


ORIGINAL RESEARCH

Short-term electrical load forecasting model based on multi-dimensional meteorological information spatio-temporal fusion and optimized variational mode decomposition

Lingyun Wang¹  | Xiang Zhou¹ | Honglei Xu² | Tian Tian¹ | Huamin Tong³

¹College of Electrical Engineering and New Energy, China Three Gorges University, Yichang, China

²School of Electrical Engineering, Computing and Mathematical Sciences, Curtin University, Perth, WA, Australia

³State Grid Yichang Power Supply Company, Yichang, China

Correspondence

Lingyun Wang, College of Electrical Engineering and New Energy, China Three Gorges University, Yichang 443002, China.
Email: wly@ctgu.edu.cn

Funding information

National Natural Science Foundation of China, Grant/Award Number: 51907104; Australian Research Council, Grant/Award Number: LP160100528; 2022 CU Building Industry Project, Grant/Award Number: CUB2022; 2022 Curtin University Science and Engineering Faculty Small Grant, Grant/Award Number: SRG2022

Abstract

This paper proposes a method to enhance the accuracy of power load forecasting by considering the variability in the impact of multi-dimensional meteorological information on power load in diverse regions. The proposed method employs spatio-temporal fusion (SF) of multi-dimensional meteorological information and applies the Copula theory to analyze the non-linear coupling of meteorological information from multiple stations with power load to achieve SF in the spatial dimension. To enhance the accuracy of load forecasting in the time dimension, this paper improves the core parameters of the variational mode decomposition (VMD) using the marine predators algorithm (MPA) and utilizes the weighted permutation entropy (WPE) to construct the MPA-VMD fitness function for the adaptive decomposition of the load sequence. Moreover, this paper constructs input sets for the long short-term memory model and the MPA-LSSVM model by combining each component of the time dimension and each meteorological information of the spatial dimension to obtain the prediction results of each component. The prediction model corresponding to each component is selected according to the evaluation index and reconstructed to obtain the overall prediction results. The analysis results demonstrate that the proposed forecasting method outperforms the traditional forecasting method and effectively enhances the accuracy of power load forecasting.

1 | INTRODUCTION

Global problems like air pollution and global warming are receiving more and more attention. Many nations now agree that reducing greenhouse gas emissions and achieving ‘carbon neutrality’ and ‘zero emissions’ are important goals. Short-term power load forecasting is vital for the balance between supply and demand in the power grid [1, 2]. According to the load forecast values, energy can be planned and dispatched effectively to decrease energy waste [3, 4]. With the increasing proportion of renewable energy, stored energy, and electric vehicles in the power system [5], electrical consumption patterns tend to be more intricate. The non-linear and non-smooth characteristics of power load data have intensified, making short-term load forecasting more demanding. Consequently, novel forecasting

methods need to be proposed to satisfy the demand for higher accuracy in short-term load forecasting.

Meteorology as a major influence on electrical load makes it highly influenced by thermal inertia and the impact varies by region [6]. In the conventional electrical load forecasting process for municipal areas, the forecasting model generally uses the city-wide meteorological information throughout the day as influencing factors, such as city-wide rainfall and maximum temperature [7]. However, meteorological information in different areas within the municipal area has certain different in effects on the electrical load. Consequently, it is essential to establish a short-term load forecasting model based on the correlation characteristics of meteorological information from different meteorological stations in the region and the electrical load in both time and space dimensions [8]. The commonly

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *IET Generation, Transmission & Distribution* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology.

used correlation analysis methods are Pearson and Kendall algorithms, but both of them are mainly applied to linear analysis, which is not efficient for non-linear power load data and meteorological data [9]. Copula theory has been applied to the study of complex problems in many fields in recent years and has many advantages such as flexible and variable forms, not being limited by marginal distribution etc. It can analyze non-linear correlations objectively, quantitatively, and accurately, and provide a methodological basis for the establishment of spatio-temporal fusion models of multi-dimensional weather information [10].

At present, machine learning techniques are introduced into load forecasting, which can be divided into two categories according to the number of forecasting models. One type of model is the single forecasting model, which first uses correlation analysis to extract the influencing factors with a high degree of coupling to the electrical load, and constructs the input feature set of the forecasting model. Then least-squares support vector machine (LSSVM) [11], long short-term memory network (LSTM) [12] etc., are used to further extract the data features and obtain the load forecasting values. LSTM can solve the long-term dependence problem of recurrent neural network (RNN), and its long- and short-term memory capability can significantly improve prediction accuracy and efficiency. LSSVM has the advantage of solving non-linear problems and is widely used in the field of power load prediction. Due to the non-linear and non-smooth nature of the power load, the ideal prediction accuracy is difficult to be achieved with a single prediction model due to structural limitations. Therefore, the second type of 'decomposition-prediction-reconstruction' model, which combines the advantages of signal processing and multiple prediction methods, has become the focus of research at this stage [13–17]. Wavelet decomposition, empirical modal decomposition, and variational modal decomposition algorithms have received much attention in short-term load forecasting [14]. Wavelet decomposition for load data processing is related to the number of mother wavelets and components and is not very adaptive. The empirical modal decomposition has some adaptiveness, but it has the defect of modal mixing in the decomposition process. The load data can be broken down using variable mode decomposition (VMD) into intrinsic mode functions (IMF) of various frequencies. It can be used to build models for short-term load forecasting because it is adaptable and non-recursive. However, parameter tuning in the VMD decomposition procedure is time-consuming and subjective, which impacts the decomposition's accuracy. References [15, 16] proposed a prediction method using an optimization algorithm for the automatic optimization of key parameters of VMD, which achieved good prediction accuracy, but the model was complex and not adapted to time series. The permutation entropy (PE) method was suggested in [17] to analyze the complexity of each modal function of VMD and reorganize the modal function to obtain subsequences, effectively increasing the efficiency of VMD in processing time series. However, the parameter setting is still arbitrary, which affects prediction accuracy.

The marine predators algorithm (MPA) chooses between the Lévy flight and Brownian motion as the optimum forag-

ing strategy by simulating the evolution of predators and prey in the ocean [18]. MPA has the advantages of minimal model parameters, difficulty in falling into local optimums, and great optimization efficiency when compared to other techniques such as particle swarm optimization (PSO).

Based on the above-described investigation, this paper proposes a short-term load combination prediction model based on the spatial-temporal fusion of multi-dimensional meteorological information and MPA to optimize the key parameters of VMD. In the spatial dimension, meteorological information from multiple meteorological stations such as wind speed, wind direction, temperature, and sunshine intensity is fused with power load information in time and space dimensions. The non-linear coupling between power load and multi-dimensional meteorological information is correlated by using the Copula theory to filter out the important features, reduce the feature dimension, and improve forecasting efficiency. In the temporal dimension, firstly, MPA was used to find the optimum for the key VMD parameters. To take into account the change magnitude of different components and to adapt to short-term load prediction, the sum of weighted permutation entropy (WPE) of modal functions with different frequencies is minimized as the objective function to decompose the original load data into components with different frequency characteristics. Moreover, to improve prediction accuracy by combining the prediction advantages of different prediction models for different frequency components, the LSTM model and the MPA-LSSVM model were used to obtain the prediction results of each component, and the corresponding prediction model for each component was selected according to the prediction evaluation index. Finally, reconstruction is used to obtain the overall prediction results. Based on actual load data, a multi-group comparison study is carried out to validate the generalization and superiority of the prediction method proposed in this paper.

The novel perspectives of the proposed model are summarized as follows: (1) To investigate the non-linear coupling of meteorological data from various meteorological stations with the power load and to accomplish spatio-temporal fusion, the Copula theory is utilized in the spatial dimension. (2) The MPA enhances the fundamental VMD features in the temporal dimension. The WPE is used to construct the MPA-VMD fitness function for adaptive decomposition of load sequences. (3) The combined forecasting method of LSTM and MPA-LSSVM is used to validate the prediction accuracy and generalization ability of the proposed model by comparing it with traditional forecasting methods, and by analyzing the working days and holidays in spring, summer, autumn, and winter, respectively.

The remainder of this paper is organized as follows: Section 2 provides an overview of the Copula theory, performs spatio-temporal fusion analysis of multi-dimensional meteorological data, and creates the matrix of influencing factors. Section 3 provides a VMD optimization technique to test the convergence of the suggested method and introduces the fundamentals of VMD and MPA. The SF-MPA-VMD combined forecasting model is established in Section 4, and the performance evaluation indicators of this paper are provided. Section 5 validates the prediction approach put out in this work, discusses the

model's superiority and effectiveness, and thoroughly analyzes the findings. The complete text is summarized in Section 6.

2 | RELATED ALGORITHM BASIS

2.1 | Correlation analysis based on Copula theory

Copula theory can conduct non-linear coupling analysis on numerous variables and does not require any linear or Gaussian assumptions. Therefore, the Copula theory is applied to perform a quantitative correlation analysis on the power load and multi-dimensional meteorological data [19].

Assume that the electrical load sequence is u and the temperature sequence is v in the example of power load and temperature. If $H(u, v)$ is defined as the joint distribution function of both, and the edge distribution function corresponds to $F_u(u)$ and $F_v(v)$, then a Copula function $C(\cdot)$ exists that associates the joint distribution and the edge distribution with two functions

$$H(u, v) = C [F_u(u), F_v(v)] \quad (1)$$

For the correlation analysis between the two in this paper, the Kendall rank correlation coefficient method with high applicability is chosen, and the expression is as follows:

$$\tau = 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1 \quad (2)$$

where τ is the outcome of Kendall correlation analysis. The closer the value of $|\tau|$ is to 1, the stronger the correlation between the variables.

There are some common Copula functions such as t-Copula, Gumbel Copula, Clayton Copula, Frank Copula, and Normal Copula. The correlation analysis of various sequences can be performed by using each different Copula function, so it is essential to choose the appropriate Copula function for accurate analysis. In this study, the best Copula function for the non-linear coupling analysis of meteorological information series and load series is chosen using the minimal Euclidean distance approach [20].

