Energy & Buildings 288 (2023) 113022

Contents lists available at ScienceDirect

Energy & Buildings

journal homepage: www.elsevier.com/locate/enb

Short-term electric load forecasting using an EMD-BI-LSTM approach for smart grid energy management system

Nada Mounir *, Hamid Ouadi, Ismael Jrhilifa

ERERA, ENSAM Rabat, Mohammed V University in Rabat, Rabat, Morocco

ARTICLE INFO

Article history:
Received 8 January 2023
Revised 15 March 2023
Accepted 23 March 2023
Available online 28 March 2023

Keywords:
EMD
IMF
Energy optimization
Power forecasting
Deep learning
LSTM

ABSTRACT

Electricity is an essential resource for human production and survival. Accurately predicting electrical load consumption can help power supply companies make informed decisions, such as peak load shifting, to maintain a reliable power supply and reduce CO2 emissions. However, forecasting electricity consumption is challenging due to the nonlinear and nonstationary time series data that is correlated with climate change. To address this challenge, this paper proposes an electricity forecasting method based on empirical mode decomposition (EMD) and bidirectional LSTM. EMD is a solid and robust instrument for time–frequency analysis and signal preprocessing, which separates the time series into components at different resolutions. The proposed model predicts the future 24 h with a resolution of 15 min by creating many stationary component sequences from the original stochastic electricity usage time series data (IMFs). To predict each Intrinsic Mode Function, a hybrid model BI-LSTM is employed. The results of each component's forecast are then merged to give the overall forecast. Two comparative studies are conducted to justify the choice of the signal processing method and the prediction algorithm. The proposed model demonstrates a minimal MAPE of 0.28% and a better R^2 close to 1 of 0.84 compared to other papers.

1. Introduction

Cities all over the world are expanding. Fast urbanization puts high pressure on energy consumption. However, because it forms the basis for making important decisions regarding the management and control of power, Power demand forecasting is a crucial component in the optimum dispatching of power systems. Load forecasts are classified into three types: short-term load forecasts (2–3 days), medium-term load forecasts (4–9 days), and long-term load forecasts (more than 10–15 days) [1].

Creating an accurate electricity demand forecasting model is challenging because the demand series frequently exhibits a highly nonlinear characteristic in addition to seasonality. This is because demand is continuously impacted by turbulent factors such as extreme weather, holidays, and more [2]. To tackle this difficulty, academics have provided a variety of models over the previous few decades. Short-term electric load forecasting involves the application of advanced machine learning algorithms, particularly deep learning techniques. Deep learning models, such as Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Gated Recurrent Unit (GRU) networks, have shown superior performance compared to traditional statistical models (Table 1.).

Many studies have been conducted to investigate deep learning approaches for modeling electrical consumption in the residential sector. Most energy modeling studies in residential buildings use aggregated data to minimize the effect of single-user nonlinear dynamics on electricity usage. To address the single home power consumption forecasting problem, a hybrid model integrating long short-term memory and convolutional neural networks is being

Abbreviations: **EMD**, Empirical mode Decomposition; **LSTM**, long short-term memory; **GRU**, gated recurrent unit; **BI-LSTM**, bidirectional long short-term memory; **EMD-BI-GRU**, Empirical mode Decomposition bidirectional gated recurrent unit; **EMD-BI-LSTM**, Empirical mode Decomposition bidirectional long short-term memory; **MAPE**, Mean Absolute Percentage Error; **RMSE**, Root Mean Squared Error; **IMFs**, intrinsic mode functions; **RNN**, recurrent neural network; $\mathbf{R^2}$, R squared; **GAP**, Global Active Power; $\mathbf{GAP_t}$, Global Active Power of instant t; $\mathbf{GAP_{t+1}}$, Global Active Power of instant t + 15 mins; $\mathbf{GAP_{t-192}}$, Global Active Power of the actual day-24h; $\mathbf{GAP_{J+1}}$, Global Active Power of the future 24h; **T**, temperature; **H**, humidity; **W**, windspeed; **V**, visibility; **DP**, dewPoint; **P**, pressure; **H(t-96)**, humidity of the last 24 h; **T(t-96)**, temperature of the last 24 h.

* Corresponding author.
E-mail addresses: nadamounir98@gmail.com, nada_mounir@um5.ac.ma (N. Mounir).

https://doi.org/10.1016/j.enbuild.2023.113022
0378-7788/

Table 1
Prediction Results of Different Algorithms.

Model	R^2	RMSE	MAE	MAPE (%)
BI-LSTM	0.70	0.43	0.27	0.42
EMD-BI-LSTM	0.84	0.31	0.20	0.28
EMD-BI-GRU	0.79	0.35	0.21	0.30

developed by authors in [3]. The proposed approach is tested on five real-world datasets of 5 min power consumption data from different households. The results demonstrate that the hybrid CNN-LSTM architecture outperforms the individual CNN and LSTM models and other benchmark models regarding forecasting accuracy, with a Mean Absolute Percentage Error (MAPE) of 13.1%, 48.8%, 2.4%, 33.2%, and 14.5% for the five households. However, the paper also highlights some limitations of the proposed method. One of the main limitations is that the proposed approach only considers historical power consumption data, and does not consider other relevant external factors such as weather that may influence power consumption. The prediction results for household power consumption data were only shown for a 4-hour time frame, which is a limitation since it doesn't cover the full 24-hour period. In [4], authors propose a multi-step short-term load forecasting method for low latitude areas with highly random household load data, utilizing the LSTNet model. The method employs CNN to capture local connections in the data, LSTM to adapt to periodic load features, and an autoregressive model as a linear component. This study has several limitations. Firstly, the research is confined to low latitude regions with erratic household load data, which may not be generalizable to other areas or scenarios. Additionally, the absence of periodic information in the utilized dataset may restrict the applicability of the results to load data displaying clear cyclic patterns. Furthermore, the proposed methodology solely relies on past power consumption data, neglecting other extrinsic factors like weather that could potentially impact power consumption patterns. This could limit the accuracy and applicability of the model to real-world settings. The model's MAPE (Mean Absolute Percentage Error) is significantly high (Between 22,43 % and 32,02 %). Similarly, a research paper [5] offers a data and predictive analytics model for short-term electricity consumption prediction in multifamily residential buildings to aid policymakers in developing effective energy supply and demand policies. The paper presents a time-series analysis module to uncover insights and characteristics of energy consumption data, followed by an ensemble prediction approach that combines LSTM and KF models based on real datasets. Advanced feature selection techniques, improved data analysis methods, and more precise input features like weather and holiday features could further enhance the study's results. Our model demonstrates superior performance compared to LSTM and KF with a MAPE of 3.264 %. (Table 2)

Recent research has focused on improving the accuracy of short-term load forecasting by developing hybrid models that combine different deep learning techniques or integrate them with traditional statistical models. Hybrid models incorporating Wavelet Transform (WT) or Empirical Mode Decomposition (EMD) techniques have also been proposed to enhance load forecasting accuracy.

