Hindawi
Scientific Programming
Volume 2022, Article ID 4755519, 8 pages
https://doi.org/10.1155/2022/4755519

Research Article

Short-Term Power Load Forecasting Model Design Based on EMD-PSO-GRU

Weihao Zhang[1] **and Ting Wang**[2]

[1]*School of International Education, Henan University, Zhengzhou 65292, Henan, China*
[2]*State Grid Hubei Electric Power Co, Ltd, Jingzhou Power Supply Company Power Dispatching Control Center, Jingzhou, Hubei, China*

Correspondence should be addressed to Weihao Zhang; 1924030003@henu.edu.cn

Received 7 July 2022; Revised 31 July 2022; Accepted 3 August 2022; Published 26 August 2022

Academic Editor: Lianhui Li

Aiming at the nonlinear, nonstationary, and time series characteristics of power load, this study proposes a load forecasting method based on empirical mode decomposition and particle swarm optimization of the gated recurrent unit neural network. First, the original power load data are decomposed into a limited number of modal components and a residual component by using empirical modal decomposition to reduce the nonstationarity and complexity of the load sequence and decrease the association between different IMFs. The subsequences build prediction models based on the gated recurrent unit neural network, respectively, and use the particle swarm algorithm to optimize the network-related hyperparameters to increase the parameter accuracy of the model; finally, superimpose the prediction results of each subsequence to obtain the final load prediction value. The results of the case study show that compared with the traditional forecasting algorithm, the proposed EMD-PSO-GRU forecasting model method can better dig the trend information of forecasting, fit the load curve better, and have higher forecasting accuracy.

1. Introduction

Accurate load forecasting is an important guarantee for stable operation of power grids, scheduling optimization, and reducing operating costs [1]. Smart grid provides a high-quality and massive database for load forecasting. With the rapid evolution of the energy Internet [2], it is more urgent and important to study algorithms with the ability to process big data and high forecasting accuracy. The accuracy of the model has important significance and high engineering application value [3, 4].

In terms of load forecasting, traditional forecasting models such as autoregression (AR), although fast in operation, have high data requirements and lack the ability to adapt and predict. The robustness is poor, and it is difficult to meet the requirements of load forecasting [5]. In recent years, the development of artificial intelligence technology has provided ideas for solving these problems, but some new problems are still derived. Xiangyu et al. [6] used a deep belief network to quickly analyze complex influencing factors, which improves the prediction accuracy, but it only targets regional loads and lacks adaptability. Wu et al. [7] used a parallelized multicore support vector machine (SVM) for load prediction, and the prediction error is reduced to a certain extent compared with a single-core SVM, but it lacks the consideration of the correlation between time series data. Shi and Zhang [8] considered the training differences of different algorithms and proposed a stacking load forecasting model embedded with various machine learning algorithms. This model ensures good accuracy in forecasting, but the cost of model integration is too high and the time is long. With the growing development of deep learning, the gated recurrent unit network as a kind of the special RNN model [9, 10] is widely used to predict events in time series due to the introduction of modules with "memory function" in its structure [11], but it has two deficiencies, namely, the

model learning rate and the number of neurons in the hidden layer, which are difficult to determine. Among them, the learning rate determines the training effect of the model, and the number of neurons in the hidden layer affects the fitting effect of the model. Usually, these parameters are determined by experience and are uncertain, which leads to a decrease in the accuracy of the model.

Similar to other neural networks, GRU's model parameters often need to be selected by human experience, and the fitting ability, training speed, and prediction effect of different model parameters are quite different. In order to more reasonably determine the model parameters of GRU and improve the stability of power load forecasting, a method is proposed based on EMD-PSO to optimize GRU network hyperparameters. Through example analysis, the prediction results verify the effectiveness of the proposed model.

2. Model Theory

2.1. Empirical Mode Decomposition. EMD is a decomposition method, which can decompose the signal according to the specific time scale features of the data and adaptively decompose the local feature signals on different time scales to obtain IMF and residual signals with different characteristics. Each IMF component obtained by decomposition has a certain physical meaning. It represents the variable components of various time scales contained in the raw load data, and the residual term represents the basic trend of the load sequence. The specific EMD steps for a given original time series $x(t)$ are as follows [12, 13]:

(1) Identify maximum and minimum points in the original sequence $x(t)$, use the cubic spline interpolation method to respectively fit $x_{up}(t)$ and $x_{low}(t)$, and calculate the mean $m(t)$ of the upper and lower envelopes.

$$m(t) = \frac{\left[x_{\text{up}}(t) + x_{\text{low}}(t)\right]}{2}. \tag{1}$$

(2) Calculate the difference between the original sequence $x(t)$ and the envelope mean $m(t)$, denoted as $h(t)$.

$$h(t) = x(t) - m(t). \tag{2}$$

(3) Determine whether $h(t)$ satisfies the IMF constraints; if not, use it as a new input sequence, and repeat steps (1) to (2) until the constraints are met; if so, $h(t)$ is the first IMF component, denoted as $c_1(t) = h(t)$, and separates $c_1(t)$ from the original sequence $x(t)$ to obtain $r_1(t)$.

$$r_1(t) = x(t) - c_1(t). \tag{3}$$

(4) $r_1(t)$ is regarded as a new original sequence, and the abovementioned smoothing steps are repeated to obtain the remaining IMF components and one remaining component. The final result of EMD can be expressed as

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t). \tag{4}$$

In the formula: $c_i(t)$ is the i-th intrinsic mode function component; $r_n(t)$ is the residual component, representing the trend term of the original series.

