International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:3, No:7, 2009

Fuzzy Time Series Forecasting Using Percentage Change as the Universe of Discourse

Meredith Stevenson and John E. Porter

Abstract—Since the pioneering work of Zadeh, fuzzy set theory has been applied to a myriad of areas. Song and Chissom introduced the concept of fuzzy time series and applied some methods to the enrollments of the University of Alabama. In recent years, a number of techniques have been proposed for forecasting based on fuzzy set theory methods. These methods have either used enrollment numbers or differences of enrollments as the universe of discourse. We propose using the year to year percentage change as the universe of discourse. In this communication, the approach of Jilani, Burney, and Ardil is modified by using the year to year percentage change as the universe of discourse. We use enrollment figures for the University of Alabama to illustrate our proposed method. The proposed method results in better forecasting accuracy than existing models.

Keywords—Fuzzy forecasting, fuzzy time series, fuzzified enrollments, time-invariant model

I. Introduction

The initial work of Zadeh concerning fuzzy set theory has been applied to a several diverse areas. Song and Chissom [17] introduced a theory for fuzzy time series and applied fuzzy time series methods [18], [19] that modeled the enrollments of the University of Alabama. In recent years, a number of techniques have been proposed for forecasting based on fuzzy set theory methods. Fuzzy forecasting methods have been used to model enrollment data for the University of Alabama ([1], [2], [3], [4], [5], [6], [7], [9], [15], [16], [18], and [19]), daily temperatures ([13]), and car fatalities ([8] and [10]).

Instead of using actual enrollments, Hwang, Chen, and Lee [7] and Sah and Degtiarev [15] proposed using year to year differences of the enrollments of the University of Alabama for the universe of discourse in their fuzzy forecasting methods producing better forecasting accuracy than those of Song and Chissom [18], [19]. While differences of enrollments can provide better forecasting accuracy, differences alone lack context for which the increase or decrease occured. For example, an increase of 100 students is treated the same whether the increase occured within a pool of 500 students or one with 50,000 students. With this possible shortcoming in mind, we propose using the percentage change of year to year enrollments as the universe of discourse. In Section 2, we modify the method Jilani, Burney, and Ardil [9] replacing the universe of discourse with the percentage change of year to year enrollments. In Section 3, we compare the proposed forecasting model with existing methods, and concluding remarks are provided in Section 4.

Meredith Stevenson is an undergraduate in the Department of Mathematics & Statistics, Murray State University, Murray, KY, 42071 USA e-mail: meredith.day.stevenson@gmail.com.

John E. Porter is with the Department of Mathematics & Statistics, Murray State University, Murray, KY, 42071 USA e-mail: ted.porter@murraystate.edu.

II. A New Method Fuzzy Time Series Forecasting Method

In this section, we modify the Jilani, Burney, and Ardil's [9] method for modeling the enrollments of the University of Alabama based fuzzy time series forecasting methods. Instead of using actual enrollments for the universe of discourse, we propose using the percentage the enrollment changed from year to year for the universe of discourse. The enrollments of the University of Alabama under consideration may found in Tables IV and V.

Step 1: Define the universe of discourse U and partition it into intervals $u_1, u_2, ..., u_n$ of equal length. The percentage change of enrollment from year to year is given in Table I and ranges from -5.83% to 7.66%. For example, take the universe of discourse to be $U = [-6, 8]$ and partition U into seven equal intervals.

TABLE I
The Year-to-year Percentage change of enrollments

Year to Year	Change	Year to Year	Change
1971-1972	3.89%	1982-1983	0.41%
1972-1973	2.24%	1983-1984	-2.27%
1973-1974	5.98%	1984-1985	0.12%
1974-1975	5.20%	1985-1986	5.41%
1975-1976	-0.96%	1986-1987	5.47%
1976-1977	1.91%	1987-1988	7.66%
1977-1978	1.65%	1988-1989	4.52%
1978-1979	5.96%	1989-1990	1.89%
1979-1980	0.67%	1990-1991	0.05%
1980-1981	-3.14%	1991-1992	-2.38%
1981-1982	-5.83%		

Step 2: Find the density based distribution of the year to year percentage change by sorting the data into the corresponding intervals shown in Table II. Then determine the number of percentage data that falls into each interval. Table II contains the density based distribution of the percentage data displayed in Table I with seven intervals.

