Advances in Intelligent Systems Research, volume 153
International Conference on Applied Mathematics, Modeling and Simulation (AMMS 2017)

Set of Time Series Forecasting Models Using the Ordered Difference

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

***Abstract*—Set of time series forecasting models using the ordered difference of historical data (SOD) is proposed. For a time series, for example, for the enrollment of the University of Alabama in 1971–1992, when simulating the prediction of historical data, the automatic optimization search method can be used to sieve the satisfactory time series forecasting model $F_w(0.00003, 0.00003)$ in SOD, and the average forecasting error rate (AFER) of the predicted values can reach AFER=0% and the mean square error (MSE) is MSE=0. The problem that the prediction accuracy of the existing fuzzy time series forecasting models is not high has been solved.**

Keywords-the fractional sum function $D_w(x,y)$ of SOD; the inverse function $E_w(x,y)$ of fractional sum function of SOD; the forecasting model $F_w(x, y)$ of SOD; automatic optimization search method; time series

I. INTRODUCTION

Time series analysis is an old problem, the research on time series can be traced back to the paper proposed by Yule in 1927 [1]. In that paper, he carried out time series analysis on the Wolfer's sunspot numbers of 1749–1924, he found that sunspot numbers would come to a maximum every 9 to 11 years. The sunspots studied by Yule [1] are time series, which are the time series with no regularity. In the past ninety years, with the continuous progress of mathematical theory and computational technique, the classical models and methods have been fully able to study and deal with regular time series.

For the time series problem with regularity, Song and Chissom [2,3] (1993) first applied fuzzy set theory [5] to study them; the famous case is: the enrollment of the University of Alabama in 1971–1992, the historical data are shown in TABLE I [2-4], and the distribution diagram is shown in FIGURE I. This is a typical time series without any rules. Up to now, although many scholars have proposed the fuzzy time series forecasting models (for example, the forecasting models proposed in [7-14]) to simulate the prediction of the enrollment of the University of Alabama in 1971–1992, the AFER and MSE of the predicted value of the enrollment are still large.

It has some practical significance to simulate the prediction of historical data of time series. For example, it can be used to study stock [8]; studying traffic accidents can be useful for studying the amount of car premiums [6], etc. This article is enlightened by the paper of Jilani, Burney, and Ardil [6,7], set of time series forecasting models using the ordered difference of historical data (SOD) is proposed. The concept of satisfactory time series forecasting model is proposed, and the automatic optimization search method is presented. For the enrollment of the University of Alabama in 1971–1992, the automatic optimization search method can be used to sieve the time series forecasting model $F_w(0.00003, 0.00003)$, the average forecasting error rate (AFER) of the predicted value of the enrollment is AFER=0% and the mean square error (MSE) is MSE=0. The problem that the prediction accuracy of the existing fuzzy time series forecasting models is not high has been solved.

II. SET OF TIME SERIES FORECASTING MODELS (SOD)

Definition 1. If the universe of discourse of historical data of a time series is $A = \{A_1, A_2, \ldots, A_n\}$. The calculation formula of the difference of historical data is $B_w = A_w - A_{w-1}$. The data in the universe of discourse $B = \{B_2, B_3, \ldots, B_n\}$ of the difference of historical data are arranged from small to large, then constitute a new set $b = \{b_2, b_3, ..., b_n\}$, which is called the universe of discourse of the ordered difference of historical data.

Definition 2. If the universe of discourse of historical data of a time series is $A = \{A_1, A_2, \ldots, A_n\}$, the universe of discourse of the difference of historical data is $B = \{B_2, B_3, \ldots, B_n\}$, the universe of discourse of the ordered difference of historical data is $b = \{b_2, b_3, ... b_n\}$, (1), (2), and (3) are defined on b, where the independent variables $x \in (0,1)$ and $y \in (0,1)$. $D_w(x,y)$ is the fractional sum function of SOD; $E_w(x,y)$ is the corresponding inverse function of fractional sum function $D_w(x,y)$; $F_w(x,y)$ is the forecasting function of SOD.

