Fuzzy Time Series Forecasting of Low Dimensional Numerical Data

Devendra Tayal[1], Shilpa Sonawani[2], Gunjan Ansari[3], Charu Gupta[4,*]

1. Associate Professor & Head of Department, Deptt. Of Computer Engg., IGIT, Delhi.
2. Assistant Professor, Deptt. Of Computer Engg., MIT, Pune.
3. Assistant Professor, Deptt. Of Computer Engg., JSSATE, Noida (U.P)
4. Assistant Professor, Deptt. Of Computer Engg., BPIT, Delhi.

Abstract: Various classical techniques such as linear regression, nearest neighbor have been used in developing predictive models in the past. But the methodologies developed using fuzzy time series includes a wide array of work that requires special attention. The time series analysis has been of great importance to engineering and economy problems. In this paper, we present a brief summary of the various infamous methodologies available in the literature for forecasting of numerical data using fuzzy time series that includes stock prediction, temperature prediction, foreign exchange daily price estimate, crop production, educational enrollments forecasting, inventory demand and also a brief mention of the limitations of fuzzy time series.

Keywords: Fuzzy time series, Forecasting, Fuzzy time series limitations.

1. Introduction

The idea of fuzzy was born in 1965 when Lofti A. Zadeh, a well- respected professor in the department of electrical engineering and computer science at University of California, Berkeley, believed that all real world problems could be solved with efficient and analytical methods. In his words, *"We need a radically different kind of mathematics, the mathematics of fuzzy or cloudy quantities which are not described in terms of probability distributions".* Initially, the concept of fuzzy sets encountered sharp criticism from the academic community. Owing to this, fuzzy logic grew in both wide and comprehensive way and established the foundation of fuzzy logic technology and led to the development of application of this technology in the following years [21]. In addition to maintain the usage legacy of fuzzy logic, fuzzy time series was used for forecasting low dimensional numerical data. Till date the research is flourishing in order to achieve a more flexible and more powerful methods and structures to solve real world problems that deal with high amount of uncertainty that grow due to the dynamics of the environment.In this paper, a brief survey of the use of fuzzy time series is highlighted. A concise description of the various methods has been provided. Section 2 provides a quick view of the fuzzy time series principles. Section 3 explains the various infamous methodologies that have been used for forecasting numerical low dimensional data using fuzzy time series. Section 4 outlines the limitations of fuzzy time series and finally in section 5 we conclude the results.

2. Fuzzy Time Series

In view of making the exposition self-contained, the various definitions and properties of fuzzy time series forecasting found in [14,15,16] are summarized and reproduced as:

Definition 1. A fuzzy set is a class of objects with a continuum of grade of membership. Let U be the Universe of discourse with U= $\{u_1, u_2, u_3 \ldots u_n\}$,where u_i are possible linguistic values of U, then a fuzzy set of linguistic variables A_i of U is defined by

$$A_i = \mu_{Ai}(u_1)/u_1 + \mu_{Ai}(u_2)/u_2 + \mu_{Ai}(u_3)/u_3 + \ldots + \mu_{Ai}(u_n)/u_n \quad (1)$$

here μ_{Ai} is the membership function of the fuzzy set A_i, such that $\mu_{Ai} : U = [0, 1]$.

If u_j is the member of A_i, then $\mu_{Ai}(u_j)$ is the degree of belonging of u_j to A_i.

Definition 2. Let Y(t) (t = . . ., 0, 1, 2, 3, . . .), is a subset of R, be the universe of discourse on which fuzzy sets $f_i(t)$ (i=1, 2, 3, . . .) are defined and F(t) is the collection of f_i, then F(t) is defined as fuzzy time series on Y(t).

Definition 3. Suppose F(t) is caused only by F(t−1) and is denoted by F(t−1)→F(t); then there is a fuzzy relationship between F(t) and F(t−1) and can be expressed as the fuzzy relational equation:

$$F(t) = F(t-1) \circ R(t, t-1) \qquad (2)$$

here " ∘ " is max–min composition operator. The relation R is called first-order model of F(t). Further, if fuzzy relation R(t, t−1) of F(t) is independent of time t, that is to say for different times t_1 and t_2, $R(t_1,t_1-1) = R(t_2, t_2-1)$, then F(t) is called a time invariant fuzzy time series.

Definition 4. If F(t) is caused by more fuzzy sets, F(t−n), F(t−n + 1), . . ., F(t−1), the fuzzy relationship is represented by

$$A_{i1}, A_{i2} \ldots A_{in} \rightarrow A_j$$

here $F(t-n) = A_{i1}$, $F(t-n+1) = A_{i2}, \ldots, F(t-1) = A_{in}$. This relationship is called n^{th} order fuzzy time series model.