Table 2
The impact of features selection.

Model	**MAPE %**	**RMSE**	R^2
EMD-BI-LSTM with temperature and humidity	0.28	0.31	0.84
EMD-BI-LSTM with humidity	0.55	0.54	0.51
EMD-BI-LSTM with temperature	0.38	0.39	0.74

The [6] paper presents a short-term electric load forecasting model that combines the EEMD decomposition method with LSTM and ELM networks. The EEMD method is employed to smooth the power load sequence and derive high-frequency and low-frequency components, which are subsequently predicted by the LSTM and ELM networks and then fused to obtain the final prediction. In contrast, our model demonstrates superior performance compared to the proposed model, as reflected by the significantly larger MAPE of 2.242 % observed in the proposed model. Another study in the same field, referred to as [7], proposes a new EMD-GRU method for short-term electricity load forecasting with feature selection. This approach uses correlation analysis to select the decomposed time series components that are highly correlated with the original load series as features and then employs a GRU forecasting model to make the final prediction. While the proposed EMD-GRU method outperforms other comparison methods, it has limitations as it only selects a subset of decomposed time series components as features. The proposed EEMD-LSTM and EMD-GRU [6 7] approaches for short-term electricity load forecasting rely solely on past power consumption data and do not take into consideration external factors, such as weather conditions, which could potentially impact power consumption patterns.

Currently, many companies use machine learning and statistical models to predict electricity demand for the coming 24 h. The methods used by these companies may vary, but they typically involve analyzing historical data, weather forecasts, and other relevant factors that may affect electricity demand. The forecasting models are usually updated regularly based on new data and other relevant information. The accuracy of the forecasts is crucial for electricity companies to optimize their operations and avoid supply shortages or overcapacity.

Forecasting electrical energy consumption is still challenging due to its time series' intrinsic complexity and irregularity [2]. Because of the instability and uncertainty of power demand, the standard forecasting technique may fail to correctly reflect the dynamic change of load curves. To fulfil this practical need, address this issue, and enhance forecast accuracy, we are proposing an innovative model based on the EMD-BI-LSTM prediction methodology.

EMD-BI-LSTM, which stands for Empirical Mode Decomposition and Bidirectional Long Short-Term Memory, has been applied in various fields such as air pollution and Diagnosing faults in wind turbines. To the best of our knowledge, there are no other implementations of the EMD-BI-LSTM approach for short-term electrical load forecasting in this field besides ours.

To create a reliable energy consumption prediction model, we start by selecting features that show a strong correlation with energy consumption, using a heatmap of correlation coefficients. In Sceaux, we have access to a database that contains detailed data on meteorological factors and historical electric load data for an individual household, including lagged values of all features. We select the lagged values of the day-ahead based on an analysis of autocorrelation. We then use EMD, a flexible and adaptable data-driven method, to decompose the signal into its intrinsic mode functions (IMFs), which allows us to analyze the data at multiple time scales and identify trends, oscillations, and noise components that may be hidden in the raw data. We use a bidirectional LSTM, which considers past and future information, to predict the simpler components (IMFs). This model is more effective than unidirectional LSTMs, which process information in only one direction, and can capture complex patterns and dependencies that are present in the data. By incorporating EMD and BI-LSTM approach, our model can capture both short-term and long-term dependencies in the electrical load data, leading to more accurate predictions. We optimize the model's hyperparameters using grid search, and the sum of the submodules' outputs provides the final forecasting result. We compare two scenarios, one with EMD and one without, and evaluate the model's performance by predicting 24 h of load data at 96 time points, forecasting under a horizon of 24 h allows for accurate and timely decision making in energy management systems. Our experiments demonstrate that the proposed model outperforms existing models for short-term load prediction, which indicates that it is a good prediction model for energy consumption.

Our findings have important practical applications for energy management systems. For example, the approach can be used by utility companies to accurately forecast the short-term electrical load demand over a 24-hour horizon, which can aid in the efficient allocation of resources, reduce energy costs, and improve the reliability of the electrical grid. By accurately predicting the electrical load demand, utilities can also minimize the risk of blackouts and brownouts, which can have significant economic and social costs.

The article adheres to a predetermined structure, with Section 2 outlining the EMD-BI-LSTM approach, research methods, dataset description, feature selection, proposed model, and scenarios. On the other hand, Section 3 details the study's results, which encompass the EMD decomposition of Global Active Power and predictions for the following 24 h. This is then compared to the BI-LSTM model without decomposition, the EMD-BI-GRU model, and the model without temperature and humidity.

2. Methodology

2.1. The proposed forecasting approach

This study employs a deep learning approach called Bi-LSTM modelling for electricity forecasting. This section describes how the prediction model is constructed. The granularity of 15 min is investigated. The overall architecture for individual household electrical forecasting is depicted in Fig. 1. The most crucial feature of home power consumption is that some households lead typical lives, which can be seen in their consumption habits. Other households, however, inconsistently use energy. As a result, different families' consumption patterns differ. Choosing the features that strongly correspond with energy consumption is more straightforward when there is a link between the input features and energy consumption [8]. Using a selection of variables related to consumption reduces the computation time needed for parameter tweaking. The power usage data is a typical nonlinear and nonstationary data set with an intense time sequence and a combination of various trends, periodicity, and other factors [9]. It is essential to employ a specific technique to stabilize the nonstationary arrangement before using the appropriate deep-learning model for the following analysis. To predict electricity consumption, an EMD-BI-LSTM hybrid model is proposed in this paper, and its MAPE, MAE, RMSE, and R^2 are assessed.

2.2. Empirical mode decomposition

A time–frequency analysis technique is called EMD [10]. It offers an alternative to standard analytical methods like the wavelength transform. EMD assumes that the data contains several oscillations that coexist naturally. EMD's essential principle is to split a given signal into several stationary signals Fig. 2. The term "IMFs" refers to these static signals. The residue, also an IMF, contains elements from several frequency bands, ranging from the highest to the lowest. The frequencies gradually drop down the line, with IMF0 being the highest frequency. Unlike the Wavelength Transform, this ensures that the load data features may be exposed at various resolutions.

The IMFs bring out the regional features already there in the data, and by examining the IMFs, the characteristics of the original data may be discovered. Additionally, with EMD, combining the appropriate IMFS and residue allows the original data to be recreated without losing any information. A function must fulfil the following two requirements to be an IMF:

- Upper and lower envelopes, determined by local maximums and minimums, must have a mean value of zero.
- There can be no more than one difference between the number of local maximums and local minimums.

2.3. Bi-directional LSTM predictor

Bidirectional Recurrent neural networks are simply two separate RNNs together Fig. 3. The networks may access forward and backward information about the sequence thanks to this structure at each time step [11].

