



Available online at www.qu.edu.iq/journalcm
JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS
 ISSN:2521-3504(online) ISSN:2074-0204(print)



A Hybrid LSTM–XGBoost Framework Optimized by Particle Swarm Optimization for Short-Term Energy Demand Forecasting Using Large-Scale Smart Meter Data

Rand Jalal^{a,*}, Intisar Al-Mejibli^b

^a Informatics Institute for Postgraduate Studies, University of Information Technology and Communications, Baghdad, Iraq.
randjalal1991@gmail.com

^b Biomedical Informatics College, University of Information Technology and Communications, Baghdad, Iraq. dr.intisar.almejibli@gmail.com

ARTICLE INFO

Article history:

Received: 11 /06/2026

Revised form: 03 /07/2026

Accepted : 05 /07/2026

Available online: 30/09/2026

Keywords:

Energy Demand Forecasting, Smart Grids, LSTM, XGBoost, Particle Swarm Optimization, Hybrid Forecasting, Walk-Forward Validation.

ABSTRACT

Rapid proliferation of cloud storage services has enlarged the issue of protecting user data from

The short term electricity demand forecast is an essential for the operation of smart grid, demand response programs, peak load management and sustainable energy planning. Given the nonlinear, dynamic and time-dependent nature of the energy demand patterns, however, predicting electricity consumption remains difficult. To tackle these challenges, the present study proposes to optimize the Hybrid LSTM–XGBoost forecasting framework by exploiting the advantage of temporal sequence learning of LSTM networks and the nonlinear regression power of Extreme Gradient Boosting (XGBoost). Moreover, Particle Swarm Optimization (PSO) algorithm is used to automatically optimize the hyperparameters of XGBoost component to improve the forecasting accuracy and eliminate the need for extensive manual tuning. A Walk-Forward Validation (WV) strategy is used during the experimental process to achieve realistic and leakage-free model assessment. The proposed framework is evaluated based on a large-scale real world data set of smart meters gathered by the Pecan Street project which produces high-resolution measurements of residential electricity consumption. A scalable pre-processing pipeline and extensive feature engineering approach are designed to capture temporal, lag-based and statistical features that represent the electricity consumption behavior well. The experimental results validate that the proposed Hybrid LSTM–XGBoost + PSO model always has better performance than the baseline XGBoost model and the non-optimized Hybrid LSTM–XGBoost model. The proposed approach gives the minimum forecasting error, maximum coefficient of determination, and consistent results in all the walk forward period validation folds. The results suggest that the deep temporal feature extraction, ensemble regression, intelligent hyperparameter tuning, and leakage-free time-aware assessment methods are an accurate, robust, and practical approach to short-term electricity demand forecasting in smart grid settings.

MSC..

<https://doi.org/10.29304/jqcm.2026.18.32974>

1. Introduction

The growth of smart grid technologies and advanced metering has resulted in larger volumes of electric power consumption data at finer time scales. Such data can now be used for new, data-driven forecasting and energy management opportunities. One of these applications is short-term electricity demand forecasting that plays an important role on grid stability, peak demand mitigation, demand response strategies and efficient energy planning in smart grid environments [20]. However, the task of forecasting electricity consumption is still difficult because of the complicated, dynamic and highly nonlinear nature of electricity consumption patterns. The use of traditional

* Rand Jalal

Email addresses: randjalal1991@gmail.com

Communicated by 'sub etitor'

statistical methods and conventional machine learning techniques has been widely used to solve electricity load forecasting problems, due to the low computation demands [7], [21]. The challenge with these approaches is that they are not always able to capture long-range temporal dependencies that may exist in sequential electricity consumption data. To overcome this restriction, LSTM networks have been widely used in recent years due to its gated memory structure which is particularly suitable for sequential modeling problems [2], [19]. Although LSTM models can be highly predictive, they can be resource-intensive and sensitive to the architecture and hyperparameters used. However, XGBoost has become one of the most successful ensemble learning algorithms for regression tasks due to its scalability, regularization properties, and powerful nonlinear modeling ability [1]. In recent years, various combinations of deep learning and ensemble learning have been studied to realize hybrid forecasting architectures to take advantage of both methods. These hybrid models have been proven to have good performance in forecasting accuracy in different energy forecasting applications [3], [4], [6], [18]. An additional key aspect that affects the forecasting outcome is the choice of the hyperparameters of the model. As a result, Hyperparameter Optimization (HPO) techniques are now a part of contemporary forecasting systems. In this context, Particle Swarm Optimization (PSO) technique has gained significant interest because of its effective search ability, fast convergence rate and capability to come up with a near-optimal parameter set to maximize the accuracy and robustness of the model [5]. Further, model evaluation methods for realistic models must take into account chronological nature of time series data to maintain the chronological nature of the data. Because of this, Walk-Forward Validation (WFOV) has been widely adopted as a valid method of assessment that avoids information leakage and offers a closer approximation to the actual operating conditions of a forecasting system [15, 16]. Although significant advances have been made in the study of electricity load forecasting, there are several limitations in the methods reported in the literature. The majority of existing studies are based on relatively small data sets, on simple preprocessing, on splits of the train–test sets that are predefined, or on evaluation procedures that do not necessarily represent realistic forecasting contexts. Furthermore, the performance of the hybrid forecasting architectures, optimization methods, and evaluation methods are often analyzed independently from each other, and not as part of a common forecasting method. Inspired by these insights, this research suggests an optimized Hybrid LSTM–XGBoost forecasting system by incorporating Particle Swarm Optimization (PSO) algorithm and testing it with the Walk-Forward Validation (WFOV) method, using a large-scale smart meter data set from the Pecan Street project. The proposed methodology combines the temporal feature learning with the ensemble-based prediction, the automated hyperparameter optimization, and the leakage-free, time-aware forecasting evaluation in a single forecasting architecture. These complementary components, along with validation with actual electricity consumption data, should provide a comprehensive and precise solution for short-term electricity demand forecasting.

