The Set of 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 autho

***Abstract*—Song and Chissom established fuzzy time series forecasting model in 1993. Stevenson and Porter improved the forecasting model of Jilani, Burney, and Ardil in 2009, and researched the forecasting problem of enrollments of the University of Alabama 1971–1992. Although they obtained the best prediction accuracy by 2009, the prediction accuracy was still not ideal. In this paper, we improved the forecasting model of Stevenson and Porter, and got the SFBODR (The Set of Fuzzy Time Series Forecasting Models Based on the Ordered Difference Rate). The forecasting model SFBODR(0.00004, 0.00003) can get the ideal state of AFER(Average Forecasting Error Rate) = 0% and MSE(Mean Square Error) = 0 in forecasting the enrollments of the University of Alabama.**

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

I. INTRODUCTION

Song and Chissom [1–3] established the fuzzy time series forecasting model in 1993, and first researched the forecasting problem of enrollments of the University of Alabama 1971–1992. A great deal of fuzzy time series forecasting models emerge as the times require, but the prediction accuracy is still not high enough. These forecasting models in [6–15] are all derived from the defuzzification technology proposed by Jilani, Burney, and Ardil [14,16]. Stevenson and Porter [11] highly improved the prediction accuracy by using the improved defuzzification technology in 2009. When studying the forecasting problem of enrollments of the University of Alabama 1971–1992, they obtained the best prediction accuracy by 2009 (See Table II). The defuzzification technology sets a precedent for exploring fuzzy time series forecasting models with high prediction accuracy. In this paper, we further extend their research achievements, and propose the SFBODR(The Set of Fuzzy Time Series Forecasting Models Based on the Ordered Difference Rate). In the study of the forecasting problem of enrollments of the University of Alabama 1971–1992, the AFER(Average Forecasting Error Rate) and MSE(Mean Square Error) of predicted values of enrollments of the forecasting models SFBODR(0.00003, 0.00004) and SFBODR(0.00004,0.00003) are both zero.

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

This paper continues to use the relevant concepts which are used by Stevenson and Porter [11]. Assuming there is a time series forecasting problem, let $E = \{E_1, E_2, \ldots, E_n\}$ be the universe of discourse for historical data. The formula for calculating the year to year difference rate of historical data is $G_p = (E_p - E_{p-1}) / E_{p-1}$. The universe of discourse for the difference rate of historical data is recorded as $G = \{G_2, G_3, \ldots, G_n\}$. If we arrange each data in $G = \{G_2, G_3, \ldots, G_n\}$ from small to large, then constitute a new set $g = \{g_1, g_2, \ldots, g_{n-1}\}$, which is called the universe of discourse for the ordered difference rate of historical data.

Definition 1. Assuming there is a time series forecasting problem, let $E = \{E_1, E_2, \ldots, E_n\}$ be the universe of discourse for historical data, let $G = \{G_2, G_3, \ldots, G_n\}$ be the universe of discourse for the difference rate of historical data, and let $g = \{g_1, g_2, \ldots, g_{n-1}\}$ be the universe of discourse for the ordered difference rate. The fuzzy number function $J_p(\mu_1, \mu_2)$ is defined on set g as

$$J_p(\mu_1,\mu_2)=\begin{cases}\dfrac{1}{g_1}+\dfrac{\mu_2}{g_2}, & if \quad p=1\\ \dfrac{\mu_1}{g_{p-1}}+\dfrac{1}{g_p}+\dfrac{\mu_2}{g_{p+1}}, & if\ 2\le p\le n-2,\\ \dfrac{\mu_1}{g_{n-1}}+\dfrac{1}{g_n}, & if \quad p=n-1.\end{cases} \tag{1}$$

Where the independent variables $\mu_1 \in [0,1)$ and $\mu_2 \in [0,1)$ are also called the membership degree of $J_p(\mu_1, \mu_2)$. When $\mu_1 = \mu_2 = \mu$, the fuzzy number function of SFBODR is recorded as $J_p(\mu)(p=1, 2, \ldots, n-1)$.