$$D_w(x,y) = \begin{cases} \frac{1}{b_2} + \frac{y}{b_3}, & w=2, \\ \frac{x}{b_{w-1}} + \frac{1}{b_w} + \frac{y}{b_{w+1}}, & 3 \le w \le n-1, \\ \frac{x}{b_{n-1}} + \frac{1}{b_n}, & w = n. \end{cases} \quad (1)$$

$$E_w(x,y)=\begin{cases}\dfrac{1+y}{\dfrac{1}{b_2}+\dfrac{y}{b_3}}, & w=2,\\ \dfrac{x+1+y}{\dfrac{x}{b_{w-1}}+\dfrac{1}{b_w}+\dfrac{y}{b_{w+1}}}, & 3\le w\le n-1,\\ \dfrac{x+1}{\dfrac{x}{b_{n-1}}+\dfrac{1}{b_n}}, & w=n.\end{cases} \quad (2)$$

$$F_w(x,y)=A_{w-1}+E_w(x,y). \quad (3)$$

Definition 3. Arbitrarily taking a $x\in(0,1)$ and $y\in(0,1)$, we can obtain a time series forecasting model. A set of all such forecasting models constructs: the set of time series forecasting models using the ordered difference (STSFMUOD), the abbreviation is further simplified as SOD.

Definition 4. For a time series, if the prediction of the historical data of the time series is simulated, and the predicted values obtain MSE=0 and AFER=0%, then the time series forecasting model is called satisfactory.

Definition 5. For a time series, the satisfactory time series forecasting model in SOD is automatically searched, this method is called automatic optimization search method. The detailed approach is: for a time series, a decimal number is used as the starting point, programming, searching, and computing, ..., until the satisfactory time series forecasting model in SOD is sieved (satisfy AFER=0%).

