

Research Article

Hybrid Time-Series Prediction Method Based on Entropy Fusion Feature

Jing Zhang,^{1,2,3,4} Yang Yang ,^{1,3} and Yong Feng^{1,3}

¹Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China

²Yunnan Key Laboratory of Computer Technology Applications, Kunming 650500, China

³Key Laboratory of Artificial Intelligence in Yunnan Province, Kunming 650500, China

⁴Yunnan Xiaorun Technology Service Co. Ltd., Kunming 650500, China

Correspondence should be addressed to Yang Yang; 18183689552@163.com

Received 1 March 2023; Revised 21 July 2023; Accepted 21 October 2023; Published 7 November 2023

Academic Editor: Alexander Hošovský

Copyright © 2023 Jing Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

High-precision time sequence forecasting is a complicated cyber-physical system (CPS) task. Due to the diversity of data scales and types, the classic time-series prediction model meets the challenge to deliver accurate prediction results for many forms of time-series data. This work proposes a hybrid model with long short-term memory (LSTM) and embedded empirical mode decomposition (EEMD) based on the entropy fusion feature. First, we apply EEMD in entropy fusion feature long short-term memory (ELSTM) to lessen pattern confusion and edge effects in traditional empirical mode decomposition (EMD). The sequence is then divided into intrinsic mode functions (IMF) by using EEMD. Then, feature vectors are constructed between IMFs and their respective information entropy for feature merging. LSTM is used to build a full connection network for each entropy fusion feature IMF subsequence for prediction and each type of IMF subsequence as the feature dimension to obtain its prediction results. Finally, the output results of all IMF subsequences are reconstructed to obtain the final prediction result. Compared with the LSTM method, the performance of the proposed method has been improved 64.33% on the evaluation metric MAPE. The proposed model has also delivered the best prediction outcomes across four different time-series datasets. The experimental results conclusively show that the proposed method outperforms other models compared.

1. Introduction

1.1. Research Motivation and Problem Statement

1.1.1. Motivation. Time-series forecasting (TSF) is a recent dynamic technique for managing CPS, intelligent processing, financial analysis [1], and equipment fault diagnosis [2, 3]. Hrabia et al. [4] pointed out that if an effective analysis model and prediction of COVID-19 can be developed, the epidemic prevention policy can be adjusted accordingly. TSF will play a constructive role in predictive decision-making. In addition, for CPS, the CPS sensor network will collect different types of data. In order to organize system resources more logically, it is necessary to develop a reliable regression prediction analysis model. Huang et al. [5] proposed an EMD method to perform modal decomposition on the original complex signal data and, finally, obtain multiple

intrinsic mode function (IMF) individuals with gentle characteristics. The EMD helps to decompose and extract the inherent information of data sequences; with the same purpose as neural networks in solving TSF problems, we consider using existing deep learning methods combined with EMD-derived methods for prediction to achieve more reliable prediction performance. Lu et al. [6] proposed a new hybrid approach Prophet-EEMD-LSTM based on decomposition methods and the prediction model. The fusion model can economically improve the prediction ability of the algorithm.

1.1.2. Problem Statement. Although the fusion model using the EMD method can improve the prediction ability of the model to a certain extent, the inherent problems of the EMD method, such as mode mixing and boundary effect, will

affect the prediction ability of the model. Therefore, it is necessary to develop a model with better decomposition ability to extract the hidden features of time series [7]. Moreover, the mode decomposition method itself cannot achieve orthogonal decomposition in the time domain decomposition and there is incomplete decomposition. Therefore, there must be a correlation between the IMF, and how to make full use of this feature to further extract data features is also the focus of our consideration.

1.2. Research Challenge and Gaps

1.2.1. Challenges. In common application scenarios, the sequence characteristics are typical nonstrictly stable and nonlinear data types, with high discretization features [8]. EMD is usually used to analyze complex signals with nonstrict stationary characteristics and a high degree of nonlinear dispersion in traditional signal processing. Zhao et al. [9] used a hybrid process based on the EMD and LSTM neural network method. By using the original EMD methodology, the feature extraction ability was improved and the prediction error was decreased. However, there is mode aliasing in signal decomposition. This problem will be particularly prominent when the original signal contains more high-frequency components. This problem can be dealt with by the VMD-LSTM method proposed by Niu et al. [10]. The intrinsic mode function (IMF) in signal decomposition has boundary effects between their components, which will affect the prediction accuracy [11]. Ensemble empirical mode decomposition (EEMD) aims to prevent the mode overlap problem that results from the EMD method's insufficient breakdown of data properties [12]. The method enhanced the extraction of periodic sequence characteristics and trends using the conventional EMD approach as a foundation. In an attempt to lessen aliasing across IMF components and EEMD-LSTM proposed superimposing Gaussian noise on the original signal [13, 14]. Based on the EEMD components [14], the data are divided into high-frequency and low-frequency sequences using the enhanced Bi-LSTM.

1.2.2. Research Gap. Existing research usually focuses on how to improve the bias of modal aliasing and boundary effects on network models through more optimized decomposition models. However, current mainstream modal decomposition fusion models ignored the correlation between IMF components obtained from pattern decomposition. At the same time, this kind of decomposition method cannot achieve complete decomposition (orthogonal decomposition). Moreover, in the framework of the current EMD-LSTM method, each IMF uses independent prediction methods to reconstruct. This step places too much emphasis on the characteristic results of the modal component itself, which will cause the prediction result to be skewed to the offset or special mode for incomplete modal decomposition. In general, the abovementioned model will inevitably have the defect of ignoring the interaction between IMFs. Furthermore, the hidden features between the time-series features of IMFs are also ignored in the construction of most fusion models.

1.2.3. Research Novelty

- (1) A hybrid time-series prediction model with EEMD and LSTM based on entropy fusion feature is proposed, namely, EEMD-ELSTM. It consists of four parts: (1) EEMD, (2) entropy feature fusion, (3) LSTM network, and (4) prediction outcome reconstruction.
- (2) EEMD method can reduce mode aliasing and boundary influence, and the entropy feature fusion method calculates the entropy of the IMF components obtained, and then adjusts the balance weight between the IMF components with high entropy and low entropy in the fully connected network by increasing the entropy sorting fusion method, thereby ensuring strong correlation between the dimensional components of the fully connected LSTM network and improving the reliability of LSTM prediction results.
- (3) The eigenvector frames between similar modal functions are processed by fully connected modal decomposition LSTM to improve the prediction accuracy of the final reconstruction data.

1.2.4. Main Contributions. The main contributions of this work are summarized as follows:

- (1) The EEMD algorithm can decompose hidden informative features from time-series data onto multiple characterized IMF components
- (2) The ELSTM phase of the algorithm can map the hidden characteristics of IMF components and pass on the same characteristics among components using entropy fusion, relying on the learning characteristics of the LSTM network
- (3) In the LSTM stage, the algorithm is able to map the hidden characteristics between IMF components through entropy fusion according to the learning characteristics of the LSTM network and transfer the commonality between the components
- (4) The proposed optimization algorithm can be applied to different datasets and has general applicability

2. Related Work

The current research in the time-series prediction field can be divided into three main categories [15, 16].

2.1. Statistical Model-Based Method. Prajapati and Kanojia [17] examined the index that has the largest influence on the fluctuation of COVID-19 for India by using autoregressive integrated moving average (ARIMA) and autoregressive (AR) models. Behzadi et al. [18] proposed a general information theory framework based on the generalized linear model (GLM), which was applied to the causal inference of heterogeneous datasets and verified on celestial data. Kap et al. [19] proposed the additive noise model (ANM) on noise level for time-series analysis, which confirmed the

effect and significance of the noise overlaid on the extraction sequence in the time-series prediction model. Hanapi et al. [20] proposed the fuzzy sliding window autoregressive conditional heteroskedasticity model for time-series prediction and applied it to aerial data. Studies based on time-series decomposition [21] that used the Prophet model can handle time-series data with default values and predict itself. However, time-series data often contain many components with cyclical characteristics and nonsignificant trends, which are not well taken into account by the aforementioned models.

2.2. Deep Learning-Based Method. Most of the time-series data streams collected by sensor networks in CPS are highly discrete and nonstationary, which makes accurate trend forecasting difficult to achieve. The spatiotemporal properties of the time-series data are not simply relative to real-time data about the current temporal nodes but are based on the traditional Markovian nature [22]. It will result in the loss of data feature information. Huang et al. [23] elaborated on related methods based on artificial neural networks (ANN), such as recursive neural networks (RNN) in the field of deep learning. Depending on the connections between hidden units in the network, it will allow the network to inherit the information characteristics of related events. Zeng et al. [24] proposed a linear combination model based on the Prophet model and LSTM, which provided a new idea for regression analysis in constructing the fusion model of diversified LSTM networks. Li et al. [25] used wavelet transform to process the signal for the regression analysis of user traffic. Zhang et al. [26] decomposed several IMF sequences through variational modal decomposition (VMD) in the study of nonstationary wind speed series. They conducted secondary denoising for each IMF and then trained in the improved residual neural network (PCA-BP-RBF), which significantly increased the prediction performance. Although the VMD method can effectively eliminate the problem of modal aliasing, there are significant differences between the decomposition results obtained by redefining the IMF function and EMD and its variants. In addition, before performing VMD decomposition on the sequence, the number of modal components K needs to be set first. However, this step needs to be based on a great deal of prior experience, and in some applications, specifying the number of IMFs is an advantage. However, for certain scenarios where the number of hidden modes of the signal is not predicted, calculating and setting this optimal K -value will actually increase the time cost of modeling. Moreover, this kind of model has a high degree of complexity and is even unsuitable for systems with low time delay requirements.

2.3. Machine Learning-Based Method. The least squares support vector machine (LSSVM) was created by Gong et al. [27]. To optimize parameters on the basis of the particle swarm optimization (PSO) algorithm, Montesinos López et al. [28] addressed the issue of high algorithm time complexity by using the sequential minimal optimization

(SMO) algorithm to optimize objectives following the quadratic programming problem of the support vector machine (SVM) and support vector regression (SVR). Pekel [29] used decision tree regression (DTR) to predict soil moisture. Jumin et al. [30] conducted the regression study of solar radiation using the boosted decision tree regression (BDTR) model. Qiu et al. [31] optimized the extreme gradient boosting (XGBoost) model and improved the prediction accuracy compared with other XGBoost-based models. This kind of method does not require a massive amount of data, but their model structures are complex.

For the three methods described above, we draw the following conclusions. First, machine learning and statistical learning have a large overlap, or machine learning is based on statistical learning. Second, statistical learning is theory-driven, making assumptions about data distribution, explaining cause and effect with strong mathematical theory support, and focusing on parameter inference. Third, machine learning is data-driven, relying on big data scale to predict the future, weakening the convergence problem, and focusing on model prediction. Fourth, deep learning is a subfield of machine learning; that is, feature extraction relies more on hidden layer models, weak explanatory properties, and tends to black boxes.