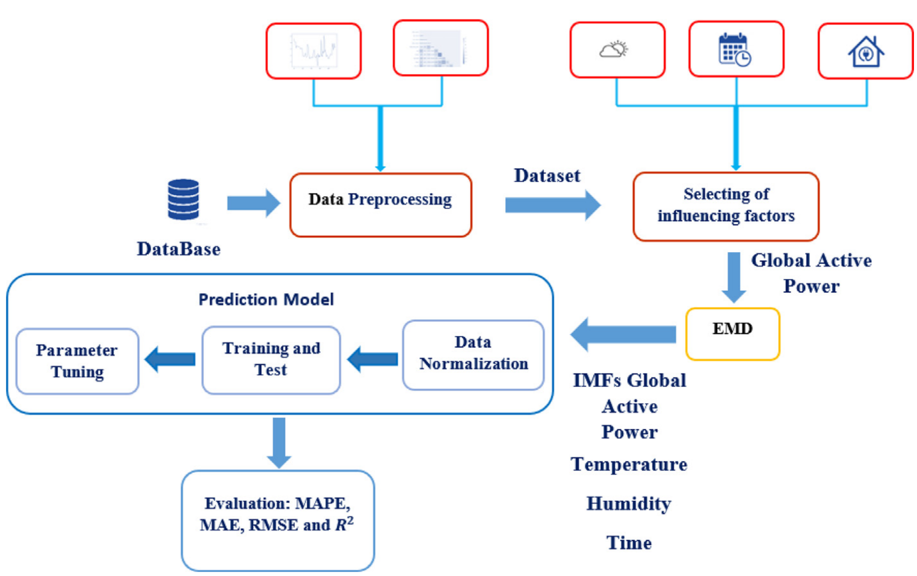

Fig. 1. Architecture of individual household electricity forecasting.

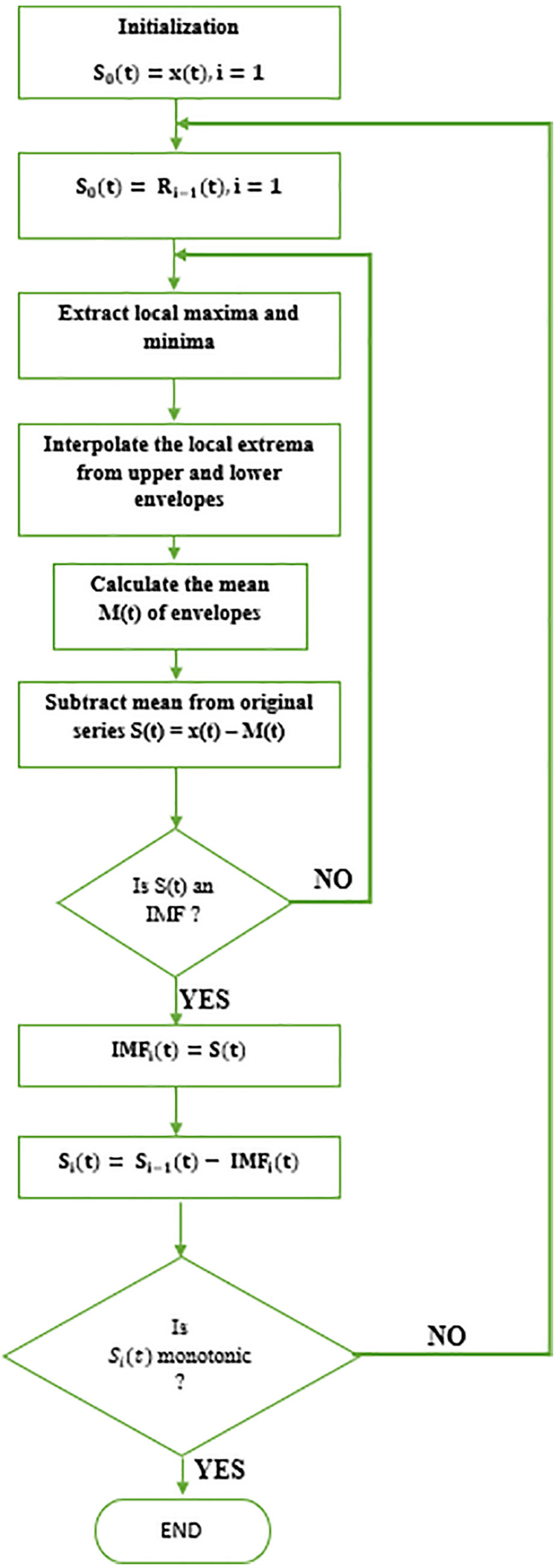

Fig. 2. The EMD algorithm.

When using bidirectional, inputs will be processed in two directions: from now to the future and from the future to the present. This method differs from the unidirectional in that information from the future is preserved in the LSTM that runs backwards, and by combining the two hidden states, it can preserve data from the present and the future at any given time.

2.4. Data preparation

2.4.1. Data description

We thoroughly reviewed the available information to construct a model that can adequately and reliably anticipate electrical usage. We examined several datasets in this sector before deciding on the Individual household electric power consumption dataset given by UCI [12] as the best fit for our research.

Meteorological data are introduced. After all, they are significant because they influence the electrical load. The database contains eight features: global active power (GAP), temperature (T), humidity (H), windspeed (W), visibility (V), dewpoint (DP), and time and pressure (P). The time is also introduced Fig. 4.

Fig. 5 displays a clear pattern of electricity consumption over a 24-hour period from 17 to 12-2006 16:45 to 18–12-2006 16:30. The x-axis denotes time, while the y-axis represents electricity consumption. The graph shows that consumption is higher in the evening and lower in the morning, indicating a discernible pattern.

2.4.2. Features selection

In Fig. 6, a heatmap of the correlation coefficient between Global Active Power, time, and meteorological features is displayed. The heatmap shows that the Global Active Power feature has a high correlation with Date and Time, Temperature, Humidity, and DewPoint. Specifically, the correlation coefficients of these features rank among the top four. The three meteorological features, temperature, humidity, and DewPoint, have a significant influence on Global Active Power values. Interestingly, the correlation coefficient between temperature and DewPoint is 0.82, which indicates a strong positive correlation between these two features. Therefore, the DewPoint feature was removed, and temperature was retained. In contrast, the correlation coefficients of Pressure and Visibility with Global Active Power were found to be less than 0.1, indicating weak correlations. Therefore, Pressure and Visibility have the weakest correlation with Global Active Power among all the meteorological features.

2.4.3. The selection of lag values

It is more efficient to use more feature lags that are selected based on their autocorrelation to predict short-term energy consumption. Specifically, the next Global Active Power (GAP) value, denoted as GAP_{t+1}, has a stronger correlation with features from the current day and the previous day. The correlation coefficient between GAP_{t+1} and GAP_t is 0.119, while the correlation coefficient between GAP_{t+1} and GAP_{t-1} is 0.112. In contrast, the correlation coefficient between GAP_{t+1} and GAP_{t-192} is 0.110, indicating a weaker correlation between the two features. By incorporating these lagged values into our selected feature, the block diagram of the proposed model is shown in the figure below Fig. 7.

2.5. Model and scenario description

Fig. 8 depicts the stages of the proposed approach.

