-world electricity demand forecasting in modern power systems.

***Keywords*:** Electrical Load Forecasting; ARIMA; SARIMA; LSTM; Hybrid Forecasting Models; Time-Series Analysis; Power Systems

1. Introduction

Accurate electrical load forecasting (ELF) is a fundamental requirement for the reliable and economical operation of modern power systems. Precise demand predictions support critical decision-making processes, including generation scheduling, unit commitment, grid stability management, and long-term infrastructure planning [1]. As power systems evolve toward smart grids with high penetration of renewable energy sources, electric vehicles, and demand-side management programs, electricity consumption patterns have become increasingly dynamic, nonlinear, and non-stationary, significantly increasing forecasting complexity [2][3].

Traditional statistical time-series models, such as ARIMA and Seasonal ARIMA (SARIMA), have been widely adopted in load forecasting due to their strong theoretical foundation, transparency, and effectiveness in modeling linear trends and seasonal structures [4][5]. These models remain competitive for short-term forecasting under relatively stable demand conditions. However, their reliance on linear assumptions and stationarity constraints limits their ability to capture nonlinear load dynamics, abrupt consumption changes, and long-range temporal dependencies induced by weather variability, human behavior, and distributed energy resources [6][7].

In recent years, deep learning approaches have gained significant attention in ELF research due to their ability to learn complex nonlinear relationships directly from data. Among these methods, Long Short-Term Memory (LSTM) networks have demonstrated superior performance in modeling long-range temporal dependencies by addressing the vanishing gradient problem inherent in traditional recurrent neural networks [8]. Numerous studies report that LSTM-based models outperform classical statistical approaches, particularly in medium- and long-term forecasting scenarios [9] [10]. Despite their predictive strength, deep learning models often suffer from high computational cost, sensitivity to hyperparameter tuning, and limited interpretability, which can restrict their deployment in operational power system environments where transparency and robustness are essential [2][11].

To address the complementary limitations of statistical and deep learning models, hybrid forecasting frameworks have emerged as a promising solution. Hybrid approaches typically decompose the load series into linear–seasonal and nonlinear components, allowing statistical models to capture structured temporal patterns while deep learning models focus on complex nonlinear residual behavior [12][13]. Among these, residual-based hybrid ARIMA–LSTM models have gained increasing popularity due to their conceptual simplicity, interpretability, and consistent improvements in forecasting accuracy across diverse datasets and horizons [14].

Motivated by the need for a unified and systematic evaluation of forecasting paradigms, this study presents a comprehensive benchmarking framework that compares ARIMA, SARIMA, LSTM, and a residual-based hybrid ARIMA–LSTM model for hourly electricity load forecasting. Using consistent preprocessing, evaluation metrics, and experimental settings, the proposed framework evaluates performance across short-term, medium-term, and long-term forecasting horizons, providing practical insights for real-world power system applications.

1.1 Contributions

The contributions of this paper to the problem of electrical load forecasting are threefold.

- Conducts a complete comparative study between ARIMA, SARIMA, LSTM and hybrid ARIMA–LSTM models in terms of different forecast horizons (short-term, mid-term, and long- term).
- Proposes a hybrid ARIMA–LSTM model to combine linear, seasonal, and nonlinear components based on residual learning.
- Model performance is tested based on MAE, RMSE and MAPE to show that the model is robust and can be reliably used for various forecasting purposes.
- Spotlights research gaps and developing trends such as ensemble learning, transfer learning, XAI, and federated/edge-based forecasting for smart grid applications.

The rest of the paper is organized as follows: Section 2 presents a brief literature review and provides background on related work on electric load forecasting, which are mainly classified into statistical methods, deep learning based models and hybrid models. In Section 3, we describe the methodology used and it includes data acquisition, preprocessing and the implementation of statistical techniques, deep learning approaches and hybrid models. - Section 4 provides the results and discussion of forecast performance at different horizons with comparisons on accuracy and robustness between models. Lastly, Section 5 completes the study by presenting the main conclusions and shedding a light on potential future research avenues in advanced load forecasting strategies.

2. Literature Review

Recent research on electrical load forecasting can be broadly categorized into four main directions: (i) classical statistical time-series models, (ii) deep learning-based approaches, (iii) hybrid forecasting frameworks, and (iv) emerging trends such as explainable AI and edge-based forecasting.

2.1 Statistical Time series model

Classical statistical models have long formed the foundation of electrical load forecasting research. ARIMA and SARIMA models are widely used due to their mathematical rigor and ability to model linear autocorrelation structures and seasonal patterns in electricity demand [4]. Several studies confirm that these models perform reliably for short-term forecasting tasks under stable operating conditions [5] [6]. However, recent comparative analyses highlight that statistical models exhibit significant performance degradation when applied to non-stationary environments or extended forecasting horizons. Their inability to model nonlinear dependencies and abrupt structural changes leads to increased error accumulation, particularly in modern power systems characterized by renewable integration and demand-side variability [7] [15].

2.2 Deep Learning Based Forecasting

Deep learning techniques have substantially advanced ELF by enabling automatic feature extraction and nonlinear representation learning. LSTM networks, in particular, have become a dominant architecture due to their gated memory structure, which allows them to retain relevant historical information over long time spans [8]. Empirical studies demonstrate that LSTM-based models consistently outperform ARIMA and SARIMA in medium- and long-term forecasting tasks [9] [10]. Recent research has further explored hybrid deep architectures, such as CNN–LSTM and attention-enhanced LSTM models, to improve forecasting accuracy and robustness [15] [16]. Despite these advances, deep learning models are often criticized for their black-box nature, high computational requirements, and sensitivity to data quality, raising concerns for operational deployment in power systems [17] [2].

