ARCHIVES OF ELECTRICAL ENGINEERING VOL. 68(4), pp. 907–923 (2019)

DOI 10.24425/aee.2019.130691

Short-term load prediction model combining FEW and IHS algorithm

MINGXING YU[1], JIAZHENG ZHU[2], LI YANG[3]

[1] *Department of Information Engineering, Chaoyang Teachers College 122000, Chaoyang, P.R. China*

[2] *Huludao Power Supply Company, State Grid Liaoning Electrical Power Co., Ltd. 125001, Huludao, P.R. China*

[3] *School of Equipment Engineering, Shenyang Ligong University 110159, Shenyang, P.R. China*
e-mail: beihesa@hotmail.com

(Received: 13.01.2019, revised: 15.07.2019)

Abstract: Accurate prediction of power load plays a crucial role in the power industry and provides economic operation decisions for the power operation department. Due to the unpredictability and periodicity of power load, an improved method to deal with complex nonlinear relation was adopted, and a short-term load forecasting model combining FEW (fuzzy exponential weighting) and IHS (improved harmonic search) algorithms was proposed. Firstly, the domain space was defined, the harmony memory base was initialized, and the fuzzy logic relation was identified. Then the optimal interval length was calculated using the training sample data, and local and global optimum were updated by optimization criteria and judging criteria. Finally, the optimized parameters obtained by an IHS algorithm were applied to the FEW model and the load data of the Huludao region (2013) in Northeast China in May. The accuracy of the proposed model was verified using an evaluation criterion as the fitness function. The results of error analysis show that the model can effectively predict short-term power load data and has high stability and accuracy, which provides a reference for application of short-term prediction in other industrial fields.

Key words: evaluation criteria, exponential fuzzy time series, fitness function, improved harmony search algorithms, load forecasting, optimal interval length

1. Introduction

Load forecasting of a power system is an important part of power system planning, the basis of the economic operation of a power system, and a necessary condition to ensure the reliability and stability of a power system [1]. Due to the unpredictability, timeliness, periodicity and multicharacteristics of power load, accurate prediction of power load has become an important and arduous task. Either underestimation or overestimation of power load will bring huge challenges to a power system [2, 3]. For example, the underestimation of load may affect the reliability of electric energy, making energy storage fail to meet the demand, while overloading will lead to unnecessary equipment operation, a low efficiency distribution rate and increased operating costs [4]. Therefore, accurate load forecasting can facilitate scheduling decisions and has engineering application value [5].

Due to the complexity and non-linear characteristics of actual load data, a model with higher prediction accuracy based on specific factors of the model is needed, which is also the primary objective of this paper. Rasul Enayatifar, a professor at the Federal University of Minas Gerais in Brazil, applied evolutionary algorithms to many engineering technologies [6]. Harmony Search (HS) is a typical heuristic evolutionary algorithm, which is favored by researchers due to its good practical application effect [7]. On the other hand, many researchers have reported a large number of fuzzy time series, which have been applied in many fields [8]. Qiang Song analyzed linguistic data using fuzzy relational Equations [9] and proposed a prediction method using fuzzy time series. Shyi-Ming Chen improved Qiang Song's method and proposed a relatively simple fuzzy time series prediction method, which requires less time in the process of addressing combination operation [10]. Furthermore, Hui-Kuang [11] proposed a weighted fuzzy time series model, which can deal with recursive and weighted problems between two fuzzy relations. Ching-Hsur Cheng [12] from National Yunlin University of Science and Technology (2009) evaluated the weighted fuzzy time series prediction method proposed by Chen and Song via a dynamic model. On the basis of literature [11], Lee [13] proposed an exponential weighted fuzzy time series model. Although this model has high application value, it is still necessary to pay attention to the problem of dividing the interval length of the theoretical domain. Kun Huarng (2001) proposed a method based on average length distribution to determine the interval length of fuzzy time series [14]. Later, he proposed an improved fuzzy time series prediction method based on the interval length of a ratio interval, and calculated the sensitivity of different percentiles [15]. Shyi-ming Chen proposed a new fuzzy time series model combined with a genetic algorithm, defined the interval length of a domain interval, and predicted the enrollment situation of the University of Alabama by using high-order fuzzy time series [10]. Furthermore, Cagdas H. Aladag proposed a new method to define high-order fuzzy time series by using a feed-forward neural network and optimized interval length by using univariate constraints [16]. Erol Egrioglu used the algorithm based on golden section search and parabola interpolation to determine the optimal interval length [17]. This algorithm was able to find the interval length of high-order fuzzy time series when optimizing the univariate constraint function. The enrollment of the University of Alabama was predicted using the proposed method, which verified the accuracy of this method.

Based on the above analysis, the following three key information can be obtained. The first one is that the fuzzy time series model is very suitable for predicting the data with unpredictable, nonlinear, periodic and multi-characteristic characteristics such as short-term power load. Secondly, it is recommended to use a harmony searchalgorithm to predict power load data, because an HS algorithm has a higher degree of fitting. The third point is to use appropriate methods to determine the interval length of the domain, which can improve the accuracy of the prediction results.

The content of this paper can be divided into three parts. The first part is to build a fuzzy exponential weighted time series model, define the domain space and initialize the harmonic memory bank, so as to prepare for the following analysis. The second part is to update the local optimum and global optimum by using optimization criteria and judging criteria. In the third part, based on the three basic features of the harmony searchalgorithm, the improved harmony searchalgorithm is applied to the FEWA model in the first part to output the optimal prediction results. The second and third parts are modified by evaluating criterion error parameters in the iterative process, the purpose of which is to reduce the prediction error of the training sample. When the iteration termination condition is met, the algorithm stops. Finally, the validity and accuracy of the proposed model are verified by different evaluation criteria.

2. Construction of prediction model

2.1. Fuzzy time framework

Fuzzy Time Series (FTS) is a statistical method for data processing of random processes of discrete indicators. To classify and describe the complete model of aspects, the definition and construction process of the FTS were described.

U is defined as a domain containing n subspaces, $U = \{u_1, u_2, \ldots, u_n\}$, then the fuzzy set [18] on the domain U is

$$A = \frac{f_A(u_1)}{u_1} + \frac{f_A(u_2)}{u_2} + \ldots + \frac{f_A(u_n)}{u_n}, \tag{1}$$

where f_A is the fuzzy membership function, $f_A(u_i)$ is the membership degree of fuzzy set A, which meets $f_A(u_k) \in [0, 1]$, $1 \leq i \leq n$.