Find the interval having the largest number of percentage data and divide it into four sub-intervals of equal length. Next, divide the interval having the second largest number of percentage data into three sub-intervals of equal length. The interval having the third largest number of percentage data should be divided into two sub-intervals of equal length. Let all subsequent intervals remain unchanged in length.

After completing this step, the universe of discourse is divided into the intervals shown in Table III.

Step 3: Define each fuzzy set X_i based on the re-divided intervals and fuzzify the historical enrollments shown in Table I, where fuzzy set X_i denotes a linguistic value of the

TABLE II
FREQUENCY DENSITY BASED DISTRIBUTION OF YEAR-TO-YEAR PERCENTAGE CHANGE DATA

Intervals	Number of Data
$[-6.0, -4.0]$	1
$[-4.0, -2.0]$	1
$[-2.0, 0.00]$	2
$[0.00, 2.00]$	7
$[2.00, 4.00]$	3
$[4.00, 6.00]$	6
$[6.00, 8.00]$	1

TABLE III
FUZZY INTERVALS USING FREQUENCY DENSITY BASED PARTITIONING

Linguistic	Intervals
X_1	$[-6.0, -4.0]$
X_2	$[-4.0, -2.0]$
X_3	$[-2.0, 0.00]$
X_4	$[0.00, 0.50]$
X_5	$[0.50, 1.00]$
X_6	$[1.00, 1.50]$
X_7	$[1.50, 2.00]$
X_8	$[2.00, 3.00]$
X_9	$[3.00, 4.00]$
X_{10}	$[4.00, 4.67]$
X_{11}	$[4.67, 5.33]$
X_{12}	$[5.33, 6.00]$
X_{13}	$[6.00, 8.00]$

year to year percentage change represented by a fuzzy set. As in [9] we use a triangular membership function to define the fuzzy sets X_i [10].

Step 4: Defuzzify the fuzzy data using the forecasting formula (see [9])

$$t_j = \begin{cases} \dfrac{1.5}{\dfrac{1}{a_1} + \dfrac{1}{a_2}} & \text{, if } j = 1 \\[2ex] \dfrac{2}{\dfrac{0.5}{a_{j-1}} + \dfrac{1}{a_j} + \dfrac{1}{a_{j+1}}} & \text{, if } 2 \le j \le n-1 \\[2ex] \dfrac{1.5}{\dfrac{0.5}{a_{n-1}} + \dfrac{1}{a_n}} & \text{, if } j = n \end{cases}$$

where a_{j-1}, a_j, a_{j+1} are the midpoints of the fuzzy intervals X_{j-1}, X_j, X_{j+1} respectively. t_j yields the predicted year to year percentage change of enrollment. Use the predicted percentage on the previous year's enrollment to determine the forecasted enrollment. The forecasted enrollment is provided in Table IV.

III. A COMPARISON OF DIFFERENT FORECASTING METHODS

As in [9], we use the average forecasting error rate (AFER) and mean square error (MSE) to compare the forecasting results of different forecasting methods:

$$AFER = \frac{|A_i - F_i|/A_i}{n} \times 100\%$$

$$MSE = \frac{\sum_{i=1}^{n}(A_i - F_i)^2}{n}$$

where A_i denotes the actual enrollment and F_i denotes the forecasting enrollment of year i, respectively.

In Table V the forecasting results of the proposed method is compared with that of the existing methods. From Table III, we can see that when the number of intervals in the universe of discourse is thirteen and the intervals are subpartitioned based on frequency density, the proposed method produces the smallest values of the MSE and AFER as compared to other methods of fuzzy time series forecasting. That is, the proposed method can produce a better accuracy when forecasting enrollments than the existing methods.

IV. CONCLUSION

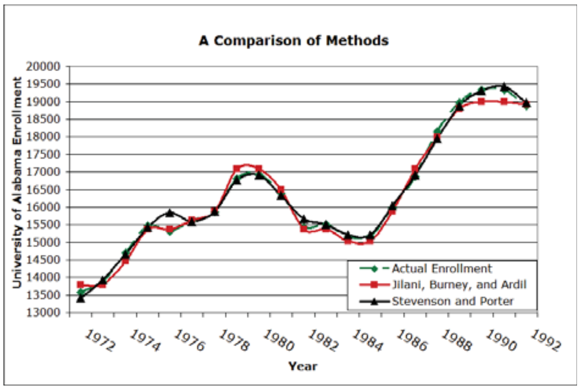

Fig. 1. A comparison between the proposed method and Jilani, Burney, Ardil [10]

In this communication, we modified Jilani, Burney, and Ardil's approach to modeling enrollments using year to year percentage change as the universe of discourse. From Table V, one sees that the proposed method provides the smallest AFER and MSE and improves on other methods using fuzzy time series forecasting methods. For future work, we will focus on how well these methods predict future enrollments.