By using the empirical mode decomposition method, different components can be decomposed from the load time series to form a series of subsequence components [14]. Compared with the original series, the subsequence has stronger stationarity to improve accuracy.

2.2. Gated Recurrent Unit Neural Network. The GRU network optimizes LSTM's three gate functions, integrates oblivion gates and input gates into one update gate [15–17], and mixes neuronal and hidden states simultaneously. This can effectively mitigate the following problems: GRU addresses the "vanishing gradient" of RNN networks and reduces the training time of the model. GRU network basic structure is shown in Figure 1.

The calculation formula of the internal structure is

$$\begin{aligned} r_t &= \sigma\left(W_{\text{r}} \cdot \left[h_{t-1}, x_t\right]\right), \\ z_t &= \sigma\left(W_{\text{z}} \cdot \left[h_{t-1}, x_t\right]\right), \\ \widetilde{h}_t &= \phi\left(W_{\widetilde{h}} \cdot \left[r_t \times h_{t-1}, x_t\right]\right), \\ h_t &= \left(\text{I} - z_t\right) \times h_{t-1} + z_t \times \widetilde{h}_t, \\ y_t &= \sigma\left(W_{\text{o}} \cdot h_t\right). \end{aligned} \tag{5}$$

x_t, h_{t-1}, r_t, z_t, $\widetilde{h}_t$, and y_t, respectively, represent the input vector, the state memory variable at the previous moment, the state of the reset gate, the state of the update gate, e-th state of the current candidate set, and the output vector; W_{r}, W_{z}, $W_{\widetilde{h}}$, and W_{o} are the weight parameters; I represents identity matrix; [ ] represents vector connection; · represents matrix dot product; × represents matrix product; σ represents sigmoid activation function; and ϕ represents tanh activation function. The mathematical description of σ and ϕ is as follows:

$$\begin{aligned} \sigma(x) &= \frac{1}{1 + \text{e}^{-x}}, \\ \phi(x) &= \frac{\text{e}^{x} - \text{e}^{-x}}{\text{e}^{x} + \text{e}^{-x}}. \end{aligned} \tag{6}$$

GRU networks use update gates and reset gates as core modules. x_t and h_{t-1} at the previous moment is input into the update gate after the sigmoid nonlinear transformation to determine the state of the previous moment. The reset gate controls the amount of information written to the candidate set, stores the information at the previous moment through h_{t-1} times $\text{I} - z_t$, and records the information at the current moment through $\widetilde{h}_t$ times z_t.

2.3. Particle Swarm Optimization Algorithm. In PSO, each particle has an initial velocity and position, and the fitness value of the particle is determined by the fitness function

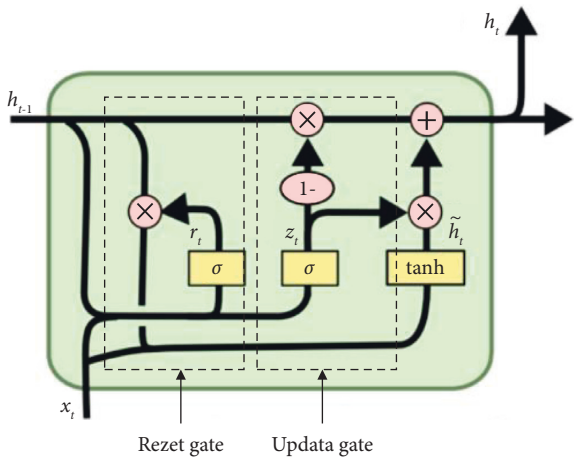

Figure 1: GRU network structure.

[18]. In the iteration, each particle can store the searched optimal position, and its velocity determines the direction and distance of flight [19]. The particles update their speed and position by comparing the fitness values, the optimal value detected by the particle itself is the individual extremum, and the optimal solution detected by the entire population is the global extremum [20]:

$$\begin{aligned} v_i(t+1) &= \omega v_i(t) + c_1 R_1 \left[R_i^b(t) - x_i(t)\right] \\ &\quad + c_2 R_2 \left[R_g^b(t) - x_i(t)\right], \\ x_i(t+1) &= x_i(t) + v_i(t+1). \end{aligned} \tag{7}$$

In the formula, t is the number of iterations; $v_i(t)$ is the speed of the i-th particle in t iterations; ω is the inertia weight; c_1, c_2 are the cognitive coefficients; R_1, R_2 are uniformly distributed random numbers; $Rb/i(t)$ is the individual historical optimal position of a particle i; $Rb/g(t)$ is the historical optimal position of the group; $x_i(t)$ is the position of the particle in t iterations.

