Energy Engineering

Tech Science Press

DOI: 10.32604/EE.2021.015145

ARTICLE

A Weighted Combination Forecasting Model for Power Load Based on Forecasting Model Selection and Fuzzy Scale Joint Evaluation

Bingbing Chen*, **Zhengyi Zhu, Xuyan Wang and Can Zhang**

State Grid Nanjing Power Supply Company, Nanjing, 210019, China

*Corresponding Author: Bingbing Chen. Email: flying387@aliyun.com

Received: 25 November 2020 Accepted: 21 December 2020

ABSTRACT

To solve the medium and long term power load forecasting problem, the combination forecasting method is further expanded and a weighted combination forecasting model for power load is put forward. This model is divided into two stages which are forecasting model selection and weighted combination forecasting. Based on Markov chain conversion and cloud model, the forecasting model selection is implanted and several outstanding models are selected for the combination forecasting. For the weighted combination forecasting, a fuzzy scale joint evaluation method is proposed to determine the weight of selected forecasting model. The percentage error and mean absolute percentage error of weighted combination forecasting result of the power consumption in a certain area of China are 0.7439% and 0.3198%, respectively, while the maximum values of these two indexes of single forecasting models are 5.2278% and 1.9497%. It shows that the forecasting indexes of proposed model are improved significantly compared with the single forecasting models.

KEYWORDS

Power load forecasting; forecasting model selection; fuzzy scale joint evaluation; weighted combination forecasting

1 Introduction

Medium and long term power load forecasting is the main basis of power special planning and distribution network planning. Its forecasting accuracy is directly related to the quality of planning scheme, and is also defined as an important index to evaluate the modernization degree of power enterprise management. In addition, medium and long term load forecasting plays an important role in the security and economy of power grid. Therefore, it has become an urgent problem to study the forecasting method and improve the load forecasting accuracy in the development of power system. However, in recent years, the power market demand has changed greatly, from the initial shortage of supply and demand to the current overall balance of supply and demand. Some new characteristics emerge from the change of power load, which brings a lot of complex factors to the forecasting work.

The single load forecasting model is restricted by the fixed scope of application, so it is difficult to be used in all cases. Selecting multiple models to combine can not only make up the limitation of the information of a single model, but also bring good properties to different models. Compared with single forecasting model, the forecasting results of various models are more effective and comprehensive [1,2]. The research focus of combination forecasting is combination model selection and combination weight determination. The existing model selection [3,4] uses analog error to determine the analog accuracy based on the error of forecasting results, and uses analog error to replace or approximate the forecasting accuracy based on the principle of continuity, so as to carry out model selection. However, there is a lack of recognition of the transfer law from analog accuracy to forecasting accuracy. The combination methods include constant weight and variable weight. Variable weight combination has good adaptability, but it is difficult to reflect the forecasting effectiveness of the model based on error theory.

At present, the commonly used combination methods include minimum variance method, variance covariance method, optimal combination method and analytic hierarchy process [5–7]. Niu et al. [8] and Xiao et al. [9] used Bayesian theory and structural risk minimization principle to establish the least squares support vector machine (LSSVM) combined forecasting model for power load. Ma et al. [10] proposed that the optimal combination forecasting technology can be divided into two parts: Model screening and combination screening. Jiang et al. [11] further analyze the advantages and disadvantages of the screening method, use grey correlation degree method to improve and establish variable weight combination forecasting model. Zhou et al. [12] used entropy method, variance covariance method and grey method to construct hierarchical structure to determine the weight of each model. To eliminate the redundant information in the prediction method, You et al. [13] tested each prediction model for redundancy one by one in combination, regarded the weight as a fuzzy number, and then obtained the optimal weight coefficient through the properties of fuzzy number.

Markov chain is a stochastic process with Markov property in probability theory and mathematical statistics and exists in discrete index set and state space, which is widely used in boundary estimation. Wilinski [14] studied the prediction in a financial time series based on a model in the form of Markov chains. Arruda et al. [15] focused on the computation of the steady state distribution of a Markov chain and made use of an embedding algorithm. Zhu et al. [16] put forward a wind power time series modeling method based on the improved Markov Chain Monte Carlo method. Wan et al. [17] optimized the foundation pit settlement prediction model of logistic curve based on Markov chain. Cloud algorithm is also a powerful tool for load forecasting. Wang et al. [18] proposed a new model with combination of cloud model and support vector machine to select the parameters of the kernel function more accurately and improve the accuracy of short-term load forecasting. Wei et al. [19] introduced a new method and theory of power emergency group decision-making based on cloud model for the power emergency evaluation system established by analytical hierarchy process (AHP). Wang et al. [20] proposed a new model which is combined by the cloud model, particle swarm optimization (PSO) and LSSVM to improve the accuracy of selecting the parameters of the kernel function, to deal with uncertainty factors and improve the accuracy of short-term load forecasting. Liu et al. [21] proposed a method based on cloud model and fuzzy Petri net to solve the problem that it is difficult to identify and control the potential hazardous trading behavior in the power market.

In order to further expand the combination forecasting method, based on previous studies, this paper applies the idea of Markov chain conversion and cloud algorithm to forecasting model selection and proposed a weighted combination forecasting model for medium and long-term power load forecasting. By forecasting the power consumption of the whole society in a certain area, the forecasting results show the effectiveness of the proposed model, and its practical value is well verified.

2 Research Structure

This paper expands the combination forecasting method for the medium and long term power load forecasting problem and proposes a weighted combination forecasting model. The flowchart of the proposed model is shown in Fig. 1.

