A HYBRID BOX-JENKINS AND DECOMPOSITION MODEL FOR DROUGHT FORECASTING IN KUALA TERENGGANU

HO MEE CHYONG

UNIVERSITI TEKNOLOGI MALAYSIA

A HYBRID BOX-JENKINS AND DECOMPOSITION MODEL FOR DROUGHT FORECASTING IN KUALA TERENGGANU

HO MEE CHYONG

A dissertation submitted in partial fulfillment of the
requirement for the award of the degree of
Master of Science (Mathematics)

Faculty of Science
Universiti Teknologi Malaysia

JANUARY 2013

To my beloved family and my dear soul mate

ACKNOWLEDGEMENT

First and foremost, I would like to grant my deepest gratitude to my supervisor, Dr. Ani Shabri who had gave me a lot of guidance and helpful suggestions throughout this project. He encouraged, helped and guided me at any time needed.

Secondly, I would like to extend my sincere appreciation to my friends especially my dear soul mate, Gabriel Ling who is kindly giving his suggestions, comments, and supports. Their suggestions, comments, and supports were invaluable to me.

Thirdly, I would like to express my greatest gratefulness to my beloved family for their moral supports. Their unstinting supports help me a lot in completing this project.

Lastly, I would like to thank to Data Information Unit, Water Resource Management and Hydrology Division (D.I.D) for providing me the monthly rainfall data of the rain gauge station, Setor JPS Kuala Terengganu in Kuala Terengganu as this study would never be completed without the provided data.

ABSTRACT

Drought is a global phenomenon which adversely affects the sustainability of one nation which encompasses three prominent aspects such as economic, social and environmental. Due to that, it has immensely attracted the awareness of environmentalists, ecologists, hydrologists, meteorologists, geologists and agricultural scientists. Therefore, drought forecasting is essential for several key players particularly the governments to evaluate the drought occurrence in order to give early warning for preparedness and mitigation measures. In this study, a hybrid Box-Jenkins and decomposition model based on standardized precipitation index (SPI) was developed to forecast drought in Kuala Terengganu. Monthly rainfall data of rain gauge station, Setor JPS Kuala Terengganu for period January 1982 to January 2012 was used in this study. Multiplicative decomposition method was employed to identify and isolate the underlying components of SPI time series for multiple time scales using Minitab 16.0. Then the isolated components were gone through the four-step iterative procedure of Box-Jenkins which are identification, estimation, diagnostic checking and forecasting. After that, the forecasted values of components were reassembled in order to gain a forecast based on the time series decomposition. The forecasting performance of the hybrid model was compared with the Box-Jenkins model. Two statistical measurements, mean absolute error (MAE) and mean squared error (MSE) were applied in this study to measure the accuracy of the forecasting models. In brief, the accuracy measure results indicated that the hybrid model can prevail over the Box-Jenkins model.

ABSTRAK

Kemarau adalah satu fenomena global yang boleh menjejaskan sector ekonomi, sosial dan alam sekitar sesebuah negara. Ini telah menarik perhatian ahli profesional seperti ahli alam sekitar, ahli ekologi, ahli hidrologi, ahli meteorologi, ahli geologi dan saintis pertanian. Oleh itu, ramalan kemarau adalah penting bagi beberapa pemain utama khususnya kerajaan untuk menilai kejadian kemarau untuk persediaan awal dan langkah –langkah penyelesaian. Dalam kajian ini, model *hybrid Box-Jenkins and decomposition* berdasarkan *standardized precipitation index* (SPI) telah digunakan untuk meramal kemarau di Kuala Terengganu. Data hujan bulanan daripada stesen tolok hujan, Setor JPS Kuala Terengganu bagi tempoh Januari 1982 hingga Januari 2012 telah digunakan dalam kajian ini. *Multiplicative decomposition* telah digunakan untuk mengenal pasti dan mengasingkan komponen asas siri masa SPI untuk pelbagai skala masa melalui *Minitab 16.0*. kemudian, komponen terpencil itu akan melalui empat prosedur *Box-Jenkins* seperti pengenalan, anggaran, pemeriksaan diagnostik dan ramalan. Selepas itu, komponen yang diramal akan digabung semula supaya nilai ramalan itu adalah berdasarkan model *decomposition*. Prestasi ramalan model *hybrid* telah dibandingkan dengan model *Box-Jenkins* melalui dua ukuran statistic iaitu *mean absolute error* (MAE) dan *mean squared error* (MSE). Secara ringkasnya, keputusan ukuran ketepatan telah menunjukkan bahawa model *hybrid* adalah lebih bagus daripada model *Box-Jenkins* untuk ramalan kemarau di Kuala Terengganu.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Natural disasters can be divided into few categories namely geological disasters, hydrological disasters, meteorological disasters, fires, health disasters and space disasters. Drought is one of the examples of meteorological disasters. Besides, compared to other kinds of disasters, drought is considered more severely and detrimentally affecting people and causing global damage approximately $6-$8 billion annually. (Wilhite, 2000) Therefore, drought can be the world's costliest catastrophe and has riveted the attention of environmentalists, ecologists, hydrologists, meteorologists, geologists and agricultural scientists.

