ISSN: 2692-7608 DOI: 10.6981/FEM.202212_3(12).0017

Interval Number Combination Forecasting Model of China Railway Passenger Volume

Lingyun Hu*, Zonghua Jiang and Huaisheng Wang

School of Management Science and Engineering, Anhui University of Finance and Economics, Bengbu, Anhui 233030, China

Abstract

There exists widespread vague information in objective prediction problems, and the use of interval number to represent fuzzy information is a simple and effective processing method. In the forecast of China's railway passenger capacity, due to frequent data fluctuations, the interval number composed of the maximum and minimum annual passenger capacity is used to characterize and the research on the range combination forecast of China's railway passenger volume is carried out, which provide important data guarantee and technical support for the management department to formulate macro policies. In this paper, a type of interval number integral distance is selected as the optimal criterion and an induced ordered weighted average operator (IOWA) is introduced to construct a combined forecast model of railway passenger traffic interval based on interval number integral distance and IOWA operator. And the data of China's railway passenger traffic from 1997 to 2017 are empirically analyzed, then the results are compared with the single prediction method, which shows that the combined prediction method constructed in this paper can effectively improve the prediction accuracy and is obviously superior to the single prediction method.

Keywords

Combination Forecast; The Integral Distance of the Interval Number; IOWA Operator; China's Railway Passenger Traffic; Optimal Model.

1. Introduction

"The Doctrine Of The Mean": In all things,success lies in previous preparations and there will be failure without previous preparations.It can be seen that in objective life, prediction is very important work, and accurate prediction values can play a key role in scientific decision-making. Since each forecasting method cannot accurately match the actual value, it is necessary to gather several forecasting methods together to build a combined forecasting model and give full play to the advantages of each forecasting method. Bates.J.M. and Granger.C.W.J. systematically proposed the combination prediction method in 1969[1], and then scholars at home and abroad carried out in-depth research on the combination prediction [2-5], and it was widely used in the prediction problems of all walks of life, and achieved good prediction results.

Nowadays, more and more prediction problems are accompanied by fuzzy phenomena, and people are increasingly willing to use fuzzy numbers to describe and portray them. Since interval numbers are one of the simplest fuzzy numbers, it is necessary to carry out research on the theory and methods of interval number combination prediction. By introducing the correlation coefficient and the COWA operator, Zhou Dai-di (2020) et al. established an optimal interval number combination forecasting model based on the correlation coefficient, while using the correlation coefficient as a measure of the contribution of single forecasting methods in combination with the Shapley value in cooperative countermeasures to give an approximate solution to the corresponding optimal combination forecasting model [6]. Hu Lingyun (2021)

ISSN: 2692-7608 DOI: 10.6981/FEM.202212_3(12).0017

et al. introduced the new generalized induced ordered weighted proportional averaging (GIOWPA) operator with the criterion of the sum of absolute values of errors and established a combined prediction model for the optimal number of intervals based on the GIOWPA operator and the sum of absolute values of errors [7]. Yuan Hongjun (2019) et al. constructed the COWG-WPA operator to develop an optimal interval number combination prediction model based on the ordered geometric weighted average Power operator for continuous intervals under the criterion of generalized absolute error subsum [8]. Jin Feifei (2013) et al. redefined the concepts of interval prediction relative error, prediction accuracy and non-inferiority combination prediction from the perspective of continuous interval ordered weighted average operator, and established an optimal interval number combination prediction model based on continuous interval ordered weighted average operator using the sum of squared interval combination prediction errors as a criterion [9]. Yu Zhiqiang (2022) et al. proposed a new induced ordered weighted continuous interval generalized ordered weighted multiple averaging (IOWC-GOWMA) operator with exponential support as a criterion, and established an optimal interval number combination prediction model based on the IOWC-GOWMA operator and exponential support [10]. Wei Xin (2020) et al. pointed out the deficiency that the first-order prediction validity only takes into account the mathematical expectation of the prediction accuracy and ignores the mathematical variance of the prediction accuracy, and established a combined prediction model for the optimal number of intervals based on the second-order prediction validity and the induced continuous ordered weighted average (ICOWA) operator [11]. Hu Lingyun (2013) et al. took the left and right endpoints of the interval number as the starting point for considering the problem, introduced the induced ordered weighted geometric average (IOWGA) operator, established the optimal combined prediction model of the IOWGA operator for the left and right endpoints respectively using the log error sum of squares as the criterion, and used the preference coefficients to transform the multi-objective optimal model into a single-objective optimal model to achieve the solution [12]. Jiang Lihui (2015) et al. combined the C-GOWA operator and the IOWA operator to propose the induced ordered weighted continuous interval generalized ordered weighted average (IOWC-GOWA) operator, and used the exponential support degree as a criterion to establish the optimal interval number combination prediction model with variable weight coefficients [13]. Yuan Hongjun (2016) et al. addressed the shortcomings of the continuous ordered weighted geometric averaging (C-OWGA) operator which can only aggregate single interval data, and proposed the induced generalized ordered weighted continuous ordered weighted geometric averaging (IGOWC-OWGA) operator with the vector entropy cosine as a criterion to establish the optimal interval number combination prediction model [14].