2.2 | Spatial-temporal fusion of multi-dimensional meteorological information

To comprehensively assess the meteorological data of various locations during different seasons, this paper employs a spatio-temporal fusion method of multi-dimensional meteorological data based on the Copula theory. It consists of three phases specifically: (1) For the 12 meteorological information shown in Figure 1, the optimal Copula function is selected for the non-linear coupling analysis based on the minimum Euclidean distance method. The primary meteorological data that influ-

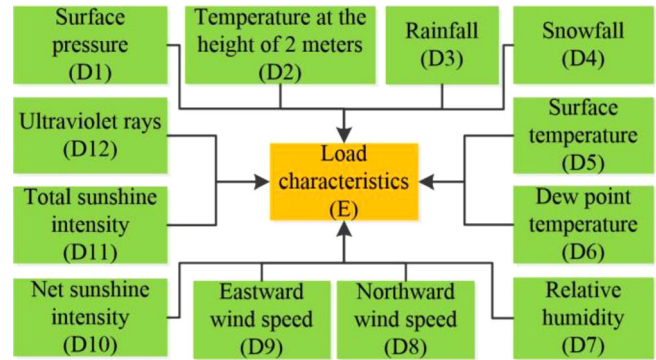


FIGURE 1 Meteorological information and load characteristics.

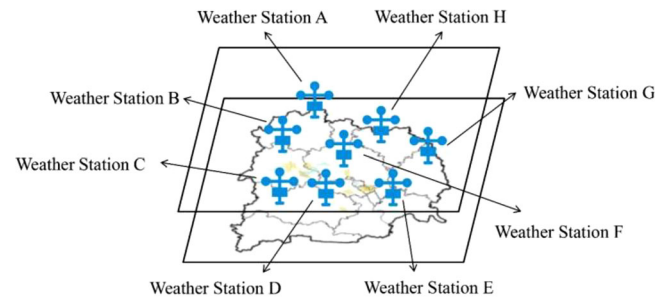


FIGURE 2 Spatio-temporal interaction diagram of multi-dimensional meteorological information.

ences the power load for each season is chosen; (2) As shown in Figure 2, based on the Copula theory, the spatial-temporal correlation analysis of the primary meteorological information from meteorological stations in various distribution areas and the power load is conducted to determine their correlation degree, respectively; (3) Choosing the meteorological station data with the highest correlation with the power load to obtain the meteorological information fusion result and form the influencing factor matrix.

2.3 | Spatio-temporal fusion of multi-dimensional meteorological information based on Copula theory

In Figure 3a, the square Euclidean distance value of each Copula function for the spatio-temporal fusion of multi-dimensional meteorological data from March to May (Spring) is displayed. The corresponding thermodynamic diagram with load is created for each meteorological information sequence by choosing the best Copula function using the least Euclidean distance method, as shown in Figure 3b. The correlation degree is indicated by each colour block in the figure; the deeper the colour block, the higher the correlation. The correlation degree, which can be employed as a load influencing factor, is larger than or equivalent to 0.3 [21].

The strongest correlation, 0.589, between surface temperature and load may be found in Figure 3b. In addition, there is a strong correlation between load and temperature at the height

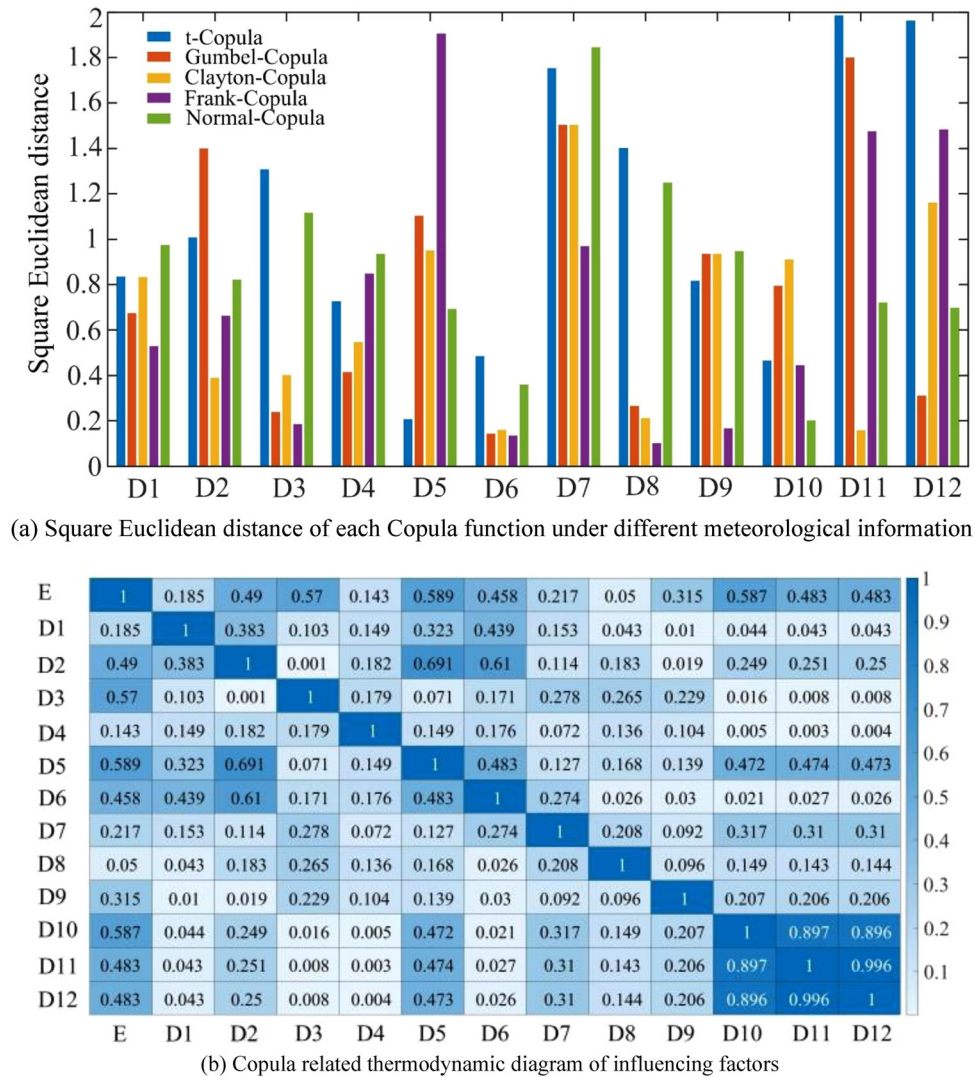


FIGURE 3 Spatio-temporal fusion of multi-dimensional meteorological information based on Copula theory. (a) Square Euclidean distance of each Copula function under different meteorological information. (b) Copula-related thermodynamic diagram of influencing factors.

of 2 m, precipitation, dew point temperature, east wind speed, net sunshine intensity, total sunshine intensity, and ultraviolet rays intensity. Since there are high mutual correlations between temperature at the height of 2 m, surface temperature, and dew point temperature, and between net sunlight intensity, total sunlight intensity, and ultraviolet rays intensity, to reduce the input dimensionality, they can be considered a class of meteorological information respectively, and represented by the meteorological information with the highest correlation to the load. As a result, the primary meteorological data are the surface temperature, precipitation, the east wind speed, and net sunshine intensity.

The data from meteorological stations F_A – F_H are then filtered and spatial-temporally merged, as indicated in Table 1, by the spatial-temporal correlation between the primary meteorological information of each meteorological station and the electrical load. An influencing factor matrix is created using the following variables: F_A surface temperature, F_C precipitation,

F_E east wind speed, and F_G net sunshine intensity. The analysis procedure is the same in the summer, fall, and winter as described above.

3 | VARIATIONAL MODAL DECOMPOSITION AND PREDICTION MODEL

3.1 | Variational mode decomposition

The electrical power demand is affected by a variety of factors, such as the weather, holidays etc., which exacerbates its non-linear and non-stationary characteristics and makes projections more challenging [22]. Power load data can be deconstructed into modal components with different frequencies using VMD to lessen their complexity and non-stationary characteristics [23, 24].

TABLE 1 Correlation coefficient of each weather station.

Meteorological stations	Surface temperature	Rainfall	Eastward wind speed	Net sunshine intensity
F_A	0.431	0.354	0.418	0.481
F_B	0.364	0.410	0.391	0.443
F_C	0.407	0.519	0.367	0.371
F_D	0.271	0.342	0.443	0.218
F_E	0.349	0.274	0.317	0.344
F_F	0.212	0.412	0.510	0.319
F_G	0.371	0.334	0.271	0.541
F_H	0.228	0.419	0.329	0.351

Assume that the input signal $f(t)$ is decomposed into k modes $u_k(t)$, and Hilbert transformation is performed for each mode to obtain the analytic signal of $u_k(t)$. We have the following calculation formula:

$$\begin{cases} \min_{\{u_k\}, \{\omega_k\}} \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \times u_k(t) \right] e^{-j\omega_k t} \right\|^2 \\ s.t. \sum_k u_k(t) = f(t) \end{cases} \quad (3)$$

where $\{u_k\}$ and $\{\omega_k\}$ are the modal functional and central frequency of the k th IMF component, and $\delta(t)$ is the unit impulse signal.

Using the augmented Lagrange function to solve Equation (3), we obtain

$$\begin{aligned} L(\{u_k\}, \{\omega_k\}, \lambda) = & \alpha \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \times u_k(t) \right] e^{-j\omega_k t} \right\|^2 \\ & + \left\| f(t) - \sum_{k=1}^K u_k(t) \right\|^2 + \left\langle \lambda(t), f(t) - \sum_{k=1}^K u_k(t) \right\rangle \end{aligned} \quad (4)$$

where α denotes a penalty parameter, $\lambda(t)$ represents a Lagrangian multiplier, and K is the number of modal decomposition.

Then, the alternating direction method of multipliers (ADMM) is adopted to solve Equation (4). The iterative formulas of which can be deduced as follows:

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \hat{\lambda}(\omega)/2}{1 + 2\alpha(\omega - \omega_k)^2} \quad (5)$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k(\omega)| d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 d\omega} \quad (6)$$

where $\hat{u}_k^{n+1}(\omega)$ is the Wiener filter results corresponding to IMF components, and ω_k^{n+1} is the centre frequency of the components.

3.2 | Marine predator algorithm

The marine predator algorithm replicates the differential mobility of active and passive predators in marine animals at different velocities to enhance the merit-seeking capability of the algorithm in a parallel architecture [25, 26]. The optimal foraging strategies of marine predators are Lévy wandering and Brownian wandering, which allow the algorithm to seek merit both locally and globally. The mathematical model of the MPA algorithm is as follows:

Firstly, the prey matrix \mathbf{P} is initialized. The following equation is used to establish each element X_{ij} in the matrix:

$$X_{ij} = X_{\min} + \text{rand}(0, 1) \cdot (X_{\max} - X_{\min}) \quad (7)$$

where X_{\max} and X_{\min} are the upper and lower bounds of the solution, respectively. $\text{rand}(0,1)$ refers to a random coefficient in the range of $(0,1)$.

