Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 6, No. 1, 2023

Interval Type Combination Prediction Model based on Improved Grey Correlation Degree IOWHA Operator

Feng Xu, Xiaowei Cai *

School of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu, 233030, China

*** Corresponding author:** Xiaowei Cai (Email: caixw555@163.com)

Abstract: In order to improve the accuracy of interval number prediction, this paper proposes a new type of interval number combination prediction model based on improved grey correlation degree and IOWHA operator. Transform the interval number into the center and radius of equivalent information, introduce the IOWHA operator, use the interval number prediction accuracy as the inducing factor, and improve the grey correlation degree as the optimal criterion to construct a variable weight interval type combination prediction model based on the improved grey correlation degree and IOWHA operator. And the variable weight coefficient interval type combination prediction model is applied to the prediction of interval number sequences, and the example proves that the model is effective.

Keywords: Interval Type Combination Prediction; Improve Grey Correlation Degree; IOWHA Operator.

1. Introduction

In the process of prediction practice, in order to achieve higher prediction accuracy, people generally use multiple prediction methods, and then combine each individual prediction method to establish a combination prediction model according to certain optimality criteria. In 1969, J M. Bates et al. innovatively proposed a combination prediction model, which enables it to fully utilize the effective information provided by all single prediction methods and reduce the risk brought by a single prediction method. In order to determine the weights of each individual prediction model in the combination prediction model, Yager proposed the Ordered Weighted Average (OWA) operator, laying the foundation for a series of subsequent information integration operator theories. In prediction practice, due to the existence of many uncertain factors in prediction problems, the predicted data is often not a single data, but within a range. Therefore, many scholars have started to use interval numbers to study prediction problems. Chen Huayou et al. proposed an interval type combination prediction model based on correlation coefficient and IOWA operator, and established a multi-objective interval combination prediction model. Zhu Jiaming et al. introduced the Power operator and the Continuous Ordered Weighted (COWA) operator, and used the sum of absolute error values as the optimal criterion to construct an interval type combination prediction model based on the uncertainty weighted Power average operator. Yuan Hongjun et al. converted interval numbers into equivalent binary connection numbers and established an interval type combination prediction model based on the optimal criterion of connection number closeness. In the above literature, the variable weight coefficient interval combination prediction with prediction accuracy as the induced value is more obvious, but the interval combination prediction is obviously based on the properties and characteristics of interval numbers, and the above prediction models cannot reflect all the information of interval number prediction. Therefore, this article proposes an interval type combination prediction model based on an improved grey correlation degree induced ordered harmonic weighted average (IOWHA) operator.

2. Basic Concepts

Definition 1: Set $H_w: R^m \to R$ is the meta function, $W = (w_1, w_2, \cdots w_m)^T$ is a weighted vector associated with H_w, $\sum_{i=1}^{m} w_i = 1, w_i \geq 0, i = 1,2,\cdots,m$, Let

$$H_w(a_1, a_2, \cdots, a_m) = \frac{1}{\sum_{i=1}^{m} \frac{w_i}{b_i}} \quad (1)$$

Among them, b_i is the i-th numerical value in descending order in $(a_1, a_2, \cdots, a_m)$, H_w is called the ordered weighted geometric mean operator, also known as the OWHA operator.

Definition 2: Assuming there are m two-dimensional arrays $\langle u_1, a_1\rangle, \langle u_2, a_2\rangle, \cdots, \langle u_m, a_m\rangle$, $a_i > 0, i = 1,2,\cdots,m$, $W = (w_1, w_2, \cdots, w_m)^T$ is a weighted vector associated with H_w, $\sum_{i=1}^{m} w_i = 1, w_i \geq 0, i = 1,2,\cdots,m$, Let

$$IOWHA_w(\langle u_1, a_1\rangle, \langle u_2, a_2\rangle, \cdots, \langle u_m, a_m\rangle) = \frac{1}{\sum_{i=1}^{m} \frac{w_i}{a_{u-index(i)}}}, \quad (2)$$

W is a weighted vector associated with H_w, Then H_w is called $u_1, u_2, \cdots, u_m$ the dimension induced ordered weighted average operator generated by m is also known as the IOWHA operator. $u-index(i)$ is the i-th numerical value in descending order in $u_1, u_2, \cdots, u_m$, u_i is the induced value of a_i .