ACKNOWLEDGMENT

The authors would like to thank Nick Hooten for his assistance with Figure 1.

REFERENCES

[1] Chen, S. M. 1996. *Forecasting enrollments based on fuzzy time series*, Fuzzy Sets and Systems, 81: 311-319.
[2] S. M. Chen, *Forecasting enrollments based on high-order fuzzy time series*, Cybernetics and Systems: An International Journal, Vol. 33: pp. 1-16, 2002.
[3] Chen, S. M. and Hsu, C.-C. 2004. *A new method to forecasting enrollments using fuzzy time series*, International Journal of Applied Science and Engineering, 2, 3: 234-244.
[4] S. M. Chen, J. R. Hwang, *Temperature prediction using fuzzy time series*, IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics, Vol. 30, pp.263-275, 2000.
[5] K. Huarng, *Heuristic models of fuzzy time series for forecasting*, Fuzzy Sets and Systems, Vol. 123, pp. 369-386, 2002.

TABLE IV
FORECASTING RESULTS OF THE PROPOSED MODEL

Year	Enrollmets	Percentage	Fuzzy Set	Predicted Precentage	Forecast	$(A_i - F_i)^2$	$\frac{\lvert A_i - F_i \rvert}{A_i}$
1971	13055						
1972	13563	3.89%	X_9	2.7229%	13410	23264	0.011246
1973	13867	2.24%	X_9	2.7229%	13932	4265	0.004710
1974	14696	5.98%	X_{12}	5.7479%	14664	1020	0.002173
1975	15460	5.20%	X_{11}	4.9537%	15424	1296	0.002329
1976	15311	-0.96%	X_4	2.5000%	15847	286760	0.034975
1977	15603	1.91%	X_8	1.7573%	15580	526	0.001470
1978	15861	1.65%	X_8	1.7573%	15877	262	0.001021
1979	16807	5.96%	X_{12}	5.7479%	16773	1178	0.002042
1980	16919	0.67%	X_6	0.5357%	16897	482	0.001298
1981	16388	-3.14%	X_2	-3.4146%	16341	2182	0.002851
1982	15433	-5.83%	X_1	-4.3750%	15671	56656	0.015423
1983	15497	0.41%	X_5	0.4800%	15507	102	0.000650
1984	15145	-2.27%	X_3	-1.9178%	15200	3003	0.003618
1985	15163	0.12%	X_5	0.04800%	15218	2992	0.003607
1986	15984	5.41%	X_{12}	5.7479%	16035	2556	0.003163
1987	16859	5.47%	X_{12}	5.7479%	16903	1914	0.002595
1988	18150	7.66%	X_{13}	6.4900%	17953	38750	0.010846
1989	18970	4.52%	X_{10}	4.0192%	18879	8193	0.004771
1990	19328	1.89%	X_8	1.7573%	19303	607	0.001275
1991	19337	0.05%	X_5	0.5357%	19432	8938	0.004889
1992	18876	-2.38%	X_3	-1.9178%	18966	8128	0.004776
						MSE=21575	AFER=0.005701