The prey matrix \mathbf{P} is obtained as follows:

$$\mathbf{P} = \begin{bmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,d} \\ X_{2,1} & X_{2,2} & \cdots & X_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n,1} & X_{n,2} & \cdots & X_{n,d} \end{bmatrix} \quad (8)$$

where n is the number of search agents and d is the position of each dimension.

The fitness value is calculated for each individual $X_i = [X_{i,1}, X_{i,2}, \dots, X_{i,d}]$ in each prey matrix \mathbf{P} . The top predator matrix \mathbf{E} is obtained by copying n copies of the individual \mathbf{X}_i^I with the best fitness

$$\mathbf{E} = \begin{bmatrix} X_{1,1}^I & X_{1,2}^I & \cdots & X_{1,d}^I \\ X_{2,1}^I & X_{2,2}^I & \cdots & X_{2,d}^I \\ \vdots & \vdots & \vdots & \vdots \\ X_{n,1}^I & X_{n,2}^I & \cdots & X_{n,d}^I \end{bmatrix} \quad (9)$$

During the optimization process, t denotes the current number of iterations, t_{\max} denotes the maximum number of iterations, and the optimization search process is divided into three stages, the locally optimal solution is disturbed using the fish aggregation devices (FADs) effect, and the post-perturbation

fitness is calculated to update the iterative process. The moving step of prey matrix \mathbf{P} and top predator matrix \mathbf{E} in the iterative process is denoted by \mathbf{d}_i . \mathbf{R}_L denotes Lévy motion, \mathbf{R}_B denotes Brownian motion, and the Kronecker product of prey and top predator under the two wandering modes and three stages is used as the influencing factor of the moving step \mathbf{d}_i .

Phase 1: When $t < t_{\max}/3$ is a high-velocity ratio stage, the top predator gives up hunting, the prey for Brownian motion, the mathematical expression is as follows:

$$\mathbf{P}_{i+1} = \mathbf{P}_i + p \cdot \mathbf{R} \otimes \mathbf{d}_i \quad (10)$$

$$\mathbf{d}_i = \mathbf{R}_B \otimes (\mathbf{E}_i - \mathbf{R}_B \otimes \mathbf{P}_i), i = 1, \dots, n \quad (11)$$

where p is taken as a constant 0.5, R is a uniform random number between 0 and 1, P_i and E_i are the prey position, and top predator position at the current moment t , respectively.

Phase 2: When $t_{\max}/3 < t < 2t_{\max}/3$ is the equal-velocity ratio phase, the population is updated in two parts: the first half of the population is the Lévy motion population, which is responsible for exploitation, and the second half of the population is the Brownian motion population, which is responsible for exploration. The parallel architecture of the MPA algorithm is reflected by this evenly divided population, and the mathematical expression of this phase is shown as follows:

$$\mathbf{P}_{i+1} = \begin{cases} \mathbf{P}_i + p \cdot \mathbf{R} \otimes \mathbf{d}_i & i = 1, \dots, n/2 \\ \mathbf{P}_i + p \cdot C_F \otimes \mathbf{d}_i & i = n/2, \dots, n \end{cases} \quad (12)$$

$$\mathbf{d}_i = \begin{cases} \mathbf{R}_L \otimes (\mathbf{E}_i - \mathbf{R}_L \otimes \mathbf{P}_i), i = 1, \dots, n/2 \\ \mathbf{R}_B \otimes (\mathbf{R}_B \otimes \mathbf{E}_i - \mathbf{P}_i), i = n/2, \dots, n \end{cases} \quad (13)$$

$$C_F = (1 - t/t_{\max})^{2t/t_{\max}} \quad (14)$$

where p is taken as 0.5 as above, and C_F is the adaptive coefficient to control the movement step of the predator.

Phase 3: When $t > 2t_{\max}/3$ is the low-velocity ratio stage, the prey adopts the same Lévy motion as the top predator for movement, and the mathematical expression at this time is shown as follows:

$$\mathbf{P}_{i+1} = \mathbf{E}_i + p \cdot C_F \otimes \mathbf{d}_i \quad (15)$$

$$\mathbf{d}_i = \mathbf{R}_L \otimes (\mathbf{R}_L \otimes \mathbf{E}_i - \mathbf{P}_i), i = 1, \dots, n \quad (16)$$

Furthermore, FADs or eddy formation effect usually alters the foraging behaviour of marine predators, and this strategy allows the algorithm to avoid local extremes as much as possible during the iterative process in order to achieve better optimization results. The following are the mathematical expressions:

$$\mathbf{P}_{i+1} = \begin{cases} \mathbf{P}_i + C_F \cdot \mathbf{P}_0 \otimes \mathbf{U}, r \leq p_F \\ \mathbf{P}_i + [p_F(1 - r) + r] (\mathbf{P}_{r1} - \mathbf{P}_{r2}), r > p_F \end{cases} \quad (17)$$

where \mathbf{U} is a randomly generated binary vector, p_F is the perturbation probability factor, which takes the value of 0.2, r is a uniform random number from 0 to 1, and the P_{r1} and P_{r2} are the randomly removed individuals in the prey matrix, respectively.

3.3 | MPA-based VMD optimization method

In the parameter initialization process of VMD, two important parameters, penalty factor α and the number of decompositions k , need to be adjusted first. If the value of k is too small, the decomposition is incomplete and will cause information loss; conversely, over-decomposition will occur [27, 28]. The penalty factor α mainly affects the bandwidth of the decomposed spectrum of each mode and the degree of convergence of the algorithm [29]. Artificially set parameters make the decomposition structure subjective; for this reason, this paper proposes to optimize the parameters of VMD using the MPA algorithm.

3.3.1 | MPA optimization algorithm tests

VMD parameter optimization can be considered a two-dimensional function optimization problem, and to verify the advantage of MPA in this optimization problem, optimization experiments are performed for the function in Equation (14):

$$f = x^2 - 10 \cos(2\pi x) + j^2 - 10 \cos(2\pi j) + 10 \quad (18)$$

PSO, Sparrow Search Algorithm (SSA), and MPA were selected to find their minimum values, respectively. The population sizes of the three algorithms are set the same to ensure the rationality and fairness of the experiments: the population size is set to 20 and the number of iterations is set to 500. Both the individual learning factor and the social learning factor of the PSO algorithm are set to 1.5, the safety threshold of Sparrow Search ST is 0.8, the discoverer is 20% of the population size, and the number of vigilantes is set to 5; the FAD of MPA algorithm is 0.2. The convergence curves corresponding to each algorithm are depicted in Figure 4.

From Figure 4b, it can be seen that compared with the PSO and SSA algorithms, the MPA algorithm has a faster search speed and the fastest convergence; in the case of 500 iterations, the PSO search result is 0.0234, the SSA search result is 0.0127, and the MPA algorithm search result is 1.2347×10^{-5} , which is closer to the global optimum. To sum up, MPA has a better optimization efficiency. Therefore, the MPA algorithm is used to optimize the key parameters of VMD to reduce the loss of the decomposition process, and the subjectivity of the decomposition structure, and improve the decomposition effect.

3.3.2 | Design of the fitness function

The convergence speed of the optimization algorithm and the solution of the optimal solution position is influenced by the fitness function. The PE algorithm can better reflect the change pattern and complexity of the time series [30–32], and the alignment entropy algorithm is used to construct the fitness function in this paper.

Let the time series be decomposed into $\{j(1), j(2), \dots, j(k)\}$ components and $\{s(t), t = 1, 2, \dots, N\}$ be the time series of IMF,

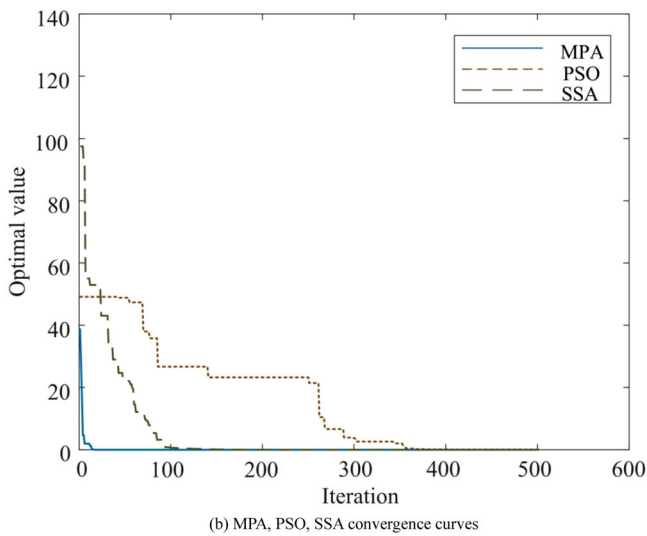
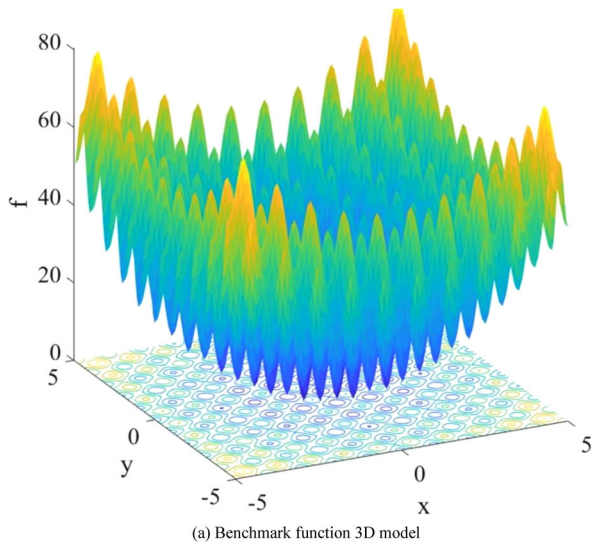


FIGURE 4 Comparison of optimization algorithm test. (a) Benchmark function 3D model. (b) MPA, PSO, SSA convergence curves. MPA, marine predators algorithm; PSO, particle swarm optimization; SSA, sparrow search algorithm.

and the series be processed using phase space reconstruction

$$x_j = [s(j), s(j + \tau), \dots, s(j + (m - 1) \tau)] \quad (19)$$

where m is the number of dimensions, $j = 1, 2, \dots, N - (m - 1) \tau$, and τ is the delay time. The sequence of the reconstructed matrix after sorting is

$$S(g) = (j_1, j_2, \dots, j_m) \quad (20)$$

where $g = 1, 2, \dots, l$. After calculating the probability values P_1, P_2, \dots, P_l for the occurrence of each symbol in $S(g)$, we then obtain the PE of the sequence $\{s(t), t = 1, 2, \dots, N\}$

$$H_p(\hat{m}) = - \sum_{g=1}^l P_g \ln P_g \quad (21)$$

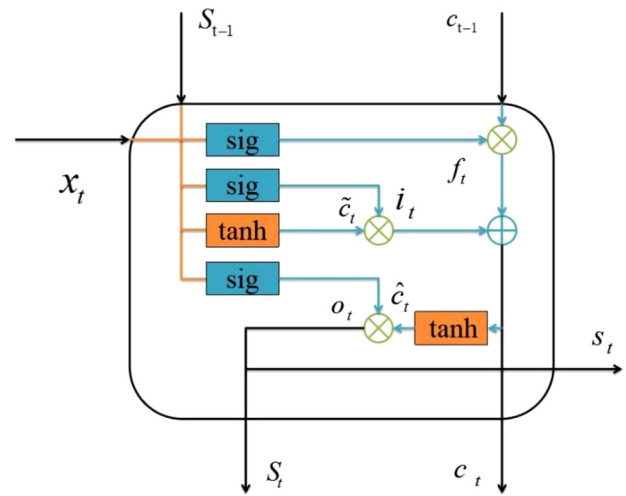


FIGURE 5 Schematic diagram of LSTM structure. LSTM, long short-term memory.