Suppose $Y(t)$ is a subset of the real number domain, where $t = \ldots, 0, 1, \ldots, f_i(t)$ is defined as a set of fuzzy sets, satisfying $F(t) = \{f_1(t), f_2(t), \ldots\}$, then $F(t)$ is the fuzzy time series defined on $Y(t)$ [19].

The logical relationship $F(t-1) \rightarrow F(t)$ represents that $F(t)$ is determined by $F(t-1)$, of which the equation is shown as [9]:

$$F(t) = F(t-1) \circ R(t, t-1), \tag{2}$$

where $F(t)$ is a first-order model, $\circ$ is the synthetic operation symbol, R is the fuzzy relation of $F(t)$.

Given $F(t-1) = A_i$, $F(t) = A_j$, then the fuzzy relation [9] is $A_i \rightarrow A_j$, where A_i is denoted as the event before fuzzy, A_j is the event after fuzzy. If the fuzzy relationship is repeated, it only persists once.

2.2. Fuzzy exponentially weight algorithm

A fuzzy exponentially weight (FEW) algorithm, proposed by H.K. Yu of Feng Chia University in 2005 [10], is an algorithm based on the traditional fuzzy time series framework, which can determine recursive fuzzy relations and assign weights to multiple fuzzy relationships. On the basis of Yu, Muhammad Hisyam Lee from Indonesia's national bureau of statistics improved the algorithm [12] and proposed a diversified weighting scheme with exponential growth of time weight and higher prediction accuracy. The parameter c in the Suhartono algorithm is randomly selected, however, some specific values selected can improve the prediction accuracy of the model. Therefore, on the basis of literature [10, 12], this paper makes appropriate modifications to the proposed method. To obtain a better estimation of parameter c, improved harmony search (IHS) was not used in the initial analysis stage, and only a basic FEW algorithm was used to study the random samples. Suhartono evaluated the influence of different c values on the prediction accuracy, and confirmed that there is an accurate prediction accuracy when $c = 1.1$. Therefore, in Equation (3) and the following studies, c is defined as constant 1.1 ($c \geq 1$).

The calculation procedure of an FEW algorithm is shown as follows:

1. Define the domain and an observation interval.
2. Establish the Fuzzy Logical Relationships (FLRs).
3. Establish the Fuzzy Logical Relationships Groups (FLRGs) of the FLR in step 2.
4. Select the best order of the FLR for prediction.
5. Solution fuzzification.
   Suppose predicting $F(t) = \{\boldsymbol{A}_{j1}, \boldsymbol{A}_{j2}, \ldots, \boldsymbol{A}_{jk}\}$, the solution fuzzification matrix $M(t) = \{\boldsymbol{M}_{j1}, \boldsymbol{M}_{j2}, \ldots, \boldsymbol{M}_{jk}\}$, $M(t)$ is the solution fuzzy sandwich matrix related to $\boldsymbol{A}_{j1}, \boldsymbol{A}_{j2}, \ldots, \boldsymbol{A}_{jk}$.
6. Considering a weighting coefficient, then the equivalent weighting coefficient of $\boldsymbol{A}_{j1}, \boldsymbol{A}_{j2}, \ldots, \boldsymbol{A}_{jk}$ is

$$\boldsymbol{W}(t) = \left[w_1', w_2', \ldots, w_k'\right] = \left[\frac{c^1}{\sum\limits_{h=1}^{k} c^{h-1}}, \frac{c^2}{\sum\limits_{h=1}^{k} c^{h-1}}, \ldots, \frac{c^{k-1}}{\sum\limits_{h=1}^{k} c^{h-1}}\right]. \tag{3}$$

7. Calculating results. The predicted value is equivalent to the product of a de-fuzzy matrix and a weighted transpose matrix.

2.3. Harmony search algorithm

The Harmony Search Algorithm (HSA) is an intelligent optimization algorithm inspired by a music phenomenon [20]. Just as a musical instrument playing discrete notes according to the players' experience, the players improve the playing notes based on aesthetic standards. In the same way, design variables allocate some discrete values according to the computational intelligence, and design variables in computer memory are continuously improved according to the objective function, so as to realize the optimization of the algorithm [21, 22].

The basic HSA has three characteristics. Considering the computational intelligence and randomness [23], the updated calculation value of the design variable is

$$x_i^N \in \begin{cases} x_i(k) \in \{x_1(1), x_1(2), \ldots, x_1(K_i)\} & p = p_R \\ x_i(k) \in \left\{x_i^1, x_i^2, \ldots, x_i^{\text{HMS}}\right\} & p = p_M \\ x_i(k \pm m) & p = p_P \end{cases}, \tag{4}$$

where design variable x_i^N is randomly selected from all candidate discrete sets $\{x_i(1), x_i(2), \ldots, x_i(K_i)\}$, and the random selection probability is p_R. A better value can be also selected from $\{x_i^1,\ x_i^2, \ldots, x_i^{\mathrm{HMS}}\}$ stored by the computer, with a search probability of p_M. Or through fine tuning the adjacent value $x_i(k \pm m)$, a better storage value $x_i(k)$ can be selected from the adjacent value group, with the tuning probability of p_P, where p is random probability.

