Forecasting power load: A hybrid forecasting method with intelligent data processing and optimized artificial intelligence

Yeming Dai[a*], Xinyu Yang[a], Mingming Leng[b]

a. School of Business, Qingdao University, Qingdao 200071, China (Yeming Dai, yemingdai@163.com; Xinyu Yang, 865050851@qq.com.)
b. Faculty of Business, Lingnan University, Hong Kong (Mingming Leng, mmleng@ln.edu.hk)

* Corresponding author.

A R T I C L E I N F O

Keywords:
Empirical mode decomposition
Minimal redundancy maximal relevance
Weighted gray relation projection algorithm
Second-order oscillation and repulsion particle swarm optimization
Power load forecasting

A B S T R A C T

An accurate power load prediction in smart grid plays an important role in maintaining the balance between power supply and demand and thus ensuring the safe and stable operation of power system. In this paper we develop a hybrid power load prediction method, which involves three main steps: data decomposition with the empirical mode decomposition method, data processes with the minimal redundancy maximal relevance method and the weighted gray relationship projection algorithm, and support vector machine prediction, whose parameters are optimized through the particle swarm optimization algorithm with a second-order oscillation and repulsive force factor. Moreover, we predict the power load with our hybrid forecasting method based on the real dataset from the electricity market in Singapore, and also compare our prediction results with those by using other forecasting methods. Our comparison results show that our novel hybrid method possesses a high accuracy in both the level and directional predictions.

1. Introduction

In recent years, people's awareness of environmental protection has continued to increase, and their interest in designing more efficient energy usage patterns has also grown (Wang et al., 2021; Zhu et al., 2020). With the development of modern technology, the real-time exchange of energy information becomes possible, and the smart grid appears as a new type of grid that can improve energy efficiency (Milchram et al., 2020). Through demand response, smart grid helps the power system reduce peak loads, achieve a balance between energy supply and demand, and encourage the users to consume more cautiously and rationally (Tang et al., 2019). Therefore, the accurate prediction of power load in smart grid plays an important role in maintaining the balance between power supply and demand, so as to ensure the safe and stable operation of power system (Yang et al., 2019).

As renewable energy tends to be increasingly popular in the global power grid, there is a large scale of renewable energy connected to the grid and a large influx of new load terminals such as electric vehicles, which has imposed a severe test on the balance of power system supply and demand and increased the difficulty of power load forecasting (Yuan et al., 2021). As a result, more advanced forecasting techniques begin to emerge for improving the accuracy of forecasting (Hong et al., 2020). But, the prediction results from traditional forecasting models——each usually involving a single prediction method——are inaccurate. It behooves us to find a novel hybrid forecasting model for the more accurate prediction of power load.

For many forecasting problems in power systems, we need to process the data decomposition (Bessec and Fouquau, 2018; Angelopoulos et al., 2019), feature selection (Ding et al., 2020; Yang et al., 2019; Niu et al., 2020), parameter optimization of forecasting model (Zhu et al., 2017; Jiang et al., 2020), and others. At present, the hybrid load forecasting models are mostly based on artificial intelligence methods. For example, support vector machine (SVM) (Li et al., 2021) and artificial neural network (ANN) (Yang et al., 2019; du Jardin, 2021) are used in many applications. However, when we deal with time series problems, we cannot obtain stable results due to the historical dependence and overfitting of ANN while SVM can effectively overcome the limitations of ANN because the SVM minimizes training and generalization errors by using empirical risk and structural risk minimization principles (Hafeez et al., 2021). In particular, SVM, capable of solving nonlinear problems and locating the global optimal solution, is widely used in power load forecasting (Ahmad and Chen, 2019). For example, Singh and Mohapatra (2021) used a hybrid forecasting model—based on wavelet transform (WT) and SVM—for power load forecasting which effectively improved the accuracy of load forecasting. Barman and Choudhury (2020) used the gray wolf optimizer (GWO) algorithm to optimize the parameters of SVM, which effectively improved the load forecasting accuracy, and the comparison results indicate that the prediction effects of SVM before and after optimization are better than that of ANN. Therefore, as an excellent prediction method, SVM has been adopted by many relevant researchers.

1.1. Data decomposition

The empirical mode decomposition (EMD) (Zhang et al., 2019; Xie et al., 2020; Liu et al., 2021) method is one of data decomposition methods most commonly used suitable for nonlinear and non-stationary time series, which can be decomposed into several intrinsic mode functions (IMFs) and a residual (Thomas et al., 2020). For example, Yaslan and Bican (2017) predicted the direction movement of power load demand based on EMD-SVR (support vector regression), and evaluates the performance of three load data sets. The results compared with a single SVR exposed that the EMD-SVR method is superior to a single SVR in directional measurement. Liu et al. (2018) proposed a short-term load forecasting model based on hybrid fuzzy combination weights-EMD and Kalman filtering-bat algorithm-SVM, and the simulation proved that the forecasting accuracy was effectively improved. A hybrid forecasting model based on ICEEMDAN (improved complete ensemble empirical mode decomposition with adaptive noise)-PSO-SVR in Al-Musaylh et al. (2018) was developed, where ICEEMDAN is an improved version of EMD with adaptive noise. The above researches apply the EMD method to the SVM power load forecasting from different aspects, and obtain more accurate prediction value than when EMD is not applied. In practical applications, EMD can be appropriately adjusted for different power markets, and also be applied in this paper.

1.2. Feature selection

The selection of features has a significant impact on the accuracy of power load forecasting. Most of existing feature selection methods are linear feature analysis methods; but, the prediction variables are non-linear mapping functions of their input variables. Therefore, one more advanced feature analysis method based on mutual information (MI) is proposed (Sharmin et al., 2019). Among the MI methods, the minimal redundancy maximal relevance (mRMR) proposed by Abedinia et al. (2016) involves the interactive modeling of feature selection, in addition to the correlation and redundancy based on information theory standards. The mRMR method has been widely used by, e.g., Che et al. (2017), Xiao et al. (2019), and Liang et al. (2019).

In addition, when we carry out the feature selection of load forecasting, it is not only necessary to select an appropriate feature selection method but also to select an appropriate method for the distinction between holidays and non-holidays (Arora and Taylor, 2018). Existing researches often select data of similar holidays as their historical data; however, using the same historical data for both holidays and non-holidays is not conducive to improving the prediction accuracy. To solve this problem, Wu et al. (2015) proposed a combination method for the short-term load forecasting, according to a gray projection-improved random forest algorithm. The set of similar days to be predicted was selected through the weighted gray relationship projection (WGRP) algorithm. The above combination method was also compared with the random forest algorithm without gray projection. The results show that the

WGRP algorithm can effectively improve the accuracy of load forecasting.

1.3. Parameter optimization

The common optimization algorithms for forecasting models include the genetic algorithm (GA), ant colony algorithms, particle swarm optimization (PSO) algorithm, artificial fish swarm algorithm, etc., among which the PSO approach and its improved versions have been the most widely used by, e.g., Zhu et al. (2018), Jiang et al. (2020), Kouziokas (2020), and Xie et al. (2020). However, the basic PSO algorithm is easy to fall into the local optimum during the particle search process, and there are shortcomings such as (i) the fast convergence speed in the initial search of the algorithm and a slower convergence speed in the later search periods, and (ii) the randomness of parameter selection (Ding et al., 2019). Hence, some researchers (e.g., Masoumi et al., 2020) have improved the PSO algorithm based on the characteristics of the power load data.

Furthermore, we note that a key to the SVM model-based prediction method is about the parameter selection of the SVM model and its kernel function. Therefore, when SVM is used for power load forecasting, we usually need to combine some optimization algorithms to find the optimal parameters and the kernel function. The improvements for the PSO method combined with SVM were made by, e.g., Zeng et al. (2018) and Jiang et al. (2020). Especially, Sun et al. (2017) proposed an improved PSO algorithm with a second-order oscillation and repulsive force factor (SecRPSO), and experimentally verified that the SVM prediction method based on this algorithm is more stable and accurate than those that are based on the network optimization and basic PSO optimization.

Through the summary of the above literatures, the mixed prediction model has better prediction effect than the single prediction model. Therefore, we develop a novel hybrid power load forecasting method. First, in order to reduce the complexity of prediction, we decompose the original load data to obtain time series that are simpler and easier to predict than the original load data. Secondly, for the IMFs and residual decomposed by EMD, the mRMR is used to select features that are highly correlated with the power load and have no redundancy. Then, the weighted gray relationship projection algorithm is applied to select the historical load sequence of the holiday for prediction, so that the historical data of the holiday is more general. Finally, the PSO with a second-order oscillation and repulsive force factor algorithm is used to optimize the parameters of SVM model to further improve the prediction accuracy of developed model.