Definition 3: Assuming the observed value is $s_t = [s_t^-, s_t^+] = (c_t, r_t), t = 1,2,\cdots,N$. $s_{it} = [s_{it}^-, s_{it}^+] = (c_{it}, r_{it})$ is the interval number prediction value of the i-th single item prediction method at time t, Among them, $i = 1,2,\cdots,m; t = 1,2,\cdots,N$。N represents the length of time, and m represents the number of single prediction methods. Assuming $w_i(i = 1,2,\cdots,m)$ is the weight obtained by the i-th single item prediction method, meet the conditions $\sum_{i=1}^{m} w_i = 1, w_i \geq 0$. $\hat{s}_t = w_i s_{it}, t = 1,2,\cdots,N$ is called the interval combination prediction value. According to the weighted average principle of interval number operation, the weighted combination prediction interval value at time t is:

$$\hat{s}_t = [\hat{s}_t^-, \hat{s}_t^+] = [w_i s_{it}^-, w_i s_{it}^+], i = 1,2,\cdots m; t = 1,2,\cdots,N \quad (3)$$

Let:

$$c_t = \frac{s_t^- + s_t^+}{2}, r_t = \frac{s_t^- - s_t^+}{2}, c_{it} = \frac{s_{it}^- + s_{it}^+}{2}, r_t = \frac{s_{it}^- - s_{it}^+}{2}, \quad (4)$$

Where c_t and r_t are the midpoint and radius of the observed values. c_{it} and r_{it} are the midpoints and radii of the interval numbers obtained by the method of single item prediction, $i = 1,2,\cdots m; t = 1,2,\cdots,N$

In order to facilitate data processing, the derivative error value of the sequence is generally considered in the harmonic average combination prediction model.

Definition 4: e_{it}^-, e_{it}^+ are referred to as the left and right precision of the number of single prediction intervals at time t for the i-th prediction method, respectively**,** Among them,

$$e_{it}^- = \frac{1}{s_{it}^-} - \frac{1}{s_t^-}, e_{it}^+ = \frac{1}{s_{it}^+} - \frac{1}{s_t^+}, \quad (5)$$

From this, the prediction accuracy sequences $e_{1t}^-, e_{2t}^-, \cdots e_{mt}^-$ and $e_{1t}^-, e_{2t}^-, \cdots e_{mt}^-$ for the left and right endpoints at the time can be obtained, by combining this sequence with prediction sequences $\hat{s}_{1t}^-, \hat{s}_{2t}^-, \cdots \hat{s}_{mt}^-$ and $\hat{s}_{1t}^+, \hat{s}_{2t}^+, \cdots \hat{s}_{mt}^+$, we can obtain two-dimensional arrays for the left and right endpoints, respectively $\{\langle e_{1t}^-, \hat{s}_{1t}^-\rangle, \langle e_{2t}^-, \hat{s}_{2t}^-\rangle, \cdots, \langle e_{mt}^-, \hat{s}_{mt}^-\rangle\}$ and $\{\langle e_{1t}^+, \hat{s}_{1t}^+\rangle, \langle e_{2t}^+, \hat{s}_{2t}^+\rangle, \cdots, \langle e_{mt}^+, \hat{s}_{mt}^+\rangle\}$. Define sequence by Definition 1 $e_{1t}^-, e_{2t}^-, \cdots e_{mt}^-$ and $e_{1t}^+, e_{2t}^+, \cdots e_{mt}^+$ as the induced value sequence, perform the following operation on the predicted sequence.