The HSA is also known as the Harmony Memory Algorithm (HMA) because of its memory storage function. Then the design variable set vector of the HMA is $(x_1^j,\ x_2^j,\ \ldots,\ x_n^j)$, the value of the objective function is also stored in the vector of each design variable. Harmony memory bank [24] M_H is a set of iteratively updated optimization design variables, of which the expression is

$$M_H = \left[\begin{array}{cccc|c} x_1^1 & x_2^1 & \cdots & x_n^1 & f\left(\boldsymbol{x}^1\right) \\ x_1^2 & x_2^2 & \cdots & x_n^2 & f\left((\boldsymbol{x}^2\right) \\ \vdots & \ldots & \ldots & \ldots & \vdots \\ x_1^{HM} & x_2^{HM} & \cdots & x_n^{HM} & f\left(\boldsymbol{x}^{HM}\right) \end{array}\right], \tag{5}$$

where $f(\boldsymbol{x}^j)$ is the objective function value (the j-th variable), $j = 1, 2, \ldots$

If the newly-generated design variable x^N is superior to the worst design variable x^W stored in M_H then x^N is exchanged with x^W, i.e.

$$\begin{aligned} x^W &\notin M_H\,, \\ x^N &\notin M_H\,. \end{aligned} \tag{6}$$

2.4. Improved harmony searchalgorithm

Since the Basic HSA is subjected to disadvantages such as slow convergence speed and unstable prediction [24], to improve the global optimal search ability, an improved harmony search (IHS) algorithm was proposed in literature [25]. On this basis, this paper further carried out optimization, and the specific steps are as follows:

1. Parameter initialization

Suppose $f(x)$ is the objective function, x is a set of decision variables, n is the number of decision variables, $\boldsymbol{X}_i$ is the range of decision variables, $\boldsymbol{X}_i^L$ is the lower limit, $\boldsymbol{X}_i^H$ is the upper limit. Therefore, the minimized unconstrained optimization problem condition of $f(x)$ is

$$x_i \in X_i, \quad i = 1, 2, \ldots, n. \tag{7}$$

During this process, harmony memory utilization, tonal tuning rate, and termination criteria are saved, and these parameters can be used for updated calculation value of the design variable. An HM process is similar to a genetic algorithm library and particle swarm optimization algorithm library. The randomly-generated vector number of a design variable group is stored randomly in harmony memory bank M_H.

2. New harmony improvisation

A harmony search algorithm consists of three basic features described in section 2.3. The new harmony generated by features is improvisation, of which the decision parameter is

$$x_i^j \in \begin{cases} x_i^j \in \left\{x_i^1,\ \ldots,\ x_i^{\mathrm{HMS}}\right\} & p = p_M \\ x_i^j \in X_i & p = 1 - p_M \end{cases}, \tag{8}$$

where p_M is the harmony storage search probability, $\boldsymbol{X}_i$ is the out of-library feasible domain.

The selected variable needs to be tested to determine whether it needs tone tuning, and its decision variable is

$$x_i^j \in \begin{cases} x_i^j \pm R(0,1) b_w & p = p_P \\ x_i^j & p = 1 - p_P \end{cases}, \tag{9}$$

where b_w is the random width, $R(0,1)$ represents the random number within [0, 1].

3. Dynamic parameter setting

p_M and p_P are two important fixed values in the basic HS algorithm, where p_M controls the improvisation mode and p_P determines the search process. To increase the search scope as far as possible and avoid falling into the local optimum, the p_M is reduced and the p_P is increased. The dynamic parameter is

$$\begin{aligned} p_M &= p_{M\,\max} - \frac{p_{M\,\max} - p_{M\,\min}}{M} k, \\ p_P &= p_{P\,\max} - \frac{p_{P\,\max} - p_{P\,\min}}{M} k, \end{aligned} \tag{10}$$

where M is the total number of iterations, k is the current iteration number, $p_{M\,\max}$ is the searching probability maximum, $p_{M\,\min}$ is the searching probability minimum, $p_{P\,\max}$ is the tuning probability maximum, $p_{P\,\min}$ is the tuning probability minimum.

4. Improved tuning algorithm

The tuning of tones is determined by b_w. The global and local harmonic positions are used to adjust a tone scalar in the real time to increase the ability to approach the optimal harmony. If b_w is reduced, the global optimization can be achieved in the search process, but it is weak to deal with problems with more local optimal values. At this time, a larger b_w can help improve the search ability of the algorithm and increase the local optimal performance. Therefore, in the iterative process, the tone tuning process relationship corresponding to the selected variables from a new harmonic library is

$$b_{wt}^{k+1} = w b_{wt}^k + c_1 \left(x_i^b - b_{wt}^k \right) + c_2 \left(x_g^b - b_{wt}^k \right), \tag{11}$$

where b_{wt}^k is the tone scalar of the selected harmony under k iterations, w is the inertia weight factor, c_1 and c_2 are the tuning factors, x_i^b is the position of local optimal harmony, x_g^b is the position of global optimum harmony, then the tone tuning updating equation is

$$x_i^j = \begin{matrix} 1 & \mathrm{Rand}() \le \mathrm{sig}(b_{wi}) \\ 0 & \mathrm{Rand}() \le \mathrm{sig}(b_{wi}) \end{matrix}, \tag{12}$$

$$\mathrm{sig}(b_{wi}) = \frac{1}{1 + e^{-b_{wi}}}, \tag{13}$$

where Rand() is a random number.

5. Universal harmonic optimization

If the optimal harmony is not found in an iteration process, the XOR (Exclusive OR) operation of the optimal harmony and poor improvisation in the harmony library is carried out to realize the general purpose of improvisation and optimization of the harmony library, so that the diversity and optimization of global search results can be ensured.

$$x_i^j = x_g^b \oplus x_i^j . \tag{14}$$

3. Application of FEW-IHS

The proposed FEW-IHS model is suitable for processing data with multiple complex influencing factors. Literature finds that power load data have the characteristics of unpredictability, timeliness, periodicity and multiple characteristics. Power load data is a kind of complex data influenced by multiple weather information such as rainfall, maximum temperature, minimum temperature, wind speed and wind direction. Using the FEW-IHS model to process power load data can give better prediction accuracy, the specific processes are shown in Figure 1.

1. The domain U of the short-term load data to be analyzed is determined to define the fuzzy set.
2. The domain U is divided into n subintervals, and n is calculated by the average criterion. Determine the adaptive coefficient α, where α is the random number satisfying $\alpha \in [0, 1]$.
3. Initializes the harmony memory bank.
4. According to the objective function, determine the local optimal harmonic position and the global optimal harmonic position, and update the harmony search probability p_M and adjust the probability p_P.
5. According to the new harmony generated by p_M, determine whether the newly generated harmony is better than the optimal harmony in the memory bank using Equation (14).
6. After the improvisation, tuning the tone according to p_P and looking for judgment and general optimal harmonic processing.
7. Determine the iteration termination condition and output the optimal decision location parameters.
8. The fuzzy power load data extracted in steps 3–7 can be used to process information related to the update interval, establish the FLR based on FTS, and adjust the optimal order of the FLR.
9. An FLRG is established on the basis of establishing the FLR.
10. In the initial prediction calculation, the initial prediction data at all times are the initial training set, which is calculated through Equations (7)–(9) described in section 2.4.
11. The final results are calculated. The calculation is completed through introducing the preliminary predicted values into the training data set.