The inertia weight ω is an important parameter of PSO. The larger the weight, the stronger the global search ability of the algorithm; the smaller the weight, the stronger the local search ability of the algorithm [21, 22]. The dynamic adjustment of the inertial weight is adopted.

$$\omega(n) = \omega_{\max} - (\omega_{\max} - \omega_{\min})\left(\frac{n}{n_{\max}}\right), \tag{8}$$

where ω_{max} represents the maximum weight, ω_{min} represents the minimum weight, and n_{max} is the maximum number of iterations.

3. EMD-PSO Optimizes the Prediction Model of GRU Hyperparameters

3.1. PSO Optimizes the GRU Hyperparameter Model. The PSO algorithm is a random search and parallel optimization algorithm, which has the characteristics of simplicity, good robustness, and fast convergence speed, and has a high probability of finding the global optimal solution to the problem [23]. The algorithm optimizes the hyperparameters of the GRU and establishes a load forecasting model with higher accuracy. In the GRU model, two hyperparameters have a positive effect on the prediction performance of the model, namely, the number of GRU neurons and the learning rate. Taking these two key parameters as the characteristics of particle optimization, the PSO algorithm is used to adjust and optimize the GRU model. The flowchart of the PSO-GRU load forecasting model is shown in Figure 2.

3.2. EMD-PSO-GRU (EPG) Load Forecasting Model. The load sequence is complex and nonstationary. The decomposed components are input to the GRU model and the PSO algorithm is used to optimize the GRU model hyperparameters. Finally, superimpose the prediction results of each component to get the final prediction value. To improve prediction accuracy, this study proposes the EMD-PSO-GRU hybrid prediction model architecture shown in Figure 3.

The specific steps for building a model are as follows.

(1) The original load time series is divided into several subsequences by EMD

(2) Divide each subsequence into a training set and a validation set and normalize the dataset. Since the GRU model is highly sensitive to the data scale, the original data have a large difference in the order of magnitude; in order to avoid that the change of larger value will cover the change in smaller value, it is necessary to constrain the input data to a similar order of magnitude to avoid affecting the effect of power load forecasting due to the large individual input value. The dataset is between [0, 1] to reinforce the convergence speed of GRU.

(3) Learn the complex relationship between input and output variables in each subsequence by building a GRU network and through the PSO algorithm to determine the optimal GRU network hyperparameters

(4) Use the trained GRU network to predict the subsequence and deformalize the prediction result to obtain the real prediction value

(5) Add the prediction results of each subsequence to obtain the final result of the load

3.3. Model Evaluation Metrics. To evaluate the performance of the prediction model, MAPE and RMSE are used as evaluation indicators [24–33], which are defined as follows:

$$\begin{aligned} y_{MAPE} &= \frac{1}{n}\sum_{i=1}^{n}\frac{\left|y_i - \hat{y}_i\right|}{y_i} \times 100\%, \\ y_{RMSE} &= \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{n}}, \end{aligned} \tag{9}$$