The normalized form of the PE $H_p(m)$ is obtained as

$$H_p = \frac{H_p}{\ln(m!)} \quad (22)$$

The PE H_p value measures the degree of randomness of the sequence, with larger values indicating a more random sequence or smaller values indicating a more regular sequence.

From Equations (17) and (18), the magnitude of each component PE is calculated as $\{P_E(1), P_E(2), \dots, P_E(k)\}$, respectively. To sum the PE values of each component, WPE is utilized. The average amplitude ratio of each component to the original sequence is used to calculate the weight value. The WPE value is calculated, and the fitness function of the MPA algorithm is set to the WPE value, so that the VMD penalty factor α and the number of decompositions k can be ideally determined [29, 30].

3.4 | LSTM neural network forecasting

In the LSTM network, the hidden layer neurons of the traditional RNN network are replaced by a memory unit [33–35]. The structure of this memory unit consists of input gates, forgetting gates, and output gates. The cell transmits data over arbitrary time periods. The gates follow the input and output data flow from the cell. The LSTM network can be used to reduce difficulties with gradient disappearing and exploding that were discovered by RNN. The structure of LSTM is shown in Figure 5.

In Figure 5, x_t denotes the current input information, s_{t-1} is the state of the implied layer at the previous moment, and the quantities are calculated as

$$f_t = \text{sigmoid}(W_{fx}x_t + W_{fs}s_{t-1} + b_f) \quad (23)$$

$$i_t = \text{sigmoid}(W_{ix}x_t + W_{is}s_{t-1} + b_i) \quad (24)$$

$$c'_t = \tanh(W_{cx}x_t + W_{cs}s_{t-1} + b_c) \quad (25)$$

$$o_t = \text{sigmoid}(W_{ox}x_t + W_{os}s_{t-1} + b_o) \quad (26)$$

$$c'_t = \tanh(c_t) \quad (27)$$

where W_{fx} , W_{ix} , W_{cx} , W_{ox} , W_{fs} , W_{is} , W_{cs} , and W_{os} are the weight matrices, b_f , b_i , b_c , and b_o are the bias terms. Moreover, f_t , i_t , and o_t are real numbers between $[0,1]$, which determine the forgetting ratio of c , thus obtaining the s_t corresponding to moment t :

$$c_t = f_t c_{t-1} + i_t c'_t \quad (28)$$

$$s_t = o_t c'_t \quad (29)$$

3.5 | MPA-LSSVM neural network

Least-Squares Support Vector Machine (LSSVM) is a better model obtained by optimizing the Support Vector Machine (SVM) model to improve the convergence speed and the accuracy of the final prediction results during the iterative process [36–38]. However, the selection of parameters in the LSSVM model has a great impact on the performance of the model, so this paper proposes MPA to optimize the parameters of the LSSVM model [39]. The optimization model of LSSVM is as follows:

$$\min_{\omega, \xi, b} J(\omega, \xi) = \frac{1}{2} \omega^T \omega + \frac{\gamma}{2} \sum_{i=1}^N \xi_i^2$$

$$s.t. y_i = \omega^T \varphi(x_i) + b + \xi_i, i = 1, 2, \dots, n \quad (30)$$

The optimal MPA-LSSVM model is obtained by iteratively finding the optimal MPA-LSSVM model based on the MPA algorithm for the regularization parameter γ and the kernel function parameter σ . The optimization steps are as follows:

1. Initialize the MPA algorithm's parameters, such as population size, position, iterations, and so forth.
2. Determine the fitness function and calculate the fitness.
3. Update the parameters and calculate the fitness, compare the historical optimum with the current fitness, and if the latter is greater, set the current top predator position as the optimum to achieve the position update.
4. Determine whether the maximum number of iterations is reached, if not, return to step 3 until the maximum number of iterations is reached.
5. Output the optimal regularization parameter γ and the kernel function parameter σ to complete the MPA-LSSVM model, and bring the optimal parameters into the LSSVM model to complete the prediction of temporal data.

4 | SF-MPA-VMD COMBINED FORECASTING MODEL

4.1 | Short-term power load forecasting model

Figure 6 depicts the general framework of multi-dimensional meteorological information spatio-temporal fusion and MPA-VMD decomposition of the short-term load combination forecasting model used in this paper. The Copula theory is used to calculate the influencing factor matrix by examining the spatio-temporal correlation between meteorological data and the power load of weather stations in various regions. When combined with MPA-VMD, the IMF components are obtained by decomposing the load data from the model input matrix, and the input matrix is divided into training and testing. The training set optimizes the LSTM and MPA-LSSVM model parameters to obtain the best prediction model and predicted value for each IMF component, the test set selects the prediction model corresponding to each frequency component based on the evaluation index, and the test set feeds each frequency component into the corresponding prediction model. The prediction model is reconstructed to obtain the final prediction result.

4.2 | Evaluation index of prediction results

In this experiment, there are three error statistics, namely root mean square error (RMSE), prediction efficiency (PE), and mean absolute percentage error (MAPE), which were used as evaluation indicators of the prediction results

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left\| \frac{\hat{y}_i - y_i}{y_i} \right\| \times 100\% \quad (31)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (32)$$

$$PE = \left(1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \right) \times 100\% \quad (33)$$

where \hat{y}_i is the predicted value of load power, y_i is the corresponding true value of load power, and n is the number of prediction points.

5 | ANALYSIS OF ALGORITHMS

5.1 | Model sources and data preparation

The experimental data include 35,040 points of power load data sampled every 15 min from 1 March 2020 to 28 February 2021, for the city of Yichang in Hubei province, in the middle of China, as well as the historical meteorological data sampled every 15 min by eight meteorological stations in the region. As shown in Figure 1, the data set contains the contents.

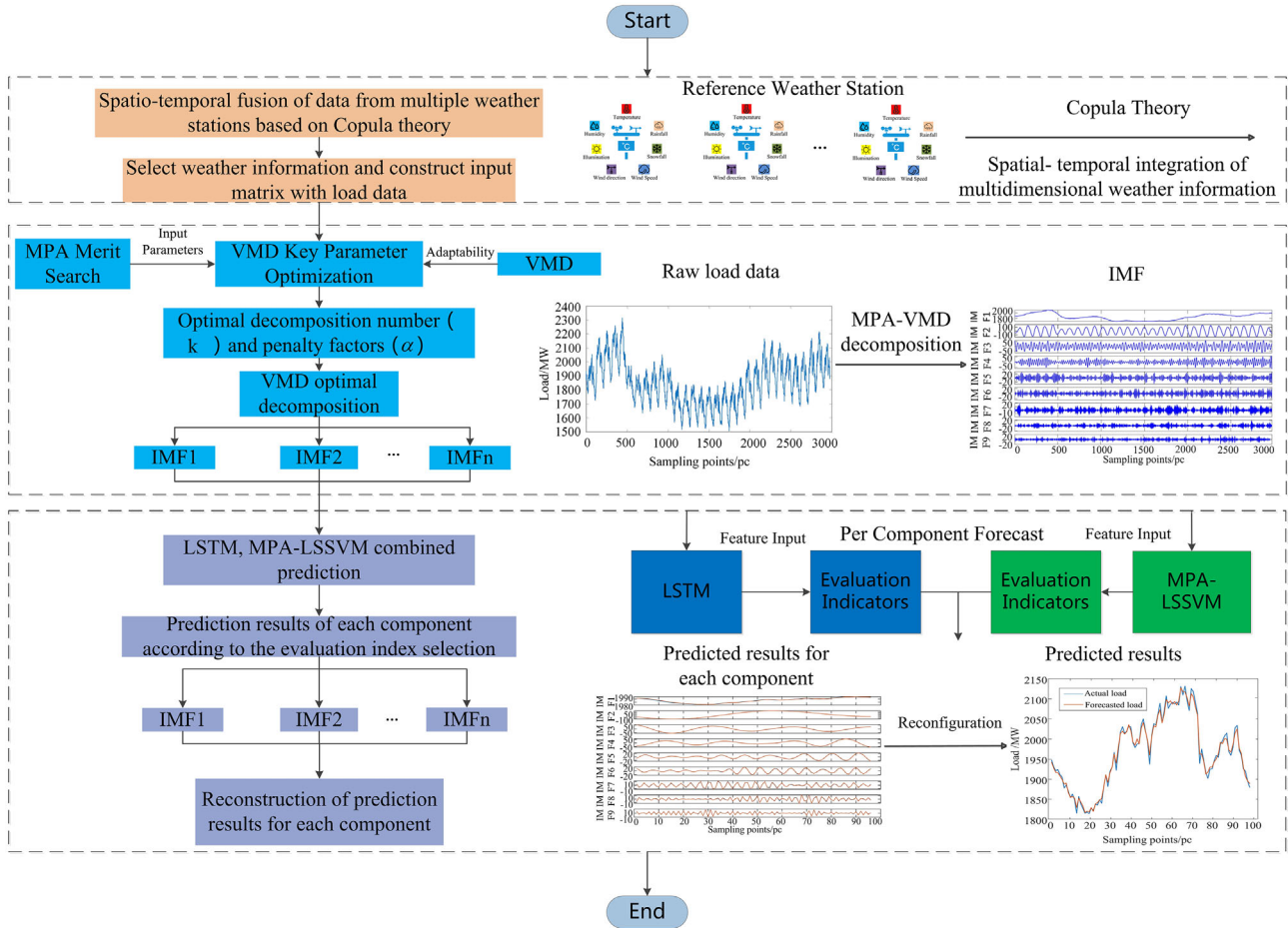


FIGURE 6 Short-term load combination forecasting model based on multi-dimensional meteorological information spatio-temporal fusion and MPA-VMD. MPA, marine predators algorithm; VMD, variational mode decomposition.