In order to measure the accuracy of the prediction model, the following evaluation criteria were used for analysis.

Mean Absolute Error

$$e_{\mathrm{MAE}} = \frac{1}{n}\sum_{i=1}^{n}\left|\overline{y}_i - y_i\right|. \tag{15}$$

Mean Absolute Percentage Error

$$e_{\mathrm{MAPE}} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|\overline{y}_i - y_i\right|}{y_1} \times 100\%. \tag{16}$$

Mean Square Error

$$e_{\mathrm{MSE}} = \frac{1}{n}\sqrt{\sum_{i=1}^{n}(\overline{y}_i - y_i)^2}. \tag{17}$$

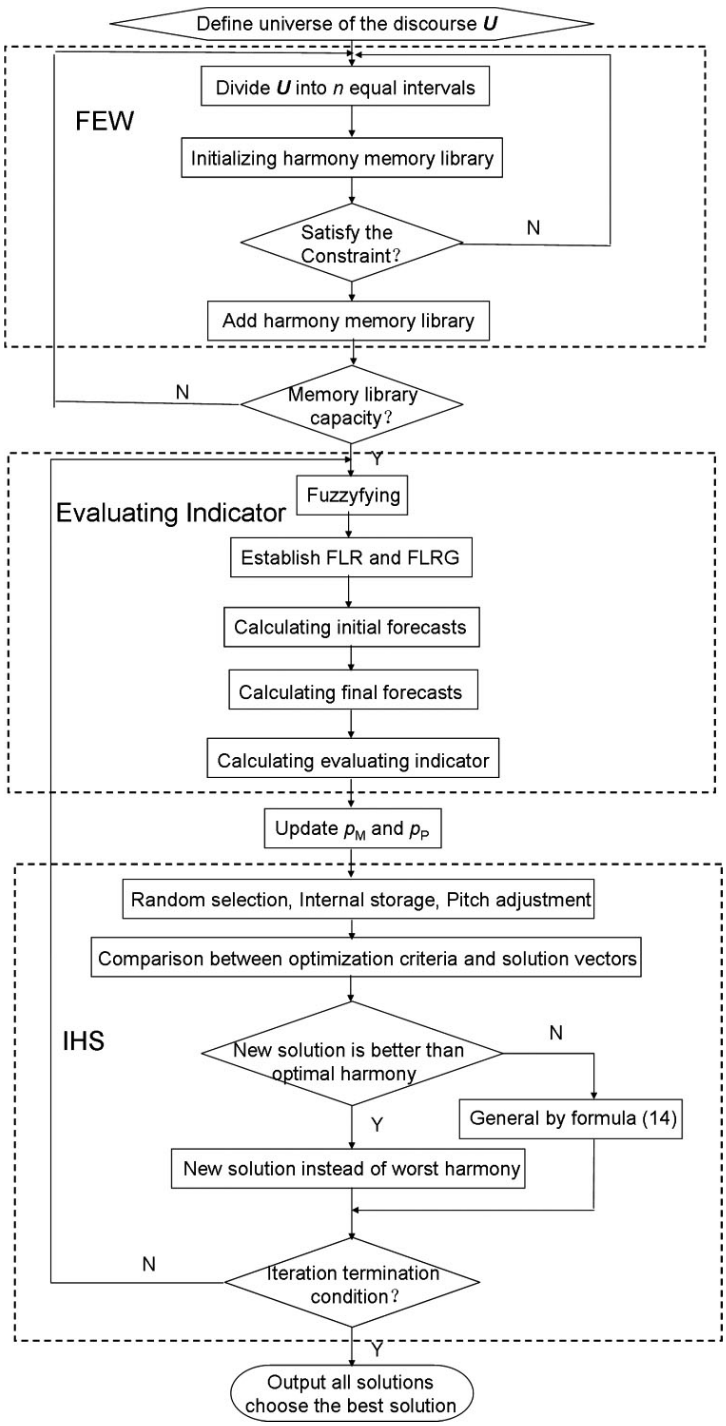


Fig. 1. FEW-IHS application flow chart

Mean Square Percent Error

$$e_{\text{MSPE}} = \frac{1}{n}\sqrt{\sum_{i=1}^{n}\left(\frac{|\overline{y}_i - y_i|}{y_1}\right)^2} \times 100\%, \tag{18}$$

where n is the number of samples, $\overline{y}_i$ is the predicted value, y_i is the actual value.

12. Analysis of the prediction results using evaluation criteria.

4. Analysis of experimental results

This section is extended from the following two perspectives:

1. To verify the effectiveness of the proposed FEW-IHS model in a power load data training set.
2. Taking the evaluation criteria as the fitness function, the FEW-IHS model was evaluated and compared with other typical algorithms to verify its stability and accuracy.

The experimental data of power load were obtained from the daily average load and meteorological conditions of the Huludao region in the Northeast China from May 1, 2013 to May 31, 2013. For the convenience of research, load data were averaged on a daily basis, and a total of 31 data cases were processed. The sampling frequency of load information was set to 15 minutes. Meteorological information included a wind direction, wind speed, maximum temperature, minimum temperature and a week. It should be noted that a wind direction, week and wind speed were converted to digital variables.

Proposed by professor Vapnik et al., from the NEC research institute in 1995 [26], an SVM (support vector machine) is an algorithm based on statistical theory to obtain the actual minimum structural risk value according to the minimum structural risk criterion. The basic idea of the algorithm is to map the original input spatial data to the high-dimensional characteristic space through nonlinear transformation. The essence of this algorithm is a classical quadratic programming problem, which can avoid local optimization. In vector machines, the selection of parameter ε, c, σ is closely related to the accuracy of the prediction model [27]. ε is used for training the fit data, and the larger its value is, the flatter the estimated regression function is; c determine the degree to which the target empirical risk is minimized; σ controls the width of Gaussian function and reflects the distribution range of training data. The three parameters affect the accuracy of the model in different ways. On the basis of Niu Dongxiao (2010) and Kavousi–Fard (2014) [4, 27], the parameter results of the SVM were determined in this paper: $[\varepsilon, c, \sigma] = [0.12, 1650, 0.48]$. The SVM fully takes into account all kinds of factors affecting the load, has a relatively fast convergence speed, and is easy to find the global optimal solution. However, due to the large demand for storage and the difficulty in programming, it cannot determine whether the knowledge in the data is redundant and the role of knowledge in the data. An SVM can achieve an ideal effect in predicting the load showing a flatter curve while it does not work well for small and medium-sized power grids with strong random fluctuations.