From the three main research categories mentioned above, it is not difficult to draw the following conclusions. Machine learning and deep learning methods have shown great advantages in related research. Moreover, the research of fusion algorithms based on modal decomposition optimization and neural networks also reflects its excellent reliability. Therefore, in this paper, we try to combine the related methods of modal decomposition and LSTM methods.

The EMD decomposition of the correlation prediction model of the sequence data with time characteristics has been developed in recent years [32–34]. Zhang et al.'s [35] derived model of EMD decomposition improved the gating cycle unit (GRU) and combined EMD with the regression prediction model of PM2.5. Ali et al. [36] pointed out that they a proposed new version of EMD based on the Akima spline interpolation technique and LSTM network, and at the same time, this method enhances the effectiveness of the improved model. Dedovic et al. [37] used EMD and ARIMA to predict air quality. Liu et al. [38] stated that combining the EMD model and ANN can improve the prediction effect. However, most works are based on single type data and there are still modal aliasing and boundary effects. In addition, the relevant algorithms did not fully use the neural network characteristics to mine data features. Therefore, this paper proposes EEMD-ELSTM to ameliorate the prediction accuracy and universality of the model.

3. Method

The global EEMD-ELSTM framework is described as follows:

- (1) EEMD using Gaussian white noise is used to deconstruct the source series data for the period

- (2) In the decomposed IMF, the entropy value is analyzed and classified according to the entropy value
- (3) The LSTM network is then used to conduct full connection prediction for IMF subcolumns of each category to obtain prediction results
- (4) All IMF subcolumns are then combined and reconstructed to obtain the final forecast result

The abovementioned process is summarized in Algorithm 1.

3.1. LSTM and EMD Relational Model

3.1.1. LSTM. Lobo Neto et al. [39] pointed out that LSTM further optimizes the network performance of RNN and impressive results have been obtained on various time-series problems. Generally, an LSTM network is composed of storage units. Three special cell structures are responsible for updating network data: output gate, input gate, and forgetting gate. Its structure is displayed in Figure 1, where \odot indicates Hadamard product and $+$ is matrix addition.

Inputs for LSTM typically consist of c^{t-1} state parameter, h^{t-1} state parameter, and x^t state parameter. The three outputs are the c^t state parameter, the h^t state parameter, and the y^t state parameter, where the x^t state parameter represents the current round's input, the h^{t-1} state parameter represents the round's state quantity output, and the c^{t-1} state parameter represents the round's global information carrier. Then, the y^t state parameter represents the current round's output, the h^t state parameter represents the round's status output, and the c^t state parameter represents the round's global information carrier.

The LSTM calculating process is as follows. First, the four states for splicing training are obtained by utilizing the current input's x^t state parameter of the LSTM and the h^{t-1} state parameter got past from the preceding state parameter, and these four states' parameters are expressed as z^f, z^i, z^o, z . After being multiplied by the weight matrix, the splicing vector is turned into a single parameter through a sigmoid activation function as a gating state. z turns the output into such a number between both -1 and 1 through using the tanh activation function.

Second is the forgetting stage, that is, the forgetting gate. Selectively forgetting the information from the preceding node is the primary task of this step. Depending on the abovementioned state parameter z^f , f is expressed as forget, which is used as the control function of the forgetting gate to decide whether the data in the previous state c^{t-1} needs to be forgotten and output a value between 0 and 1 , where 1 means to be completely retained and 0 means to be completely forgotten. The z^f function is expressed as follows:

$$z^f = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f). \quad (1)$$

The third stage is selecting the memory stage and inputting the gate. This stage will selectively memorize the input time-series data. It is mainly used to selectively memorize the input x^t . The selected gating signal is represented by the state parameter z , the current input is

represented by the state parameter z^i , and i is information. The next stage of transmission c^t can be obtained by adding the abovementioned two steps.

$$z^i = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad (2)$$

$$z = \tanh(W_z \cdot [h_{t-1}, x_t] + b_z), \quad (3)$$

$$c^t = z^f \odot c^{t-1} + z^i \odot z. \quad (4)$$

The final stage is the output stage and this stage will decide whether the data will be output as the current state. The decision is mainly made through the state parameter z^o . In addition, the c^t obtained in the previous stage is scaled through the tanh activation function mentioned above. Compared with the traditional RNN, y^t output is also obtained through the change of h^t .

$$z^o = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \quad (5)$$

$$h^t = z^o \odot \tanh(c^t), \quad (6)$$

$$y^t = \sigma(W' h^t). \quad (7)$$

On the basis of the traditional RNN chain structure, LSTM uses a special gate structure memory unit to replace the originally hidden nodes, which enhances the overall network's ability to retain time-series data information and extends the network's long-term memory ability.

3.1.2. EMD. It disintegrates the sequence data into multiple subcolumns according to the characteristics of the time scale itself and does not need other basis functions at the same time. In the structure of the decomposition algorithm, the main part is empirical decomposition.

Usually for a data sequence with temporal characteristics, $M = \{m_1, m_2, \dots, m_n\}$, EMD can decompose multiple subcolumns of IMF and a residual component, stack staggers through multiple IMF subcolumns, and reconstruct the original sequence $x(t)$ as

$$X = \sum_{i=1}^N C_i(t) + r_n(t). \quad (8)$$

Among them, $r_n(t)$ is able to be such stagger component, displaying the general current of such sequence resources.

The EMD method is used with the aim of extracting, from the raw signal, the high- and low-frequency sequences of such signal, as well as the various scale components, and arranging them in the plain sequence of frequency from high to low, so as to obtain the plain sequencer of IMF.

The main problems of the EMD algorithm mentioned in Section 2 can be described as follows [40–42]:

- (1) Boundary effects: the endpoints of the entire time-series data are usually not extreme points. Therefore, the constructed envelope function diverges at the beginning and at the end of the sequence due to the uneven distribution of the endpoints. The deviation

Input: data steam $M = \{m_1, m_2, \dots, m_n, \dots\}$, $m_n \in R$, epoch I , the number of iterations K , error parameters σ , cycle index N , number of decomposition n , white Gaussian noise W , and entropy classification number E

Output: Predicted result O

- (1) $M \leftarrow m_1, m_2, \dots, m_n, \dots, m_n \in R$ //input data
- (2) $O \leftarrow \{\}$ //output data
- (3) **for** each $i \in N$ **do**
- (4) $H_i \leftarrow m + W_i$
- (5) **for** $j \in I$ **do** $H_{ij} \leftarrow EEMD(H_i)$
- (6) **end for**
- (7) **end for**
- (8) **for** $j \in n$ **do**
- (9) $E_t \leftarrow (Entropy\ feature\ fusion(IMF_j))$
- (10) **end for**
- (11) **for** each $t \in E$ **do**
- (12) $O \leftarrow \sum ELSTM(IMF_j, E_t)$
- (13) **end for**

ALGORITHM 1: EEMD-ELSTM approach.

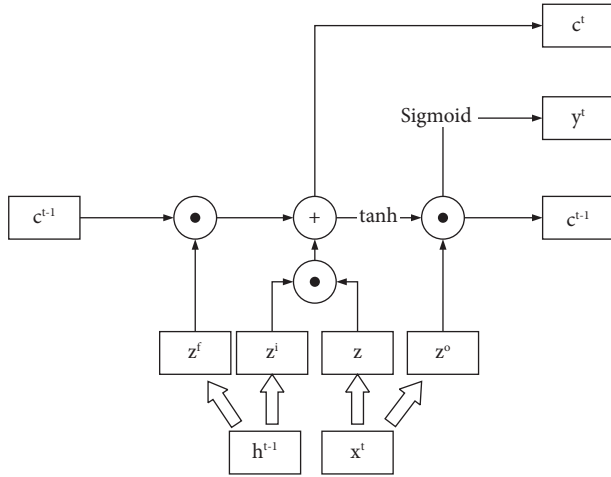


FIGURE 1: LSTM structure.

caused by this phenomenon will continue to be superimposed in the decomposition course, which will eventually interfere with the decomposition reliability.

- (2) Mode aliasing: IMF after overall sequence data decomposition has incomplete decomposition and components of different scales and frequencies are mixed in a subcolumn. When the abovementioned phenomena occur in multiple subcolumns at the same time, the EMD algorithm will lose its physical meaning.

3.1.3. EEMD. To address the abovementioned issues, we use superimposed noise to extend the EEMD of the sequence. The EEMD used in this paper superimposes Gaussian white noise. These are the precise processes of breakdown.

Given the input time-series data $M = \{m_1, m_2, \dots, m_n\}$, the Gaussian white noise $W_i, i = 1, 2, \dots, n$ is loaded with an additive mean of 0 into the sequence M .

$$H_i = M + W_i, \quad (9)$$

where i represents the amount of Gaussian white noise added. Comparisons between EMD and EEMD based on the UCI power network dataset are shown in Figure 2. The EEMD method extends the extremes of the original EMD process during decomposition to alleviate the extreme trailing effect and mode overlap at the endpoints.

In addition, for the IMF subcolumns that can be obtained by the decomposition of the abovementioned two methods, it is obvious that the EEMD method can obtain more subcolumns. At the same time, the IMF subsequences obtained by the modal decomposition algorithm are generally different and random.

- (a) EEMD model
- (b) EMD model

The EEMD method extends the extremes of the original EMD process during decomposition to alleviate the extreme trailing effect and mode overlap at the endpoints. It can be seen from Figure 2 that the boundary effect and modal aliasing problems that arise during the decomposition of the EMD algorithm are effectively improved in the EEMD model after superimposing Gaussian white noise.

$$IMF_i = \{imf_{i1}, imf_{i2}, \dots, imf_{ij}\}, \quad (10)$$

$$j = 1, \dots, n.$$

The abovementioned steps are then repeated until the entire time-series data are decomposed N times. Each time, a new Gaussian white noise is added to the series, and finally, the entire IMF subcolumn can be obtained.

$$IMF_i = \{imf_{i1}, imf_{i2}, \dots, imf_{ij}\}, \quad (11)$$

$$i = 1, \dots, N; j = 1, \dots, n.$$

Integrated averaging on imf_{ij} is performed as mentioned in the above steps to obtain imf of the overall sequence M .