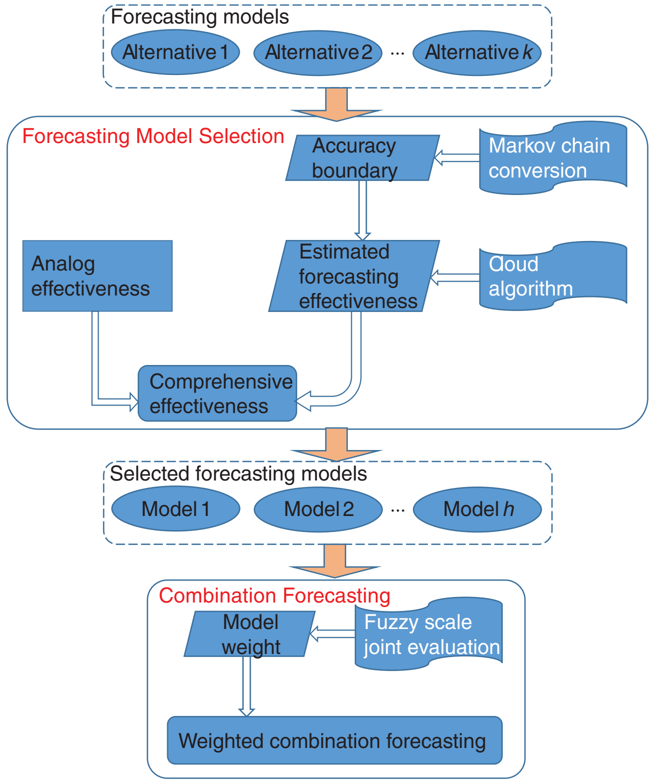

Figure 1: The flowchart of the proposed weighted combination forecasting model

As shown in Fig. 1, the proposed model is divided into two stages: Forecasting model selection and combination forecasting. In the first stage, the forecasting accuracy boundary of forecasting model is determined by Markov chain conversion, and then the estimated forecasting effectiveness is determined by cloud algorithm. After this, the comprehensive effectiveness of forecasting model can be obtained by integrating the analog effectiveness and estimated forecasting effectiveness. According to the comprehensive effectiveness of forecasting model several outstanding models are selected for the combination forecasting. In the second stage, fuzzy scale joint evaluation is carried out to determine the weight of selected forecasting model. Based on the model weight and single model forecasting value, weighted combination forecasting is implemented.

3 Forecasting Model Selection

It is assumed that there are m history years and the power load of history year i is ρ_i, here $i = 1, 2, \ldots, m$. There are n forecasting years and the forecasting power load of forecasting year j is ϑ_j, here $j = 1, 2\ldots n$. There are k alternatives of forecasting models. Through forecasting model l, the analog value of power load of history year i is $\rho'_{l,i}$ while the forecasting value of power load of forecasting year j is $\vartheta'_{l,j}$, here $l = 1, 2\ldots k$.

For forecasting model l, the analog value relative error $\eta_{l,i}$ of history year i and forecasting value relative error $\eta_{l,j}$ of forecasting year j are:

$$\eta_{l,i} = (\rho_i - \rho'_{l,i})/\rho_i \tag{1}$$

$$\eta_{l,j} = (\vartheta_j - \vartheta'_{l,j})/\vartheta_j \tag{2}$$

If $0 \le |\eta_{l,i}| \le 1$, the analog accuracy of history year i is $\theta_{l,i} = 1 - |\eta_{l,i}|$; If $|\eta_{l,i}| > 1$, the analog accuracy of history year i is $\theta_{l,i} = 0$. Similarly, if $0 \le |\eta_{l,j}| \le 1$, the forecasting accuracy of forecasting year j is $\theta_{l,j} = 1 - |\eta_{l,j}|$; If $|\eta_{l,j}| > 1$, the forecasting accuracy of history year i is $\theta_{l,j} = 0$.

Next the analog effectiveness μ_l and forecasting effectiveness v_l of forecasting model l are:

$$\mu_l = \frac{\sum_{i=1}^{m} \theta_{l,i}}{m} (1 - \sigma(\theta_{l,i})) \tag{3}$$

$$v_l = \frac{\sum_{j=1}^{n} \theta_{l,i}}{n} (1 - \sigma(\theta_{l,j})) \tag{4}$$

Here $\sigma(\theta_{l,i})$ is the standard deviation of $\theta_{l,i}$ and $\sigma(\theta_{l,j})$ is the standard deviation of $\theta_{l,j}$.

The comprehensive effectiveness of forecasting model l is defined by Eqs. (3) and (4), which characterizes the credibility of the forecasting model. In the future forecasting interval, the practical power load value has not yet appeared and the forecasting value relative error (Eq. (2)) cannot be obtained. The accuracy and the effectiveness of a forecasting model can only be estimated based on its inherent property. After that, the forecasting models are screened and the better forecasting models are selected to implement combination forecasting.

The accuracy of forecasting model is an inherent property. The analog accuracy, which can be obtained by the virtual forecasting for the multi-time power load, is a performance of the accuracy of forecasting model. Through forecasting model l the power load of history year i is forecasted. Then the analog accuracy sequence is obtained as $\{\theta_{l,1}, \theta_{l,2}, \ldots, \theta_{l,m}\}$. In $\{\theta_{l,1}, \theta_{l,2}, \ldots, \theta_{l,m}\}$, the expectation and the standard deviation of the analog accuracy of each history year show the property of the forecasting model. As everyone knows, randomness and discreteness always appear in analog accuracy sequence $\{\theta_{l,1}, \theta_{l,2}, \ldots, \theta_{l,m}\}$. In view of this, Markov chain conversion matrix [22] can be adopted to describe the conversion principle. The accuracy boundary is estimated based on Markov chain by the following steps.

Step 1: For forecasting model l the distribution interval of analog accuracy of history year i can be equally divided into m_l subintervals which are $\varpi_{l,1}, \varpi_{l,2}, \ldots, \varpi_{l,m_l}$, here $m_l \le m$. Each subinterval can be treated as an analog accuracy status. All analog accuracy statuses form a status sequence $\{\varpi_{l,1}, \varpi_{l,2}, \ldots, \varpi_{l,m_l}\}$.

Step 2: Based on the analog accuracy of forecasting model l of history year i, it is assumed that the appearance number of analog accuracy status $\varpi_{l,s}$ is $AN_{l,s}$, here $s = 1, 2, \ldots, m_l$ and $AN_{l,s} < m$. It means that there are $AN_{l,s}$ times belonging to analog accuracy status $\varpi_{l,s}$. Assuming that the conversion times from status $\varpi_{l,g}$ to status $\varpi_{l,s}$ is $C_l(g, s)$. Therefore, the conversion probability of forecasting model l from status $\varpi_{l,g}$ to status $\varpi_{l,s}$ is obtained as:

$$P_l(g, s) = \frac{C_l(g, s)}{AN_{l,s}} \tag{5}$$

Step 3: The 1st order status conversion matrix of forecasting model l is constructed as:

$$P_l^{(1)} = [P_l(g, s)]_{m' \times m'} \tag{6}$$

Then the q^{th} order status conversion matrix of forecasting model l is obtained as:

$$P_l^{(q)} = (P_l^{(1)})^q \tag{7}$$

Step 4: For forecasting model l, the appearance numbers of every analog accuracy status form an initial vector AN_l as:

$$AN_l = \left[AN_{l,1}, AN_{l,2}, \ldots, AN_{l,m_l}\right] \tag{8}$$

A new status matrix of forecasting model l can be obtained by multiplying initial vector AN_l and q^{th} order status conversion matrix $P_l^{(q)}$:

$$P_l = AN_l \cdot P_l^{(q)} \tag{9}$$

Step 5: The sum of every column of P_l is calculated one by one. Assuming that the column with the maximum sum is column s, the forecasting accuracy belongs to accuracy status $\varpi_{l,s}$. Therefore, the accuracy boundary estimation of forecasting model l is obtained.