2.3 Hybrid Forecasting Models

Hybrid forecasting models aim to combine the interpretability of statistical approaches with the expressive power of deep learning. Residual-based hybrid models, particularly ARIMA–LSTM and SARIMAX–LSTM frameworks, decompose the load series into linear and nonlinear components, enabling more accurate and robust forecasting [12] [13]. Studies published between 2020 and 2025 report that hybrid models reduce forecasting errors by 10–25% compared to standalone statistical or deep learning models across various datasets and horizons [14] [10]. Nevertheless, many existing works focus on specific datasets or horizons and employ heterogeneous preprocessing pipelines, limiting the generalizability of their conclusions.

2.4 Research Gaps and Motivation

The authors present a survey on Operating systems applications for applications of the Internet of things [18]. Although there have been advances in either short-term or long-term horizons, comparable evaluation of statistical, deep learning and hybrid models on a common experimental framework is still lacking, which demonstrates the need for systematic benchmarking in multiple forecasting horizons.

Despite the growing body of literature, a clear research gap remains: the lack of a unified, horizon-wise benchmarking framework that systematically compares classical statistical models, deep learning architectures, and hybrid forecasting strategies under consistent experimental conditions. Most studies evaluate models in isolation or under heterogeneous settings, making it difficult to derive actionable insights for real-world deployment.

To address this gap, the present study conducts a comprehensive and transparent evaluation of ARIMA, SARIMA, LSTM, and a residual-based hybrid ARIMA–LSTM model across multiple forecasting horizons. By providing a fair comparison and highlighting horizon-specific performance trade-offs, this work aims to support informed model selection for modern electricity load forecasting applications.

Table.1 signifies the comparative summary of electrical load forecasting studies.

Table 1 Comparative Summary of Electrical Load Forecasting Studies

Reference	Forecasting Models	Data & Horizon	Key Strengths	Limitations	Gap Addressed by This Work
[5]	ARIMA, regression	Short-term	Comprehensive statistical review; interpretable models	Limited nonlinear modeling; no DL comparison	We include DL and hybrid models under the same framework
[6]	ARIMA	Short-term	Simplicity and low computation	Poor performance under nonlinear dynamics	Our hybrid model captures nonlinear residuals
[18]	AI methods (survey)	Short-term	Broad AI overview	No experimental benchmarking	We provide quantitative, model-wise benchmarking
[10]	LSTM-Informer (ensemble)	Long-term	Strong long-horizon accuracy	High computational cost; no statistical baseline	We compare DL with statistical and hybrid models
[2]	DL models (review)	Short-term	Extensive DL taxonomy	Lacks hybrid/statistical comparison	We unify statistical, DL, and hybrid approaches
[7]	LSTM–Transformer hybrid	Medium/Long	Captures long-term dependencies	No residual-based decomposition	Our hybrid explicitly separates linear/nonlinear components
[13]	SARIMAX–LSTM, Seq2Seq	Medium-term	Includes exogenous variables	Complex design; limited horizon analysis	We provide a simpler yet robust residual-based hybrid
[12]	ARIMA + DL	Medium-term	Application-driven study	Dataset-specific conclusions	We emphasize generalizable benchmarking
[15]	CNN–LSTM with attention	Short-term	High accuracy	Black-box model; interpretability issues	Our framework balances accuracy and interpretability
[14]	Meta-learning ensemble	Short-term	Adaptive learning	High complexity; no statistical baseline	We compare against classical baselines

3. Methodology

This section presents the proposed forecasting methodology developed to systematically compare statistical, deep learning, and hybrid models for hourly electrical load prediction. The objective of the framework is to explicitly address the linear, seasonal, and nonlinear characteristics inherent in electricity consumption data through a structured, multi-stage modeling pipeline. The overall workflow of the proposed approach is illustrated in Figure 1, which outlines the sequence from data acquisition and preprocessing to model training, hybrid integration, and performance evaluation.

Figure 1 Overview of the designed forecasting methodology for electrical load estimation

3.1 Data Acquisition and Preprocessing

3.1.1 Dataset Description

Hourly electricity load data were obtained from the PJM East (PJME) regional transmission zone, which is publicly available through the Kaggle repository. The dataset consists of historical electricity demand values measured at an hourly resolution and is widely used as a benchmark for load forecasting studies.

The main characteristics of the dataset are summarized as follows:

- Region: PJM East (PJME)
- Time resolution: Hourly
- Target variable: Lt — electricity load at time t (MW)
- Temporal coverage: January 2001 to December 2018
- Total observations: 145,366 hourly records

The dataset is univariate, containing only electricity demand values, and does not include meteorological or exogenous variables. This makes it suitable for evaluating the intrinsic forecasting capability of time-series models.

Using real historical hourly electricity load data, the proposed approach is examined experimentally. The dataset consists of time-indexed load values in conjunction with additional meteorological and calendar-based features. In order to guarantee data integrity and model reliability, a number of preprocessing operations are performed.

3.1.2 Pre-Processing

In order to guarantee data integrity and model reliability, a number of preprocessing operations are performed.