To verify the effectiveness of the proposed model in short-term power load prediction,, the predicted results of EM-IHS was compared with the predicted results of particle swarm optimisation (PSO) [28], improved particle swarm optimization (IPSO) [29], the artificial neural network ANN [30] and the SVM [31], respectively. Figure 2 shows the curve comparison between the predicted load of different algorithms and the actual load in May. It can be preliminarily seen that the FEW-IHS model has a high similarity with the actual load curve at week 2 and week 4, showing a good tracking ability. Next, the effectiveness and accuracy of the proposed model were analyzed through evaluation indexes.

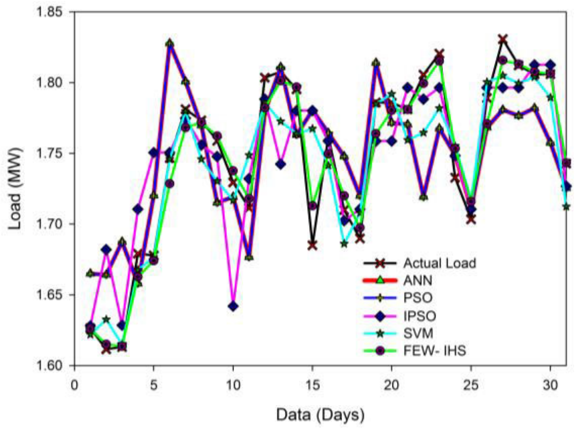
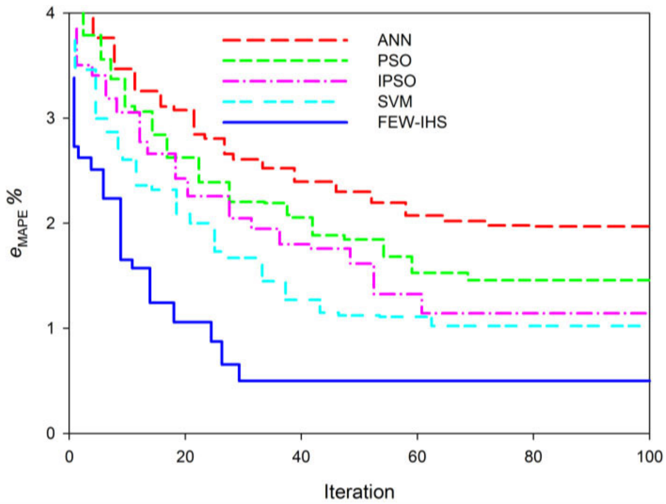


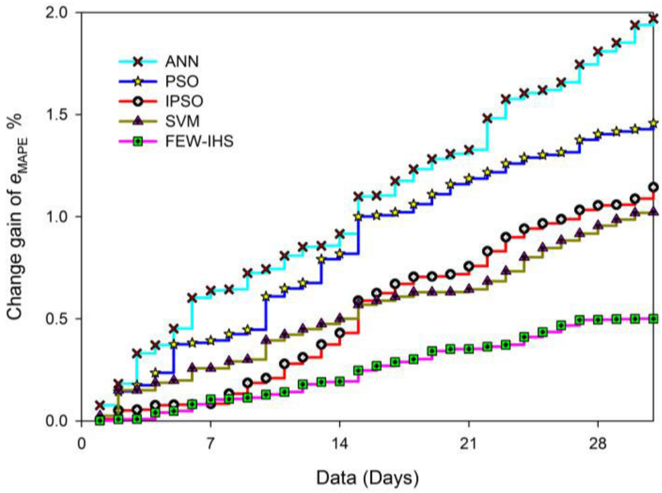
Fig. 2. Comparison of actual load data with prediction data of various algorithms

The mean absolute percentage error (e_{MAPE}) is one of the important indicators for error analysis, which represents the degree of dispersion of predicted values. The smaller the value is, the better the prediction model is than the strategy of using the mean value for prediction. In order to show the high convergence and strong searching ability of the FEW-IHS model, Figure 3(a) shows the e_{MAPE} distribution within 100 stable iterations of various models. It can be seen that as the number of iterations increased, the e_{MAPE} gradually decreased, and the ANN algorithm had the largest number of iterations when it converged to 65. The SVM algorithm was iterated 63 times, with the e_{MAPE} value reduced to 1.02. The FEW-IHS algorithm converged after only 28 iterations, indicating its fast convergence ability. By the 11th iteration, the e_{MAPE} of the FEW-IHS model had been decreased by 72%. Figure 3(b) shows the comparison chart of e_{MAPE} gain among various algorithms within a month. It can be found that in the initial stage, the difference in the e_{MAPE} among various algorithms was not significant. As the number of days increased, the ANN algorithm had the most significant increase in the e_{MAPE}; the e_{MAPE} of the PSO algorithm, increased significantly on the 4th, 10th and 15th days; the e_{MAPE} of IPSO algorithm suddenly overtook SVM on day 15; the e_{MAPE} of the proposed FEW-IHS model maintained a slow growth, with a small fluctuation only on the 15th day, and the e_{MAPE} was maintained at a low level all the time. In conclusion, the IHS method has strong search ability and is reasonable in optimizing power load data. However, it is difficult to estimate the difference between the predicted value and the observed value through the e_{MAPE} alone, so the error analysis of the predicted value is still needed.