In m history years, affected by various kinds of factors, analog accuracy sequence $\{\theta_{l,1}, \theta_{l,2}, \ldots, \theta_{l,m}\}$ of forecasting model l obviously has the features of random variables. Therefore, the forecasting accuracy is equivocal in the accuracy boundary (obtained in Step 5). The non-determinacy in the accuracy boundary can be described by the concepts of expectation, entropy and hyper-entropy in cloud model theory [23]. Then the accuracy can be estimated quantitatively by the following steps.

Step 6: Based on reverse cloud algorithm, analog accuracy sequence $\{\theta_{l,1}, \theta_{l,2}, \ldots, \theta_{l,m}\}$ is treated as the input of cloud model while expectation α_l, entropy β_l and hyper-entropy γ_l is the output of cloud model. The reverse cloud algorithm is:

(1) Expectation:

$$\alpha_l = \frac{\sum_{i=1}^{m} \theta_{l,i}}{m} \tag{10}$$

(2) 1^{st} order absolute center distance:

$$\Omega_1 = \frac{1}{m}\sum_{i=1}^{m} \left|\theta_{l,i} - \alpha_l\right| \tag{11}$$

(3) 2^{nd} order absolute center distance:

$$\Omega_2 = \frac{1}{m-1}\sum_{i=1}^{m} (\theta_{l,i} - \alpha_l)^2 \tag{12}$$

(4) Entropy:

$$\beta_l = \sqrt{\frac{\pi}{2}}\Omega_1 \tag{13}$$

(5) Hyper-entropy:

$$\gamma_l = \sqrt{\Omega_2 - (\beta_l)^2} \tag{14}$$

Step 7: By treating expectation α_l, entropy β_l and hyper-entropy γ_l as the input of forward cloud generator and treating accuracy boundary $\varpi_{l,s}$ as constraint, the accuracy is estimated quantitatively by forward cloud model. The forward cloud algorithm is:

(1) A normal random number is generated with an expectation of β_l and a variance of γ_l:

$$\Psi = NORM(\beta_l, \gamma_l) \tag{15}$$

(2) In accuracy boundary $\varpi_{l,s}$, normal random forecasting accuracy is generated with an expectation of α_l and a variance of Ψ:

$$\theta_{l,j} = NORM(\alpha_l, \Psi) \tag{16}$$

Next forecasting model selection is carried out based on the comprehensive effectiveness of each forecasting model. In m history years and n forecasting years, the comprehensive effectiveness of forecasting model l is defined based on Eqs. (3) and (4) as:

$$\varepsilon_l = \lambda \cdot \mu_l + \eta \cdot v_l \tag{17}$$

Here, μ_l is the analog effectiveness of forecasting model l in m history years and v_l is the estimated forecasting effectiveness of forecasting model l in n forecasting years. λ and η are the importance factors of analog effectiveness and forecasting effectiveness respectively which satisfy:

$$\begin{cases} \lambda + \eta = 1 \\ \lambda > 0, \eta > 0 \end{cases} \tag{18}$$

The threshold value of forecasting model selection is determined by:

$$\bar{\varepsilon} = \frac{\sum_{l=1}^{k} \varepsilon_l}{k} \tag{19}$$

When ε_l is greater than or equal to threshold value $\bar{\varepsilon}$, forecasting model l is selected for the subsequent combination forecasting.

4 Combination Forecasting

It is assumed that there are h selected forecasting models, here $h < k$. How to determine the weight of each model is a key problem in combination forecasting. In the weight determination problem, expert evaluation method can make full use of the experience and wisdom of experts. In this paper, the expert evaluation method is introduced into the determination of forecasting model weight. Experts make a judgment on the principle of the model, the degree of agreement with the actual situation and the forecasting effect, so as to determine the weight of the model. In the traditional expert evaluation method, the accurate scale value is used to express the experts' evaluation on the importance of different objects [24]. Compared with the accurate scale value, the fuzzy scale value can better reflect the uncertainty of expert evaluation. In addition, compared with the single expert evaluation, the joint evaluation of multiple experts is more reasonable. Therefore, this paper proposes a fuzzy scale joint evaluation method to determine the weight of forecasting model. In proposed method, trapezoid fuzzy number is used to express the experts' evaluation on the importance of different forecasting models and rough boundary interval is used to integrate the judgments of different experts. Its detailed process is as following.

Step 1: It is assumed that there are t experts. For h selected forecasting models, each expert evaluates the relative importance of any two selected forecasting models. The fuzzy reciprocal evaluation matrix given by expert r (r = 1, 2, …, t) is as:

$$\tilde{E}^r = \left[\tilde{e}^r_{x,y}\right]_{h\times h} \tag{20}$$

Here $\tilde{e}^r_{x,y}$ is the evaluation score of model x relative to model y given by expert r. $\tilde{e}^r_{x,y}$ is a trapezoid fuzzy number scale and $\tilde{e}^r_{x,y} = (\xi^r_{x,y}, \varsigma^r_{x,y}, \psi^r_{x,y}, \zeta^r_{x,y})$. If $x = y$, $\tilde{e}^r_{x,y} = (1,1,1,1)$. The next step can only be carried out after t evaluation matrices are qualified in consistency inspection. Otherwise, the corresponding expert will adjust his evaluation matrix.

Step 2: The joint evaluation matrix is constructed as:

$$\tilde{E} = \left[\tilde{e}_{x,y}\right]_{h\times h} \tag{21}$$

Here $\tilde{e}_{x,y} = \{\tilde{e}^1_{x,y}, \tilde{e}^2_{x,y}, \ldots, \tilde{e}^t_{x,y}\}$.

