

Short-Term Load Forecasting based on Variational Mode Decomposition and Least Squares Support Vector Machine by Improved Artificial Fish Swarm-Shuffled Frog Jump Algorithms

Haizhu Yang^a, Zhaoyang Jiang^{a,*}, Menglong Li^a, and Peng Zhang^b

^a*School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo, 454000, China*

^b*School of Electrical and Information Engineering, Tianjin University, Tianjin, 300072, China*

Abstract

Short-term load forecasting plays a key role in the safe dispatching and economic operation of the power system. The least square support vector machine (LSSVM) has the power system. The least square support vector machine (LSSVM) has great potential in forecasting problems, particularly by employing an appropriate algorithm to determine the values of its two parameters. In order to improve LSSVM load prediction accuracy, this paper proposes a LSSVM based on the Variational mode decomposition(VMD) electric load forecasting model that uses an artificial fish swarm-shuffled frog leaping algorithm to determine the appropriate values of the two parameters. The historical data such as load and weather in the first 15 days of the forecast day are the input into LSSVM. The AFSA-SFLA-LSSVM forecasting model, the LAVAFSA-SFLA-LSSVM forecasting model, the AFSA-LSSVM forecasting model, and the VMD-LAVAFSA-SFLA-LSSVM forecasting model were established for electrical load forecasting in a certain area within 24 hours of a specific day. The results of the example show that the accuracy of the VMD-LAVAFSA-SFLA-LSSVM forecasting model was higher than the other three forecasting models and the prediction error was smaller as well.

Keywords: short-term load forecasting; Variational mode decomposition; least squares support vector machine; improve artificial fish swarm-shuffled frog leaping algorithm

(Submitted on November 19, 2019; Revised on December 20, 2019; Accepted on December 24, 2019)

© 2019 Totem Publisher, Inc. All rights reserved.

1. Introduction

With the continuous development of the social economy in recent years, the demand for power resources in China is increasing. However, there is not enough power supply. For this reason, the power sector needs to study the data of load forecasting in detail to ensure the safe and stable operation of the power system. It is necessary to clearly indicate the power load in different regions, plan reasonably according to the load, balance the power supply and demand, and meet the actual power demand of each user. Power load forecasting mainly refers to the comprehensive forecasting and calculation of power load, power demand, power situation and power consumption in the future period based on historical data and social factors [1-2]. Power load data is affected by many factors. The data not only shows periodic changes with residents and production activities, but also shows random changes under the constraints of various environmental factors such as temperature, precipitation, humidity, holidays, market transactions, etc. The cycle of power load forecasting includes short-term, medium-term and long-term. The short-term forecasting not only provides an important guarantee for the safe and economic operation of the power system, but also provides the basis for scheduling, power supply planning and trading planning [3-4]. Over the years, theories and methods of power load forecasting have emerged such as neural network, time series [5], bayesian, fuzzy theory, wavelet analysis, regression analysis [6], support vector machine [7] and other technologies, which all provide powerful tools for power load forecasting. Li et al. [8] combines the grey theory with the Elman neural network to improve the prediction accuracy and convergence speed. Wang et al. [9] uses a BP neural network and an improved variable learning rate backpropagation to predict power load. In addition, some scholars have proposed a support vector machine (SVM) prediction model. The SVM method based on the statistical theory can solve nonlinear and small sample problems better and has superiority for power load forecasting [10-11]. The least squares support vector

* Corresponding author.

E-mail address: zhaoyangjiang1994@163.com

machine (LSSVM) replaces the inequality constraint with the equality constraint on the basis of SVM, which avoids solving the time-consuming quadratic programming problem. Reference [12] used SVM with a simulated annealing algorithm to forecast Taiwan's electricity load. Reference [13-14] introduced LSSVM to predict the annual load in China with the rolling mechanism. The good results verify the applicability of LSSVM in load forecasting. The prediction effect of the LSSVM model relies heavily on the regularization of parameter "C" and the bandwidth of kernel function "σ". Therefore, the key to using LSSVM is to choose optimal parameters. Reference [15] used a genetic algorithm (GA) for optimal parameters, but GA programming is complex and search speed is slow. Reference [16-17] used particle swarm optimization (PSO) to select optimal parameters. PSO operation is simple and easy to implement, but it will appear premature and easy to fall into local optimum. In this paper, the Levy artificial fish swarm-shuffled frog leaping algorithm based on adaptive vision (LAVAFSA-SFLA) is proposed to optimize LSSVM parameters. At the same time, the load data is decomposed by the variational mode decomposition (VMD), and the LSSVM with a proper kernel function optimized by LAVAFSA-SFLA is selected to predict according to the decomposed subsequence. The results show that the prediction accuracy of load data becomes higher after VMD.

2. Variational Mode Decomposition

After the daily load data is decomposed by the variational mode decomposition (VMD), the trend component and fluctuation component of different frequencies can be obtained. VMD and EEMD have different recursive screening principles. VMD uses non-recursive and variable modal principles to decompose signals into a series of finite bandwidth subsequences. VMD has better harmonic separation ability and each subsequence has better regularity [18].