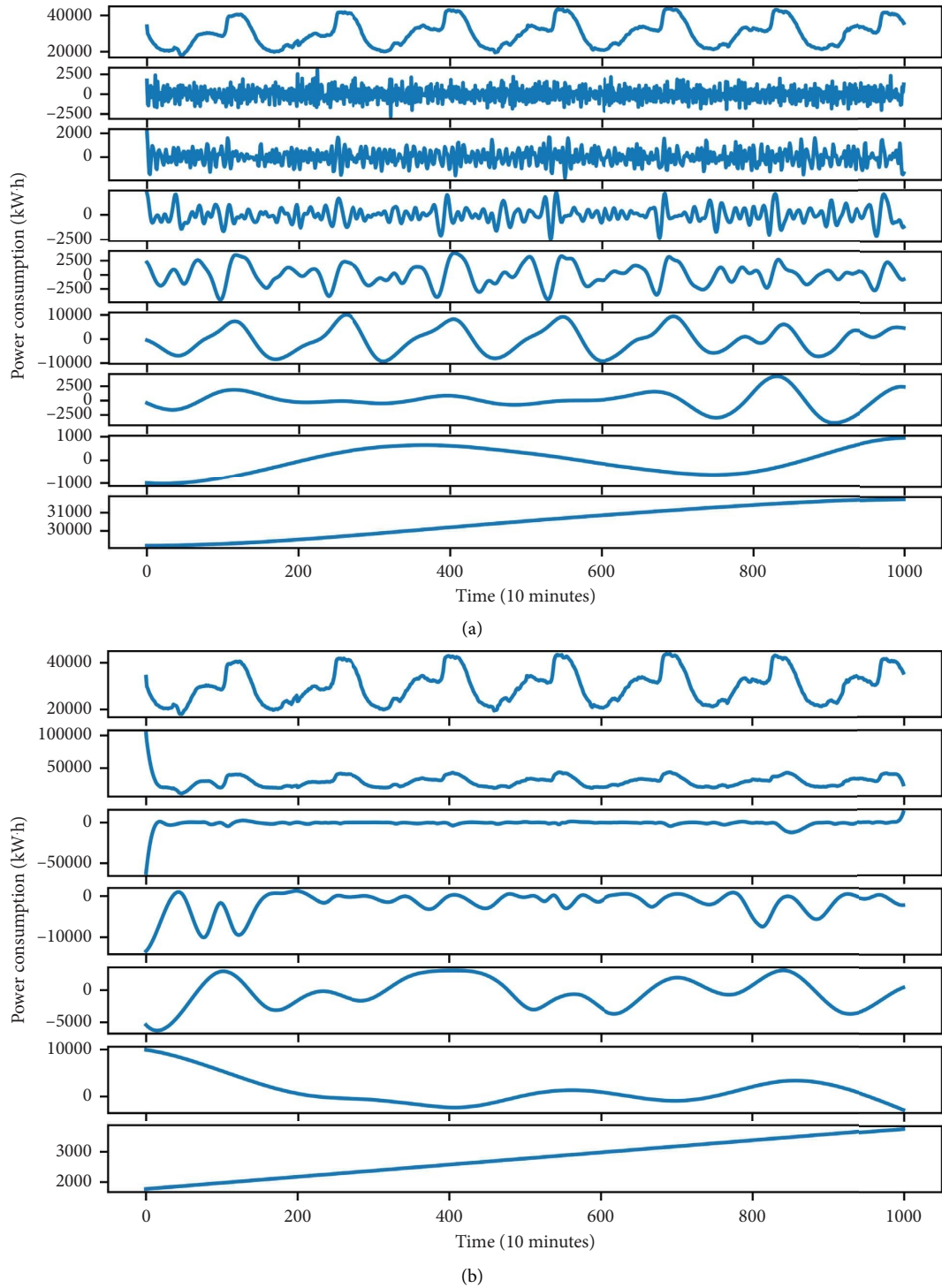


FIGURE 2: Ensemble empirical mode decomposition examples: (a) EEMD model and (b) EMD model.

$$IMF = \{imf_j\} \quad j = 1, \dots, n, \quad (12)$$

$$imf_j = \sum_{i=1}^N ((imf_{ij})/2N). \quad (13)$$

3.2. *EEMD-ELSTM Algorithm.* By integrating the EEMD and ELSTM models, the EEMD-ELSTM model can be created. Overall, EEMD-ELSTM can be divided into three steps: (1) EEMD data decomposition, (2) entropy feature fusion, and (3) combining the LSTM prediction results of

each IMF subcolumn. The EEMD-ELSTM model algorithm flow structure presented in this paper is depicted in Figure 3.

In Section 3.1, to break down the data series, we employ the EEMD approach, $m = \{m_1, m_2, \dots, m_n, \dots\}$, and gain the $imf = \{imf_j\}$, $j = 1, n$. Then, we use the ELSTM algorithm to anticipate each imf_j , $j = 1, \dots, n$, providing each IMF series' forecast outcomes.

$$L_j = ELSTM(imf_j), \quad j = 1, \dots, n. \quad (14)$$

After the calculation iteration, the L_j component of each $\{imf_j\}$ subsequence is obtained. We then perform entropy analysis on the abovementioned imf subsequence and take the two highest entropy score subsequences of imf subsequence as the feature dimension, and further construct the feature vector of the ELSTM model. The specific implementation steps of the entropy value feature fusion model defined in this section are as follows: first, calculate the entropy value of each IMF component and then screen the two maximum entropy values after obtaining the entropy value score. Second, for the model of this part, based on modal decomposition, we will get a residual under the ideal decomposition condition since a good residual is usually close to a linear function of one variable. At the same time, based on entropy theory. We can know that under the ideal condition of decomposition, the residual will be a maximum entropy sequence. However, in order to fully exploit other hidden data features after decomposition, we further select an IMF component with the second highest score as an auxiliary dimension. Third, the specific step of constructing the feature vector is to add two high entropy components as the feature vector for the LSTM sequence structure with three inputs and one output for the low entropy sequence training, while when constructing the LSTM sequence structure for the high entropy components, all IMF components are calculated as their auxiliary dimensions. For the low entropy sequence, the feature trend for the high entropy sequence is increased, while the high entropy sequence acquires the high-frequency features of the rest of the sequences. So as to achieve the model optimization, we ultimately expect and improve the robustness of information physical system modeling.

The value of the final predicted sequence data m can be expressed as

$$F_m = \sum_{j=1}^n (L_j). \quad (15)$$

The prediction formula for $M = \sum(imf_j)$ series is as described above so that the abovementioned L can be regarded as the prediction value of IMF of each subcolumn. The amount of import data n determines the algorithm complexity of the traditional EMD model, which is $O(n^2)$. Our EEMD model performs multiple iterations and cycles on the original EMD model with Gaussian white noise added and subtracted.

As a result, the EEMD method has an $O(k * n^2 + k)$ complexity, in which k seems to be the number of cycles. In the entropy fusion model, calculating the entropy once and

dividing the threshold according to the entropy value is approximate to the K -means classification algorithm. Consequently, this decomposition algorithm complicity is defined as $O(n)$. Since such complexity that belongs to the LSTM depends on the input size n and the hidden size m , its complicity is approximately $4(m * n + m^2 + m)$, and each prediction of the ELSTM network will increase by $n - 1$ dimensions; thus, ELSTM has a complexity of $(m * n + m^2 + m * (n - 1)) + n$. When it is reduced to $m^2 + m * n + m + m * (n - 1)$, where $m * n$ and $m(n - 1)$ are equal levels, the complicity of the method is $O(k * n^2 + m^2 + m * n)$.

4. Experiment

4.1. Experimental Settings. The datasets used in this work are provided by the University of California. The experimental results show that the proposed approach overcomes the drawbacks of the conventional EEMD-LSTM method and further improves its accuracy compared with the state-of-the-art (SOTA) algorithms. Four different types of datasets of Tetouan and Morocco from the UCI database (2017) are used in this work: (1) the power grid power consumption, (2) the temperature, (3) the wind speed, and (4) the humidity. The datasets are listed in Table 1. The schematic design is shown in Figure 4. The dataset that is used in this comparative study is related to power distribution networks of Tetouan city which is located in north Morocco. The historical data used have been taken from the Supervisory Control and Data Acquisition System (SCADA) every 10 minutes for the period between 2017-01-01 and 2017-12-31. We used 1000 sample sequences from this dataset in our study.

It can be seen from Figure 4 that the characteristics of nonstrict stability and nonlinearity are in line with the problem we expect to solve. In order to better visualize the distribution of the aforementioned dataset we used, we also conducted box plot analysis on the data samples. The results are shown in Figure 5, and it can be seen from the box plot analysis results that there are great nonperiodic and discrete characteristics in the wind speed data in the abovementioned data distribution. In summary, the temporal characteristics of the four research objects are consistent with the problems that our proposed research model expects to solve in cyber-physical systems.

From Figure 6, we know that, from the decomposition results, the power consumption data have the most typical periodic characteristics, followed by local periodic characteristics of wind speed and temperature. Wind speed data have strong discreteness and nonperiodicity, and there are outliers in the data sequence.

Based on the decomposition results of Figure 5, we perform the entropy value calculation operation on the IMF component in Table 2, and combined with our previous description of the entropy fusion feature model, we can clearly know from the entropy theory that when there is a sequence with strong periodicity IMF, the entropy calculation result will inevitably be a low value, and the amount of hidden information contained in it itself is also low. In

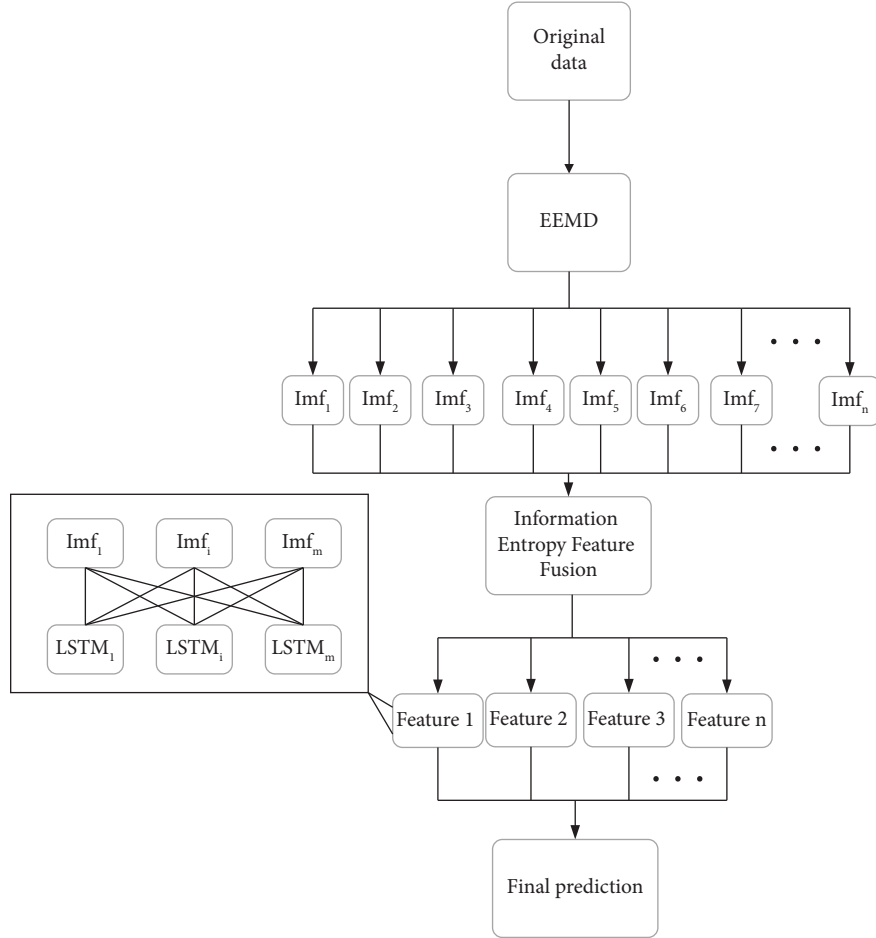


FIGURE 3: Proposed EEMD-ELSTM algorithm.

