The Set of Improved Fuzzy Time Series Forecasting Models Based on the Ordered Difference Rate

Chengguo Yin[1], Hongxu Wang[2,*], Hao Feng[3] and Xiaoli Lu[2]
[1]School of Ocean Information Engineering, Hainan Tropical Ocean University, Sanya, Hainan, 572022, China
[2]School of Ocean Business, Hainan Tropical Ocean University, Sanya, Hainan, 572022, China
[3]School of Ocean Science and Technology, Hainan Tropical Ocean University, Sanya, Hainan, 572022, China
[*]Corresponding author

Abstract—**Song and Chissom first proposed the fuzzy time series forecasting model in 1993. In this paper, we improved the forecasting model proposed by Stevenson and Porter, and dug out the SIFBODR (The Set of Improved Fuzzy Time Series Forecasting Models Based on the Ordered Difference Rate). In the research on the forecasting problem of enrollments of the University of Alabama 1971–1992, the forecasting model SIFBODR(0.00002, 0.00004) of SIFBODR can obtain AFER (Average Forecasting Error Rate) = 0% and MSE(Mean Square Error) = 0. The problem that the prediction accuracy of fuzzy time series forecasting models is not high for many years is basically solved.**

Keywords-fuzzy time series forecasting method; fuzzy number function of SIFBODR; inverse fuzzy number function of SIFBODR; forecasting function of SIFBODR

I.   INTRODUCTION

Zadeh [1] established the fuzzy set theory in 1965, which had become an indispensable important theory in the study of uncertainty problem. Song and Chissom [2–4] proposed the fuzzy time series forecasting model by using the fuzzy set theory in 1993, and discussed the forecasting problem of enrollments of the University of Alabama 1971–1992. A large number of fuzzy time series forecasting models have been proposed, such as [5–17], but the problem of low prediction accuracy still perplexes people. Stevenson and Porter [17] improved the forecasting model of Jilani, Burney, and Ardi [8], and proposed a new forecasting model based on the defuzzification technology. They used the forecasting model to predict the enrollments of the University of Alabama, and obtained the AFER(Average Forecasting Error Rate) = 0.57% and MSE(Mean Square Error) = 21575 of predicted values of the enrollments, which were unprecedented high prediction accuracy. The reason is that this technology provides a key idea to improve the prediction accuracy. In this paper, we further develop their research achievements, and introduce the concepts of fuzzy number function, inverse fuzzy number function and forecasting function, and then propose the SIFBODR(The Set of Improved Fuzzy Time Series Forecasting Models Based on the Ordered Difference Rate). In the discussion of the forecasting problem of enrollments of the University of Alabama 1971–1992, the forecasting models SIFBODR(0.00002, 0.00004) and SIFBODR(0.00004,0.00002) of SIFBODR can obtain AFER = 0% and MSE = 0. We basically solve the problem that the prediction accuracy of

fuzzy time series forecasting models is not high, which has perplexes people for many years.

II.   THE SET OF IMPROVED FUZZY TIME SERIES FORECASTING MODELS BASED ON THE ORDERED DIFFERENCE RATE

In this paper, we continue to use the relevant concepts which are used by Stevenson and Porter [17]. If there is a time series forecasting problem, the universe of discourse for historical data is $M = \{M_1, M_2, \ldots, M_n\}$. The formula for the year to year difference rate of historical data is $P_q = (M_q - M_{q-1})/M_{q-1}$. If the universe of discourse for the difference rate of historical data is recorded as $P = \{P_2, P_3, \ldots, P_n\}$. Each data in $P = \{P_2, P_3, \ldots, P_n\}$ is arranged from small to large, then they constitute a new set $p = \{p_1, p_2, \ldots, p_{n-1}\}$; the set p is called the universe of discourse for the ordered difference rate of historical data.