The VMD consists of three steps: establishing a constrained variational model, Lagrangian transformation, and alternating updates:

Step 1 Establishing a constrained variational model. VMD can decompose a real valued signal into a series of band-limited Intrinsic Mode Functions (IMF) u_k in a spectral domain owning specific sparsity properties. It is assumed that each IMF u_k can be concentrated around a center frequency ω_k which is determined along with the decomposition process [19]. The objective function explores the K IMFs of the original signal of which the sum of the estimated bandwidth is the minimum. Furthermore, the aggregation of the IMFs should be equal to the input signal $f(t)$, which is the constraint condition. The constrained variational problems are shown in Equations (1) and (2):

$$\min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \cdot u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \quad (1)$$

$$s.t. \sum_{k=1}^K u_k = f(t) \quad (2)$$

Where $f(t)$ is the original signal to be decomposed, t implicates the time script, $\{u_k\}$ represent the k th IMF, and $\{u_k\}$ stands for the set of IMFs $\{u_1, u_2, \dots, u_k\}$; $\{\omega_k\}$ represents the corresponding center frequency; $\{\omega_k\}$ denotes the set of center frequencies of the IMFs $\{u_1, u_2, \dots, u_k\}$; K denotes the total number of the IMFs. ∂_t is the Dirac distribution.

Step 2 Lagrangian transformation. In order to translate the constraint variational problem from above into an unconstrained optimization problem, both quadratic penalty term and Lagrangian multipliers are considered. The augmented Lagrange equation is shown in Equation (3):

$$L(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \cdot u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \|f(t) - \sum u_k(t)\|_2^2 + \langle \lambda(t), f(t) - \sum u_k(t) \rangle \quad (3)$$

Where λ is the Lagrange multiplication factor for tightening restraint, α refers to the balance parameter of the data fidelity constraint.

Step 3 Alternate update. The optimization problem in step 2 can be solved according to Equations (4) and (5).

$$\hat{u}_k^{n+1}(\omega) = \left(\hat{f}(\omega) - \sum_{i \neq k}^K \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2} \right) \frac{1}{1 + 2\alpha(\omega - \omega_k)^2} \quad (4)$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 d\omega} \quad (5)$$

$\hat{u}_k^{n+1}(\omega)$ is the Wiener filtering of IMF. it can be obtained by inverse Fourier transform. $\omega_k^{n+1}(\omega)$ is the frequency center of IMF where n represents the number of iterations and ω is the frequency value. The VMD algorithm flow is summarized as follows:

- The original signal $f(t)$ is input.
- $\{u_k^1\}$, $\{\omega_k^1\}$ and $\{\hat{\lambda}_k^1\}$ are initialized. $n = 0$.
- $\hat{u}_k^{n+1}(\omega)$ and ω_k^{n+1} are updated by Equations (4) and (5).
- $\hat{\lambda}$ is updated by the Equation (6).
- $\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \tau \left(f(\omega) - \sum_k^K \hat{u}_k^{n+1} \right)$ (6)
- Perform steps 3 and 4 until the Equation (7) is satisfied.
- $\sum_{k=1}^K \|\hat{u}_k^{n+1} - \hat{u}_k^n\|_2^2 / \|\hat{u}_k^n\|_2^2 < \varepsilon$ (7)
- IMFs are output.

3. Least Squares Support Vector Machine (LSSVM)

The basic idea of SVM is to use a nonlinear mapping to map the data of the input space to a high-dimensional feature space and transform the actual problem into a quadratic programming problem with inequality constraints [20]. The LSSVM is an extension of SVM. The specific principle is as follows: set a set of samples $\{x_i, y_i\}_{i=1}^m$, where m is the sample set size, $x_i \in R^n$ is the input vector and $y_i \in R^n$ is the corresponding output value for sample i . The samples are mapped to higher dimensional spaces by using a nonlinear function φ , which performs linear regression [21]. The regression function is shown in Equation (8):

$$f(x) = \omega^T \varphi(x) + b \quad (8)$$

Where ω is the weight vector and b denotes the offset value. The LSSVM equation with the equality constraints can be described as Equations (9) and (10):

$$\min J(\omega, \xi) = \frac{1}{2} \omega^T \omega + \frac{1}{2} C \sum_{i=1}^m \xi_i^2 \quad (9)$$

$$s.t. y_i = \omega \varphi(x) + b + \xi_i, \quad i = 1, 2, \dots, m \quad (10)$$

Where C is the penalty parameter; and ξ_i is the slack variable; the Lagrangian function L can be constructed by the Equation (11):

$$L(\omega, b, \xi, a) = \frac{1}{2} \omega^T \omega + \frac{1}{2} C \sum_{i=1}^m \xi_i^2 - \sum_{i=1}^m a_i \{ \omega^T \varphi(x_i) + b + \xi_i - y_i \} \quad (11)$$

Where a_i is the Lagrangian multiplier. The Karush-Kuhn-Tucker (KKT) conditions for optimality are given by the Equation (12):

$$\begin{cases} \frac{\partial L}{\partial \omega} = 0 \rightarrow \omega = \sum_{i=1}^m a_i \varphi(x_i) \\ \frac{\partial L}{\partial b} = 0 \rightarrow \sum_{j=1}^m a_j = 0 \\ \frac{\partial L}{\partial \xi_i} = 0 \rightarrow a_i = C \xi_i \\ \frac{\partial L}{\partial a_i} = 0 \rightarrow \omega^T \varphi(x_i) + b + \xi_i - y_i = 0 \end{cases} \quad (12)$$

Eliminating the variables ω and ξ_i is shown in Equation (13):

$$\begin{bmatrix} 0 & Q^T \\ Q & K + C^{-1}I \end{bmatrix} \begin{bmatrix} b \\ A \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix} \quad (13)$$

Where $Q = [1, \dots, 1]^T$, $A = [a_1, a_2, \dots, a_m]^T$, $Y = [y_1, y_2, \dots, y_m]^T$.