TABLE 1: Information from four categories of statistical data sources.

| Data | Power consumption (kW · h) | Temperature (°C) | Wind speed (M/s) | Humidity (%) |
|--------|-------------------------------|------------------|------------------|--------------|
| Max | 44831.89 | 18.40 | 0.16 | 88.86 |
| Min | 18045.57 | 4.21 | 0.07 | 50.65 |
| Mean | 30100.37 | 11.34 | 0.09 | 73.57 |
| Var | 49432194.31 | 12.05 | 0.18 | 87.39 |
| Std | 7030.80 | 3.47 | 0.42 | 9.35 |
| Median | 29966.58 | 11.53 | 0.08 | 75.21 |

addition, through the entropy calculation score, we can also clarify the periodic characteristics of local cycles and problems between power consumption data and temperature and humidity, but the frequency bands are different, and the wind speed data has a modal gradient completely different from the above three, so the four data sets can represent the complex and changeable real environment in the cyber-physical system to a certain extent. This is one of the reasons why we chose the abovementioned four datasets as the model test objects.

To evaluate the reliability and dependability of the proposed model, the root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error

(MAPE) functions, and Pearson correlations (CORR) are used as the indicators [43–45].

$$\text{RMSE} = \frac{1}{N} \sqrt{\sum_{i=1}^N (y_i - \hat{y}_i)^2}, \quad (16)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - x_i|, \quad (17)$$

$$\text{MAPE} = \frac{100\%}{N} \sum_{i=1}^N \frac{|y_i - x_i|}{x_i}, \quad (18)$$

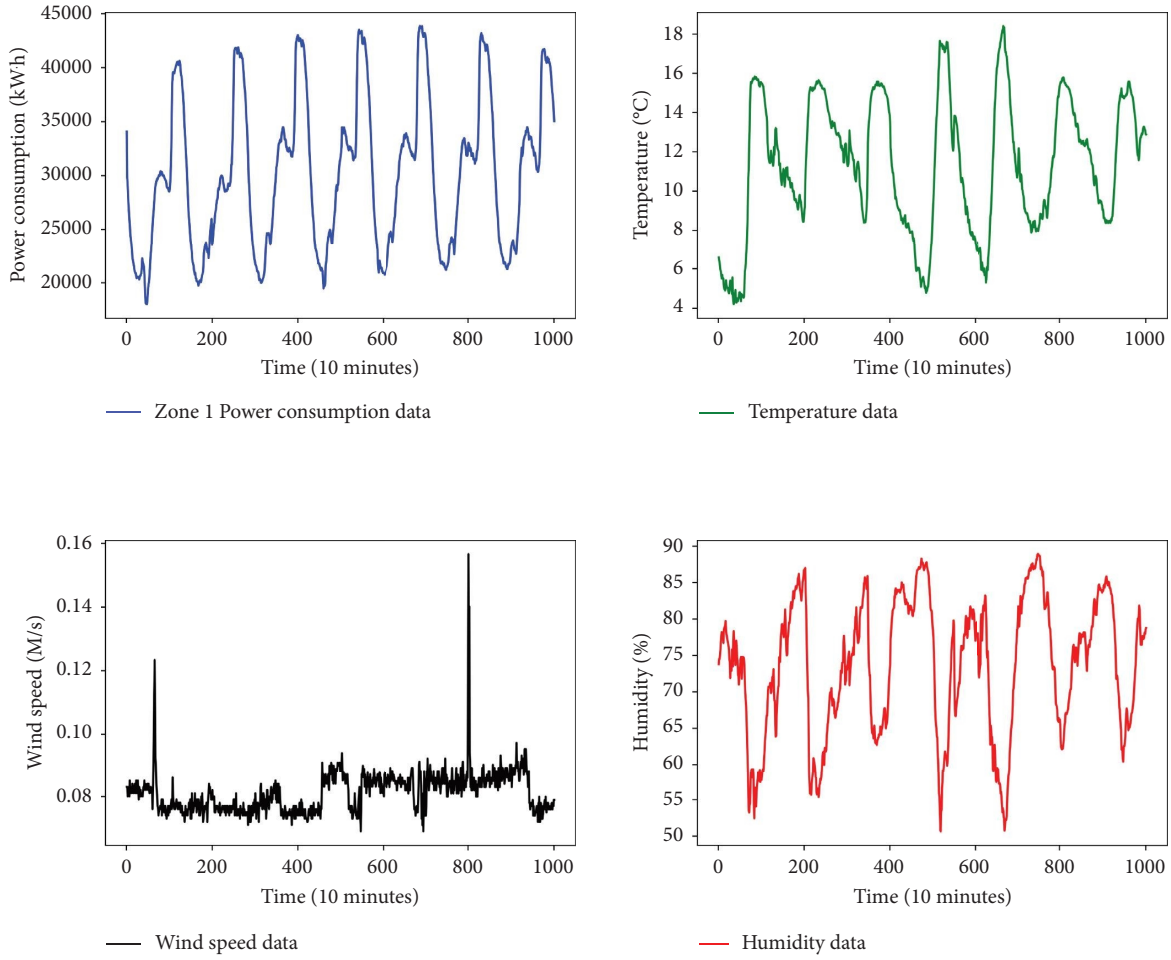


FIGURE 4: Time-series data flow statistics of four different datasets.

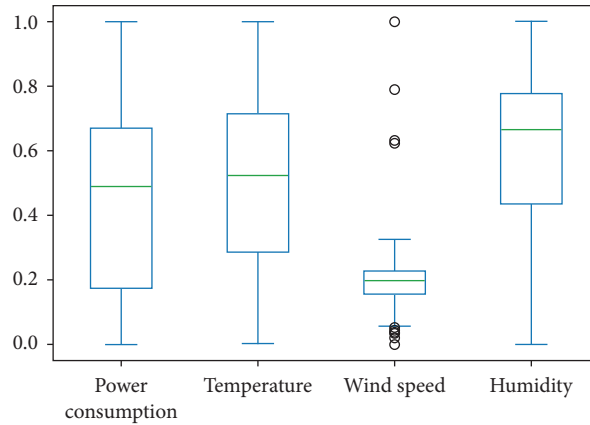


FIGURE 5: Box plot distribution of data samples.

$$\text{CORR(Pearson)} = \frac{\text{cov}(y_i, \hat{y}_i)}{N \sqrt{\text{var}(y_i)\text{var}(\hat{y}_i)}} \quad (19)$$

where y_i is the forecasting outcome, \hat{y}_i is the authentic value resource, $\text{cov}(y_i, \hat{y}_i)$ is the covariance, and $\text{var}(y_i)$ and $\text{var}(\hat{y}_i)$ represent the variance.

4.2. Parameter Setting. The noise added to the experiment is assumed to be zero-mean Gaussian noise. In addition, an unlimited amount of data decompositions is allowed to pursue total decomposition. The EEMD-ELSTM model in this paper is validated by shuffle-split cross-validation and all the result parameters are the optimal value of the result of 10-fold validation.

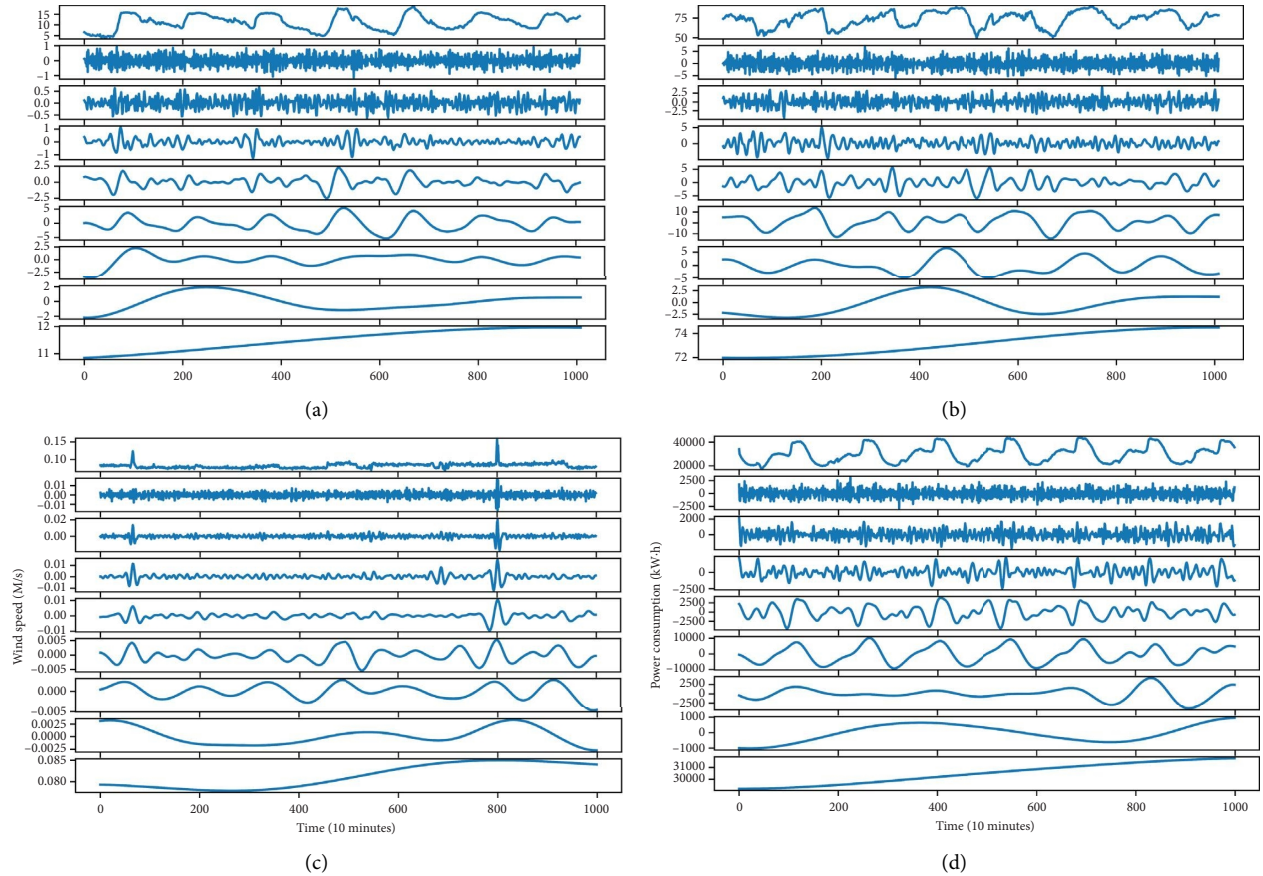


FIGURE 6: Decomposition results of the EEMD method for four different datasets: (a) temperature analysis, (b) humidity analysis, (c) wind speed analysis, and (d) power consumption analysis.

