

**A DUAL-SEQUENCE HYBRID STACKING MODEL FOR ENHANCED SHORT-TERM LOAD FORECASTING**

**Dr. Alsharif Mohamed<sup>1\*</sup>**

<sup>1\*</sup>Department of Computer Science, Haql University Collage, University of Tabuk, Tabuk-71491, Saudi Arabia, amohamed@ut.edu.sa

**Abstract:**

Accurate short-term electric load forecasting is essential for effective power system planning and operational reliability, particularly in the face of demand variability introduced by intermittent energy sources. This paper introduces a novel Dual-Sequence hybrid model for short-term load forecasting (STLF) that integrates LSTM/BiLSTM networks with Extreme Gradient Boosting (XGBoost) using stacking technique. Initially, LSTM, BiLSTM, and XGBoost models are developed independently to forecast electricity load. Subsequently, a combined structure is constructed by embedding LSTM and BiLSTM generated features into the XGBoost model, allowing the integration of temporal learning capabilities with robust feature selection and optimization. While LSTM networks are proficient in modeling nonlinear temporal dependencies, their performance can be hindered by increased computational complexity as model parameters scale. To address this, XGBoost is employed to reduce input dimensionality and mitigate overfitting risks. The proposed hybrid model was validated using real-world datasets from Malaysia. Experimental results demonstrate that the combined model significantly outperforms individual models, as well as single-hybrid models, reducing MAPE by 5.7% and 4.3%, respectively, achieving a mean absolute percentage error (MAPE) of 1.15%. This notable reduction in forecasting error highlights the potential of the Dual-Sequence hybrid approach for advancing short-term load forecasting applications.

**Keywords:** Short-term load forecasting; deep learning; Extreme Gradient Boosting; stacking approach; multivariate time series.

**1. Introduction**

Accurate Short-Term Load Forecasting (STLF) is essential for maintaining the balance between electricity supply and demand, enabling efficient resource management and reducing operational costs. In competitive energy markets, even a 1% increase in STLF error can lead to over 10 million United States Dollar (USD) in additional expenses annually (Peng et al., 2021). To improve forecast accuracy, researchers have explored various approaches, typically categorized into statistical, data-driven, and hybrid models. Hybrid methods are particularly promising, as they integrate the strengths of multiple techniques while mitigating individual limitations (J. Li et al., 2020)(Tayab et al., 2020). This paper proposes a dual-sequence hybrid Long Short-Term Memory and Bidirectional Long Short-Term Memory with Extreme Gradient Boosting (LSTM/BiLSTM-XGBoost) Model for next-day hourly load forecasting. While LSTM and BiLSTM effectively model sequential dependencies, and XGBoost excels at learning nonlinear relationships, each approach alone has limitations. This motivates a hybrid model that integrates LSTM and BiLSTM features into XGBoost, combining temporal learning and nonlinear predictive strength to achieve more accurate and robust load forecasts. The remainder of this section reviews the literature review of related work and also some problems and gaps are discussed.

### ***1.1. Literature Review***

#### ***1.1.1. Statistical models***

Statistical models use mathematical functions and historical data to uncover patterns and forecast future behavior. Common approaches include the AutoRegressive Integrated Moving Average (ARIMA), its extension Autoregressive Integrated Moving Average with eXogenous variables (ARIMAX), and Multiple Linear Regression (MLR) models (Moradzadeh et al., 2021). ARIMA has proven effective for peak load forecasting, often outperforming the simpler AutoRegressive Moving Average (ARMA) model (Moradzadeh et al., 2021)(Yang, 2019). The ARIMAX model, which incorporates external factors such as weekday, weekend, and holiday effects, has also shown improved accuracy (Shilpa & Sheshadri, 2020)(Xiao et al., 2022). To address seasonal variations and nonlinear dependencies, the Seasonal AutoRegressive Integrated Moving Average (SARIMA) model has been proposed as a more robust alternative (Maldonado et al., 2019). However, as the number of input variables increases, statistical models may face limitations due to longer computation times, higher processing demands, and limited generalization capabilities(Massaoudi et al., 2021). In addition, a time-varying state space model employing the Kalman filter has been used to estimate hourly load demand (Al-Hamadi & Soliman, 2004). More recently, a fuzzy linear regression approach was introduced, where model parameters were estimated using a mixed linear programming method based on three years of historical data (Song et al., 2005).

#### ***1.1.2. Artificial intelligence (AI) models***

Artificial intelligence (AI) models are increasingly favored in short-term load forecasting (STLF) due to their ability to capture complex, nonlinear relationships between load demand and influencing factors (Solyali, 2020). Among these, Artificial Neural Networks (ANNs) are widely used for their self-learning and error-tolerant nature. However, limitations in generalizability led to the development of support vector machines (SVMs) as an alternative (Ye et al., 2012). Comparative studies have demonstrated that XGBoost often outperforms traditional models such as backpropagation neural networks, especially under seasonal conditions(Engineering & 2020, 2020). Variants like a window-based XGBoost incorporating real-time pricing achieved a Mean Absolute Percentage Error (MAPE) as low as 0.35% (Zhao et al., 2022), and other versions have effectively modeled peak loads during extreme weather events (Deng et al., 2022). Recent advancements in deep learning have emphasized LSTM models, with applications in both single-step and multi-step intraday forecasting. These models have outperformed Generalized Regression Neural Networks (GRNN) and Extreme Learning Machines (ELM) in accuracy (Hossain et al., 2020). Bi-LSTM, in particular, has shown a lower MAPE of 0.22% and superior peak load prediction capabilities compared to stacked LSTM, although some studies report the opposite(Atef et al., 2020)(Mughees et al., 2021)(Ren et al., 2022). Although model performance can vary due to differences in datasets and implementation, LSTM-based architectures consistently demonstrate strong performance across a wide range of STLF scenarios.