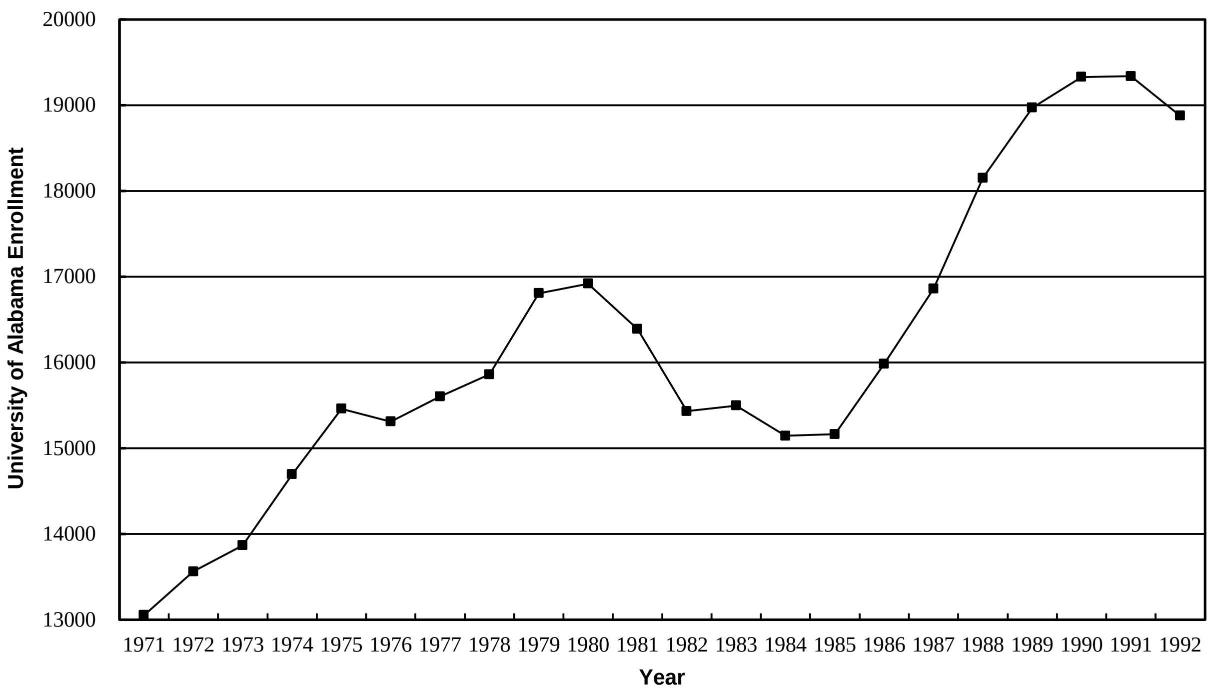

FIGURE I. THE DISTRIBUTION DIAGRAM OF THE ENROLLMENT OF THE UNIVERSITY OF ALABAMA IN 1971–1992.

Example 1.In simulating the prediction of the enrollment of the University of Alabama in 1971–1992, the whole process of applying automatic optimization search method to sieve the satisfactory time series forecasting model in SOD is: 1) Input the universe of discourse A of historical data of the time series, the universe of discourse B of the difference, and the universe of discourse b of the ordered difference. 2) Using x = y = 0.003 as starting point, successively select x = y = 0.003, x = y = 0.0003, x = y = 0.00003, ..., through programming, searching, and computing, ..., until the satisfactory time series forecasting model is sieved. We first get TABLE I, because AFER ≠ 0% and MSE ≠ 0, continue searching, calculating, then get TABLE II; because AFER ≠ 0% and MSE ≠ 0, continue searching, calculating, then get TABLE III; because AFER=0% and MSE=0, F_w(0.00003,0.00003) in TABLE III is a satisfactory time series forecasting model.

III. CONCLUSIONS

For a time series, the automatic optimization search method can be used to sieve ideal and satisfactory time series forecasting model in SOD, in such a fashion that the prediction accuracy of predicted values can reach AFER=0% and MSE=0. The problem that the prediction accuracy of the existing fuzzy time series forecasting models is not high has been solved. Such time series forecasting model also has certain practical significance.

TABLE I. PARTIAL FORECASTING MODEL OF F_W (0.003,0.003) IS APPLIED TO FORECAST THE ENROLLMENT OF THE UNIVERSITY OF ALABAMA

Year	Enrollment A_{uw}	The ordered difference b_w	Forecast F_{uw}	$F_{uw} - A_{uw}$	$(F_{uw}-A_{uw})^2$	$\lvert F_{uw}-A_{uw}\rvert/A_{uw}$
1971	13055	-	-	-	-	-
1972	13563	d_1, C_{1982}, -955	13563	0	0	0.000000
1973	13867	d_2, C_{1981}, -531	13867	0	0	0.000000
1974	14696	d_3, C_{1992}, -461	14696	0	0	0.000000
1975	15460	d_4, C_{1984}, -352	15459	-1	1	0.000065
1976	15311	d_5, C_{1976}, -149	15302	-9	81	0.000588
1977	15603	d_6, C_{1991}, 9	15603	0	0	0.000000
1978	15861	d_7, C_{1985}, 18	15860	-1	1	0.000063
1979	16807	d_8, C_{1983}, 64	16808	1	1	0.000059
1980	16919	d_9, C_{1980}, 112	16919	0	0	0.000000
1981	16388	d_{10},C_{1978}, 258	16388	0	0	0.000000
1982	15433	d_{11},C_{1977}, 292	15435	2	4	0.000130
1983	15497	d_{12},C_{1973}, 304	15497	0	0	0.000000
1984	15145	d_{13},C_{1990}, 358	15146	1	1	0.000066
1985	15163	d_{14},C_{1972}, 508	15163	0	0	0.000000
1986	15984	d_{15},C_{1975}, 764	15984	0	0	0.000000
1987	16859	d_{16},C_{1989}, 820	16859	0	0	0.000000
1988	18150	d_{17},C_{1986}, 821	18149	-1	1	0.000055
1989	18970	d_{18},C_{1974}, 829	18970	0	0	0.000000
1990	19328	d_{19},C_{1987}, 875	19328	0	0	0.000000
1991	19337	d_{20},C_{1979}, 946	19337	0	0	0.000000
1992	18876	d_{21},C_{1988},1291	18876	0	0	0.000000
AFER						0.0049%
MSE					4.2857	

Note. In TABLE II: MSE (Mean Square Error): $MSE = \frac{1}{n-1}\sum_{uw=2}^{n}(A_{uw} - F_{uw})^2$; AFER (Average Forecasting Error Rate): $AFER = \left(\frac{1}{n-1}\sum_{uw=2}^{n}\lvert A_{uw} - F_{uw}\rvert / A_{uw}\right)\times 100\%$.