2.5.1. EMD and its hyperparameters

The widely used EMD-based prediction approach involves creating separate prediction models for each decomposed global active power sub-series. However, this approach can increase the level of model complexity, and it is not efficient when dealing with large amounts of data. EMD does not rely on a predefined basis set. Instead, it adaptively generates a basis set based on the signal characteristics.

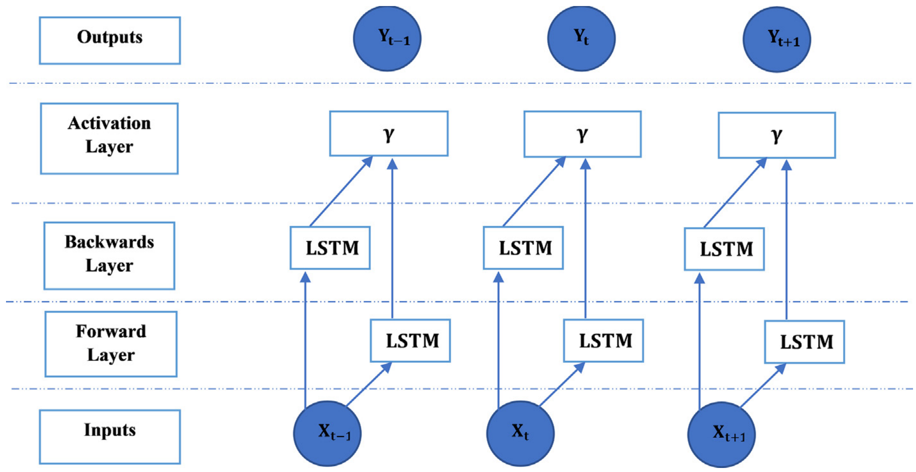

Fig. 3. The architecture of the BI-LSTM.

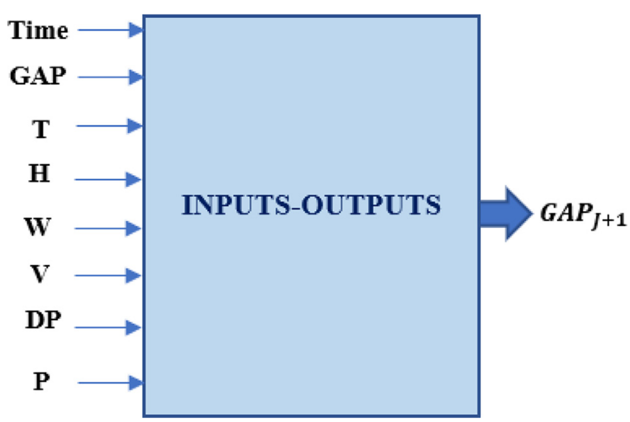

Fig. 4. Bloc diagram of features contains in the dataset and the output.

To address this issue, we propose to combine all the intrinsic mode functions (IMFs) and the residue together for prediction. By decomposing the original signal using EMD, the noise components can be removed, and the resulting IMFs and residue can capture the underlying trends and patterns in the data.

To create the prediction model, we built 19 databases, each of which combines the IMF/residue and its lag values with other features such as weather and time. They used the PyEMD [13] library in Python to perform the EMD decomposition. Before performing the prediction, they normalized each database because each feature belongs to a different range.

By using all the IMFs and residue together, we were able to capture the high autocorrelation between these components and the original signal. This approach reduces the complexity of the prediction model while improving prediction accuracy.

2.5.2. BI-LSTM and its hyperparameters

The main objective of the paper is to predict the 24-hour load value in the short term. The proposed method to achieve this is

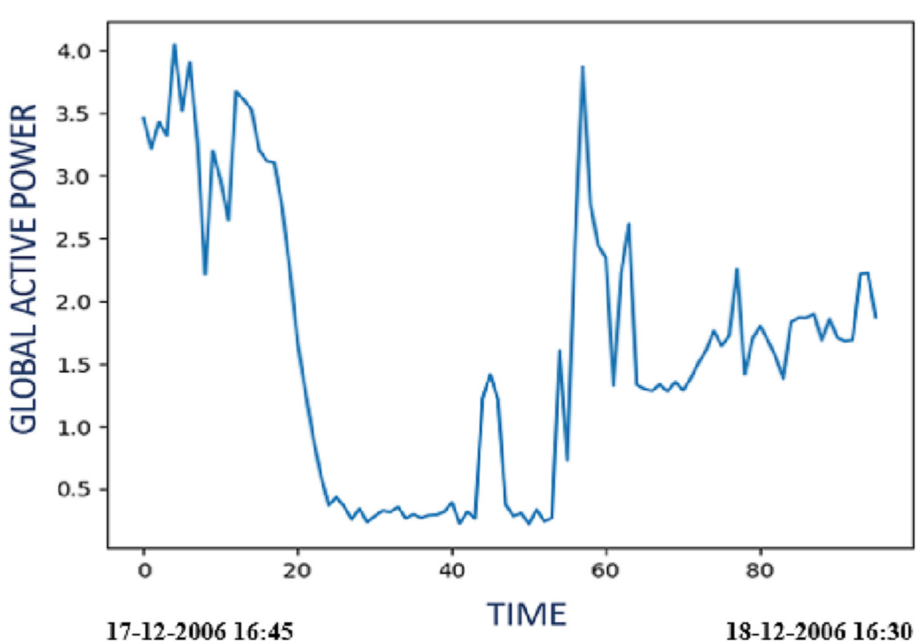

Fig. 5. Example of electricity consumption data for one day.