We apply our method to the prediction of power load in Singapore, using an electricity market dataset in Singapore (Mendeley Data 2020). According to our calculation of the prediction accuracy for the method, we conclude that our method is significantly effective to forecasting. The main contributions of this paper are listed as follows:

(1) We propose a novel hybrid forecasting method for power load, combining the EMD, mRMR, WGRP, and SVM prediction models with the optimized PSO algorithm.

(2) We use the EMD method to decompose the original load data. The mRMR and WGRP algorithms are used to extract features from original data and process holiday load, respectively.

(3) The parameters of SVM prediction model are optimized by the PSO algorithm with a second-order oscillation and repulsive force factor.

(4) We measure the accuracy of the level prediction and direction prediction to demonstrate the effectiveness of our hybrid prediction method.

The rest of this paper is organized as follows: Section 2 introduces the methods involved in the model and describes the detailed process of power load forecasting in the proposed model. Section 3 verifies the prediction accuracy of the model through case experiments. Finally conclusions are drawn and future research direction is pointed out in Section 4.

2. Methodology

In this section, the proposed hybrid forecasting method for power load is introduced. Section 2.1 briefly describes the EMD data decomposition approach. Then, the basic knowledge of mRMR for feature selection is provided in Section 2.2. Section 2.3 gives the process of the weighted gray relationship projection algorithm for finding out the historical sample set similar to the holiday sample to be predicted. Section 2.4 introduces the knowledge of the original SVM forecasting model in this paper. Finally, the developed second-order oscillation and repulsive force factor algorithm is introduced to improve the diversity of the particle population and avoid premature stagnation.

2.1. The empirical mode decomposition (EMD) method

The EMD is an adaptive data analysis method that can effectively obtain valid information of nonlinear and non-stationary time series and decompose the time series into several intrinsic mode functions (IMFs) and a residual (Thomas et al., 2020). IMF is a function that satisfies the following two conditions. First, in the entire dataset, the number of extreme points and the number of zero points are identical or differ by at most 1. Second, at any point, the mean of the envelope defined by the local maximum and local minimum is zero (Thomas et al., 2020). Given a time series, the specific decomposition process of EMD involves the following steps:

(1) Determine the maximum and minimum points of the original sequence $x(t)$, and use the cubic spline function to generate the upper and lower envelopes of $x(t)$.

(2) Compute the average of these two envelopes to obtain the average envelope $m_1(t)$.

(3) Subtract $m_1(t)$ from the original sequence $x(t)$ to find a new sequence $h_1(t)=x(t)-m_1(t)$.

(4) Check the character of $h_1(t)$: if $h_1(t)$ is an IMF, then $h_1(t)$ is the first IMF. Otherwise, if $h_1(t)$ is not an IMF, then repeat (1)-(3) until there is no extreme value in $h_k(t)$ after k iterations. That is, $h_k(t)=h_{k-1}(t)-m_k(t)$ is the first IMF.

(5) Subtract the first IMF from the original sequence $x(t)$, and repeat steps (1)-(4) for the remaining sequence until the sequence obtained after multiple decompositions is a monotonic function or the amplitude of remaining sequence is less than preset value. After all IMFs are removed from the original sequence, the remaining residual value represents the overall trend of the original sequence. This indicates that the decomposition of the EMD algorithm is completed, and original sequence $x(t)$ is decomposed into a superposition of several IMFs and a residual.

The EMD method has been widely used in practice. To enhance the predictive accuracy of tourism demand forecasting, Xie et al. (2020) developed a decomposition-ensemble approach based on the complete/ensemble EMD with adaptive noise, data characteristic analysis, and the Elman's neural network model. The empirical results of Hong Kong tourism demand showed that the proposed model outperformed other models in both point and interval forecasts for different prediction horizons. Liu et al. (2021) proposed a wind speed prediction model combining EMD with some novel recurrent neural networks (RNN) and the autoregressive integrated moving average (ARIMA). In the model, the EMD was used to decompose the wind speed sequence to reduce the complexity and non-stationary of the series. The prediction results indicate that the EMD method combined with long short-term memory network (LSTM) can improve the wind speed prediction performance. The hybrid forecasting model by Al-Musaylh et al. (2018) is based on ICEEMDAN-PSO-SVR, where ICEEMDAN is the abbreviation of improved complete ensemble empirical mode decomposition with adaptive noise, and it is an improved version of EMD with adaptive noise. The above indicates that with the EMD method, we can obtain a more accurate prediction than when EMD is not used.

2.2. The minimal redundancy maximal relevance (mRMR) method

The selection of features has a significant impact on the forecasting accuracy. Many researchers have used some selection methods to choose features. When the prediction variable is a non-linear mapping function of input variables, the mutual information-based methods are mostly used. The mutual information refers to the intersection of two or more random variables which reflect the information contained between the variables. That is, the mutual information can characterize the degree of correlation between two random variables x and y, which can be computed as

$$I(x,y)=\iint p(x,y)\log\frac{p(x,y)}{p(x)p(y)}dxdy \tag{1}$$

where x and y are random variables, $p(x)$ and $p(y)$ are marginal probability density functions, and $p(x,y)$ is joint probability density function. Using the mutual information, Abedinia et al. (2016) proposed the minimal redundancy maximal relevance (mRMR) method, which involves the interactive modeling of feature selection, in addition to the correlation and redundancy based on information theory standards. As the mRMR method can capture the linear and non-linear relationship between two variables during prediction, it has been widely used in practice. For example, Liang et al. (2019) applied the mRMR method for the feature selection of short-term power load forecasting, and obtained accurate prediction results.

2.2.1. The maximal relevance

The maximal relevance in mRMR requires the largest mean of MI between the selected feature and the target variable, i.e.,

$$\max\ D=\frac{1}{|F|}\sum_{x_i\in F} I(x_i,y) \tag{2}$$

where D means the mean of MI, $|F|$ represents the number of features in the feature set F, y is the target variable, and x_i ($i=1,2,\cdots,|F|$) represents the i-th feature in F.

2.2.2. The minimal redundancy

The features selected only by maximal relevance may contain duplicate information, which means that there is a certain degree of redundancy. This increases the prediction error and the calculation complexity. Therefore, the mRMR method introduces a minimum redundancy to the selected features such that the correlation between features x_i and x_j ($i,j=1,2,\cdots,|F|$) is the smallest, i.e.,

$$\min\ R=\frac{1}{|F|^2}\sum_{x_i,x_j\in F} I(x_i,x_j) \tag{3}$$

where R represents the dependency between x_i and x_j.

Integrating (2) and (3) and transforming them, we obtain the objective function of mRMR as

$$mRMR:\ \max_{x_j\in J_m-F_{n-1}}\left[I(x_j,y)-\frac{1}{n-1}\sum_{x_i\in F_{n-1}} I(x_j,x_i)\right] \tag{4}$$

where J_m represents the candidate feature set, and m is the number of features in J_m.The objective function in (4) can be solved by using the incremental search, in which we can select the features that maximize it from the candidate sample feature set J_m one by one. That is, if n features need to be selected, then F_{n-1} is the set of n-1 features that are selected from J_m, and then the n-th feature is the feature x_j that maximizes the objective function (4) in set $\{J_m-F_{n-1}\}$.

2.3. The weighted hray telationship projection (WGRP) algorithm

For the power load forecasting, the generality of historical data has a great impact on prediction results. Using the same historical data selection method for both holidays and non-holidays is not conducive to improving the prediction accuracy.

The WGRP algorithm introduces the concept of weighting and projection based on gray correlation coefficient. Through appropriate weighting methods, the WGRP algorithm finds the key factors with the greatest influence. Then, this algorithm obtains the correlation between the historical samples and the samples to be predicted, by combining the projection of historical samples on the samples to be predicted. This helps find a historical sample set similar to the samples to be predicted (Dai and Zhao, 2020). Such approach not only ensures the data generality but also reduces the time to select the training data.

We describe the WGRP algorithm below.

(1) The feature vector of the sample to be predicted is Y_0. Select historical data of n_1 samples before the sample to be predicted, and the feature vector of the i-th sample is Y_i.

$$Y_0=[y_{01},y_{02},\cdots,y_{0m_1}] \tag{5}$$

$$Y_i=[y_{i1},y_{i2},\cdots,y_{im_1}],\quad i=1,2,\cdots,n_1 \tag{6}$$

where m_1 means the number of influencing factors.