The kernel function can be set according to the Mercer condition as shown in Equation (14):

$$K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j) \quad (14)$$

Then, the function of LSSVM is estimated as Equation (15):

$$f(x) = \sum_{i=1}^m a_i K(x, x_i) + b \quad (15)$$

Where $K(x, x_i)$ is a kernel function.

In this paper, we need to use two kernels: the linear kernel function as shown in Equation (16) and the radial basis function (RBF) as shown in Equation (17).

$$K(x, x_i) = x \cdot x_i^T \quad (16)$$

$$K(x, x_i) = \exp\{-\|x - x_i\|^2 / 2\sigma^2\} \quad (17)$$

Where x is the m -dimensional input vector, and x_i is the center of the i^{th} radial basis function. $\|x - x_i\|$ is the norm of vector $x - x_i$, which represents the distance between x and x_i . σ is the width of the kernel function.

4. Optimization of LSSVM Parameters based on the Improved Artificial Fish Swarm-Shuffled Frog Leaping Algorithm

The artificial fish is easy to fall into local optimum in the early stage, resulting in a slower search speed in the later stage [22]. In this paper, the Levy variation and the adaptive visual field are added to improve the complexity of the fish swarm and the speed of optimization. Next, it is matched with the shuffled frog leaping algorithm with advantages in late local search. The goal is to be able to quickly output optimal solutions and reduce load forecasting time.

4.1. The Levy Variation is Added to Artificial Fish Swarm Algorithm

The step size of the Levy variation is better than the Gaussian variation and the Cauchy variation and it can produce more

disturbances. The probability density function of the Levy distribution is shown as Equation (18):

$$L_{\varsigma,\gamma}(z) = \frac{1}{\pi} \int_0^{\infty} \exp(-\gamma q^{\varsigma}) \cos(qz) dq \quad (18)$$

Where ς ($0 < \varsigma < 2$) represents the sharpness of the levy distribution and γ ($\gamma > 0$) is the scale unit used to control the Levy distribution.

When the artificial fish is trapped in the local optimum or the iteration has not changed many times, the poorest quality fish in the current fish swarm is exchanged with the highest quality fish in the bulletin board to obtain a new middle fish swarm. Then, the artificial fish that falls into the local solution jumps out and searches for global optimization again. A random disturbance factor based on Levy distribution is added to the fish swarm and the behavior of artificial fish does not change. The mathematical model is shown as Equation (19):

$$X'_i = X_i + X_i \times L_j(t) \quad (19)$$

4.2. The Adaptive Field of Visual is Added to Artificial Fish Swarm Algorithm

To some extent, the vision of artificial fish determines the optimization speed. The experimental results show that when the field of vision is large, although the convergence speed is fast in the early stage, it will slow down in the later stage and the optimization accuracy is not enough. On the contrary, if the field of vision is very small, the convergence speed is slow in the early stage and it is easy to fall into local optimum. In order to solve the problems mentioned above, an adaptive visual field method for artificial fish is proposed [23]. The artificial fish has many behaviors, but foraging behavior is the most basic. Here, only foraging behavior is analyzed. First, a larger field of vision is set in the initialization stage of the algorithm. With the iteration of the algorithm, the field of vision of the artificial fish is reduced with the area of the higher food concentration. In order to improve the search accuracy, it needs to be reduced to a certain extent. If the field of vision is too small, the probability of searching the optimal solution will be reduced. As a result, the lower limit should be set to prevent it from being too small. The equation of the distance between individual fish and the fish swarm with a higher food concentration is shown in Equation (20):

$$\begin{cases} \text{Disn}(i) = \frac{1}{M_1 \cdot S} \sum_{i=1}^{M_1} \sqrt{\sum_{j=1}^L (x_{ij}^n - X_j^n)} \\ X_j^n = \frac{1}{M_1} \sum_{i=1}^{M_1} x_{ij}^n \end{cases} \quad (20)$$

Where M_1 represents the number of artificial fish in the field of vision and S is the longest diagonal length of the whole search space. L represents the dimension of spatial individuals and n is the number of iterations. $j = 1, 2, \dots, L$; x_{ij}^n represents the artificial fish of the j^{th} dimension of the i^{th} fish in n iterations; x_j^n is the center of the j^{th} dimension of the fish swarm in n iterations.