Definition 2. Assuming there is a time series forecasting problem, let $E = \{E_1, E_2, \ldots, E_n\}$ be the universe of discourse for historical data, let $G = \{G_2, G_3, \ldots, G_n\}$ be the universe of discourse for the difference rate of historical data, and let $g = \{g_1, g_2, \ldots, g_{n-1}\}$ be the universe of discourse for the ordered difference rate. If the ordered difference rate $g_p(p \in \{1, 2, \ldots, n-1\})$ corresponds to the difference rate $G_b(b \in \{2, 3, \ldots, n\})$, then the inverse fuzzy number function $I_p(\mu_1, \mu_2)$ of the fuzzy number function $J_p(\mu_1, \mu_2)$ is defined on set g as

$$I_p(\mu_1,\mu_2)=\begin{cases}\dfrac{1+\mu_2}{\dfrac{1}{g_1}+\dfrac{\mu_2}{g_2}}, & if\ p=1,\\ \dfrac{\mu_1+1+\mu_2}{\dfrac{\mu_1}{g_{p-1}}+\dfrac{1}{g_p}+\dfrac{\mu_2}{g_{p+1}}}, & if\ 2\le p\le n-2,\\ \dfrac{\mu_1+1}{\dfrac{\mu_1}{g_{n-1}}+\dfrac{1}{g_n}}, & if\ p=n-1.\end{cases} \quad (2)$$

Where the independent variables $\mu_1\in[0,1)$ and $\mu_2\in[0,1)$ are also called the membership degree of $I_p(\mu_1, \mu_2)$. When $\mu_1=\mu_2=\mu$, the inverse fuzzy number function of SFBODR is recorded as $I_p(\mu)$(p=1, 2, …, n-1). If the ordered difference rate g_p corresponds to the difference rate G_b of year b, then $I_p(\mu_1, \mu_2)$ represents the approximation of the difference rate G_b of year b.

Definition 3. Assuming there is a time series forecasting problem, let E = {E_1, E_2, …, E_n} be the universe of discourse for historical data, let G = {G_2, G_3, …, G_n} be the universe of discourse for the difference rate of historical data, and let g = {g_1, g_2, …, g_{n-1}} be the universe of discourse for the ordered difference rate. For each selected p $\in$ {1, 2, …, n-1}, the forecasting function $H_p(\mu_1, \mu_2)$ is defined as

$$H_p(\mu_1,\mu_2)=E_{b-1}\left(1+I_p(\mu_1,\mu_2)\right), \quad (3)$$

where the independent variables $\mu_1\in[0,1)$ and $\mu_2\in[0,1)$ are also called the membership degree of the forecasting function $H_p(\mu_1, \mu_2)$. E_{b-1} is the historical data of year b-1, $I_p(\mu_1, \mu_2)$ is the inverse fuzzy number function (2) of year p. When $\mu_1=\mu_2=\mu$, the forecasting function of SFBODR is recorded as $H_p(\mu)$(p=1, 2, …, n-1). When μ_1 and μ_2 fetch the specific membership degree in [0,1), then $H_p(\mu_1, \mu_2)$ is called the forecasting formula of SFBODR, the forecasting formula can be also recorded as SFBODR(μ_1, μ_2). If the ordered difference rate g_p corresponds to the difference rate G_b of year b, then $H_p(\mu_1, \mu_2)$ represents the predicted value of year b.

Assuming there is a time series forecasting problem, let E = {E_1, E_2, …, E_n} be the universe of discourse for historical data, let G = {G_2, G_3, …, G_n} be the universe of discourse for the difference rate of historical data, and let g = {g_1, g_2, …, g_{n-1}} be the universe of discourse for the ordered difference rate. If the membership degree μ_1 and μ_2 fetch specific values in [0,1), a forecasting formula SFBODR(μ_1, μ_2) can be established by (3). In the study of the time series forecasting problem, the application steps are as follows:

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

2). Write out the universe of discourse H for the historical data, the universe of discourse G for the difference rate and the universe of discourse g for the ordered difference rate;

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

4). Use SFBODR(μ_1, μ_2) to calculate the predicted values of the historical data.

Thus, the forecasting formula SFBODR(μ_1, μ_2) is a fuzzy time series forecasting model SFBODR(μ_1, μ_2).

Definition 4. For a time series forecasting problem, let E = {E_1, E_2, …, E_n} be the universe of discourse for historical data, let G = {G_2, G_3, …, G_n} be the universe of discourse for the difference rate of historical data, and let g = {g_1, g_2, …, g_{n-1}} be the universe of discourse for the ordered difference rate. When μ_1 and μ_2 fetch each value in the semi-closed and semi-open interval [0,1), infinite fuzzy time series forecasting models SFBODR(μ_1, μ_2) can be obtained. All the time series forecasting models SFBODR(μ_1, μ_2) are taken as elements, and they constitute a set, which is called SFTSFMBODR(The Set of Fuzzy Time Series Forecasting Models Based on the Ordered Difference Rate), the abbreviation is simplified as SFBODR. The general element of SFBODR is SFBODR(μ_1, μ_2). SFBODR(μ_1, μ_2) represents the forecasting formula of a fuzzy time series as well as a fuzzy time series forecasting model taking SFBODR(μ_1, μ_2) as the forecasting formula.