(2) Treating Y_0 as the parent sequence and Y_i as the subsequence, we calculate the correlation coefficient between Y_i and Y_0, and construct a gray correlation matrix as follows:

$$G=\begin{bmatrix} G_{01} & \cdots & G_{0m_1} \\ \vdots & \ddots & \vdots \\ G_{n_1 1} & \cdots & G_{n_1 m_1} \end{bmatrix} \tag{7}$$

where G_{jk} ($j=0,1,\cdots,n_1$ and $k=1,2,\cdots,m_1$) represents the gray correlation coefficient of the k-th factor in the j-th sample.

(3) We calculate the weight of each influencing factor using the entropy method, obtain the weight vector γ,

$$\gamma=[\gamma_1,\gamma_2,\cdots,\gamma_{m_1}] \tag{8}$$

and then weight the gray correlation matrix as

$$G'=G\gamma^{\mathrm{T}}=\begin{bmatrix} \gamma_1 & \cdots & \gamma_{m_1} \\ \vdots & \ddots & \vdots \\ \gamma_1 G_{n_1 1} & \cdots & \gamma_{m_1} G_{n_1 m_1} \end{bmatrix} \tag{9}$$

(4) We regard each row in G' as a row vector. The first row is the row vector of the sample to be predicted, denoted as $G_0^{'}$, and other historical sample row vector are denoted as $G_i^{'}$, for $i=1,2,\cdots,n_1$. The angle between each $G_i^{'}$ and $G_0^{'}$ is viewed as the gray projection angle of the sample, denoted as θ_i. We then have

$$\cos\theta_i=\frac{\sum_{j=1}^{m_1}\gamma_j G_{ij}\gamma_j}{\sqrt{\sum_{j=1}^{m_1}(\gamma_j G_{ij})^2}\sqrt{\sum_{j=1}^{m_1}\gamma_j^2}} \tag{10}$$

Hence, the weighted gray correlation projection value between each historical sample and the sample to be predicted is calculated as

$$D_i=\frac{\sum_{j=1}^{m_1}\gamma_j G_{ij}\gamma_j}{\sqrt{\sum_{j=1}^{m_1}\gamma_j^2}} \tag{11}$$

(5) We sort the weighted gray correlation projection values of each historical vector from the largest to the smallest. A number of samples with the largest projection value are selected to form a similar sample set.

2.4. The support vector machine (SVM) method

The SVM method can be used to solve the classification and regression problems in machine learning with optimization methods (Li et al., 2021). A key to the SVM model-based prediction is about the parameter selection of the SVM model and its kernel function. Therefore, when we use the SVM model for prediction, we usually combine some optimization algorithms to find optimal parameters.

Solving regression problems such as a load forecasting is to find the function $y=f(x)=\omega^T x+b$ for the purpose of inferring the value of y corresponding to any x, where ω is the weight and b is the deviation. When the SVM method solves regression problems, the objective function is

$$\min \quad \varepsilon$$
$$\text{s.t.} \quad -\varepsilon < \left(\omega^T x_i\right) + b - y_i < \varepsilon \quad i = 1,\cdots,l \tag{12}$$

As for nonlinear regression problems, nonlinear mapping is usually used to map the input space into a linear and separable Hilbert space. Then, we can transform the minimization problem in (12) to a binary classification problem. Suppose that after nonlinear mapping, the prediction model is

$$f(x) = \omega^T \phi(x) + b \tag{13}$$

In order to avoid complex operations in a high-dimensional space, the kernel function is introduced, and the prediction model (13) can be transformed into

$$y = f(x) = \sum_{i=1}^{l} \left(\bar{\alpha}_i - \bar{\alpha}_i^*\right) K\left(x_i, x\right) + \bar{b} \tag{14}$$

where α_i and α_i^* denote Lagrange multipliers which are used to solve the binary classification problem converted from (12) and obtain the optimal solution $\bar{\alpha}_i$ and $\bar{\alpha}_i^*$.

The radial basis kernel function is a widely-used kernel function with calculation convenience and strong stability to solve nonlinear problems. The expression is

$$K\left(x, y\right) = \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right) \tag{15}$$

where σ denotes a width coefficient. The function in (15) can transform the prediction model (13) to (14).

2.5. The particle swarm optimization (PSO) algorithm with a second-order oscillation and repulsive force factor

The common optimization algorithms include the GA, ant colony algorithms, PSO algorithm, artificial fish swarm algorithm, etc., among which the PSO approach and its improved versions have been the most widely used (Xie et al., 2020). However, the basic PSO algorithm is easy to fall into the local optimum during the particle search process, and there are shortcomings such as (i) the fast convergence speed in the initial search of the algorithm and a slower convergence speed in the later search periods, and (ii) the randomness of parameter selection (Ding et al., 2019).

Hence, in the power load prediction, some researchers have improved the PSO algorithm based on the characteristics of the power load data. For example, Zhao et al. (2020) improved the PSO algorithm from both cognitive/social coefficients and additional recovery operators to enhance its search ability. Masoumi et al. (2020) used a weighted, improved PSO algorithm to solve an optimization problem in the renewable energy load forecasting.

Hence, a second-order oscillating and repulsion factor on the basis of the basic PSO algorithm is developed to improve the diversity of particle population and avoid premature stagnation and local optimization. In the M dimensional search space, we generate a population of n particles randomly. Letting V and U denote the speed and position of particles in the search space, respectively, we can calculate the corresponding fitness value by solving an objective function. $P_i = (p_{i1}, p_{i2}, \cdots, p_{iM})$ and $P_g = (p_{g1}, p_{g2}, \cdots, p_{gM})$ mean the individual and global extremes, respectively. In the iteration process, the particle speed depends on the position information of the current particle and the particle in the previous iteration. We develop the equation for particle speed as follows:

$$\begin{aligned} V_{im}^{k+1} &= wV_{im}^k + c_1 r_1 (P_{im}^k - (1+\beta_1)U_{im}^k + \beta_1 U_{im}^{k-1}) \\ &\quad + c_2 r_2 (P_{gm}^k - (1+\beta_2)U_{im}^k + \beta_2 U_{im}^{k-1}) \end{aligned} \tag{16}$$

where $m = 1, 2, \cdots, M$; $i = 1, 2, \cdots, n$; k is the current number of iteration; c_1, $c_2 > 0$ are acceleration factors; r_1, r_2 are random numbers between 0 and 1; and $\beta_j < 2\sqrt{c_j - 1} / c_j$ in the early period and $\beta_j \geq 2\sqrt{c_j - 1} / c_j$ in the later period. Moreover, in (16), w denotes the inertia weight coefficient, which decreases non-linearly with the number of iterations, and it is computed by

$$w = w_{ini} - (w_{ini} - w_{end})(k / k_{\max})^2 \tag{17}$$

where w_{ini} and w_{end} represent the initial value and the termination value of w, respectively. The particle updates its position by

$$U_{im}^{k+1} = \begin{cases} U_{im}^k + V_{im}^{k+1} + 2d^k, & \sum\limits_{i \neq j, j=1}^{n} \| u_{im}^k - u_{jm}^k \| < d^k \\ U_{im}^k + V_{im}^{k+1}, & \sum\limits_{i \neq j, j=1}^{n} \| u_{im}^k - u_{jm}^k \| \geq d^k \end{cases} \tag{18}$$

where $d^k = d_{ini} - (d_{ini} - d_{end})(k / k_{\max})^2$ is the allowed minimum distance between particles; and, d_{ini} and d_{end} represent the initial and termination value of d^k, respectively.

2.6. The proposed power load forecasting method

In this subsection, the framework of the proposed hybrid forecasting model for power load is presented. The proposed power load prediction method shown in Figure 1 is divided into four stages.