2. Related Work

In this section, recent research related to short-term electricity load forecasting is summarized, focusing on deep learning-based models, hybrid forecasting architectures, optimization methods, and the evaluation methods adapted to the temporal context of electricity demand forecasts. The aim is to examine the pros and cons of the existing methods and to outline the research gap to be filled by the developed forecasting framework.

2.1 Deep Learning and Hybrid Forecasting Models

The use of smart grids and advanced metering infrastructures has grown enormously and as a result, the amount of electricity consumption data has also increased, which has motivated the application of deep learning techniques on short-term electricity load forecasting. Unlike many of the popular statistical and machine learning methods that are used in forecasting, these models have the ability to learn complex, non-linear relationships between the time series of energy consumptions as well as temporal dependencies. Within the different deep learning architectures, the Long Short-Term Memory (LSTM) networks have gained a lot of interest because they are able to process the sequential nature of the data and achieve the long-term temporal dependency with the gated architecture of their memory. Many studies have shown that the LSTM model is effective in the load forecasting problems for residential, commercial, and smart grid applications, and LSTM models are able to surpass traditional load forecasting methods in some cases [2, 4]. Although the standalone deep learning model has achieved success in recent years, the research has paid more and more attention to the hybrid forecasting model, which fuses several learning paradigms to make up for the weaknesses of individual models, so that the advantage can be achieved by combining several models. Of these methods, Hybrid LSTM–XGBoost frameworks are very popular. For these architectures, the LSTM is used to

capture intricate temporal features and relationships from past consumption data, while the XGBoost is used to model nonlinear relationships and improve the accuracy of the forecasting model. For instance, a model based on Bi-LSTM-XGBoost was proposed by Semmelmann et al. for the load forecasting of energy communities, and they have obtained satisfactory performance in terms of forecasting accuracy and peak-load prediction [18]. Similarly, recent work has demonstrated that incorporating temporal representations learnt by an LSTM network with the ensemble regression power of XGBoost can lead to better forecasting than relying on an LSTM network alone or an XGBoost ensemble alone. Most of the previous research, however, is to design hybrid models for higher forecasting accuracy, instead of using intelligent optimization techniques and realistic time aware evaluation strategies in a unified forecasting framework. As a result, the application and generalization in the real world of many hybrid forecasting models are not fully explored.

2.2 Optimization Techniques for Energy Forecasting

The selection of the hyperparameters of forecasting models has a significant impact on the ability to predict and generalize. Recently, energy forecasting applications have been a main application of metaheuristic optimization algorithms that have been widely used to automate parameter tuning and enhance energy forecasting model robustness. The techniques are designed to find near-optimal parameter configurations in large search spaces that might not be found by trial-and-error, but that are still useful for a human user. The particle swarm optimization (PSO) is one of the optimization techniques that has become very popular because of its simplicity, fast convergence and balancing between exploring and exploiting in the search process. There have been several studies reporting that when applied to the optimization of machine learning and deep learning models, PSO can significantly improve the forecasting accuracy, especially in the case of forecasting problems related to nonlinear and high dimensional systems [5]. Although these benefits are available, the majority of previous research has focused on applying PSO in the optimization of a single forecasting model. However, there have been only a few studies that have explored the use of PSO in hybrid systems that integrate deep learning and ensemble learning methods in forecasting systems. Thus, the potential value of hybridizing forecasting models and intelligent hyperparameter optimization in the context of short-term electricity load forecasting has not been thoroughly investigated.

2.3 Time-Aware Evaluation Strategies and Research Gap

The performance of a forecasting system cannot just be judged by its predictive architecture; it is also very much dependent on the assessment framework employed during the assessment of the forecasting system. Inappropriate validation procedures can lead to misleading results in time-series forecasting, since information from future observations can be used to indirectly influence the training of the model. Therefore, the forecasting performance reported may not be representative of the true forecasting power of the model. To address this challenge, the time-aware validation techniques such as Walk-Forward Validation (WFOV) and Rolling-Origin Evaluation are popular methods for electricity load forecasting studies [15] and [16]. Such methods differ from traditional methods of validation by preserving the temporal order of observations during the entire process of evaluation. Models are learned from past data and then evaluated on future data not seen during learning. This evaluation setting is more representative of operational forecasting settings and allows for more credible evaluation of model generalization performance.