Let $W = (w_1, w_2, \cdots, w_m)^T$ weighted vectors for IOWHA in combination prediction using various prediction methods, Compare the prediction accuracy sequence e of the left and right endpoints of each single item prediction method at the moment of prediction $e_{1t}^-, e_{2t}^-, \cdots e_{mt}^-$ and $e_{1t}^+, e_{2t}^+, \cdots e_{mt}^+$ arrange in ascending order, Let $u - index(it)$ be the index of the i-th largest prediction accuracy, according to Definition 2, Let

$$h_w^-(\langle e_{1t}^-, \hat{s}_{1t}^-\rangle, \langle e_{2t}^-, \hat{s}_{2t}^-\rangle, \cdots, \langle e_{mt}^-, \hat{s}_{mt}^-\rangle) = \frac{1}{\sum_{i=1}^m \frac{w_i}{s_{u-index(it)}^-}}, \quad (6)$$

$$h_w^+(\langle e_{1t}^+, \hat{s}_{1t}^+\rangle, \langle e_{2t}^+, \hat{s}_{2t}^+\rangle, \cdots, \langle e_{mt}^+, \hat{s}_{mt}^+\rangle) = \frac{1}{\sum_{i=1}^m \frac{w_i}{s_{u-index(it)}^+}}, \quad (7)$$

Equations (6) and (7) are referred to as the prediction accuracy sequence e $e_{1t}^-, e_{2t}^-, \cdots e_{mt}^-$ and $e_{1t}^+, e_{2t}^+, \cdots e_{mt}^+$ the IOWHA combination prediction value of the left and right endpoints of the interval number generated.

Definition 5: If

$$\gamma_i^- = \frac{1}{N-1}\sum_{t=1}^{N-1}\frac{1}{1+\left|a^{(1)}(s_t^-) - a^{(1)}(s_{it}^-)\right| \Big/ (\sigma_{s_t^-}\sigma_{s_{it}^-})}, \quad (8)$$

$$\gamma_i^+ = \frac{1}{N-1}\sum_{t=1}^{N-1}\frac{1}{1+\left|a^{(1)}(s_t^+) - a^{(1)}(s_{it}^+)\right| \Big/ (\sigma_{s_t^+}\sigma_{s_{it}^+})}, \quad (9)$$

Then γ_i^+ and γ_i^- are referred to as the improved grey correlation between the i-th prediction method's predicted value reciprocal sequence and the actual value reciprocal sequence.

Definition 6: If

$$\gamma^- = \frac{1}{N-1}\sum_{t=1}^{N-1}\frac{1}{1+\left|a^{(1)}(s_t^-) - a^{(1)}(\hat{s}_t^-)\right| \Big/ (\sigma_{s_t^-}\sigma_{\hat{s}_t^-})}, \quad (10)$$

$$\gamma^+ = \frac{1}{N-1}\sum_{t=1}^{N-1}\frac{1}{1+\left|a^{(1)}(s_t^+) - a^{(1)}(\hat{s}_t^+)\right| \Big/ (\sigma_{s_t^+}\sigma_{\hat{s}_t^+})}, \quad (11)$$

Then γ^- and γ^+ are referred to as the improved grey correlation between the inverse sequence of IOWHA combined predicted values and the inverse sequence of actual values.

In definitions 5 and 6, $a^{(1)}(s_t^-), a^{(1)}(s_t^+), a^{(1)}(s_{it}^-), a^{(1)}(s_{it}^+), a^{(1)}(\hat{s}_t^-), a^{(1)}(s_{it}^+)$ and $\sigma_{s_t^-}, \sigma_{s_t^+}, \sigma_{s_{it}^-}, \sigma_{s_{it}^+}, \sigma_{\hat{s}_t^-}, \sigma_{s_{it}^+}$ are the cumulative generated sequences and standard deviations of $\frac{1}{s_t^-}, \frac{1}{s_t^+}, \frac{1}{s_{it}^-}, \frac{1}{s_{it}^+}, \frac{1}{\hat{s}_t^-}, \frac{1}{s_{it}^+}$, respectively.