Definition 1. If there is a time series forecasting problem, the universe of discourse for historical data is $M = \{M_1, M_2, \ldots, M_n\}$, the universe of discourse for the year to year difference rate is $P = \{P_2, P_3, \ldots, P_n\}$, and the universe of discourse for the ordered difference rate is $p = \{p_1, p_2, \ldots, p_{n-1}\}$. For each selected $q \in \{1, 2, \ldots, n-1\}$, the fuzzy number function $S_q(\mu_1, \mu_2)$, the inverse fuzzy number function $R_q(\mu_1, \mu_2)$, and the forecasting function $Q_q(\mu_1, \mu_2)$ are respectively defined on set p as

$$R_q(\mu_1, \mu_2) = \begin{cases} \dfrac{1 + \mu_2}{\dfrac{1}{p_1} + \dfrac{\mu_2}{p_2}}, & \text{if } q = 1, \\[4mm] \dfrac{\mu_1 + 1}{\dfrac{\mu_1}{p_{q-1}} + \dfrac{1}{p_q}}, & \text{if } 2 \leq q \leq n-1. \end{cases} \tag{1}$$

$$Q_q(\mu_1, \mu_2) = M_{q-1}\left(1 + R_q(\mu_1, \mu_2)\right), \tag{2}$$

$$S_q(\mu_1,\mu_2)=\begin{cases}\dfrac{1}{p_1}+\dfrac{\mu_2}{p_2}, & if \quad q=1\\[2mm]\dfrac{\mu_1}{p_{q-1}}+\dfrac{1}{p_q}, & if \; 2\le q\le n-1.\end{cases} \tag{3}$$

where the independent variables $\mu_1\in[0,1)$ and $\mu_2\in[0,1)$ are also called the membership degree. When $\mu_1=\mu_2=\mu$, the fuzzy number function of SIFBODR is recorded as $S_q(\mu)$(q=1, 2, …, n-1); the inverse fuzzy number function of SIFBODR is recorded as $R_q(\mu)$(q=1, 2, …, n-1); the forecasting function of SIFBODR is recorded as $Q_q(\mu)$(q=1, 2, …, n-1). When μ_1 and μ_2 fetch the specific membership degree in [0,1), then $Q_q(\mu_1,\mu_2)$ is called the forecasting formula of SIFBODR, the forecasting formula can be also recorded as SIFBODR(μ_1,μ_2). If the ordered difference rate p_q(q $\in$ {1, 2, …, n-1}) corresponds to the difference rate p_c(c $\in$ {2, 3, …, n}), then $Q_q(\mu_1,\mu_2)$ represents the predicted value of year c.

Once the membership degree μ_1 and μ_2 fetch specific values in [0,1), a forecasting formula SIFBODR(μ_1,μ_2) can be established by (3). In the research on the time series forecasting problem, the application steps are as follows:

(1). Write out the historical data table for the time series forecasting problem;

(2). Write out the universe of discourse M for the historical data, the universe of discourse P for the year to year difference rate and the universe of discourse p for the ordered difference rate;

(3). Write out the forecasting formula SIFBODR(μ_1,μ_2);

(4). Write out the predicted values of the historical data by using SIFBODR(μ_1,μ_2) to calculate.

It is clear that the forecasting formula SIFBODR(μ_1,μ_2) is exactly the forecasting model SIFBODR(μ_1,μ_2). The technical route we use is: the fuzzy number function $S_q(\mu_1,\mu_2)$ → the inverse fuzzy number function $R_q(\mu_1,\mu_2)$ → the forecasting model SIFBODR(μ_1,μ_2)(or the forecasting formula SIFBODR(μ_1,μ_2)). Since the forecasting model SIFBODR is obtained from the fuzzy set theory, SIFBODR is called fuzzy time series forecasting model, but in fact, it is the set of time series forecasting models.