Table 1. Summary of Representative Studies on Electricity Load Forecasting

Reference	Year	Model	Dataset	Forecast Horizon	Main Findings	Limitation
Semmelmann et al.	2022	Bi-LSTM-XGBoost	Smart Meter Data	Day-ahead	Improved forecasting accuracy and peak-load prediction	No optimization algorithm
Lu et al.	2024	Hybrid LSTM	Electrical Load Data	Day-ahead	Improved forecasting performance using hybrid learning	No ensemble regression model

Reference	Year	Model	Dataset	Forecast Horizon	Main Findings	Limitation
Dakheel et al.	2025	LSTM–XGBoost	Smart Grid Data	Short-term	Enhanced prediction accuracy through hybrid modeling	No walk-forward validation
Mazibuko et al.	2025	LSTM–XGBoost	Energy Consumption Data	Short-term	Outperformed standalone forecasting models	Limited evaluation strategy
Liu et al.	2025	BiLSTM–RF–XGBoost	Building Electricity Data	Short-term	Achieved superior forecasting accuracy using stacked hybrid learning	Different application domain
Proposed Study	—	Hybrid LSTM–XGBoost + PSO	Pecan Street Smart Meter Data	Short-term	Integrates deep temporal feature extraction, ensemble regression, PSO-based hyperparameter optimization, leakage-free walk-forward validation, and large-scale real-world smart meter data within a unified forecasting framework	—

Table 1 offers a comparative review of some representative studies for short-term electricity load forecasting, which includes the forecasting technique used, the datasets used, the main contributions of the studies and the limitations identified. The literature review shows that deep learning algorithms, hybrid learning frameworks, and metaheuristic optimization techniques have helped improve the accuracy of forecasting a considerable amount. However, there are still some methodological problems that constrain the application and generalization of many current approaches. A key finding from the literature is that many studies have focused on improving forecast accuracy by developing a hybrid model, while on the other hand a lesser number of studies have paid attention to the inclusion of intelligent hyper parameter optimization mechanisms. Others have investigated optimization methods separately, in many cases to stand-alone forecasting systems, or in limited experimental scenarios. Furthermore, the lack of care taken in preserving the temporal dependencies of electricity consumption time series in the construction of fixed train–test partitions and the traditional validation process may impact the reliability of the reported forecasting results. While Hybrid LSTM–XGBoost models have proven to be promising for temporal representation learning, ensemble-based regression, automated hyperparameter optimization and leakage-free evaluation, few studies have integrated all the above components into a single framework for temporal forecasting. In addition, although extensive validation with large-scale real-world smart meter datasets is crucial to gauge the robustness of the model in a realistic operating environment, this is relatively under explored. According to the above observations, a huge research potential can be seen for developing an integrated forecasting system that can integrate temporal feature learning, ensemble prediction, intelligent optimization and chronologically consistent evaluation procedures. In this work, an optimized Hybrid LSTM–XGBoost framework has been proposed to meet this requirement by optimizing the configuration of the proposed model using Particle Swarm Optimization (PSO), and evaluating their performance in a realistic and leakage-free way using Walk-Forward Validation (WV). The framework is tested with large-scale smart meter data attained from the Pecan Street project with the goal of providing a forecasting solution that is both accurate and robust, and can be used in real-world short-term electricity demand forecasting applications.

3. Methodology

A methodological system of the proposed forecasting system is described here, including the description of the dataset, data preprocessing, feature engineering, predictive modelling, hyperparameter optimization, system validation strategy and performance evaluation.

3.1 Dataset Description

In this study, Pecan Street Smart Grid Dataset was used as one of the most widely used benchmark datasets for smart grid analytics and electricity load forecasting studies [17]. The data set consists of high resolution real residential electricity consumption data from smart meters. Each household has its own dataid, allowing analysis of individual consumption patterns, and development of data-driven forecasting models. The electricity demand from the power grid (grid_x) was used as the target variable in this study. Since the dataset was a time series and the consumption pattern was complicated, the following data pre-processing procedures are applied before the model is developed. These were timestamp organization, data quality checks, working with missing data, and creating features such as temporal and lag features to reflect historical consumption patterns. The selection of the dataset was based on the availability of realistic and high resolution energy consumption data for residential buildings, being representative of the electricity consumption of households in a smart grid environment. These features make it ideal for testing short-term load forecasting models and understanding their performance in capturing time-of-day variations and demand fluctuations in real-world operating conditions.

3.2 Data Preprocessing

The raw data of the smart meter is huge and has very high time resolution, in order to prevent high consumption of memory and instability of calculation, the scalable preprocessing pipeline was implemented according to the principle of chunk-wise processing. The timestamp data was first demystified and converted to a standard datetime format. The observations were then sorted chronologically, maintaining the time sequential nature of the electricity consumption records. Inaccuracies and non-physical consumption values were checked and eliminated. To smooth the data and remove noise, and to get a more regular temporal resolution, the data was then aggregated to a 15-minute time step using a mean aggregation method. Handling of missing observations was done systematically to improve data quality and model reliability. Lastly, the cleaned data was saved in the efficient Parquet columnar format, which is suitable for later feature engineering and model training phases.

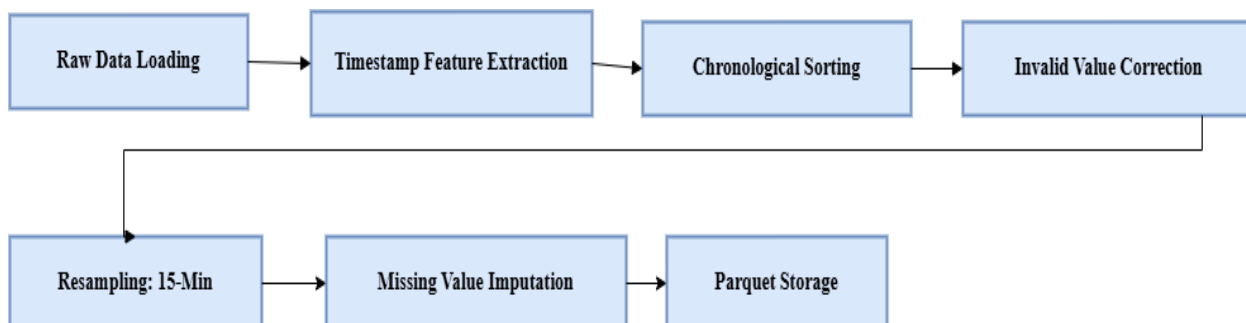


Fig. 1 .Data preprocessing pipeline applied to the raw smart meter data prior to feature engineering and model training.