In a study on the forecasting problem of railway passenger traffic, Ge Ling (2018) et al. used monthly data on railway passenger traffic in China from 2005-2016, modelled using two multiplicative seasonal models, and forecast railway passenger traffic for January-June 2017 [15]. Based on the railway passenger volume data of Shaanxi Province from 2005-2016, Wang Lisa (2018) et al. applied a Markov process to correct the forecast values and forecast the railway passenger volume of Shanxi Province from 2017-2022 based on the grey GM(1,1) model forecast [16]. Wang Zhihong(2013) used the idea of mobile holiday effect and Genhol procedure to improve the previous understanding that the impact of mobile holidays on passenger traffic is a uniform variation, and designed a bimodal variation process of the impact of mobile holidays on passenger traffic, and established a three-period seasonal adjustment model of Chinese railway passenger traffic based on the X-12-ARIMA model [17].Based on system dynamics, Qi Shan (2016) et al. analyzed the main factors affecting passenger traffic, extracted three key independent variables such as GDP, railway business mileage and civilian car ownership using stepwise regression methods, and established a medium- and long-term prediction model for railway passenger traffic under uncertain environment using fuzzy

ISSN: 2692-7608 DOI: 10.6981/FEM.202212_3(12).0017

multiple regression prediction theory, and the corresponding passenger traffic prediction results were changed from traditional real numbers to interval numbers [18]. The existing literature basically focuses on the forecasting of railway passenger traffic for real data, often using a single statistical forecasting method, and the research mainly focuses on empirical analysis, lacking the construction and application of fuzzy combination forecasting models. To address the characteristics of frequent data fluctuations in Chinese railway passenger traffic forecasting, this thesis adopts the interval number composed of the maximum and minimum values of annual passenger traffic to characterize Chinese railway passenger traffic, constructs a class of interval number integral distance as a new optimal criterion, introduces the induced ordered weighted average operator (IOWA), and establishes a combined railway passenger traffic interval number forecasting model based on interval number integral distance and IOWA operator, which is a new combined forecasting method of interval number with variable weight coefficients, circumventing the complex fuzzy data processing process and improving the forecasting accuracy. Finally, an empirical analysis is conducted with the interval number data of Chinese railway passenger traffic from 1997 to 2017, which again verifies that the proposed interval number combination forecasting model is a feasible and effective forecasting method. By carrying out research on the interval number combination forecasting of Chinese railway passenger traffic, it not only combines the innovative fuzzy combination forecasting method with practical application, but also provides important data guarantee and technical support for management departments to formulate macro policies.

2. Basic Concepts

Definition 1 If two real numbers satisfy $a \le b$, then $X=[a,b]$ is said to be an interval number on the set of real numbers R. Let $c=\frac{a+b}{2}$, $r=\frac{b-a}{2}$, then , c,r is said to be the midpoint and radius of the interval number X , respectively.

In definition 1, if $a=b$, then the interval number X will degenerate into a real number, so the real number is a special form of interval number, and the interval number is an extended form of the real number.

Definition 2 Let the two interval numbers $X_1=[a_1,b_1]$and $X_2=[a_2,b_2]$, they have the following operation rules:

(1) Equality operation: $X_1=X_2 \Leftrightarrow a_1=a_2 且 b_1=b_2$;

(2) Addition operation: $X_1+X_2=[a_1+a_2,b_1+b_2]$;

(3) Subtraction operation: $X_1-X_2=[a_1-b_2,b_1-a_2]$;

(4) Number multiplication operation: $\lambda X_1=\begin{cases}[\lambda a_1,\lambda b_1] & \lambda\ge 0\\ [\lambda b_1,\lambda a_1] & \lambda<0\end{cases}$.

Definition 2 can be used to perform different operations on two or more finite interval numbers, which provides a rule basis for the interval number data set knot in the interval number combination prediction process.

Definition 3 Let the two interval numbers $X_1=[a_1,b_1]$ and $X_2=[a_2,b_2]$, the midpoints are $c_1=\frac{a_1+b_1}{2}$and $c_2=\frac{a_2+b_2}{2}$, the radii are $r_1=\frac{b_1-a_1}{2}$ and $r_2=\frac{b_2-a_2}{2}$, let [19]:

ISSN: 2692-7608 DOI: 10.6981/FEM.202212_3(12).0017

$$D^2(X_1,X_2)=\int_{-\frac{1}{2}}^{\frac{1}{2}}\left\{[\frac{b_1+a_1}{2}+(b_1-a_1)x]-[\frac{b_2+a_2}{2}+(b_2-a_2)x]\right\}^2 dx$$

$$=(\frac{b_1+a_1}{2}-\frac{b_2+a_2}{2})^2+\frac{1}{3}\left(\frac{b_1-a_1}{2}-\frac{b_2-a_2}{2}\right)^2$$

$$=(c_1-c_2)^2+\frac{1}{3}(r_1-r_2)^2$$

$$D(X_1,X_2)=\sqrt{(c_1-c_2)^2+\frac{1}{3}(r_1-r_2)^2}=\sqrt{\frac{1}{3}\left[(a_1-a_2)^2+(a_1-a_2)(b_1-b_2)+(b_1-b_2)^2\right]}$$

$D(X_1,X_2)$ is called the integral distance between the number of intervals X_1 and the number of intervals X_2.