According to Mishra and Singh (2010), definitions of drought can be depicted in manifold ways. For example, the World Meteorological Organization (WMO, 1986) defined drought is an endless and prolonged shortage in precipitation. Besides, UN Secretariat General (1994) defined drought is a natural phenomenon that takes place when the precipitation is below normal the recorded level which has triggered serious hydrological imbalances that adversely affect land resource production systems. Drought as the smallest annual value of daily streamflow is defined by Gumbel (1963). Moreover, the Food and Agriculture Organization (FAO, 1983) of the United Nations defined a drought is perilous as it suffers crops from producing due to water paucity. Last but not least, the encyclopedia of climate and weather (Schneider, 1996) defined a drought as a lengthy period (a season, a year, or several years) of deficient rainfall relative to the statistical multi-year mean for a region.

Wilhite and Glantz stated that there are four categories of drought namely meteorological drought, hydrological drought, agricultural drought and socio-economic drought (Mishra and Singh, 2011). Meteorological drought means the lack of rainfall over a region for a period and gravely affects the biotic element such as flora and fauna. Under given water resources management system, streamflows are inadequate to supply established uses over a period and this induces hydrological drought. The agricultural drought is associated with declining soil moisture over a period and causing crop failure without any reference to surface water resources. The socio-economic drought can be regarded as water resources systems fail to meet the water demands and thus it should be associating drought with the supply and demand of an economic good (water).

A drought's period can be varied where it can be in short or protracted term. Sometimes it may last for few months or more grievously it will take years. Hence, this logically causes devastating effects on agriculture and water supplies which yield a great impact on economic beyond the affected area. A drought that lasting for 4 to 9 months in 1991 in Malacca, Malaysia affected more than 170,000 people comprising an area of $2797 km^2$. Furthermore, $1580 km^2$ suffered a wild fire which $100\ km^2$ were agricultural lands and affected 7200 farmers. The estimated economical loss was RM 7 million. In addition to that, several areas were suffered from water rationing and rice crops were totally wiped out and consequently it forced the public send food supplies in order to mitigate the situation. North Eastern part of Sarawak near Miri region was also involved.

According to Frechtling (2001), forecasting is a process of organizing information about past phenomenon in order to anticipate future. Forecasting can be applied in myriad fields throughout the world in various fields such as sales, financial, weather, agricultural and others. So, undeniably it will also be employed in the field of weather in this study and drought is the goal of forecasting. This chapter will carry out a brief understanding of the main purpose of this study. The outcome of this study is hoped to contribute in Malaysia's meteorological sector and to the body of knowledge in the related areas.

1.2 Background of the Study

Drought can give a significant impact on sustainability which includes economic, social and environmental for both developing and developed countries. Drought can be a hazard and a disaster. Drought is hazardous because it is a natural accident of unpredictable occurrence but of recognised recurrence and simultaneously drought is a disaster because it corresponds to the failure of precipitation regime and thus causes the inadequate water supply for agricultural ecosystems and human activities. Therefore, as Wilhite (2000) stated, it was important to develop prediction tools in order to support early warning for timely implementation of preparedness and mitigation measures.

A number of different indices have been developed to quantify a drought. The standardized precipitation index (SPI) is originally introduced by McKee *et al.* (1993) and it is one of the well-known drought indices. SPI index is one of the simplest indicators used for drought assessment. The use of SPI index gives a number of advantages. Firstly, the SPI index is simple and only based on the amount of precipitation. Secondly, the SPI index can calculate the precipitation deficit on different time scales. Thirdly, the calibration of SPI index can assure the independence from the geographical position as the index is calculated with respect to the average precipitation in the same place and the ability of SPI index to describe both dry and wet periods in the same way. (Rossi *et al.*, 2007)

An accurate drought prediction enables optimal operation of irrigation systems. Many different methods have been proposed to forecast drought. According to Mishra and Singh (2011), Gabriel and Neumann and Torranin were the first researchers applied Markov and regression models in drought forecasting. Besides, using the geometric probability distribution in forecasting of properties of droughts was first attempted by Yevjevich (1967). He defined a drought of *k* years as *k* consecutive years when there are no adequate water resources. In addition, the linear stochastic models, ARIMA and SARIMA were used in drought forecasting which is based on the procedure of model development (Mishra and Desai, 2005). The models were also applied in drought forecasting, standardized precipitation

index (SPI) which it had been tested at the Kansabati river basin in India. (Mishra and Desai, 2005)

