ATLANTIS PRESS

Advances in Intelligent Systems Research, volume 153

International Conference on Applied Mathematics, Modeling and Simulation (AMMS 2017)

A Set of Time Series Forecasting Models Based on 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*—A set of time series forecasting models based on the ordered difference of historical data (ASOD) is proposed. For a time series, the automatic optimization search method can be applied to sieve standard time series forecasting model $C_p(k,h)$ in ASOD, so that in simulating the prediction of historical data of the time series, the predicted values can reach AFER (Average Forecasting Error Rate) = 0% and MSE (Mean Square Error) = 0. For instance, for the enrollment of the University of Alabama in 1971–1992, the automatic optimization search method can be applied to sieve standard time series forecasting model $C_p(0.0003,0.0003)$, the problem that the prediction accuracy of fuzzy time series forecasting model is not ideal for many years has been solved.**

Keywords-time series; automatic optimization search method; the fractional sum function of ASOD; the inverse function of fractional sum function of ASOD; the forecasting function of ASOD

I. INTRODUCTION

In 1993, Song and Chissom [1,2] proposed the first fuzzy time series forecasting model, they pioneered the application of fuzzy set theory [3] to study time series, and used the model to simulate the prediction of the enrollment of the University of Alabama in 1971–1992, its historical data is shown in TABLE I [2], its distribution diagram is shown in FIGURE I. As seen from TABLE I and FIGURE I, this example is actually a time series, it is rolling and irregular time series. Song and Chissom [1,2] applied fuzzy set theory to study time series. As early as 1927, Yule [4] studied the Wolfer's sunspot numbers of 1749–1924; in that paper the time series was called disturbed series. Yule [4] found the sunspot numbers will reach a maximum value every nine to eleven years. It is an influential paper on time series analysis. So far, scholars have tried to use various advanced theories and techniques to study time series continuously, for the regular time series, scholars have got comprehensive method of time series analysis. Since Song and Chissom [1,2] proposed the first fuzzy time series forecasting model, although a large number of fuzzy time series forecasting models are born (e.g., models provided by [5-18]), when simulating the prediction of the enrollment of the University of Alabama in 1971–1992, the mean square error (MSE) and the average forecasting error rate (AFER) of the predicted value of the enrollment are still not ideal. For a time series, simulating the prediction of historical data, this problem has some practical significance. For example, it can be used to study traffic accidents in various situations and provide a reference for setting the amount of car premiums [5]; it also can be used to study stock [7], the potential application is to study stock software, etc. With the help of the thought of fuzzy time series forecasting model proposed by Jilani, Burney, and Ardil [5,6], this paper presents a set of time series forecasting models based on the ordered difference of historical data (ASOD), and presents the concept of standard time series forecasting model. For the enrollment of the University of Alabama in 1971–1992, the automatic optimization search method can be applied to sieve time series forecasting model $C_p(0.0003,0.0003)$, so that, the average forecasting error rate of the predicted value of the enrollment can reach AFER=0% and MSE=0. The problem that the prediction accuracy of the fuzzy time series forecasting model is not ideal for twenty-four years has been solved.

II. SET OF TIME SERIES FORECASTING MODELS (ASOD)

Definition 1. Suppose that the universe of discourse of historical data of a time series is $U = \{ U_1, U_2, \ldots, U_n\}$. The calculation formula of the difference of historical data is $V_p= U_p–U_{p-1}$. The data in the universe of discourse $V=\{V_2, V_3, \ldots, V_n\}$ of the difference of historical data are arranged from small to large, then constitute a new set $v = \{v_2, v_3, ..., v_n\}$, which is called the universe of discourse of the ordered difference of historical data.

Definition 2. Suppose that the universe of discourse of historical data of a time series is $U = \{ U_1, U_2, \ldots, U_n\}$, the universe of discourse of the difference of historical data is $V=\{V_2, V_3, \ldots, V_n\}$, the universe of discourse of the ordered difference of historical data is $v = \{v_2, v_3, ..., v_n\}$, (1), (2), and (3) are defined on v, where the independent variables $k \in (0,1)$ and $h \in (0,1)$. $T_p(k,h)$ is the fractional sum function of ASOD; $S_p(k,h)$ is the corresponding inverse function of fractional sum function $T_p(k,h)$; $C_p(k,h)$ ($p \in \{2, 3, \ldots, n\}$) is the forecasting function of ASOD.