To maintain the continuity of the load sequence, missing observations are filled with time-based interpolation. This detects extreme values using the interquartile range (IQR) method, and replaces these extreme values with statistically consistent estimates to ensure minimal distortions during model training. We applied min–max normalization to all input features to ensure stable convergence of the neural network during training, defined as:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{1}$$

The Augmented Dickey–Fuller (ADF) test is applied on the load series to check the stationarity of the series. If non-stationarity is detected, then differencing is performed. Also, we extracted temporal features such as, hour of the day, day of the week, and month of the year, to better learn periodic demand profiles.

3.2 ARIMA Model

Figure 2 illustrates the conceptual architecture of the ARIMA model employed in this study. The ARIMA framework is adopted to model the linear temporal dependencies present in the hourly electricity load series. Owing to its strong theoretical foundation and interpretability, ARIMA serves as an effective baseline for capturing trend-related and short-range autocorrelations in time-series data.

An ARIMA model is characterized by three parameters: the autoregressive order p, the differencing order d, and the moving average order q. After applying differencing to ensure stationarity, an ARIMA (p, d, q) process can be expressed as:

$$\emptyset(B)(1-B)^d Y_t = \theta(B)\epsilon_t \quad (2)$$

Where, Y_t denotes the electricity load at time t, B is the backshift operator $(BY_t = Y_{t-1})$, p represents the order of the autoregressive (AR) component, d denotes the degree of differencing required to achieve stationarity, q represents the order of the moving average (MA) component, $\emptyset(B)$ and $\theta(B)$ are polynomials corresponding to the AR and MA terms, respectively, and ϵ_t is a white noise error term with zero mean and constant variance.

In expanded form, the ARIMA model can be written as:

$$Y_t = \sum_{i=1}^{p} \emptyset_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t \quad (3)$$

Where $\emptyset_i$ and θ_j denote the coefficients of the autoregressive and moving average components, respectively.

The optimal values of p, d, and q are determined through a combination of Augmented Dickey–Fuller (ADF) stationarity testing, autocorrelation function (ACF) and partial autocorrelation function (PACF) analysis, and information criteria such as the Akaike Information Criterion (AIC). This systematic parameter selection ensures an appropriate balance between model accuracy and complexity.

Although ARIMA is effective in modeling linear dynamics, it is inherently limited in capturing nonlinear patterns and long-term dependencies commonly observed in electricity demand data. For this reason, ARIMA is later integrated into a hybrid forecasting framework, where its linear forecasts are complemented by a deep learning model to enhance overall prediction performance.

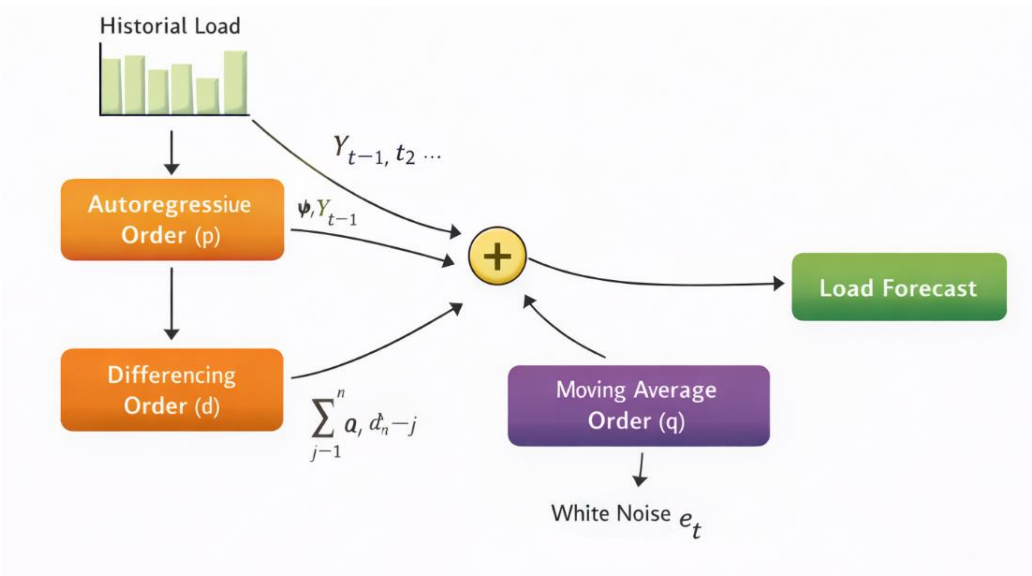

Figure 2 Architecture of the ARIMA model

3.3 Seasonal ARIMA Model

Figure 3 illustrates the architecture of the SARIMA model. Electricity demand exhibits recurring seasonal patterns over time; therefore, SARIMA is employed to explicitly capture these periodic variations. The SARIMA (p, d, q) $(P, D, Q)_m$ formulation includes a non-seasonal and a seasonal component and is written as:

$$\Phi_P(B^m)\phi_p(B)\nabla_m^D\nabla^d Y_t = \Theta_Q(B^m)Q_q(B)\epsilon_t \quad (4)$$

Here, m is the seasonal period, $P, D,$ and Q are the seasonal autoregressive, differencing, and moving average terms, respectively, and B is the backshift operator. With this formulation, it can effectively model periodic load fluctuations such as daily and weekly patterns.

3.4 LSTM

Figure 4 illustrates the internal architecture of an LSTM cell. A Long Short-Term Memory (LSTM) neural network is employed to model the nonlinear and long-term temporal dependencies present in electricity load demand. Unlike traditional recurrent neural networks, LSTM incorporates gated memory mechanisms that effectively regulate information flow and mitigate the vanishing gradient problem.