Definition 5. Suppose F(t) is caused by an F(t−1), F(t−2), . . ., and F(t−m) (m> 0) simultaneously and the relations are time variant. The F(t) is said to be time variant fuzzy time series and the relation can be expressed as the fuzzy relational equation:

$$F(t) = F(t-1) \circ R^w(t, t-1) \qquad (3)$$

here w > 1 is a time (number of years) parameter by which the forecast F(t) is being affected. Various complicated computational methods are available to for the computations of the relation $R^w(t, t-1)$.

3. Forecasting Using Fuzzy Time Series

The area of forecasting numerical data has been explored with forecasting stock, temperature, foreign exchange daily price, crop production, educational enrollments (Song & Chissom [16]; Sullivan & Woodall [15]; Chen [7] ; Hwang, Chen & Lee [12]; Chen & Hwang [8]; Huarng [11]; Chen[9]; Chen & Hsu [4,6]; Lee, Wang, Chen & Leu (2006); Chen & Chung [5]; Huarng & Yu [10]; Singh [14]; Xihio & Yimin [19]; Zarandi, Molladavoudi & Beigi Ali [20]; Tanuwijaya & Chen [17]; Chen, Wang & Pan [3]; Chen & Chen [3]; Rajaram & Sakthisree [13]). The properties, definitions of fuzzy time series are defined in Zadeh [21]; Wu & W [18], Song & Chissom [16].

Using fuzzy time series, Song & Chissom [16] proposed two divisions of time series, namely time-invariant and time-variant fuzzy time series that worked on linguistic values. They emphasized the fact that with traditional forecasting methods it is difficult to work with observations that consists of linguistic values. The fuzzy time series model was applied to forecast the enrollments of the University of Alabama that deals with fuzzifying the historical data, then developing a fuzzy time series model and finally calculating and interpreting the results. The fact that while dealing with forecasting based on fuzzy time series, the length of intervals are of utmost importance that greatly affect the forecasting accuracy rate (Chen & Chung [5]). Later on, a time-invariant Markov model was proposed that used linguistic labels with probability distributions (Sullivan & Woodall [15]). The proposed approach was compared with time variant and time invariant fuzzy time series and gave better results. Hwang, Chen, Lee [12] in 1998 presented an approach of forecasting university enrollments using fuzzy time series and the concept of window basis. The method was based on the concept, that variation of enrollment of the current year is related to the trend in enrollments of the past years. Furthermore, a new approach used arithmetic operations (Chen [9]) that simplified the complicated max min composition operator presented by Song and Chissom, 1993-1994. The method made good forecast on the enrollments of the University of Alabama and also provided a robust way of forecasting even when the data is not accurate.

In time series analysis, length of interval plays a vital role in determining the effectiveness of forecasting. The issue of determining the effective length of intervals was studied by Huarng [11] in 2001. He pointed out that length of intervals in the universe of discourse greatly affect the forecasting accuracy rate. He also stated that, proper choice of the length of intervals can greatly improve the forecasting accuracy results. So he presented a distribution- and average-length based method to deal with forecasting problem. Working in the area of variation in the interval length (Huarng, [11]), Tanuwijaya and Chen [17], 2009 presented a new method of forecasting, that partitioned the universe of discourse into different length of intervals. A method for forecasting enrollments, inventory demand, and the Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) based on ratio based lengths of intervals was also presented to improve the forecasting accuracy rate (Huarng and Yu, [10]).

Chen and Hsu in 2004 presented a method to forecast the enrollments of the University of Alabama based on first order time-variant fuzzy time series model along with Lee, Wang, Chen & Leu (2004). The results of forecasting using first order fuzzy time series were not good enough. So, Chen [7] in 2002 presented a high order fuzzy time series forecast model to overcome the drawbacks of loose bound on the data. Working on the same front, Chen and Hsu [4] presented method for forecasting the enrollments of the University of Alabama using high order fuzzy time series. To determine the forecasting trend, the proposed method used second order differences of the previous year's data.

The impreciseness and uncertainty is best modeled using a fuzzy approach as it can easily accommodate the ambiguities of the real world and logic. The time series assumes that there is a functional relationship with the current and past values. This assumption has been put to use in the methodologies developed so far.

4. Limitation of Fuzzy Time Series

Time series also suffers from some limitations that are under consideration and require a mention. A disadvantage with time series analysis is that the time series should be converted to stationary and periodic series in order to analyse it and should be examined in a phase space in order to get interesting pattern from it [1]. R.J. Povinelli in 1999 studied time series with applications such as earthquake prediction, sharp fall of stock prices. In the first step, the time series was transformed to phase space and different event characterization functions were selected for each application. Despite the novelty, several problems limited its application such as:

(i) The temporal patterns were defined by rigid regions that were hard to adjust when there is noise in phase space,
(ii) It often generated false-positive prediction,
(iii) It also incurred high computational complexity and lacked stability.