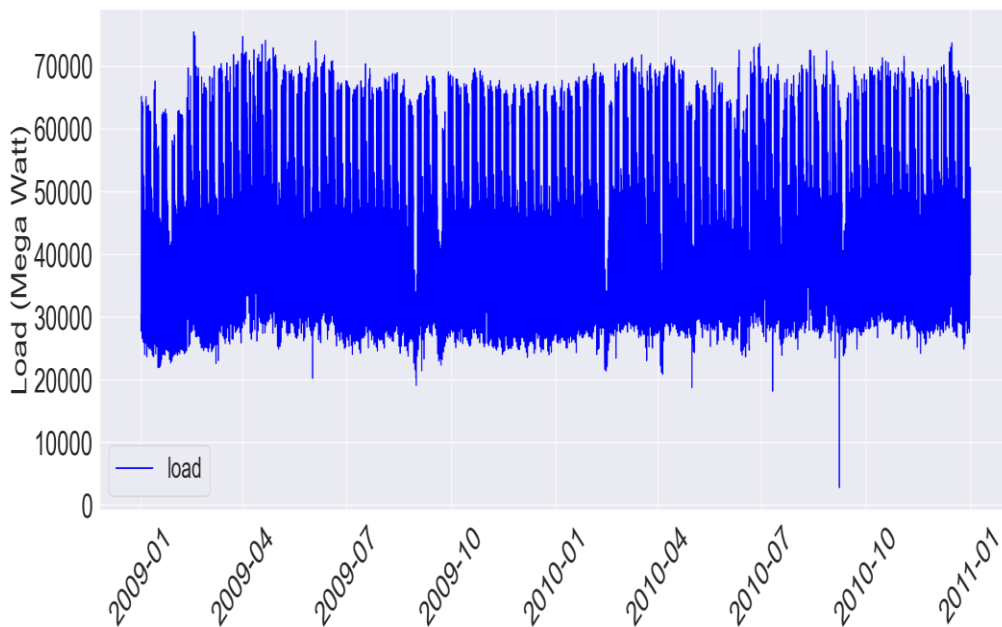
#### ***1.1.3. Hybrid models***

Hybrid models, which integrate multiple predictive algorithms, are increasingly employed in short-term load forecasting (STLF) to leverage the strengths of individual techniques(Wu et al., 2019)(Moradzadeh et al., 2020). By combining data processing methods with neural networks, these models aim to improve forecasting accuracy. For example, a hybrid model using a Grasshopper Optimization Algorithm (GOA) and a Support Vector Network (SVN) outperformed standard Support Vector Machines (SVMs) when accounting for variables like temperature and humidity, though it left room for incorporating additional factors(Barman et al., 2018). Another approach utilized Wavelet Neural Networks (WNNs) to decompose inputs, followed by ANN for prediction, achieving lower MAPE than standalone ANNs (El-Hendawi et al., 2020). LSTM-XGBoost hybrids

have also shown promise, with one model achieving a MAPE of 0.57 using a reciprocal error technique (C. Li et al., 2019). While stacking techniques may offer even greater integration potential, research in this area remains limited (Massaoudi et al., 2021). An advanced hybrid model employed stacking alongside an improved artificial fish swarm algorithm to integrate multiple Support Vector Regression (SVR) models using prior-day temperature data (Tan et al., 2020). Another ensemble combined Deep Neural Networks (DNNs) using Principal Component Regression (PCR) to construct a robust metamodel (Moon et al., 2020). These examples highlight the effectiveness of ensemble learning in boosting prediction performance. This review explores these advancements, focusing on hybrid LSTM/BiLSTM–XGBoost methodologies and their implications for improving STLF accuracy.

**1.1.4. Difficulties in Building the STLF Model**

Developing a high-performance STLF model requires careful consideration of several influencing factors, including historical load data and meteorological variables such as wind speed, temperature, humidity, and atmospheric pressure. Due to the complexity and nonlinearity inherent in STLF data characterized by intra-seasonal variations and dynamic patterns-accurate prediction using a single model remains challenging (Taylor et al., 2008)(Society & 2003, 2003) (see Fig. 1). Traditional linear models like ARIMA, SARIMA, and Seasonal AutoRegressive Fractionally Integrated Moving Average (SARFIMA) demand extensive data preprocessing and rely on autocorrelation and partial autocorrelation analyses to identify suitable hyperparameters. However, even with such tools, optimal parameter selection is often uncertain and difficult to generalize (Gooijer et al., 2006) (see Fig. 2). In contrast, hybrid deep learning approaches, particularly LSTM/BiLSTM–XGBoost models, offer improved accuracy through automated feature extraction and the ability to model complex temporal dynamics.



**Fig. 1:** Hourly load data of year 2009 and 2010

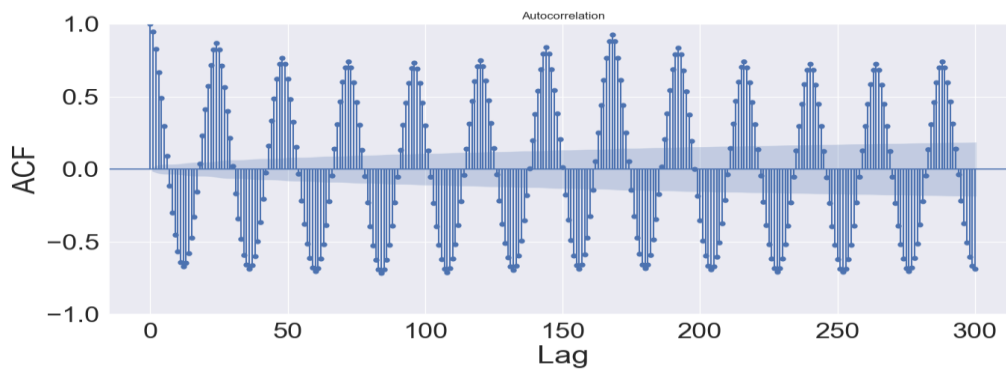


Fig. 2: Auto-correlation of the load data of Malaysia

## 2. Materials and Methods

This section outlines the methodology employed and the structure of the proposed model.