TABLE V
FORECASTING RESULTS OF DIFFERENT FORECASTING MODELS

Year	Enrollmets	Song Chissom [18]	Song Chissom [19]	Chen [1]	Hwang Chen & Lee [7]	Huarng [5]	Chen [2]	Jilani and Burney [8]	Jilani Burney & Ardil [9]	Jilani Burney & Ardil [10]	Proposed Method
1971	13055	–	–	–	–	–	–	–	14464	13579	–
1972	13563	14000	–	14000	–	14000	–	–	14464	13798	13410
1973	13867	14000	–	14000	–	14000	–	–	14464	13798	13932
1974	14696	14000	–	14000	–	14000	14500	14730	14710	14452	14664
1975	15460	15500	14700	15500	–	15500	15500	15615	15606	15373	15423
1976	15311	16000	14800	16000	16260	15500	15500	15614	15606	15373	15847
1977	15603	16000	15400	16000	15511	16000	15500	15611	15606	15623	15580
1978	15861	16000	15500	16000	16003	16000	15500	15611	15606	15883	15877
1979	16807	16000	15500	16000	16261	16000	16500	16484	16470	17079	16773
1980	16919	16813	16800	16833	17407	17500	16500	16476	16470	17079	16897
1981	16388	16813	16200	16833	17119	16000	16500	16469	16470	16497	16341
1982	15433	16789	16400	16833	16188	16000	15500	15609	15606	15737	15671
1983	15497	16000	16800	16000	14833	16000	15500	15614	15606	15737	15507
1984	15145	16000	16400	16000	15497	15500	15500	15612	15606	15024	15200
1985	15163	16000	15500	16000	14745	16000	15500	15609	15606	15024	15218
1986	15984	16000	15500	16000	15163	16000	15500	15606	15606	15883	16035
1987	16859	16000	15500	16000	16384	16000	16500	16477	16470	17079	16903
1988	18150	16813	16800	16833	17659	17500	18500	18482	18473	17991	17953
1989	18970	19000	19300	19000	19150	19000	18500	18481	18473	18802	18879
1990	19328	19000	17800	19000	19770	19000	19500	19158	19155	18994	19303
1991	19337	19000	19300	19000	19928	19500	19500	19155	19155	18994	19432
1992	18876	–	19600	19000	15837	19000	18500	18475	18473	18916	18966
MSE	423027	775687	407507	321418	226611	86694	86694	82269	227194	41426	21575
AFER	3.2238%	4.3800%	3.1100%	3.1169%	2.4452%	1.5294%	1.5294%	1.4064%	2.3865%	1.0242%	0.5701%

[6] K. Huarng, *Effective lengths of intervals to improve forecasting in fuzzy time series*, Fuzzy Sets and Systems, Vol. 12, pp. 387-394, 2001.

[7] J. R. Hwang, S. M. Chen, C. H. Lee, *Handling forecasting problems using fuzzy time series*, Fuzzy Sets and Systems, Vol. 100, pp. 217-228, 1998.

[8] T. A. Jilani, S. M. A. Burney, *M-factor high order fuzzy time series forecasting for road accident data*, In IEEE-IFSA 2007, World Congress, Cancun, Mexico, June 18-21, Forthcoming in Book series Advances in Soft Computing, Springer-Verlag, 2007.

[9] 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. 23, pp.333-338., 2007.

[10] T. A. Jilani, S. M. A. Burney, 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] G. J. Klir, T. A. Folger, *Fuzzy Sets, Uncertainty, and Information*, Prentice-Hall, New Jersey, U.S.A, 1988.

[12] G. J. Klir, B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, New Jersey, U.S.A, 2005.

[13] L. W. Lee, L. W. Wang, S. M. Chen, *Handling forecasting problems based on two-factors high-order time series*, IEEE Transactions on Fuzzy Systems, Vol. 14, No. 3, pp.468-477, 2006.

[14] H. Li, R. Kozma, *A dynamic neural network method for time series prediction using the KIII model*, Proceedings of the 2003 International Joint Conference on Neural Networks, 1: 347-352, 2003.

[15] S. Melike, K. Y. Degtiarev, *Forecasting Enrollment Model Based on First-Order Fuzzy Time Series*, Proceedings of World Academy of Science, Engineering and Technology, Vol. 1, pp. 1307-6884, 2005.

[16] S. Melike, Y. D. Konstsntin, *Forecasting enrollment model based on*

first-order fuzzy time series, in proc. International Conference on Computational Intelligence, Istanbul, Turkey, 2004.
[17] Q. Song, B. S. Chissom, *Fuzzy time series and its models*, Fuzzy Sets and Systems, Vol. 54, pp. 269-277, 1993.
[18] Q. Song, B. S. Chissom, *Forecasting enrollments with fuzzy time series Part I*, Fuzzy Sets and Systems, 54: 1-9.
[19] Q. Song, B. S. Chissom, *Forecasting enrollments with fuzzy time series: Part II*, Fuzzy Sets and Systems, Vol. 62: pp. 1-8, 1994.