In definition 3, a distance formula is given with the help of definite integrals, and some common distance formulas are often calculated by the left and right endpoints of the interval number, ignoring the internal points of the interval number, in contrast, the integral distance in definition 3 can more effectively measure the proximity of the two interval numbers, which can be used as a measure index to measure the closeness of the two interval numbers.

Definition 4 Let OWA and IOWA are n-ary functions of $R^n \to R$, and $W=(w_1,w_2,\cdots,w_n)^T$ is weighted vectors related to the function, satisfying $\sum_{i=1}^{n} w_i = 1$, $w_i \geq 0$, $i=1,2,\cdots,n$, and the data sequence $\{(\theta_1,a_1),(\theta_2,a_2),\cdots,(\theta_n,a_n)\}$ contains *n* two-dimensional array, such that[3]:

$$OWA(a_1,a_2,\cdots,a_n)=\sum_{i=1}^{n} w_i b_i,$$

$$IOWA((\theta_1,a_1),(\theta_2,a_2),\cdots,(\theta_n,a_n))=\sum_{i=1}^{n} w_i a_{\theta-index(i)},$$

Arrange $a_1,a_2,\cdots,a_n$ from largest to smallest, b_i is the ith largest number, and the function OWA is said to be a n dimensional ordered weighted average operator; Arrange $\theta_1,\theta_2,\cdots,\theta_n$ from largest to smallest, $\theta-index(i)$ is the subscript of the ith largest number, θ_i called the induced value of a_i, and says that the function IOWA is the n dimensionally induced ordered weighted average operator.

Definition 4 gives a method of data set knot for one-dimensional data series and two-dimensional array sequence, OWA operator and IOWA operator have excellent properties such as monotonicity, permutation invariance, idempotency and mesovalence, and the specific properties and their proofs can be detailed in the literature [3].

3. Constructing the Interval Number Combination Forecasting Model of Interval Number Integration Distance and IOWA Operator

Suppose that in an uncertain prediction problem, the sequence of the number of actual value intervals is $\{X_t | X_t=[a_t,b_t], t=1,2,\cdots,N\}$. A single prediction method is used to predict the actual value data series, and the sequence group of each single prediction value interval obtained is $\{X_{it} | X_{it}=[a_{it},b_{it}], i=1,2,\cdots,m; t=1,2,\cdots,N\}$, the weight coefficient of each single forecasting

ISSN: 2692-7608 DOI: 10.6981/FEM.202212_3(12).0017

method in the combined forecasting is denoted as $w_1, w_2, \cdots, w_m$, and meet $w_i \in [0,1]$, $\sum_{i=1}^{m} w_i = 1$.

The interval number of combined forecast value formed by the combination of m single forecasting methods in the form of arithmetic weighting is denoted as $\hat{X}_t = [\hat{a}_t, \hat{b}_t], t = 1,2,\cdots,N$, according to the relevant operations in Definition 2, that is:

$$\hat{X}_t = \sum_{i=1}^{m} w_i X_{it} = \sum_{i=1}^{m} w_i [a_{it}, b_{it}] = \left[\sum_{i=1}^{m} w_i a_{it}, \sum_{i=1}^{m} w_i b_{it}\right],$$

$$\hat{a}_t = \sum_{i=1}^{m} w_i a_{it}, \hat{b}_t = \sum_{i=1}^{m} w_i b_{it},$$

In the above combined prediction value interval number, the left and right end points of various single prediction methods at each time point are combined according to the fixed weight coefficient. This method cannot effectively distinguish the difference between high and low accuracy, so it needs to be improved and optimized.

Definition 5 Let:

$$\theta_{it} = \begin{cases} 1 - \left|(a_t - a_{it})/a_t\right| & when\ \left|(a_t - a_{it})/a_t\right| < 1 \\ 0 & when\ \left|(a_t - a_{it})/a_t\right| \geq 1 \end{cases},$$

$$\delta_{it} = \begin{cases} 1 - \left|(b_t - b_{it})/b_t\right| & when\ \left|(b_t - b_{it})/b_t\right| < 1 \\ 0 & when\ \left|(b_t - b_{it})/b_t\right| \geq 1 \end{cases},$$

θ_{it} is the left accuracy of the number of single prediction intervals of the i prediction method at time t, and δ_{it} is the right accuracy of the number of single prediction intervals of the i prediction method at time t. Obviously, $\theta_{it} \in [0,1]$, $\delta_{it} \in [0,1]$, $i = 1,2,\cdots,m; t = 1,2,\cdots,N$.

For each single prediction method, the left precision θ_{it} of the single prediction interval number and the left endpoint a_{it} of the interval number are combined together, they can form m two-dimensional arrays $(\theta_{1t}, a_{1t}), (\theta_{2t}, a_{2t}), \cdots, (\theta_{mt}, a_{mt})$, and the left precision θ_{it} is taken as the induced value; Similarly, m two-dimensional arrays $(\delta_{1t}, b_{1t}), (\delta_{2t}, b_{2t}), \cdots, (\delta_{mt}, b_{mt})$ can be formed by combining the right precision δ_{it} of the single prediction interval number and the right endpoint b_{it} of the interval number, and the right precision δ_{it} is taken as the induced value.