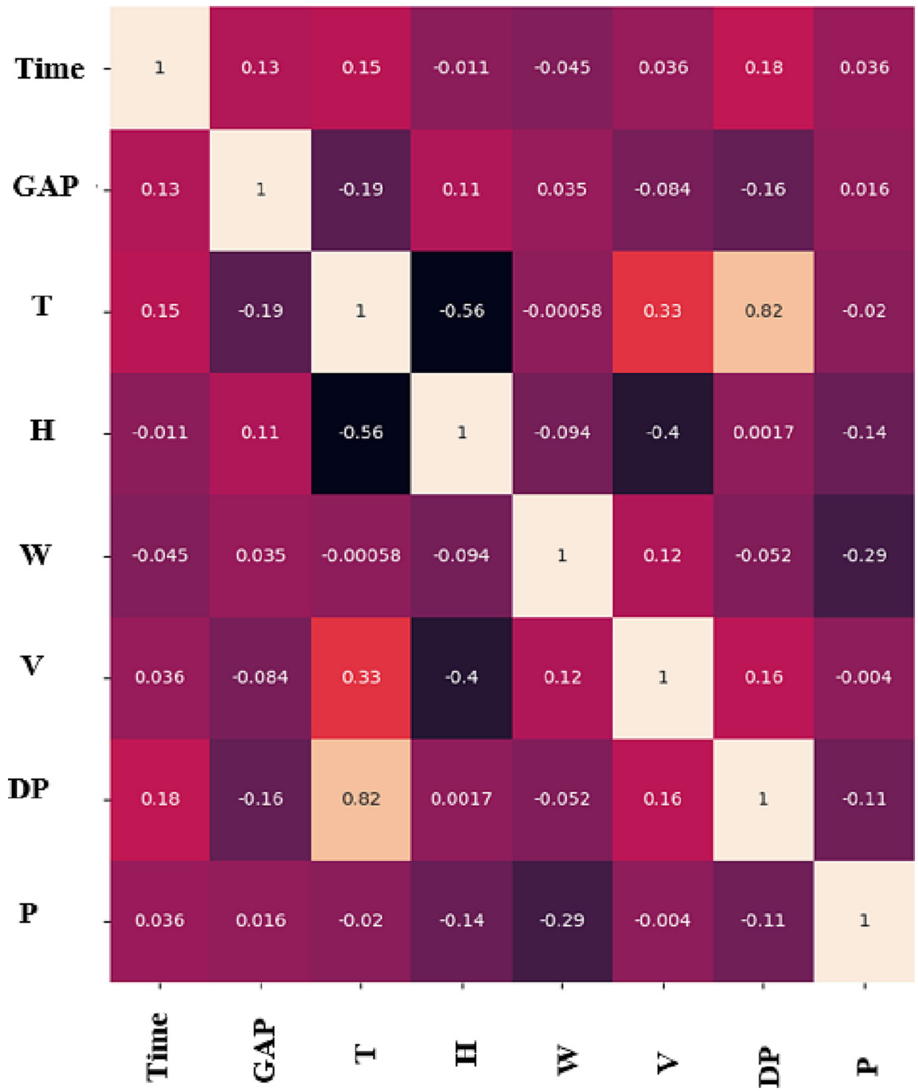

Fig. 6. Heat map of the correlation coefficient between all of the features.

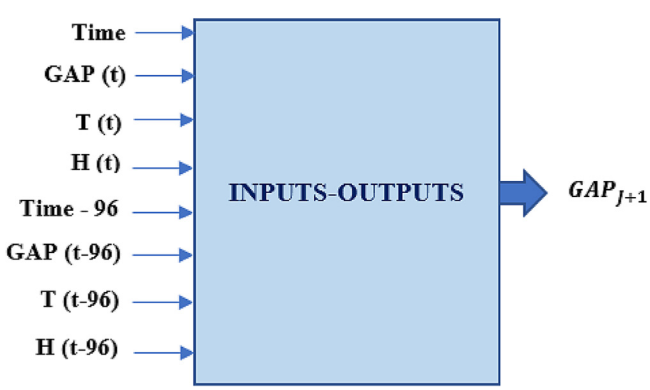

Fig. 7. Bloc diagram of the features selected.

through an EMD-Bi-LSTM model. The BI-LSTM network was used to predict each residue or IMF. The BI-LSTM network is a type of recurrent neural network that can process sequential data in both forward and backward directions.

The LSTM neural network for bus load prediction has two categories of hyperparameters: structural and training. Structural hyperparameters determine the expressive ability of the network and include the number of hidden neurons. Optimal hidden layer neurons improve performance, but too many can cause over-fitting and slow prediction. Bi-LSTM improves learning but can affect prediction time. The choice of curves can save prediction time while ensuring accuracy. Training hyperparameters, such as learning rate parameter is used to prevent over-fitting and improve the network's convergence speed and accuracy. To optimize the network's hyperparameters, this paper proposes the use of the grid search to determine the number of hidden layer neurons, whether to use Bi-LSTM and initial learning rate parameter.

The hyperparameters used for the BI-LSTM network were the same for each residue or IMF, including a hidden layer of 500 and a mini-batch size of 8. The optimization of the objective function was carried out using the adaptive moment estimation (Adam) algorithm, which is a popular optimization algorithm used for deep learning models.

One significant advantage of this method is that it can prevent multiple random errors, which ultimately enhances the accuracy of the predictions. By using a single final prediction model instead of establishing multiple models for each residue or IMF, the complexity of the model is reduced.

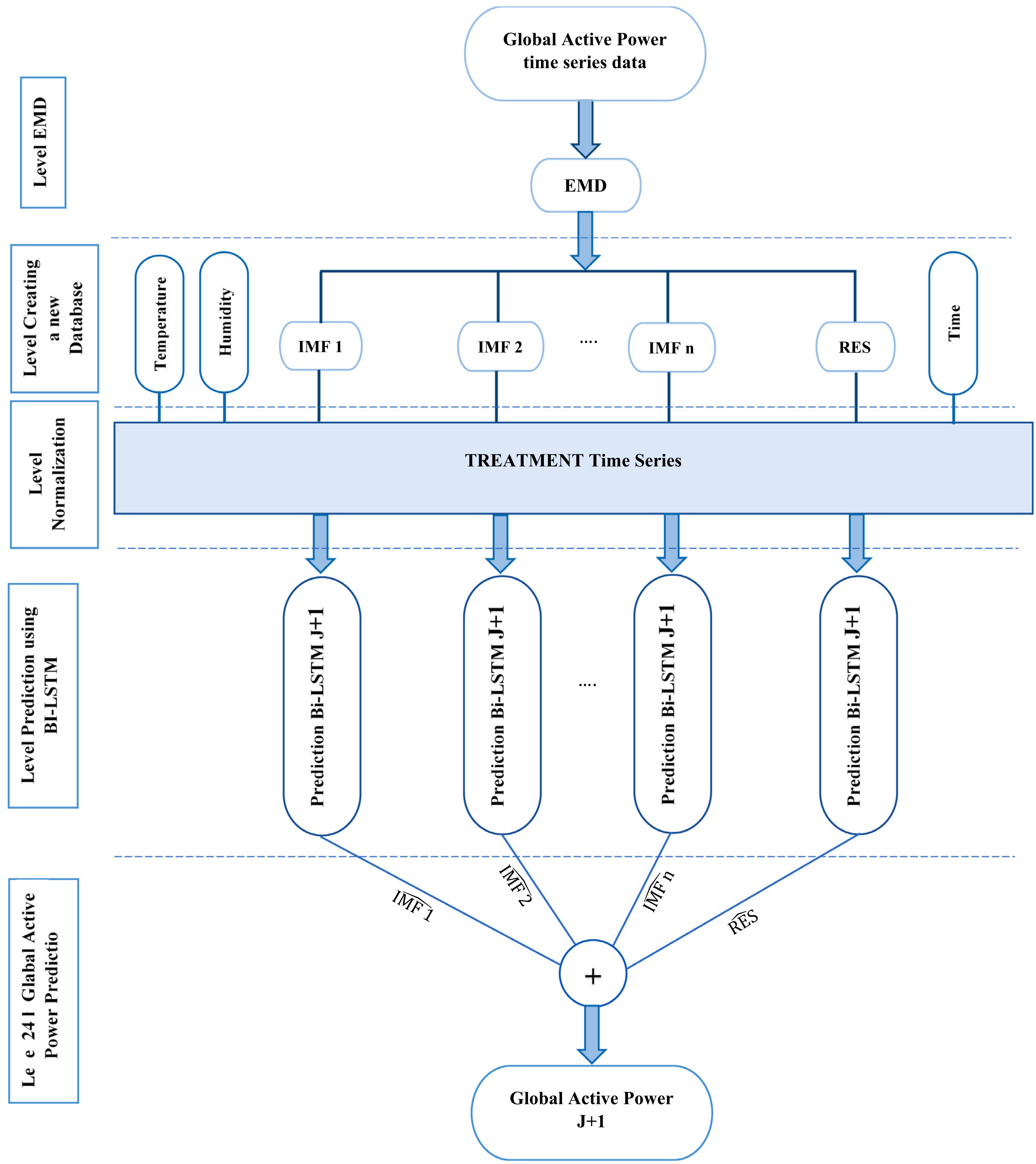

Fig. 8. The full flowchart of the EMD-BI-LSTM model.

