

## Application of F5 (0.000001) in Forecasting Tourist Quantity of Sanya

Kun Zhang<sup>1, a</sup>, HongXu Wang<sup>2, b, \*</sup>, HaiFeng Wang<sup>1, 3,</sup>, Zhuang Li<sup>1, 3,</sup>

<sup>1</sup>College of Computer Engineering, Hainan Tropical Ocean University, Sanya, Hainan, 572022, China

<sup>2</sup> College of China-Austria, Hainan Tropical Ocean University, Sanya, Hainan, 572022, China

<sup>3</sup> Sanya Key Laboratory of Computer Vision, Sanya, Hainan, 572022, China

<sup>a</sup>zk0588@163.com, <sup>b,\*</sup>Corresponding author: whx16233@126.com

**Keywords:** Fuzzy time series forecasting model; F5 (0.000001); inverse fuzzy number; percentage of percentage; tourist quantity

**Abstract.** The work proposed F5 (0.000001)—an element in the set of FTFSM 5 (Fuzzy Time Series Forecasting Model 5). It is a time series forecasting model established based on the variables of historical data, percentage of historical data, the percentage of percentage of historical data, and inverse fuzzy number. F5 (0.000001) was used to forecast historical data of tourism revenue in Sanya, with higher accuracy. Meanwhile, the data of unknown years were also predicted and studied in the work. F5 (0.000001) has become a new time series forecasting method.

### Introduction

Due to widespread application of time series forecasting methods and diversity of objective things, researchers have tried to expand the existing time series forecasting methods for years through various techniques. One of the aims is to seek new time series forecasting methods with smaller gap between the predicted and actual values. Song et al[1] used fuzzy technology to propose the first fuzzy time series forecasting model in 1993; meanwhile, they studied the classical case "forecasting registration quantity from 1971 to 1992 of University of Alabama". This is an important beginning. Saxena et al. [2] proposed high-accuracy forecasting model using the concept of fuzzy number [3] in 2012. In terms of classic cases, the average forecasting error rate of the model is AFER = 0.3406%, and the mean square error MSE=9169, demonstrating an unprecedented high level. These data are used to propose Definition 1 of fuzzy time series forecasting model with high accuracy. Wang et al[4-8] also developed several forecasting models with the basic variable of inverse fuzzy number. In the work, the forecasting models of Literature[4-8] were improved to simplify their expression forms of formulas. Thus, the work proposed the set of fuzzy time series forecasting models-FTFSM 5 (Fuzzy Time Series Forecasting Model 5). One element of the collection F5(0.000001) was called common model. F5(0.000001) is a fuzzy time series forecasting model with high accuracy — AFER=0.0251% and MSE=167—while using F5(0.000001) to study classical cases

### FTFSM 5—set of forecasting models

**Construction of forecasting model set—FTFSM 5.** The formulas of forecasting models proposed in Literature [4-8] are improved to establish the forecasting formula F5 ( $\mu$ ):

$$S_i = T_{i-1} \times [1 + (1 + Z_i)P_{i-1}], \quad Z_i = \frac{\mu + 1}{\frac{\mu}{P_{i-1}} + \frac{1}{Q_i}}, \quad (\mu \in (0, 1)), \quad (1)$$

where  $S_i$  is the predicted value of historical data for Year  $i$ ;  $T_{i-1}$  the historical data of Year  $i-1$ ;  $Z_i$  the inverse fuzzy number of Year  $i$ ;  $P_{i-1}$  the percentage of historical data for Year  $i-1$ ;  $Q_i$  the percentage of percentage of historical data for Year  $i$ ;  $\mu$  the membership (of Element  $P_{i-1}$ ). Forecasting Formula (3) is denoted as F5 ( $\mu$ ).

If  $\mu \in (0,1)$  is determined, we can obtain a forecasting model based on Formula (1). Its application process is similar to Section 2.2 as follows. The set consisting of all forecast models of  $F5(\mu)$  is denoted as FTFSM 5. General elements of FTFSM 5 are the set of fuzzy time series forecasting models with  $F5(\mu)$  ( $\mu \in (0,1)$ ) and Formula (1).

**Definition 1** Fuzzy time series forecasting model is called high-accuracy fuzzy time series forecasting model. When the model is used to forecast registration quantity of University of Alabama from 1971 to 1992, it is obtained that  $AFER \leq 0.3406\%$  and  $MSE \leq 9169$ .