Considering the characteristics of power load in each season, we divide the data set into four periods: March to May (Spring), June to August (Summer), September to November (Autumn), and December to February (Winter). Each period data was divided into weekdays and holidays according to weekdays or not. The precede 80% of the data in each period is taken as the training set, and the subsequent 20% is taken as the test set. Each data point is normalized to [0,1] to remove the influence of varying factor magnitudes. The influence factor matrix is then created using the multi-dimensional meteorological information fusion method described in Section 1.3. The input set is made up of MPA-VMD load decomposition components. The single-step rolling forecast is then used to predict the load value at 96 points for the next 24 h. The model is updated with the projected load values as new features, and the predicted load values are then acquired and back-normalized.

5.2 | MPA-VMD load series decomposition

To demonstrate the benefits of MPA-VMD load data decomposition, the VMD parameters were obtained by optimizing the objective function constructed in Section 3.3.2 using PSO, SSA,

and MPA, and the load sequence was decomposed into the corresponding components, and the resulting parameter values and load decomposition loss values are shown in Table 2.

As shown in Table 2, compared with the empirical setting parameters and the PSO and SSA setting parameters, the decomposition loss is smaller using MPA setting parameters, which is reduced by 85.39%, 80.79%, and 69.22%, respectively. From the above analysis, it can be concluded that the MPA-VMD decomposition can reduce the decomposition loss and improve the decomposition effect by adaptively determining the optimal parameters of VMD while avoiding the randomness brought by the empirical setting parameters, and the decomposition sequence obtained by decomposing the load data using the MPA-VMD method is shown in Figure 7.

5.3 | Short-term load forecasting results

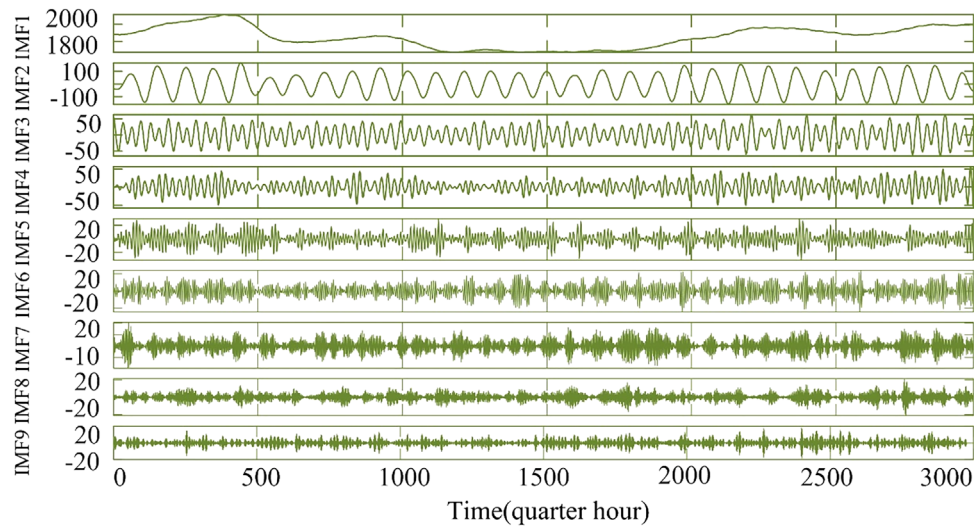
5.3.1 | SF-MPA-VMD combined forecasting

The high- and low-frequency components of the training set were brought into the combined LSTM and MPA-LSSVM prediction models, and the prediction results were obtained by

TABLE 2 Decomposition parameter setting and experimental results of VMD.

Method of decomposition	Decomposition parameters		Results of decomposition
	Number of decomposition k	Penalty factor α	Decomposition of loss(MW)
VMD	5	1000	20.33
PSO-VMD	15	2546	15.46
SSA-VMD	13	2358	9.65
MPA-VMD	9	1443	2.97

MPA, marine predators algorithm; PSO, particle swarm optimization; SSA, sparrow search algorithm; VMD, variational mode decomposition.

**FIGURE 7** Load data decomposition.**TABLE 3** Evaluation index of each prediction model.

IMF	LSTM			MPA-LSSVM		
	RMSE/MW	MAPE /%	PE /%	RMSE /MW	MAPE /%	PE /%
IMF ₁	7.169	3.358	83.413	5.073	3.147	87.374
IMF ₂	1.493	2.599	89.475	0.158	0.672	98.453
IMF ₃	0.573	1.866	94.264	0.701	1.975	92.716
IMF ₄	0.349	1.347	96.724	10.967	3.523	85.134
IMF ₅	0.350	1.494	95.721	11.262	3.597	85.472
IMF ₆	0.524	1.948	93.176	7.872	3.367	86.436
IMF ₇	0.461	1.765	94.192	6.002	3.231	87.154
IMF ₈	0.481	1.895	94.656	3.716	2.962	89.575
IMF ₉	0.290	0.738	98.773	1.051	2.256	89.726

IMF, intrinsic mode functions; LSTM, long short-term memory; MAPE, mean absolute percentage error; PE, prediction efficiency; RMSE, root mean square error.

using the LSTM and MPA-LSSVM models for each component, and the evaluation indexes RMSE and MAPE values of each component prediction result are shown in Table 3.

As shown in Table 3, for the IMF1 and IMF2 components, the RMSE, MAPE, and PE values of the predicted values obtained using MPA-LSSVM forecasting are smaller

than those of the LSTM forecasting model, and for the IMF3-IFM9 components, the RMSE and MAPE values of the predicted values obtained using the LSTM forecasting model are smaller than those of the MPA-LSSVM, so these two evaluation indexes are considered together to obtain IMF1-IMF2 by using the MPA-LSSVM prediction model, and

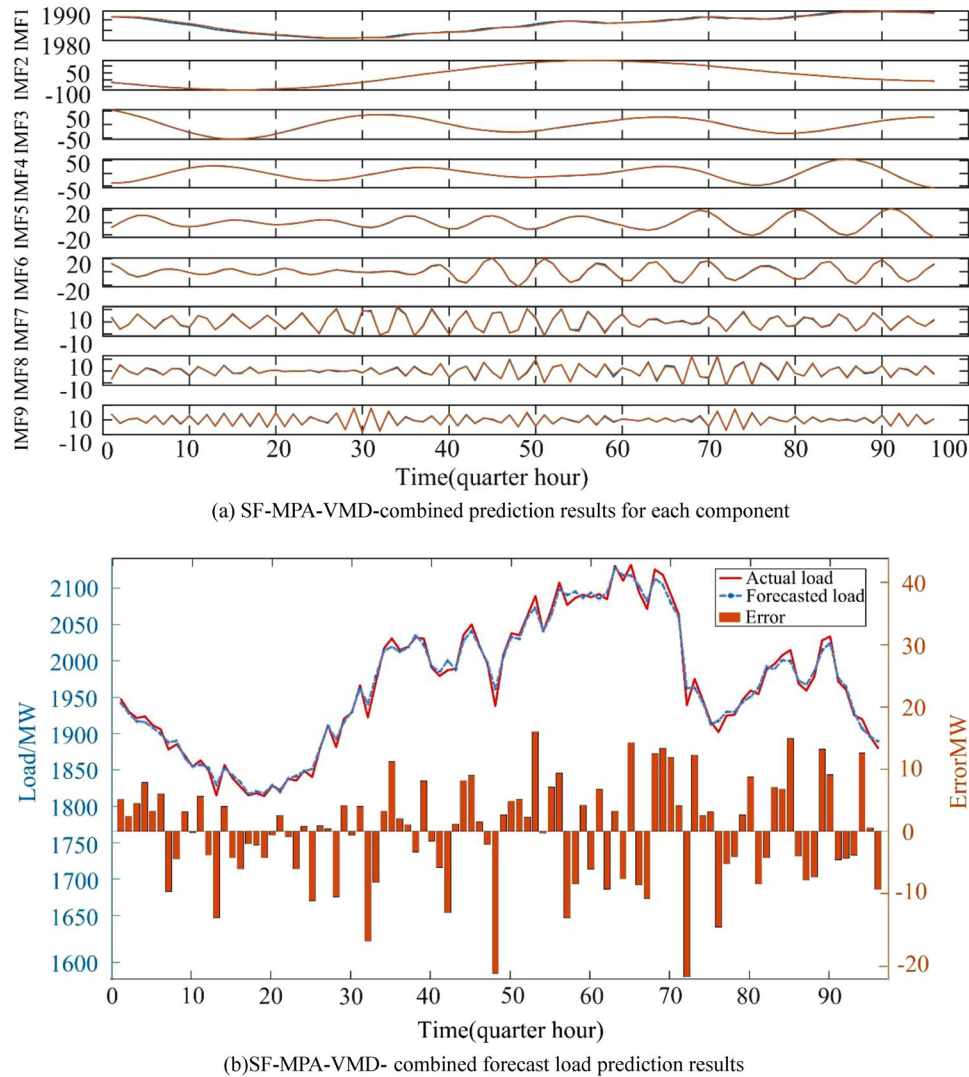


FIGURE 8 SF-MPA-VMD combined forecasting component and load forecasting result. (a) SF-MPA-VMD-combined prediction results for each component. (b) SF-MPA-VMD-combined forecast load prediction results. MPA, marine predators algorithm; VMD, variational mode decomposition; SF, spatio-temporal fusion.

IMF3-IMF9 by using the LSTM prediction model. Based on the above method, the prediction results of each component on 30 April 2021 are obtained as shown in Figure 8a, and the overall load prediction results are obtained by reconstructing the component prediction results as shown in Figure 8b.

5.3.2 | Comparison with other methods

To verify the effectiveness of the proposed method, the following three sets of comparative experiments are designed to predict the validation set load data and conduct comparative analysis.

Comparison experiment 1

To verify the influence of the spatio-temporal fusion of multi-dimensional meteorological information on the accuracy of

TABLE 4 Comparison of prediction evaluation indexes between single meteorological station and spatio-temporal fusion.