Overall, the proposed EMD-Bi-LSTM model provides an effective solution for short-term load forecasting. The method can effectively capture the complex and nonlinear relationships between the historical load data and the future load values, leading to more accurate predictions.

3. Forecasting results

This section will present a comprehensive analysis of our model's effectiveness. We will cover four key areas. Firstly, we will demonstrate the efficacy of our model by showcasing the number of IMFs generated through decomposition. Secondly, we will compare the performance of our EMD-BI-LSTM model against a BI-LSTM model, highlighting the benefits of incorporating EMD. Thirdly, we will compare the EMD-BI-LSTM model against the EMD-BI-GRU model, showcasing the advantages of the BI-LSTM model. Finally, we will examine the impact of integrating additional features into the electricity consumption data on the model's accuracy.

3.1. Extracting signal components with the EMD

This section presents the outcomes of the empirical mode decomposition (EMD) process applied to the electric load data.

The decomposition results will be discussed in terms of their components and their relevance to the analysis of the load data. Then to validate it, we will do a signal reconstruction to know the error between this signal and the original signal.

In the case of Global Active Power, EMD decomposes it into 18 IMF components and a residual component, which represent the different scales at which influencing variables impact the power load sequence, from high to low frequency. The residual component shows the long-term changing tendency of the Global Active Power sequence.

It is evident from the figure below that the initial IMF components were random with no clear change rule, indicating that meteorological factors had a significant impact on them. The middle IMF components, however, fluctuated regularly and were similar to the original load series, with smooth day-to-day cycles, suggesting that these components were primarily determined by daily fixed electricity consumption habits. The low-frequency periodic components represented slow-changing processes influenced by meteorological factors on load changes.

By decomposing the Global Active Power sequence into its constituent IMFs, EMD provides a way to analyze the signal at different time scales and extract meaningful information about the underlying processes that affect the power load. Each IMF captures a different frequency band of the signal, and together they provide a comprehensive representation of the signal's temporal and spectral characteristics.

Moreover, the decomposed intrinsic mode functions are often more stable than the original power load sequence, which makes them useful for forecasting and modeling applications. Fig. 9 illustrates the decomposition result and highlights the contribution of each IMF component to the overall signal.

After decomposing the Global Active Power signal into its constituent IMFs using EMD, the signal can be reconstructed by summing the individual IMFs and the residual component. The reconstructed signal closely approximates the original signal, as evidenced by the high correlation coefficient of 0.99 between the two signals.

Fig. 10 shows the reconstructed signal overlaid on the original signal, and the small difference between the two signals is only 6.51×10^{-17}, which is negligible in practical applications. This high degree of accuracy in signal reconstruction demonstrates the effectiveness of EMD as a signal processing technique for analyzing power load data. The reconstructed signal can be used for further analysis and modeling to gain insights into the underlying processes that affect the power grid.

3.2. Exploring the benefits of combining Empirical mode decomposition (EMD) with the BI-LSTM

In this section, we will investigate the advantages of integrating Empirical Mode Decomposition (EMD) with Bidirectional Long Short-Term Memory (BI-LSTM) networks.

The EMD-BI-LSTM model's performance is superior to that of the BI-LSTM model, as evidenced by the reduction in MAE by

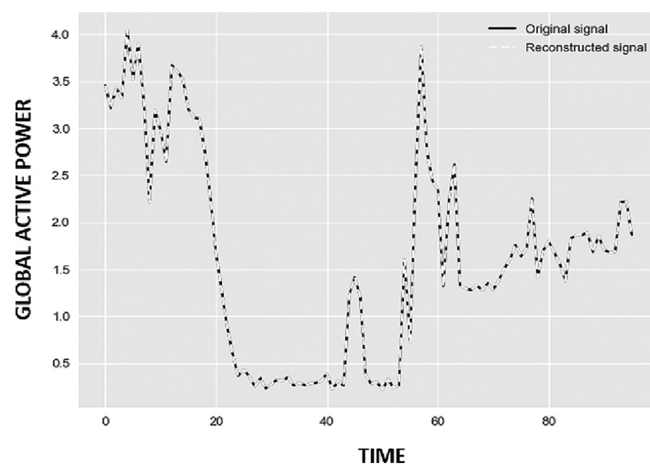

Fig. 10. The original electric load time series and the reconstructed signal.

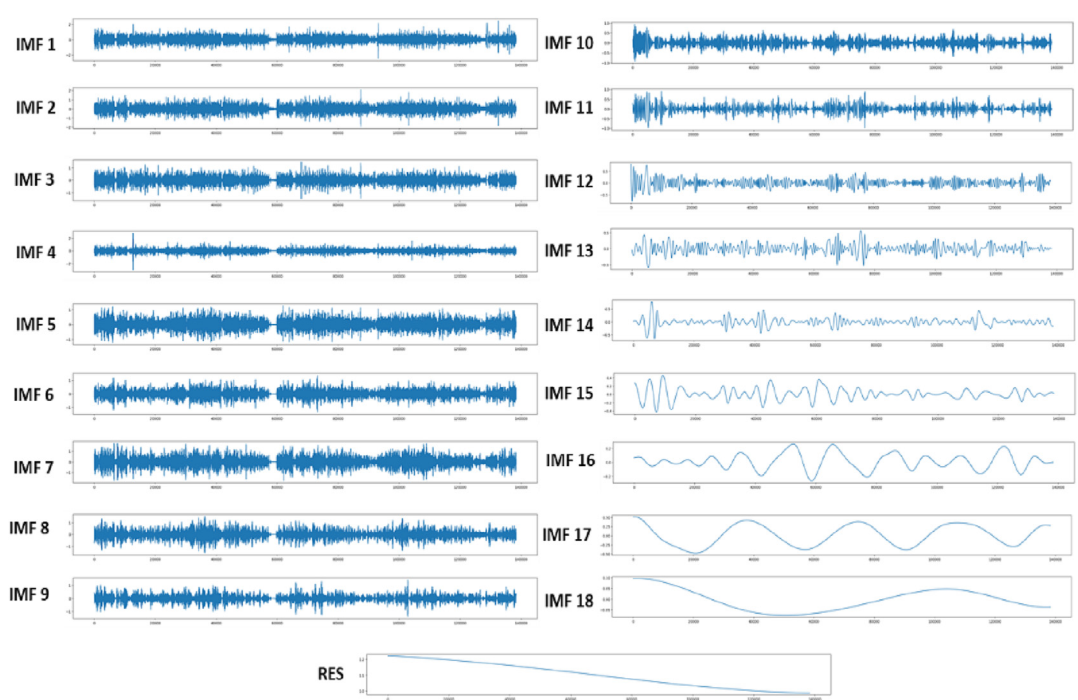

Fig. 9. IMFs and residual plot after the decomposition of the electric load series.