When the model overtrains the performance of a group of data, sometimes overfitting occurs. In other words, the model has memorized the training data, and overfitting occurs instead of summarizing from learning and changing trends. Overfitting often occurs in prediction models

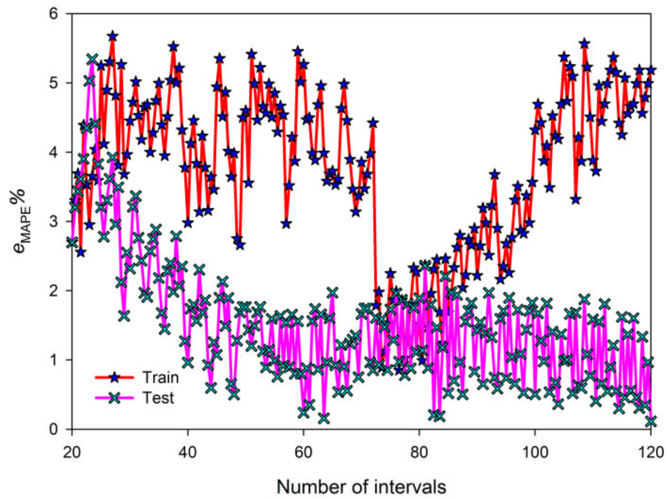


(a)

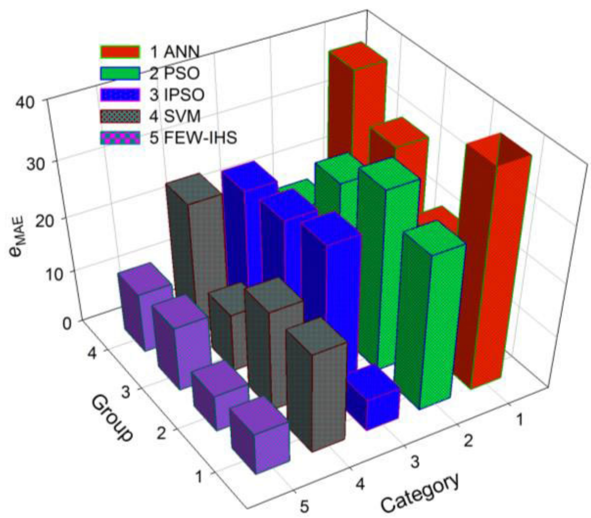


(b)

of optimization algorithms. In order to control the overfitting problem of the proposed FEW-IHS model [32], the determination time of fitting in the FTS model should first be controlled. In the process of dividing the domain space U, the length of each subinterval should not be too small; otherwise, different data in the training set may have the same fuzzy set. In the training set of each FTS model, if the determined interval length is zero, the total error will be close to zero, but the prediction result error rate will be very large. Figure 3(c) shows the change of the e_{MAPE} with the change of interval. It can be seen that when the interval is above 90, the larger the interval, the higher the prediction error rate is; within the range of an interval between 30 and 70, the error value is still large and unacceptable. In this paper, based on the mean value method proposed in literature [29], the effective length of the initial interval was calculated to be 460 and the interval



(c)



(d)

Fig. 3. Contrast charts of evaluation criteria: (a) convergence velocity; (b) e_{MAPE} gain; (c) variation of e_{MAPE} in different intervals; (d) comparison of e_{MAE} distribution

value was calculated to 82. The number of interval length was kept at the set value, but the interval length of each interval was constantly changed in the iterative process, that is to say, all the intervals in the domain space could not become very small, so the phenomenon of overfitting could be effectively controlled.

The mean absolute error (e_{MAE}) is one of the comprehensive indexes of error analysis, which synthesizes the absolute value of error and takes the average, so as to avoid the influence caused by the cancellation of the positive and negative prediction errors. To determine the accuracy of weekly load more clearly, short-term load cases of the eastern region in May were grouped according to Table 1. The e_{MAE} distribution of the mean absolute error of power load in the four groups is shown in Figure 3(d). It can be seen that in the first group the e_{MAE} of the IPSO algorithm was the least, which was 6.08, followed by the that (8.01) of the FEW-IHS algorithm; in the third group, the e_{MAE} of the SVM algorithm was 10.91, followed by that (12.19) of the FEW-IHS algorithm; in the second and fourth group, the e_{MAE} of FEW-IHS algorithm was the lowest. This shows that the actual error value of the power load predicted by the FEW-IHS algorithm is small. The e_{MAE} alone could not completely judge the fit of the model, and then the accuracy was analyzed through the mean square error (e_{MSE}).

Table 1. Grouping table of short-term load cases of eastern region in May

Group number	**Start**	**End**	**Testing day**
1	May 1st	May 7th	May 8th
2	May 8th	May 14th	May 15th
3	May 15th	May 21st	May 22nd
4	May 22nd	May 28th	May 29th

According to the classification method in Figure 3(d), the grouping form in Table 1 calculates the mean absolute percentage error (e_{MAPE}) of power load of the four groups, and the distribution cloud diagram is shown in Figure 4. It can be seen that the IPSO algorithm had the lowest e_{MAPE} in the first group, which was 0.37; the SVM algorithm had the lowest e_{MAPE} in the third group,

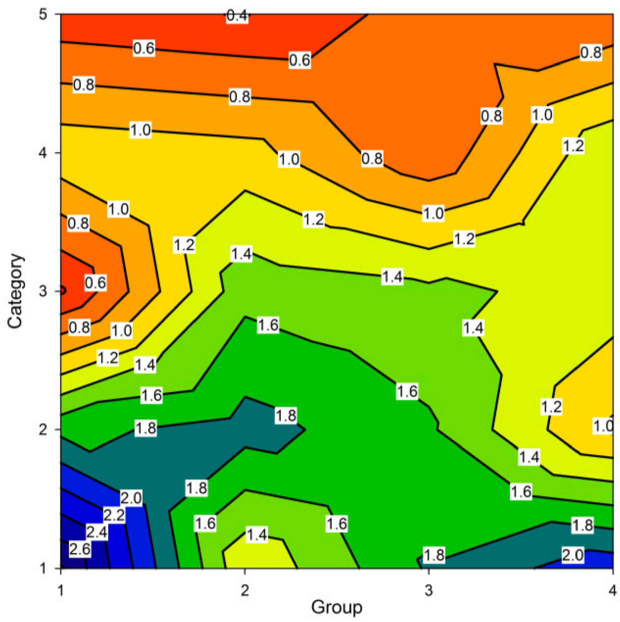
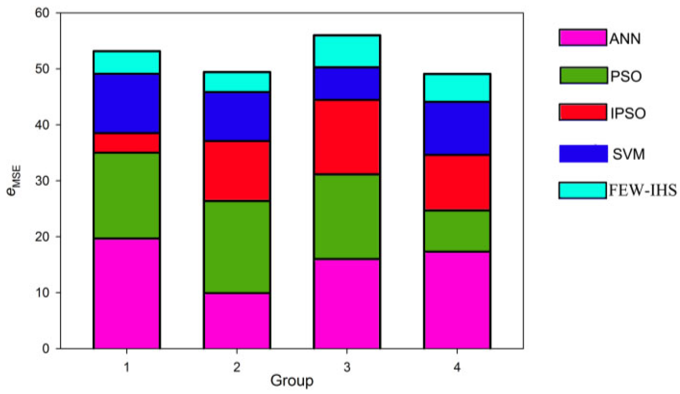