$$T_p(k,h)=\begin{cases}\dfrac{1}{v_1}+\dfrac{h}{v_2}, & p=2,\\ \dfrac{k}{v_{p-1}}+\dfrac{1}{v_p}, & 3\le p\le n,\end{cases} \tag{1}$$

$$S_p(k,h)=\begin{cases}\dfrac{1+h}{\dfrac{1}{v_1}+\dfrac{h}{v_2}}, & p=2,\\ \dfrac{k+1}{\dfrac{k}{v_{p-1}}+\dfrac{1}{v_p}}, & 3\le p\le n,\end{cases} \tag{2}$$

$$C_p(k,h)=U_{p-1}+S_p(k,h). \tag{3}$$

Definition 3. Arbitrarily taking a $k\in(0,1)$ and $h\in(0,1)$, we can obtain a time series forecasting model. A set of all such forecasting models is called: a set of time series forecasting models based on the ordered difference (ASTSFMBOD), the abbreviation is further simplified as ASOD.

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

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

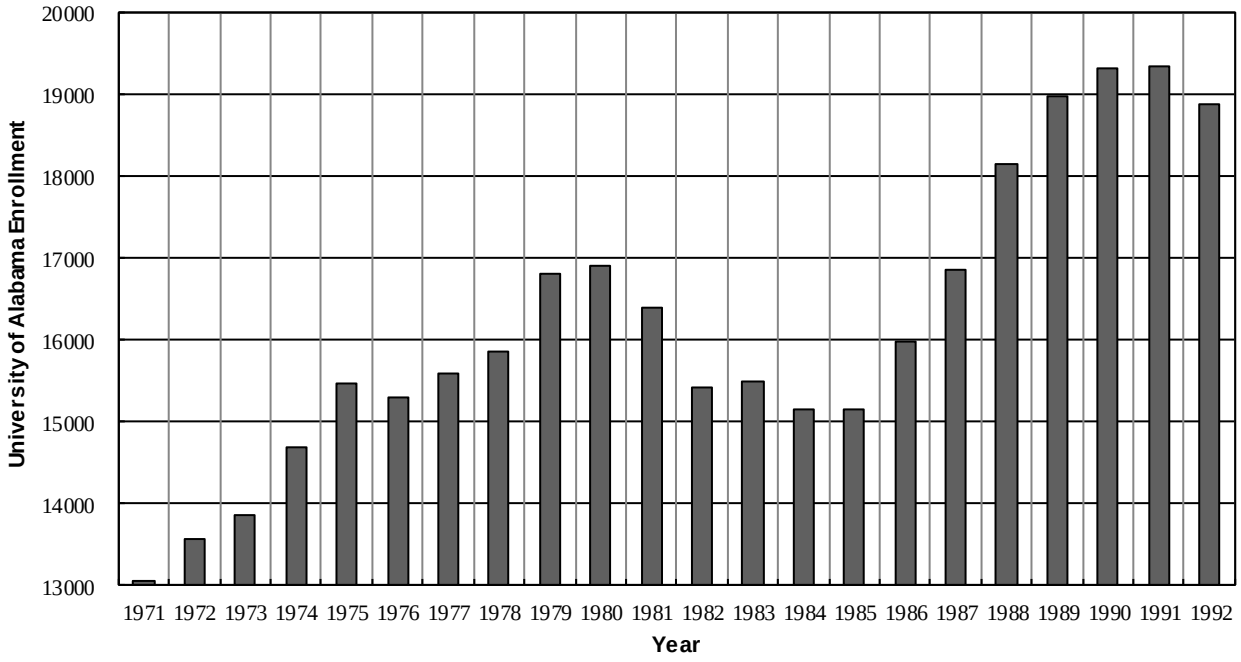

FIGURE I. THE DISTRIBUTION DIAGRAM OF THE ENROLLMENT OF THE UNIVERSITY OF ALABAMA IN 1971–1992.

Example 1.In simulating the prediction of the enrollment of the University of Alabama in 1971–1992, this paper applies automatic optimization search method to sieve the standard time series forecasting model in ASOD, the whole process is: 1) Input the universe of discourse U of historical data of the time series, the universe of discourse V of the difference, and the universe of discourse v of the ordered difference; 2) Using k = h = 0.03 as the starting point of the calculation, successively select k = h = 0.03, k = h = 0.003, k = h = 0.0003, ..., through programming, searching, computing, ..., until the standard time series forecasting model is sieved. We get TABLE I, because AFER ≠ 0% and MSE ≠ 0, continue searching and calculating, then get TABLE II; because AFER ≠ 0% and MSE ≠ 0, continue searching and calculating, then get TABLE III; because AFER=0% and MSE=0, C_p(0.0003,0.0003) in TABLE III is standard time series forecasting model.