Step 3: According to rough sets theory [25,26], the rough boundary interval of $\tilde{e}^r_{x,y}$ in $\tilde{e}_{x,y}$ can be expressed as:

$$RBI(\tilde{e}^r_{x,y}) = \left[\underline{L}(\tilde{e}^r_{x,y}), \overline{L}(\tilde{e}^r_{x,y})\right] \tag{22}$$

$$\underline{L}(\tilde{e}^r_{x,y}) = \frac{\sum\limits_{\tilde{e}^{\underline{r}}_{x,y} \le \tilde{e}^r_{x,y}} \tilde{e}^{\underline{r}}_{x,y}}{\underline{N}(\tilde{e}^r_{x,y})} \tag{23}$$

$$\overline{L}(\tilde{e}^r_{x,y}) = \frac{\sum\limits_{\tilde{e}^{\overline{r}}_{x,y} \ge \tilde{e}^r_{x,y}} \tilde{e}^{\overline{r}}_{x,y}}{\overline{N}(\tilde{e}^r_{x,y})} \tag{24}$$

Here $\underline{L}(\tilde{e}^r_{x,y})$ is the rough lower boundary of $RBI(\tilde{e}^r_{x,y})$ and $\overline{L}(\tilde{e}^r_{x,y})$ is the rough upper boundary of $RBI(\tilde{e}^r_{x,y})$, $\underline{N}(\tilde{e}^r_{x,y})$ means the number of the elements which smaller than or equal to $\tilde{e}^r_{x,y}$ and $\overline{N}(\tilde{e}^r_{x,y})$ means the number of the elements which bigger than or equal to $\tilde{e}^r_{x,y}$.

Then the rough boundary interval of $\tilde{e}_{x,y}$ can be expressed as:

$$RBI(\tilde{e}_{x,y}) = \left[\underline{L}(\tilde{e}_{x,y}), \overline{L}(\tilde{e}_{x,y})\right] \tag{25}$$

$$\underline{L}(\tilde{e}_{x,y}) = \frac{\sum\limits_{r=1}^{t} \underline{L}(\tilde{e}^r_{x,y})}{t} \tag{26}$$

$$\overline{L}(\tilde{e}_{x,y}) = \frac{\sum\limits_{r=1}^{t} \overline{L}(\tilde{e}^r_{x,y})}{t} \tag{27}$$

Step 4: The rough boundary interval evaluation matrix is constructed as:

$$\tilde{\Theta} = \left[RBI(\tilde{e}_{x,y})\right]_{h\times h} \tag{28}$$

Then $\tilde{\Theta}$ is divided into two matrices as:

$$\underline{\tilde{\Theta}} = \left[\underline{L}(\tilde{e}_{x,y})\right]_{h\times h} \tag{29}$$

$$\overline{\tilde{\Theta}} = \left[\overline{L}(\tilde{e}_{x,y})\right]_{h\times h} \tag{30}$$

Here $\underline{\tilde{\Theta}}$ is the rough lower boundary matrix and $\overline{\tilde{\Theta}}$ is the rough upper boundary matrix.

Based on the gravity formula of trapezoid fuzzy number [24,27], $\underset{\sim}{\tilde{\Theta}}$ and $\bar{\tilde{\Theta}}$ are converted to $\underline{\Theta}$ and $\bar{\Theta}$ which are real number form respectively. Then the eigenvectors of $\underline{\Theta}$ and $\bar{\Theta}$ corresponding to maximum eigenvalues are obtained respectively as:

$$\underline{\omega} = [\underline{\omega}_1, \underline{\omega}_2, \ldots, \underline{\omega}_h]^{\mathrm{T}} \tag{31}$$

$$\bar{\omega} = [\bar{\omega}_1, \bar{\omega}_2, \ldots, \bar{\omega}_h]^{\mathrm{T}} \tag{32}$$

Step 5: After averaging the two eigenvectors obtained by Eqs. (31) and (32), the weight vector of h selected forecasting models is obtained as:

$$\omega = [\omega_1, \omega_2, \ldots, \omega_h]^{\mathrm{T}} \tag{33}$$

Here $\omega_x = \dfrac{\underline{\omega}_x + \bar{\omega}_x}{2}$ is the initial weight value of selected forecasting model x. Then the initial weight value is normalized as:

$$\omega_x = \frac{\omega_x}{\sum\limits_{x=1}^{h} \omega_x} \tag{34}$$

Assuming that the forecasting value of power load of selected forecasting model x for forecasting year j is $\vartheta'_{x,j}$, the weighted combination forecasting is implemented as:

$$\vartheta'_j = \sum_{x=1}^{h} \omega_x \vartheta'_{x,j} \tag{35}$$

5 Case Study

The rural electricity consumption data of Jiangsu Province from 2005 to 2016 are selected to establish the model. The data are from China Statistical Yearbook 2017 (Tab. 1), and the unit of electricity consumption in this paper is 100 million kWh. The years 2005 to 2015 are history years and 2016 is forecasting year. The electricity consumption of 2016 shown in Tab. 1 is taken as validation data.

Table 1: The rural electricity consumption data of Jiangsu Province from 2005 to 2016

Year	Electricity consumption
2005	825.10
2006	1011.79
2007	1159.03
2008	1234.14
2009	1316.62
2010	1472.89
2011	1606.83
2012	1696.41
2013	1801.86
2014	1834.93
2015	1836.19
2016	1869.27

The initial forecasting models are (1) Hyperbola model; (2) COMPERTZ model; (3) Exponential model; (4) Power function model; (5) Cubic curve model; (6) S-curve model; (7) Logarithm model; (8) Para-curve model. Their analog and forecasting electricity consumption values are shown in Tabs. 2 and 3.

Table 2: The analog and forecasting electricity consumption values of eight forecasting models

Year	Forecasting models							
	1	2	3	4	5	6	7	8
2005	822.69	821.12	826.76	863.99	840.05	858.84	882.74	855.96
2006	1023.33	1027.77	984.71	1026.35	995.07	1101.78	1016.71	982.05
2007	1141.42	1139.85	1130.48	1170.52	1127.80	1197.08	1186.62	1111.89
2008	1235.14	1231.57	1264.06	1249.16	1248.20	1240.51	1260.73	1241.93
2009	1338.19	1337.61	1385.45	1352.43	1362.85	1403.59	1409.02	1368.64
2010	1460.63	1463.49	1494.66	1481.78	1474.96	1504.77	1470.10	1488.47
2011	1592.99	1595.85	1591.68	1637.57	1584.38	1628.02	1605.40	1597.88
2012	1713.36	1712.79	1676.52	1722.48	1687.55	1749.42	1724.59	1693.34
2013	1797.78	1794.20	1749.17	1886.36	1777.57	1798.29	1857.53	1771.30
2014	1833.67	1832.09	1809.64	1923.04	1844.14	1855.39	1912.38	1828.22
2015	1836.47	1840.90	1857.92	1890.13	1873.59	1928.27	1899.87	1860.57

Table 3: The forecasting electricity consumption values of eight forecasting models

Year	Forecasting models							
	1	2	3	4	5	6	7	8
2016	1869.39	1869.54	1894.01	1867.82	1848.89	1868.28	1864.81	1869.39

Assuming that analog effectiveness and forecasting effectiveness are equally important, $\lambda = \eta = 0.5$. According to the comprehensive effectiveness calculation approach based on Markov chain conversion and cloud model (Eq. (17)), the comprehensive effectiveness of each forecasting model is obtained as shown in Tab. 4.