TABLE II. PARTIAL FORECASTING MODEL OF F_W (0.0003,0.0003) IS APPLIED TO FORECAST THE ENROLLMENT OF THE UNIVERSITY OF ALABAMA

Year	Enrollment A_{uw}	The ordered difference b_w	Forecast F_{uw}	$F_{uw} - A_{uw}$	$(F_{uw}-A_{uw})^2$	$\lvert F_{uw}-A_{uw}\rvert/A_{uw}$
1971	13055	-	-	-	-	-
1972	13563	d_1, C_{1982}, -955	13563	0	0	0.000000
1973	13867	d_2, C_{1981}, -531	13867	0	0	0.000000
1974	14696	d_3, C_{1992}, -461	14696	0	0	0.000000
1975	15460	d_4, C_{1984}, -352	15460	0	0	0.000000
1976	15311	d_5, C_{1976}, -149	15310	1	1	0.000065
1977	15603	d_6, C_{1991}, 9	15603	0	0	0.000000
1978	15861	d_7, C_{1985}, 18	15861	0	0	0.000000
1979	16807	d_8, C_{1983}, 64	16807	0	0	0.000000
1980	16919	d_9, C_{1980}, 112	16919	0	0	0.000000
1981	16388	d_{10},C_{1978}, 258	16388	0	0	0.000000
1982	15433	d_{11},C_{1977}, 292	15433	0	0	0.000000
1983	15497	d_{12},C_{1973}, 304	15497	0	0	0.000000
1984	15145	d_{13},C_{1990}, 358	15145	0	0	0.000000
1985	15163	d_{14},C_{1972}, 508	15163	0	0	0.000000
1986	15984	d_{15},C_{1975}, 764	15984	0	0	0.000000
1987	16859	d_{16},C_{1989}, 820	16859	0	0	0.000000
1988	18150	d_{17},C_{1986}, 821	18150	0	0	0.000000
1989	18970	d_{18},C_{1974}, 829	18970	0	0	0.000000
1990	19328	d_{19},C_{1987}, 875	19328	0	0	0.000000
1991	19337	d_{20},C_{1979}, 946	19337	0	0	0.000000
1992	18876	d_{21},C_{1988},1291	18876	0	0	0.000000
AFER						0.0003%
MSE					0.0476	

TABLE III. PARTIAL FORECASTING MODEL OF F_W (0.00003,0.00003) IS APPLIED TO FORECAST THE ENROLLMENT OF THE UNIVERSITY OF ALABAMA

Year	**Enrollment A_{uw}**	**The ordered difference b_w**	**Forecast F_{uw}**	**$F_{uw}-A_{uw}$**	**$(F_{uw}-A_{uw})^2$**	**$\|F_{uw}-A_{uw}\|/A_{uw}$**
1971	13055	-	-	-	-	-
1972	13563	d_1, C_{1982}, -955	13563	0	0	0.000000
1973	13867	d_2, C_{1981}, -531	13867	0	0	0.000000
1974	14696	d_3, C_{1992}, -461	14696	0	0	0.000000
1975	15460	d_4, C_{1984}, -352	15460	0	0	0.000000
1976	15311	d_5, C_{1976}, -149	15311	0	0	0.000000
1977	15603	d_6, C_{1991}, 9	15603	0	0	0.000000
1978	15861	d_7, C_{1985}, 18	15861	0	0	0.000000
1979	16807	d_8, C_{1983}, 64	16807	0	0	0.000000
1980	16919	d_9, C_{1980}, 112	16919	0	0	0.000000
1981	16388	d_{10},C_{1978}, 258	16388	0	0	0.000000
1982	15433	d_{11},C_{1977}, 292	15433	0	0	0.000000
1983	15497	d_{12},C_{1973}, 304	15497	0	0	0.000000
1984	15145	d_{13},C_{1990}, 358	15145	0	0	0.000000
1985	15163	d_{14},C_{1972}, 508	15163	0	0	0.000000
1986	15984	d_{15},C_{1975}, 764	15984	0	0	0.000000
1987	16859	d_{16},C_{1989}, 820	16859	0	0	0.000000
1988	18150	d_{17},C_{1986}, 821	18150	0	0	0.000000
1989	18970	d_{18},C_{1974}, 829	18970	0	0	0.000000
1990	19328	d_{19},C_{1987}, 875	19328	0	0	0.000000
1991	19337	d_{20},C_{1979}, 946	19337	0	0	0.000000
1992	18876	d_{21},C_{1988},1291	18876	0	0	0.000000
AFER						0%
MSE					0	