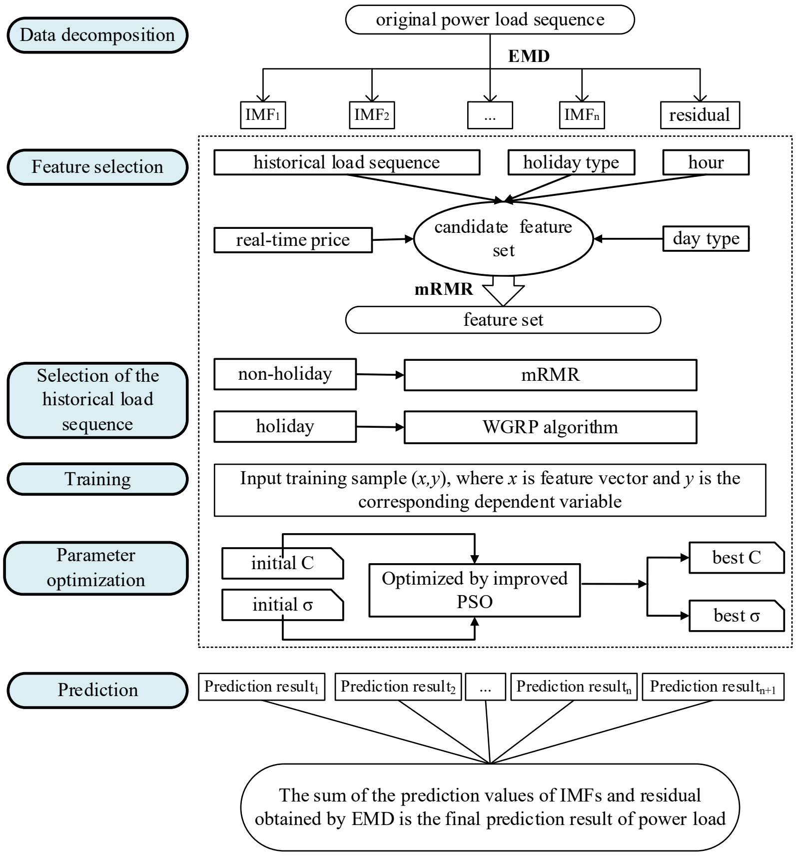

Fig. 1. The process of load forecasting based on the proposed model.

2.6.1. Data preprocess

To simplify the feature selection and reduce the difficulty in solving the prediction model, we consider three time attributes: (1) day attribute set $W=\{1,2,3,4,5,6,7\}$, where the values correspond to 7 days in one week; (2) holiday attribute set $H=\{0,1\}$, where 0 means non-holiday and 1 means holiday; (3) hour attribute set $h=\{1,2,\cdots,24\}$, where the values represent 24 hours in one day.

In view of the fact that the historical data of holidays is not general, we use the WGRP algorithm to process the holiday data, so as to eliminate the impact of holidays on load forecasting.

2.6.2. Prediction steps

Stage 1: Data decomposition. The original power load time series is decomposed by the EMD technique into several IMFs and residual with different frequencies. The number of modes for the power load sequence can be adjusted based on experimental evidence.

Stage 2: Component identification. The original power load sequence is decomposed into different data characteristics. To distinguish these components, the mRMR algorithm is used to selected features and apply the WGRP algorithm to process the holiday data.

(1) The mRMR method is applied to select the features that are highly correlated with the prediction variable for the IMFs and residual decomposed by the EMD method, which results in no redundancy. Then the features that contain rich information from two aspects of relevance and redundancy are selected.

(2) The WGRP algorithm is used to process the holiday data, mainly by selecting the similar sample set to be predicted, and the load data of similar samples constitute the historical load sequence of the holiday, which is made more general.

Stage 3: Component prediction. The abovementioned components could be determined by the SVM model. The SVM model is applicable to forecasting nonlinear sequences, so the component can be determined by the SVM model. We train the SVM model by using the historical data, and also optimize the parameters of the SVM model by using the PSO algorithm with second-order oscillating and repulsion factor.

Stage 4: Aggregating the prediction results. Each component is predicted based on the identified data characteristics. Then, the prediction results of all components are aggregated to obtain the final forecasting result. Finally, the power load forecasting is carried out via the proposed prediction method. We plot Figure 1 to depict the specific process for the forecasting method.

2.6.3. Model evaluation

To decide on whether our forecasting model can accurately predict the power load or not, we comprehensively evaluates the prediction accuracy of developed the model from directional prediction and level prediction. Mean absolute percent error (MAPE) and mean absolute error (MAE) are used as evaluation criteria for model level prediction, directional prediction statistic (D_{state}) is used as the evaluation criterion for direction prediction. Specifically,

$$\mathrm{MAPE}=\frac{1}{n}\sum_{i=1}^{n}\frac{\left|Y_i-\hat{Y}_i\right|}{\left|Y_i\right|} \tag{19}$$

$$\mathrm{MAE}=\frac{1}{n}\sum_{i=1}^{n}\left|Y_i-\hat{Y}_i\right| \tag{20}$$

$$\mathrm{D}_{state}=\frac{1}{n}\sum_{i=1}^{n}a_i,\quad a_i=\begin{cases}1, & if\left(Y_i-Y_{i-1}\right)\left(\hat{Y}_i-Y_{i-1}\right)\geq 0\\ 0, & otherwise\end{cases} \tag{21}$$

where Y_i means the real value of power load, and $\hat{Y}_i$ represents the predicted value.

In addition, we compare the prediction performances of different models from the perspective of statistics. The Diebold-Mariano (DM) test (Diebold and Mariano, 1995) is used to further analyze the prediction performance. In the DM test, the null hypothesis is that the prediction accuracy of the two models is equal. Furthermore, we use the mean square prediction error (MSPE) as the loss function. The DM test is defined as

$$Z_{DM}=\frac{\bar{g}}{\sqrt{\hat{V}_{\bar{g}}/T}}\sim N(0,1),\ \ T\to\infty \tag{22}$$

where

$\bar{g}=\frac{1}{T}\Sigma_{t=1}^{T}g_t$,

$g_t=\sum_{t=1}^{T}\left(y_t-\hat{y}_{te,t}\right)^2-\left(y_t-\hat{y}_{re,t}\right)^2$,

$\hat{V}_{\bar{g}}=\gamma_0+2\sum_{j=1}^{\infty}\gamma_j$,

and

$y_j=\text{cov}(g_t,g_{y-1})$.

Among them, $\hat{y}_{te,t}$ and $\hat{y}_{re,t}$ represent the predicted values calculated according to y_t at time t by the test model (TE) and the reference model (RE), respectively. $\hat{V}_{\bar{g}}$ is the consistent estimation usually defined as the progressive long-term variance. In the DM test, if the absolute value of Z_{DM} is less than 1.96 or the significance level is less than 5%, then the null hypothesis is accepted. Otherwise, the hypothesis is rejected.

We also use the rate test (RT) and the Pesaran-Timmermann test (PT) to test the accuracy of the direction test. The RT test is

$$Z_{RT}=\frac{d_{te}-d_{re}}{\sqrt{\frac{d_{te}(1-d_{te})}{T}+\frac{d_{re}(1-d_{re})}{T}}}\sim N(0,1),\quad T\to\infty \tag{23}$$

where d_{te} and d_{re} represent the direction prediction accuracy of RE and TE, respectively. The null hypothesis of the RT test is that the direction prediction accuracy of TE and RE is equal. Using the two-sided test, the null hypothesis was accepted at the 5% significance level when the absolute value of Z_{RT} is below 1.96.

The PT test is

$$Z_{PT}=\left[\frac{P^*(1-P^*)}{T}\right]^{-0.5}\left(\hat{P}-P^*\right)\sim N(0,1),\quad T\to\infty \tag{24}$$

where

$\hat{P}=\frac{1}{T}\sum_{t=1}^{T}H_t\left[(Y_i-Y_{i-1})(\hat{Y}_i-Y_{i-1})\right]$,

$P^*=\hat{P}_1P_1+(1-\hat{P}_1)(1-P_1)$,

$P_1=\frac{1}{T}\sum_{t=1}^{T}H_t\left[(Y_i-Y_{i-1})\right]$,

$\hat{P}_1=\frac{1}{T}\sum_{t=1}^{T}H_t\left[(\hat{Y}_i-Y_{i-1})\right]$,

and

$$H(Y)=\begin{cases}1, & Y\geq 0\\ 0, & Y<0\end{cases}$$

In the PT test, the null hypothesis is the assumption that the predicted and actual directions of the model are independent of each other. For a two-sided test, the null hypothesis is accepted if the absolute value of Z_{PT} is below 1.96 or the corresponding p-value is below 5%.

3. Case study

In this section, the case study using the time series of power load is described. To perform the illustration and verification, the prediction results and the comparison analysis are also presented.