3.3 Feature Engineering

To improve the predictive capability of the proposed forecasting framework, a comprehensive feature engineering process was performed to extract informative patterns from the electricity consumption time series. Since electricity demand exhibits strong temporal characteristics, a range of calendar-based attributes was generated from the timestamp information, including the hour of the day, day of the week, month of the year, and weekend status. These attributes provide valuable contextual information that helps the forecasting models recognize repetitive consumption behaviors and seasonal variations occurring at different temporal scales.

In addition to calendar-related attributes, historical dependency information was incorporated through the creation of lag-based features. These variables capture the relationship between current electricity demand and previously observed consumption values, allowing the models to exploit temporal correlations embedded within the data. To further enrich the feature space, statistical descriptors were generated using Rolling Mean and Exponentially Weighted Moving Average (EWMA) calculations. Such features summarize recent consumption dynamics, highlight underlying trends, and mitigate the influence of short-term irregular fluctuations.

By integrating calendar-based, lag-derived, and statistical features within a unified feature set, a richer representation of electricity consumption behavior was obtained. This enhanced representation supports both machine learning and deep learning algorithms in capturing complex temporal structures and contributes to improved forecasting accuracy for short-term electricity demand prediction.

Table 2. Summary of Engineered Features

Feature Category	Description
Temporal Features	Hour of Day, Day of Week, Month, Weekend Indicator
Lag Features	Historical electricity consumption values used to capture temporal dependencies
Statistical Features	Rolling Means and Exponentially Weighted Moving Averages (EWMA)
Target Variable	Grid electricity demand (grid_x)

3.4 Predictive Models

This study developed three forecasting models.

3.4.1 XGBoost Baseline

The engineered time-series features were used to train a baseline XGBoost regression model because it is scalable, efficient and performs well on nonlinear regression tasks [1]. In XGBoost, a series of regression trees are created and the prediction is made by aggregating results of weak learners, following the gradient boosting approach. The XGBoost prediction function is mathematically described in Equation (1):

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i) \quad (1)$$

where $\hat{y}_i$ denotes the predicted electricity demand, f_k represents the k -th regression tree, and K is the total number of trees used in the ensemble model.

3.4.2 Hybrid LSTM-XGBoost

An LSTM network was used as a deep temporal feature extractor for capturing the complex temporal dependency in the electricity consumption pattern. The LSTM model takes in the sequential input data and outputs the latent representations that encode the short-term and long-term temporal dynamics. The hidden state generation process in the LSTM model is shown as in Equation (2).

$$h_t = LSTM(x_t, h_{t-1}) \quad (2)$$

where x_t denotes the input sequence at time step t , h_t represents the hidden state at time step t , and h_{t-1} denotes the hidden state from the previous time step.

The extracted latent features generated by the LSTM network were combined with the manually engineered temporal and statistical features and subsequently provided as input to the XGBoost model. The hybrid design combines the forecasting power of LSTM and the nonlinear regression power of XGBoost to achieve a better forecasting result, and offers better representation of a complex electricity consumption behavior [4].

Table 3. LSTM Architecture Configuration

Parameter	Value
Sequence Length	24
LSTM Layers	1
Hidden Units	64
Optimizer	Adam
Loss Function	Mean Squared Error (MSE)
Epochs	12
Batch Size	64

3.4.3 Hybrid LSTM-XGBoost + PSO

To improve the forecast results, Particle Swarm Optimization (PSO) was used: To choose the best hyperparameters for the XGBoost part without changing the LSTM architecture. The particle positions are iteratively changed in the search space until the particle reaches the optimum position in the search space. Best parameters to set to reduce forecasting error.

The particle velocity update mechanism is expressed in Equation (3):

$$vi(t + 1) = wvi(t) + c1r1(pbesti - xi(t)) + c2r2(gbest - xi(t)) \quad (3)$$

The particle position update mechanism is given in Equation (4).

$$xi(t + 1) = xi(t) + vi(t + 1) \quad (4)$$

where vi is the velocity of the particle, xi is the position of the particle, $pbesti$ is the personal best solution for the particle and $gbest$ is the global best solution for the swarm. To avoid manual tuning procedures and to automate the hyperparameter selection, the PSO algorithm was used. The optimization process aimed at finding combinations of parameters that reduce the error in the forecasts, while preserving the ability of the model to generalize the forecasts within the framework of the time-series validation.

Table 4. PSO Configuration

Parameter	Value
Swarm Size	12
Iterations	12
Inertia Weight (w)	0.7
Cognitive Coefficient (c1)	1.5

Parameter	Value
Social Coefficient (c2)	1.5

The hyperparameters of XGBoost were optimized using the PSO algorithm, such as the learning rate. The number of rate, maximum tree depth, subsampling ratio, and column sampling ratio. estimators' goal of reducing the error in forecasting and enhancing the generalizability of the model

3.5 Walk-Forward Validation

Throughout the experimental evaluation process, a Walk-Forward Validation (WFOV) framework was used to get a reliable evaluation of the performance of the forecasts [15,16]. For this reason, WFOV retains the time structure in the data and tests the models in a way that matches real forecasting applications, where models are evaluated based on the data available until a certain time. An expanding-window strategy with four folds was used for implementing the evaluation procedure. Firstly, 60% of the data was used for model development and the remaining 10% was reserved for testing. The training window was expanded by adding more historical observations after each evaluation stage while the testing window remained the same as it represented unseen future observations. The same process was iterated for each fold; the forecasting models were given more and more training data, while at the same time the training and test sets were clearly separated in time. Use of this evaluation framework improves the credibility of the experimental results as it would not allow for subsequent information to affect model training. The reported performance metrics are thus more realistic to reflect the forecasting capability and closer to real-world short-term electricity demand forecasting environments.