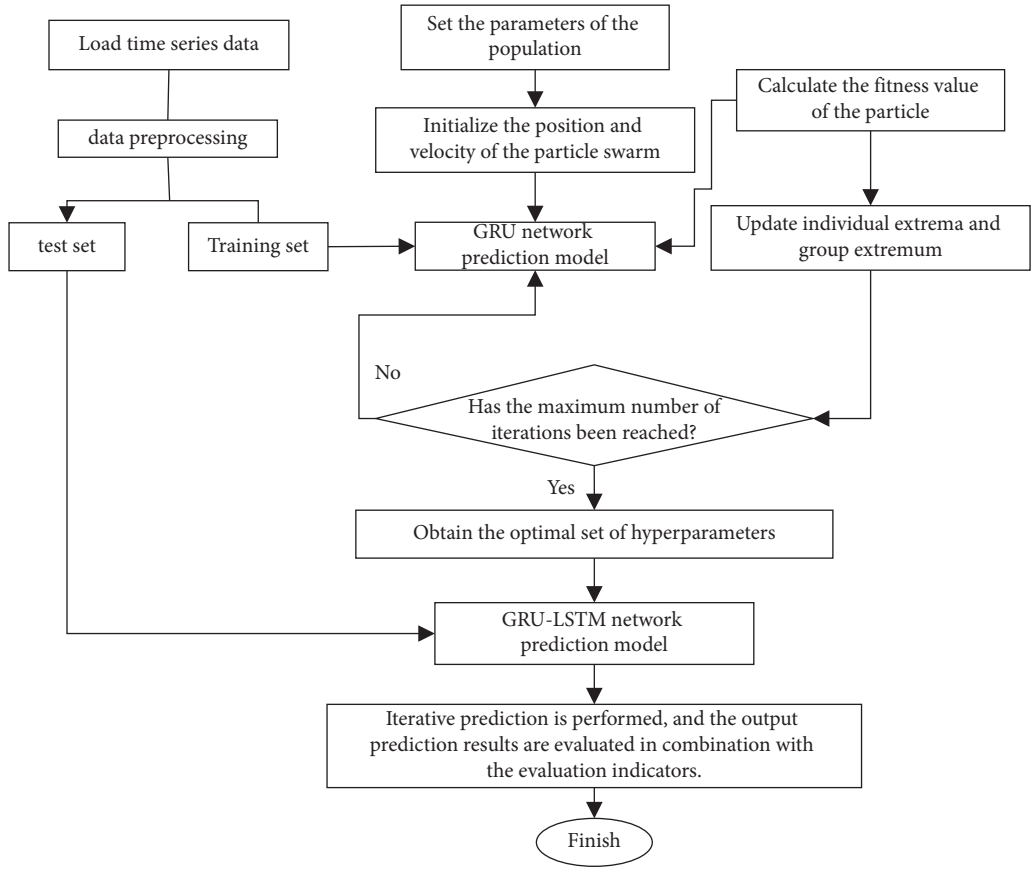

Figure 2: Flowchart of the PSO-LSTM load prediction model.

where y_i is the actual load value on the i-th day; $\hat{y}_i$ is the load forecast value on the i-th day; n is the number of samples in the test set.

4. EMD-PSO Optimizes the Prediction Model of GRU Hyperparameters

To verify the scientificity and reliability of the load forecasting model based on EMD-PSO-GRU proposed in this study, the power load data of a specific area from January 1, 2016, to May 8, 2020, are used. Figure 4 shows the raw data of the average daily load, with a total of 1588 data samples divided into a training set and test set with a ratio of 9 : 1, in which the training set data has 1430 data and the test set has 158 data.

4.1. EMD Decomposition. The load data are decomposed by EMD, and it is decomposed step by step from high frequency to low frequency into 9 IMF components and a residual component. The decomposition result is shown in Figure 5. Compared with the original power load sequence, the decomposed components become more stable in turn. The nine IMF components, respectively, reflect the influence of different influencing factors on the load data at different scales, and the residual component represents the long-term change trend of the load sequence.

4.2. Model Parameter Settings. The number of PSO population sizes is set to 8, the dimension of population optimization is set to 2, the maximum number of iterations is set to 10, the learning factors $c_1 = 2$, $c_2 = 2$, and the inertia weights are set to $w_{\max} = 1.2$, $w_{\min} = 0.8$. The range of the number of neurons is set to [1, 60], and the range of the learning rate is set to [0.001, 0.01].

4.3. Hyperparameter Optimization Results. The predictive performance of a GRU network depends on the choice of parameters for building the network, such as the learning rate and the number of hidden layer nodes. The learning rate and the number of hidden layer nodes in the GRU network are determined by the PSO optimization. Table 1 shows the optimal learning rate and optimal number of hidden layer nodes for the GRU network for each IMF.

In order to verify the superiority of the EMD-PSO-GRU hybrid prediction model, PSO-GRU, GRU, SVM, and RNN models were established under the same prediction process for comparative analysis.

4.4. Load Forecasting Result Analysis. Each model is trained with the data of the training set, the prediction results are compared and verified with the data of the test set, and the MAPE and RMSE evaluation indicators are selected to evaluate the accuracy of the prediction model. The

