DEVELOPING DEFUZZIFYING METHOD OF FUZZY TIME-VARIANT SERIES FOR FORECASTING PRODUCT DEMAND

SAEED AKBARNATAJBISHEH

A thesis submitted in fulfilment of the

requirements for the award of the degree of

Master of Engineering (Industrial Engineering)

Faculty of Mechanical Engineering

Universiti Teknologi Malaysia

JUNE 2013

It is my honor to dedicate this thesis to my kind parents who have always been my motivator and supporter in each step of my life. I also dedicate this work respectfully to my elder brother, Abed, and my dear sister, Yeganeh.

ACKNOWLEDGEMENT

All praises are to Allah Al-Mighty for his grace and blessing and bestowing me with health and opportunity to gain this treasure of knowledge. Apart from the efforts of author, the success of any research project depends largely on the encouragement and guidelines of many others. I take this opportunity to officially express my gratitude to the people who have been instrumental in the successful completion of this thesis.

Foremost, I would like to express my sincere gratitude to my supervisor, Dr. Syed Ahmad Helmi Bin Syed Hassan for the continuous support of my Master study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of the thesis. Besides my supervisor, I would like to appreciate one of my best friends, Seyed Navid Seyedi for his assistance and comments on some parts of this thesis.

Last but not the least; I would like to thank parents for their selfless love and their supports throughout my life.

ABSTRACT

This thesis is about proposing a method of defuzzifying fuzzy time-variant series on forecasting demand by developing time-variant model considering Mean Absolute Error (MAE) and trend of primary forecasting (MAE&Trend). The previous method of defuzzifying was based on Artificial Neural Network (ANN). Defuzzifying by ANN needs model identification which is time consuming. To justify the robustness of the proposed method, it is compared with Song's and Chen's fuzzy time-invariant methods as well as Autoregressive Integrated Moving Average (ARIMA) time series method. Advantage of fuzzy method is its robustness in covering large range of data including linguistic data as well as numerical and its ability to forecast by small amount of data. Accuracy of the methods is achieved by comparing Mean Absolute Percentage Error (MAPE). The MAPE results are modified by omitting outranged data to have more logical evaluations. Based on the evaluations, the proposed model of MAE&Trend shows better accuracy in forecasting demand.

ABSTRAK

Kajian ini mengkaji kaedah defuzzifying siri masa fuzzy varian untuk meramal permintaan dengan menghasilkan model siri masa yang mempertimbangkan purata ralat mutlak (MAE) dan dan corak ramalan utama (MAE&Trend). Kaedah sebelum ini menggunakan artificial neural network (ANN). Namun begitu, defuzzifying menggunakan ANN memerlukan proses pengenal-pastian model dan ini memakan masa. Bagi mengenalpasti keteguhan model yang dikaji, ia dibandingkan dengan siri masa fuzzy invariant oleh Song dan Chen dan Autoregressive Integrated Moving Average (ARIMA) model. Kelebihan kaedah fuzzy ialah ia mampu meramal untuk pelbagai jenis data termasuk data nombor dan data bertulis dan ia mampu meramal walaupun data input adalah sedikit. Ketepatan kaedah yang dikaji dicapai dengan membandingkan purata peratusan ralat mutlak (MAPE). Keputusan MAPE diubahsui dengan mengecualikan data di luar lingkungan untuk memperoleh penilaian yang lebih logical. Berdasarkan keputusan MAPE, model yang dikaji, iaitu (MAE& Trend) menunjukkan ketepatan yang lebih baik dalam meramal permintaan.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ANN	Artificial Neural Network
ARIMA	Autoregressive Integrated Moving Average
ACF	Autocorrelation Function
PACF	Partial Autocorrelation Function

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

.

1.1 Introduction

In this chapter firstly background of study is explained which is mostly about anticipating future demand with historical data. Secondly, potential problems are stated which is generally about limitations of using neural network for defuzzifying fuzzy time-variant method. In addition, reaching new way of defuzzifying fuzzy time-variant method with mean absolute error (MAE) and trend of primary forecasting with better accuracy is stated for the objective. After defining scope of this thesis, a brief summary of literature as well as conceptual framework and significance of study is explained. Finally, in thesis organization, the overview of each chapter is given.