Table 5. Walk-Forward Validation Configuration

Fold	Training Data	Testing Data
Fold 1	60%	10%
Fold 2	70%	10%
Fold 3	80%	10%
Fold 4	90%	10%

3.6 Performance Metrics

The models proposed were quantitatively tested by three common performance metrics including Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and coefficient of determination (R^2). These metrics reflect the overall accuracy of the predictions, the size of the errors in the predictions, and the closeness of the fit between the actual and predicted electricity demand values [15], [16]. The RMSE is defined as in Equation (5).

$$RMSE = \sqrt{(1/n) \times \sum (y_i - \hat{y}_i)^2} \quad (5)$$

The MAE metric is defined in Equation (6).

$$MAE = (1/n) \times \sum |y_i - \hat{y}_i| \quad (6)$$

The coefficient of determination (R^2) is defined in Equation (7).

$$R^2 = 1 - [\sum (y_i - \hat{y}_i)^2 / \sum (y_i - \bar{y})^2] \quad (7)$$

where y_i denotes the actual electricity demand, $\hat{y}_i$ represents the predicted value, $\bar{y}$ is the mean of the observed values, and n denotes the total number of observations.

4. Results and Discussion

4.1 Overall Walk-Forward Forecasting Performance

This section evaluates the predictive performance of the developed forecasting models using the Walk-Forward Validation framework. The analysis focuses on examining the contribution of temporal representation learning, hybrid model integration, and PSO-based hyperparameter optimization to short-term electricity demand forecasting accuracy.

Table 6. Comparison of forecasting performance using walk-forward validation.

Model	RMSE (Mean $\pm$ Std)	MAE (Mean)	R^2 (Mean)	Description
XGBoost (Baseline)	0.0935 $\pm$ 0.0108	0.0691	0.9062	Baseline model using handcrafted time-series features only
Hybrid LSTM-XGBoost	0.0500 $\pm$ 0.0085	0.0362	0.9733	Hybrid model combining deep temporal features from LSTM with XGBoost
Hybrid LSTM-XGBoost + PSO (Proposed)	0.0483 $\pm$ 0.0083	0.0337	0.9751	Optimized hybrid model with the best overall forecasting performance

Table 6 summarizes the forecasting results obtained from all evaluated models across the walk-forward validation folds. Among the investigated approaches, the proposed Hybrid LSTM-XGBoost + PSO framework delivered the most favorable results, achieving the lowest prediction errors and the highest coefficient of determination. Specifically, the proposed model attained an RMSE of 0.0483 ± 0.0083 , an MAE of 0.0337, and an R^2 value of 0.9751, indicating superior predictive capability and strong agreement between the predicted and observed electricity demand values.

It is notable that introducing the deep temporal feature extraction into the forecasting process has benefits as observed by comparing the baseline XGBoost model with the Hybrid LSTM-XGBoost architecture. The LSTM-XGBoost hybrid framework could successfully combine the good of the sequence learning property of LSTM and the nonlinear modeling power of XGBoost, leading to a significant decrease in forecasting errors and a remarkable improvement in predictive accuracy. The effects of increasing the performance were further obtained by using Particle Swarm Optimization. PSO algorithm identified better hyperparameter values for XGBoost component which resulted in better improvements for all the evaluation metrics. The improvements made over the non-optimized hybrid model were relatively small, but the results show the potential of automated hyperparameter optimization to enhance model performance and make the process more independent of human intervention for parameter tuning. The results obtained are in accordance with the results reported in the previous studies where hybrid forecasting systems based on deep learning and ensemble-based machine learning method generally maintain better predictive performance than the individual forecasting systems in short-term forecasting of electricity demand [4]. In addition,

these findings confirm the performance of the proposed time feature learning, ensemble regression, and intelligent hyperparameter optimization within a single forecasting system.

4.2 Temporal Robustness Across Walk-Forward Folds

The proposed Hybrid LSTM-XGBoost + PSO framework was tested for model robustness using a four-fold cross-validation test scheme of the walk-forward validation approach. The aim was to examine the consistency and generalization performance of the suggested forecasting model across various time evaluation windows.

Table 7. Performance of the proposed Hybrid LSTM-XGBoost + PSO model over the all walk-forward validation folds.

Fold	RMSE	MAE	R ²
Fold 1	0.0406	0.0291	0.9773
Fold 2	0.0444	0.0289	0.9774
Fold 3	0.0485	0.0357	0.9754
Fold 4	0.0598	0.0410	0.9702

Forecasting results show a strong consistency of results over all validation folds. The coefficient of determination (R^2) was greater than 0.97 for all of the evaluation periods; slight variations can be seen between the results of Fold 1 and Fold 4 in terms of RMSE and MAE values. The results showed that the proposed forecasting framework retains the high predictive power as seen with progressively unseen future observations. The smaller differences in forecasting error across the various folds in the validation set are an indicator of good temporal stability and strong generalisation abilities. This fluctuation in electricity consumption is especially significant in the real world Smart Grid, where user operating habits and conditions can shift with time. The good performance in all the walk forward folds, then, offers further evidence of the reliability and robustness of the proposed forecasting framework.