Model	RMSE/MW	MAPE/%	PE/%
MPA-VMD-LSTM	25.321	0.964	96.457
SF-MPA-VMD-LSTM	11.946	0.494	98.981

LSTM, long short-term memory; MAPE, mean absolute percentage error; MPA, marine predators algorithm; PE, prediction efficiency; RMSE, root mean square error; SF, spatio-temporal fusion; VMD, variational mode decomposition.

load forecasting, the meteorological data for which feature selection and spatio-temporal fusion were used as two sets of influencing factors, and the components obtained from the MPA-VMD decomposition formed the input matrix of the LSTM forecasting model to obtain the MPA-VMD-LSTM and SF-MPA-VMD-LSTM forecasting results, and the evaluation indicators are shown in Table 4.

TABLE 5 Comparison of MPA optimized VMD decomposition load forecasting evaluation indexes.

Model	RMSE/MW	MAPE/%	PE/%
SF-VMD-LSTM	23.622	0.883	97.378
SF-PSO-VMD-LSTM	21.564	0.821	97.448
SF-SSA-VMD-LSTM	18.463	0.762	97.846
SF-MPA-VMD-LSTM	11.946	0.494	98.981

LSTM, long short-term memory; MAPE, mean absolute percentage error; MPA, marine predators algorithm; PE, prediction efficiency; PSO, particle swarm optimization; RMSE, root mean square error; SF, spatio-temporal fusion; SSA, sparrow search algorithm; VMD, variational mode decomposition.

From Table 4, the influencing factors matrix formed through the spatio-temporal fusion of multi-dimensional meteorological information, compared with the influencing factors matrix formed by feature selection only; the prediction evaluation indexes of the proposed model such as the value of RMSE are reduced by 13.375 MW, the value of MAPE is reduced by 0.47%. The PE value is improved by 2.52%. It is hereby proved that the accuracy of load forecasting can be enhanced by spatio-temporal fusion of multi-dimensional meteorological information effectively.

Comparison experiment 2

To verify the effectiveness of the MPA-VMD decomposition signal, the meteorological data after the spatio-temporal fusion of multi-dimensional meteorological information is used as the input matrix of influence factors, and the load data are processed by PSO-VMD, SSA-VMD, MPA-VMD decomposition, and empirically set VMD parameter decomposition respectively. Four groups of decomposition results are obtained, and each constitutes an input matrix with influence factors and is input to the LSTM prediction model. The prediction results of the four decomposition methods, SF-PSO-VMD-LSTM, SF-SSA-VMD-LSTM, SF-MPA-VMD-LSTM, and SF-VMD-LSTM, were obtained, and the evaluation indexes were compared as shown in Table 5.

From the evaluation results in Table 5, it can be seen that the prediction ability of the prediction model is further improved after optimizing the VMD decomposition parameters by PSO, SSA, and MPA. Compared with the traditional empirical setting of VMD parameters, PSO-VMD and SSA-VMD for load data decomposition prediction, the MPA-VMD method proposed in this paper is optimal for all three evaluation indicators, indicating that MPA has the best finding effect and effectively improves the prediction accuracy.

Comparison experiment 3

To verify the influence of the combined forecasting model on the accuracy of load forecasting, the meteorological data after the spatio-temporal fusion of multi-dimensional meteorological information is used as the influencing factor, and the load components decomposed by MPA-VMD form the input matrix, which is the input to the LSTM forecasting model, MPA-LSSVM forecasting model, and combined

TABLE 6 Comparative results of different forecasting methods.

Model	RMSE /MW	MAPE /%	PE /%
MPA-VMD-LSTM	25.321	0.964	96.457
SF-VMD-LSTM	23.622	0.883	97.378
SF-PSO-VMD-LSTM	21.564	0.821	97.448
SF-SSA-VMD-LSTM	18.463	0.762	97.846
SF-MPA-VMD-MPA-LSSVM	18.335	0.625	98.435
SF-MPA-VMD-LSTM	11.946	0.494	98.981
The proposed	8.207	0.333	99.356

LSTM, long short-term memory; MAPE, mean absolute percentage error; MPA, marine predators algorithm; PE, prediction efficiency; PSO, particle swarm optimization; RMSE, root mean square error; SF, spatio-temporal fusion; SSA, sparrow search algorithm; VMD, variational mode decomposition.

LSTM and MPA-LSSVM forecasting model, and the forecasting results of each model are obtained. The prediction results of SF-MPA-VMD-LSTM, SF-MPA-VMD-MPA-LSSVM, and the method SF-MPA-VMD-combination of this paper are shown in Figure 9.

Summarizing the above three sets of experiments, the comparison of prediction evaluation indexes for a total of seven prediction methods in three groups of comparison tests was obtained as shown in Table 6.

As shown in Figure 9 and Table 6, compared with other forecasting models, the RMSE of the SF-MPA-VMD-combined forecasting model proposed in this paper decreases by 17.114, 15.415, 13.357, 10.256, 10.128, and 3.739 MW, respectively, from the perspective of the average absolute percentage error. In terms of the average absolute percentage error, the MAPE values of the proposed model decrease by 65.46%, 62.29%, 59.44%, 56.30%, 46.72%, and 32.59%, respectively. From the PE perspective, the forecasting model proposed in this paper is better than other models. In summary, the multi-dimensional meteorological information spatio-temporal fusion and MPA-VMD decomposition short-term load combination forecasting model proposed in this paper has better stability and the accuracy of load forecasting has been significantly improved.

Predictive model performance comparison experiments

To verify the impact of the forecasting model on the accuracy of the load forecast, the prediction models proposed in this paper are compared with XGBoost, CNN-BiLSTM, LSTM, and GBDT models. The prediction results of each model are obtained as shown in Figure 10, and the comparison of prediction evaluation indicators is shown in Figure 11.

As can be seen from Figures 10 and 11, compared to other prediction models, the RMSE of the prediction models proposed in this paper decreased by 4.3799, 1.4869, 1.4737, and 0.2464 MW, respectively. MAPE values decreased by 32.7678%, 15.0158%, 14.6469%, and 3.3675% respectively. From a PE value perspective, the proposed method in this paper has the largest PE value. In summary, the short-term electricity load forecasting method proposed in this paper has better stability and the accuracy of load forecasting has been significantly

FIGURE 9 Comparison of load forecasting effect.

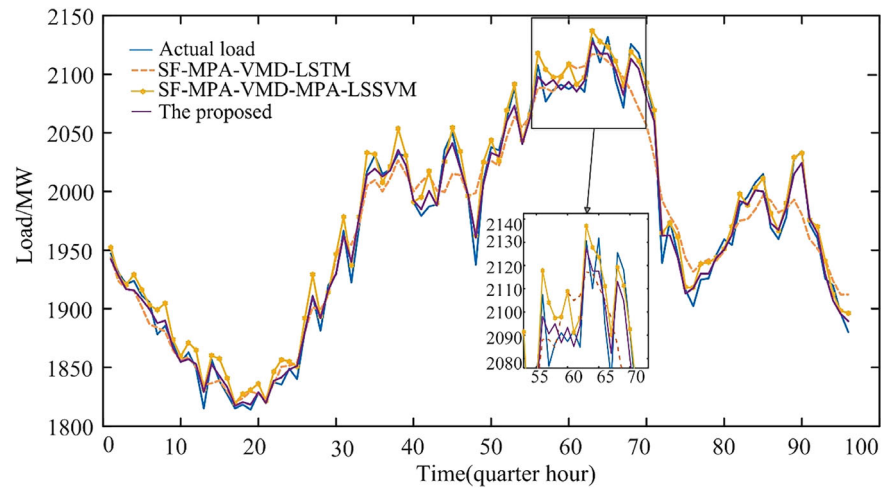


FIGURE 10 Comparison of load forecasting effect.

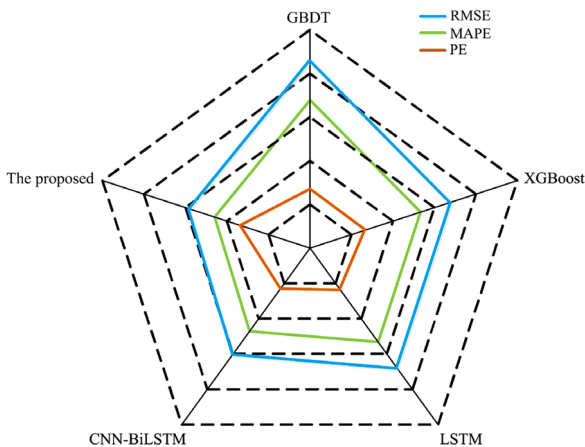
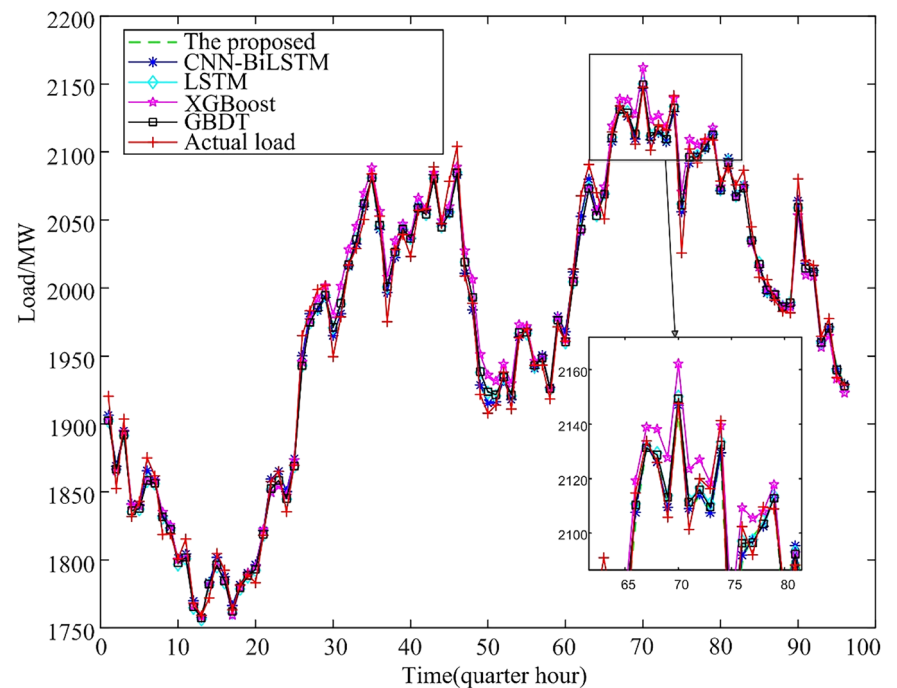


FIGURE 11 Different prediction model evaluation indicators.

improved. New methods are provided for power load forecasting to ensure safe and stable operation of the power system.