When  $F5(0.000001)$  is applied to forecast classical cases, it is obtained that  $AFER=0.0251\% < 0.3406\%$  and  $MSE=167 < 9169$ . Therefore,  $F5(0.000001)$  is a high-accuracy fuzzy time series forecasting model, called common model of FTFSM 5.

**Application process of Forecasting Model  $F5(\mu)$ .** The application process of  $F5(\mu)$  includes following steps:

Step 1: Inputting historical data; Step 2: Establishing 3 discourse domains of historical data; Step 3: Building forecasting formula  $F5(\mu)$ ; Step 4: Applying  $F5(\mu)$  to forecast historical data; Step 5: Using  $F5(\mu)$  to forecast the data of unknown years.

### Application of $F5(0.000001)$ in forecasting tourist quantity of Sanya

In this section,  $F5(0.000001)$  was applied to study tourist quantity of Sanya from 2006 to 2014, expecting to describe the whole application process of  $F5(0.000001)$ .

**Inputting historical data.** Table 1 shows historical data of Sanya from 2006 to 2014.

Table 1 Application of  $F5(0.000001)$  in forecasting Sanya's tourist quantity in 2006-2014

Year	Tourist quantity $T_i$ (Ten thousand)	Percentage $P_i$	Percentage of percentage $Q_i$	Predicted value $S_i$ (Ten thousand)	$(S_i - T_i)^2$	$ S_i - T_i /T_i$
2006	454.90	-	-	-	-	-
2007	538.43	0.1836	-	-	-	-
2008	604.15	0.1221	-0.3350	604.17	0.0004	0.000033
2009	669.05	0.1074	-0.1204	669.04	0.0001	0.000015
2010	882.65	0.3193	1.9730	882.68	0.0009	0.000034
2011	1021.07	0.1568	-0.5089	1021.06	0.0001	0.000009
2012	1102.22	0.0795	-0.4930	1102.24	0.0004	0.000018
2013	1228.40	0.1145	0.4403	1228.43	0.0009	0.000024
2014	1352.76	0.1012	-0.1162	1352.71	0.0025	0.000037
AFER						0.0024%
MSE					0.000757	

**Constructing 3 discourse domains.** Discourse domains of Sanya's tourist quantity can be constructed based on Table 1:

$$T = \{T_{2006}=454.90, T_{2007}=538.43, \dots, T_{2013}=1228.40, T_{2014}=1352.76\}.$$

According to Formula  $P_i = (T_{i-1} - T_{i-2})/T_{i-1}$ , the percentage domain of Sanya's tourist quantity for every year can be constructed as:

$$P = \{P_{2007}=0.1836, W_{2008}=0.1221, \dots, W_{2013}=0.1145, W_{2014}=0.1012\}.$$

According to Formula  $Q_i = (P_i - P_{i-1})/P_{i-1}$ , the domain for percentage of percentage can be constructed as:

$$Q = \{Q_{2008} = -0.3350, Q_{2009} = -0.1204, \dots, Q_{2013} = 0.4403, U_{2014} = -0.1162\}.$$

**Establishing forecasting formula of  $F5(0.000001)$ .**  $F5(0.000001)$  is an element of FTFSM 5, with the forecasting formula:

$$S_i = T_{i-1} \times [1 + (1 + Z_i)P_{i-1}], \quad Z_i = \frac{0.000001 + 1}{\frac{0.000001}{P_{i-1}} + \frac{1}{Q_i}}, \quad (2)$$

where  $S_i$  is the predicted value of tourist quantity for Year  $i$ ;  $T_{i-1}$  the tourist quantity of Year  $i-1$ ;  $Z_i$  the inverse fuzzy number of Year  $i$ ;  $P_{i-1}$  the percentage of tourist quantity for Year  $i-1$ ;  $Q_i$  the percentage of percentage of tourist quantity for Year  $i$ ; 0.000001 the membership.

**Using F5(0.000001) to Forecast tourist quantity of Sanya from 2008 to 2014.** The predicted values of Sanya's tourist quantity from 2008 to 2014 are calculated in Table 1.

**Using F5(0.000001) to Forecast tourist quantity of Sanya from 2014 to 2015.** The predicted value of tourist quantity for 2015 is only studied from perspective of historical data. Formula(2) cannot be directly used to calculate the predicted value of tourist quantity for 2015, due to the lack of  $Q_{2015}$  (percentage of percentage of tourist quantity). Therefore, it is necessary to construct rules of using F5 (0.000001) to forecast tourist quantity of unknown years.