### 2.1. Long Short-Term Memory (LSTM)

LSTM was suggested in 1997 by Hochreiter and Schmidhuber as a standard type of Recurrent Neural Networks (RNN) for learning and processing long-term information, time-series data, feature extraction, and pattern recognition(Hochreiter et al., 1997)(Choi et al., 2020). The problems of gradient vanishing and the explosion of long-term dependencies have been improved by replacing the LSTM with basic hidden neurons in the RNN structure. As shown in Fig. 3, the LSTM includes the forget gate, input gate, update gate, and output gate in the principle structure(Wang et al., 2019). The LSTM network implements temporary memory through switch gates to prevent gradient vanishing. The main computation formula of the LSTM is as follows (Wang et al., 2019)(Guo et al., 2019):

$$f_t = \sigma(W_f(h_{t-1}, x_t) + b_f)$$

(1)

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

$$\tilde{c}_t = \tanh(W_c((h_{t-1}, x_t) + b_c)$$

(3)

$$c_t = f_t c_{t-1} + i_t \tilde{c}_t$$

(4)

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

(5)

$$h_t = o_t \tanh(c_t)$$

(6)

where  $f_t$ ,  $i_t$ ,  $\tilde{c}_t$ , and  $o_t$  determine the output values of the forget, input, update, and output gates, respectively,  $W_f$ ,  $W_i$ ,  $W_c$ , and  $W_o$  demonstrate the weight metrics,  $b_f$ ,  $b_i$ ,  $b_c$ , and  $b_o$  illustrate the bias vectors,  $c_t$  and  $\sigma$  show the memory cell and sigmoid activation function, respectively. In addition, the inputs of the four gates contain the LSTM target value  $h_{t-1}$  at a past time step  $t - 1$ .

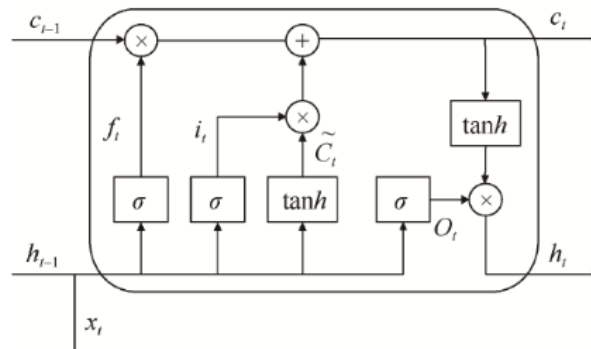


Fig. 3: The structure of LSTM.

**2.2. Bidirectional Long Short-Term Memory (BiLSTM)**

The Bidirectional Long Short-Term Memory (BiLSTM) network extends the conventional LSTM by integrating both forward and backward information flows. As illustrated in Fig. 4, BiLSTM comprises a forward LSTM layer and a backward LSTM layer. In the forward pass, the forward LSTM sequentially processes the input from the start to the end, effectively capturing past dependencies. Simultaneously, the backward LSTM traverses the sequence in reverse, capturing future dependencies. The outputs from both directions are then combined to form the final output sequence. This bidirectional mechanism enables the network to leverage contextual information from both past and future time steps, reducing dependency on any single point in the sequence and enhancing overall model robustness(Atef et al., 2020)(Mughees et al., 2021).

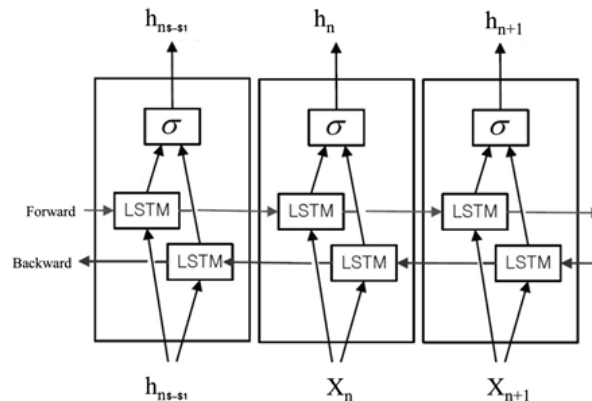


Fig. 4: The structure of BiLSTM

**2.3. Extreme Gradient Boosting (XGBoost)**

XGBoost, which stands for Extreme Gradient Boosting(Chen & Guestrin, 2016), is a highly efficient and powerful machine learning algorithm designed to enhance the performance of gradient boosting models. It excels in accurately capturing dynamic trends in short-term power loads, thus making it particularly effective for STLF problems(Engineering & 2020, 2020). It primarily utilizes gradient boosting decision trees to enhance speed and performance. The technique works by iteratively adding new weak learners to correct the residuals of previous ones. This iterative process addresses the limitations of individual learners, which might not achieve satisfactory results on their own. By combining these weak learners, XGBoost produces a final prediction that is more accurate than the predictions of the individual models. In essence, the final prediction generated by XGBoost is the sum of the scores of each weak learner(Suo et al., 2019)(Nobre et al., 2019) . The computational procedure of the XGBoost is as follows:

- 1: The sample weights and model parameters are first initialized by assigning the same weights to all samples in the training set.

2: We need to use Equation (1) to calculate the classification error rate at each iteration:

$$err_c = \frac{\sum w_i I(y_i \neq G_c x_i)}{\sum w_i}$$

(7)

where  $w_i$  denotes the weight of the  $i^{th}$  sample, and  $G_c x_i$  is the prediction of the  $c^{th}$  classifier.

3: In the following iteration, the sample weights are updated to place more emphasis on misclassified instances, as shown in Equation (2):

$$w_{i+1} = w_i e^{a_c I(y_i \neq G_c x_i)}$$

(8)

where  $a_c = \log((1 - err_c)/err_c)$

4: To optimize model performance, XGBoost minimizes a regularized objective function that balances prediction accuracy with model complexity. The objective function is expressed in Equation (3):

$$O_{bj} = \sum_{i=1}^n L(y_i, \hat{y}_i) + \sum_{k=1}^k \Omega(f_k)$$

(9)

Here,  $L(y_i, \hat{y}_i)$  denotes the loss between the true value  $y_i$  and the predicted value  $\hat{y}_i$ , and  $\Omega(f_k)$  is the predicted value, and  $\Omega(f_k)$  is a regularization term applied to the  $k^{th}$  tree to control its complexity.