TABLE 2: IMF component entropy score obtained from time-series data decomposition.

| IMFs | Power consumption | Temperature | Humidity | Wind speed |
|------|-------------------|-------------|----------|------------|
| IMF1 | 0.030 | 0.069 | 0.060 | 0.006 |
| IMF2 | 0.037 | 0.099 | 0.037 | 0.011 |
| IMF3 | 0.023 | 0.043 | 0.041 | 0.014 |
| IMF4 | 0.084 | 0.075 | 0.055 | 0.023 |
| IMF5 | 0.158 | 0.159 | 0.092 | 0.032 |
| IMF6 | 0.183 | 0.078 | 0.165 | 0.051 |
| IMF7 | 0.220 | 0.198 | 0.248 | 0.138 |
| IMF8 | 0.266 | 0.280 | 0.303 | 0.203 |

In our comparative experiments and related ablation validation experiments, we set the following parameters.

The batch scale is set to 25, the epoch scale is set to 50, the validation split is 0.1, and the loss function is a mean square error (MSE), and the optimizer chose the Adam and Dense value as 1. There are 40 neurons in the LSTM and ELSTM variable parameter settings. The SVR parameters are $C = 100$, degree 3, gamma 1, and the kernel is *rbf*. The ARIMA parameters are $p = 2$, $q = 2$, and $d = 0$. For the Prophet model and the XGBoost model, we also give their hyperparameter definitions in Table 3. The data are divided into two halves for training (fifty percent) and testing (fifty percent). All experiments are made in the Anaconda with TensorFlow 2.1 platform.

4.3. Ablation Experiment. The ablation experiments are compared and analyzed on the dataset power consumption for LSTM, ELSTM, EEMD-LSTM [46], and EEMD-ELSTM algorithms. The results are listed in Table 4.

The results show that the entropy fusion model and EEMD approach have enhanced the LSTM computation optimization. The prediction precision of the LSTM model can be improved by introducing the EEMD-ELSTM algorithm. In addition, it can be known from the indicated parameters in Table 4 that the EEMD-ELSTM method has the most optimized results among the abovementioned methods, and the ELSTM using the entropy fusion method is also optimized to a certain extent.

TABLE 3: Experiment settings of each control model.

| Method | Hyperparameter setting |
|------------|---|
| EEMD-ELSTM | Hidden layer: 1 |
| EEMD-LSTM | Hidden layer neuron: 40, 50 |
| ELSTM | Learning rate: 0.0001 |
| LSTM | Batch size: 25, 5 |
| Prophet | Changepoint prior scale: 64 Interval width: 3 Growth: linear Changepoints: 25 |
| XGBoost | Estimators: 50 Max depth: 6 Min child weight: 0.5 Eta: 0.2 |
| ARIMA | P-autoregressive term: 2 Q-moving average term: 2 D-differential times: 0 |
| SVR | Kernel: RBF Degree: 3 C: 100 Gamma: 1 |
| LSTM-TCN | Filters: 128 Kernel size: 3 Hidden layer: 1 Learning rate: 0.002 Dilation rate: 1, 2, 4 LSTM units: 64, 128, 256 |

TABLE 4: Prediction evaluation at different iteration times.

| Index | Algorithm | epoch10 | epoch15 | epoch20 | epoch25 | epoch50 | epoch75 | epoch100 |
|---------------|------------|---------|---------|---------|---------|---------|---------|----------|
| RMSE | EEMD-ELSTM | 1931.74 | 1165.97 | 909.42 | 841.73 | 575.16 | 550.01 | 556.59 |
| | EEMD-LSTM | 1951.56 | 1634.48 | 1184.47 | 1067.86 | 651.13 | 609.33 | 630.51 |
| | ELSTM | 5438.14 | 5367.18 | 5295.41 | 4956.89 | 4632.53 | 4376.05 | 4533.58 |
| | LSTM | 5582.33 | 5582.28 | 5301.50 | 5009.36 | 4739.33 | 4497.84 | 4634.36 |
| MAE | EEMD-ELSTM | 1650.04 | 942.45 | 702.52 | 640.11 | 449.06 | 411.75 | 439.26 |
| | EEMD-LSTM | 1654.95 | 1392.24 | 962.65 | 863.52 | 528.75 | 480.41 | 497.95 |
| | ELSTM | 4598.59 | 4565.94 | 4541.52 | 4372.06 | 4222.25 | 3909.26 | 4286.08 |
| | LSTM | 4662.28 | 4660.97 | 4542.76 | 4433.95 | 4366.71 | 4189.13 | 4325.04 |
| MAPE (%) | EEMD-ELSTM | 4.77 | 2.77 | 2.03 | 1.85 | 1.29 | 1.22 | 1.31 |
| | EEMD-LSTM | 4.81 | 4.01 | 2.81 | 2.55 | 1.55 | 1.54 | 1.55 |
| | ELSTM | 12.66 | 12.60 | 12.51 | 12.14 | 12.01 | 11.80 | 11.92 |
| | LSTM | 12.75 | 12.72 | 12.53 | 12.33 | 12.45 | 11.89 | 12.32 |
| CORR(Pearson) | EEMD-ELSTM | 0.96 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| | EEMD-LSTM | 0.96 | 0.97 | 0.98 | 0.99 | 0.99 | 0.99 | 0.99 |
| | ELSTM | 0.91 | 0.91 | 0.92 | 0.94 | 0.96 | 0.97 | 0.97 |
| | LSTM | 0.88 | 0.88 | 0.91 | 0.94 | 0.96 | 0.97 | 0.97 |

4.4. Results and Discussion. From the experimental validation results in Figure 7, the prediction results show that the EEMD-ELSTM model gives the regression closer to the truth data than the LSTM model. Combined with the comparison of the iteration results in Table 4, it is claimed that the proposed approach outperforms the LSTM model in terms of prediction accuracy.

Figure 8 shows the comparison with the statistical method ARIMA, machine learning method SVM, deep learning method LSTM, Prophet, and decision tree approach XGBoost models.

Among the five models, it can be concluded that the prediction results of EEMD-ELSTM are the closest to the truth. The prediction error of the EEMD-ELSTM model in Table 5 is also at the smallest level. Six models are compared in terms of RMSE, MAE, and MAPE in Table 5.

In order to further discuss the influence of LSTM network parameters on the model results, we adjust the Batch size and the number of neurons to verify the effectiveness of the model, and in order to prove that the baseline model LSTM has certain advantages over the traditional ARIMA model, we add the following experiments, which are

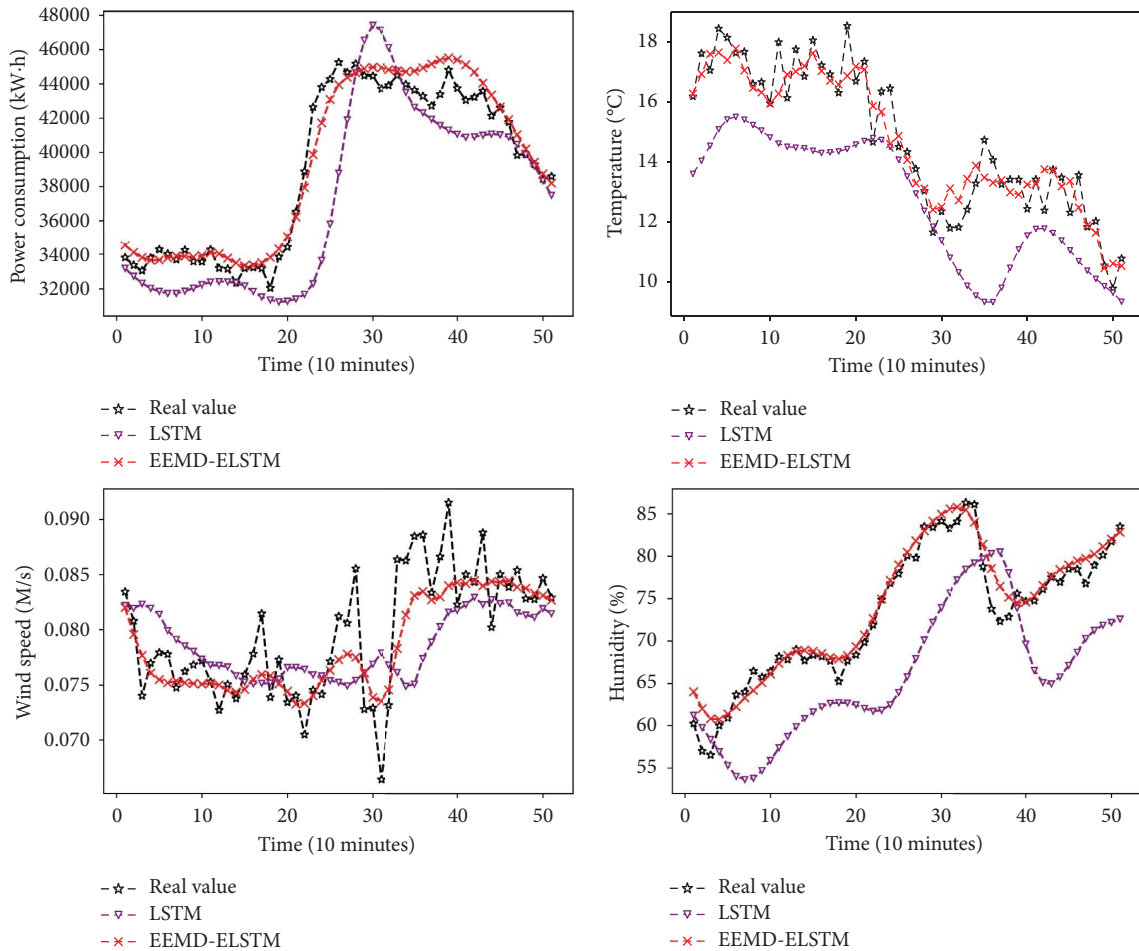


FIGURE 7: Prediction results of EEMD-ELSTM and LSTM models.

performed when the Batch size=5 and the number of neurons is 50. In addition, we have added a set of SOTA from 2023 as a comparison model of LSTM-TCN [47].

From Figure 9 and Table 6, through the comparison of indicators, it is not difficult to find that the LSTM-TCN model has certain advantages over XGBoost in the performance of periodically significant time-series data prediction ability, and after parameter adjustment, the LSTM model has certain advantages over the traditional ARIMA model, but it can also be seen that the baseline model always has some disadvantages over the fusion model and ensemble algorithm in terms of the versatility and validity of different data.

It can be seen from Table 5 that the EEMD-ELSTM model outperforms other compared models on all datasets in terms of accuracy. Moreover, compared with other models, the EEMD-ELSTM method has good universality. The best evaluation values were obtained on four different datasets. At the same time, EEMD-ELSTM also has certain advantages over the evaluation value of the current advanced algorithm XGBoost model.

From the comparison of evaluation indicators in Tables 5 and 6, it can be seen that compared with the LSTM-TCN model, the proposed model has better applicability in the practical application of CPS. In addition, the results of the

ARIMA model are not reliable in the presence of complex stationarity for the data.

The visual comparison is presented in histograms in Figure 10. The information is formed into a histogram to more effectively represent the prediction outcomes of the approaches in Tables 5 and 6. It can be seen from the histogram that the EEMD-ELSTM method gains the most optimized values of RMSE, MAE, and MAPE.