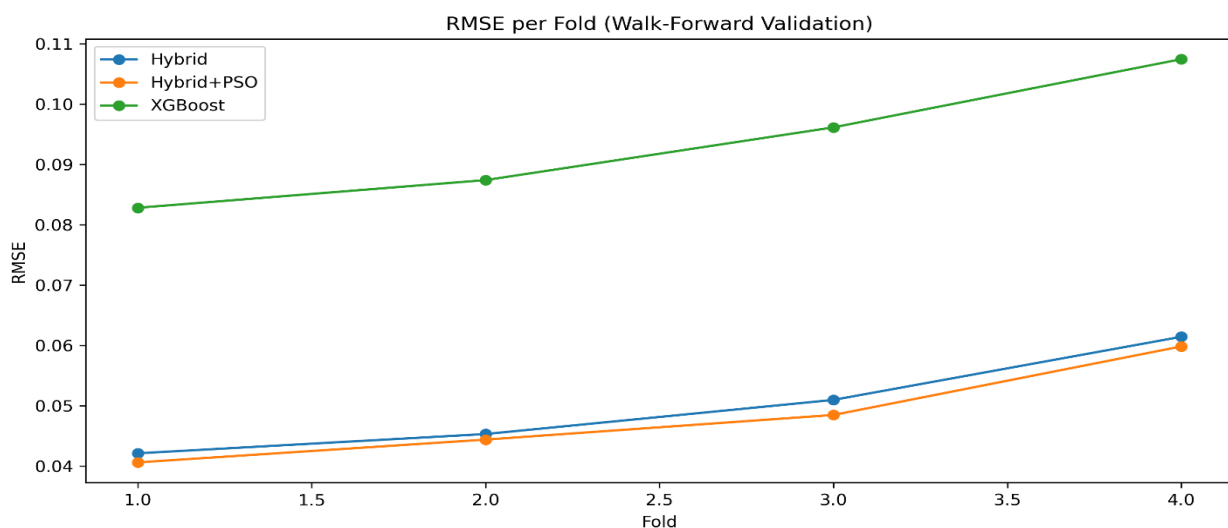


Fig. 2. Comparisons of the RMSE between the different walk-forward validation folds for XGBoost, Hybrid LSTM-XGBoost, and Hybrid LSTM-XGBoost + PSO.

By analyzing the results shown in Figure 2, it can be seen that the proposed framework achieved lower forecasting errors in all validation folds than the baseline XGBoost model and the non-optimized Hybrid LSTM–XGBoost model, further verifying the effectiveness of the proposed framework.

4.3 Visual Comparison of Forecasting Results

A visual comparison was made to evaluate the forecasting behavior of the models further. The plots in Figure 3 show the actual and predicted electricity consumption history for the last validation fold in the walk-forward test.

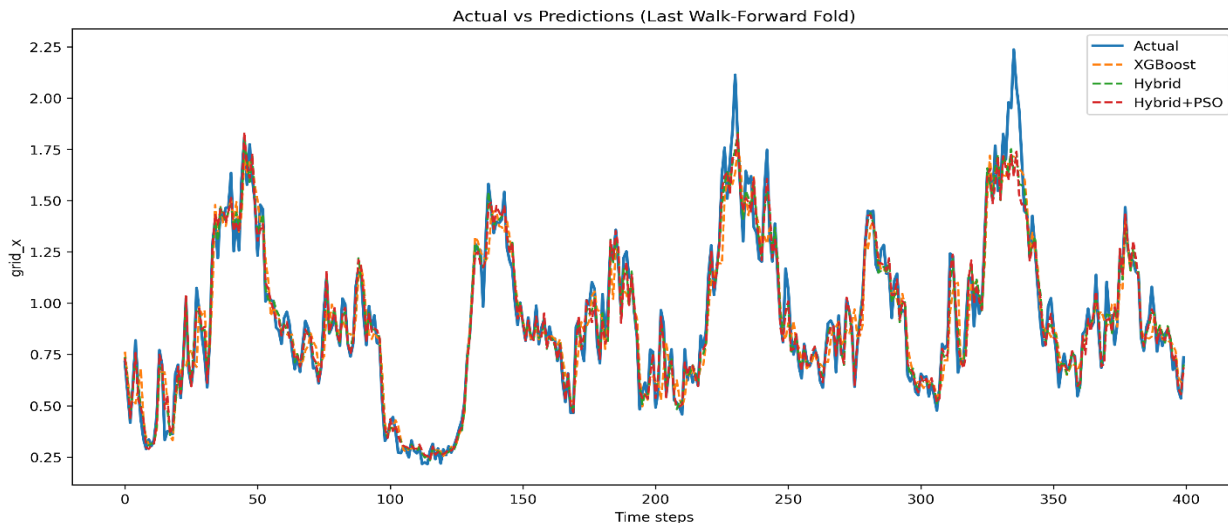


Fig. 3. The actual and predicted electricity consumption in the last validation fold (walk forward).

The results of the proposed Hybrid LSTM–XGBoost + PSO model are shown in Figure 3, which closely resembles the actual electricity consumption profile and outperforms the baseline XGBoost model and the non-optimized Hybrid LSTM–XGBoost model. Throughout the evaluation period, there appears to be a better picture of short-term fluctuations and overall consumption trends, as indicated by visual inspection. Using the predictions produced by the proposed forecasting framework, a scatter plot analysis was carried out to further explore the agreement between the observed and predicted values.