5: The complexity term  $\Omega(f)$  is calculated using Equation (4):

$$\Omega(f) = \gamma N + \frac{1}{2} \lambda \sum_{i=1}^N w_i^2$$

(10)

where  $N$  is the number of leaf nodes, and  $\gamma$  is a penalty term for each additional leaf node, and  $\lambda$  is the L2 regularization coefficient. This regularization strategy ensures that the model remains generalizable and less prone to overfitting by penalizing overly complex tree structures.

**Table 1:** List of best parameters of XGBoost

Class	Parameter	Value
General Parameters	Booster	Gbtree
Booster Parameters	learning rate	0.05
Booster Parameters	max_depth	5
Booster Parameters	min_child_weight	1
Booster Parameters	Subsample	0.7
Booster Parameters	colsample_bytree	0.7
Booster Parameters	n-estimators	1000

#### 2.4. Stacking Technique

The stacking technique is a form of ensemble learning that combines multiple predictive models to enhance overall performance. Its main objective is to capitalize on the strengths of individual models while mitigating their respective weaknesses. Typically, stacking consists of two layers: the first layer, comprising several base models, and the second layer, a single meta-model. The predictions generated by the base models serve as new input features for the meta-model, whose output constitutes the final prediction. Employing multiple base models in the first layer is advantageous for feature learning, as it allows the ensemble to capture diverse patterns from different modeling perspectives. For the meta-model layer, XGBoost has proven particularly effective due to its capacity to handle complex nonlinear relationships and optimize predictive performance (Moon et al., 2020).

#### 2.5. The Structure of the Proposed Model

The proposed model is a stacked ensemble approach for short-term load forecasting, combining LSTM and BiLSTM predictions as input features for an XGBoost meta-model. In this architecture, LSTM and BiLSTM serve as base models in the first computational layer, each capturing distinct

temporal patterns from the data. Their predictions are then integrated by the XGBoost meta-model, forming the second layer, which produces the final load forecast. By leveraging the complementary strengths of the base models, the stacking framework provides a more robust and accurate prediction, capturing diverse patterns and nonlinear relationships in the load data Fig. 5.

The development of the proposed model follows a systematic process:

1. Data normalization: The training dataset is transformed using min–max scaling to standardize feature ranges, reducing the impact of heterogeneous data and improving model stability.
2. Base model prediction: LSTM and BiLSTM are applied independently to generate initial predictions, which are then combined to form a comprehensive feature set for the meta-model.
3. Meta-model prediction: The XGBoost meta-model leverages the processed base model outputs to produce the final forecast.
4. Evaluation: The predictive performance is assessed on the testing dataset using MAPE and RMSE, providing quantitative measures of accuracy and robustness.

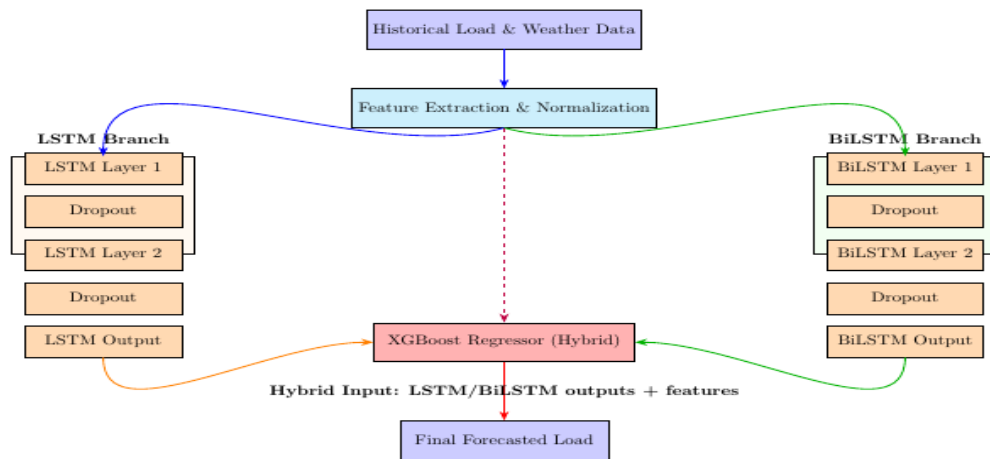


Fig. 5: Structure of the proposed hybrid LSTM/BiLSTM-XGBoost model

### 3. Results

The experimental analysis employed the hourly-load-data dataset, made available on Kaggle. This dataset, originally compiled by the power supply company of Johor, Malaysia, encompasses hourly electricity load records from the years 2009 and 2010 (see Fig. 1). Temperature time series was coupled with hourly load data, to improve the accuracy of the model. The temperature time series were plotted in Fig. 6. The data is split into train and test sets as detailed in Table 2.

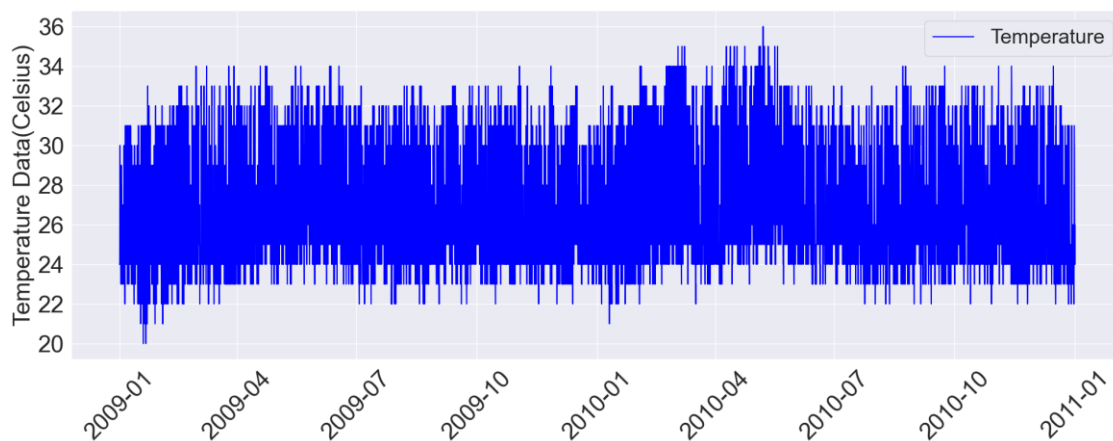


Fig. 6: Temperature data of year 2009-2010

**Table 2:** Train and test sets for year 2009 to 2010.

Type of Data	Start and end period	Number of items
Train set	Start from 2009-01-01 01:00 a.m.	17422
	To 2010-12-27 23:00 a.m.	
Test set	Start from 2010-12-28 00:00 a.m.	72
	To 2010-12-30 23:00 a.m.	