Model generalization ability test

To verify the generalization ability of the model proposed in this paper, choose 12 and 16 May 2020 as spring weekday and holiday forecast samples, 27 and 29 August 2020 as summer weekday and holiday forecast samples, 24 and 28 November 2020 as autumn weekday and holiday forecast samples, and 22 and 26 February 2021 as winter weekday and holiday forecast samples, respectively, the load forecast results are obtained as shown in Figure 12, and the evaluation indexes are shown in Figure 13.

From Figures 12 and 13, it can be seen that the proposed forecasting model in this paper has the best RMSE, MAPE,

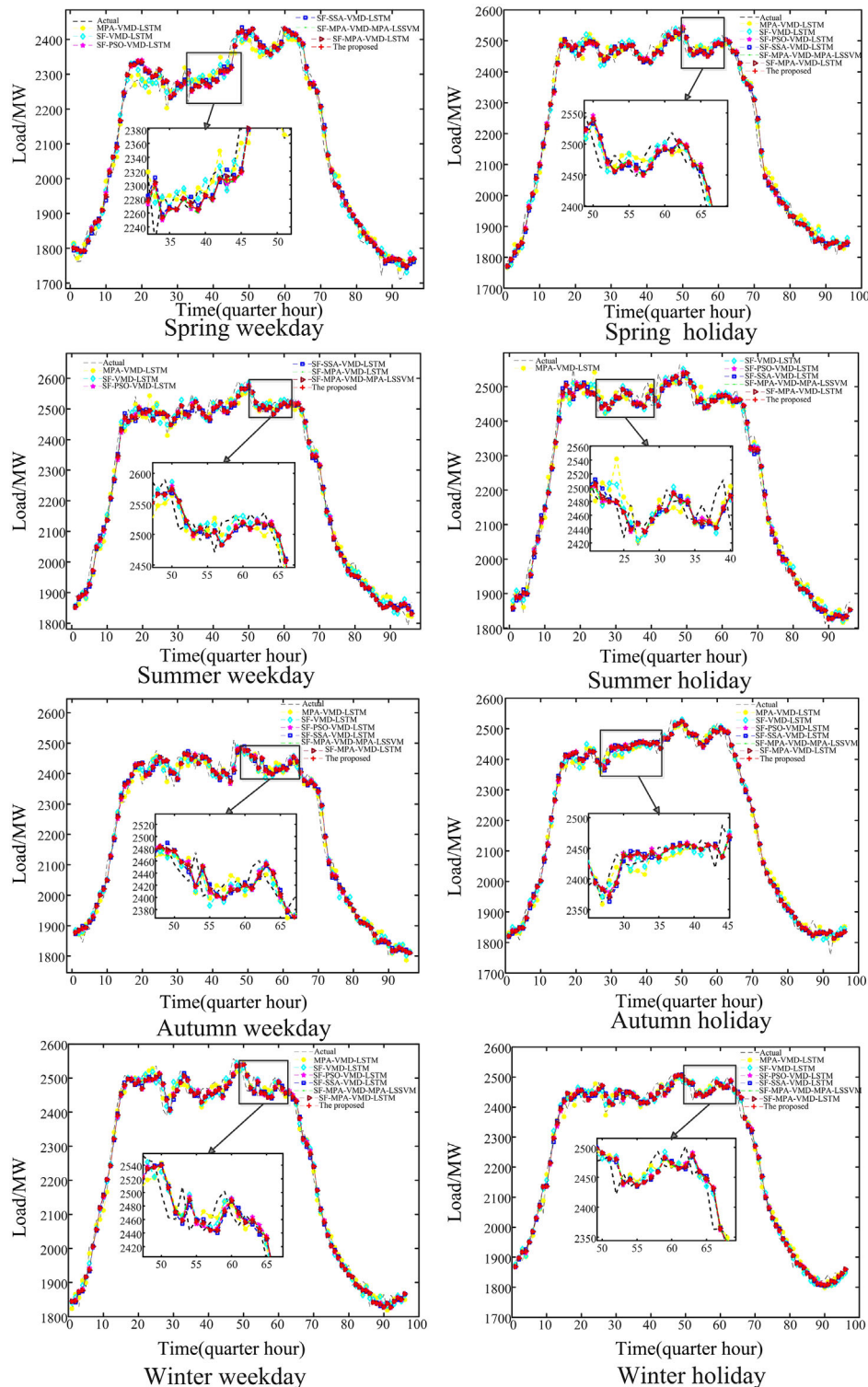


FIGURE 12 Load forecast results of working days and rest days by season in Spring, Summer, Autumn, and Winter.

and PE values for weekday and holiday forecasting in Spring, Summer, Autumn, and Winter seasons, which shows that the model is suitable for electrical load forecasting in different seasons and fully reflects the model generalization ability and forecasting performance.

In summary, the proposed multidimensional meteorological information spatio-temporal fusion and MPA-VMD short-term load combination prediction model has good stability and the generalization ability and the load prediction accuracy have been significantly improved.

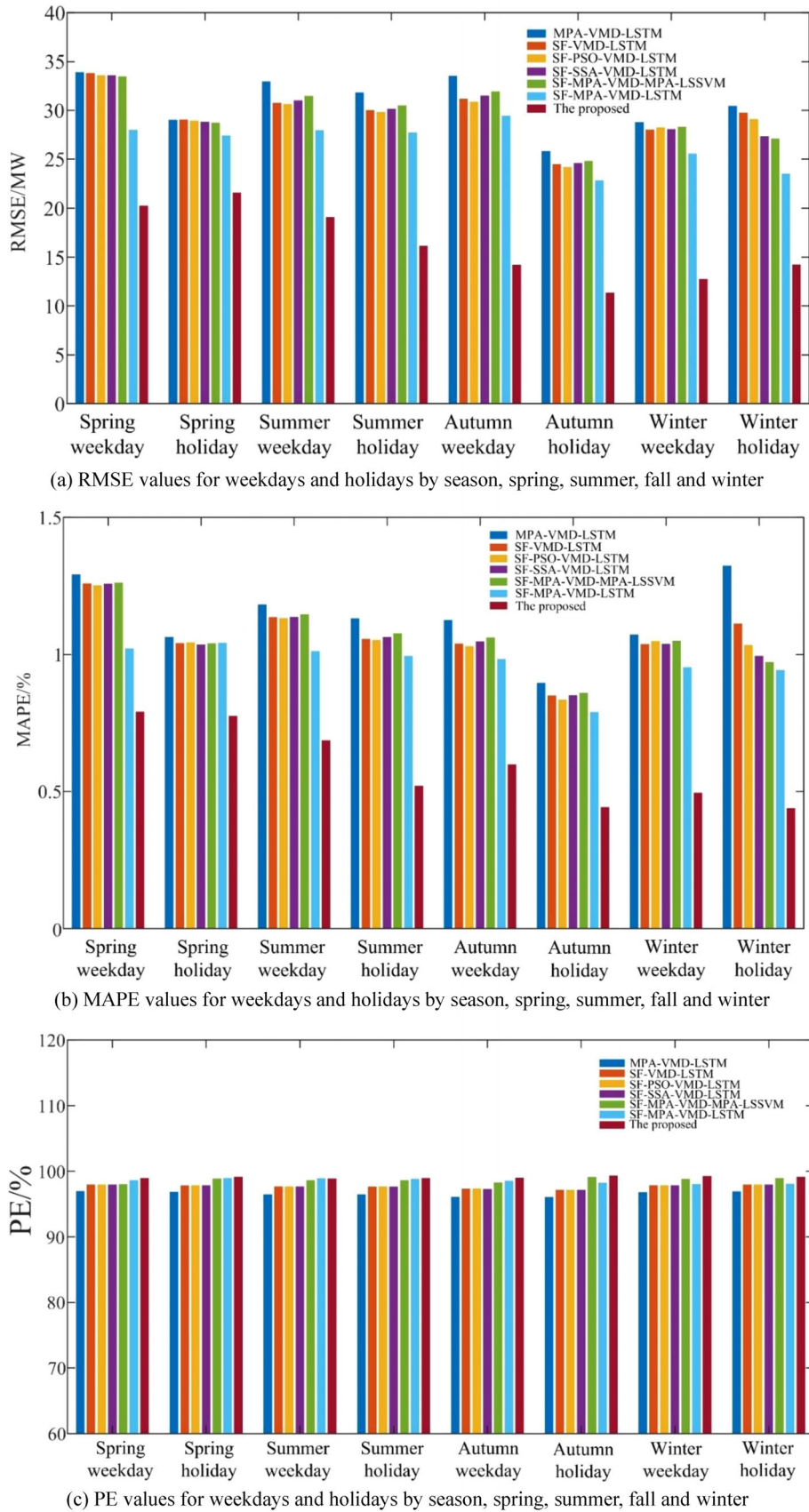


FIGURE 13 Forecast evaluation of working days and rest days by season in spring, summer, autumn, and winter. (a) RMSE values for weekdays and holidays by season, spring, summer, fall, and winter. (b) MAPE values for weekdays and holidays by season, spring, summer, fall, and winter. (c) PE values for weekdays and holidays by season, spring, summer, fall, and winter. MAPE, mean absolute percentage error; PE, prediction efficiency; RMSE, root mean square error.



The forecasting method is applied to the local power grid load forecasting, which has achieved excellent results, improved the power grid operation efficiency, and reduced the work intensity of the operation and maintenance personnel.

6 | CONCLUSIONS

To improve the prediction accuracy of power load, this paper considers the characteristics of the load and meteorological information in both time and space dimensions and designs a combined MPA-VMD short-term load prediction model in the time dimension. In the spatial dimension, the meteorological information such as wind speed, wind direction, temperature, and sunshine intensity from multiple meteorological stations are temporally and spatially fused with the electrical load information to forecast the load according to the meteorological factors in the region, and the following conclusions are obtained based on the comparative analysis of experimental results.

1. Based on the Copula theory, the spatio-temporal fusion of power load and meteorological data from multiple meteorological stations, and the minimum Euclidean distance method to select the optimal Copula function for non-linear coupling analysis of meteorological information series and load series in each season can effectively reduce the feature dimension and improve the PE and accuracy.
2. The MPA search algorithm has better search capability and speed than the PSO algorithm and the SSA algorithm, and can effectively avoid falling into local optimum in the search process. MPA is used for VMD key parameter optimization, compared with PSO, SSA optimized VMD parameters, and empirically set VMD parameters, the RMSE and MAPE values are optimal, fully demonstrating the necessity of optimizing VMD parameters, and also proving the effectiveness of constructing the objective function according to the ranked entropy algorithm, and the superiority of MPA optimized VMD parameters to achieve adaptive decomposition of load sequences.
3. Using a combination of LSTM and MPA-LSSVM forecasting methods, the experimental results show that the proposed model has better forecasting accuracy and generalization ability by comparing with traditional forecasting methods and forecasting weekdays and holidays in Spring, Summer, Autumn, and Winter.