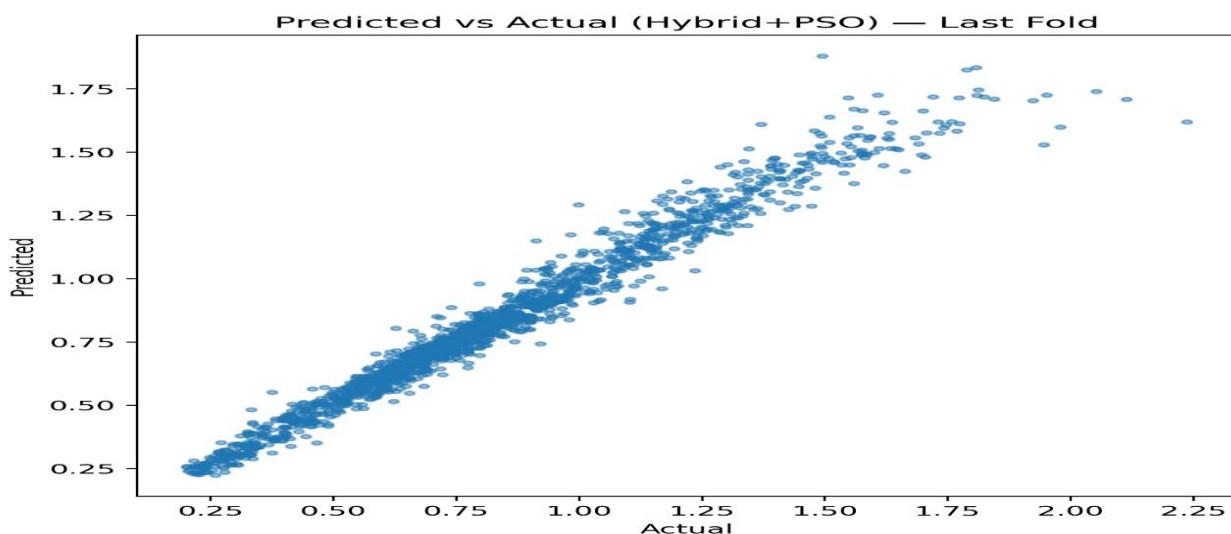


Fig. 4. Scatter plot of the actual and predicted electricity consumption of the proposed Hybrid LSTM – XGBoost + PSO model.

The data points are concentrated along the diagonal line, which suggests that the actual and predicted electricity consumption values are in good agreement. This finding is aligned with the high coefficient of determination (R^2) obtained in Table 6, and further establishes the predictive power of the proposed framework. The other benefit of the proposed method is the usage of Walk-Forward Validation that maintains the sequence of the data points and reduces the possibility of information leakage as compared to the traditional train–test splitting scheme [15, 16]. Therefore, the forecasting results presented here are more realistic than the results obtained under idealized model deployment. Overall, the experimental results show that the proposed Hybrid LSTM–XGBoost + PSO framework not only produces accurate short-term electricity demand forecast results but also outperforms the baseline XGBoost model and the non-optimized Hybrid LSTM–XGBoost model in all the various scenarios evaluated.

4.4 Discussion Summary

The experimental results highlight the success of integrating the deep learning, ensemble learning and intelligent optimization in a single framework for short-term electricity demand forecasting. The findings indicated that the Hybrid LSTM–XGBoost model achieved significantly better performance than the baseline XGBoost model, underscoring the importance of integrating the ability of deep temporal feature extraction to capture intricate electricity consumption patterns and temporal dependencies. The Particle Swarm Optimization (PSO) integration added another boost to the forecasting accuracy by optimizing the hyperparameters of the XGBoost component for a better performance. While the gains from PSO optimization were not as substantial as those of the non-optimized hybrid model, the PSO consistently produced better performance with all of the evaluation metrics. The result indicates that intelligent optimization can help improve forecasting accuracy and minimize expert hyperparameter tuning. The results that are presented across all the walk forward validation folds further validate the robustness and generalization capability of the proposed framework in different temporal evaluation scenarios. This is especially crucial for smart grid applications in real life, where the electricity consumption continually changes with time. The other significant contribution of this study is the use of Walk-Forward Validation as a leakage-free evaluation strategy. This approach retains the sequential nature of the observations throughout the forecasting process, thus reducing the risk of leakage of information, and is a more realistic measure of model performance than conventional train–test splitting approaches [15, 16]. In general, the results demonstrate that the proposed Hybrid LSTM–XGBoost + PSO framework is able to integrate deep temporal learning, ensemble regression, intelligent hyper-parameter optimization, and time-aware evaluation in a single forecasting architecture. In addition, the proposed approach can be further validated in practical applications with large-scale real world smart meter data for accurate short-term electricity demand forecasting in smart grid networks.