5. Conclusion

Time series enjoyed fruitful applications in forecasting educational enrollments, stock prediction, foreign exchange price estimation, temperature, agricultural production etc. [1,3,4,8,14,20]. In spite of tremendous work done in time series, researchers are constantly focussing on other methods or techniques that can be used as an alternative for forecasting. The hunger for developing an efficient and robust tool is the motivation behind the on-going work. Conclusively this paper provides an outline and concise description of the work done using fuzzy time series for forecasting both algebraic (such as enrollment data) and exponential (such as inventory demand data) data.

To sum up, not only the time variant and time invariant analysis affect the forecasting of data but the length of intervals (for dividing the universe of discourse) also have a vital role in the analysis of the problem under consideration. With the tremendous work done, there is also a scope of improvement in the way the intervals are formed. Fuzzy interval division is one such aspect that can be investigated more deeply for achieving a high amount of accuracy in forecasting. Innovative approach to reveal the time ordered structure, called temporal patterns can be identified.

In time series analysis, event characterization function is employed that correlates the occurrence of a future event with the occurrence of some event in the past. The variation of temporal patterns with effect to the noise also degrades the stability. The approach that can be put to use before determining the event characterization function is to select the appropriate function that can model the behaviour of a non-linear real world problem in a more specific manner.

References

1. Aydin, I., Karakose, M. and Akin, E., 2009. The prediction algorithm based on fuzzy logic using time series data mining model. World Academy of Science, Engineering and Technology, 51.

2. Chen, S. M. and Chen, C.D., 2010. Handling forecasting problems based on high-order fuzzy logical relationships. Expert Systems with Applications, 38(4), 3857-3864.

3. Chen, S.M., Wang, N.Y. and Pan, J.S., 2009. Forecasting enrollments using automatic clustering techniques and fuzzy

logical relationships. Expert Systems with Applications, 36(8).

4. Chen, S.M. and Hsu, C.C., 2008. A new approach for handling forecasting problems using high order fuzzy time series. Intelligent Automation and Soft Computing, 14(1), 29-43.

5. Chen, S.M. and Chung, N.Y., 2006. Forecasting enrollments using high order fuzzy time series and genetic algorithms. International Journal of Intelligent Systems, 21(5), 485-501.

6. Chen, S.M. and Hsu, C.C., 2004. A new method to forecast enrollments using fuzzy time series. International Journal of Applied Science and Engineering, 2(3), 234–244.

7. Chen S.M., 2002. Forecasting enrollments based on high-order fuzzy time series. Cybernetics and Systems, 33(1), 1–16.

8. Chen, S. M. and Hwang, J. R., 2000. Temperature prediction using fuzzy time series. IEEE Transactions on systems, man, and cybernetics—Part B: cybernetics, 30(2), 263–275.

9. Chen, S. M., 1996. Forecasting enrollments based on fuzzy time series. Fuzzy Sets and Systems, 81(3), 311–319.

10. Huarng, K. and Yu, H.K., 2006. Ratio-based lengths of intervals to improve fuzzy time series forecasting. IEEE Transactions on systems, man, and cybernetics—Part B: cybernetics, 36(2), 328–340.

11. Huarng, K., 2001. Effective lengths of intervals to improve forecasting in fuzzy time series. Fuzzy Sets and Systems, 123(3), 387–394.

12. Hwang, J.R. , Chen, S.M. and Lee, C.H., 1998. Handling forecasting problems using fuzzy time series. Fuzzy Sets and Systems, 100(2), 217–228.

13. Rajaram, S. and Sakthisree, P., 2010. Forecasting enrollments using interval length, Weightage factor and fuzzy logical relationships. International Journal of Statistics and Systems, 5(2), 173-182.

14. Singh, S. R . , 2008. A computational method of forecasting based on fuzzy time series. Mathematics and Computers in Simulation, 79, 539–554.

15. Sullivan, J. and Woodall, W.H., 1994. A comparison of fuzzy forecasting and Markov modeling. Fuzzy Sets and Systems, 64(3), 279–293.

16. Song, Q. and Chissom, B.S., 1993. Forecasting enrollments with fuzzy time series—Part I. Fuzzy Sets and Systems, 54(1), 1–9.

17. Tanuwijaya, K. and Chen, S.M., 2009. A new method to forecast enrollments using fuzzy time series and clustering techniques. In proceedings of 8th International Conference on Machine learning and Cybernetics, Baoding, China.

18. Wu and W., 1986, Fuzzy reasoning and fuzzy relation equations. Fuzzy sets and Systems, 20, 67-68.

19. Xihao, S. and Yimin, L.,2008. Average-based fuzzy time series models for forecasting Shanghai compound index. World Journal of Modelling and Simulation, 4(2), 104-111.

20. Zarandi, M.H.F., Molladavoudi, A. and Beigi Ali, M.H., 2009. A new method for temperature prediction and the TAIFEX Forecasting based on fuzzy logical relationship and double interval division. Industrial Engineering and Engineering Management, IEEM, 1, 1543-1547.

21. Zadeh, L.A., 1975. The concept of linguistic variable and its application to a approximate reasoning. Information Sciences, 8, 199-249.