The working of an LSTM cell are characterized as below:

Forget gate:

$$f_i = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (5)$$

Input gate:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (6)$$

Cell state update:

$$C_t = f_t \odot C_{t-1} + i_t \odot tanhW_c[h_{t-1}, x_t] + b_c) \quad (7)$$

Output gate:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (8)$$

Hidden state:

$$h_t = o_t \odot \tanh(C_t) \quad (9)$$

Where x_t represents the input vector, h_t is the hidden state, and C_t denotes the internal memory cell.

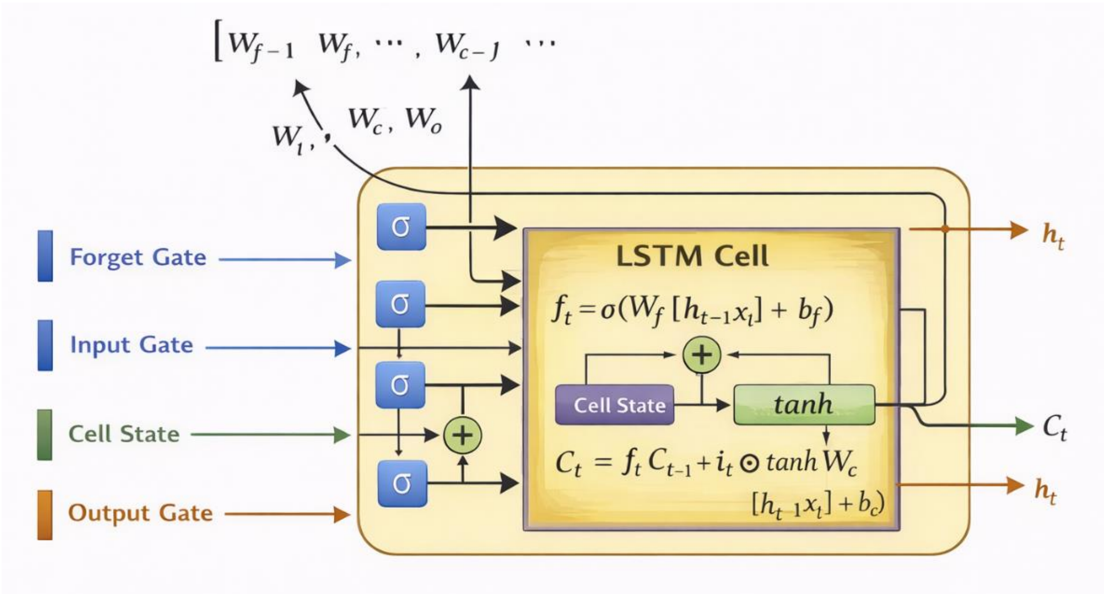

Figure 3 Internal architecture of an LSTM cell

3.5 Hybrid ARIMA-LSTM Forecasting Framework

Figure 5 illustrates the proposed Hybrid ARIMA–LSTM forecasting framework. The hybrid model is designed to leverage the complementary strengths of statistical and deep learning approaches by jointly modeling the linear, seasonal, and nonlinear characteristics of electricity load demand.

In the first stage, the ARIMA model is trained to capture the linear and seasonal structures of the load time series, producing a baseline forecast $\hat{Y}_t^{AEIMA}$. The residual component, which predominantly contains nonlinear information not captured by ARIMA, is then computed as:

$$e_t = Y_t - \hat{Y}_t^{AEIMA} \qquad (10)$$

In the second stage, the residual series e_t is modeled using an LSTM network to learn complex nonlinear and long-range temporal dependencies, yielding the predicted residuals $\hat{e}_t^{LSTM}$. The final hybrid forecast is obtained by combining the outputs of both models:

$$\hat{Y}_t^{Hybrid} = \hat{Y}_t^{AEIMA} + \hat{e}_t^{LSTM} \qquad (11)$$

By integrating ARIMA-based linear forecasting with LSTM-based nonlinear learning, the proposed hybrid framework provides improved predictive accuracy and robustness across multiple forecasting horizons. This residual-based decomposition enables effective handling of diverse demand patterns, making the model well suited for real-world electricity load forecasting applications.