AUTHOR CONTRIBUTIONS

Lingyun Wang: Writing – original draft; Conceptualization; Methodology; Writing – review & editing; Supervision. **Xiang Zhou:** Writing –original draft; Conceptualization; Methodology; Software; Data curation. **Honglei Xu:** Writing – review & editing; Investigation; Supervision. **Tian Tian:** Writing – review & editing; Investigation; Supervision. **Huamin Tong:** Investigation; Software.

ACKNOWLEDGEMENTS

This work was partially supported by the National Natural Science Foundation of China (51907104), Australian Research Council (LP160100528), 2022 CU Building Industry Project (CUB2022), and the 2022 Curtin University Science and Engineering Faculty Small Grant (SRG2022).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Lingyun Wang  <https://orcid.org/0000-0002-7501-7910>

REFERENCES

1. Ahajjam, M.A., et al.: Experimental investigation of variational mode decomposition and deep learning for short-term multi-horizon residential electric load forecasting. *Appl. Energy* 326, 119963 (2022). <https://doi.org/10.1016/J.APENERGY.2022.119963>
2. Moradzadeh, A., et al.: Electric load forecasting under False Data Injection Attacks using deep learning. *Energy Rep.* 8, 9933–9945 (2022). <https://doi.org/10.1016/J.EGYR.2022.08.004>
3. Huang, M., Jinghan Y.: Research on adversarial domain adaptation method and its application in power load forecasting. *Mathematics* 10(18), 3223 (2022). <https://doi.org/10.3390/MATH10183223>
4. Alsharekh, M. F., et al.: Improving the efficiency of multistep short-term electricity load forecasting via R-CNN with ML-LSTM. *Sensors* 22(18), 6913 (2022). <https://doi.org/10.3390/S22186913>
5. Zhou, F., et al.: Multi-step ahead short-term electricity load forecasting using VMD-TCN and error correction strategy. *Energies* 15(15), 5375 (2022). <https://doi.org/10.3390/EN15155375>
6. Zheng, C., et al.: GA– Reinforced deep neural network for net electric load forecasting in microgrids with renewable energy resources for scheduling battery energy storage systems. *Algorithms* 15(10), 338 (2022). <https://doi.org/10.3390/A15100338>
7. Dudek, G., Paweł P., Slawek S.: A hybrid residual dilated LSTM and exponential smoothing model for midterm electric load forecasting. *IEEE Trans. Neural Netw. Learn. Syst.* 33(7), 2879–2891 (2021). <https://doi.org/10.1109/TNNLS.2020.3046629>
8. Tong, X., et al.: LS-LSTM-AE: Power load forecasting via long-short series features and LSTM-Autoencoder. *Energy Rep.* 8, 596–603 (2022). <https://doi.org/10.1016/J.EGYR.2021.11.172>
9. Feng, Q., Suping Q.: Research on power load forecasting model of economic development zone based on neural network. *Energy Rep.* 7, 1447–1452 (2021). <https://doi.org/10.1016/J.EGYR.2021.09.098>
10. Yuan, J., et al.: Short-term electric load forecasting based on improved extreme learning machine mode. *Energy Rep.* 7, 1563–1573 (2021). <https://doi.org/10.1016/J.EGYR.2021.09.067>
11. Lee, E., Wonjong R.: Individualized short-term electric load forecasting with deep neural network based transfer learning and meta learning. *IEEE Access.* 9, 15413–15425 (2021). <https://doi.org/10.1109/ACCESS.2021.3053317>
12. Huang, J., et al.: A decomposition-based multi-time dimension long short-term memory model for short-term electric load forecasting. *IET Gener. Transm. Distrib.* 15(24), 3459–3473 (2021). <https://doi.org/10.1049/GTD2.12265>
13. Wang, G., et al.: A VMD-CISSA-LSSVM based electricity load forecasting model. *Mathematics* 10(1), 28 (2021). <https://doi.org/10.3390/MATH10010028>

14. Li, M., Xiaoming X., Du Z.: Improved deep learning model based on self-paced learning for multiscale short-term electricity load forecasting. *Sustainability* 14(1), 188 (2022). <https://doi.org/10.3390/SU14010188>
15. Bashir, T., et al.: Short term electricity load forecasting using hybrid prophet-LSTM model optimized by BPNN. *Energy Rep.* 8, 1678–1686 (2022). <https://doi.org/10.1016/J.EGYR.2021.12.067>
16. Brusaferrri, A., et al.: Probabilistic electric load forecasting through Bayesian mixture density networks. *Appl. Energy* 309, 118341 (2022). <https://doi.org/10.1016/J.APENERGY.2021.118341>
17. Hu, Y., et al.: Industrial artificial intelligence based energy management system: Integrated framework for electricity load forecasting and fault prediction. *Energy* 244, 123195 (2022). <https://doi.org/10.1016/J.ENERGY.2022.123195>
18. Obst, D., Joseph D.V., Yannig G.: Adaptive methods for short-term electricity load forecasting during COVID-19 lockdown in France. *IEEE Trans. Power Syst.* 36(5), 4754–4763 (2021). <https://doi.org/10.1109/TPWRS.2021.3067551>
19. Wu, J., et al.: Support vector regression with asymmetric loss for optimal electric load forecasting. *Energy* 223, 119969 (2021). <https://doi.org/10.1016/J.ENERGY.2021.119969>
20. Oreshkin, B.N., et al.: N-BEATS neural network for mid-term electricity load forecasting. *Appl. Energy* 293, 116918 (2021). <https://doi.org/10.1016/J.APENERGY.2021.116918>
21. Serrano Guerrero, X., et al.: A new interval prediction methodology for short-term electric load forecasting based on pattern recognition. *Appl. Energy* 297, 117173 (2021). <https://doi.org/10.1016/J.APENERGY.2021.117173>
22. Xinyu, W., Dou C., Yue D.: Electricity load forecast considering search engine indices. *Electr. Power Syst. Res.* 199 (2021). <https://doi.org/10.1016/J.EPSR.2021.107398>
23. Jawad, M., et al.: Machine learning based cost effective electricity load forecasting model using correlated meteorological parameters. *IEEE Access* 8, 146847–146864 (2020). <https://doi.org/10.1109/ACCESS.2020.3014086>
24. Zeng, P., et al.: Short-term power load forecasting based on cross multi-model and second decision mechanism. *IEEE Access.* 8, 184061–184072 (2020). <https://doi.org/10.1109/ACCESS.2020.3028649>
25. Jeong, D., et al.: Short-term electric load forecasting for buildings using logistic mixture vector autoregressive model with curve registration. *Appl. Energy.* 282, 116249 (2021). <https://doi.org/10.1016/j.apenergy.2020.116249>
26. Sulandari, W., et al.: Indonesian electricity load forecasting using singular spectrum analysis, fuzzy systems and neural networks. *Energy.* 190, 116408 (2020). <https://doi.org/10.1016/j.energy.2019.116408>
27. Hafeez, G., et al.: Electric load forecasting based on deep learning and optimized by heuristic algorithm in smart grid. *Appl. Energy* 269, 114915 (2020). <https://doi.org/10.1016/j.apenergy.2020.114915>
28. Bouktif, S., et al.: Multi-sequence LSTM-RNN deep learning and meta-heuristics for electric load forecasting. *Energies* 13(2), 391 (2020). <https://doi.org/10.3390/en13020391>
29. Ge, Q., et al.: Industrial power load forecasting method based on reinforcement learning and PSO-LSSVM. *IEEE Trans Cybern* 52(2), 1112–1124 (2020). <https://doi.org/10.1109/TCYB.2020.2983871>
30. Kong, X., et al.: Short-term electrical load forecasting based on error correction using dynamic mode decomposition. *Appl. Energy* 261, 114368 (2020). <https://doi.org/10.1016/j.apenergy.2019.114368>
31. Song, W., et al.: Multifractional Brownian motion and quantum-behaved particle swarm optimization for short term power load forecasting: An integrated approach. *Energy* 194 (2020). <https://doi.org/10.1016/j.energy.2019.116847>
32. Wu, Di, et al.: Multiple kernel learning-based transfer regression for electric load forecasting. *IEEE Trans. Smart Grid* 11(2), 1183–1192 (2019). <https://doi.org/10.1109/tsg.2019.2933413>
33. Wei, D., et al.: Research and application of a novel hybrid model based on a deep neural network combined with fuzzy time series for energy forecasting. *Energies* 12(18), 3588 (2019). <https://doi.org/10.3390/en12132467>
34. Zhang, J., et al.: Short term electricity load forecasting using a hybrid model. *Energy* 158, 774–781 (2018). <https://doi.org/10.1016/j.energy.2018.06.012>
35. Ghadimi, N., et al.: Two stage forecast engine with feature selection technique and improved meta-heuristic algorithm for electricity load forecasting. *Energy* 161, 130–142 (2018). <https://doi.org/10.1016/j.energy.2018.07.088>
36. Li, Y., et al.: Subsampled support vector regression ensemble for short term electric load forecasting. *Energy* 164, 160–170 (2018). <https://doi.org/10.1016/j.energy.2018.08.169>
37. Huang, N., et al.: A permutation importance-based feature selection method for short-term electricity load forecasting using random forest. *Energies* 9(10), 767 (2016). <https://doi.org/10.3390/en9100767>
38. Li, C., et al.: A least squares support vector machine model optimized by moth-flame optimization algorithm for annual power load forecasting. *Appl. Intell.* 45, 1166–1178 (2016). <https://doi.org/10.1007/s10489-016-0810-2>
39. Shen, Z., et al.: Model-independent approach for short-term electric load forecasting with guaranteed error convergence. *IET Control Theory Appl.* 10(12), 1365–1373 (2016). <https://doi.org/10.1049/iet-cta.2015.0818>

How to cite this article: Wang, L., Zhou, X., Xu, H., Tian, T., Tong, H.: Short-term electrical load forecasting model based on multi-dimensional meteorological information spatio-temporal fusion and optimized variational mode decomposition. *IET Gener. Transm. Distrib.* 17, 4647–4663 (2023). <https://doi.org/10.1049/gtd2.12992>