**3.1. Evaluation Metrics**

The Absolute Percentage Error (APE), the Mean Absolute Percentage Error (MAPE), and the Root Mean Squared Error (RMSE), as mentioned in Eq.11-Eq.13, represent the most important criteria utilized when evaluating the results of the models in STLF. In this study, the comparison relies on these metrics, which are described as follows:

$$APE_i = \left| \frac{(Actual_i - Predicted_i)}{Actual_i} \right| \times 100$$

(11)

Where  $Actual_i$  denote the observed value at instance  $i$ , and  $Predicted_i$  represent the corresponding forecast generated by the model.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{(Actual_i - Predicted_i)}{Actual_i} \right| \times 100$$

(12)

Where  $Actual_i$  denote the observed value at instance  $i$ , and  $Predicted_i$  represent the corresponding forecast generated by the model. A larger value signifies a greater deviation between the predicted and actual loads, while minimizing this value is essential to attain optimal prediction accuracy.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Actual_i - Predicted_i)^2}$$

(13)

where  $Actual_i$  is the observed value,  $Predicted_i$  is the model’s estimated value for observation  $i$ , and  $n$  is the total number of observations. RMSE measures the standard deviation of prediction errors, indicating how closely the predicted values match the actual values

**3.2. System Configuration**

To ascertain the efficacy and authenticity of the proposed methodology, a Keras version: 3.8.0 platform, accompanied by TensorFlow version: 2.18.0, is employed on a personal computer featuring an Intel Core i5, 2.3GHz central processing unit, 8GB of RAM, and a 500GB solid-state drive, operating under a Windows 10 environment.

**3.3. Experimental Evaluations**

The method is implemented in Python. Keras is used to implement LSTM, and Python XGBoost is used to implement XGBoost. The sampling period of the electrical load data was set to 24 hours, and 24 sets of data were generated each day. The training data of the LSTM, BiLSTM, and XGBoost models is from 01/012009 to 27/12/2010. The test data is from 28/12/2010 to 30/12/2010, and the power load from 29/12/2010 to 30/12/2010 is forecasted. The forecasted and actual comparison among the six methods is shown in Fig. 7. The forecast residuals are shown in Fig. 8. The evaluation of average MAPE, and RMSE of the all methods are shown in Table 3.

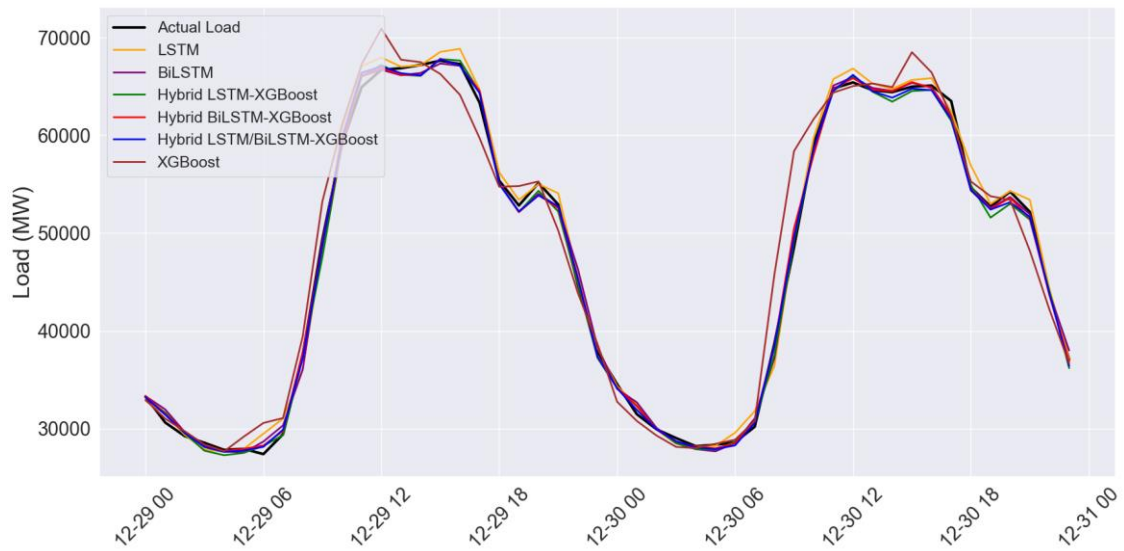


Fig. 7: Forecasted and actual time series comparison among all methods for test dataset of year 2009-2010.

Table 3: Evaluation of average MAPE and average RMSE for the proposed method and some other models in STLF.