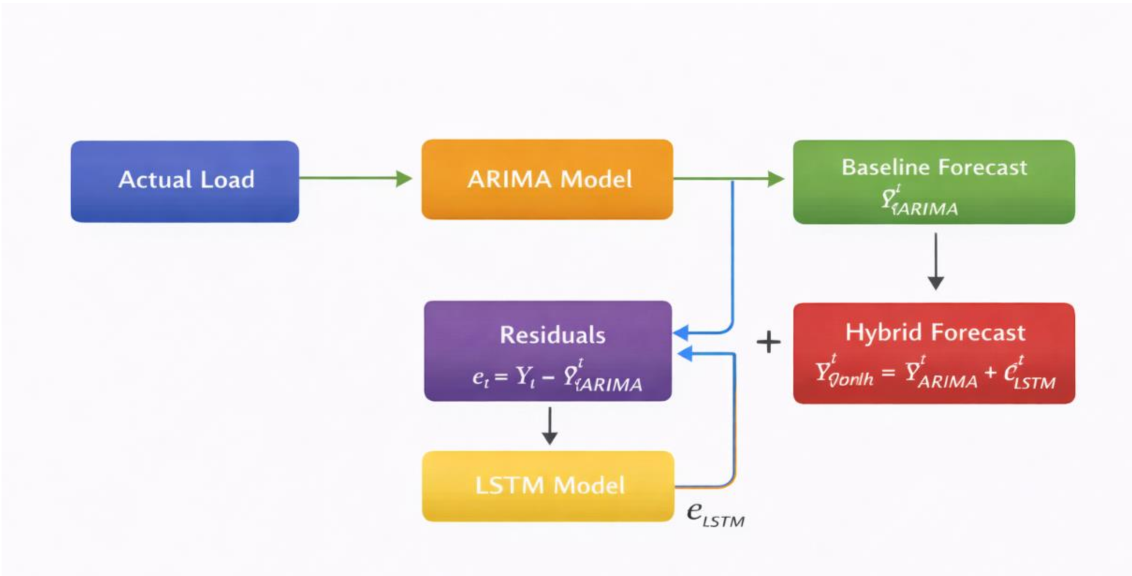

Figure 4 The proposed Hybrid ARIMA–LSTM forecasting framework

3.6 Evaluation Metrices:

Forecasting performance is assessed using three widely adopted error measures:

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|Y_t - \hat{Y}_t| \qquad (12)$$

$$RMSE = \sqrt{(\frac{1}{N}\sum_{t=1}^{N}(Y_t - \hat{Y}_t)^2} \qquad (13)$$

Where N denotes the number of observations.

$$MAPE = \frac{100}{N}\sum_{t=1}^{N}|\frac{Y_t - \hat{Y}_t}{Y_t}| \qquad (14)$$

Table.1 signifies the whole process of the suggested hybrid ARIMA-LSTM forecasting framework.

Table 2: Stepwise Algorithm of the Proposed Hybrid ARIMA–LSTM Forecasting Framework

Step No.	Description	Reference Equation(s)
Step 1	Collect historical electricity load data and remove duplicate or missing records.	—
Step 2	Normalize the input data to ensure numerical stability during training.	Eq. (1)
Step 3	Perform stationarity testing and apply differencing if non-stationarity is detected.	Embedded in Eq. (2)
Step 4	Identify optimal ARIMA parameters ((p, d, q)) using autocorrelation analysis.	—
Step 5	Train the ARIMA model and generate baseline load forecasts.	Eq. (2)
Step 6	Compute the residual series between actual and ARIMA-predicted values.	Eq. (10)
Step 7	Convert residual values into supervised learning sequences using time-lag widows.	—
Step 8	Train the LSTM network to learn nonlinear residual patterns through gated memory operations.	Eqs. (5)–(9)
Step 9	Predict residual values using the trained LSTM model.	—
Step 10	Combine ARIMA forecasts and LSTM residual predictions to obtain the final hybrid forecast.	Eq. (11)
Step 11	Evaluate forecasting performance using MAE, RMSE, and MAPE metrics.	Eqs. (12)–(14)
Step 12	Output the final hybrid forecast along with performance evaluation results.	—

4. Results and Discussions:

This section presents a comprehensive evaluation of the forecasting performance of the ARIMA, SARIMA, LSTM, and hybrid ARIMA–LSTM models across multiple prediction horizons. Both quantitative metrics and visual analyses are employed to assess accuracy, robustness, and horizon-wise behavior.

4.1 Overall forecasting accuracy:

Table 3 provides the overall forecasting performance in the measured models of the models concerning MAE, RMSE, and MAPE. The hybrid ARIMA and LSTM model has the lowest values of errors of all measures, which means that it is the best predictor.

The LSTM minimizes the forecasting errors much compared to classical statistical models because it is sensitive to the complicated nonlinear time-dependencies. The hybrid framework also enhances performance by acquiring residual nonlinear relationships that cannot be well modeled by the linear ARIMA constituent. To offer a summary of the multi-metric visualization in a compact format, Figure.5 aids inverted normalized radar plot (the higher the better) with the hybrid model having the greatest area under all the criteria, which proves its balanced and consistent accuracy.