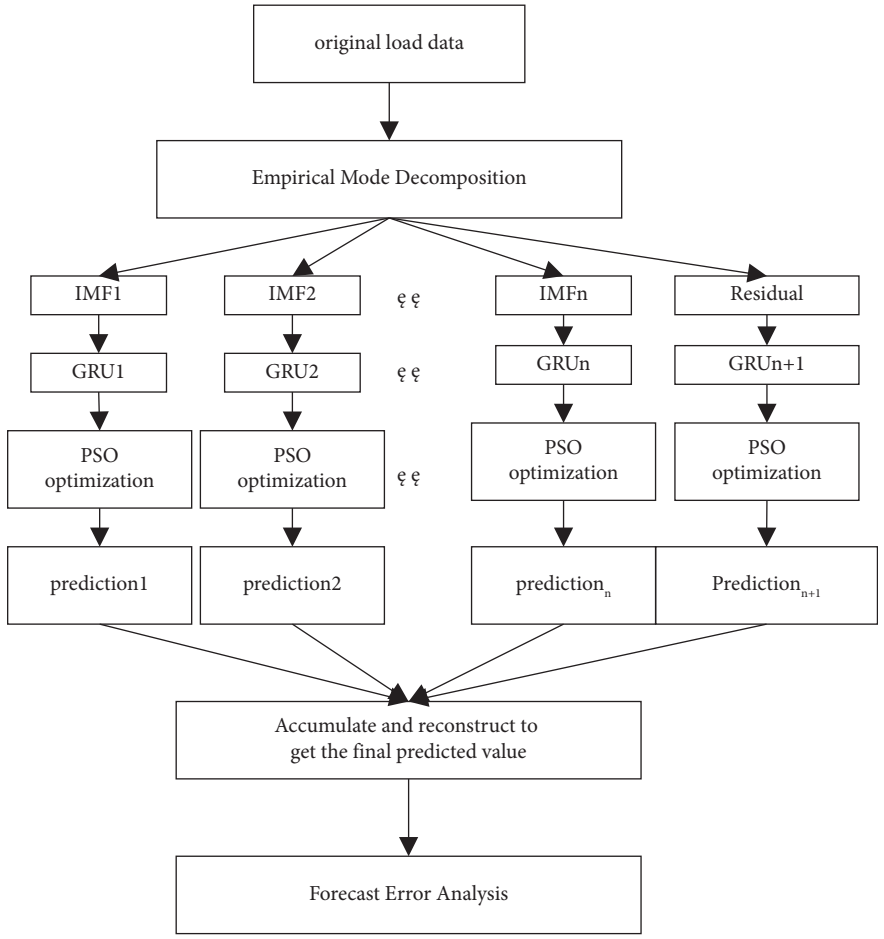

Figure 3: EMD-PSO-GRU combined prediction flowchart.

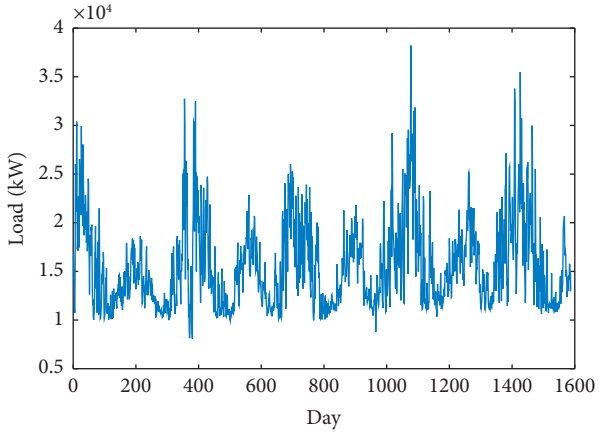

Figure 4: Load raw data.

comparison results of the test set are shown in Figure 6. In order to more clearly see the prediction performance of each model, the error rates of the last two months on the test set are selected for comparison, and the comparison chart is shown in Figure 7. It can be seen from Figures 6 and 7 that among the five models of EMD-PSO-GRU, PSO-GRU, GRU, SVM, and RNN, the overall fitting effect of the EMD-PSO-GRU model and the fitting effect at the peaks and troughs of the waves, The combined effect is better, and the error rate is obviously smaller. It can be seen that the empirical mode decomposition can stabilize the data with strong volatility and then predict it, which can effectively improve the prediction accuracy of the model.

The prediction accuracy evaluation results of each model are shown in Figure 8, and the quantitative indicators of each evaluation result are shown in Table 2. The accuracy evaluation results of the evaluation indicators of the EMD-PSO-GRU model are the best, and the prediction effect of the SVM model is the worst. The MAPE index of the EMD-PSO-GRU model is 1.678%, and the RMSE index is 259.32. Compared with the PSO-GRU prediction model, the MAPE is reduced by 24.92%, and the RMSE is reduced by 28.85%. Strong data stabilization before prediction can effectively improve the prediction accuracy of the model. The MAPE of PSO-GRU is 2.235%, and the RMSE is 364.48. Compared with the GRU model, its MAPE is reduced by 19.78%, and the RMSE is reduced by 25.60%. Due to the influence of experience, adaptively finding the optimal solution of hyperparameters is very important to improve the prediction performance of GRU, which can better improve the prediction accuracy.