4.5. Friedman Validation and Post Hoc Nemenyi Test. We conducted the Friedman test [48] to investigate the advantages and disadvantages of the proposed model. Table 7 presents the Friedman validation result for the six models on four different datasets, and the result of Friedman test verification is $\text{stat} = 22.691$, $p = 0.001$.

On the basis of the order value in Table 7, we calculate the critical difference by the post hoc Nemenyi algorithm and we get the result in Figure 11.

Each algorithm's mean value is different, and the EEMD-ELSTM model has the lowest mean value. The mean values of other methods' are the XGBoost decision tree approach, ARIMA method (statistical methods), LSTM model and LSTM-TCN model, Prophet model (deep learning method), and the SVR model (machine learning method) from low to

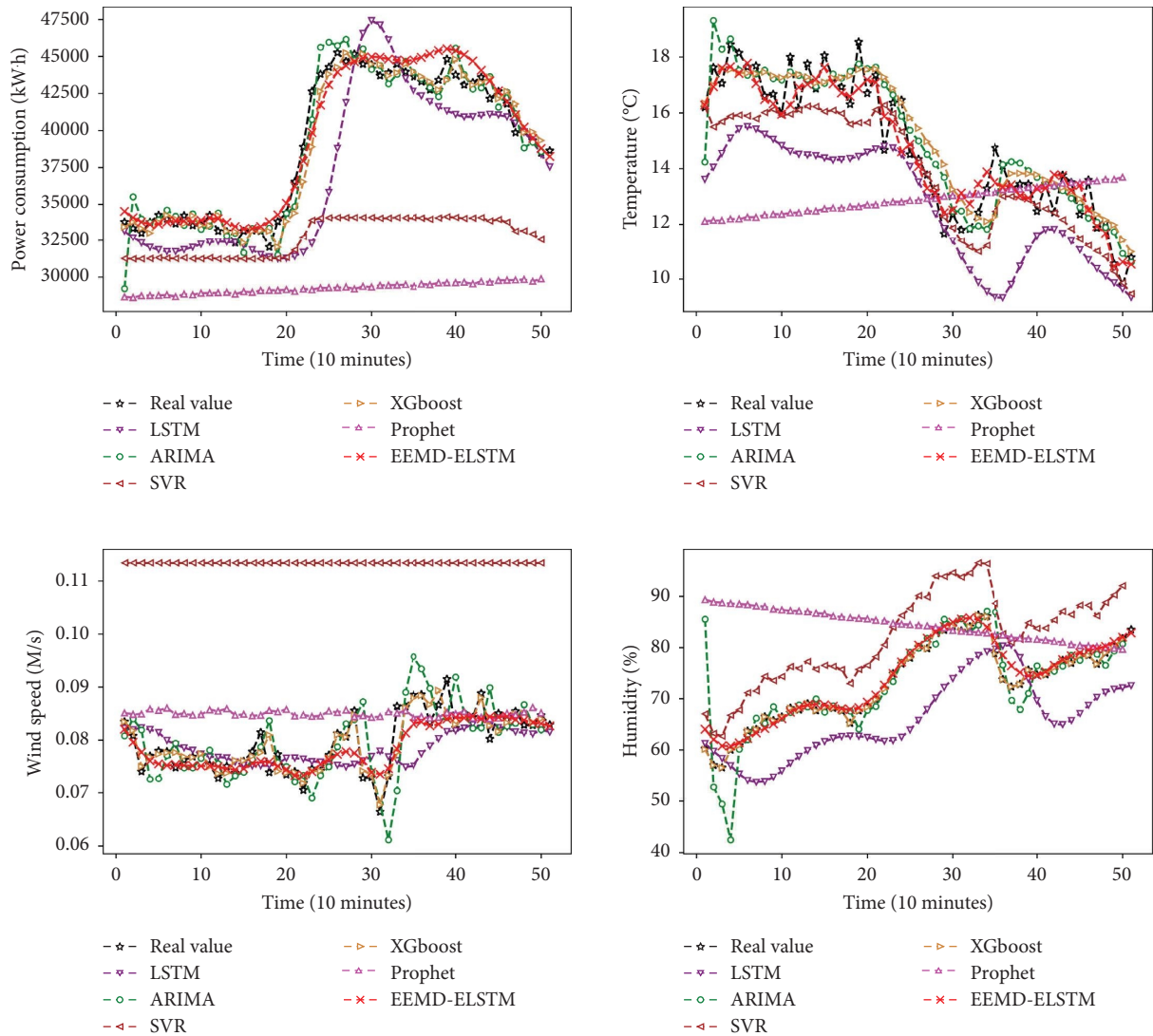


FIGURE 8: Comprehensive Comparison Experiment of EEMD-ELSTM Algorithm Part 1.

TABLE 5: Comparative experimental results of different models.

| Index | Model | Power consumption (kW · h) | Temperature (°C) | Wind speed (M/s) | Humidity (%) |
|-------|------------|----------------------------|------------------|------------------|--------------|
| RMSE | EEMD-ELSTM | 575.15 | 0.51 | 0.0038 | 1.29 |
| | LSTM | 4739.34 | 4.34 | 0.0068 | 7.85 |
| | ARIMA | 1217.73 | 0.72 | 0.0054 | 4.98 |
| | SVR | 7206.11 | 3.42 | 0.0344 | 8.91 |
| | XGBoost | 866.60 | 0.55 | 0.0051 | 1.91 |
| | Prophet | 10793.95 | 6.67 | 0.0079 | 15.25 |
| MAE | EEMD-ELSTM | 449.06 | 0.35 | 0.0026 | 0.99 |
| | LSTM | 4366.71 | 4.43 | 0.0042 | 7.38 |
| | ARIMA | 928.45 | 0.49 | 0.0041 | 2.62 |
| | SVR | 6270.58 | 3.32 | 0.0339 | 8.83 |
| | XGBoost | 666.23 | 0.38 | 0.0032 | 1.43 |
| | Prophet | 9777.84 | 6.08 | 0.0065 | 11.85 |

TABLE 5: Continued.

| Index | Model | Power consumption (kW · h) | Temperature (°C) | Wind speed (M/s) | Humidity (%) |
|----------|------------|-------------------------------|------------------|---------------------|--------------|
| MAPE (%) | EEMD-ELSTM | 1.29 | 2.47 | 3.04 | 1.27 |
| | LSTM | 12.44 | 27.41 | 4.73 | 9.08 |
| | ARIMA | 2.42 | 4.92 | 5.04 | 3.89 |
| | SVR | 15.15 | 18.38 | 43.48 | 12.06 |
| | XGBoost | 2.10 | 2.48 | 3.623 | 1.79 |
| | Prophet | 33.31 | 48.03 | 7.68 | 13.71 |

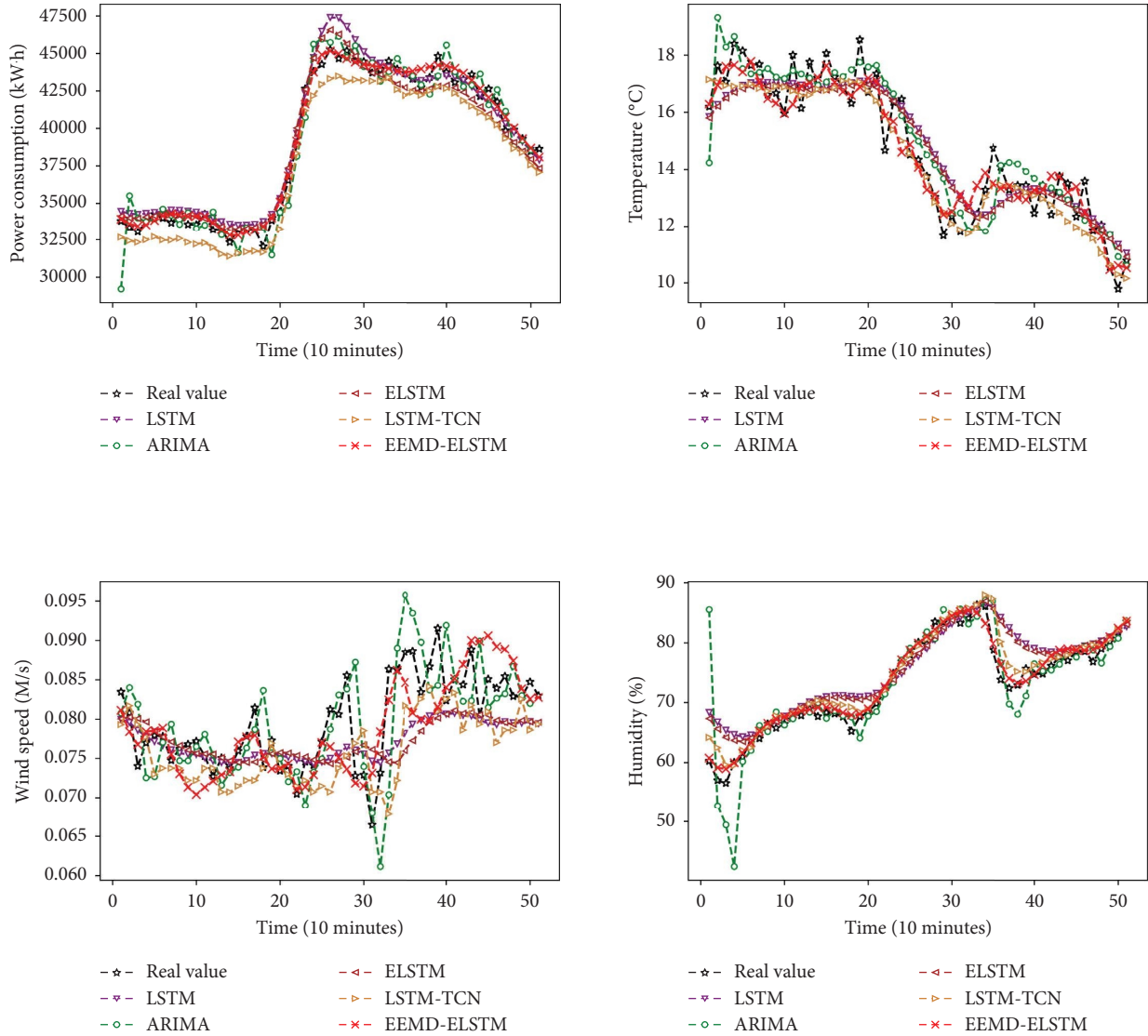


FIGURE 9: Comprehensive Comparison Experiment of EEMD-ELSTM Algorithm Part 2.