Model	MAPE (%)	RMSE
LSTM	1.81	0.0140
BiLSTM	1.55	0.0105
XGBoost	3.81	0.0342
Hybrid LSTM-XGBoost	1.28	0.0104
Hybrid BiLSTM-XGBoost	1.18	0.0089
Proposed LSTM/BiLSTM-XGBoost	1.15	0.0084

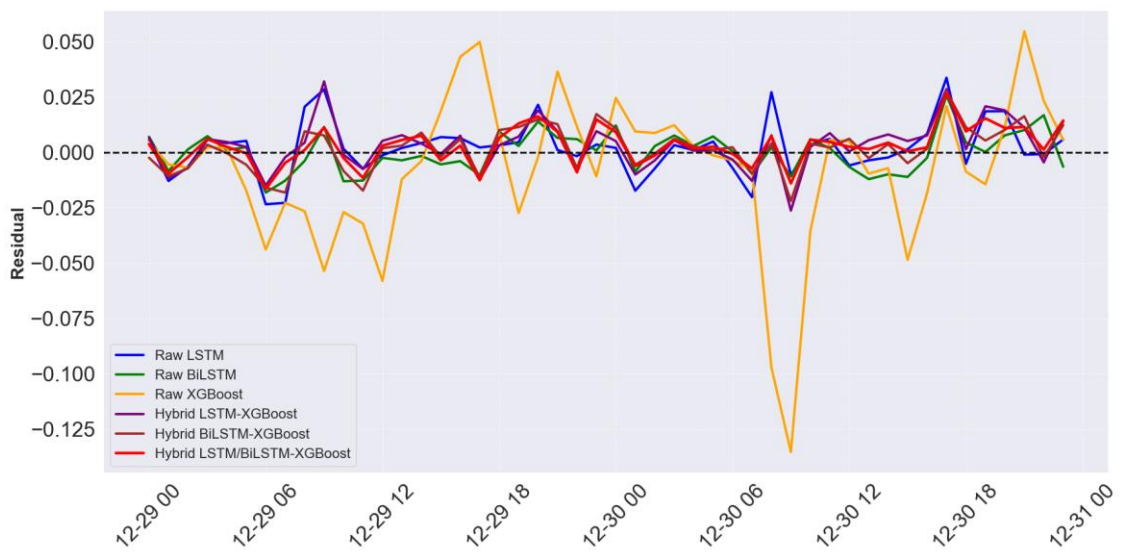


Fig. 8: Comparison of Residuals for all Models

#### **4. Discussion**

The experimental results demonstrate the comparative performance of individual and hybrid models for short-term load forecasting. Among the standalone models, BiLSTM outperformed LSTM and XGBoost, reflecting its ability to capture bidirectional temporal dependencies in the load data. XGBoost alone, while effective for nonlinear regression on tabular features, yielded the highest error, indicating that purely sequential patterns were not fully exploited. Integrating LSTM and BiLSTM outputs into XGBoost markedly improved accuracy, with hybrid models outperforming their individual counterparts. Notably, the dual-sequence hybrid LSTM/BiLSTM-XGBoost achieved the lowest errors, benefiting from complementary insights of unidirectional and bidirectional sequences and the nonlinear learning capability of XGBoost. These findings underscore the effectiveness of hybrid stacking strategies, demonstrating that combining deep sequential networks with ensemble regressors produces more accurate and robust load forecasts suitable for operational power system applications.

#### **5. Conclusion**

The proposed LSTM/BiLSTM-XGBoost hybrid model demonstrates superior performance for short-term load forecasting compared to both standalone and simpler hybrid models. By integrating features from LSTM and BiLSTM into XGBoost, the model effectively combines sequential learning with gradient-boosted regression. The dual-hybrid approach leverages complementary insights from unidirectional and bidirectional sequences, producing richer feature representations that enhance forecast accuracy and robustness. These results highlight that combining deep recurrent networks with ensemble-based regressors is a powerful strategy for reliable short-term load forecasting, offering practical value for improved operational decision-making and grid stability.

#### **Acknowledgement:**

The author received no financial support for this research, authorship, or publication, and no technical assistance was provided.

#### **Funding Information**

The author received no financial support for the research, authorship, or publication of this article.

#### **Author's Contributions**

The author is responsible for all aspects of this work, including conception, methodology, data collection, analysis, and manuscript preparation.

#### **Conflict of Interest**

The author declares no competing interests, financial or non-financial.

#### **Ethics**

This manuscript faithfully reflects the author's own research and analysis. The findings are situated within the existing scholarly record, and all sources and related work are cited appropriately.

#### **References**

1. Al-Hamadi, H. M., & Soliman, S. A. (2004). Short-term electric load forecasting based on Kalman filtering algorithm with moving window weather and load model. *Electric Power Systems Research*, 68(1), 47–59. [https://doi.org/10.1016/S0378-7796\(03\)00150-0](https://doi.org/10.1016/S0378-7796(03)00150-0)
2. Atef, S., Research, A. E.-E. P. S., & 2020, U. (2020). Assessment of stacked unidirectional and bidirectional long short-term memory networks for electricity load forecasting. *ElsevierS Atef, AB EltawilElectric Power Systems Research*, 2020•Elsevier. <https://www.sciencedirect.com/science/article/pii/S0378779620302923>
3. Barman, M., Choudhury, N., Energy, S. S.-, & 2018, U. (2018). A regional hybrid GOA-SVM