Table 4: The comprehensive effectiveness of each forecasting model

Forecasting model	Comprehensive effectiveness
1	87.45%
2	88.25%
3	87.84%
4	74.54%
5	86.61%
6	75.33%
7	77.09%
8	84.66%

According to Eq. (19), the threshold value of forecasting model selection is 82.72%. Therefore forecasting models 1, 2, 3, 5 and 8 are selected for the subsequent combination forecasting. These models correspond to selected forecasting models (SFM) 1, 2, 3, 4 and 5 in turn. Based on the fuzzy scale joint evaluation method proposed in Section 4, it is assumed that there are four experts to determine the weight of selected forecasting model. This paper uses the trapezoid fuzzy number scale proposed in Reference [24,27] to express the evaluation score of expert. The traditional nine-level comparison scale is improved as shown in Tab. 5. For example, the scale value of comparison scale "level 2: Strong inferior" is 2, which can firstly be converted to 2/8. Then real number "2" corresponds to trapezoid fuzzy number (1, 1.5, 2.5, 3) while real number "8" corresponds to trapezoid fuzzy number (7, 7.5, 8.5, 9). Lastly based on the operation rules of trapezoid fuzzy number the trapezoid fuzzy number scale of "level 2: strong inferior" can be obtained as (0.11, 0.18, 0.33, 0.43), which is shown in Tab. 5.

Table 5: The relationship between traditional nine-level comparison scale and improved fuzzy scale

Nine-level comparison scale	Traditional scale	Trapezoid fuzzy number scale
level 1: extremely inferior	1	(0.11,0.11,0.18,0.25)
level 2: strong inferior	2	(0.11,0.18,0.33,0.43)
level 3: obviously inferior	3	(0.25,0.33,0.54,0.67)
level 4: slightly inferior	4	(0.43,0.54,0.82,1.00)
level 5: identical	5	(1.00,1.00,1.00,1.00)
level 6: slightly superior	6	(1.00,1.22,1.86,2.33)
level 7: obviously superior	7	(1.50,1.86,3.00,4.00)
level 8: strongly superior	8	(2.33,3.00,5.67,9.00)
level 9: extremely superior	9	(4.00,5.67,9.00,9.00)

The fuzzy reciprocal evaluation matrices given by the four experts are shown in Tabs. 6, 7, 8 and 9.

Table 6: The fuzzy reciprocal evaluation matrix given by expert 1 ($\tilde{E}^1 = \left[\tilde{e}^1_{x,y}\right]_{5\times 5}$)

	SFM 1	SFM 2	SFM 3	SFM 4	SFM 5
SFM 1	/	level 1	level 3	level 4	level 5
SFM 2	/	/	level 7	level 8	level 8
SFM 3	/	/	/	level 6	level 7
SFM 4	/	/	/	/	level 6
SFM 5	/	/	/	/	/

Table 7: The fuzzy reciprocal evaluation matrix given by expert 2 ($\tilde{E}^2 = \left[\tilde{e}^2_{x,y}\right]_{5\times 5}$)

	SFM 1	SFM 2	SFM 3	SFM 4	SFM 5
SFM 1	/	level 5	level 6	level 7	level 7
SFM 2	/	/	level 7	level 7	level 6
SFM 3	/	/	/	level 6	level 6
SFM 4	/	/	/	/	level 8
SFM 5	/	/	/	/	/

Table 8: The fuzzy reciprocal evaluation matrix given by expert 3 ($\tilde{E}^3 = \left[\tilde{e}^3_{x,y}\right]_{5\times5}$)

	SFM 1	SFM 2	SFM 3	SFM 4	SFM 5
SFM 1	/	level 9	level 6	level 7	level 7
SFM 2	/	/	level 1	level 3	level 3
SFM 3	/	/	/	level 5	level 7
SFM 4	/	/	/	/	level 1
SFM 5	/	/	/	/	/

Table 9: The fuzzy reciprocal evaluation matrix given by expert 4 ($\tilde{E}^4 = \left[\tilde{e}^4_{x,y}\right]_{5\times5}$)

	SFM 1	SFM 2	SFM 3	SFM 4	SFM 5
SFM 1	/	level 5	level 8	level 8	level 5
SFM 2	/	/	level 6	level 7	level 6
SFM 3	/	/	/	level 7	level 5
SFM 4	/	/	/	/	level 3
SFM 5	/	/	/	/	/

Evaluation matrices $\tilde{E}^1$, $\tilde{E}^2$, $\tilde{E}^4$ and $\tilde{E}^4$ are qualified in consistency inspection according to the consistency inspection approach [24,27]. Then the joint evaluation matrix is constructed as $\tilde{E} = \left[\tilde{e}_{x,y}\right]_{5\times5}$ and $\tilde{e}_{x,y} = \{\tilde{e}^1_{x,y}, \tilde{e}^2_{x,y}, \tilde{e}^3_{x,y}, \tilde{e}^4_{x,y}\}$. For example, $\tilde{e}^1_{1,2}$ = (0.11,0.11,0.18,0.25), $\tilde{e}^2_{1,2}$ = (1.00,1.00,1.00,1.00), $\tilde{e}^3_{1,2}$ = (4.00,5.67,9.00,9.00), $\tilde{e}^4_{1,2}$ = (1.00,1.00,1.00,1.00), so $\tilde{e}_{1,2}$ = {(0.11,0.11,0.18,0.25), (1.00,1.00,1.00,1.00), (4.00,5.67,9.00,9.00), (1.00,1.00,1.00,1.00)}.

Next the rough boundary interval of $\tilde{e}_{1,2}$ is calculated according to Eqs. (22)–(27).

For $\tilde{e}^1_{1,2}$ in $\tilde{e}_{1,2}$, the rough upper boundary is:

$\underline{L}(\tilde{e}^1_{1,2})$ = ((0.11,0.11,0.18,0.25) + (1.00,1.00,1.00,1.00) + (4.00,5.67,9.00,9.00) + (1.00,1.00,1.00,1.00))/4 = (1.53,1.94,2.79,2.81).