We use an electricity market dataset in Singapore (from the Mendeley Data 2020) for our prediction, with an aim to illustrate the accuracy of our model. The dataset contains the power load, real-time electricity price, legal holiday time, and other data of the Singapore power market from January 1, 2017 to March 31, 2018. The data from January 1 to December 31 in the year of 2017 is used for model training, and the data from January 1 to March 31 in the year of 2018 is used for model verification. Next, we use our forecasting model to predict the power load of Singapore in the time interval from March 25, 2018 to March 31, 2018, during which March 31, 2018 is a legal holiday in Singapore. It is worth noting that we consider an one-step ahead forecasting, which is good to the short-term prediction.

3.1. Data

We first use the EMD method to decompose the original power load data, and obtains 10 IMFs and a residual as shown in Figure 2. We learn from Figure 2 that as a result of using the EMD method, the complex time series become simple and features become obvious. This thereby reduces the difficulty and complexity for prediction. We then perform the feature selection on the 11 sequences that follows the EMD-based decomposition. Based on previous relevant researches and the actual situation in Singapore (Dai and Zhao, 2020), we choose historical load, real-time price, and time attributes as candidate features, as described in Table 1.

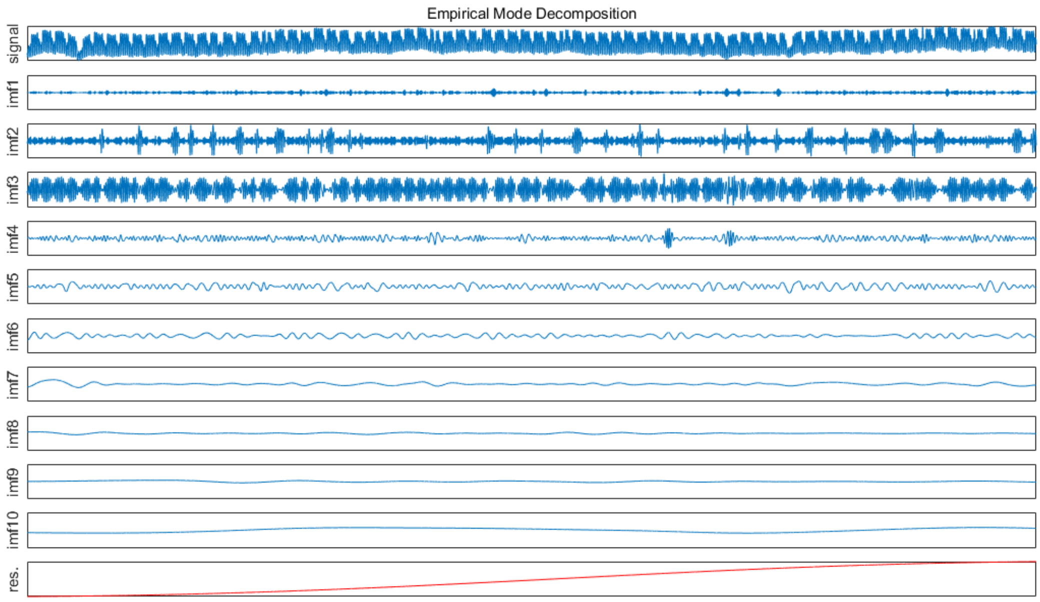

Fig. 2. Decomposition result of original load data based on EMD.

For the purpose of reducing the complexity of algorithm, this paper takes 1 hour as the time interval, and the time interval of real-time price sampling value in Singapore is 0.5 hour, so the first value of the hourly sampling value is selected as the real-time electricity price at that moment, namely,

$$P_t=p_{2t\text{-}1} \tag{22}$$

where P_t represents the price at time t, and p_{2t-1} is the sample value.

For the 11 time series after decomposition, we perform the feature selection separately through mRMR, and show our final results as in Table 2. We learn from Table 2 that the feature sets of each sequence obtained after the power load data decomposition are mostly composed of historical load sequences and real-time prices. When the time series is simpler, the number of features becomes fewer. Furthermore, the time series after the sixth IMF has only one feature (i.e., the load data in the hour before prediction). For each hour of the holiday to be predicted, the historical load sequence selected by mRMR is not general. Therefore, we apply the WGRP algorithm to select the same number of similar samples in the above-mentioned feature set to form the historical load sequence. For example, the number of historical load data in the feature set of IMF5 is 5. Then, we select five similar samples in the hour to be predicted to form the historical load sequence in the prediction hour.

Table 1
Candidate features.

Symbol	Description
$L_i, i=1,2,\cdots,168$	Power load at each hour of the 7 days before the hour to be predicted
H、$H_i, i=1,2,\cdots,168$	Holiday type of the hour to be predicted and holiday type of the 7 days before it
W、$W_i, i=1,2,\cdots,168$	Day type of the hour to be predicted and day type of the 7 days before it
P、$P_i, i=1,2,\cdots,168$	Real-time price of the hour to be predicted and real-time price of the 7 days before it
h、$h_i, i=1,2,\cdots,168$	The hour to be predicted and hour of the 7 days before it

Table 2
Features of each IMF and residual.

Time series	Feature set
IMF_1	$L_{168}, P_7, L_{166}, L_{25}, L_1, L_{163}, L_{10}, L_{145}, L_{127}, L_{97}, H_{165}, L_{19}, L_{112}, L_{161}, L_{88}$
IMF_2	$L_{163}, P_{100}, L_{73}, L_{155}, H_{152}, L_{109}, L_{165}, L_{157}, L_{11}, L_{101}, L_{164}, L_{128}, L_{121}, L_{168}$
IMF_3	$L_{51}, P_{11}, L_{17}, L_{147}, L_{127}, L_{157}, L_{161}, L_{80}, L_{110}, L_{158}, L_{13}, L_{77}, L_{124}, L_{164}$
IMF_4	$L_{168}, P_{25}, L_{148}, L_{102}, L_{167}, L_{144}, L_{137}, L_{17}, L_{34}, L_{166}, L_{94}, L_{157}, L_{121}$
IMF_5	$L_{168}, P_{90}, L_{165}, L_{88}, L_{167}, L_{128}$
IMF_6	$L_{168}, P_{69}, L_{166}, L_{81}, L_{167}, L_{128}, L_{34}$
IMF_7- IMF_{10}	L_{168}
residual	L_{168}

3.2 Results comparison and error analysis

To further verify our forecasting model, we compare the prediction results by using our hybrid forecasting method (which can be viewed as the "base" Model for our comparisons) with those by using the five forecasting models as described as follows: (a) Model 1 (Dai and Zhao, 2020) that uses the mRMR-WGRP and SecRPSO-SVM methods; (b) Model 2 that contains the EMD and mRMR-WGRP-BPNN methods; (c) Model 3 that involves the mRMR-WGRP-BPNN method in which BPNN is the abbreviation of the term "back propagation neural network"; (d) Model 4 that applies the mRMR-WGRP and GA-SVM methods; and (e) Model 5 that consists of the EMD, mRMR-WGRP, and GA-SVM methods.

We depict our prediction results in Figure 3 to present the difference between the results by our hybrid forecasting model and those by Models 1-5. Moreover, we evaluated the prediction performance of different models from both level and directional prediction perspectives, as given in Table 3. Tables 4-6 are the calculation results of RT test, DM test and PT test for different models.

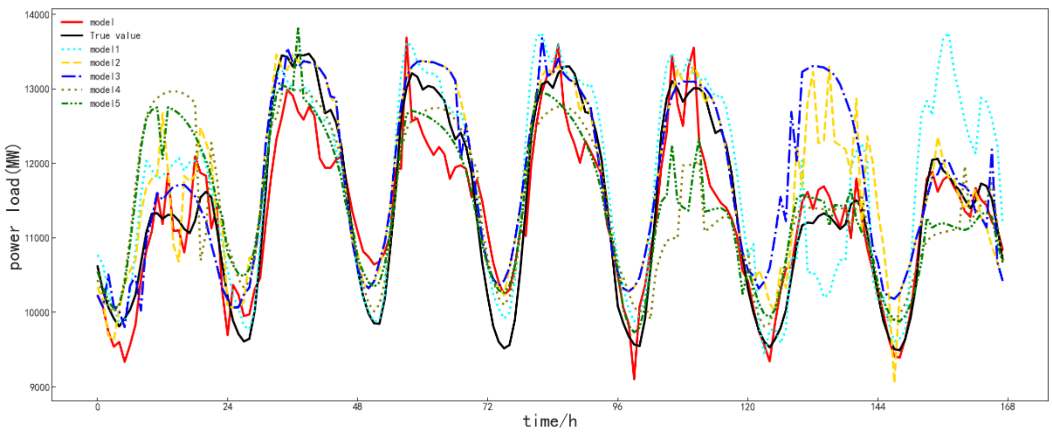

Fig. 3. The prediction results by using our base Model and Models 1-5.