$E^- = (E_{ij}^-)_{m\times m}$ is a combined prediction covariance information matrix of the m-order IOWHA operator, where: $E_{ij}^- = \sum_{t=1}^N e_{a-index(it)}^- e_{a-index(jt)}^-, i,j = 1,2,\cdots m, W = (w_1, w_2, \cdots, w_m)^T$ represents the weighted series vector of combination prediction, then

$$\gamma^- = \frac{1}{N-1}\sum_{t=1}^{N-1}\frac{1}{1+N\left|(e_{t+1}^- - e_t^-) - \sum_{i=1}^m w_i(e_{a-index(i(t+1))}^- - e_{a-index(it)}^-)\right| \Big/ \sqrt{\sum_{t=1}^N (e_t^-)^2}\sqrt{W^T EW}}, \quad (12)$$

Similarly definable

$$\gamma^+ = \frac{1}{N-1}\sum_{t=1}^{N-1}\frac{1}{1+N\left|(e_{t+1}^+ - e_t^+) - \sum_{i=1}^m w_i(e_{a-index(i(t+1))}^+ - e_{a-index(it)}^+)\right| \Big/ \sqrt{\sum_{t=1}^N (e_t^+)^2}\sqrt{W^T EW}} \quad (13)$$

From equations (12) and (13), it can be seen that the improved grey correlation degree γ^- and γ^+ between the interval type combination prediction sequence of the IOWHA operator and the interval type actual value sequence are functions of $W = (w_1, w_2, \cdots, w_m)^T$, denoted as $\gamma^-(W)$ and $\gamma^+(W)$, respectively. And according to the definition formula of the improved grey correlation degree, the larger the value, the better. When the improved grey correlation degree reaches its maximum value of 1, the combination prediction sequence and the actual value sequence are exactly the same. At this point, the prediction is completely accurate and error free. However, in prediction practice, errors are inevitable. Therefore, an optimal interval type combination prediction model based on improved grey correlation degree IOWHA operator is established to maximize its improved grey correlation degree. The established model represents the following model (I).

$$max\,\gamma(W) = \lambda\gamma^-(W) + (1-\lambda)\gamma^+(W)$$

$$\begin{cases} \sum_{i=1}^m w_i = 1 \\ w_i \geq 0, i = 1,2,\cdots,m \end{cases}$$

Among them, $\lambda \in [0,1]$ is the preference coefficient, which refers to the preference for the accuracy of the left and right endpoints of the predicted interval number.

3. Application Example of Interval Type Combination Prediction Model based on Improved Grey Correlation Degree IOWHA Operator

To verify the effectiveness of the interval type combination prediction model based on the improved grey correlation degree IOWHA operator, the following evaluation index system is selected:

Sum of squared average interval position errors: $MSEP = \sum_{t=1}^n (c_t - \hat{c}_t)^2 \Big/ n$,

Sum of squared errors in average interval length:

$$MSEL = \sum_{t=1}^{n}(r_t - \hat{r}_t)^2 \Big/ n,$$

Sum of squared mean interval error:

$$MSEI = MSEP + MSEL = \sum_{t=1}^{n}(c_t - \hat{c}_t)^2 \Big/ n + \sum_{t=1}^{n}(r_t - \hat{r}_t)^2 \Big/ n$$

This article applies the data from reference 4 (a paper by student Yuan Hongjun), where the preference coefficient takes a value of 0.5, indicating that the preference for the accuracy of the predicted interval number's left and right endpoints is the same. The data in reference 4 are the predicted values obtained by using three single prediction methods, namely exponential smoothing, support vector regression, and long short-term memory network, to predict the original interval number sequence. They are denoted as: x_{1t}, x_{2t}, x_{3t}

Table 1. Actual Interval Number Sequence and Prediction Sequence of Each Single Item Prediction Method

t	x_t	x_{1t}	x_{2t}	x_{3t}
1	[21.50,28.90]	[18.58,25.04]	[21.36,26.98]	[20.34,27.41]
2	[21.20,31.30]	[20.38,28.28]	[21.49,27.46]	[21.18,30.26]
3	[20.70,32.30]	[20.54,29.50]	[21.20,28.58]	[20.19,30.68]
4	[22.40,33.80]	[21.44,32.32]	[22.05,29.78]	[20.21,32.98]
5	[24.40,33.00]	[22.41,31.69]	[22.09,30.16]	[21.22,31.63]
6	[25.80,31.20]	[23.39,38.93]	[22.56,31.33]	[23.22,28.53]
7	[25.20,28.10]	[23.47,26.84]	[22.94,30.99]	[23.66,26.69]
8	[24.40,30.60]	[22.56,29.32]	[23.26,29.90]	[21.52,28.83]
9	[24.60,32.80]	[23.35,30.45]	[23.24,29.16]	[22.65,29.97]
10	[23.41,28.77]	[23.41,28.77]	[22.95,30.28]	[22.55,27.86]

According to Definition 3, the actual daily extreme temperature interval sequence and the predicted results of each individual prediction method in Table 1 are transformed into center and radius sequences, as shown in Table 2.