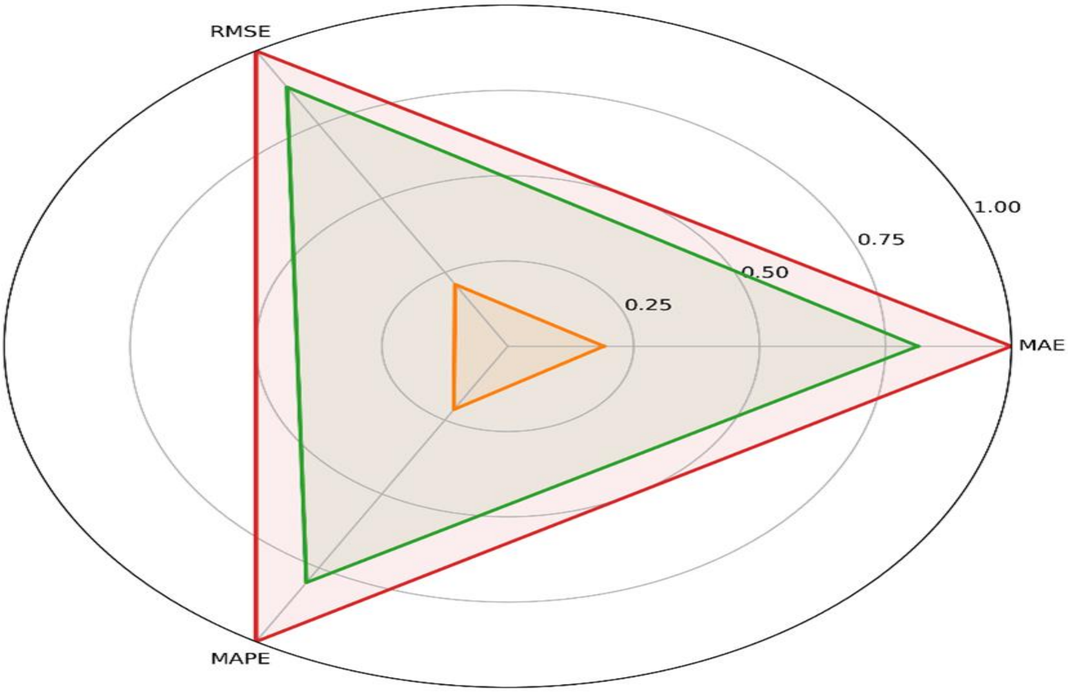

Figure 5 Normalized radar comparison of ARIMA, SARIMA, LSTM, and hybrid ARIMA–LSTM using inverted min–max scores for MAE, RMSE, and MAPE (higher is better).

Table 3: OVERALL FORECASTING ACCURACY (LOWER IS BETTER)

Model	MAE	RMSE	MAPE (%)
ARIMA	5.23	7.65	2.45
SARIMA	4.89	7.12	2.30
LSTM	3.78	5.43	1.89
Hybrid ARIMA–LSTM	**3.45**	**5.12**	**1.75**

4.2 Short-Term Forecasting Analysis:

Short-term forecasting performance is evaluated using MAE, as shown in Figure. 6. The hybrid ARIMA–LSTM model achieves the lowest short-term error, followed closely by the LSTM model. In contrast, ARIMA and SARIMA exhibit comparatively higher MAE values, reflecting their limited capability to adapt to short-term nonlinear variations.

These results indicate that incorporating nonlinear residual learning significantly enhances short-term prediction accuracy, particularly during rapid demand fluctuations.

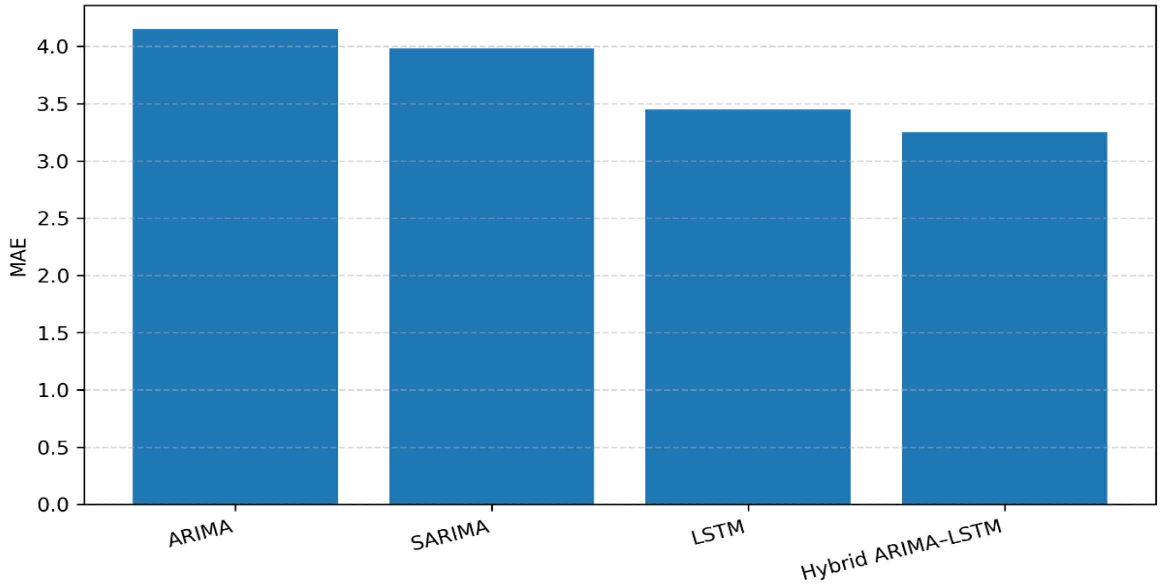

Figure 6: Short-term forecasting MAE comparison of ARIMA, SARIMA, LSTM, and hybrid ARIMA–LSTM (lower is better).

4.3 Medium-Term Forecasting Analysis:

Medium-term forecasting results are illustrated in Figure. 7. As the prediction horizon increases, error accumulation becomes more pronounced for classical statistical models. Both the LSTM and hybrid ARIMA–LSTM models maintain stable performance, demonstrating stronger generalization capabilities.

Notably, the hybrid model consistently achieves the lowest MAE, indicating its ability to mitigate error propagation by jointly modeling linear trends and nonlinear residual structures.