**(1)F5 (3-3-1-4) Forecasting Rule.** It is supposed that the percentage of percentage of tourist quantity for the first three years ( $h-3, h-2, h-1$ ) are  $Q_{h-3}, Q_{h-2}$  and  $Q_{h-1}$ , respectively. Then, it is calculated that  $\tau = \{\min\{Q_{h-3}, Q_{h-2}, Q_{h-1}\}, \min\{Q_{h-3}, Q_{h-2}, Q_{h-1}\}\}$ . According to Formula(2) of F5(0.000001), calculation results are obtained with the parameters  $P_{h-4}$  and  $Q_1 = \min\{Q_{h-3}, Q_{h-2}, Q_{h-1}\}$ ;  $P_{h-3}$  and  $Q_{h-3}, Q_{h-2}, Q_{h-1} + \tau$ ;  $P_{h-2}$  and  $Q_3 = \min\{Q_{h-3}, Q_{h-2}, Q_{h-1}\} + 2\tau$ ;  $P_{h-1}$  and  $Q_4 = \max\{Q_{h-3}, Q_{h-2}, Q_{h-1}\}$ ; and  $T_h$ . The calculated values, ranged in ascending orders, are small predicted value, smaller predicted value, larger predicted value and large predicted value of tourist quantity for the unknown year  $h$ , respectively. This forecasting rule is called F5(3-3-1-4) Forecasting Rule.

**(2)Decision scheme of predicted values.** The first decision scheme is to determine the recommend ranks of small, smaller, larger and large predicted values for 2015 according to the predicted values of 2014, which were sorted by forecasting error rate with real values. The second decision scheme is to determine the sorting method by decision-makers.

**(3)Using F5 (3-3-1-4) to calculate predicted value of unknown years**

According to Table 1,  $\tau = \{\max\{Q_{2011}, Q_{2012}, Q_{2013}\}, \min\{Q_{2011}, Q_{2012}, Q_{2013}\}\} / 3 = 0.3164$ . According to Formula(2) of F5(0.000001), calculation results are obtained with the parameters  $P_{2013} = 0.1145$  and  $Q_1 = \min\{Q_{2011}, Q_{2012}, Q_{2013}\} = -0.5089$ ;  $P_{2013} = 0.1145$  and  $Q_2 = \min\{Q_{2011}, Q_{2012}, Q_{2013}\} + \tau = -0.1925$ ;  $P_{2013} = 0.1145$  and  $Q_3 = \min\{Q_{2011}, Q_{2012}, Q_{2013}\} + 2\tau = 0.1239$ ;  $P_{2013} = 0.1145$  and  $Q_4 = \max\{Q_{2011}, Q_{2012}, Q_{2013}\} = 0.4403$ ; and  $T_{2013} = 1228.40$ . (See Table 2).

Table 2 Application of F5(3-3-1-4) in forecasting Sanya's tourist quantity of 2014

Year	Forecasting Type	Real data (2014) (ten thousand)	Predicted data (2014) (ten thousand)	$ S_{2014} - T_{2014}  / T_{2014}$	Rank of forecasting error
2014	Small	1297.47	1352.76	4.0872%	3
2014	Smaller	1341.98	1352.76	0.7969%	1
2014	Larger	1386.48	1352.76	2.4927%	2
2014	Large	1430.98	1352.76	5.7823%	4

Table 2 shows each predicted value of tourist quantity in 2014, "forecasting error with real data" and their ranks in ascending order. No.1 is smaller predicted value 1341.98, with prediction error rate of 0.7969%; No.2 is larger predictive value 1386.48, with prediction error rate of 2.4927%; No.3 is small predicted value 1297.47, with error rate of 4.0872%; and No.4 is large predicted value 1430.88, with error rate of 5.7823%.