And the rough lower boundary is $\overline{L}(\tilde{e}_{1,2})$ = (0.11,0.11,0.18,0.25)/1 = (0.11,0.11,0.18,0.25).

Then the rough boundary interval of $\tilde{e}^1_{1,2}$ in $\tilde{e}_{1,2}$ is obtained as $RBI(\tilde{e}^1_{1,2})$ = [(0.11,0.11,0.18,0.25), (1.53,1.94,2.79,2.81)]. By same way:

$RBI(\tilde{e}^2_{1,2})$ = [(0.70,0.70,0.73,0.75),(2.00,2.56,3.67,3.67)],

$RBI(\tilde{e}^3_{1,2})$ = [(1.53,1.94,2.79,2.81),(4.00,5.67,9.00,9.00)],

$RBI(\tilde{e}^4_{1,2})$ = [(0.70,0.70,0.73,0.75),(2.00,2.56,3.67,3.67)].

According to Eqs. (25)–(27), The rough boundary interval of $\tilde{e}_{1,2}$ is obtained as $RBI(\tilde{e}_{1,2})$ = [(0.76,0.86,1.11,1.14),(2.38,3.18,4.78,4.79)]. After calculating the rough boundary interval of all elements in $\tilde{E} = \left[\tilde{e}_{x,y}\right]_{5\times5}$, rough boundary interval evaluation matrix is constructed as shown in Tab. 10.

Table 10: The rough boundary interval evaluation matrix $\left(\tilde{\Theta}=\left[RBI(\tilde{e}_{x,y})\right]_{5\times 5}\right)$

	SFM 1	SFM 2	SFM 3	SFM 4	SFM 5
SFM 1	[(1,1,1,1),(1,1,1,1)]	[(0.76,0.86,1.11,1.14), (2.38,3.18,4.78,4.79)]	[(0.72,0.91,1.46,1.95), (1.59,2.02,3.60,5.42)]	[(1.04,1.30,2.12,2.88), (1.83,2.32,4.14,6.21)]	[(1.13,1.21,1.50,1.75), (1.38,1.64,2.50,3.25)]
SFM 2	[(0.76,0.87,1.11,1.14), (2.38,3.18,4.78,4.79)]	[(1,1,1,1),(1,1,1,1)]	[(0.68,0.83,1.30,1.71), (1.34,1.66,2.66,3.52)]	[(0.95,1.20,1.99,2.72), (1.82,2.31,4.12,6.19)]	[(0.72,0.91,1.46,1.95), (1.59,2.02,3.60,5.42)]
SFM 3	[(0.34,0.45,0.72,0.91), (0.92,1.15,1.83,2.40)]	[(0.51,0.70,1.11,1.24), (2.17,3.05,4.84,4.92)]	[(1,1,1,1),(1,1,1,1)]	[(1.03,1.16,1.52,1.80), (1.24,1.51,2.35,3.05)]	[(1.13,1.27,1.71,2.08), (1.40,1.71,2.71,3.57)]
SFM 4	[(0.23,0.31,0.52,0.66), (0.60,0.75,1.16,1.45)]	[(0.26,0.35,0.59,0.76), (0.84,1.05,1.71,2.25)]	[(0.38,0.47,0.70,0.84), (0.69,0.75,0.89,0.98)]	[(1,1,1,1),(1,1,1,1)]	[(0.42,0.51,0.86,1.21), (1.53,1.95,3.54,5.43)]
SFM 5	[(0.44,0.50,0.65,0.75), (0.81,0.83,0.88,0.92)]	[(0.34,0.45,0.72,0.91), (0.92,1.15,1.83,2.40)]	[(0.32,0.40,0.61,0.74), (0.67,0.72,0.84,0.92)]	[(0.64,0.86,1.40,1.64), (2.56,3.54,5.64,5.94)]	[(1,1,1,1),(1,1,1,1)]

According to Eqs. (29) and (30), $\tilde{\Theta}=\left[RBI(\tilde{e}_{x,y})\right]_{5\times 5}$ is divided into two matrices as: $\underline{\tilde{\Theta}}=\left[\underline{L}(\tilde{e}_{x,y})\right]_{5\times 5}$ and $\bar{\tilde{\Theta}}=\left[\overline{L}(\tilde{e}_{x,y})\right]_{5\times 5}$. As shown in Tabs. 11 and 12, $\underline{\tilde{\Theta}}$ and $\bar{\tilde{\Theta}}$ are converted to $\underline{\Theta}$ and $\bar{\Theta}$ which are real number form, respectively.

Table 11: The rough lower boundary matrix with real number form ($\underline{\Theta}$)

	SFM 1	SFM 2	SFM 3	SFM 4	SFM 5
SFM 1	1	0.97	1.27	1.85	1.40
SFM 2	0.97	1	1.14	1.73	1.27
SFM 3	0.61	0.89	1	1.38	1.55
SFM 4	0.43	0.49	0.60	1	0.76
SFM 5	0.59	0.61	0.52	1.14	1

Table 12: The rough upper boundary matrix with real number form ($\bar{\Theta}$)

	SFM 1	SFM 2	SFM 3	SFM 4	SFM 5
SFM 1	1	3.77	3.21	3.68	2.21
SFM 2	3.77	1	2.31	3.66	3.21
SFM 3	1.59	3.73	1	2.05	2.36
SFM 4	0.99	1.47	0.83	1	3.17
SFM 5	0.86	1.59	0.79	4.41	1

The eigenvectors of $\underline{\Theta}$ and $\bar{\Theta}$ corresponding to maximum eigenvalues are obtained respectively as $\underline{\omega}$ = [0.56,0.53,0.46,0.28,0.33]^{T} and $\bar{\omega}$ = [0.56,0.55,0.45,0.29,0.31]^{T}. According to Eqs. (33) and (34), the weight vector of five selected forecasting models is obtained as ω = [0.26,0.25,0.21,0.13,0.15]^{T}. Then according to Eq. (35), the weighted combination forecasting result is obtained as shown in Tab. 13.

Table 13: The weighted combination forecasting result

Year	Electricity consumption
2005	825.07
2006	1011.77
2007	1160.16
2008	1229.39
2009	1325.56
2010	1465.80
2011	1603.85
2012	1709.03
2013	1788.64
2014	1842.19
2015	1834.06
2016	1869.54

The percentage error (PE) of single forecasting models (SFM 1, 2, 3, 4 and 5) and weighted combination forecasting model are compared as shown in Tab. 14, while the mean absolute percentage error (MAPE) of single forecasting models (SFM 1, 2, 3, 4 and 5) and weighted combination forecasting model are compared as shown in Tab. 15.