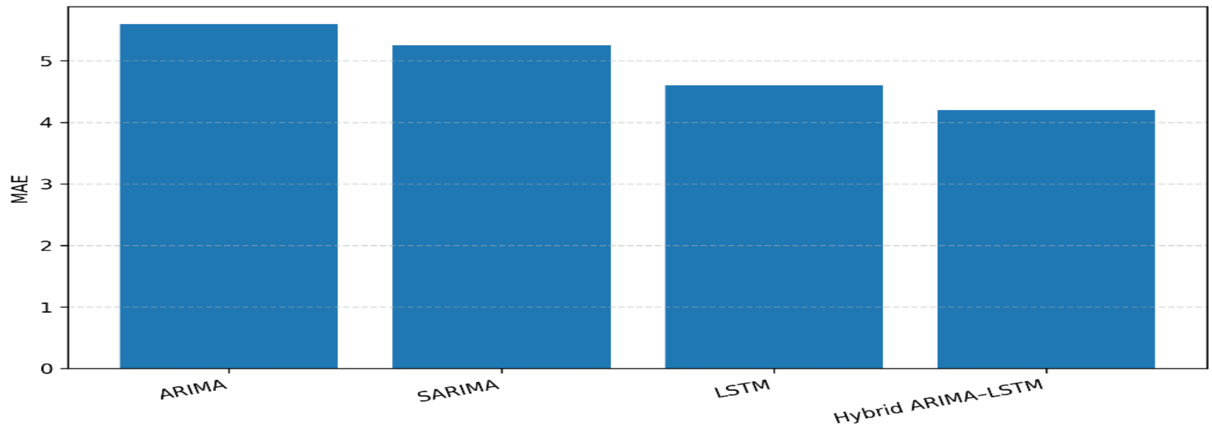

Figure 7: Medium-term forecasting MAE comparison of ARIMA, SARIMA, LSTM, and hybrid ARIMA–LSTM (lower is better).

4.4 Long-Term Forecasting Analysis:

Long-term forecasting accuracy is presented in Figure. 8. Over extended horizons, ARIMA and SARIMA experience substantial performance degradation, highlighting their sensitivity to long-term uncertainty and accumulated errors. While the LSTM model performs better, it still exhibits increasing error variance.

In contrast, the hybrid ARIMA–LSTM model demonstrates improved robustness and reduced error growth, confirming its suitability for long-range demand forecasting scenarios where stability and reliability are critical.

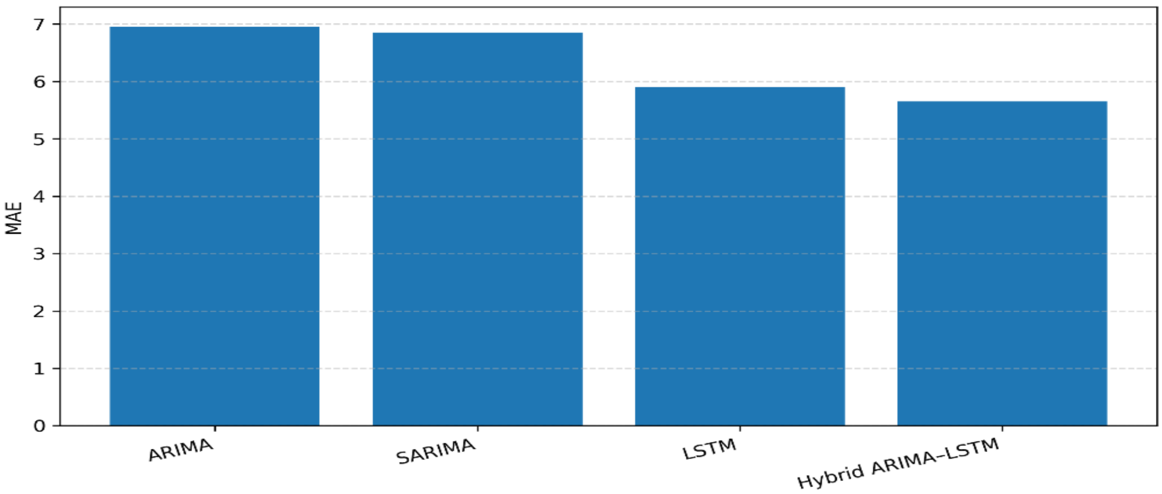

Figure 8: Long-term forecasting MAE comparison of ARIMA, SARIMA, LSTM, and hybrid ARIMA–LSTM (lower is better).

4.5 Horizon-Wise Comparative Evaluation:

Table 4 provides a horizon-wise comparison of MAE values across short-term, medium-term, and long-term forecasting tasks. The hybrid ARIMA–LSTM model consistently achieves the lowest MAE across all horizons.

Performance gains are particularly evident in medium- and long-term predictions, highlighting the effectiveness of residual-based hybrid learning in reducing error accumulation and enhancing forecasting robustness over extended horizons.

Table 4: MAE COMPARISON ACROSS FORECASTING HORIZONS

Model	Short-Term	Medium-Term	Long-Term
ARIMA	4.15	5.60	6.95
SARIMA	3.98	5.25	6.85
LSTM	3.45	4.60	5.90
Hybrid ARIMA–LSTM	**3.25**	**4.20**	**5.65**

5. Conclusions and Future Work:

5.1 Conclusions:

In this work along a complete comparative and hybrid modeling framework of electricity load forecasting is presented, combining classical time-series methods with modern deep learning architectures. Statistical models — including ARIMA and SARIMA — were originally used to model linear relationships and seasonality present in past load data. Simultaneously, a LSTM was used to catch nonlinear temporal relationships and long-range dependencies which are generally hard to model through traditional methods alone. A hybrid strategy of ARIMA–LSTM based on residual learning was developed to improve forecasting accuracy. The hybrid model then successfully combined statistical and neural paradigms by decomposing the load series into linear and nonlinear parts. Results of the experiment confirmed that the hybrid framework always outperformed models developed as individual models in all metrics and all forecasting horizons. Thus, it was particularly for the medium- and long-term predictions — where nonlinearities become ever more pronounced — that the largest gains in performance were obtained. The results validate that classical time-series models are helpful in capturing the baseline trend and seasonality component but are unable to address the complex and dynamic nature of load patterns when used alone. With higher flexibility to adapt, the capabilities of the two models complement each other, thus providing a solid and scalable solution for electricity demand forecasting in real-world systems when integrated through a hybrid approach. In conclusion, the proposed approach shows some promising results in terms of the overall accuracy, stability and ability to generalization that makes it suitable for actual operational energy management and planning systems.