Definition 6 Let:

$$\hat{a}_{\theta-index(t)} = IOWA((\theta_{1t}, a_{1t}), (\theta_{2t}, a_{2t}), \cdots, (\theta_{mt}, a_{mt})) = \sum_{i=1}^{m} w_i a_{\theta-index(it)},$$

$$\hat{b}_{\delta-index(t)} = IOWA((\delta_{1t}, b_{1t}), (\delta_{2t}, b_{2t}), \cdots, (\delta_{mt}, b_{mt})) = \sum_{i=1}^{m} w_i b_{\delta-index(it)},$$

$\hat{a}_{\theta-index(t)}$ is the predicted value of the combination of the left end points of the interval number based on the IOWA operator at time t, and $\hat{b}_{\delta-index(t)}$ is the predicted value of the combination of the right end points of the interval number based on the IOWA operator at time t.

According to Definition 6, it can be seen that the combined prediction value $\hat{a}_{\theta-index(t)}$ of the left endpoint of interval number based on IOWA operator is closely related to the left accuracy of each single prediction method at each time point. At the same time, the combined prediction

ISSN: 2692-7608 DOI: 10.6981/FEM.202212_3(12).0017

value $\hat{b}_{\delta-index(t)}$ of the right endpoint of interval number based on IOWA operator is also closely related to the right accuracy of each single prediction method at each time point. According to this rule, the weight coefficients of the left and right end points of the corresponding interval number to each single prediction method at each time point are not always unchanged, but will change with the change of its left and right accuracy. In this case, the interval number of combined predicted value based on IOWA operator is:

$$\hat{X}_t = [\hat{a}_{\theta-index(t)}, \hat{b}_{\delta-index(t)}] = \left[\sum_{i=1}^{m} w_i a_{\theta-index(it)}, \sum_{i=1}^{m} w_i b_{\delta-index(it)}\right].$$

The midpoint and radius of the combined predicted value interval number based on IOWA operator respectively are:

$$\hat{c}_t = \frac{\hat{a}_{\theta-index(t)} + \hat{b}_{\delta-index(t)}}{2} = \frac{\sum_{i=1}^{m} w_i a_{\theta-index(it)} + \sum_{i=1}^{m} w_i b_{\delta-index(it)}}{2},$$

$$\hat{r}_t = \frac{\hat{b}_{\delta-index(t)} - \hat{a}_{\theta-index(t)}}{2} = \frac{\sum_{i=1}^{m} w_i b_{\delta-index(it)} - \sum_{i=1}^{m} w_i a_{\theta-index(it)}}{2}.$$

Definition 7 Let:

$$D_i(X_t, X_{it}) = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left[(c_t - c_{it})^2 + \frac{1}{3}(r_t - r_{it})^2\right]},$$

$$D(X_t, \hat{X}_t) = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left[(c_t - \hat{c}_t)^2 + \frac{1}{3}(r_t - \hat{r}_t)^2\right]},$$

$D_i(X_t, X_{it})$ is the interval number integration distance between the sequence of the i single prediction method and the actual value sequence. $D(X_t, \hat{X}_t)$ is the interval number integration distance between the combined predicted value sequence and the actual value sequence based on IOWA operator.

The midpoint and radius of the actual value interval number and the combined predicted value interval number based on IOWA operator are substituted into the formula of Definition 7, then:

$$D(X_t, \hat{X}_t) = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left[\left(\frac{a_t + b_t}{2} - \frac{\sum_{i=1}^{m} w_i a_{\theta-index(it)} + \sum_{i=1}^{m} w_i b_{\delta-index(it)}}{2}\right)^2 + \frac{1}{3}\left(\frac{b_t - a_t}{2} - \frac{\sum_{i=1}^{m} w_i b_{\delta-index(it)} - \sum_{i=1}^{m} w_i a_{\theta-index(it)}}{2}\right)^2\right]}$$

$$= \sqrt{\frac{1}{4N}\sum_{t=1}^{N}\left[\left(\sum_{i=1}^{m} w_i(a_t - a_{\theta-index(it)}) + \sum_{i=1}^{m} w_i(b_t - b_{\delta-index(it)})\right)^2 + \frac{1}{3}\left(\sum_{i=1}^{m} w_i(b_t - b_{\delta-index(it)}) - \sum_{i=1}^{m} w_i(a_t - a_{\theta-index(it)})\right)^2\right]}$$

$$= \sqrt{\frac{1}{3N}\sum_{t=1}^{N}\left[\left(\sum_{i=1}^{m} w_i(a_t - a_{\theta-index(it)})\right)^2 + \left(\sum_{i=1}^{m} w_i(a_t - a_{\theta-index(it)})\right)\left(\sum_{i=1}^{m} w_i(b_t - b_{\delta-index(it)})\right) + \left(\sum_{i=1}^{m} w_i(b_t - b_{\delta-index(it)})\right)^2\right]}$$