In order to accelerate the convergence speed, the attenuation factor α is introduced. In order to prevent the visual field from being too small, the lower limit factor β is introduced according to the reference [24]. The algorithm equations are shown in Equation (21):

$$\begin{cases} \text{Visual}_{k+1} = \alpha \cdot \text{Disn}(i), & \alpha \geq \beta \\ \text{Visual}_{k+1} = \beta \cdot \text{Visual}_0, & \alpha < \beta \end{cases} \quad (21)$$

Where Visual_0 is the initial value of visual field and Visual_{k+1} is the visual field after dynamic adjustment. The ranges of values of α and β are both $[0, 1]$.

4.3. Parameter Optimization of LSSVM based on LAVAFSA-SFLA

The shuffle frog leaping algorithm is a new intelligent algorithm which is simple in concept, easy to implement, and fast in local optimization. At first, the artificial fish is used to find the optimal solution region. Then, the shuffle frog leaping algorithm is used to optimize it.

The objective function is shown in Equation (22):

$$\begin{aligned} \min F(C, \sigma) &= \sum_{i=1}^M (J_i - Q_i)^2 \\ \text{s.t. } C &\in [C_{\min}, C_{\max}] \\ \text{s.t. } \sigma &\in [\sigma_{\min}, \sigma_{\max}] \end{aligned} \quad (22)$$

Where J_i is the output value of the i^{th} known sample; Q_i represents the predicted value of the i^{th} sample, which can be calculated by the Equation (15).

There are two parameters that need to be chosen in the LSSVM model which are the bandwidth of the Gaussian RBF kernel " σ " and the regularization parameter " C ". In this paper, the LAVAFSA-SFLA is used to optimize the parameters. The position of each artificial fish contains parameter " σ " and parameter " C ". The fitness function of the artificial fish swarm algorithm is obtained by the Equation (24). The training sample set and prediction sample set were established based on the historical load data. The specific steps of optimizing parameters are as follows:

Step 1 Initialization settings. M is the total number of artificial fish. Visual represents the vision of artificial fish. Step is the mobile step length. δ represents the congestion factor. The try-number is the number of attempts in the movement of artificial fish. x_i represents the position of artificial fish and contains C and σ parameters.

Step 2 Bulletin board assignment. Each artificial fish corresponds to a fitness function value and puts the optimal value on the bulletin board.

Step 3 Behavior choice. Each artificial fish was evaluated and the best behavior was selected. The state of high quality artificial fish was recorded on the bulletin board. The steps of the four execution actions are as follows:

1) Foraging behavior

Artificial fish move in the direction of high food concentration. If $F_i < F_j$, the x_i will move a step to x_j . Conversely, it will reselect the position of the artificial fish. When the artificial fish reaches the set maximum number of attempts in the optimization process, the Equation (25) is not satisfied and Step 4 is executed. The concrete expression is shown in Equation (23):

$$x_j = x_i + \text{rand}() \times \text{Step} \times \frac{x_j - x_i}{\|x_j - x_i\|}, \quad F_i < F_j \quad (23)$$

Where x_i represents the location of an artificial fish and the corresponding fitness function is F_i ; x_j represents an artificial fish randomly selected in the artificial fish field of vision, and the corresponding fitness function is F_j ; $\text{rand}()$ is a random number of (0,1).

2) Clustering behavior

When the artificial fish move to the region with high concentration, there will be clustering behavior. The central position of the visual field is x_c . If $F_c / M_1 > \delta Z_i$, the artificial fish moves step units to x_c . Otherwise, the foraging behavior will be performed. The concrete expression is shown in Equation (24):

$$x_j = x_i + rand() \times Step \times \frac{x_c - x_i}{\|x_c - x_i\|}, \quad F_c / M_1 > \delta F_i \quad (24)$$

Where Z_i is the concentration of food in the visual field of artificial fish, and F_c represents the fitness function value of x_c .

3) Rear end behavior

If $M_2 / M < \delta$, the artificial fish in this field of view are not crowded and x_i can move one step to x_b . Otherwise, the foraging behavior will be performed. The concrete expression is shown in Equation (25):

$$x_j = x_i + rand() \times Step \times \frac{x_b - x_i}{\|x_b - x_i\|}, \quad F_b / M_2 > \delta F_i \quad (25)$$

Where x_b represents the best position in x_i field of vision; and M_2 is the number of fish in the field of view of it.

4) Stochastic behaviour

When the artificial fish cannot satisfy the moving condition in the region, it moves one step in the field of vision randomly until it is satisfied. The concrete expression is shown in Equation (26):

$$x_j = x_i + rand() \times Step, \quad F_j \leq F_i \quad (26)$$

Step 4 Bulletin board update. After the iteration, the high-quality artificial fish will be compared with the bulletin board. If it is superior to the bulletin board, it will be replaced and displayed on the bulletin board.

Step 5 Judgment of Variation Conditions. If the artificial fish does not iterate for three consecutive times or the change in iteration is very small, step 6 will be executed. Otherwise, step 7 will be executed.

Step 6 Mutation operation. If the artificial fish have not been iterated for three consecutive times or the change in iteration is very small, the fish will be mutated according to Equation (20).