Theorem. For a time series forecasting problem, let E = {E_1, E_2, …, E_n} be the universe of discourse for historical data, let G = {G_2, G_3, …, G_n} be the universe of discourse for the difference rate of historical data, and let g = {g_1, g_2, …, g_{n-1}} be the universe of discourse for the ordered difference rate. If the ordered difference rate g_p(p$\in${1, 2, …, n-1}) corresponds to the difference rate G_b(b$\in${2, 3, …, n}), for each selected p$\in$ {1, 2, …, n-1}, then

1). The fuzzy number function $J_p(\mu_1, \mu_2)$, inverse fuzzy number function $I_p(\mu_1, \mu_2)$, and forecasting function $H_p(\mu_1, \mu_2)$ of SFBODR are all continuous functions;

2). When $\mu_1\to 0$ and $\mu_2\to 0$, the inverse fuzzy number function $I_p(\mu_1, \mu_2)$ converges to the difference rate G_b of historical data of year b:

$$\lim_{\mu_1\to 0,\mu_2\to 0} I_p(\mu_1,\mu_2)=G_b;$$

3). When $\mu_1\to 0$ and $\mu_2\to 0$, then the forecasting function $H_p(\mu_1, \mu_2)$ converges to the historical data N_b of year b:

$$\lim_{\mu_1\to 0,\mu_2\to 0} H_p(\mu_1,\mu_2)=N_b;$$

4). When the membership degree μ_1 and μ_2 are small enough, the forecasting function value $H_p(\mu_1, \mu_2)$ of SFBODR is equal to the historical data N_b of year b.

In the study of the forecasting problem of enrollments of the University of Alabama 1971–1992, Table I and II give the prediction results of the fuzzy time series forecasting models based on the defuzzification technology. The AFER and MSE of the prediction results in Table I and II are relatively small in all fuzzy time series forecasting models. Especially, Table I gives the prediction results by using SFBODR (0.00004, 0.00003) and SFBODR (0.00003, 0.00004) to forecast the enrollments of the University of Alabama 1971–1992; the MSE=0 and AFER=0.0%, which are the most ideal prediction accuracy.

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

$$MSE = \frac{1}{n}\sum_{b=1}^{n}(E_b - H_b)^2;$$

$$AFER = \left(\frac{1}{n}\sum_{b=1}^{n}\left|E_b - H_b\right| / E_b\right)\times 100\%.$$

TABLE I. COMPARISONS OF DIFFERENT FORECASTING MODELS

Year	Enrollments	Wang, Guo, Feng, Jin.[12]	Wang, Guo, Feng, Jin.[13]	Jilani, Burney, Ardil. [14]	Wang, Guo, Feng, Jin.[15]	SFBODR (0.00004,0.00003)	SFBODR (0.00003,0.00004)
1971	13055	-	-	13579	-	13055	13055
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	em, le	0	0

TABLE II. COMPARISONS OF DIFFERENT FORECASTING MODELS

Year	Enrollments	Feng, Guo, Wang, Zhang.[6]	Saxena, Sharma, Easo. [7]	Wang, Guo, Feng, Jin.[8]	Wang, Guo, Wang, Feng.[9]	Wang, Guo, Feng, Zhang.[10]	Stevenson, Porter. [11]
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

III. Conclusions

In the study of the forecasting problem of enrollments of the University of Alabama 1971–1992, the forecasting models SFBODR(0.00004, 0.00003) and SFBODR(0.00003, 0.00004) given in this paper can obtain MSE=0 and AFER=0.0%(See Table I), which are already the most ideal prediction accuracy. The history of unsatisfactory prediction accuracy of fuzzy time series forecasting models is terminated. The following work is to dig out forecasting models with high accuracy to predict the unknown data.

Acknowledgments

This work is supported by the Science and Technology Cooperation Project of the Academy and Government of Sanya (2016YD04, 2014YD33) and Hainan Province Nature Science Foundation Project (20156222, 714283).

References

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

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

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

[4] 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.

[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] 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.

[7] 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.

[8] 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].

[9] 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.

[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] 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.

[12] 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].

[13] 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].

[14] 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.

[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] 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.