From the above results, if the interval number integration distance is smaller, the two interval number sequences will be closer, especially when $D(X_t, \hat{X}_t) = 0$, the two interval number

ISSN: 2692-7608 DOI: 10.6981/FEM.202212_3(12).0017

sequences X_t and $\hat{X}_t$ completely coincide. The above equation $D(X_t, \hat{X}_t)$ can be regarded as a function of the weight coefficient $w_1, w_2, \cdots, w_m$ of each single prediction method. When the problem of interval number combination prediction is investigated from the perspective of IOWA operator and interval number integration distance, the smaller the value of $D(X_t, \hat{X}_t)$ is, the better it is. The following optimal interval number combination prediction model can be established:

$$\min D = \sqrt{\frac{1}{3N}\sum_{t=1}^{N}\left[\left(\sum_{i=1}^{m} w_i(a_t - a_{\theta-index(it)})\right)^2 + \left(\sum_{i=1}^{m} w_i(a_t - a_{\theta-index(it)})\right)\left(\sum_{i=1}^{m} w_i(b_t - b_{\delta-index(it)})\right) + \left(\sum_{i=1}^{m} w_i(b_t - b_{\delta-index(it)})\right)^2\right]}$$

$$s.t.\begin{cases} \sum_{i=1}^{m} w_i = 1 \\ w_i \ge 0, i = 1,2,\cdots,m \end{cases} \qquad (\#)$$

Definition 8 Record $D_{\min} = \min_{1\le i\le m}\{D_i\}$, $D_{\max} = \max_{1\le i\le m}\{D_i\}$, $D(X_t, \hat{X}_t)$ as the interval number integration distance between the combined predicted value sequence and the actual value sequence based on IOWA operator, then:

(1) When $D(X_t, \hat{X}_t) < D_{\min}$, the model (#) is the combination prediction of optimal interval number;

(2) When $D_{\min} \le D(X_t, \hat{X}_t) \le D_{\max}$, the model (#) is called the combination prediction of non-inferior intervals;

(3) When $D(X_t, \hat{X}_t) > D_{\max}$, the model (#) is the combination prediction of inferior intervals.

4. Interval Number Combination Prediction of Chinese Railway Passenger Volume based on Interval Number Integral Distance and IOWA Operator

To test the effectiveness of interval number combination forecast model based on the interval Numbers integral distance and the IOWA operator , we do empirical analysis on China's railway passenger traffic data, considering the practical prediction problems existing in the uncertain phenomena, so the maximum and the minimum passenger traffic data a year, forming China's annual passenger traffic interval data to study.Data from china economic network statistical database, time from 1997 to 2017, respectively using three single forecast methods ,the nonlinear time series prediction method, gray prediction, exponential smoothing prediction , put the actual value of interval number of China's railway passenger traffic and three kinds of single forecasting interval Numbers data in the table, as shown in Table 1.

Table 1. Actual value intervals number and three single predicted value intervals number (unit: million people)

time	actual value intervals Number	Single-item Method 1 Predict intervals number	Single-item Method 2 Predict intervals number	Single-item Method 3 Predict of intervals number
1997	[67,83]	[67.00,83.00]	[53.85,67.38]	[67.00,83.00]
1998	[68,82]	[68.00,82.00]	[57.68,72.73]	[70.20,88.20]

ISSN: 2692-7608 DOI: 10.6981/FEM.202212_3(12).0017

1999	[72,91]	[71.08,87.50]	[61.77,78.50]	[71.30,85.91]
2000	[72,94]	[74.54,92.80]	[66.16,84.72]	[74.50,94.05]
2001	[72,95]	[76.06,90.69]	[70.85,91.44]	[74.85,98.61]
2002	[73,96]	[79.99,93.79]	[75.88,98.70]	[74.11,98.85]
2003	[76,96]	[76.54,96.11]	[81.27,106.52]	[74.16,98.14]
2004	[77,105]	[80.59,102.57]	[87.03,114.97]	[76.69,96.81]
2005	[85,112]	[83.27,112.87]	[93.21,124.09]	[78.32,106.85]
2006	[93,122]	[89.96,121.55]	[99.83,133.94]	[86.13,117.46]
2007	[99,135]	[99.16,130.76]	[106.91,144.56]	[96.16,130.00]
2008	[103,141]	[105.95,140.73]	[114.50,156.02]	[104.44,145.44]
2009	[109,150]	[111.89,150.95]	[122.63,168.40]	[109.50,152.07]
2010	[122,162]	[121.13,161.55]	[131.33,181.76]	[115.02,159.27]
2011	[131,182]	[132.35,174.07]	[140.65,196.17]	[127.53,171.19]
2012	[142,185]	[142.80,184.21]	[150.64,211.73]	[138.82,194.28]
2013	[140,203]	[147.78,206.85]	[161.33,228.53]	[150.89,198.71]
2014	[160,235]	[156.91,232.23]	[172.78,246.65]	[150.49,214.72]
2015	[150,255]	[170.23,241.78]	[185.04,266.21]	[166.26,252.26]
2016	[204,280]	[183.21,269.17]	[198.18,287.33]	[160.31,279.11]
2017	[226,307]	[198.17,288.74]	[212.24,310.12]	[206.67,305.15]