TABLE 6: Improved comparative experiment with tuning parameters.

| Index | Model | Power consumption (kW · h) | Temperature (°C) | Wind speed (M/s) | Humidity (%) |
|----------|------------|----------------------------|------------------|------------------|--------------|
| RMSE | EEMD-ELSTM | 501.79 | 0.51 | 0.0040 | 1.08 |
| | LSTM | 1140.19 | 0.72 | 0.0072 | 2.78 |
| | ARIMA | 1217.73 | 0.72 | 0.0054 | 4.98 |
| | ELSTM | 1108.35 | 0.68 | 0.0060 | 2.52 |
| | LSTM-TCN | 680.20 | 0.43 | 0.0052 | 2.08 |
| MAE | EEMD-ELSTM | 399.93 | 0.24 | 0.0028 | 0.87 |
| | LSTM | 758.65 | 0.45 | 0.0042 | 2.05 |
| | ARIMA | 928.45 | 0.49 | 0.0041 | 2.62 |
| | ELSTM | 727.27 | 0.43 | 0.0036 | 1.86 |
| | LSTM-TCN | 513.91 | 0.34 | 0.0032 | 1.51 |
| MAPE (%) | EEMD-ELSTM | 1.27 | 1.61 | 3.58 | 1.16 |
| | LSTM | 2.41 | 4.57 | 5.88 | 2.99 |
| | ARIMA | 2.42 | 4.92 | 5.04 | 3.89 |
| | ELSTM | 2.39 | 4.48 | 5.75 | 2.70 |
| | LSTM-TCN | 1.71 | 3.02 | 4.62 | 2.55 |

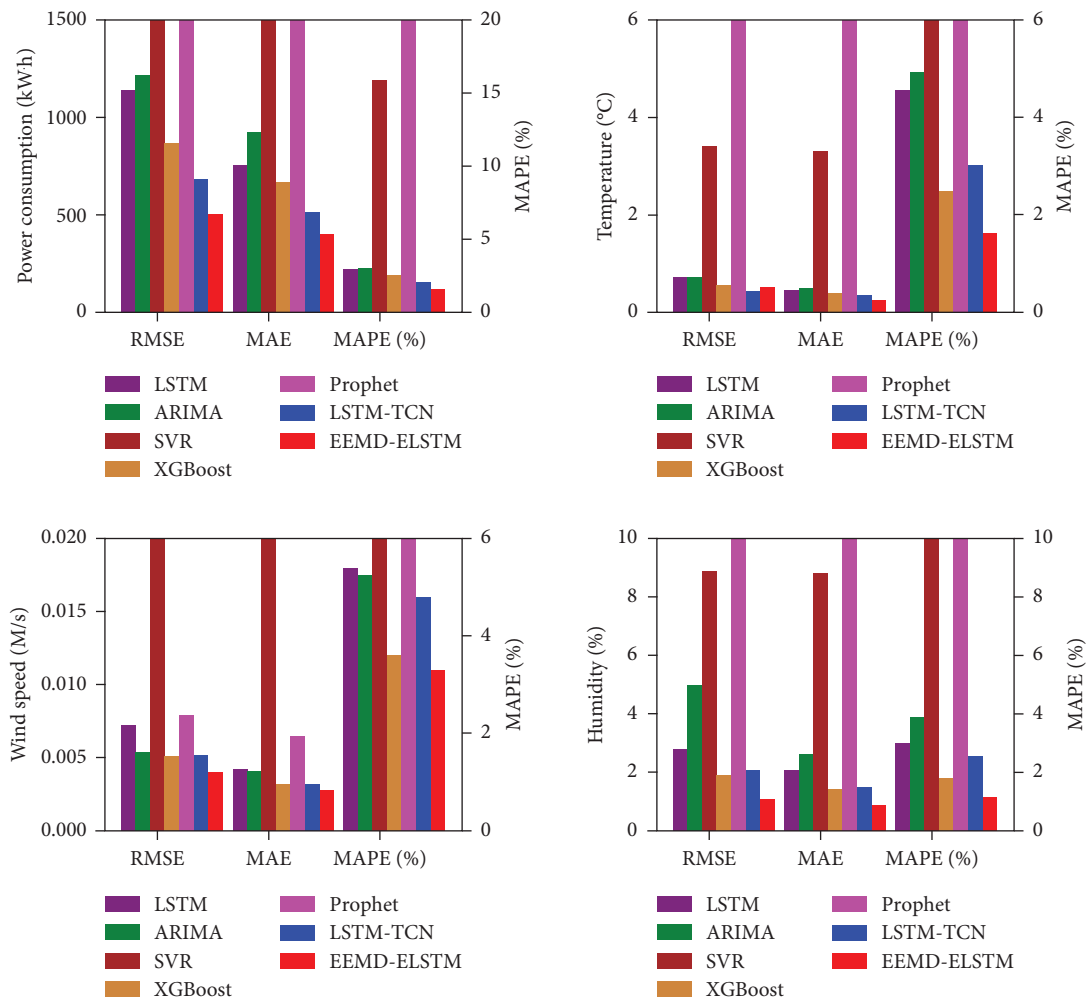


FIGURE 10: Comparison of prediction results of different models.

TABLE 7: Friedman validation results.

| | EEMD-ELSTM | LSTM | ARIMA | SVR | XGBoost | Prophet | LSTM-TCN |
|-------------------|------------|------|-------|------|---------|---------|----------|
| Power consumption | 1 | 4 | 3 | 6 | 3 | 7 | 2 |
| Temperature | 1 | 4 | 3 | 6 | 2 | 7 | 3 |
| Wind speed | 1 | 5 | 4 | 7 | 2 | 6 | 3 |
| Humidity | 1 | 4 | 3 | 6 | 2 | 7 | 3 |
| Mean value | 1 | 4.25 | 4.75 | 6.25 | 2.25 | 6.75 | 2.75 |

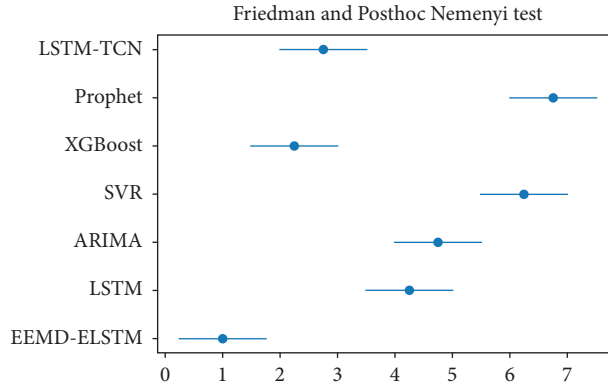


FIGURE 11: Friedman and posthoc Nemenyi test.

high. The critical difference (CD) shows that the proposed algorithm has obvious differences compared with the existing SOTA. These Friedman and post hoc Nemenyi outcomes explain that the EEMD-ELSTM method is more practical and trustworthy than the generic forecasting model.

5. Conclusion

The entropy fusion feature-based hybrid time-series prediction approach for EEMD and LSTM is proposed in this paper. First, a modal decomposition method EEMD based on Gaussian white noise is constructed. Then, the IMF subcolumn components of some input data are obtained by using this algorithm. The IMF subcolumn components generated by decomposition are analyzed by entropy. The component of the IMF subcolumn with the largest entropy value is extracted as the auxiliary dimension, the eigenvector of the ELSTM model is constructed, and the final prediction result is obtained by model prediction and reconstruction. In addition, four time-series stream datasets with varying data sizes and six SOTA prediction models are chosen for comparison in the experimental verification section. The experimental results showed that the proposed model outperforms others in reliability and validity. Moreover, based on the MAPE index, we calculated the optimization percentage of the results under each dataset and different parameter configurations by comparing the results of the baseline model LSTM obtained by the experiment and the results of our improved model, and finally, we improved the MAPE parameters by 66.43% compared with the baseline model.

The abovementioned experimental verification shows that the EEMD-ELSTM model has the following characteristics:

- (1) Validity: from the analysis of evaluation indices in Tables 5 and 6 and Figures 10, and 11, the optimization EEMD-ELSTM method clearly outperforms the generic nonfused network prediction model in forecasting validity.
- (2) Universality: from the evaluation index results and Friedman test results of the abovementioned comparison experiments, it can be clearly known that among the four different types of datasets, the EEMD-ELSTM algorithm always has a good prediction effect, has good universality, and can be applied to datasets of different scales.
- (3) Causality: the effectiveness of the optimization part and the improved strategy mechanism proposed in the article can be seen from the comparison of ablation experiments set in the article. Ablation research is the most direct way to understand the causality in the system and generate reliable predictions. The causality of each part of the model was proved by ablation.

6. Further Study

- (1) Such time complicity of the LSTM method is regarded as generally large, and it depends on the machine computing power, increasing network neurons, and iteration times to improve the operation effect. In future research, it is

considered to further optimize the operation process of relevant machine learning models to improve the model effect or to replace the traditional neural network part of the current common hybrid models. (2) This article only aims at one-dimensional time-series data flow, which has certain limitations. Predictive studies with a wealth of multidimensional data are among the hot studies of the present time series, as a result, we are considering including multidimensional data decomposition prediction as the next topic for research.

Data Availability

The data used to support the findings of the study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest associated with this work or significant financial support that could have appeared to influence its outcome.

Acknowledgments

This work was partially supported by the Yunnan Fundamental Research Key Projects (202101AS070016), the Major Science and Technology Project of Yunnan Province (202302AD080002), the Open Fund of Yunnan Key Laboratory of Computer Technology Application (CB22144S073A), and the Yunnan Province "Xingdian Talents Support Plan" Industrial Innovation Talents Project (Yfgr [2019] no. 1096).