5. Conclusion

In this study, the authors proposed an optimized short-term electricity demand forecasting model based on Hybrid LSTM–XGBoost model with the help of large-scale real-world smart meter data from the Pecan Street project. The suggested framework combined the deep temporal feature extraction method, the ensemble regression method, the PSO hyperparameter optimization and the leakage-free Walk-Forward Validation method, into a single forecasting architecture. The experimental results showed that the Hybrid LSTM–XGBoost framework significantly outperformed the baseline XGBoost model. Moreover, the use of Particle Swarm Optimization (PSO) yielded extra and uniform enhancement in forecasting accuracy by optimal hyperparameter selections. The performance improvement with PSO optimization was not very high, but the results showed that the intelligent optimization can improve the model performance and reduce the manual model parameter adjustment. It was observed that the proposed architecture had low forecasting error, high coefficient of determination and achieved stable performance in all the folds of walk forward validation, which reflects the ability model in forecasting and temporal robustness and generalization performance. Furthermore, by incorporating Walk-Forward Validation, there was a realistic evaluation strategy that maintained the chronological structure in the data while reducing the amount of information leakage. Overall, the results validate the benefits of deep temporal learning, ensemble regression, intelligent optimization, and time-aware evaluation when integrated into a single forecasting system. Finally, the employment of massive real-world smart meter data further demonstrates the usefulness of the proposed approach for short-term forecasting in smart grid environments. The obtained forecasting accuracy confirms the potential of the proposed framework for the support of modern smart grid demand side management, load balancing, peak demand mitigation and operational planning [20]. Therefore, the proposed model can be used practically by the utility providers and intelligent energy management systems as a decision support tool. A multi-step forecasting horizon, attention-based neural networks, transformer architectures, and Reinforcement learning techniques in adaptive energy management applications can be explored in future work. Further research can also test the

proposed framework with varying scales of electricity consumption data, such as at the household, building, community and city level. Incorporating these external data sources, like weather forecasts, renewable energy production and social and economic data, can also enhance the accuracy of forecasts and the ability of models to adapt.

References

- [1] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, San Francisco, CA, USA, 2016, pp. 785–794.
- [2] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [3] C. Sun, Q. Lv, S. Zhu, W. Zheng, Y. Cao, and J. Wang, “Ultra-short-term power load forecasting based on double-layer XGBoost considering multi-feature effects,” *High Voltage Technology*, vol. 48, pp. 2885–2898, 2021.
- [4] Z. Chen, J. Liu, C. Li, X. Ji, D. Li, Y. Huang, F. Di, X. Gao, and L. Xu, “Ultra-short-term power load forecasting based on LSTM and XGBoost combined model,” *Power Grid Technology*, vol. 22, pp. 614–620, 2020.
- [5] A. Banks, J. Vincent, and C. Anyakoha, “A review of particle swarm optimization. Part I: Background and development,” *Natural Computing*, vol. 6, no. 4, pp. 467–484, 2007.
- [6] H. Tan, Q. Yang, J. Xing, K. Huang, S. Zhao, and H. Hu, “Photovoltaic power prediction based on XGBoost–LSTM combined model,” *Journal of Solar Energy*, vol. 43, pp. 75–81, 2022.
- [7] R. Masini, M. Medeiros, and E. Mendes, “Machine learning advances for time series forecasting,” *Journal of Economic Surveys*, vol. 35, no. 1, pp. 201–243, 2021.
- [8] M. Lafuente-Cacho et al., “State of the art for solar and wind energy forecasting methods for sustainable grid integration,” *Current Sustainable/Renewable Energy Reports*, vol. 12, no. 1, 2025.
- [9] J. Kim, H. Kim, D. Lee, and S. Yoon, “A comprehensive survey of deep learning for time series forecasting: Architectural diversity and open challenges,” *Artificial Intelligence Review*, vol. 58, no. 7, pp. 1–95, 2025.
- [10] I. Hussain et al., “Predicting monthly wind speeds using XGBoost: A case study for renewable energy optimization,” *Processes*, vol. 13, no. 6, p. 1763, 2025.
- [11] M. Rhafes, O. Moussaoui, and M. Raboaca, “Machine learning models in renewable energy forecasting: A systematic literature review,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 37, no. 3, pp. 1874–1891, 2025.
- [12] M. Baltruszewicz et al., “Social outcomes of energy use in the United Kingdom,” *Ecological Economics*, vol. 205, p. 107686, 2023.
- [13] D. Lorente et al., “Renewable energy, economic complexity and biodiversity risk,” *Environmental and Sustainability Indicators*, vol. 18, p. 100244, 2023.
- [14] A. Androniceanu and O. Sabie, “Overview of green energy as a strategic option for sustainable development,” *Energies*, vol. 15, p. 8573, 2022.
- [15] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 3rd ed. Melbourne, Australia: OTexts, 2021.
- [16] C. Bergmeir, R. J. Hyndman, and B. Koo, “On the validity of cross-validation for time series prediction,” *Computational Statistics & Data Analysis*, vol. 120, pp. 70–83, 2018.
- [17] Pecan Street Inc., “Dataport,” Available: <https://dataport.pecanstreet.org>. Accessed: 2026.
- [18] L. Semmelmann, S. Henni, and C. Weinhardt, “Load forecasting for energy communities: A novel LSTM-XGBoost hybrid model based on smart meter data,” *Energy Informatics*, vol. 5, Suppl. 1, 2022.
- [19] E. Demir and S. Gunal, “Short-term electricity consumption forecasting with deep learning,” *The Journal of Supercomputing*, vol. 81, 2025.
- [20] T. A. Rajaperumal and C. C. Columbus, “Transforming the electrical grid: The role of AI in advancing smart, sustainable, and secure energy systems,” *Energy Informatics*, vol. 8, 2025.
- [21] M. Huotari, A. Malhi, and K. Främling, “Machine Learning Applications for Smart Building Energy Utilization: A Survey,” *Archives of Computational Methods in Engineering*, vol. 31, pp. 2537–2556, 2024.