ACKNOWLEDGMENT

This work is supported by Hainan Province Nature Science Foundation Project (Fund number: 714283), the Science and Technology Cooperation Project of the Academy and Government of Sanya (Fund number: 2016YD04, 2015YD33), and Hainan Tropical Ocean University Research Project (Fund number: RDJGb2017–14).

REFERENCES

[1] G. Udny Yule, “On a Method of Investigating Periodicities in Disturbed Series, with Special Reference to Wolfer's Sunspot Numbers,” Philosophical Transactions of the Royal Society. A: Mathematical, Physical and Engineering Sciences, vol. 226, pp. 267–298, 1927.

[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 I,” Fuzzy Sets and Systems. vol. 54, pp. 1–9, 1993.

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

[5] L. A. Zadeh, “Fuzzy set,” Fuzzy Sets and Systems. vol. 8, pp. 338–353, 1965.

[6] 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.

[7] Tahseen Ahmed Jilani, Syed Muhammad Aqil Burney, and Cemal Ardil, “Fuzzy Metric Approach for Fuzzy Time Series Forecasting based on Frequency Density Based Partitioning,” World Academy of Science, Engineering and Technology. vol. 34, pp. 1–6, 2007.

[8] Bhagawati P. Joshi and Sanjay Kumar, “A computational method for fuzzy time series forecasting based on difference parameters,” International Journal of Modeling, Simulation, and Scientific Computing. vol. 4, No. 1, pp. 1250023–1–1250023–12, 2013.

[9] M. Stevenson and J. 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.

[10] 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, No. 3, pp. 957–961, 2012.

[11] R. C. Tsaur, J. C. O. Yang, and H. F. Wang, “Fuzzy relation analysis in fuzzy time series model,” Comput. Math. Appl. vol. 49, pp. 539–548, 2005.

[12] T. A. Jilani, S. M. A. Burney, and C. Ardil, “ Multivariate high order fuzzy time series forecasting for car road accidents,” International Journal of Computational Intelligence. vol. 4, No. 1, pp. 15–20, 2007.

[13] S. R. Singh, “A simple method of forecasting based on fuzzy time series,” Appl. Math. Comput. vol. 186, pp. 330–339, 2007.

[14] S. R. Singh, “A robust method of forecasting based on fuzzy time series,” Appl. Math. Comput. vol. 188, pp. 472–484, 2007.

[15] 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.

[16] 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.

[17] Ming Liu, Hongxu Wang, Youming Li, and Xuebing Huang, “Fuzzy Time Series Forecasting Model Based on DRDDR and Application,” 2016 2nd International Conference on Electronics, Network and Computer Engineering (ICENCE2016), Advances in Computer Science Research, Part C, ISSN 2352-538X. vol. 67, pp. 574–578, 2016.

[18] Ming Liu, Hongxu Wang, Youming Li, and Xuebing Huang, “Analysis of Proportion of Possession of Tourism Income in Tertiary Industry in Hainan Province based on DDD,” Advances in Economics, Business and Management Research, (IEESASM 2016), Part D, ISSN 2352-538X. vol. 67, pp. 691–694, 2016.