Table 3

Predictive performance evaluation of different models.

Model \ Error	MAE	MAPE	D_{state}
Model 1	468.0968	0.0412	0.6310
Model 2	462.2923	0.0425	0.6845
Model 3	484.0083	0.0446	0.5774
Model 4	541.9373	0.0475	0.5893
Model 5	516.4112	0.0454	0.6964
The Proposed Hybrid Model	403.4315	0.0351	0.7917

Table 4

RT test results for different models.

Reference model	Test model				
	Model	Model 1	Model 2	Model 3	Model 4
Model 1	3.3028(0.0012)				
Model 2	2.2502(0.0257)	-1.0365(0.3015)			
Model 3	4.3433(0.0000)	1.0055(0.3161)	2.0477(0.0422)		
Model 4	4.1120(0.0000)	0.7837(0.4343)	1.8241(0.0700)	-0.2213(0.8251)	
Model 5	2.0122(0.0458)	-1.2733(0.2047)	-0.2360(0.8137)	-2.2865(0.0235)	-2.0623(0.0407)

Note： Z_{RT} (p-value)

Table 5

DM test results for different models.

Reference model	Test model				
	Model	Model 1	Model 2	Model 3	Model 4
Model 1	2.2408(0.0264)				
Model 2	2.2479(0.0259)	0.0412(0.9672)			
Model 3	2.9163(0.0040)	1.4447(0.1504)	1.9697(0.0505)		
Model 4	4.0758(0.0000)	2.2342(0.0268)	2.0697(0.0400)	0.4919(0.6234)	
Model 5	3.6586(0.0003)	1.0697(0.2863)	0.9443(0.3464)	-0.6050(0.5460)	-2.8829(0.0045)

Note: Z_{DM} (p-value)

Table 6

PT test results for different models.

	Model	Model 1	Model 2	Model 3	Model 4	Mode 5
Z_{PT}	7.3051	3.0614	4.5016	1.6465	2.4175	4.5559
p-value	0.0000	0.0026	0.000	0.1015	0.0167	0.0000

We learn from Figure 3 that in terms of changing trends, the prediction values generated by each model are roughly identical to the real values. In the early period and the middle to end periods, our prediction values given by our proposed method (i.e., the proposed hybrid forecasting method) are significantly close to real values. Especially, according to the prediction results on March 31, 2018, depicted in Figure 4, we find that the prediction values by our proposed model are the closest to the real data.

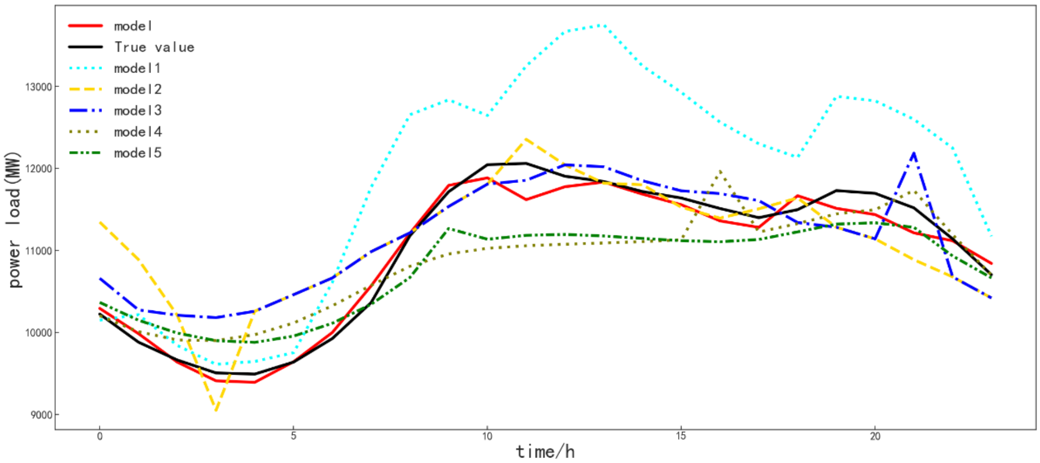

Fig. 4. The prediction results generated by using base Model and Models 1-5 on March 31, 2018.

In addition, we plot Figure 5 to visually show the prediction results on March 30, 2018 (which is the holiday during our forecasting period). Observing the predictions, we reveal that the prediction results by our proposed model in holidays are, by and large, more accurate than other models. It is thus concluded that the model proposed in this paper dominates Models 1-5 in predicting power load in holidays.

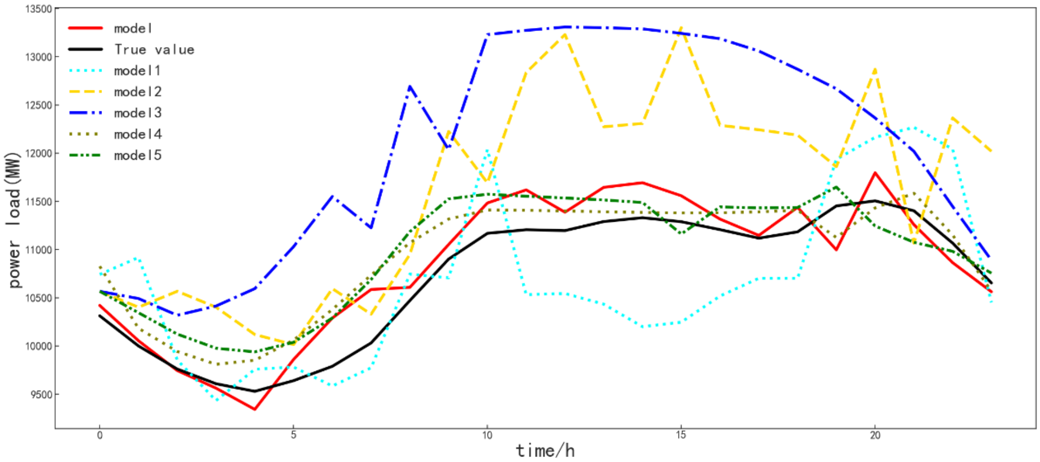

Fig. 5. The prediction results generated by using the base Model and Models 1-5 on March 30, 2018 (which is a holiday).

In spite of the above observations, the differences between the prediction and real values differ for the forecasting models. Specifically, the difference between the prediction and real values by the proposed hybrid forecasting model is slightly larger than that by other models from the early to middle periods. Thus, it would be difficult to use our observations only to confirm which model has a higher prediction accuracy. According to the level forecasting demonstrated by MAPE and MAE and direction forecasting demonstrated by D_{state}, we find that, if the EMD method is used to decompose the original load data, we can obtain more accurate prediction values than when we do not use the EMD method for decomposition. According to the comparison between the results from Models 1 and 4 as well as the comparison between our hybrid forecasting model and Model 5, we conclude that the improved PSO algorithm is better than the GA algorithm in the parameter optimization. Moreover, by comparing the results by Models 1 and 3 as well as those by Model 2 and our hybrid forecasting model, we reveal that when the same data decomposition and feature selection methods are adopted, the prediction accuracy from the SVM method is higher than that from the BPNN method. In addition, Dai and Zhao (2020) had shown the superiority of the mRMR-WGRP method in feature selection and holiday data processing.

From the DM test results shown in Table 5, the following conclusions can be obtained. First, when the proposed hybrid method is used as the test target, the *p*-values are all less than 5%, which indicates that the statistical performance of proposed hybrid method is better than other models. Secondly, when the EMD method is introduced based on original model, it outperforms the original model at the 5% significance level, which indicates that the prediction performance can be improved by data decomposition. Similarly, Table 6 shows the calculation results of the PT test for each model. It is not difficult to find that except for model 3, Z_{PT} of other models is all greater than 1.96. The proposed method with the largest Z_{PT} achieves better directional prediction accuracy compared with other models. In addition, according to the RT test results of each model given in Table 4, we can obtain the conclusions similar to those in the DM test. First, since the obtained *p*-values are less than 5%, the proposed method has the best directional prediction accuracy compared with other models. Meanwhile, compared with the model without EMD, we find that the EMD method improves the directional prediction accuracy.

Therefore, the above observations and discussions indicate that the proposed method has a higher level prediction accuracy and direction prediction accuracy compared to other forecasting models.

We also present a boxplot as in Figure 6 to show the degree of dispersion of the absolute prediction error $|\hat{Y}_i - Y_i|$. It can be noticed that, if we consider the factors such as the number of outliers and the value of the quartile, the model proposed in this paper outperforms other models with respect to the statistic $|\hat{Y}_i - Y_i|$.