Table 2. Actual number of intervals and center and radius sequences of each individual prediction method

t	x_t	x_{1t}	x_{2t}	x_{3t}
1	(21.20,3.70)	(21.810,3.230)	(24.170,2.810)	(23.875,3.535)
2	(26.25,5.05)	(24.330,3.950)	(24.475,2.985)	(25.720,4.540)
3	(26.50,5.80)	(25.020,4.408)	(24.890,3.690)	(25.435,5.245)
4	(28.10,5.70)	(26.880,5.440)	(25.915,3.865)	(26.595,6.385)
5	(28.70,4.30)	(27.050,4.640)	(26.125,4.035)	(26.425,5.205)
6	(28.50,2.70)	(31.160,7.770)	(26.945,4.385)	(25.875,2.655)
7	(26.65,1.45)	(25.155,1.685)	(26.965,4.025)	(25.175,1.515)
8	(27.50,3.10)	(25.940,3.380)	(26.580,3.320)	(25.175,3.665)
9	(28.70,4.10)	(26.900,3.550)	(26.200,2.960)	(26.310,3.660)
10	(26.09,2.68)	(26.090,2.680)	(26.615,3.665)	(25.205,2.655)

For the actual number of intervals and the center and radius sequences of each individual prediction method mentioned above, construct an interval type combination prediction model based on an improved grey correlation degree IOWHA operator, namely the interval number combination prediction model (I). In the combination prediction model (I), a preference coefficient value of 0.5 is taken, which means that the preference for the accuracy of the left and right endpoints of the predicted interval is considered to be the same. A combination prediction model is established, and Lingo software is used to solve it. The weight coefficients of each individual prediction method in the combination prediction model are obtained as follows: $w_1 = 0.824, w_2 = 0.176, w_3 = 0.000$

Table 3. Center and radius sequences of actual interval numbers, individual prediction methods, and combination prediction methods

t	x_t	x_{1t}	x_{2t}	x_{3t}	$\hat{x}_t$
1	(21.20,3.70)	(21.810,3.230)	(24.170,2.810)	(23.875,3.535)	(23.997,3.572)
2	(26.25,5.05)	(24.330,3.950)	(24.475,2.985)	(25.720,4.540)	(25.579,4.729)
3	(26.50,5.80)	(25.020,4.408)	(24.890,3.690)	(25.435,5.245)	(25.728,5.523)
4	(28.10,5.70)	(26.880,5.440)	(25.915,3.865)	(26.595,6.385)	(26.718,5.638)
5	(28.70,4.30)	(27.050,4.640)	(26.125,4.035)	(26.425,5.205)	(27.139,4.524)
6	(28.50,2.70)	(31.160,7.770)	(26.945,4.385)	(25.875,2.655)	(27.842,3.121)
7	(26.65,1.45)	(25.155,1.685)	(26.965,4.025)	(25.175,1.515)	(26.526,2.125)
8	(27.50,3.10)	(25.940,3.380)	(26.580,3.320)	(25.175,3.665)	(26.674,3.132)
9	(28.70,4.10)	(26.900,3.550)	(26.200,2.960)	(26.310,3.660)	(27.814,3.724)
10	(26.09,2.68)	(26.090,2.680)	(26.615,3.665)	(25.205,2.655)	(26.274,2.756)

By calculating the weight coefficient results and combining them with the optimal interval type combination prediction model based on the improved grey correlation degree IOWHA operator, the combination prediction model results are obtained and recorded as $\hat{x}_t$ .So the actual interval number center and radius sequence, the center and radius sequence of each individual prediction method, and the center and radius sequence of the combined prediction method are shown in Table 3.

Based on the actual number of intervals in Table 3, the center and radius sequences of each individual prediction method and combination prediction method, combined with various error calculation formulas, the MSEP, MSEL, and MSEI of each individual prediction method and combination prediction method can be calculated. The specific results are shown in Table 4.