Definition 2. If there is a time series forecasting problem, the universe of discourse for historical data is M = {M₁, M₂, …, Mₙ}, the universe of discourse for the year to year difference rate is P = {P₂, P₃, …, Pₙ}, and the universe of discourse for the ordered difference rate is p = {p₁, p₂, …, pₙ₋₁}. When the membership degree μ_1 and μ_2 fetch each value in the semi-open and semi-closed interval [0,1), infinite fuzzy time series forecasting models SIFBODR(μ_1,μ_2) can be obtained. The fuzzy time series forecasting models SIFBODR(μ_1,μ_2) are taken as elements, then the set they constitute is called

SIFTSFMBODR(The Set of Improved Fuzzy Time Series Forecasting Models Based on the Ordered Difference Rate), the abbreviation is simplified as SIFBODR. Its general element is SIFBODR(μ_1,μ_2). SIFBODR(μ_1,μ_2) represents the forecasting formula of a fuzzy time series as well as a fuzzy time series forecasting model taking SIFBODR(μ_1,μ_2) as the forecasting formula.

Theorem. If there is a time series forecasting problem, the universe of discourse for historical data is M = {M₁, M₂, …, Mₙ}, the universe of discourse for the year to year difference rate is P = {P₂, P₃, …, Pₙ}, and the universe of discourse for the ordered difference rate is p = {p₁, p₂, …, pₙ₋₁}. For each selected q $\in$ {1, 2, …, n-1}, if the ordered difference rate p_q(q $\in$ {1, 2, …, n-1}) corresponds to the difference rate p_c(c $\in$ {2, 3, …, n}), then

(1). The fuzzy number function $S_q(\mu_1,\mu_2)$, inverse fuzzy number function $R_q(\mu_1,\mu_2)$, and forecasting function $Q_q(\mu_1,\mu_2)$ of SIFBODR are all continuous functions;

(2). When $\mu_1\to 0$ and $\mu_2\to 0$, the inverse fuzzy number function $R_q(\mu_1,\mu_2)$ converges to the difference rate P_c of year c:

$$\lim_{\mu_1\to 0,\mu_2\to 0} R_q(\mu_1,\mu_2)=P_c;$$

(3). When $\mu_1\to 0$ and $\mu_2\to 0$, then the forecasting function $Q_q(\mu_1,\mu_2)$ converges to the historical data M_c of year c;

(4). When the membership degree μ_1 and μ_2 are small enough, the forecasting function value $Q_q(\mu_1,\mu_2)$ of SIFBODR is equal to the historical data M_c of year c.

In the research on the forecasting problem of enrollments of the University of Alabama 1971–1992, this paper simply lists the comparisons between partial forecasting models and other different prediction methods by using the defuzzification technology, as shown in Table I and II. Table II also gives the prediction results by using SIFBODR(0.00002, 0.00004) and SIFBODR(0.00004, 0.00002) proposed in this paper to forecast the enrollments of the University of Alabama 1971–1992. The MSE = 0 and AFER = 0.0%, which are the highest prediction accuracy, and also verify the correctness of the partial conclusions of the theorem.

In Table I and II: the formulas of MSE and AFER are:

$$MSE=\frac{1}{n}\sum_{c=1}^{n}(M_c-Q_c)^2;$$

$$AFER=\left(\frac{1}{n}\sum_{c=1}^{n}|M_c-Q_c|/M_c\right)\times100\%.$$

TABLE I. COMPARISONS OF DIFFERENT FORECASTING MODELS

Year	Enrollments	Feng, Guo, Wang, Zhang.[16]	Saxena, Sharma, Easo. [13]	Wang, Guo, Feng, Jin.[9]	Wang, Guo, Wang, Feng.[12]	Wang, Guo, Feng, Zhang.[10]	Stevenson, Porter. [17]
1971	13055	-	-	-	-	-	-
1972	13563	13563	13486	-	13563	-	13410
1973	13867	13867	13896	13809	13867	13867	13932
1974	14696	14696	14698	14610	14695	14691	14664
1975	15460	15461	15454	15422	15460	15461	15423
1976	15311	15312	15595	15299	15311	15311	15847
1977	15603	15604	15600	15642	15603	15607	15580
1978	15861	15860	15844	15901	15861	15861	15877
1979	16807	16804	16811	16782	16796	16797	16773
1980	16919	16920	16916	16935	16919	16920	16897
1981	16388	16387	16425	16328	16388	16375	16341
1982	15433	15430	15657	15362	15432	15436	15671
1983	15497	15496	15480	15496	15497	15497	15507
1984	15145	15143	15214	15077	15144	15136	15200
1985	15163	15163	15184	15274	15163	15163	15218
1986	15984	15976	15995	15966	15982	15861	16035
1987	16859	16858	16861	16849	16859	16858	16903
1988	18150	18150	17965	18312	18150	18154	17953
1989	18970	18974	18964	18974	18970	18976	18879
1990	19328	19326	19329	19236	19327	19329	19303
1991	19337	19338	19378	19299	19337	19338	19432
1992	18876	18872	18984	18951	18874	18766	18966
AFER		0.0099%	0.3406%	0.273%	0.0026%	0.085%	0.57%
MSE		7	9169	2912	1	1384	21575

TABLE II. COMPARISONS OF DIFFERENT FORECASTING MODELS

Year	Enrollments	Wang, Guo, Feng, Jin.[11]	Wang, Guo, Feng, Jin.[14]	Jilani, Burney, Ardil. [8]	Wang, Guo, Feng, Jin.[15]	SIFBODR (0.00002,0.00004)	SIFBODR (0.00004,0.00002)
1971	13055	-	-	13579	-	-	-
1972	13563	-	-	13798	-	13563	13563
1973	13867	13813	13845	13798	13745	13867	13867
1974	14696	14681	14729	14452	14531	14696	14696
1975	15460	15525	15412	15373	15575	15460	15460
1976	15311	15189	15317	15373	15446	15311	15311
1977	15603	15685	15620	15623	15555	15603	15603
1978	15861	15895	15895	15883	15901	15861	15861
1979	16807	16878	16786	17079	16933	16807	16807
1980	16919	16839	16961	17079	16950	16919	16919
1981	16388	16505	16334	16497	16601	16388	16388
1982	15433	15349	15461	15737	15456	15433	15433
1983	15497	15511	15497	15737	15544	15497	15497
1984	15145	15026	15094	15024	15165	15145	15145
1985	15163	15051	15133	15024	15187	15163	15163
1986	15984	15980	15972	15883	15953	15984	15984
1987	16859	16805	16805	17079	16849	16859	16859
1988	18150	18246	18183	17991	18211	18150	18150
1989	18970	18926	18990	18802	19077	18970	18970
1990	19328	19275	19338	18994	19344	19328	19328
1991	19337	19428	19346	18994	19200	19337	19337
1992	18876	19046	18822	18916	18851	18876	18876
AFER		0.442%	0.171%	1.02%	0.462%	0.0%	0.0%
MSE		6825	1121	41426	8963	0	0

III. CONCLUSIONS

In the research on the forecasting problem of enrollments of the University of Alabama 1971–1992, the forecasting models SIFBODR(0.00002, 0.00004) and SIFBODR(0.00004, 0.00002) given in this paper obtain MSE=0 and AFER=0.0%(See Table II). The history of low prediction accuracy of fuzzy time series forecasting models is terminated. Because the forecasting models of SIFBODR can be used for time series analysis, the

following work needs to find the direct practical application of this kind of forecasting models.

REFERENCES

[1] L. A. Zadeh, "Fuzzy set," Fuzzy Sets and Systems. vol. 8, pp. 338-353, 1965.

[2] Q. Song and B. S. Chissom, "Forecasting enrollments with fuzzy time series—Part II," Fuzzy Sets and Systems. vol. 62, pp. 1–8, 1994.