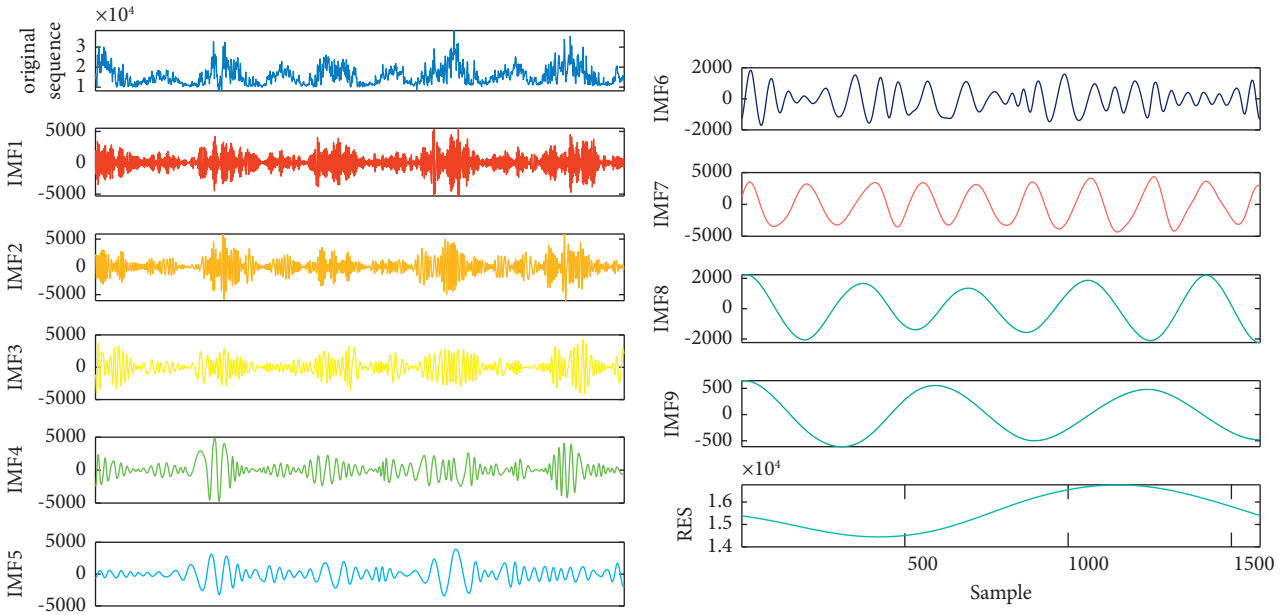

Figure 5: EMD exploded view.

Table 1: Hyperparameter optimization results for each IMF.

Sequence components	The number of neurons in the hidden layer	Learning rate
IMF1	36	0.0092
IMF2	52	0.0082
IMF3	42	0.0063
IMF4	38	0.0086
IMF5	40	0.0055
IMF6	55	0.0037
IMF7	56	0.0036
IMF8	38	0.0035
IMF9	53	0.0086
RES	38	0.0065

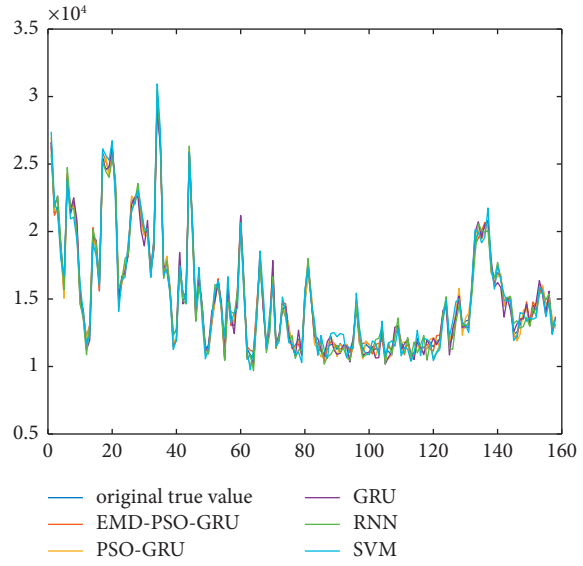

Figure 6: Comparison of prediction results of each model.

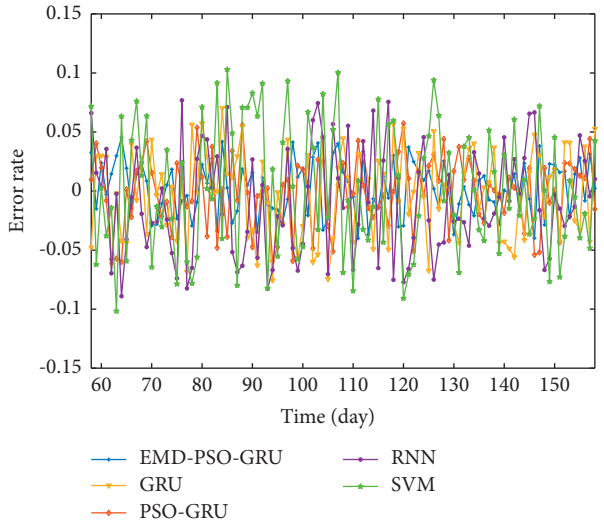

Figure 7: Comparison of the prediction results error rates of each model in the last two months on the test set.

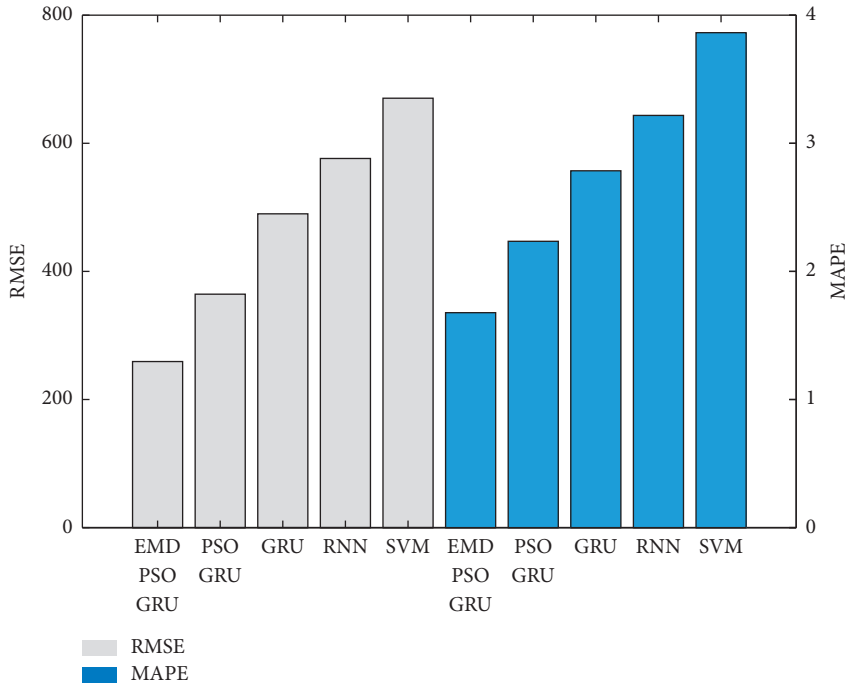

Figure 8: Comparison of prediction performance of each model.