Step 7 Handover Algorithm. The artificial fish meeting the requirements of Equation (27) are selected to form a new fish swarm M_n . On the contrary, the artificial fish is returned to step 2.

$$\begin{cases} F_{k+1} < \mu \\ |F_{k+1} - F_k| < \lambda \end{cases} \quad (27)$$

Where F_k is the optimal fitness function of artificial fish after the k th iteration, and F_{k+1} is the optimal fitness function of artificial fish after the $(k + 1)^{\text{th}}$ iteration; μ and λ are allowable error values.

Step 8 Shuffle the frog leaping algorithm. According to the fitness function value, these artificial fishes of M_n are arranged from large to small and divided into two parts. The first half of M_n are assigned to the frog population and named as N subpopulation, and the rest as H subpopulation.

Step 9 Local search. The local search of N subpopulations was carried out and the subpopulations with a large fitness function value jumped to the subpopulations with small fitness function value. The concrete expression is shown in Equation (28):

$$\begin{cases} D_i = rand() \cdot (X_a - X_w) \\ X_{w-n} = X_{w-c} + D_i \end{cases} \quad (28)$$

Where i represents the i^{th} frog in the subpopulation, and D_i is the frog instant update step; $i=1,2,\dots,n$; X_{w-n} represents the updated position of frog individuals whose fitness function value is smaller than X_{w-c} .

Each frog position corresponds to a fitness function value. When the updated position is better than before, the update is successful. Otherwise, the best position X_{best} in the current frog population is selected, and X_a in Equation (30) is replaced. If the position update is stick, the position will be moved randomly into space until it is updated [25].

Step 10 Final optimization. The frog subpopulation of N is redefined as G after being updated. The frog subpopulation G and H are mixed to form a new frog subpopulation Z . The local search of frog subpopulation Z is carried out until the fitness function value is the maximum. At this time, the objective function Equation (24) is the minimum, and the algorithm is over. The optimal parameter solution C and σ is output.

5. Prediction Model of VMD-LAVAFSA-SFLA-LSSVM

The short-term load forecast is affected by the date type, weather condition and daily temperature. The input data shall include the load, weather and date type of the 15 days before the forecast date, as well as the daily minimum temperature and average temperature.

Prediction steps:

Step 1 The input data shall include the 24-hour load data and daily environment data in the first 15 days of the forecast date. VMD is used to decompose the daily load data and four subsequences of different frequency scales are obtained.

Step 2 The 15d load subsequences were combined according to the frequency scale to obtain the load matrix of 15 rows \times 24 columns (" 15×24 ") with four different frequency scales.

Step 3 Four groups of matrix and environment data are the input to four LSSVM models for load forecasting, and the parameters of LSSVM are determined by LAVAFSA-SFLA.

Step 4 The final daily load forecasting results are obtained by adding up four groups of load forecasting results with different frequency scales. The time interval of load data is one hour, and environment data is one day.

5.1. Specific Steps of Prediction Method of VMD-LAVAFSA-SFLA-LSSVM

Step 1 Bad data correction. In the actual data collection process, there are often some recorded data that deviate too much from the actual data, and the accuracy of prediction is seriously affected. Thus, it is necessary to correct the bad data before decomposition.

Step 2 VMD and matrix. The daily load sequence is divided into four 1×24 subsequences by VMD. The number of modal functions is four, i.e. $K = 4$. Taking the daily load data of a county in Jiaozuo City, Henan Province as an example, the four subsequences of the original data decomposed by VMD are shown in Figure 1.

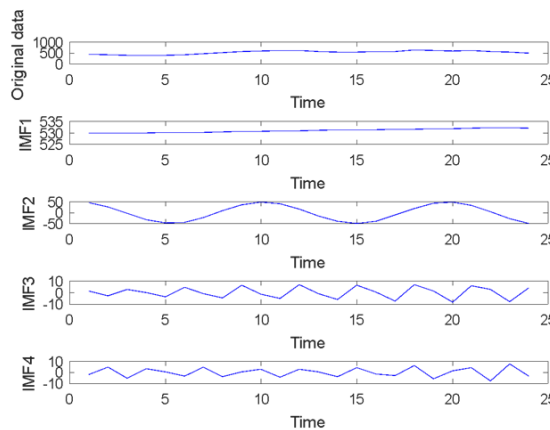


Figure 1. The daily load data is decomposed by VMD

It can be seen from Figure 1 that the four subsequences show the change of the daily load curve on different frequency scales. IMF1 represents the trend component which reflects the overall trend of the curve. IMF2-4 is the fluctuation component. IMF2 reflects the changing shape of the curve. IMF3 and IMF4 reflect the details of the random fluctuation of the curve. Because the fluctuation of IMF1 and IMF2 is gentle and the trend is strong, the linear kernel function (Equation (16)) is used in LSSVM 1 and LSSVM 2. IMF3 and IMF4 fluctuate strongly, and the RBF kernel function (Equation (17)) is selected in LSSVM3 and LSSVM4. The daily load data sequences of the first 15 days of the forecast day were decomposed by VMD, and four sets of 15×24 load matrices were formed according to four frequency scales.