0.07, MAPE by 0.14%, RMSE by 0.12, and an increase in R^2 by 0.14. These results indicate that incorporating the EMD technique into the BI-LSTM model enhances its ability to capture the complex temporal and spectral characteristics of power load data, leading to better forecasting accuracy.

Furthermore, Fig. 11 shows that the prediction curve of the EMD-BI-LSTM model closely tracks the real curve, indicating the model's ability to capture the underlying patterns in the power load data. Overall, these results demonstrate the effectiveness of the proposed EMD-BI-LSTM model for power load forecasting and highlight the importance of EMD in improving the accuracy of forecasting models.

Fig. 12 shows the prediction results of the BI-LSTM model for power load forecasting. It is observed that the predicted curve has a slight lag compared to the actual data, particularly during peak power usage periods. This lag is likely due to the nonstationary of the power load time series, which makes it challenging to capture the underlying patterns in the data accurately.

Moreover, the predicted curve deviates from the actual curve, indicating that the BI-LSTM model may not be able to capture the complex temporal and spectral characteristics of the power load data effectively. To improve the accuracy of power load forecasting, it is necessary to use techniques that can stabilize the non-stationary time series and extract meaningful information from the data. One such technique is EMD, which decomposes the time series into its constituent intrinsic mode functions and enables the analysis of the signal at different time scales. The results presented earlier demonstrate that incorporating EMD into the BI-LSTM model significantly improves its forecasting accuracy and reduces the deviation between the predicted and actual curves.

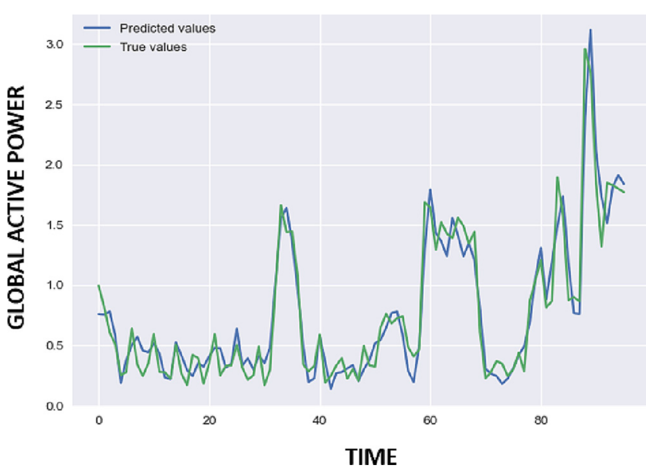

Fig. 11. A household 15 mins prediction vs actual observation EMD-BI-LSTM model.

3.3. Analyzing the performance of EMD-BI-LSTM and EMD-BI-GRU models

This section aims to compare the performance of two Empirical Mode Decomposition (EMD) based Bidirectional Recurrent Neural Networks (BRNNs), namely the EMD-BI-LSTM and EMD-BI-GRU approaches, to assess the benefits of the BI-LSTM model. We will evaluate and compare their performance in terms of their ability to accurately predict time-series data.

The BI-GRU architecture is similar to the BI-LSTM (Bidirectional Long Short-Term Memory) architecture in that it also allows for capturing dependencies in sequential data in both forward and backward directions. However, the BI-GRU has fewer parameters than the BI-LSTM and is therefore faster to train and requires less computational resources. The GRU architecture also uses fewer gates than the LSTM, which can lead to better performance in some cases where the data has less complex temporal dependencies.

Although EMD-BI-GRU produced some reasonable results, its performance was not flawless. Its MAPE value of 0.30% indicates that the model's predictions had a relatively high level of error in comparison to the actual values. Similarly, the mean absolute error of 0.21 suggests that, on average, the model's predictions were off by a significant amount. Additionally, the R-squared (R^2) value of 0.79 indicates that the model's ability to explain the variance in the data was somewhat limited. Finally, the root mean square error (RMSE) of 0.35 indicates that the model's predictions had some degree of error relative to the actual values. Therefore, while the EMD-BI-GRU model achieved reasonable performance, it may not be the best model choice for this particular dataset Fig. 13.

Table **I** displays the distribution of errors in predicting the original values for the BI-LSTM, EMD-BI-LSTM, and EMD-BI-GRU models. The errors are calculated as the absolute difference between the actual and predicted values. For each model, the table shows the number of predictions that fall within a certain error range. The EMD-BI-LSTM model had the smallest errors on average, with the majority of its predictions falling within a very small error range. The errors for the BI-LSTM and EMD-BI-GRU models are lar-

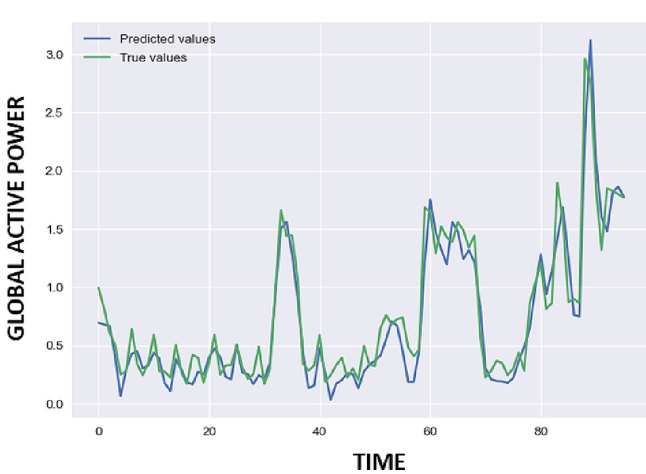

Fig. 12. A household 15 mins prediction vs actual observation BI-LSTM model.

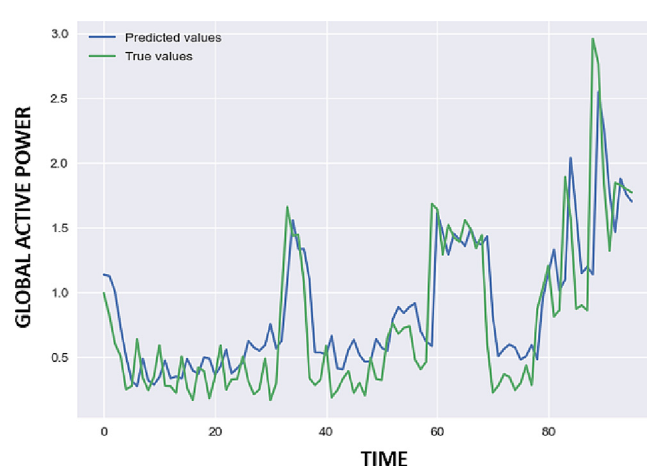

Fig. 13. A household 15 min prediction vs actual observation EMD-BI-GRU model.

ger and more spread out, with a higher number of predictions falling outside of the ideal error range.