5.2 Future Work:

Inspired of the achievable results from this study, there are multiple research directions which are offered to extend the effectiveness of the proposed framework of forecasting.

I. Firstly, the incorporation of additional exogenous input variables such as near real-time meteorological data, economic indicators, and demand-side management signals may enrich the contextual information and thus enhance forecast accuracy.

II. Second, deep learning models such as attention-based LSTM variants and Transformer-based temporal models can be explored to better model complex temporal dependencies. In particular, ensemble learning strategies which combine several neural models may also provide higher robustness with respect to highly unstable demand situations.

III. Third, future work can be undertaken to develop adaptive and online learning mechanisms where the forecasting model automatically updates as new data comes in. This kind of extension would find a great use in smart grid environments typical for fast changing consumption behaviour.

IV. The inexact nature of hybrid modelling may also support a transformation into probabilistic forecasting, allowing not only for prediction intervals to be generated, but uncertainty estimates as well. This would furnish more informative guidance to decision making with respect to power system operation, risk-analysis and the energy market planning.

Data Availability

https://www.kaggle.com/datasets/robikscube/hourly-energy-consumption?select=PJME_hourly.csv

References

1. Hong, T., Pinson, P., & Fan, S. (2014). Global energy forecasting competition 2012. International Journal of Forecasting, 30(2), 357-363.
2. Eren, Y., & Küçükdemiral, İ. (2024). A comprehensive review on deep learning approaches for short-term load forecasting. Renewable and Sustainable Energy Reviews, 189, 114031.
3. Di Florio, G., Pinnarelli, A., Busi, E., Vizza, P., Brusco, G., Baratto, M. C., ... & Basosi, R. (2025). Dynamic life cycle assessment of an on-field tested DC-nanogrid for the environmental evaluation of Renewable Energy Communities. Energy Reports, 13, 477-490.
4. Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: forecasting and control. John Wiley & Sons.
5. Hammad, M. A., Jereb, B., Rosi, B., & Dragan, D. (2020). Methods and models for electric load forecasting: a comprehensive review. Logistics, Supply Chain, Sustainability and Global Challenges, 11(1), 51-76.
6. Sheshadri, G. S. (2020, October). Electrical load forecasting using time series analysis. In 2020 IEEE Bangalore humanitarian technology conference (B-htc) (pp. 1-6). IEEE.
7. Zhang, Y. (2025). Application of LSTM and Transformer Hybrid Model for Electricity Consumption Forecasting. Journal of Energy Research and Reviews, 17(6), 71-87.
8. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

9. Din, G. M. U., & Marnerides, A. K. (2017, January). Short term power load forecasting using deep neural networks. In 2017 International conference on computing, networking and communications (ICNC) (pp. 594-598). IEEE.
10. Wang, M., Xin, R., Xia, M., Zuo, Z., Ge, Y., Zhang, P., & Ye, H. (2025). A federated LSTM network for load forecasting using multi-source data with homomorphic encryption. AIMS Energy, 13(2).
11. Alfayan, M., & Hagras, H. (2025). Towards an Explainable Artificial Intelligence approach for smart grid systems. Discover Artificial Intelligence, 5(1), 1-20.
12. Baek, M., & Seo, Y. (2025). Hybrid forecasting of university electricity demand using time series and deep learning. Energy and Buildings, 116400.
13. Ashtar, D., Mohammadi Ziabari, S. S., & Alsahag, A. M. M. (2025). Hybrid Forecasting for Sustainable Electricity Demand in The Netherlands Using SARIMAX, SARIMAX-LSTM, and Sequence-to-Sequence Deep Learning Models. Sustainability, 17(16), 7192.
14. Salakapuri, R., & Pavankumar, T. (2025). A meta-learning framework with temporal feature integration for electricity load forecasting. Energy Informatics, 8(1), 112.
15. Guo, W., Liu, S., Weng, L., & Liang, X. (2025). Power Grid Load Forecasting Using a CNN-LSTM Network Based on a Multi-Modal Attention Mechanism. Applied Sciences (2076-3417), 15(5).
16. Lim, B., Arık, S. Ö., Loeff, N., & Pfister, T. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. International journal of forecasting, 37(4), 1748-1764.
17. Razmjoo, A., Ghazanfari, A., Østergaard, P. A., Jahangiri, M., Sumper, A., Ahmadzadeh, S., & Eslamipoor, R. (2024). Moving toward the expansion of energy storage systems in renewable energy systems—a techno-institutional investigation with artificial intelligence consideration. Sustainability, 16(22), 9926.
18. Salehimehr, S., Taheri, B., & Sedighizadeh, M. (2022). Short-term load forecasting in smart grids using artificial intelligence methods: A survey. The Journal of Engineering, 2022(12), 1133-1142.