Fig. 4. Variation of e_{MAPE} in different intervals

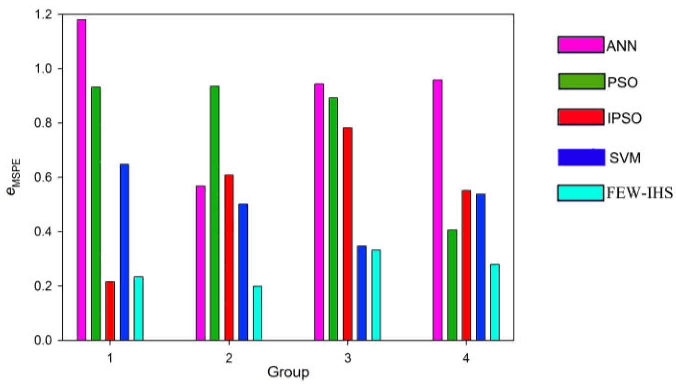
which was 0.64; in the second and fourth group, the FEW-IHS algorithm had the lowest e_{MAPE}, which was 0.39 and 0.63, respectively. The overall e_{MAPE} of the proposed algorithm was close to 0.5, which indicates its good prediction ability.

The mean square error (e_{MSE}) is used to evaluate the degree of data change. The smaller the value is, the more accurate the prediction model is to describe the experimental data. Figure 5(a) shows the e_{MSE} comparison diagram of different algorithms. It can be seen that except the e_{MSE} of the IPSO algorithm in the first group, the e_{MSE} of the FEW-IHS algorithm in other groups was the smallest. This shows that the proposed FEW-IHS algorithm is more accurate than other algorithms in predicting short-term power load.

The mean square percentage error (e_{MSPE}) is also one of the comprehensive indexes for error analysis, which is used to evaluate the degree of fitting between the predicted value and the original value. The closer this value is to 0, the higher the fitting degree of the data prediction model is. Figure 5(b) shows the e_{MSPE} distributions of different algorithms. It can be found that in the first group the e_{MSPE} of the IPSO algorithm was slightly lower than that of the FEW-IHS



(a)



(b)

Fig. 5. Distribution of evaluation criteria among different algorithms: (a) comparison of e_{MSE}; (b) comparison of e_{MSPE}

algorithm, and the e_{MSPE} of the SVM algorithm in the third group was slightly higher than that of the FEW-IHS algorithm. However in the second and fourth group, the e_{MSPE} value of the FEW-IHS algorithm was the lowest, which was significantly different from that of other algorithms. This indicates that the fitting degree of the FEW-IHS model is higher than other algorithms.

Table 2 shows the comparison results of the mean absolute error (e_{MAE}), mean absolute percentage error (e_{MAPE}), mean square error (e_{MSE}) and mean square percentage error (e_{MSPE}) of the five models in power load data prediction in May. This indicates that this analysis method is suitable for analyzing the network load data, and the analysis results in this paper are highly consistent, which once again proves the validity and accuracy of the proposed model.

Table 2. Comparison table of evaluation criteria among different algorithms

Method	**ANN**	**PSO**	**IPSO**	**SVM**	**FEW-IHS**
e_{MAE}	34.1751	25.2028	20.0232	17.7423	8.7468
e_{MAPE} (%)	1.9704	1.4580	1.1445	1.0226	0.5004
e_{MSE}	7.5239	6.3720	4.6612	4.0803	2.0966
e_{MSPE} (%)	0.4359	0.3728	0.2681	0.2386	0.1201

5. Conclusion

Accurate prediction of power load is beneficial for planning the power market, so this paper proposes a new short-term power load data prediction based on the FEW-IHS model. The following conclusions can be drawn:

1. An IHS algorithm can improve the search ability and convergence of an HSA, reduce the probability of falling into local optimal, expand the search scope, and improve the search efficiency.
2. An IHS algorithm can improve the optimization performance of load prediction. The IHS can be used to determine the optimal interval length, improve the accuracy of power load data prediction results, and facilitate the search for the optimal results.
3. The overall and grouping evaluation criteria of the FEW-IHS model are small, and the experimental results are highly consistent, showing high stability of the FEW-IHS model. From the perspective of optimization, the proposed algorithm is more advantageous than ANN, PSO, IPSO and SVM algorithms.

References

[1] Hagan M.T., Behr S.M., *The time series approach to short term load forecasting*, IEEE Transactions on Power Systems, vol. 2, no. 3, pp. 785–791 (1987).

[2] Fallah S.N., Ganjkhani M., Shamshirband S., Chau K.W., *Computational Intelligence on Short-Term Load Forecasting: A Methodological Overview*, Energies, vol. 12, no. 3, 393 (2019).

[3] Fallah S.N., Deo R.C., Shojafar M., Conti M., Shamshirband S., *Computational intelligence approaches for energy load forecasting in smart energy management grids: state of the art, future challenges, and research directions*, Energies, vol. 11, no. 3, 596 (2018).

[4] Kavousi-Fard A., Samet H., Marzbani F., *A new hybrid modified firefly algorithm and support vector regression model for accurate short term load forecasting*, Expert Systems with Applications, vol. 41, no. 13, pp. 6047–6056 (2014).