From the data in Table 4, it can be seen that the prediction error indicators of the proposed combination prediction model, except for MSEL in the LSTM method, are significantly lower than the prediction results of the original interval number sequence by the three single prediction methods of exponential smoothing, SVR, and LSTM. Further analysis of the prediction results of the LSTM method and the combination prediction model ($\lambda = 4$) proposed in reference 4 reveals that although these two methods perform well in predicting the radius, the central error is relatively large because the interval number is composed of the center and radius, indicating that the LSTM prediction method and the combination prediction model ($\lambda = 4$) proposed in reference 4 have not comprehensively considered the internal connection and overall nature of the interval number. The combination prediction method proposed in this article performs well on MSEI that integrates center and radius information, further demonstrating that the proposed combination prediction model can significantly improve prediction accuracy and is an effective prediction method.

Table 4. Error Indicators of Individual and Combination Prediction Methods

Prediction methods	MSEP	MSEL	MSEI
Exponential Smoothing	3.656	2.950	6.606
SVR	2.810	2.473	5.283
LSTM	3.158	0.239	3.397
Combined Forecasting Model	1.021	0.997	1.149
The Combination Prediction Model of Reference 4($\lambda = 4$)	1.324	1.219	2.543

The grey correlation degree between the actual number of intervals and the center and radius intervals of each prediction method can be obtained from Definition 6. The specific calculation results are shown in Table 5.

Table 5. Grey correlation between the actual number of intervals and the centers and radii of various predicted results

Prediction methods	Exponential Smoothing	SVR	LSTM	Combined Forecasting Model
γ	0.912	0.934	0.941	0.974

From Table 5, it can be seen that the improved grey correlation between the prediction results of interval number sequences and the actual interval number of the proposed combination prediction model is better than that of individual prediction results, indicating that the proposed combination prediction model is an optimal interval number combination prediction model.

4. Conclusion

By combining the improved grey correlation degree with the IOWHA operator, a new type of interval number combination prediction model is constructed, and it is applied to the prediction of interval number sequences in reference 4. To ensure the integrity of interval number sequences, interval number sequences are represented by centers and radii, interval number prediction accuracy is used as the inducing factor, and improved grey correlation degree is used as the optimality criterion to construct an interval type combination prediction model based on improved grey correlation degree and induced ordered harmonic weighted average (IOWHA) operator. Then, it was applied to the data in reference 4 for prediction, proving the effectiveness and superiority of the proposed combination prediction model.

Acknowledgments

This study was funded by Key Projects of Anhui Province's Higher Education Research Program (Project No: 2022 AH0 50565).

References

[1] Yuan Hongjun, Du Kang, Hu Lingyun. Interval type combination prediction model based on triangular fuzzy number correlation [J]. Statistics and Decision, 2019, 35 (6): 24-29.

[2] Du Kang, Yuan Hongjun, Zheng Yanan. Interval type combination prediction model based on triangular fuzzy numbers and GIOWA operator [J]. Statistics and Decision, 2019, 35 (16): 22-28.

[3] Zhang Na, Ge Leijiao. Short term photovoltaic output range combination prediction based on SOA optimization [J]. Journal of Solar Energy, 2021, 42 (5): 252-259.

[4] Yang Ye, Yuan Hongjun, Chen Qianru. A Variable Weight Interval Combination Prediction Model for a Class of Daily Extreme Temperature [J]. Journal of Datong University, Shanxi, 2023, 39 (4): 49-55.

[5] Yuan Hongjun, Wei Chenjunwa, Zhong Mei. Interval type combination prediction model based on the closeness of connection numbers and its effectiveness [J]. Forum on Statistics and Information, 2017, 32 (6): 31-37.

[6] Jiang Lihui, Chen Huayou, Ding Fangqing, et al. Interval combination prediction model based on IOWC-GOWA operator [J]. Computer Engineering and Applications, 2015, 51 (3): 50-54.

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

[8] Ding Xiaosong, Yang Guiyuan. Exchange rate combination prediction model based on GIOWHA operator [J]. Statistics and Decision, 2018, 34 (12): 9-13.