Step 3 Data processing. The input data includes load, weather, date and temperature. In order to make power load forecasting more accurate, these data need to be normalized. The specific operation is as follows:

1) Load data normalization

$$x_{ij}^L = \lg x_{ij} \quad (29)$$

Where x_{ij} represents the original load, and x_{ij}^L represents the normalized load in Equation (29).

2) Date normalization

The date type variable $D(d)$ is shown in Table 1.

Table 1. Date type variable

Date type	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
D(d)	0.8	0.9	0.9	0.9	0.9	0.6	0.5

3) Temperature normalization

$$T_i^L = (T_i - T_{i\min}) / (T_{i\max} - T_{i\min}) \quad (30)$$

Where T_i is the original temperature; $T_{i\min}$ and $T_{i\max}$ denote the minimal and maximal value of the original temperature; $i = 1, 2, \dots, n$; T_i^L is the normalized temperature in Equation (30).

4) Weather normalization

The definition of weather type variable $C(d)$ is shown in Table 2.

Table 2. Definition of daily types of weather $C(d)$

Types of weather	$C(d)$
Sunny	0
Cloudy	0.06
Overcast	0.12
Light rain	0.18
Moderate rain	0.25
Heavy rain	0.31
Sleet	0.43
Light snow	0.52
Moderate snow	0.56
Heavy snow	0.62
Frost	0.75
Fog	0.81
Breeze	0.87
Gale	0.93

Step 4 Load forecast based on LAVAFSA-SFLA-LSSVM. The Four groups of normalized load matrix and the environment data in the first 15 days were input into four groups of LSSVM to predict the load value of 24 hours on January 16 and February 16, 2019. The parameters of LSSVM in the prediction process were optimized by LAVAFSA-SFLA. The final prediction result is the sum of four LSSVM results.

6. Case Simulation and Prediction Evaluation

The actual historical data of a county in Jiaozuo City, Henan Province is used as an example. The power load of 24 hours on January 16 and February 16, 2019 are predicted. The AFSA-LSSVM, AFSA-SFLA-LSSVM, LAVAFSA-SFLA-LSSVM and VMD-LAVAFSA-SFLA-LSSVM are established. The maximum number of iterations is 100. The predicted values of the four models are compared with the actual loads as shown in Figure 2. $\gamma \in [0.1, 100]$; $\sigma \in [0.01, 10]$; $Visual_0 = 4$; $try-number = 50$; $\delta = 0.618$; $Step = 1$; $\beta = 1/2$.

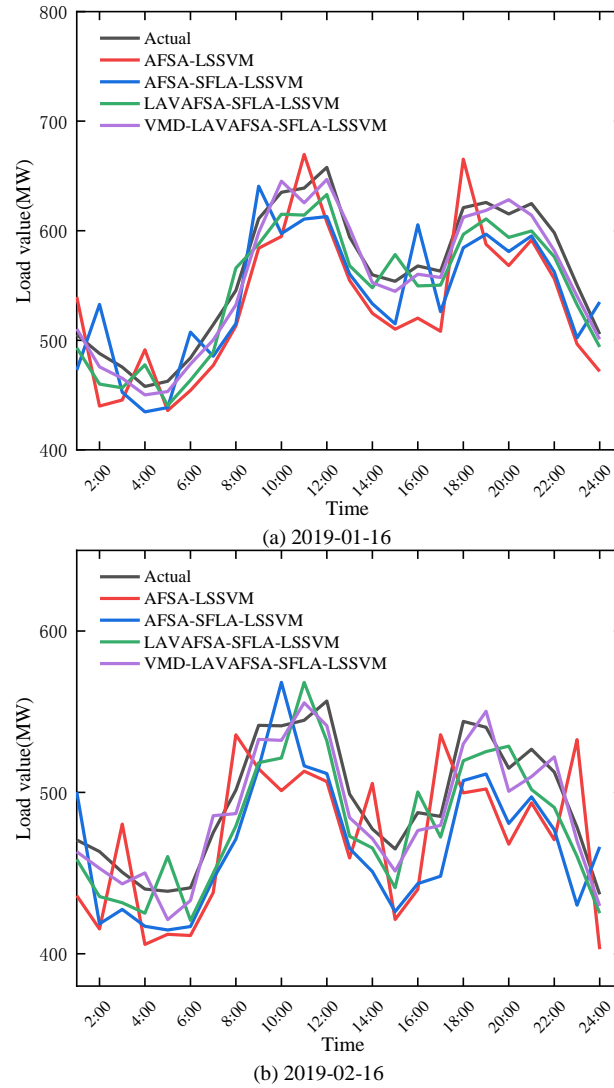


Figure 2. Original load and forecasting results of each combined model

The relative error and average error in the error evaluation index are shown in Equations (31) and (32):

$$E_{MAE} = \left| \frac{Y_i - Y_{Li}}{Y_i} \right| \times 100\% \quad (31)$$

$$E_{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{Y_i - Y_{Li}}{Y_i} \right| \times 100\% \quad (32)$$

Where E_{MAE} represents the relative error and E_{MAPE} represents the average error; Y_i is the true value; Y_{Li} is the predicted value; N is the number of predicted values.