Table 14: The comparison of PE of single forecasting models (SFM 1, 2, 3, 4 and 5) and weighted combination forecasting model

Year	SFM 1	SFM 2	SFM 3	SFM 4	SFM 5	Weighted combination forecasting model
2005	1.8119%	3.7402%	0.2012%	–0.2921%	–0.4824%	–0.0036%
2006	–1.6525%	–2.9393%	–2.6764%	1.1406%	1.5794%	–0.0020%
2007	–2.6945%	–4.0672%	–2.4633%	–1.5194%	–1.6548%	0.0975%
2008	1.1393%	0.6312%	2.4244%	0.0810%	–0.2082%	–0.3849%
2009	3.5113%	3.9510%	5.2278%	1.6383%	1.5942%	0.6790%
2010	0.1405%	1.0578%	1.4780%	–0.8324%	–0.6382%	–0.4814%
2011	–1.3972%	–0.557%	–0.9429%	–0.8613%	–0.6833%	–0.1855%
2012	–0.5223%	–0.181%	–1.1725%	0.9992%	0.9656%	0.7439%
2013	–1.3481%	–1.696%	–2.9242%	–0.2264%	–0.4251%	–0.7337%
2014	0.5019%	–0.3657%	–1.3783%	–0.0687%	–0.1548%	0.3957%
2015	2.0368%	1.3277%	1.1834%	0.0152%	0.2565%	–0.1160%
2016	–1.0903%	–0.2386%	1.3235%	0.0064%	–0.0776%	0.0144%

Table 15: The comparison of MAPE of single forecasting models (SFM 1, 2, 3, 4 and 5) and weighted combination forecasting model

SFM 1	SFM 2	SFM 3	SFM 4	SFM 5	Weighted combination forecasting model
1.4872%	1.7294%	1.9497%	0.6401%	0.7267%	0.3198%

Figs. 2 and 3 are error analysis of single forecasting models (SFM 1, 2, 3, 4 and 5) and weighted combination forecasting model. As can be seen in Fig. 2, the maximum PEs of SFM 1, 2, 3, 4 and 5 are 3.5113%, –4.0672%, 5.2278%, 1.6383% and –1.6548%, respectively, while the maximum PE of weighted combination forecasting model is only 0.7439%, which is greatly reduced. In Fig. 3, the MAPE of weighted combination forecasting model is only 0.3198%. Compared with single forecasting models, the MAPE of weighted combination forecasting model is also greatly reduced.

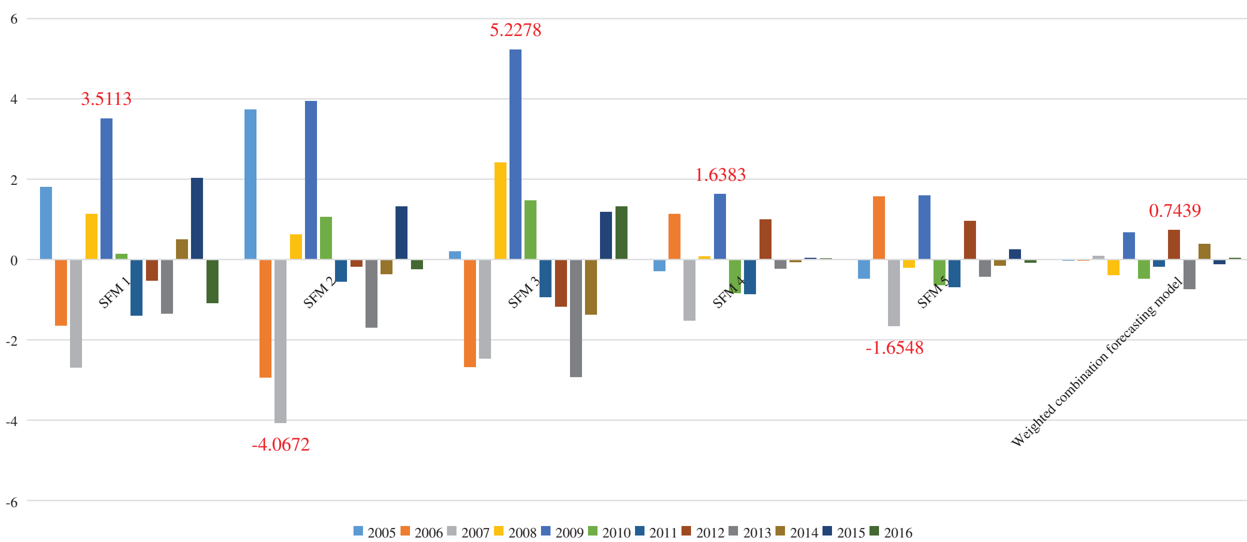

Figure 2: PE analysis

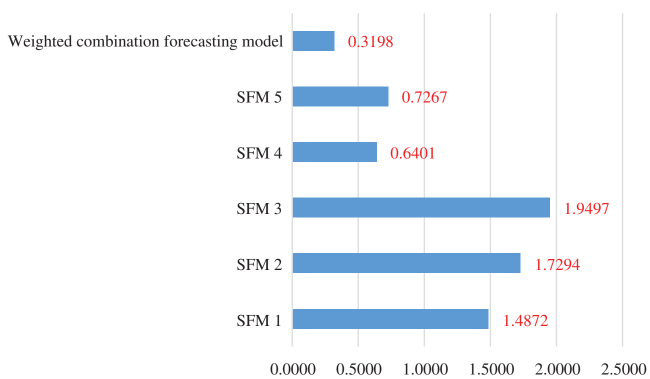

Figure 3: MAPE analysis

It can be seen that this weighted combination forecasting model can effectively improve the forecasting accuracy and increase the credibility of the model. In load forecasting with abundant historical information, combination forecasting model can synthesize information from various single models, which not only improves the accuracy of load forecasting, but also has strong operability.