- model based on similar day approach for short-term load forecasting in Assam, India. *ElsevierM Barman, NBD Choudhury, S SutradharEnergy, 2018•Elsevier, 145, 710–720.* <https://www.sciencedirect.com/science/article/pii/S0360544217321990>
4. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-Aug, 785–794.* <https://doi.org/10.1145/2939672.2939785>
  5. Choi, E., Cho, S., Sustainability, D. K.-, & 2020, U. (2020). Power demand forecasting using long short-term memory (LSTM) deep-learning model for monitoring energy sustainability. *Mdpi.ComE Choi, S Cho, DK KimSustainability, 2020•mdpi.Com, 12(3), 1109.* <https://www.mdpi.com/2071-1050/12/3/1109>
  6. Deng, X., Ye, A., Zhong, J., Xu, D., Yang, W., Reports, Z. S.-E., & 2022, U. (2022). Bagging–XGBoost algorithm based extreme weather identification and short-term load forecasting model. *ElsevierX Deng, A Ye, J Zhong, D Xu, W Yang, Z Song, Z Zhang, J Guo, T Wang, Y Tian, H PanEnergy Reports, 2022•Elsevier.* <https://www.sciencedirect.com/science/article/pii/S2352484722012124>
  7. El-Hendawi, M., Research, Z. W.-E. P. S., & 2020, U. (2020). An ensemble method of full wavelet packet transform and neural network for short term electrical load forecasting. *ElsevierM El-Hendawi, Z WangElectric Power Systems Research, 2020•Elsevier, 182, 106265.* <https://www.sciencedirect.com/science/article/pii/S0378779620300729>
  8. Engineering, J. X.-2020 7th I. F. on E., & 2020, U. (2020). Research on power load forecasting based on machine learning. *Ieeexplore.Ieee.OrgJ Xu2020 7th International Forum on Electrical Engineering and, 2020•ieeexplore.Ieee.Org.* <https://ieeexplore.ieee.org/abstract/document/9356855/>
  9. Gooijer, J. De, Forecasting, R. H.-I. journal of, & 2006, U. (2006). 25 years of time series forecasting. *Elsevier, 22(3), 443–473.* <https://www.sciencedirect.com/science/article/pii/S0169207006000021>
  10. Guo, J., Xie, Z., Qin, Y., Jia, L., Access, Y. W.-I., & 2019, U. (2019). Short-term abnormal passenger flow prediction based on the fusion of SVR and LSTM. *Ieeexplore.Ieee.OrgJ Guo, Z Xie, Y Qin, L Jia, Y WangIeee Access, 2019•ieeexplore.Ieee.Org, 7, 42946–42955.* <https://ieeexplore.ieee.org/abstract/document/8675737/>
  11. Hochreiter, S., Computation, J. S.-N., & 1997, U. (1997). Long short-term memory. *Ieeexplore.Ieee.OrgS Hochreiter, J SchmidhuberNeural Computation, 1997•ieeexplore.Ieee.Org, 9(8), 1735–1780.* <https://ieeexplore.ieee.org/abstract/document/6795963/>
  12. Hossain, M., Energy, H. M.-2020 I. P. and, & 2020, U. (2020). Short-term load forecasting using an LSTM neural network. *Ieeexplore.Ieee.OrgMS Hossain, H Mahmood2020 IEEE Power and Energy Conference at Illinois (PECI), 2020•ieeexplore.Ieee.Org.* <https://ieeexplore.ieee.org/abstract/document/9064654/>
  13. Li, C., Chen, Z., Liu, J., Li, D., Gao, X., Di, F., Li, L., & Ji, X. (2019). Power load forecasting based on the combined model of LSTM and XGBoost. *Dl.Acm.OrgC Li, Z Chen, J Liu, D Li, X Gao, F Di, L Li, X JiProceedings of the 2019 the International Conference on Pattern Recognition, 2019•dl.Acm.Org, 46–51.* <https://doi.org/10.1145/3357777.3357792>
  14. Li, J., Deng, D., Zhao, J., Cai, D., ... W. H.-I. transactions on, & 2020, U. (2020). A novel hybrid short-term load forecasting method of smart grid using MLR and LSTM neural network. *Ieeexplore.Ieee.OrgJ Li, D Deng, J Zhao, D Cai, W Hu, M Zhang, Q HuangIEEE Transactions on Industrial Informatics, 2020•ieeexplore.Ieee.Org, 17, 2443–2452.* <https://ieeexplore.ieee.org/abstract/document/9109705/>
  15. Maldonado, S., Gonzalez, A., Computing, S. C.-A. S., & 2019, U. (2019). Automatic time series analysis for electric load forecasting via support vector regression. *ElsevierS Maldonado, A Gonzalez, S CroneApplied Soft Computing, 2019•Elsevier, 83.* <https://doi.org/10.1016/j.asoc.2019.105616>

16. Massaoudi, M., Refaat, S., Chihi, I., Energy, M. T.-, & 2021, U. (2021). A novel stacked generalization ensemble-based hybrid LGBM-XGB-MLP model for Short-Term Load Forecasting. *ElsevierM Massaoudi, SS Refaat, I Chihi, M Trabelsi, FS Oueslati, H Abu-RubEnergy, 2021•Elsevier*. <https://www.sciencedirect.com/science/article/pii/S0360544220319812>
17. Moon, J., Jung, S., Rew, J., Rho, S., Buildings, E. H.-E. and, & 2020, U. (2020). Combination of short-term load forecasting models based on a stacking ensemble approach. *ElsevierJ Moon, S Jung, J Rew, S Rho, E HwangEnergy and Buildings, 2020•Elsevier, 216*, 109921. <https://www.sciencedirect.com/science/article/pii/S0378778819327938>
18. Moradzadeh, A., Mansour-Saatloo, A., Nazari-Heris, M., Mohammadi-Ivatloo, B., & Asadi, S. (2021). Introduction and Literature Review of the Application of Machine Learning/Deep Learning to Load Forecasting in Power System. *Power Systems*, 119–135. [https://doi.org/10.1007/978-3-030-77696-1\\_6](https://doi.org/10.1007/978-3-030-77696-1_6)
19. Moradzadeh, A., Zakeri, S., Shoaran, M., Mohammadi-Ivatloo, B., & Mohammadi, F. (2020). Short-term load forecasting of microgrid via hybrid support vector regression and long short-term memory algorithms. *Mdpi.ComA Moradzadeh, S Zakeri, M Shoaran, B Mohammadi-Ivatloo, F MohammadiSustainability, 2020•mdpi.Com, 12*, 7076. <https://doi.org/10.3390/su12177076>
20. Mughees, N., Mohsin, S., ... A. M.-E. S. with, & 2021, U. (2021). Deep sequence to sequence Bi-LSTM neural networks for day-ahead peak load forecasting. *ElsevierN Mughees, SA Mohsin, A Mughees, A MugheesExpert Systems with Applications, 2021•Elsevier*. <https://www.sciencedirect.com/science/article/pii/S0957417421002852>
21. Nobre, J., Applications, R. N.-E. S. with, & 2019, undefined. (2019). Combining principal component analysis, discrete wavelet transform and XGBoost to trade in the financial markets. *ElsevierJ Nobre, RF NevesExpert Systems with Applications, 2019•Elsevier, 125*, 181–194. <https://doi.org/10.1016/j.eswa.2019.01.083>
22. Peng, L., Lv, S., Wang, L., Artificial, Z. W.-E. A. of, & 2021, U. (2021). Effective electricity load forecasting using enhanced double-reservoir echo state network. *ElsevierL Peng, SX Lv, L Wang, ZY WangEngineering Applications of Artificial Intelligence, 2021•Elsevier, 99*, 104132. <https://www.sciencedirect.com/science/article/pii/S0952197620303699>
23. Ren, H., Li, Q., Wu, Q., Zhang, C., Dou, Z., Reports, J. C.-E., & 2022, U. (2022). Joint forecasting of multi-energy loads for a university based on copula theory and improved LSTM network. *ElsevierH Ren, Q Li, Q Wu, C Zhang, Z Dou, J ChenEnergy Reports, 2022•Elsevier, 8*, 605–612. <https://www.sciencedirect.com/science/article/pii/S2352484722010447>
24. Shilpa, N. G., & Sheshadri, G. S. (2020). Electrical Load Forecasting Using Time Series Analysis. *Proceedings of B-HTC 2020 - 1st IEEE Bangalore Humanitarian Technology Conference*. <https://doi.org/10.1109/B-HTC50970.2020.9297986>
25. Society, J. T.-J. of the O. R., & 2003, undefined. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *Taylor & FrancisJW TaylorJournal of the Operational Research Society, 2003•Taylor & Francis, 54(8)*, 799–805. <https://doi.org/10.1057/PALGRAVE.JORS.2601589>
26. Solyali, D. (2020). A comparative analysis of machine learning approaches for short-/long-term electricity load forecasting in Cyprus. *Mdpi.ComD SolyaliSustainability, 2020•mdpi.Com, 12(9)*, 3612. <https://doi.org/10.3390/su12093612>
27. Song, K., Baek, Y., ... D. H.-I. transactions on, & 2005, U. (2005). Short-term load forecasting for the holidays using fuzzy linear regression method. *Ieeexplore.Ieee.OrgKB Song, YS Baek, DH Hong, G JangIEEE Transactions on Power Systems, 2005•ieeexplore.Ieee.Org, 20(1)*, 96–101. <https://doi.org/10.1109/pes.2005.1489152>
28. Suo, G., Song, L., Dou, Y., International, Z. C.-2019 18th, & 2019, U. (2019). Multi-dimensional short-term load forecasting based on XGBoost and fireworks algorithm. *Ieeexplore.Ieee.OrgG*