Table 2. Interval number left accuracy and interval number right accuracy of each single prediction method

time	Precision of the left θ_{1t}	Precision of the left θ_{2t}	Precision of the left θ_{3t}	The right accuracy δ_{1t}	The right accuracy δ_{2t}	The right accuracy δ_{3t}
1997	1.0000	0.8037	1.0000	1.0000	0.8118	1.0000
1998	1.0000	0.8482	0.9676	1.0000	0.8870	0.9244
1999	0.9872	0.8579	0.9903	0.9615	0.8626	0.9441
2000	0.9647	0.9189	0.9653	0.9872	0.9013	0.9995
2001	0.9436	0.9840	0.9604	0.9546	0.9625	0.9620
2002	0.9042	0.9605	0.9848	0.9770	0.9719	0.9703
2003	0.9929	0.9307	0.9758	0.9989	0.8904	0.9777
2004	0.9534	0.8697	0.9960	0.9769	0.9050	0.9220
2005	0.9796	0.9034	0.9214	0.9922	0.8921	0.9540
2006	0.9673	0.9266	0.9261	0.9963	0.9021	0.9628
2007	0.9984	0.9201	0.9713	0.9686	0.9292	0.9630
2008	0.9714	0.8883	0.9860	0.9981	0.8935	0.9685
2009	0.9735	0.8750	0.9954	0.9937	0.8773	0.9862
2010	0.9929	0.9235	0.9428	0.9972	0.8780	0.9831
2011	0.9897	0.9263	0.9735	0.9564	0.9221	0.9406
2012	0.9944	0.9392	0.9776	0.9957	0.8555	0.9498
2013	0.9444	0.8476	0.9222	0.9810	0.8742	0.9789
2014	0.9807	0.9201	0.9406	0.9882	0.9504	0.9137
2015	0.8651	0.7664	0.8916	0.9482	0.9560	0.9893
2016	0.8981	0.9715	0.7858	0.9613	0.9738	0.9968
2017	0.8769	0.9391	0.9145	0.9405	0.9898	0.9940

ISSN: 2692-7608 DOI: 10.6981/FEM.202212_3(12).0017

According to definition 5, the interval number left accuracy θ_{it} and interval number right accuracy δ_{it} of the first single item prediction method at the t time are respectively calculated, as shown in Table 2.

According to the data in table 1 and table 2,combine interval number left precision θ_{it} and interval number left endpoint $a_{it}$of the each single prediction , then combine interval number right precision $\delta_{it}$and interval number right endpoints b_{it} of the each single prediction, use definition 6 regrouped into The left and right endpoints based on combination forecast interval number of the IOWA operator, again Generation into the optimal interval number combination forecast model (#), Through the LINGO software, the optimal weight coefficient is obtained as follows:

$$w_1 = 1.0000, w_2 = 0, w_3 = 0$$

Table 3. The Number of actual value intervals and the number of combined predicted value intervals of IOWA Operator (unit: million people)

time	Number of actual value intervals		The number of combined predicted value intervals of the IOWA operator	
	$[a_t, b_t]$	(c_t, r_t)	$[\hat{a}_t, \hat{b}_t]$	$(\hat{c}_t, \hat{r}_t)$
1997	[67,83]	(75.0,8.0)	[67.00,83.00]	(75.000,8.000)
1998	[68,82]	(75.0,7.0)	[68.00,82.00]	(75.000,7.000)
1999	[72,91]	(81.5,9.5)	[71.30,87.50]	(79.400,8.100)
2000	[72,94]	(83.0,11.0)	[74.50,94.05]	(84.275,9.775)
2001	[72,95]	(83.5,11.5)	[70.85,91.44]	(81.145,10.295)
2002	[73,96]	(84.5,11.5)	[74.11,93.79]	(83.950,9.840)
2003	[76,96]	(86.0,10.0)	[76.54,96.11]	(86.325,9.785)
2004	[77,105]	(91.0,14.0)	[76.69,102.57]	(89.630,12.940)
2005	[85,112]	(98.5,13.5)	[83.27,112.87]	(98.070,14.800)
2006	[93,122]	(107.5,14.5)	[89.96,121.55]	(105.755,15.795)
2007	[99,135]	(117.0,18.0)	[99.16,130.76]	(114.960,15.800)
2008	[103,141]	(122.0,19.0)	[104.44,140.73]	(122.585,18.145)
2009	[109,150]	(129.5,20.5)	[109.50,150.95]	(130.225,20.725)
2010	[122,162]	(142.0,20.0)	[121.13,161.55]	(141.340,20.210)
2011	[131,182]	(156.5,25.5)	[132.35,174.07]	(153.210,20.860)
2012	[142,185]	(163.5,21.5)	[142.80,184.21]	(163.505,20.705)
2013	[140,203]	(171.5,31.5)	[147.78,206.85]	(177.315,29.535)
2014	[160,235]	(197.5,37.5)	[156.91,232.23]	(194.570,37.660)
2015	[150,255]	(202.5,52.5)	[166.26,252.26]	(209.260,43.000)
2016	[204,280]	(242.0,38.0)	[198.18,279.11]	(238.645,40.465)
2017	[226,307]	(266.5,40.5)	[212.24,305.15]	(258.695,46.455)