[3] Q. Song and B. S. Chissom, "Forecasting enrollments with fuzzy time series—Part I," Fuzzy Sets and Systems. vol. 54, pp. 1–9, 1993.

[4] Q. Song and B. S. Chissom, "Fuzzy series and its models," Fuzzy Sets and Systems. vol. 54, pp. 269–277, 1993.

[5] Wang Hongxu and Wu Zhenxing, "Preliminary Theory of Set DR of Fuzzy Time Series Forecasting Model," 2016 International Conference on Mathematical, Computational and Statistical Sciences and Engineering (MCSSE2016). pp. 256–260, 2016.

[6] Wang Hongxu and Wu Zhenxing, "Preliminary Theory of Set SDR of Fuzzy Time Series Forecasting Model," 2016 International Conference on Mathematical, Computational and Statistical Sciences and Engineering (MCSSE2016). pp. 261–265, 2016.

[7] Tahseen A. Jilani, S. M. Aqil Burney, and C. Ardil, "Multivariate high order fuzzy time series forecasting for car road accidents," World Academy of Science, Engineering and Technology. vol. 1, pp. 288–293, 2007.

[8] T. A. Jilani, S. M. A. Burney, and 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, pp. 333–338, 2007.

[9] Wang Hongxu, Guo Jianchun, Feng Hao, and Jin Hailong, "A new forecasting model of fuzzy time series," Applied Mechanics and Materials. vol. 678, pp. 59–63, 2014 [2014 3rd International Conference on Mechatronics and Control Engineering (ICMCE 2014). 2014].

[10] H. X. Wang, J. C. Guo, H. Feng, and F. J. Zhang, "A new model of forecast enrollment using fuzzy time series," 2014 International Conference on Education Management and Management Science (ICEMMS 2014). Tianjin, China, pp. 95–98, 2014.

[11] HongXu Wang, JianChun Guo, Hao Feng, and HaiLong Jin, "A fuzzy time series forecasting model based on data differences," 1 ICT IN EDUCATION. Frontiers in Computer Education. pp. 15–18, 2014 [2nd International Conference on Frontiers in Computer Education (ICFCE2014). Wuhan, China, 2014].

[12] Hongxu Wang, Jianchun Guo, Hui Wang, and Hao Feng*, "A fuzzy time series forecasting model based on yearly difference of the student enrollment number," 2014 2nd International Conference on Soft Computing in Information Communication Technology (SCICT2014). The Authors– Published by Atlantis Press, pp. 172–175, 2014.

[13] Preetika Saxena, Kalyani Sharma, and Santhosh Easo, "Forecasting enrollments based on fuzzy time series with higher forecast accuracy rate," Int. J. Computer Technology & Applications. vol. 3 (3), pp. 957–961, 2012.

[14] Wang Hongxu, Guo Jianchun, Feng Hao, and Jin Hailong, "An improved forecasting model of fuzzy time series," Applied Mechanics and Materials. vol. 678, pp. 64–69, 2014 [2014 3rd International Conference on Mechatronics and Control Engineering (ICMCE 2014). 2014].

[15] HongXu Wang, JianChun Guo, Hao Feng, and HaiLong Jin, "A fuzzy time series forecasting model based on percentages," 1 ICT IN EDUCATION.Frontiers in Computer Education. pp. 11–14, 2014 [2nd International Conference on Frontiers in Computer Education (ICFCE2014). Wuhan, China, 2014].

[16] Hao Feng, Jianchun Guo, Hongxu Wang*, and Fujin Zhang, "A modified method of foreecasting enrollments based on fuzzy time series," 2014 2nd International Conference on Soft Computing in Information Communication Technology (SCICT2014). The Authors– Published by Atlantis Press, pp. 176–179, 2014.

[17] Meredith Stevenson and John E. Porter, "Fuzzy time series forecasting using percentage change as the universe of discourse," Proceedings of World Academy of Science, Engineering and Technology. vol. 55, pp. 154–157, 2009.