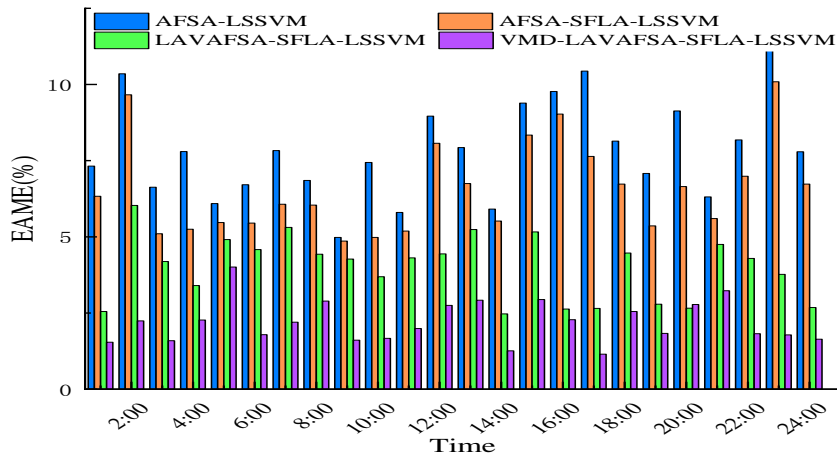


Figure 3. Prediction error values of four models on February 16, 2019

It can be seen from figure 3 that the relative error value of VMD-LAVAFSA-SFLA-LSSVM is the minimum within 24 hours of the forecast day. In addition, the average error of AFSA-LSSVM is 7.84%. The average error of AFSA-SFLA-LSSVM is 6.58%. The average error of LAVAFSA-SFLA-LSSVM is 4.36%. The average error of VMD-LAVAFSA-SFLA-LSSVM is 2.23%.

7. Conclusions

In order to improve the prediction accuracy, an improved artificial fish leapfrog algorithm based on variational mode decomposition is proposed to optimize the LSSVM prediction method. The main conclusions are as follows:

- 1) The improved artificial fish swarm leapfrog algorithm is proposed, and two complex functions are used to verify the superiority of the algorithm in the optimization accuracy and convergence speed.
- 2) The parameters of LSSVM are optimized by AFSA、AFSA-SFLA and LAVAFSA-SFLA respectively. The results show that the prediction effect of LAVAFSA-SFLA-LSSVM is better.
- 3) VMD was used to decompose the nonlinear and volatile daily load sequence. The obtained load matrix of different frequency scales was input into different LSSVM for prediction. The parameters of LSSVM are optimized by LAVAFSA-SFLA. The results show that the accuracy of load forecasting is improved and the prediction error is reduced to some extent by adding VMD into LAVAFSA-SFLA-LSSVM.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (No.61703144) and the department of Education project of Guizhou Province(NO.KY[2015]468); The authors gratefully acknowledge the helpful comments and suggestions of the reviewers.

References

1. A. H. Sanstad, S. McMenamin, A. Sukenik, G. L. Barbose, and C. A. Goldman, "Modeling an Aggressive Energy-Efficiency Scenario in Long-Range Load Forecasting for Electric Power Transmission Planning," *Applied Energy*, Vol. 128, pp. 265-276, September 2014
2. A. Zakariazadeh, S. Jadid, and P. Siano, "Stochastic Operational Scheduling of Smart Distribution System Considering Wind Generation and Demand Response Programs," *International Journal of Electrical Power Energy Systems*, Vol. 63, pp. 218-225, December 2014
3. J. J. Wang, L. Li, D. X. Niu, and Z. F. Tan, "An Annual Load Forecasting Model based on Support Vector Regression with Differential Evolution Algorithm," *Applied Energy*, Vol. 94, pp. 65-70, June 2012
4. C. Hamzacebi and H. A. Es, "Forecasting the Annual Electricity Consumption of Turkey using an Optimized Grey Model," *Energy*, Vol. 70, pp. 165-171, June 2014
5. E. Paparoditis and T. Sapatinas, "Short-Term Load Forecasting: The Similar Shape Functional Time-Series Predictor," *IEEE Transactions on Power Systems*, Vol. 28, No. 4, pp. 3818-3825, November 2013
6. R. Azimi, M. Ghofrani, and M. Ghayekhloo, "A Hybrid Wind Power Forecasting Model based on Data Mining and Wavelets