References

- [1] M. Ali, D. M. Khan, M. Aamir, A. Ali, and Z. Ahmad, "Predicting the direction movement of financial time series using artificial neural network and support vector machine," *Complexity*, vol. 2021, Article ID 2906463, 2021.
- [2] X. Kong and C. Luo, "A novel convlstm with multifeature fusion for financial intelligent trading," *International Journal of Intelligent Systems*, vol. 37, no. 11, pp. 8855–8877, 2022.
- [3] Y. Liu, D. Li, S. Wan et al., "A long short-term memory-based model for greenhouse climate prediction," *International Journal of Intelligent Systems*, vol. 37, no. 1, pp. 135–151, 2022.
- [4] A. Hrabia, J. Kozak, and P. Juszczak, "Machine learning in the context of covid-19 pandemic data analysis," in *Advances in Computational Collective Intelligence*, K. Wojtkiewicz, J. Treur, E. Pimenidis, and M. Maleszka, Eds., pp. 361–374, Springer International Publishing, Cham, 2021.
- [5] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.
- [6] Y. Lu, B. Sheng, G. Fu, R. Luo, G. Chen, and Y. Huang, "Prophet-eemd-lstm based method for predicting energy consumption in the paint workshop," *Applied Soft Computing*, vol. 143, Article ID 110447, 2023.
- [7] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2014.
- [8] Z. Chen, Z. Peng, X. Zou, and H. Sun, "Deep learning based anomaly detection for multi-dimensional time series: a survey," in *Cyber Security*, W. Lu, Y. Zhang, W. Wen, H. Yan, and C. Li, Eds., pp. 71–92, Springer Nature, Singapore, 2022.
- [9] W. Zhao, H. Yang, J. Li, L. Shang, L. Hu, and Q. Fu, "Network traffic prediction in network security based on emd and lstm," in *Proceedings of the 9th International Conference on Computer Engineering and Networks*, Q. Liu, X. Liu, L. Li, H. Zhou, and H.-H. Zhao, Eds., pp. 509–518, Springer, Singapore, 2021.
- [10] H. Niu, K. Xu, and W. Wang, "A hybrid stock price index forecasting model based on variational mode decomposition and lstm network," *Applied Intelligence*, vol. 50, no. 12, pp. 4296–4309, 2020.
- [11] Y.-Q. Gong, "The volatility and cycle of emerging industry stock market in China: an empirical study based on eemd," in *Proceedings of 20th International Conference on Industrial Engineering and Engineering Management*, E. Qi, J. Shen, and R. Dou, Eds., pp. 279–289, Springer, Berlin, Heidelberg, 2013.
- [12] P. Dong, X. Bin, M. Jun et al., "Short-term load forecasting based on eemd-approximate entropy and elm," in *Proceedings of the 2019 IEEE Sustainable Power and Energy Conference*, pp. 1772–1775, Beijing, China, November 2019.
- [13] M. H. D. M. Ribeiro, R. G. da Silva, J. H. K. Larcher, V. C. Mariani, and L. S. Coelho, *Ensemble Learning Models Coupled with Urban Mobility Information Applied to Predict COVID-19 Incidence Cases*, Springer International Publishing, Cham, 2022.
- [14] B. C. Kovoov, "Eemd-based wind speed forecasting system using bidirectional lstm networks," in *Proceedings of the 2021 International Conference on Computer Communication and Informatics*, Coimbatore, India, January 2021.
- [15] Y. Yang, C. Fan, and H. Xiong, "A novel general-purpose hybrid model for time series forecasting," *Applied Intelligence*, vol. 52, no. 2, pp. 2212–2223, 2022.
- [16] F. Riaz, A. Hassan, S. Rehman, I. K. Niazi, and K. Dremstrup, "Emd-based temporal and spectral features for the classification of eeg signals using supervised learning," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 1, pp. 28–35, 2016.
- [17] D. Prajapati and M. Kanojia, "Forecasting of covid-19 cases in India using arima and ar time-series algorithm," in *Proceedings of the 13th International Conference on Soft Computing and Pattern Recognition (SoCPaR 2021)*, A. Abraham, A. Engelbrecht, F. Scotti et al., Eds., pp. 361–370, Springer International Publishing, Cham, 2022.
- [18] S. Behzadi, B. Schelling, and C. Plant, "Itgh: information-theoretic granger causal inference on heterogeneous data," in *Advances in Knowledge Discovery and Data Mining*, H. W. Lauw, R. C.-W. Wong, A. Ntoulas, E.-P. Lim, S.-K. Ng, and S. J. Pan, Eds., pp. 742–755, Springer International Publishing, Cham, 2020.
- [19] B. Kap, M. Aleksandrova, and T. Engel, "The effect of noise level on the accuracy of causal discovery methods with additive noise models," in *Artificial Intelligence and Machine Learning*, L. A. Leiva, C. Pruski, R. Markovich, A. Najjar, and C. Schommer, Eds., pp. 120–140, Springer International Publishing, Cham, 2022.
- [20] A. L. M. Hanapi, M. Othman, R. Sokkalingam, and H. Sakidin, "A comparison study on developed fswgarch, swgarch and garch models in time series forecasting: an application to airline passenger volume," in *Proceedings of the 6th International Conference on Fundamental and Applied Sciences*, S. A. Abdul Karim, M. F. Abd Shukur, C. Fai Kait,

- H. Soleimani, and H. Sakidin, Eds., pp. 623–636, Springer Nature, Singapore, 2021.
- [21] U. K. Yusof, M. N. A. Khalid, A. Hussain, and H. Shamsudin, “Financial time series forecasting using prophet,” in *Innovative Systems for Intelligent Health Informatics*, F. Saeed, F. Mohammed, and A. Al-Nahari, Eds., pp. 485–495, Springer International Publishing, Cham, 2021.
- [22] S. Behzadi, K. Hlaváčková-Schindler, and C. Plant, “Granger causality for heterogeneous processes,” in *Advances in Knowledge Discovery and Data Mining*, Q. Yang, Z.-H. Zhou, Z. Gong, M.-L. Zhang, and S.-J. Huang, Eds., pp. 463–475, Springer International Publishing, Cham, 2019.
- [23] L. Huang, B. Jiang, S. Lv, Y. Liu, and D. Li, “Research review of recommendation system based,” *Journal of Computer Science*, vol. 41, no. 7, p. 29, 2018.
- [24] Q. Zeng, Y. Liang, G. Chen, H. Duan, and C. Li, “Noise prediction of chemical industry park based on multi-station prophet and multivariate lstm fitting model,” *EURASIP Journal on Applied Signal Processing*, vol. 2021, no. 1, p. 106, 2021.
- [25] Y. Li, Z. Ma, Z. Pan, N. Liu, and X. You, “Prophet model and Gaussian process regression based user traffic prediction in wireless networks,” *Science China Information Sciences*, vol. 63, no. 4, Article ID 142301, 2020.
- [26] Y. Zhang, B. Chen, G. Pan, and Y. Zhao, “A novel hybrid model based on vmd-wt and pca-bp-rbf neural network for short-term wind speed forecasting,” *Energy Conversion and Management*, vol. 195, pp. 180–197, 2019.
- [27] Q. Gong, W. Lu, W. Gong, and X. Wang, “Short-term load forecasting of lssvm based on improved pso algorithm,” in *Pattern Recognition*, S. Li, C. Liu, and Y. Wang, Eds., pp. 63–71, Springer, Berlin, Heidelberg, 2014.
- [28] O. A. Montesinos López, A. Montesinos López, and J. Crossa, *Support Vector Machines and Support Vector Regression*, Springer International Publishing, Cham, 2022.
- [29] E. Pekel, “Estimation of soil moisture using decision tree regression,” *Theoretical and Applied Climatology*, vol. 139, no. 3–4, pp. 1111–1119, 2020.
- [30] E. Jumin, F. B. Basaruddin, Y. B. M. Yusoff, S. D. Latif, and A. N. Ahmed, “Solar radiation prediction using boosted decision tree regression model: a case study in Malaysia,” *Environmental Science and Pollution Research*, vol. 28, no. 21, pp. 26571–26583, 2021.
- [31] Y. Qiu, J. Zhou, M. Khandelwal, H. Yang, P. Yang, and C. Li, “Performance evaluation of hybrid woa-xgboost, gwo-xgboost and bo-xgboost models to predict blast-induced ground vibration,” *Engineering with Computers*, vol. 38, no. 5, pp. 4145–4162, 2021.
- [32] Q. Yang, D. Liu, Y. Fang, D. Yang, Y. Zhou, and Z. Sheng, “Research on a hybrid emd-svr model for time series prediction,” in *Security and Privacy in New Computing Environments*, D. Wang, W. Meng, and J. Han, Eds., pp. 137–150, Springer International Publishing, Cham, 2021.
- [33] S. K. Panda, P. Ray, and D. P. Mishra, “Short term load forecasting using empirical mode decomposition (emd), particle swarm optimization (pso) and adaptive network-based fuzzy interference systems (anfis),” in *Innovations in Bio-Inspired Computing and Applications*, A. Abraham, M. Panda, S. Pradhan, L. Garcia-Hernandez, and K. Ma, Eds., pp. 161–168, Springer International Publishing, Cham, 2021.
- [34] X. Zhang, D. Zhao, B. Duan, and W. Qiao, “A novel groundwater burial depth prediction model—based on the combined vmd-wsd-elman model,” *Environmental Science and Pollution Research*, vol. 29, no. 50, pp. 76310–76320, 2022.
- [35] Q. Zhang, G. Yang, and E. Yuan, “Pm2.5 spatial-temporal long series forecasting based on deep learning and emd,” in *Knowledge and Systems Sciences*, J. Chen, T. Hashimoto, X. Tang, and J. Wu, Eds., pp. 3–19, Springer Nature, Singapore, 2022.
- [36] M. Ali, D. M. Khan, H. M. Alshambari, and A. A.-A. H. El-Bagoury, “Prediction of complex stock market data using an improved hybrid emd-lstm model,” *Applied Sciences*, vol. 13, no. 3, p. 1429, 2023.
- [37] M. M. Dedovic, S. Avdaković, A. Mujezinović, and N. Dautbasic, “The Hybrid EMD-SARIMA Model for Air Quality Index Prediction,” *Case of Canton Sarajevo*, Springer International Publishing, Cham, pp. 139–150, 2021.
- [38] H. Liu, S. Yin, C. Chen, and Z. Duan, “Data multi-scale decomposition strategies for air pollution forecasting: a comprehensive review,” *Journal of Cleaner Production*, vol. 277, Article ID 124023, 2020.
- [39] V. C. Lobo Neto, L. A. Passos, and J. P. Papa, “Evolving long short-term memory networks,” in *Computational Science-ICCS 2020*, V. V. Krzhizhanovskaya, G. Závodszyk, M. H. Lees et al., Eds., pp. 337–350, Springer International Publishing, Cham, 2020.
- [40] A. Zülfiqar and A. Mehmet, “Empirical mode decomposition and convolutional neural network-based approach for diagnosing psychotic disorders from eeg signals,” *Applied Intelligence*, vol. 52, 2022.
- [41] H. Qin, M. Liu, J. Wang, Z. Guo, and J. Liu, “Adaptive diagnosis of dc motors using r-wdcnn classifiers based on vmd-svd,” *Applied Intelligence*, vol. 51, no. 7, pp. 4888–4907, 2021.
- [42] B. Huang, Y. Zhu, Y. Gao et al., “The analysis of isolation measures for epidemic control of covid-19,” *Applied Intelligence*, vol. 51, no. 5, pp. 3074–3085, 2021.
- [43] H. Liu, Y. Lin, X. Liu, J. Zhao, and H. Ruan, “Dual-stage time series analysis on multifeature adaptive frequency domain modeling,” *International Journal of Intelligent Systems*, vol. 37, no. 10, pp. 7837–7856, 2022.
- [44] L. Song, L. Ding, T. Wen, M. Yin, and Z. Zeng, “Time series change detection using reservoir computing networks for remote sensing data,” *International Journal of Intelligent Systems*, vol. 37, no. 12, pp. 10845–10860, 2022.
- [45] M. Ali, D. M. Khan, I. Saeed, and H. M. Alshambari, “A new approach to empirical mode decomposition based on akima spline interpolation technique,” *IEEE Access*, vol. 11, pp. 67370–67384, 2023.
- [46] L. An, Y. Hao, T.-C. J. Yeh, Y. Liu, W. Liu, and B. Zhang, “Simulation of karst spring discharge using a combination of time–frequency analysis methods and long short-term memory neural networks,” *Journal of Hydrology*, vol. 589, Article ID 125320, 2020.
- [47] T. Limouni, R. Yaagoubi, K. Bouziane, K. Guissi, and E. H. Baali, “Accurate one step and multistep forecasting of very short-term pv power using lstm-tcn model,” *Renewable Energy*, vol. 205, pp. 1010–1024, 2023.
- [48] B. Vidya and P. Sasikumar, “Parkinson’s disease diagnosis and stage prediction based on gait signal analysis using emd and cnn-lstm network,” *Engineering Applications of Artificial Intelligence*, vol. 114, Article ID 105099, 2022.