- Suo, L Song, Y Dou, Z Cui* 2019 18th International Symposium on Distributed Computing and, 2019•*ieeexplore.Ieee.Org*, 245–248. <https://ieeexplore.ieee.org/abstract/document/8921302/>
29. Tan, Z., Zhang, J., He, Y., Zhang, Y., Xiong, G., Access, Y. L.-I., & 2020, U. (2020). Short-term load forecasting based on integration of SVR and stacking. *Ieeexplore.Ieee.Org* Z Tan, J Zhang, Y He, Y Zhang, G Xiong, Y Liueee Access, 2020•*ieeexplore.Ieee.Org*. <https://ieeexplore.ieee.org/abstract/document/9274433/>
  30. Tayab, U., Zia, A., Yang, F., Lu, J., Energy, M. K.-, & 2020, U. (2020). Short-term load forecasting for microgrid energy management system using hybrid HHO-FNN model with best-basis stationary wavelet packet transform. *Elsevier* UB Tayab, A Zia, F Yang, J Lu, M Kashif Energy, 2020•*Elsevier*, 203, 117857. <https://www.sciencedirect.com/science/article/pii/S0360544220309646>
  31. Taylor, J., Systems, P. M.-I. T. on P., & 2007, undefined. (2008). Short-term load forecasting methods: An evaluation based on european data. *Ieeexplore.Ieee.Org* JW Taylor, PE McSharry *IEEE Transactions on Power Systems*, 2007•*ieeexplore.Ieee.Org*, 22, 2213–2219. <https://ieeexplore.ieee.org/abstract/document/4349130/>
  32. Wang, K., Qi, X., Energy, H. L.-, & 2019, U. (2019). Photovoltaic power forecasting based LSTM-Convolutional Network. *Elsevier* K Wang, X Qi, H Liu Energy, 2019•*Elsevier*, 189, 116225. <https://www.sciencedirect.com/science/article/pii/S0360544219319206>
  33. Wu, Z., Zhao, X., Ma, Y., Energy, X. Z.-A., & 2019, U. (2019). A hybrid model based on modified multi-objective cuckoo search algorithm for short-term load forecasting. *Elsevier* Z Wu, X Zhao, Y Ma, X Zhao *Applied Energy*, 2019•*Elsevier*, 237, 896–909. <https://www.sciencedirect.com/science/article/pii/S0306261919300467>
  34. Xiao, X., Mo, H., Zhang, Y., Energy, G. S.-, & 2022, U. (2022). Meta-ANN–A dynamic artificial neural network refined by meta-learning for Short-Term Load Forecasting. *Elsevier* X Xiao, H Mo, Y Zhang, G Shan energy, 2022•*Elsevier*, 246, 123418. <https://www.sciencedirect.com/science/article/pii/S0360544222003218>
  35. Yang, T. (2019). *Electric Load Forecasting Using Long Short-term Memory Algorithm*. <https://scholarscompass.vcu.edu/cgi/viewcontent.cgi?article=7126&context=etd>
  36. Ye, N., Liu, Y., 2012, Y. W.-W. A. C., & 2012, U. (2012). Short-term power load forecasting based on SVM. *Ieeexplore.Ieee.Org* N Ye, Y Liu, Y Wang *World Automation Congress 2012, 2012•ieeexplore.Ieee.Org*, 47–51. <https://ieeexplore.ieee.org/abstract/document/6321271/>
  37. Zhao, X., Li, Q., Xue, W., Zhao, Y., Zhao, H., Energies, S. G.-, & 2022, U. (2022). Research on ultra-short-term load forecasting based on real-time electricity price and window-based XGBoost model. *Mdpi.Com* X Zhao, Q Li, W Xue, Y Zhao, H Zhao, S Guo *Energies*, 2022•*mdpi.Com*. <https://www.mdpi.com/1996-1073/15/19/7367>