- Analysis,” *Energy Conversion and Management*, Vol. 127, pp. 208-225, November 2016
7. Y. L. Yang, J. X. Che, Y. Y. Li, Y. J. Zhao, and S. L. Zhu, “An Incremental Electric Load Forecasting Model based on Support Vector Regression,” *Energy*, Vol. 113, pp. 796-808, October 2016
 8. P. H. Li, Y. G. Li, Q. Y. Xiong, Y. Chai, and Y. Zhang, “Application of a Hybrid Quantized Elman Neural Network in Short-Term Load Forecasting,” *International Journal of Electrical Power and Energy Systems*, Vol. 55, pp. 749-759, February 2014
 9. Y. L. Wang, D. X. Niu, and L. Ji, “Short-Term Power Load Forecasting based on IVL-BP Neural Network Technology,” *Systems Engineering Procedia*, Vol. 4, pp. 168-174, 2012
 10. Q. Wu, “Hybrid Model based on Wavelet Support Vector Machine and Modified Genetic Algorithm Penalizing Gaussian Noises for Power Load Forecasts,” *Expert Systems with Applications*, Vol. 38, No. 1, pp. 379-385, January 2011
 11. E. Ceperic, V. Ceperic, and A. Baric, “A Strategy for Short-Term Load Forecasting by Support Vector Regression Machines,” *IEEE Transactions on Power Systems*, Vol. 28, No. 4, pp. 4356-4364, November 2013
 12. P. F. Pai and W. C. Hong, “Support Vector Machines with Simulated Annealing Algorithms in Electricity Load Forecasting,” *Energy Conversion and Management*, Vol. 46, No. 17, pp. 2669-2688, October 2005
 13. C. Li, S. Li, and Y. Liu, “A Least Squares Support Vector Machine Model Optimized by Moth-Flame Optimization Algorithm for Annual Power Load Forecasting,” *Applied Intelligence*, Vol. 45, No. 4, pp. 1166-1178, December 2016
 14. C. A. D. A. Padilha, D. A. C. Barone, and A. D. D. Neto, “A Multi-Level Approach using Genetic Algorithms in an Ensemble of Least Squares Support Vector Machines,” *Knowledge-based Systems*, Vol. 106, pp. 85-95, August 2016
 15. M. J. Mahjoob, M. Abdollahzade, and R. Zarringhalam, “GA based Optimized LS-SVM Forecasting of Short Term Electricity Price in Competitive Power Markets,” in *Proceedings of the Industrial Electronics and Applications*, pp. 3-5, Singapore, June 2008
 16. W. M. Lin, C. S. Tu, R. F. Yang, and M. T. Tsai, “Particle Swarm Optimisation Aided Least-Square Support Vector Machine for Load Forecast with Spikes,” *IET Generation Transmission and Distribution*, Vol. 10, No. 5, pp. 1145-1153, April 2016
 17. R. Dong, J. Xu, and B. Lin, “ROI-based Study on Impact Factors of Distributed PV Projects by LSSVM-PSO,” *Energy*, Vol. 124, pp. 336-349, April 2017
 18. A. A. Abdoos, “A New Intelligent Method based on Combination of VMD and ELM for Short Term Wind Power Forecasting,” *Neurocomputing*, Vol. 203, pp. 111-120, August 2016
 19. K. J. He, Y. H. Chen, and G. K. F. Tso, “Forecasting Exchange Rate using Variational Mode Decomposition and Entropy Theory,” *Physica A: Statistical Mechanics and its Applications*, Vol. 501, pp. 15-25, November 2018
 20. L. Xiang, Z. Q. Deng, and A. J. Hu, “Forecasting Short-Term Wind Speed based on IEWT-LSSVM Model Optimized by Bird Swarm Algorithm,” *IEEE Access*, Vol. 7, pp. 2169-3536, May 2019
 21. X. Yan and N. A. Chowdhury, “Mid-Term Electricity Market Clearing Price Forecasting: A Hybrid LSSVM and ARMAX Approach,” *International Journal of Electrical Power and Systems*, Vol. 53, pp. 20-26, December 2013
 22. W. Shen, X. P. Guo, C. Wu, and D. S. Wu, “Forecasting Stock Indices using Radial Basis Function Neural Network Optimized by Artificial Fish Swarm Algorithm,” *Knowledge-based Systems*, Vol. 24, No. 3, pp. 378-385, April 2011
 23. X. T. Hu, H. Q. Zhang, Y. A. Huang, and Z. P. Yin, “A Novel Self-Adaptation Hybrid Artificial Fish-Swarm Algorithm,” *IFAC Proceedings Volumes*, Vol. 46, No. 5, pp. 583-588, April 2013
 24. J. M. Xiao, X. M. Zheng, X. H. Wang, and Y. F. Huang, “A Modified Artificial Fish-Swarm Algorithm,” in *Proceedings of the World Congress on Intelligent Control and Automation (WCICA)*, pp. 3456-3460, Dalian, China, June 2006
 25. X. M. Cheng, X. D. Zhang, L. Zhao, A. D. Deng, Y. Q. Bao, Y. Liu, et al., “The Application of Shuffled Frog Leaping Algorithm to Wavelet Neural Networks for Acoustic Emission Source Location,” *Comptes Rendus Mécanique*, Vol. 342, No. 4, pp. 229-233, April 2014

Haizhu Yang received the Ph.D. from Beijing Jiaotong University in 2005. He is an associate professor working in the School of Electrical Engineering and Automation, Henan Polytechnic University. His research interests include power electronics and electrical drive and power systems and automation.

Zhaoyang Jiang is currently pursuing the M.S degree in the School of Electrical Engineering and Automation, Henan Polytechnic University. His research interests include power system and automation and power load forecasting.

Menglong Li is currently pursuing the M.S degree in the School of Electrical Engineering and Automation, Henan Polytechnic University. His research interest includes power system and automation.

Peng Zhang is a lecturer working in the School of Electrical and Information Engineering, Tianjin University. His research interest includes Integrated energy system planning and operation.