1.2 Background

Predicting future demand is one of the most crucial concerns of managements in a company. Moreover, accuracy of the forecast and adapting the prediction to the real world play a vital role in forecasting. To have more accurate forecast, more historical data should be gathered. Since data ,in this case, are categorized into

linguistic and numerical data the most useful approach to deal with data is fuzzy logic (Song and Chissom, 1993b) which is able to use both types of data as an input and produce fuzzy data as output. Fuzzy time series are classified into two categories, fuzzy time-invariant and fuzzy time-variant series(Song and Chissom, 1993b). Defuzzifying of fuzzy time-invariant series is applied by two methods, Song's method(Song and Chissom, 1993a) and Chen's method(Chen, 1996); but the only defuzzification method for fuzzy time-variant series is by neural network (Song and Chissom, 1994). Neural network uses fuzzy memberships as input as well as actual data for target values. The major disadvantage of neural network is the process of model identification. It means that you may not reach the best result in the first train. Hence, it is proposed a new way of defuzzification for fuzzy time-variant series to cope with this problem.

Forecasting by various methods gives us the chance to evaluate that which method has better accuracy. ARIMA method is selected to be compared with the proposed method because of its robustness and dealing with demand fluctuation over time (Wang, 2011). It is also able to cover time series with seasonality and nonseasonality. Making the best decision among results could be achieved by suitable comparison between methods. Hence, the results of methods would be compared by MAPE (mean absolute percentage error). Output of MAPE is percentage of error which is difference between actual data and forecasted data.

1.3 Problem Statement

Defuzzification of time-variant model was by applying neural network so far. One of the neural network disadvantages in defuzzifying is that it needs model identification. Identifying the best model in neural network is time consuming. In addition, MAPE of defuzzifying by ANN in time-variant series is higher than time-invariant series.

In this paper, reaching a method to defuzzifying time-variant series is targeted that lead to less error in comparison with ANN. Dealing with this problem it is proposed new defuzzification method based on MAE and trend of forecasting. Considering MAE as well as trend of time series would affect positively on accuracy that means reaching less error.

1.4 Objective

In this project, it is proposed a new way of defuzzifying fuzzy time-variant method which firstly proposed by Song and Chissom (1994). The previous method has utilized neural network for defuzzifying that has some limitations and disadvantageous such as model identification which would be explained in further sections. Proposed defuzzification method of fuzzy time-variant method is developed by consideration of MAE and trend of primary forecasting.

Comparison of proposed model by ARIMA and previous method of defuzification would be applied by MAPE to determine the accuracy of proposed model.

1.5 Scope

The scope of this project is as follow:

a) Data has been collected in paint manufacturing company "BINA Integrated".
b) Fuzzy time series and ARIMA model are applied on historical data.
c) MATLAB and MINITAB software is used for getting results.

1.6 Summary of Literature

Fuzzy time series was proposed in 1993 by Song et al. as "a special dynamic process with linguistic values" (Song and Chissom, 1993b) that is able to forecast future by use of fuzzy sets and variables. Two main categories of fuzzy time series are fuzzy time-invariant and fuzzy time-variant series: In fuzzy time-invariant series, for example, forecasting month i is done by use of data of $i-1$, but in fuzzy time-variant series sets of data $(1,2,\dots,i-1)$ is considered for forecasting month i. In addition defuzzifying methods of these two categories are different; two methods of defuzzifying fuzzy time-invariant series, Song's method and Chen's method, are defined in this project as well as one method of defuzzifying fuzzy time-variant series which is defuzzification by neural network. Neural network is a system, in which an input is presented to this system, and a corresponding desired response is set at the output, an error is extracted from the difference of the desired response and the real system output.

One of the issues in forecasting is comparison various methods to determine that which one has the best accuracy or lower MAPE (mean absolute percentage error); hence, another method which is considered to be compared with fuzzy time series, is ARIMA (Autoregressive integrated moving average) model. This method is proposed by Box and Jenkins (1970) and it is a linear combination of pervious points (past observations or historical data) and random errors.(GE, 1970)

1.7 Conceptual Framework

The proposed method starts by defining parameter of time-variant method (w); then fuzzy membership of fuzzy time-variant is extracted. Fuzzy memberships are used to calculating primary forecasting (PF_i) by applying defuzzifying rules of time-invariant (Song's method). Next, calculate mean absolute error (MAE) of time-invariant forecasting in period of w. Trend (T_r) of the primary forecasting is another

factor which is considered. Finally, the results are defuzzified values (forecasting) for time-variant method by multiplying PF_i, T_r and MAE.