Table 2: Evaluation indicators of the prediction performance of each prediction model.

Model	RMSE	MAPE/%
EMD-PSO-GRU	259.32	1.678
PSO-GRU	364.48	2.235
GRU	489.92	2.786
RNN	576.32	3.218
SVM	670.38	3.864

5. Conclusions

A load forecasting method based on the EMD-PSO-GRU is proposed. First, the data are subjected to variational modal decomposition, preprocessing, and division of the sample dataset, and then, a GRU neural network forecasting model is established for each subsequence. The PSO algorithm is used to optimize the hyperparameters of the GRU neural network, and finally, the respective sequence prediction results are accumulated to obtain the final load prediction value. The method proposed has the following merits:

(1) The EMD decomposition method can well mine the essential features such as cycles and trends in the sequence

(2) Show good advantages of the GRU network in the field of data mining. The GRU network has a unique network structure and strong learning ability for time series data.

(3) The PSO algorithm has stronger global optimization ability and convergence speed by adding nonlinearly changing inertia weights, avoiding the drawbacks of traditional manual selection of hyperparameters, and can independently optimize the hyperparameters of the network model

(4) The prediction accuracy of the "decomposition-prediction-reconstruction" method is higher. The proposed method demonstrates better prediction accuracy compared with the existing artificial intelligence prediction methods.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors gratefully acknowledge the support of Henan University and the authors of the literature cited in the paper.

References

[1] Z. Ruicheng, G. U. Jie, and J. Zhijian, "Research on short-term load forecasting variable selection based on fusion of data driven method and forecast error driven method," *Proceedings of the CSEE*, vol. 40, no. (2), pp. 487–500, 2020.

[2] I. K. Nti, M. Teimeh, O. Nyarko-Boateng, and A. F. Adekoya, "Electricity load forecasting: a systematic review," *Journal of Electrical Systems and Information Technology*, vol. 7, no. 1, pp. 1–19, 2020.

[3] A. Al Mamun, M. Sohel, N. Mohammad, M. d. S. H. Sunny, D. R. Dipta, and A. E. Hossain, "comprehensive review of the

load forecasting techniques using single and hybrid predictive models," *IEEE Access*, vol. 8, pp. 134911–134939, 2020.

[4] J. Zhang, Y. M. Wei, D. Li, Z. Tan, and J. Zhou, "Short term electricity load forecasting using a hybrid model," *Energy*, vol. 158, pp. 774–781, 2018.

[5] Z. Wang, Z. H. A. O. Bing, and J. I. Weijia, "Short-term load forecasting method based on GRU-NN model," *Automation of Electric Power Systems*, vol. 43, no. 5, pp. 53–58, 2019.

[6] K. Xiangyu, Z. Feng, and E. Zhijun, "Short-term load forecasting based on deep belief network," *Automation of Electric Power Systems*, vol. 42, no. 5, pp. 133–139, 2018.

[7] Q. Wu, J. Gao, G. S. Hou, B. Han, and K. Y. Wang, "Short-term load forecasting support vector machine algorithm based on multi-source heterogeneous fusion of load factors," *Automation of Electric Power Systems*, vol. 40, no. 15, pp. 67–72, 2016.

[8] J. Shi and J. Zhang, "Load forecasting based on multi-model by stacking ensemble learning," *Proceedings of the CSEE*, vol. 39, no. 14, pp. 4032–4042, 2019.

[9] T. Li, Y. Xu, J. Luo, J. He, and S. Lin, "A Method of Amino Acid Terahertz Spectrum Recognition Based on the Convolutional Neural Network and Bidirectional Gated Recurrent Network model," *Scientific Programming*, vol. 2021, 2021.

[10] Y. Gao, R. Wang, and E. Zhou, "Stock Prediction Based on Optimized LSTM and GRU Models," *Scientific Programming*, vol. 2021, 2021.

[11] X. Wang, W. Li, and Q. Li, "A New Embedded Estimation Model for Soil Temperature prediction," *Scientific Programming*, vol. 2021, 2021.

[12] A. J. Quinn, V. Lopes-dos-Santos, D. Dupret, A. C. Nobre, and M. W. Woolrich, "EMD: empirical mode decomposition and Hilbert-Huang spectral analyses in Python," *Journal of open source software*, vol. 6, no. 59, 2021.