III. CONCLUSIONS

For a time series, the automatic optimization search method can be applied to sieve standard time series forecasting model in ASOD. TABLE III shows that the automatic optimization search method can be applied to sieve standard time series forecasting model C_p(0.0003, 0.0003) for the enrollment of the University of Alabama in 1971–1992, the average forecasting error rate and mean square error of the predicted value of the enrollment can obtain AFER=0% and MSE=0. The problem that the prediction accuracy of the fuzzy time series forecasting model is not ideal for twenty-four years has been solved.

TABLE I. $C_P(0.03,0.03)$ IS APPLIED TO FORECAST THE ENROLLMENT OF THE UNIVERSITY OF ALABAMA

Year	Enrollment U_p	The ordered difference V_{jp}	Forecast C_{jp}	$U_{jp}-C_{jp}$	$(U_{jp}-C_{jp})^2$	$\lvert U_{jp}-C_{jp}\rvert/U_{jp}$
1971	13055	-	-	-	-	-
1972	13563	d_1, C_{1982}, -955	13557	6	36	0.000442
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	15449	11	121	0.000712
1976	15311	d_5, C_{1976}, -149	15308	3	9	0.000196
1977	15603	d_6, C_{1991}, 9	15602	1	1	0.000064
1978	15861	d_7, C_{1985}, 18	15852	9	81	0.000567
1979	16807	d_8, C_{1983}, 64	16805	2	4	0.000119
1980	16919	d_9, C_{1980}, 112	16917	2	4	0.000118
1981	16388	d_{10},C_{1978}, 258	16381	7	49	0.000427
1982	15433	d_{11},C_{1977}, 292	15455	-22	484	0.001426
1983	15497	d_{12},C_{1973}, 304	15493	4	16	0.000258
1984	15145	d_{13},C_{1990}, 358	15143	2	4	0.000132
1985	15163	d_{14},C_{1972}, 508	15162	1	1	0.000066
1986	15984	d_{15},C_{1975}, 764	15984	0	0	0.000000
1987	16859	d_{16},C_{1989}, 820	16858	1	1	0.000059
1988	18150	d_{17},C_{1986}, 821	18136	14	196	0.000771
1989	18970	d_{18},C_{1974}, 829	18968	2	4	0.000105
1990	19328	d_{19},C_{1987}, 875	19326	2	4	0.000103
1991	19337	d_{20},C_{1979}, 946	19337	0	0	0.000000
1992	18876	d_{21},C_{1988},1291	18874	2	4	0.000106
AFER						0.027%
MSE					48.5238	

Note. In TABLE II: MSE (Mean Square Error): $MSE = \frac{1}{n-1}\sum_{jp=2}^{n}(U_{jp}-C_{jp})^2$; AFER (Average Forecasting Error Rate): $AFER = \left(\frac{1}{n-1}\sum_{jp=2}^{n}\left|U_{jp}-C_{jp}\right|/U_{jp}\right)\times 100\%.$

TABLE II. $C_P(0.003,0.003)$ IS APPLIED TO FORECAST THE ENROLLMENT OF THE UNIVERSITY OF ALABAMA

Year	Enrollment U_p	The ordered difference V_{jp}	Forecast C_{jp}	$U_{jp}-C_{jp}$	$(U_{jp}-C_{jp})^2$	$\lvert U_{jp}-C_{jp}\rvert/U_{jp}$
1971	13055	-	-	-	-	-
1972	13563	d_1, C_{1982}, -955	13562	1	1	0.000074
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	15311	0	0	0.000000
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	16807	0	0	0.000000
1980	16919	d_9, C_{1980}, 112	16919	0	0	0.000000
1981	16388	d_{10},C_{1978}, 258	16387	1	1	0.000061
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	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	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						0.0021%
MSE					0.4286	

TABLE III. $C_P(0.0003,0.0003)$ IS APPLIED TO FORECAST THE ENROLLMENT OF THE UNIVERSITY OF ALABAMA

Year	Enrollment U_p	The ordered difference V_{jp}	Forecast C_{jp}	$U_{jp}–C_{jp}$	$(U_{jp}–C_{jp})^2$	$\|U_{jp}–C_{jp}\|/U_{jp}$
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] Q. Song and B. S. Chissom, “Fuzzy series and its models,” Fuzzy Sets and Systems. vol. 54, pp. 269–277, 1993.

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

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[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] 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.

[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] S. R. Singh, “A robust method of forecasting based on fuzzy time series,” Appl. Math. Comput. vol. 188, pp. 472–484, 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] T. A. Jilani and S. M. A. Burney, “M-factor high order fuzzy time series forecasting for road accident data,” IEEE-IFSA 2007, World Congress, Cancun, Mexico, June 18-21, Forthcoming in Book series Advances in Soft Computing, Springer-Verlag, 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] 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.

[17] 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.

[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.