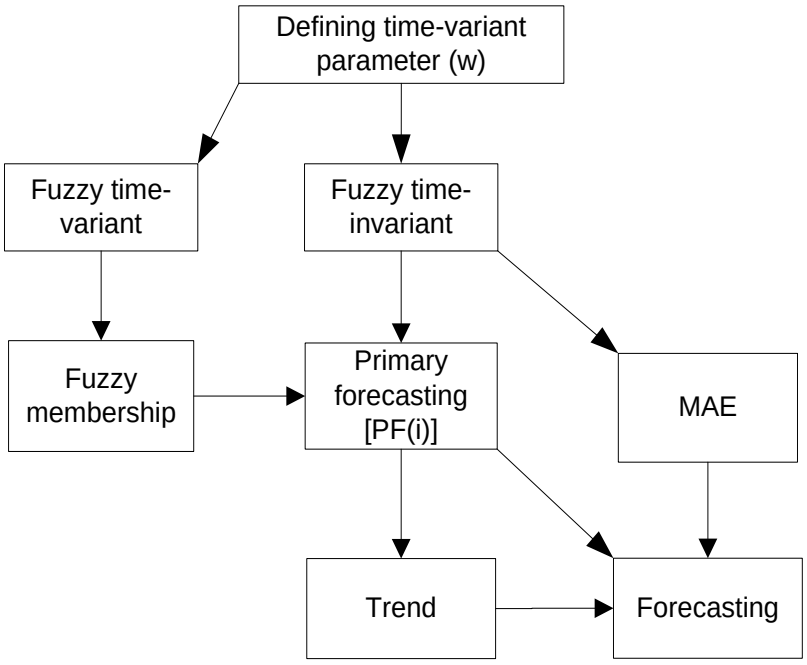

Figure 1.1: Conceptual Framework

1.8 Significance of the Study

This study is mainly determined some peroperties of fuzzy time variant method and proposed a new defuzzification method of fuzzy time-variant series. The proposed model is based on mean absolute error (MAE) and trend of primary forecasting while the previous method of defuzzifying fuzzy time-variant was by use of neural network. In addition, it is found that there is a limitation of defining parameter of time-variant method (w), as well as the easier way of calculating R (matrix) which lead to shorten the time of calculation in fuzzy forecasting. For calculating MAPE of each method, it is asserted that if outranged data be omitted, the comparison is more logical; because by omitting outranged data, diverse effect of external factors are reduced.

1.9 Overview and Conclusion

This thesis is inclusive of 5 chapters; the first one is generally gives an overview of project; in this chapter introduction and background of project are stated then set the objective due to stated problem. The objective is proposing new method of defuzzification for fuzzy time-variant series. To manage the project, some methods and software for the scope of this thesis are considered. Significance of study and theoretical framework are brought to give us the image of what have been done in this thesis.

Chapter 2 is about reviewing the literatures for the topic; in this part, background of project is explained more and definitions and application of each method is clarified. Some advantages and disadvantages of each method are discussed according to related papers.

In Chapter 3 the properties of the proposed model are explained. Then the methodology of applying new way of defuzzification for fuzzy time-variant series is clarified. Finally, the result of applying proposed model for enrollment for University of Alabama are compared with neural network defuzzification method.

Chapter 4 is dedicated to data collection and analysis. Demand of paint manufacturing company is gathered to justify the robustness of the proposed method. Since this data is fluctuated, performance of any methods would be measured by comparing the MAPE of results.

In the last chapter, results of each method are calculated. In addition, some constraints of fuzzy time series such as determining w (parameter of time-variant method) are defined. To reduce the diverse effect of external factors outranged data are omitted. The meaning of outranged data and corresponding reasons are stated.

Finally results of methods are compared by MAPE (mean absolute percentage error) in which the lower MAPE of method illustrates its better accuracy in forecasting.

1.10 Conclusion

The overall view of this thesis is mentioned in the first chapter. Defining objective and problem statement are two important parts of this chapter. The conceptual framework shows then map of proposed method of defuzzifying time-variant series. In next chapter definitions of each method as well as literature review will explain more to clarify application of methods.

REFERENCES

Azadeh, A., Saberi, M., Asadzadeh, S. & Khakestani, M. 2011. A hybrid fuzzy mathematical programming-design of experiment framework for improvement of energy consumption estimation with small data sets and uncertainty: The cases of USA, Canada, Singapore, Pakistan and Iran. *Energy*.

Chen, S. M. 1996. Forecasting enrollments based on fuzzy time series. *Fuzzy sets and Systems,* 81**,** 311-319.

Chu, H. H., Chen, T. L., Cheng, C. H. & Huang, C. C. 2009. Fuzzy dual-factor time-series for stock index forecasting. *Expert Systems with Applications*, 36**,** 165-171.

Day, M. 1995. Research of Applying Genetic Algorithms to Fuzzy Forecasting—Focus on Sales Forecasting. MS thesis, Tamkang Univ., Taipei, Taiwan, ROC.