This detailed table provide a more granular view of the prediction results, showing the distribution of errors for each model. It reinforces the conclusion that EMD-BI-LSTM outperforms the other models in accuracy and may be the most suitable model for the given dataset.

3.4. Understanding the importance of feature selection for accurate predictions

The aim of this section is to compare the predictive accuracy of our EMD-BI-LSTM model in two scenarios - one that considers both temperature and humidity as input features, and another that relies solely on energy consumption.

The table **II** compares the performance of three different EMD-BI-LSTM models with varying input features. The model that utilized both temperature and humidity achieved the best results, with the lowest MAPE (0.28), highest R^2 value (0.84), and lowest RMSE (0.31). On the other hand, the model that only used humidity as input had a significantly higher MAPE (0.55), lower R^2 value (0.51), and higher RMSE (0.54). Similarly, the model that only used temperature as input had the highest MAPE (0.38), lowest R^2 value (0.74), and highest RMSE (0.39).

Based on these results, it is clear that incorporating both temperature and humidity as input features improves the performance of the EMD-BI-LSTM model in predicting energy consumption. This is likely due to the fact that temperature and humidity are important environmental factors that affect energy consumption, and incorporating them as input features provides the model with more relevant information to make accurate predictions. On the other hand, using only one input feature (either temperature or humidity) results in poorer performance, which suggests that both input features are important for accurate predictions.

In conclusion, the EMD-BI-LSTM model with both temperature and humidity input features is the most accurate in predicting energy consumption, and using only one input feature results in inferior performance. Incorporating relevant environmental factors as input features is crucial for accurately predicting energy consumption, and these findings can be useful for developing more efficient energy management systems.

4. Conclusion

In conclusion, the proposed EMD-BI-LSTM approach for short-term electrical load forecasting has demonstrated significant improvements in accuracy with an MAPE of 0.28% and RMSE of 0.31, indicating high precision in predicting the electrical load demand over a 24-hour horizon. This is achieved through the integration of several advanced techniques, including empirical mode decomposition (EMD) and bidirectional long short-term memory (BI-LSTM) approach, and features selection.

The EMD-BI-LSTM approach has significant implications for energy management systems and the development of smart grids. Accurate short-term load forecasting is essential for efficient allocation of resources, reduction of energy costs, and improved reliability of the electrical grid. The EMD-BI-LSTM approach can provide accurate predictions that aid in the decision-making process for energy management systems, such as demand response and peak shaving.

Future developments of this study will involve the implementation of peak shaving to balance the load demand during peak periods, which can help to avoid costly peak demand charges and ensure a more stable electrical grid. Additionally, the proposed approach can be extended to address other challenges in short-term load forecasting, such as handling missing data and incorporating external factors such as holidays and events that affect electricity usage.

Data availability

The authors are unable or have chosen not to specify which data has been used.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the Ministry of Higher Education, Scientific Research and Innovation, the Digital Development Agency (DDA) and the CNRST of Morocco (Alkhawarizmi/2020/39).

References

[1] A. Almalaq et G. Edwards, « Comparison of Recursive and Non-Recursive ANNs in Energy Consumption Forecasting in Buildings », in 2019 IEEE Green Technologies Conference(GreenTech), Lafayette, LA, USA, avr. 2019, p. 1 5. 10.1109/GreenTech.2019.8767130.

[2] Z. Liu, X. Wang, Q. Zhang, C. Huang, Empirical mode decomposition based hybrid ensemble model for electrical energy consumption forecasting of the cement grinding process, Measurement 138 (mai 2019,) 314–324, https://doi.org/10.1016/j.measurement.2019.02.062.

[3] K.e. Yan, X. Wang, Y. Du, N. Jin, H. Huang, H. Zhou, Multi-Step Short-Term Power Consumption Forecasting with a Hybrid Deep Learning Strategy, Energies 11 (11) (nov. 2018,) 3089.

[4] X. Guo, Y. Gao, Y. Li, D. Zheng, D. Shan, Short-term household load forecasting based on Long- and Short-term Time-series network, Energy Rep. 7 (avr. 2021,) 58–64, https://doi.org/10.1016/j.egyr.2021.02.023.

[5] A.N. Khan, N. Iqbal, R. Ahmad, D.-H. Kim, Ensemble Prediction Approach Based on Learning to Statistical Model for Efficient Building Energy Consumption Management, Symmetry 13 (3) (mars 2021,) 405.

[6] J. Yuan, L. Wang, Y. Qiu, J. Wang, H. Zhang, Y. Liao, Short-term electric load forecasting based on improved Extreme Learning Machine Mode, Energy Rep. 7 (nov. 2021,) 1563–1573, https://doi.org/10.1016/j.egyr.2021.09.067.

[7] X. Gao, X. Li, B. Zhao, W. Ji, X. Jing, Y. He, Short-Term Electricity Load Forecasting Model Based on EMD-GRU with Feature Selection, Energies vol. 12, no 6, Art. no 6 (2019) janv, https://doi.org/10.3390/en12061140.

[8] W. Niu, Z. Feng, S. Li, H. Wu, J. Wang, Short-term electricity load time series prediction by machine learning model via feature selection and parameter optimization using hybrid cooperation search algorithm, Environ. Res. Lett. 16 (5) (mai 2021,), https://doi.org/10.1088/1748-9326/abeeb1.

[9] S. Taheri, B. Talebjedi, T. Laukkanen, Electricity Demand Time Series Forecasting Based on Empirical Mode Decomposition and Long Short-Term Memory, Energy Eng. 118 (6) (2021) 1577–1594.

[10] V. A. O. and Z. Ismail, « A New Approach to Peak Load Forecasting based on EMD and ANFIS », Indian J. Sci. Technol., vol. 6, n° 12, p. 1 7, juin 2013, 10.17485/ijst/2013/v6i12.9.

[11] Md. A. Istiake Sunny, M. M. S. Maswood, et A. G. Alharbi, « Deep Learning-Based Stock Price Prediction Using LSTM and Bi-Directional LSTM Model », in 2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES), oct. 2020, p. 87 92. 10.1109/NILES50944.2020.9257950.

[12] « UCI Machine Learning Repository: Individual household electric power consumption Data Set ». https://archive.ics.uci.edu/ml/datasets/individual+household+electric+power+consumption# (consulté le 7 décembre 2022).

[13] G. Rilling, P. Flandrin, et P. Goncalves, « ON EMPIRICAL MODE DECOMPOSITION AND ITS ALGORITHMS », p. 5.