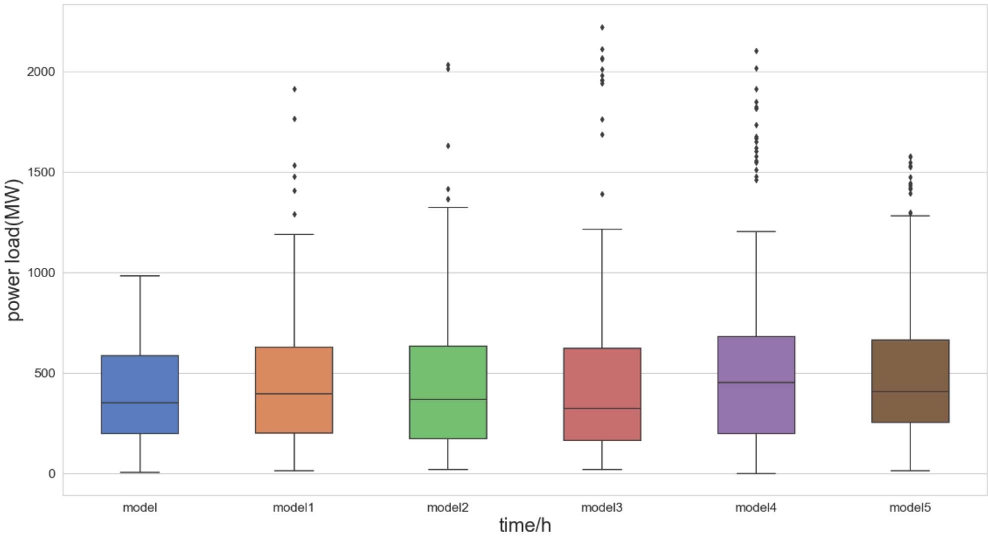

Fig. 6. The boxplot of the absolute forecast errors by our base Model and Models 1-5.

4. Conclusion

Precise and timely short-term power load forecasting is critical for accurate decision making by most power industry practitioners, researchers and policy makers. Numerous power load prediction models have been extensively developed in recent years. However, the unexpected COVID-19 outbreak caused unprecedented economic and social disruptions, which produced negative effects on the uncertainty of electricity consumption. Then more novel durable forecasting methods are being explored. Generally, the prediction accuracy is closely related to the quality of the datasets and forecasting model. In this paper, we develop a novel hybrid forecasting method which are effective approaches to data decomposition, feature selection, holiday historical load sequence selection, and parameter optimization. We aim to develop a combined hybrid forecasting method for power load. To improve the forecasting performance, we use the EMD method to decompose original power load data and apply the mRMR method to select the features that are highly correlated with the prediction variable for the IMFs and residual decomposed by the EMD method, which results in no redundancy. Then, we select the historical data sequence of a holiday on which we process through the WGRP algorithm. Finally, we use the PSO algorithm with second-order oscillating and repulsion factor to optimize the parameters of SVM prediction model for a further improving the prediction accuracy. The results of forecasting the power load using a real electricity market dataset in Singapore are compared with the mRMR-WGRP-SecRPSO-SVM model, the mRMR-WGRP-BPNN model, the EMD-mRMR-WGRP-BPNN model, the mRMR-WGRP-GA-SVM model, and the EMD-mRMR-WGRP-GA-SVM methods.

The following conclusions are obtained via the practical verification of Singapore electricity market. First, the prediction after decomposing the original data with the EMD method can reduce the computational complexity and also improve the prediction accuracy. Accordingly, the proposed hybrid method has the potential to conduct power load forecasting. Secondly, ceteris paribus, the prediction result from the optimized SVM model is more accurate than that from the BPNN. In particular, parameter optimization and setting the key parameter values are the first steps to investigate and enhance the prediction performance. Thirdly, ceteris paribus, the prediction error generated by the SVM model with the improved PSO algorithm is smaller than that by the SVM model with the GA algorithm. Fourthly, the proposed hybrid forecasting method is superior to other forecasting methods according to their level and direction prediction accuracies. In conclusion, the proposed hybrid forecasting method is accurate, effective, and feasible prediction method.

According to the above conclusions, it can be seen that the proposed hybrid power load forecasting method with intelligent data process and parameter optimization is accurate and effective, and has practical application value and operability. However, when we use the improved PSO algorithm for the SVM parameter optimization, the computation time would be long. Therefore, in the future, we will continue to improve the PSO algorithm to reduce the calculation time. Moreover, the EMD method is improved to enhance the accuracy of data decomposition, so as to further improve the accuracy of load prediction. In addition, the inclusion of carbon emission as an impact factor in the future electricity load forecasting studies could be considered, which may make the load forecasting more relevant to the current social changes.

Declaration of Competing Interest

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

Acknowledgments

This work was supported by the National Natural Science Foundation of China [No.72171126], Ministry of Education Project of Humanities and Social Science [No.20YJA630009], Faculty Research Grant of Lingnan University under the grant number DB21B1.

References

Abedinia, O., Amjady, N. and Zareipour, H. (2016). A new feature selection technique for load and price forecast of electrical power systems. IEEE Transactions on Power Systems, 32(1), 62-74. 10.1109/TPWRS.2016.2556620.

Ahmad, T. and Chen, H. (2019). Deep learning for multi-scale smart energy forecasting. Energy, 175, 98-112. https://doi.org/10.1016/j.energy.2019.03.080.

Al-Musaylh, M. S., Deo, R. C., Li, Y. and Adamowski, J. F. (2018). Two-phase particle swarm optimized-support vector regression hybrid model integrated with improved empirical mode decomposition with adaptive noise for multiple-horizon electricity demand forecasting. Applied energy, 217, 422-439. https://doi.org/10.1016/j.apenergy.2018.02.140.

Angelopoulos, D., Siskos, Y. and Psarras, J. (2019). Disaggregating time series on multiple criteria for robust forecasting: The case of long-term electricity demand in Greece. European Journal of Operational Research, 275(1), 252-265. https://doi.org/10.1016/j.ejor.2018.11.003.

Arora, S. and Taylor, J. W. (2018). Rule-based autoregressive moving average models for forecasting load on special days: A case study for France. European Journal of Operational Research, 266(1), 259-268. https://doi.org/10.1016/j.ejor.2017.08.056.

Barman, M. and Choudhury, N. B. D. (2020). A similarity based hybrid GWO-SVM method of power system load forecasting for regional special event days in anomalous load situations in Assam, India. Sustainable Cities and Society, 61, 102311. https://doi.org/10.1016/j.scs.2020.102311.

Bessec, M. and Fouquau, J. (2018). Short-run electricity load forecasting with combinations of stationary wavelet transforms. European Journal of Operational Research, 264(1), 149-164. https://doi.org/10.1016/j.ejor.2017.05.037.

Che, J., Yang, Y., Li, L., Bai, X., Zhang, S. and Deng, C. (2017). Maximum relevance minimum common redundancy feature selection for nonlinear data. Information Sciences, 409, 68-86.

https://doi.org/10.1016/j.ejor.2017.05.037.

Dai, Y. and Zhao, P. (2020). A hybrid load forecasting model based on support vector machine with intelligent methods for feature selection and parameter optimization. Applied Energy, 279, 115332. https://doi.org/10.1016/j.apenergy.2020.115332.

Diebold, F. X. and Mariano, R. S. (2002). Comparing predictive accuracy. Journal of Business and economic statistics, 20(1), 134-144. https://doi.org/10.1198/073500102753410444.

Ding, J., Wang, M., Ping, Z., Fu, D. and Vassiliadis, V. S. (2020). An integrated method based on relevance vector machine for short-term load forecasting. European Journal of Operational Research, 287(2), 497-510. https://doi.org/10.1016/j.ejor.2020.04.007.

Ding, Y., Zhang, W., Yu, L. and Lu, K. (2019). The accuracy and efficiency of GA and PSO optimization schemes on estimating reaction kinetic parameters of biomass pyrolysis. Energy, 176, 582-588. https://doi.org/10.1016/j.energy.2019.04.030.

du Jardin, P. (2021). Forecasting corporate failure using ensemble of self-organizing neural networks. European Journal of Operational Research, 288(3), 869-885. https://doi.org/10.1016/j.ejor.2020.06.020.