Du, T. C.-T. & Wolfe, P. M. 1997. Implementation of fuzzy logic systems and neural networks in industry. *Computers in industry,* 32**,** 261-272.

Ediger, V. Ş. & Akar, S. 2007. ARIMA forecasting of primary energy demand by fuel in Turkey. *Energy Policy,* 35**,** 1701-1708.

Ge, P. 1970. Box., and G. M, Jenkins.,"Time series analysis, forecasting and control". San Francisco, CA: Holden Day.

Hong, W. C., Dong, Y., Chen, L. Y. & Wei, S. Y. 2011. SVR with hybrid chaotic genetic algorithms for tourism demand forecasting. *Applied Soft Computing,* 11**,** 1881-1890.

Huarng, K. 2001. Heuristic models of fuzzy time series for forecasting. *Fuzzy sets and Systems,* 123**,** 369-386.

Hwang, J. R., Chen, S. M. & Lee, C. H. 1998. Handling forecasting problems using fuzzy time series. *Fuzzy sets and Systems,* 100**,** 217-228.

Jilani, T. A. & Burney, S. M. A. 2008. A refined fuzzy time series model for stock market forecasting. *Physica A: Statistical Mechanics and its Applications,* 387**,** 2857-2862.

Klir, G. J., St Clair, U. & Yuan, B. 1997. *Fuzzy set theory: foundations and applications*, Prentice-Hall, Inc.

Li, S. T. & Cheng, Y. C. 2007. Deterministic fuzzy time series model for forecasting enrollments. *Computers & Mathematics with Applications,* 53**,** 1904-1920.

Liu, G. Q. 2011. Comparison of regression and ARIMA models with neural network models to forecast the daily streamflow of White Clay Creek, University of Delaware.

Morley, C. L. 1993. Forecasting tourism demand using extrapolative time series methods. *Journal of Tourism Studies,* 4**,** 19-25.

Negnevitsky, M. 2005. Artificial intelligence: a guide to intelligent systems, Addison-Wesley Longman.

Nihan, N. L. & Holmesland, K. O. 1980. Use of the Box and Jenkins time series technique in traffic forecasting. *Transportation,* 9**,** 125-143.

Pareto, V. & Page, A. N. 1971. Translation of Manuale di economia politica ("Manual of political economy"). *AM Kelley, ISBN,* 1037475820.

Qiu, W., Liu, X. & Li, H. 2011. A generalized method for forecasting based on fuzzy time series. *Expert Systems with Applications,* 38**,** 10446-10453.

Shi, J., Guo, J. & Zheng, S. 2012. Evaluation of hybrid forecasting approaches for wind speed and power generation time series. *Renewable and Sustainable Energy Reviews,* 16**,** 3471-3480.

Singh, S. 2007a. A robust method of forecasting based on fuzzy time series. *Applied mathematics and computation,* 188**,** 472-484.

Singh, S. 2007b. A simple method of forecasting based on fuzzy time series. *Applied mathematics and computation,* 186**,** 330-339.

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

Song, Q. & Bortolan, G. 1994. Some properties of defuzzification neural networks. *Fuzzy sets and Systems,* 61**,** 83-89.

Song, Q. & Chissom, B. S. 1993a. Forecasting enrollments with fuzzy time series—part I. *Fuzzy sets and Systems,* 54**,** 1-9.

Song, Q. & Chissom, B. S. 1993b. Fuzzy time series and its models. *Fuzzy sets and Systems,* 54**,** 269-277.

Song, Q. & Chissom, B. S. 1994. Forecasting enrollments with fuzzy time series—part II. *Fuzzy sets and Systems,* 62**,** 1-8.

Sullivan, J. & Woodall, W. H. 1994. A comparison of fuzzy forecasting and Markov modeling. *Fuzzy sets and Systems,* 64**,** 279-293.

Tsaur, R. C., O Yang, J. C. & Wang, H. F. 2005. Fuzzy relation analysis in fuzzy time series model. *Computers & Mathematics with Applications,* 49**,** 539-548.

Wang, C. C. 2011. A comparison study between fuzzy time series model and ARIMA model for forecasting Taiwan export. *Expert Systems with Applications,* 38**,** 9296-9304.

Yu, H. K. 2005. Weighted fuzzy time series models for TAIEX forecasting. *Physica A: Statistical Mechanics and its Applications,* 349**,** 609-624.

Zadeh, L. A. 1965. - 8**,** - 353.

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

Zhang, G., Eddy Patuwo, B. & Y Hu, M. 1998. Forecasting with artificial neural networks:: The state of the art. *International Journal of Forecasting,* 14**,** 35-62.

Zhang, G. P. 2003. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing,* 50**,** 159-175.