It should be noted that the optimal interval number combination prediction model (#) is a variable weight coefficient combination prediction method, that is, using the assembly characteristics of IOWA operator data , the single prediction method with the maximum prediction accuracy at each time point is given a large weight, and the single prediction method with the minimum prediction accuracy is given a small weight.In the above optimal weight coefficient, the left endpoint corresponding to the maximum left accuracy and the right endpoint corresponding to the maximum right accuracy of the three single prediction methods

ISSN: 2692-7608 DOI: 10.6981/FEM.202212_3(12).0017

at each time point are respectively taken as the combined prediction value interval number, which is not only considering the prediction results of the first single prediction method.

The method of combining the left and right end points of a predicted value according to the IOWA operator in DEFINITION 6,Using the above optimal weight coefficient,the number of combined predicted value intervals at each time point can be calculated,as shown in Table 3.

In order to objectively judge the combined prediction intervals number of the IOWA operator, and three kinds of forecasting are interval Numbers, which one Predictive value interval number sequence closer to the actual value interval number sequence, then we can decide which kind of prediction method is more effective.we choose the following four evaluation index to construct the interval combination forecast effect evaluation index system:

Mean interval position error sum of squares: MSEP= $\frac{1}{N}\cdot\sum_{t=1}^{N}(c_t-\hat{c}_t)^2$

Mean interval length error sum of squares: MSEL= $\frac{1}{N}\cdot\sum_{t=1}^{N}(r_t-\hat{r}_t)^2$

Mean interval error sum of squares: MSEI= MSEP+ MSEL= $\frac{\sum_{t=1}^{N}(c_t-\hat{c}_t)^2+\sum_{t=1}^{N}(r_t-\hat{r}_t)^2}{N}$

Sum of mean interval relative errors: MRIE= $\frac{1}{N}\cdot\sum_{t=1}^{N}(\frac{|c_t-\hat{c}_t|}{(r_t+\hat{r}_t)})$

Where (c_t, r_t) represents the array formed by the midpoints and radii of the actual value interval number, and $(\hat{c}_t, \hat{r}_t)$ represents the array formed by the midpoints and radii of the predicted value interval number.The above evaluation indexes measure the proximity of two interval number series from the perspective of error. Obviously, the smaller the value of each evaluation index, the more accurate the prediction value, and the more effective the corresponding prediction method is.The four prediction effect evaluation index values of various methods can be obtained through calculation, as shown in Table 4.

Table 4. Values of prediction effect evaluation indexes of various methods

	MSEP	MSEL	MSEI	MRIE
Single item Forecasting Method 1	41.4928	19.9141	61.4069	0.0608
Single item Forecasting Method 2	156.7872	19.4968	176.2840	0.3290
Single item Forecasting Method 3	53.5512	38.3181	91.8692	0.1209
Optimal interval Number Combination Prediction Model (#)	9.2217	8.3670	17.5887	0.0464

According to the evaluation index data of each prediction method in Table 4, the optimal interval number combination prediction model (#) is far better than the corresponding evaluation index of the three single prediction methods in terms of MSEP, MSEL, MSEI and other evaluation indexes of the absolute error, and MRIE evaluation index of the relative error. Especially, it is significantly superior to the second single forecasting method,we can see clearly in this paper,interval number combination forecast model based on interval Numbers integral distance and IOWA operator , can effectively predict a path for China's railway passenger traffic, can get more accurate forecast, for the railway administrative department of the macroscopic policy to provide important data security and technical support.

ISSN: 2692-7608 DOI: 10.6981/FEM.202212_3(12).0017

Table 5. Interval number integral distance of various forecasting methods

	Single Item Method 1	Single Item Method 2	Single Item Method 3	Optimal Interval Number Combination Prediction Model (#)
Interval number integration distance	6.9376	12.7783	8.1439	3.4659

From table 5 all kinds of forecasting methods of interval Numbers integral distance data, in the optimal interval number combination prediction model (#),the interval number integration distance D of the optimal combination predicted value sequence and the actual value sequence is much smaller than the interval number integration distance D_1, D_2, D_3 of the actual value sequence and each single method predicted value sequence, that is, $D < \min(D_1, D_2, D_3)$,according to the superior combination forecasting and Inferior combination forecasting of definition 8 , The optimal interval number combination prediction model of China railway passenger volume based on interval number integration distance and IOWA operator is the superior combination forecasting model, which further indicates that the interval number combination prediction model proposed in this paper is a reasonable and effective prediction method.