**Using F5 (3-3-1-4) Rule to forecast Sanya's tourist quantity of the unknown year 2015.** According to Table 1,  $\tau = \{\max\{Q_{2012}, Q_{2013}, Q_{2014}\}, \min\{Q_{2012}, Q_{2013}, Q_{2014}\}\} / 3 = 0.3111$ . According to Formula (2) of F5(0.000001), calculation results are obtained with the parameters  $P_{2014} = 0.1012$  and  $Q_1 = \min\{Q_{2012}, Q_{2013}, Q_{2014}\} = -0.4930$ ;  $P_{2014} = 0.1012$  and  $Q_2 = \min\{Q_{2012}, Q_{2013}, Q_{2014}\} + \tau = -0.1819$ ;  $P_{2014} = -0.1012$  and  $Q_3 = \min\{Q_{2012}, Q_{2013}, Q_{2014}\} + 2\tau = 0.1292$ ;  $P_{2014} = 0.1012$  and  $Q_4 = \max\{Q_{2012}, Q_{2013}, Q_{2014}\} = 0.4403$ ; and  $T_{2014} = 1352.76$ . (See Table 3).

Table 3 Application of F5(3-3-1-4) in forecasting Sanya's tourist quantity of 2015

Year	Forecasting Type	Real data (2014) (ten thousand)	Predicted data (2015) (ten thousand)	$ A_{2015} - B_{2014}  / B_{2014}$	Recommended rank
2014	Small	1352.76	1422.17	5.1310%	3
2014	Smaller	1352.76	1464.76	8.2794%	1
2014	Larger	1352.76	1507.35	11.4277%	2
2014	Large	1352.76	1549.94	14.5761%	4

According to the recommended ranks in Table 3, the first is smaller predicted value of 1464.76, with an increase of 8.2794% than that of 2014; followed by larger predicted value 1507.35, with increase of 11.4277% than that of 2014; the third is small predicted value 1422.17, with growth of 5.1310% than that of 2014; the fourth is large predicted value 1549.94, with 14.5761% than that of 2014. The ranks can also be determined according to the second decision scheme—decision makers determine the ranks based on their years of work experience.

### Summary

F5 (0.000001) is a common model in the set of forecasting models FTSFM 5, with high forecasting accuracy. In forecasting the data of unknown years, four predicted values can be calculated according to F5 (3-3-1-4) Rule. There are two decision schemes available for decision makers, where the first one is to obtain the recommended sorting numbers based on the sorting of predicted values in 2014. The smaller predicted value of tourist quantity in 2014 ranks first in terms of prediction accuracy (0.7696%). Thus, the smaller predicted value 1464.76 of 2015 that ranks first is expected to have high accuracy. The second decision scheme helps to make use of the experience of decision-makers. F5(0.000001) has increased a method for time series forecasting.

### Acknowledgements

This work was financially supported by the key project of scientific research of Hainan Province (No.Hnky2015ZD-14), the Hainan Planning projects of philosophy and Social Sciences (No.HNSK (QN)15-60), the scientific and technological cooperative project for college and region of Sanya (2015YD16, 2015YD43, 2014YD30), the key Laboratory of Sanya Project (No.L1410).

### References

- [1] Q Song, B S Chissom. *Forecasting enrollments with fuzzy time series*. Fuzzy Sets and Systems, vol.54(1993), pp.269-277.
- [2] Preetika Saxena, Kalyani Sharma, Santhosh Easo. *Forecasting enrollments based on fuzzy time series with higher forecast accuracy rate*. Int. J. Computer Technology & Applications, Vol.3, No.3(2012), pp: 957-961.
- [3] T A Jilani, S M A Burney, C Ardil. *Fuzzy metric approach for fuzzy time series forecasting based on frequency density based partitioning*. Proceedings of World Academy of Science, Engineering and Technology, Vol. 34(2007), pp: 333-338.
- [4] Hongxu Wang, Jianchun Guo, Hao Feng, and Hailong Jin. *A new forecasting model of fuzzy time series*. Applied Mechanics and Materials, Vol .678(2014), pp.59-63.
- [5] Hongxu Wang, Jianchun Guo, Hao Feng, and Hailong Jin. *An improved forecasting model of fuzzy time series*. Applied Mechanics and Materials, Vol.678(2014), pp.64-69.
- [6] H X Wang, J C Guo, H Feng, F J Zhang. *A new model of forecast enrollment using fuzzy time series*. Education Management and Management Science, 2014 International Conference on Education Management and Management Science (ICEMMS 2014), 7-8 August, (2014), pp: 95-98.

- [7] Wang H X, Guo J C, Feng H, et al. *A fuzzy time series forecasting model based on data differences*[J]. *Frontiers in Computer Education*, (2014),pp.15-18.
- [8] HongXu Wang, JianChun Guo, Hao Feng, HaiLong Jin. *A fuzzy time series forecasting model based on percentages*. *Frontiers in Computer Education*. (2014),pp:11-14.