6 Conclusions

The forecasting model selection is applied to weighted combination forecasting, and the rural electricity consumption data of a province in China is forecasted. The weighted combination forecasting results show that the forecasting indexes PE and MAPE are improved significantly compared with the single forecasting models. The forecasting accuracy has been greatly improved, proving that the forecasting method is effective and feasible, and has played a role in expanding the combined forecasting method. Compared with other combination methods, the weighted combination forecasting based on fuzzy scale joint evaluation method, which is based on the experience and wisdom of experts, is more practical and easier to implement. However, all experts are treated equal in the fuzzy scale joint evaluation. Obviously, the ability of experts is different and this problem will be considered in future work.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Chen, Q., Jin, X., Yao, J., Yang, S., Gong, L. et al. (2014). Improved artificial bee colony algorithm applied to medium and long-term load combination forecasting. *Power System Protection and Control, 23,* 113–117.
2. Zhou, S., Zhou, R., Li, H., Li, S. (2015). Method of combine load forecast via information transformation and language evaluation. *Proceedings of the Chinese Society of Universities for Electric Power System and Its Automation, 27(2),* 20–26.
3. Jin, X., Luo, D., Sun, G., Zhang, H., Zheng, D. et al. (2012). Sifting and combination method of medium and long term load forecasting model. *Proceedings of the Chinese Society of Universities for Electric Power System and Its Automation, 24(4),* 150–156.
4. Sun, G., Yao, J., Xie, Y., Bu, H. (2009). Combination forecast of medium and long term load using fuzzy adaptive variable weight based on fresh degree function and forecasting availability. *Power System Technology, 33(9),* 103–107.
5. Yang, J., Ouyang, S., Shi, Y., Huang, R., Liu, Z. (2014). Combined membership function and its application on fuzzy evaluation of power quality. *Advanced Technology of Electrical Engineering and Energy, 33(2),* 63–69.
6. Huang, Y., Xu, Z., Liu, J., Zhou, R. (2005). Combining forecast based on fuzzy optimal selection of multi attributes semi structural decision-making. *Electric Power, 38(9),* 61–65.
7. Li, C., Niu, D., Meng, L. (2009). A comprehensive model for long- and medium-term load forecasting based on analytic hierarchy process and radial basis function neural network. *Power System Technology, 33(2),* 99–104.
8. Niu, D., Lu, H., Zhang, Y. (2008). Combination method of mid-long term load forecasting based on support vector machine within the Bayesian evidence framework. *Journal of North China Electric Power University, 35(6),* 62–66.
9. Xiao, X., Ge, J., He, D. (2008). Combination method of mid-long term load forecasting based on support vector machine. *Proceedings of the Chinese Society of Universities for Electric Power System and Its Automation, 20(1),* 84–88.
10. Ma, X., Yan, B., Tang, Y., Zhang, J. (2015). Medium and long-term load forecasting based on optimized combination forecast technology. *Proceedings of the Chinese Society of Universities for Electric Power System and its Automation, 27(6),* 62–67.
11. Jiang, Y., Wang, S., Feng, Y. (2012). Medium-long term power load forecasting based on recursive right combination model. *Proceedings of the Chinese Society of Universities for Electric Power System and Its Automation, 24(1),* 151–155.
12. Zhou, Q., Ren, H., Li, J., Zhang, Y., Zhou, Y. et al. (2010). Variable weight combination method for mid-long term power load forecasting based on hierarchical structure. *Proceedings of the Chinese Society for Electrical Engineering, 30(16),* 47–52.
13. You, S., Cheng, H., Xie, H., Guo, W., Lu, J. (2004). Application of combination forecasting by Fuzzy method in mid-and long-term load forecasting. *Proceedings of the Chinese Society of Universities for Electric Power System and Its Automation, 16(3),* 53–56.
14. Wilinski, A. (2019). Time series modeling and forecasting based on a Markov chain with changing transition matrices. *Expert Systems with Applications, 133(7),* 163–172. DOI 10.1016/j.eswa.2019.04.067.
15. Arruda, E., Fragoso, M., Ourique, F. (2019). A multi-cluster time aggregation approach for Markov chains. *Automatica, 99,* 382–389. DOI 10.1016/j.automatica.2018.10.027.

16. Zhu, C., Zhang, Y., Yan, Z., Zhu, J., Zhao, T. (2020). A wind power time series modeling method based on the improved Markov Chain Monte Carlo Method. *Transactions of China Electrotechnical Society, 35(3),* 577–589.
17. Wan, H., Ren, G. (2020). Study on the foundation pit settlement prediction model of logistic curve optimized by Markov Chain. *Journal of Geomatics Science and Technology, 37(2),* 166–170.
18. Wang, H., Liu, K., Zhu, H. (2016). Load forecasting of power system based on cloud model and support vector machine with RBF. *Journal of Lanzhou University of Technology, 42(4),* 85–88.
19. Li, H., Wei, M., Chen, Y., Lin, Z. (2016). Assessment method for power emergency group decision-making based on cloud model. *Electric Power, 49(4),* 42–48.
20. Wang, H., Liu, K., Yang, S. (2016). Load forecasting of power system based on cloud model SVM. *Computer Systems & Applications, 25(5),* 209–212.
21. Liu, D., Zhang, Q., Li, X., Gu, Y., Tan, Z. (2019). Identification of potential harmful behaviors in electricity market based on cloud model and fuzzy petri net. *Automation of Electric Power Systems, 43(2),* 25–33.
22. Yoder, M., Hering, A. S., Navidi, W. C., Larson, K. (2014). Short-term forecasting of categorical changes in wind power with Markov chain models. *Wind Energy, 17,* 1425–1439.
23. Wang, G., Xu, C., Li, D. (2014). Generic normal cloud model. *Information Sciences, 280,* 1–15. DOI 10.1016/j.ins.2014.04.051.
24. Li, L., Mao, C., Lei, B., Gao, Y., Liu, Y. et al. (2020). Decision-making of product-service system solution selection based on integrated weight and technique for order preference by similarity to an ideal solution. *IET Collaborative Intelligent Manufacturing, 2(3),* 102–108. DOI 10.1049/iet-cim.2020.0003.
25. Greco, S., Matarazzo, B., Slowinski, R. (2001). Rough sets theory for multicriteria decision analysis. *European Journal of Operational Research, 129(1),* 1–47. DOI 10.1016/S0377-2217(00)00167-3.
26. Li, J., Fang, H., Song, W. (2019). Failure mode and effects analysis using variable precision rough set theory and TODIM method. *IEEE Transactions on Reliability, 99,* 1–15.
27. He, Y., Wei, C., Long, H., Raza Ashfaq, R., Huang, J. (2018). Random weight network-based fuzzy nonlinear regression for trapezoidal fuzzy number data. *Applied Soft Computing Journal, 70(3),* 959–979. DOI 10.1016/j.asoc.2017.08.006.