Hafeez, G., Khan, I., Jan, S., Shah, I. A., Khan, F. A. and Derhab, A. (2021). A novel hybrid load forecasting framework with intelligent feature engineering and optimization algorithm in smart grid. Applied Energy, 299, 117178. https://doi.org/10.1016/j.apenergy.2021.117178.

Hong, T., Pinson, P., Wang, Y., Weron, R., Yang, D. and Zareipour, H. (2020). Energy forecasting: A review and outlook. IEEE Open Access Journal of Power and Energy, 7, 376-388. 10.1109/OAJPE.2020.3029979.

Jiang, P., Li, R., Liu, N. and Gao, Y. (2020). A novel composite electricity demand forecasting framework by data processing and optimized support vector machine. Applied Energy, 260, 114243. https://doi.org/10.1016/j.apenergy.2019.114243.

Kouziokas, G. N. (2020). A new W-SVM kernel combining PSO-neural network transformed vector and Bayesian optimized SVM in GDP forecasting. Engineering Applications of Artificial Intelligence, 92, 103650. https://doi.org/10.1016/j.engappai.2020.103650.

Li, R., Hu, Y., Heng, J. and Chen, X. (2021). A novel multiscale forecasting model for crude oil price time series. Technological Forecasting and Social Change, 173, 121181. https://doi.org/10.1016/j.techfore.2021.121181.

Liang, Y., Niu, D. and Hong, W. C. (2019). Short term load forecasting based on feature extraction and improved general regression neural network model. Energy, 166, 653-663. https://doi.org/10.1016/j.energy.2018.10.119.

Liu, M. D., Ding, L. and Bai, Y. L. (2021). Application of hybrid model based on empirical mode decomposition, novel recurrent neural networks and the ARIMA to wind speed prediction. Energy Conversion and Management, 233, 113917. https://doi.org/10.1016/j.enconman.2021.113917.

Liu, Q., Shen, Y., Wu, L., Li, J., Zhuang, L. and Wang, S. (2018). A hybrid FCW-EMD and KF-BA-SVM based model for short-term load forecasting. CSEE Journal of Power and Energy Systems, 4(2), 226-237. 10.17775/CSEEJPES.2016.00080.

Masoumi, A., Ghassem-zadeh, S., Hosseini, S. H. and Ghavidel, B. Z. (2020). Application of neural network and weighted improved PSO for uncertainty modeling and optimal allocating of renewable energies along with battery energy storage. Applied Soft Computing, 88, 105979. https://doi.org/10.1016/j.asoc.2019.105979.

Mendeley Data (2020). Dataset of Singapore's power market. https://data.mendeley.com/datasets/hvz7g6r3mw/2 (URL last accessed on August 7, 2021).

Milchram, C., Künneke, R., Doorn, N., van de Kaa, G. and Hillerbrand, R. (2020). Designing for justice in electricity systems: A comparison of smart grid experiments in the Netherlands. Energy Policy, 147, 111720. https://doi.org/10.1016/j.enpol.2020.111720.

Niu, T., Wang, J., Lu, H., Yang, W. and Du, P. (2020). Developing a deep learning framework with two-stage feature selection for multivariate financial time series forecasting. Expert Systems with Applications, 148, 113237. https://doi.org/10.1016/j.eswa.2020.113237.

Sharmin, S., Shoyaib, M., Ali, A. A., Khan, M. A. H. and Chae, O. (2019). Simultaneous feature selection and discretization based on mutual information. Pattern Recognition, 91, 162-174. https://doi.org/10.1016/j.patcog.2019.02.016.

Singh, S. N. and Mohapatra, A. (2021). Data driven day-ahead electrical load forecasting through repeated wavelet transform assisted SVM model. Applied Soft Computing, 111, 107730. https://doi.org/10.1016/j.asoc.2021.107730.

Sun, H., Xie, B., Tian, Y. and Li, Z. (2017). Forecasting of short-term power load of SecRPSO-SVM based on data-driven (In Chinese). Journal of System Simulation, 29(8), 1829-1836. 10.16182/j.issn1004731x.joss.201708025.

Tang, R., Wang, S. and Li, H. (2019). Game theory based interactive demand side management responding to dynamic pricing in price-based demand response of smart grids. Applied Energy, 250, 118-130. https://doi.org/10.1016/j.apenergy.2019.04.177.

Thomas, S. R., Kurupath, V. and Nair, U. (2020). A passive islanding detection method based on K-means clustering and EMD of reactive power signal. Sustainable Energy, Grids and Networks, 23, 100377. https://doi.org/10.1016/j.segan.2020.100377.

Wang, X., Cho, S. H. and Scheller-Wolf, A. (2021). Green technology development and adoption: competition, regulation, and uncertainty—a global game approach. Management Science, 67(1), 201-219. https://doi.org/10.1287/mnsc.2019.3538.

Wu, X. Y., He, J. H., Zhang, P. and Hu, J. (2015). Power system short-term load forecasting based on improved random forest with grey relation projection. Automation of Electric Power Systems, 39(12), 50-55. 10.7500/AEPS20140916005.
Wang, X., Cho, S. H. and Scheller-Wolf, A. (2021). Green technology development and adoption: competition, regulation, and uncertainty—a global game approach. Management Science, 67(1), 201-219. https://doi.org/10.1287/mnsc.2019.3538.
Xiao, L., Wang, C., Dong, Y. and Wang, J. (2019). A novel sub-models selection algorithm based on max-relevance and min-redundancy neighborhood mutual information. Information Sciences, 486, 310-339. https://doi.org/10.1016/j.ins.2019.01.075.

Xie, G., Qian, Y. and Wang, S. (2020). A decomposition-ensemble approach for tourism forecasting. Annals of Tourism Research, 81, 102891. https://doi.org/10.1016/j.annals.2020.102891.

Xie, K., Yi, H., Hu, G., Li, L. and Fan, Z. (2020). Short-term power load forecasting based on Elman neural network with particle swarm optimization. Neurocomputing, 416, 136-142. https://doi.org/10.1016/j.neucom.2019.02.063.

Yang, A., Li, W. and Yang, X. (2019). Short-term electricity load forecasting based on feature selection and Least Squares Support Vector Machines. Knowledge-Based Systems, 163, 159-173. https://doi.org/10.1016/j.knosys.2018.08.027.

Yang, Y., Hong, W. and Li, S. (2019). Deep ensemble learning based probabilistic load forecasting in smart grids. Energy, 189, 116324. https://doi.org/10.1016/j.energy.2019.116324.

Yaslan, Y. and Bican, B. (2017). Empirical mode decomposition based denoising method with support vector regression for time series prediction: A case study for electricity load forecasting. Measurement, 103, 52-61. https://doi.org/10.1016/j.energy.2019.116324.

Yuan, G., Gao, Y. and Ye, B. (2021). Optimal dispatching strategy and real-time pricing for multi-regional integrated energy systems based on demand response. Renewable Energy, 179, 1424-1446. https://doi.org/10.1016/j.renene.2021.07.036.

Zeng, N., Qiu, H., Wang, Z., Liu, W., Zhang, H. and Li, Y. (2018). A new switching-delayed-PSO-based optimized SVM algorithm for diagnosis of Alzheimer's disease. Neurocomputing, 320, 195-202. https://doi.org/10.1016/j.neucom.2018.09.001.

Zhang, X., Wang, J. and Gao, Y. (2019). A hybrid short-term electricity price forecasting framework: Cuckoo search-based feature selection with singular spectrum analysis and SVM. Energy Economics, 81, 899-913. https://doi.org/10.1016/j.eneco.2019.05.026.

Zhao, E., Zhang, Z. and Bohlooli, N. (2020). Cost and load forecasting by an integrated algorithm in intelligent electricity supply network. Sustainable Cities and Society, 60, 102243. https://doi.org/10.1016/j.scs.2020.102243.

Zhu, B., Han, D., Wang, P., Wu, Z., Zhang, T. and Wei, Y. M. (2017). Forecasting carbon price using empirical mode decomposition and evolutionary least squares support vector regression. Applied energy, 191, 521-530. https://doi.org/10.1016/j.apenergy.2017.01.076.

Zhu, B. Z., Jiang, M. X., Zhang, S. F. and Jin, Y. (2020). Resource and environment economic complex system: models and applications (In Chinese). Science Press, Beijing, China.

Zhu, B., Ye, S., Wang, P., Chevallier, J. and Wei, Y. M. (2022). Forecasting carbon price using a multi-objective least squares support vector machine with mixture kernels. Journal of Forecasting, 41(1), 100-117. https://doi.org/10.1002/for.2784

.