5. Conclusion

There are many uncertain phenomena in objective real life, and fuzzy data are used to describe and quantitatively analyze, and fuzzy prediction and fuzzy combination prediction are carried out. In the process of forecasting China's railway passenger traffic, the real data can no longer meet the analysis of the problem, and then the interval number composed of the maximum and minimum values of the annual passenger traffic is used to characterize, and a new type of correlation index - interval number integral distance, together with the induced ordered weighted average operator (IOWA), is introduced into the research of interval number combination prediction, and a combined forecast model of railway passenger traffic interval number based on interval number integral distance and IOWA operator is established. Through the empirical analysis of China's railway passenger traffic data from 1997 to 2017, and the results are showed by two aspect which are the evaluation index of the prediction effect of the method and the advantages and disadvantages of the method. The optimal interval number combination prediction model (#) proposed in this paper is a feasible and effective prediction method. This method has generalizability, can use the model to deal with fuzzy prediction problems in other economic management fields, the model itself avoids the complex operation of fuzzy data, and brings great convenience in the process of use. Since the research on interval number combination prediction model is still very imperfect, how to construct a new optimal criterion formula, how to introduce a new more efficient set settler, and how to establish a higher precision interval number combination prediction method need to be studied in depth.

Acknowledgments

This work is supported by Planning Project for Philosophy and Social Science in Anhui Province (Project number: AHSKY2020D42), Major Research Fund Project of Anhui University of Finance and Economics (Project number: ACKYA21004).

ISSN: 2692-7608 DOI: 10.6981/FEM.202212_3(12).0017

References

[1] Bates J M, Granger C W J. Combination of Forecasts[J]. Operations Research Quarterly, 1969,20 (4): 451-468.

[2] Tang Xiaowo, Ma Yongkai, Zeng Yong, et al. Modern portfolio forecasting and portfolio investment decision-making methods and application research [M]. Beijing:Science Press,2003.

[3] Chen Huayou. The theory of effectiveness of portfolio forecasting methods and its application [M]. Beijing: Science Publishing House, 2008.

[4] Wang Yingming. A study of correlation-based portfolio forecasting methods [J]. Forecasting, 2002, 21 (2):58-62.

[5] Yuan Hongjun, Yang Guiyuan. An optimality combination forecasting model based on a closeness IOWA operator[J]. Statistics and Information Forum,2010,25(2):32-37

[6] Zhou Daidi, Chen Huayou, Zhang Qianyun, et al. Optimization model of correlation interval combination prediction based on COWA operator and its approximate solution[J]. Journal of the University of Science and Technology of China,2020,50(5):673-681.

[7] Hu Lingyun,Yang Wei,Wang Hengna. A new high-precision interval number combination prediction method[J]. Statistics and Decision Making,2021,37(24):5-10.

[8] Yuan Hongjun, Hu Lingyun. An interval-type portfolio forecasting model based on the ordered geometric weighted average Power operator for continuous intervals[J]. Mathematical Statistics and Management,2019,38(2):270-280.

[9] Jin Feifei,Li Jie,Chen Huayou,et al. Interval combination prediction based on continuous ordered weighted average operator[J]. Journal of Wuhan University of Technology (Information and Management Engineering Edition),2013,34(5):668-672.

[10] Yu Zhiqiang, Zhou Xue, Zhou Ligang, et al. An interval-type combination forecasting model based on IOWC-GOWMA operator and exponential support degree[J]. Fuzzy Systems and Mathematics, 2022, 36 (3):96-107.

[11] Wei Xin, Hao Jiangfeng, Chen Huayou, et al. A combined prediction model of ICOWA operator interval based on prediction validity[J]. Journal of Hefei University of Technology (Natural Science Edition),2020,43(11):1578-1584.

[12] Hu Lingyun,Yuan Hongjun. An interval-type combination prediction model based on the IOWGA operator with left and right endpoints[J]. Statistics and decision making,2013,29(11):22-25.

[13] Jiang LH, Chen HY, Ding FQ, et al. An interval combination prediction model based on the IOWC-GOWA operator[J]. Computer Engineering and Applications,2015,50(3):50-54.

[14] Yuan Hongjun,Zhang Chao. IGOWC-OWGA operator and its application in interval combination prediction[J]. Journal of Anhui University (Natural Science Edition),2016,40(1):11-17.

[15] Ge Ling,Zhang Jie. Railway passenger volume prediction based on multiplicative seasonal model[J]. Journal of Chongqing University of Technology and Business (Natural Science Edition), 2018, 35 (3): 18-25.

[16] Wang Lisha,Liang Kaihong. Forecasting railway passenger traffic in Shaanxi Province based on grey Markov model[J]. The practice and understanding of mathematics,2018,48(8):166-172.?

[17] Wang Zhihong,Wang Qianyuan. Research on the seasonal adjustment model of monthly passenger traffic of Chinese railways based on the Spring Festival factor[J]. Journal of Railway,2013,35(7):9-13.

[18] Qi Suan, Li Xia Miao, Wu Huishan, et al. A method for forecasting railway passenger traffic under uncertainty [J]. Journal of Railway Science and Engineering,2016,13(1):168-175.

[19] Liu Huawen. Fuzzy number ranking based on distance measure[J]. Journal of Shandong University (Science Edition),2004,45(2):30-36.

[20] Yager R.R. Families of OWA operators[J]. Fuzzy Sets and Systems,1993,59:125-148.