[13] J. J. Ruiz-Aguilar, I. Turias, J. González-Enrique, D. Urda, and D. Elizondo, "A permutation entropy-based EMD–ANN forecasting ensemble approach for wind speed prediction," *Neural Computing & Applications*, vol. 33, no. 7, pp. 2369–2391, 2021.

[14] D. Han, N. Zhao, and P. Shi, "Gear fault feature extraction and diagnosis method under different load excitation based on EMD, PSO-SVM and fractal box dimension[J]," *Journal of Mechanical Science and Technology*, vol. 33, no. 2, pp. 487–494, 2019.

[15] Y. Zhang, T. Zhou, X. Huang, L. Cao, and Q. Zhou, "Fault diagnosis of rotating machinery based on recurrent neural networks," *Measurement*, vol. 171, Article ID 108774, 2021.

[16] Y. Suo, W. Chen, C. Claramunt, and S. Yang, "A ship trajectory prediction framework based on a recurrent neural network," *Sensors*, vol. 20, no. 18, p. 5133, 2020.

[17] S. Yang, X. Yu, and Y. Zhou, "Lstm and gru neural network performance comparison study: taking yelp review dataset as an example," in *Proceedings of the 2020 International workshop on electronic communication and artificial intelligence (IWECAI)*, pp. 98–101, Shanghai, China, June2020.

[18] W. Deng, J. Xu, H. Zhao, and Y. Song, "A Novel Gate Resource Allocation Method Using Improved PSO-Based QEA," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 3, pp. 1737–1745, 2020.

[19] X. Ren, S. Liu, X. Yu, and X. Dong, "A method for state-of-charge estimation of lithium-ion batteries based on PSO-LSTM," *Energy*, vol. 234, Article ID 121236, 2021.

[20] Y. Zhu, G. Li, R. Wang, S. Tang, H. Su, and K. Cao, "Intelligent fault diagnosis of hydraulic piston pump combining improved LeNet-5 and PSO hyperparameter optimization," *Applied Acoustics*, vol. 183, Article ID 108336, 2021.

[21] W. Zhu, H. N. Rad, and M. Hasanipanah, "A chaos recurrent ANFIS optimized by PSO to predict ground vibration generated in rock blasting[J]," *Applied Soft Computing*, vol. 108, Article ID 107434, 2021.

[22] S. S. Band, S. Janizadeh, S. Chandra Pal et al., "Novel ensemble approach of deep learning neural network (DLNN) model and particle swarm optimization (PSO) algorithm for prediction of gully erosion susceptibility," *Sensors*, vol. 20, no. 19, p. 5609, 2020.

[23] L. T. Le, H. Nguyen, J. Dou, and J. Zhou, "A comparative study of PSO-ANN, GA-ANN, ICA-ANN, and ABC-ANN in estimating the heating load of buildings' energy efficiency for smart city planning," *Applied Sciences*, vol. 9, no. 13, p. 2630, 2019.

[24] H. Eskandari, M. Imani, and M. P. Moghaddam, "Convolutional and recurrent neural network based model for short-term load forecasting," *Electric Power Systems Research*, vol. 195, Article ID 107173, 2021.

[25] L. Yin and J. Xie, "Multi-temporal-spatial-scale temporal convolution network for short-term load forecasting of power systems," *Applied Energy*, vol. 283, 2021.

[26] R. Zhang, S. Chen, Z. Zhang, and W. Zhu, "Genetic Algorithm in Multimedia Dynamic Prediction of Groundwater in Open-Pit Mine," *Computational Intelligence and Neuroscience*, vol. 2022, 2022.

[27] L. Li, C. Mao, H. Sun, Y. Yuan, and B. b. Lei, "Digital twin driven green performance evaluation methodology of intelligent manufacturing: hybrid model based on fuzzy rough-sets AHP, multistage weight synthesis, and PROMETHEE II," *Complexity*, vol. 2020, no. 6, pp. 1–24, 2020.

[28] E. Pekel, M. Gul, E. Celik, and S. Yousefi, "Metaheuristic approaches integrated with ANN in forecasting daily emergency department visits," *Mathematical Problems in Engineering*, vol. 2021, p. 2021.

[29] L. Li and C. Mao, "Big data supported PSS evaluation decision in service-oriented manufacturing," *IEEE Access*, vol. 8, pp. 154663–154670, 2020.

[30] R. Kafieh, R. Arian, N. Saeedizadeh et al., "COVID-19 in Iran: forecasting pandemic using deep learning," *Computational and Mathematical Methods in Medicine*, vol. 2021, 2021.

[31] L. Li, T. Qu, Y. Liu et al., "Sustainability assessment of intelligent manufacturing supported by digital twin," *IEEE Access*, vol. 8, pp. 174988–175008, 2020.

[32] L. Li, B. Lei, and C. Mao, "Digital twin in smart manufacturing," *Journal of Industrial Information Integration. January*, vol. 26, no. 9, 2022.

[33] Y. Mai, Z. Sheng, H. Shi, and Q. Liao, "Using improved XGBoost algorithm to obtain modified atmospheric refractive index," *International Journal of Antennas and Propagation*, vol. 2021, 2021.