[5] Torabi M., Hashemi S., Saybani M. R., Shamshirband S., Mosavi A., *A Hybrid clustering and classification technique for forecasting short term energy consumption*, Environmental Progress and Sustainable Energy, vol. 38, no. 1, pp. 66–76 (2019).

[6] Enayatifar R., Yousefi M., Abdullah A.H., Darus A.N., *LAHS: a novel harmony search algorithm based on learning automata*, Communications in Nonlinear Science and Numerical Simulation, vol. 18, no. 12, pp. 3481–3497 (2013).

[7] Salman A., Ahmad I., Hanaa A.L.R., Hamdan S., *Solving the task assignment problem using Harmony Search algorithm*, Evolving Systems, vol. 4, no. 3, pp. 153–169 (2013).

[8] Cheng C.H., Chen T.L., Teoh H.J., Chiang C.H., *Fuzzy time-series based on adaptive expectation model for TAIEX forecasting*, Expert systems with applications, vol. 34, no. 2, pp. 1126–1132 (2008).

[9] Song Q., Chissom B.S., *Fuzzy time series and its models*, Fuzzy Sets and Systems, vol. 54, no. 3, pp. 269–277 (1993).

[10] Yu H.K., *Weighted fuzzy time series models for TAIEX forecasting*, Physica A: Statistical Mechanics and its Applications, vol. 349, no. 3–4, pp. 609–624 (2005).

[11] Cheng C.H., Chen Y.S., Wu Y.L., *Forecasting innovation diffusion of products using trend-weighted fuzzy time-series model*, Expert Systems with Applications, vol. 36, no. 2, pp. 1826–1832 (2009).

[12] Lee M.H., Javedani H., *A weighted fuzzy integrated time series for forecasting tourist arrivals*, International Conference on Informatics Engineering and Information Science, Berlin, Heidelberg, pp. 206–217 (2011).

[13] Huarng K., *Effective lengths of intervals to improve forecasting in fuzzy time series*, Fuzzy sets and systems, vol. 123, no. 3, pp. 387–394 (2001).

[14] Huarng K., Yu T.H.K., *Ratio-based lengths of intervals to improve fuzzy time series forecasting*, IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 36, no. 2, pp. 328–340 (2006).

[15] Chen S.M., *Forecasting enrollments based on high-order fuzzy time series*, Cybernetics and Systems, vol. 33, no. 1, pp. 1–16 (2002).

[16] Aladag C.H., Basaran M.A., Egrioglu E., Yolcu U., Uslu V.R., *Forecasting in high order fuzzy times series by using neural networks to define fuzzy relations*, Expert Systems with Applications, vol. 36, no. 3, pp. 4228–4231 (2009).

[17] Egrioglu E., Aladag C.H., Basaran M.A., Yolcu U., Uslu V.R., *A new approach based on the optimization of the length of intervals in fuzzy time series*, Journal of Intelligent and Fuzzy Systems, vol. 22, no. 1, pp. 15–19 (2011).

[18] Qiu W., Liu X., *Survey on theory and application of fuzzy time series forecasting models*, Fuzzy Systems and Mathematics (in Chinese), vol. 28, no. 3, pp. 173–181 (2011).

[19] Chen G., Qu H., Chu Y., *A method of multi-scale ratios interval partitioning about fuzzy time series model*, Control and Decision (in Chinese), vol. 29, no. 2, pp. 251–256 (2014).

[20] Geem Z.W., *Novel derivative of harmony search algorithm for discrete design variables*, Applied Mathematics and Computation, vol. 199, no. 1, pp. 223–230 (2008).

[21] Su H., *Chaos quantum-behaved particle swarm optimization based neural networks for short-term load forecasting*, Procedia Engineering, vol. 15, pp. 199–203 (2011).

[22] Song C., Liu D., Wu J., Wang H., Zhao Y., Pan X., Dong F., *Multi-objective differential planning based on improved harmony search algorithm for power network*, Electric Power Automation Equipment, vol. 34, no. 11, pp. 142–148 (2008).

[23] Mahdavi M., Fesanghary M., Damangir E., *An improved harmony search algorithm for solving optimization problems*, Applied Mathematics and Computation, vol. 188, no. 2, pp. 1567–1579 (2007).

[24] Cortes C., Vapnik V., *Support-vector networks*, Machine Learning, vol. 20, no. 3, pp. 273–297 (1995).

[25] Niu D., Wang Y., Wu D.D., *Power load forecasting using support vector machine and ant colony optimization*, Expert Systems with Applications, vol. 37, no. 3, pp. 2531–2539 (2010).

[26] Poli R., Kennedy J., Blackwell T., *Particle swarm optimization*, Swarm intelligence, vol. 1, no. 1, pp. 33–57 (2007).

[27] Bahrami S., Hooshmand R.A., Parastegari M., *Short term electric load forecasting by wavelet transform and grey model improved by PSO (particle swarm optimization) algorithm*, Energy, vol. 72, pp. 434–442 (2014).

[28] Park D.C., El-Sharkawi M.A., Marks R.J., Atlas L.E., Damborg M.J., *Electric load forecasting using an artificial neural network*, IEEE Transactions on Power Systems, vol. 6, no. 2, pp. 442–449 (1991).

[29] Li Y.C., Fang T.J., Yu E.K., *Study of support vector machines for short-term load forecasting*, Proceedings of the Chinese Society of Electrical Engineering (in Chinese), vol. 23, no. 6, pp. 55–59 (2003).

[30] Enayatifar R., Sadaei H.J., Abdullah A.H., Gani A., *Imperialist competitive algorithm combined with refined high-order weighted fuzzy time series (RHWFTS–ICA) for short term load forecasting*, Energy Conversion and Management, vol. 76, pp. 1104–1116 (2013).

[31] Nazari-Heris M., Mohammadi-Ivatloo B., Asadi S., Kim J.H., Greem Z.W., *Harmony search algorithm for energy system applications: an updated review and analysis*, Journal of Experimental and Theoretical Artificial Intelligence, pp. 1–27 (2018).

[32] Kassim N., Sulaiman S.I., Othman Z., Musirin I., *Harmony search-based optimization of artificial neural network for predicting AC power from a photovoltaic system*, Proceedings of 2014 IEEE 8th International Power Engineering and Optimization Conference (PEOCO2014), Langkawi, Malaysia, pp. 504–507(2014).