Time series models have also been widely applied in scientific, economic and engineering field. Time series models have several advantages over other models as time series models have the systematic search capability for identification, estimation, and diagnostic check for model development. Therefore, the time series approach, a hybrid Box-Jenkins and decomposition model is proposed in this study. And the standardized precipitation index (SPI) is used as the drought quantifying parameter in this study. Kuala Terengganu is proposed as the study area as the eastern areas can be regarded as the driest areas during Southwest Monsoon (SWM) period which is from late May to September. The rainfall station, Setor JPS Kuala Terengganu in Kuala Terengganu is chosen as drought did occur in Kuala Terengganu. According to local historians, about 100 years of "Low They Well" was dug during a period of severe drought. Furthermore, lots of livestock especially buffalo were in slim condition due to the lack of food and water during drought in year of 1998.

1.3 Statement of the Problem

Drought is different from most other natural disasters as it can be developed slowly and has prolonged existence. Drought can cause water shortage due to the lack of precipitation and people need to suffer from water rationing. Agricultural areas are also affected during drought as plants cannot survive due to insufficient water. In addition, the worst drought can also cause people to perish from starvation. In brief, drought has the significant impact on economic, societal and environmental losses. Therefore the forecasting of drought becomes crucial.

The following questions have been explored in this study:

(i) How to forecast drought in Kuala Terengganu?

(ii) Which approach between the Box-Jenkins model and a hybrid Box-Jenkins and decomposition model performs better in drought forecasting?

1.4 Objective of the Study

The objectives of this study are:

(i) To develop a hybrid Box-Jenkins and decomposition model for drought forecasting.

(ii) To compare the performance of the hybrid model with Box-Jenkins model.

1.5 Scope of the Study

This study focuses on the hybrid Box-Jenkins and decomposition model that is applied to forecast drought in Kuala Terengganu based on standardized precipitation index (SPI). Comparison between Box-Jenkins model and the hybrid model was considered in order to provide the high accuracy and reliability in drought forecasting. Monthly rainfall data from the rain gauge station, Setor JPS Kuala Terengganu in Kuala Terengganu from January 1982 to January 2012 was collected. The data was collected from Data Information Unit, Water Resource Management and Hydrology Division.

1.6 Significance of the Study

The significance of study is to investigate the appropriate approach for drought forecasting in Kuala Terengganu by using standardized precipitation index (SPI). Comparison between Box-Jenkins model and the hybrid model can help to provide the high accuracy and reliability in drought forecasting. This study can help government in the risk evaluation of drought occurrence. From this, government can actually render an early warning for timely implementation of preparedness based on the prediction. Hence, the negative impacts of drought can be prevented. Furthermore, good water, agricultural planning and other mitigation measures can be implemented as well.

1.7 Thesis Organization

This thesis contains five chapters. The first chapter consists introduction. Introduction, background of the study, statement of the problem, objectives, scopes and the significances of the study are presented in this chapter.

Then, chapter 2 is the literature review. In this chapter, reviews on the previous researches of models used in drought forecasting, time series models, simulations of precipitation distributions, parameter estimation, goodness-of-fit (GOF) test and forecasting performance evaluation are given.

Chapter 3 concerns the methodology. This chapter discusses three types of distribution namely Gamma distribution, Normal distribution and Beta distributions to fit the monthly rainfall data. Besides, two time series approaches (Box-Jenkins and decomposition) are introduced in chapter III. This chapter also presents forecasting error measurements, i.e., mean absolute error (MAE) and mean square error (MSE) to measure the accuracy of a forecasting model.

Chapter 4 presents the implementation of methodology that had discussed in chapter 3. The test for the best fitted distribution in compute SPIs series for drought forecasting in Terengganu using EasyFit 5.5 software and the SPI result which computed by SPI SL 6.exe is revealed in this chapter. This chapter carries out the implementation of the proposed time series approaches (Box-Jenkins and the hybrid) as well. The performance of these two models on drought forecasting for Kuala Terengganu will also be compared via the forecasting error measurements, MAE and MSE.

Finally, chapter 5 discusses the conclusions and some suggestions for the future research.

REFERENCES

1. Abaurrea, J. and Cebrian,A. C. Drought Analysis Based On A Cluster Poisson Model: Distribution Of The Most Severe Drought, *Journal of the Climate Research*. 2002. 22, 227-235.
2. Ani Shabri. *Time Series Course: Lecture Notes*. Department of Mathematics, Faculty of Science, Universiti Teknologi Malaysia. 2011.
3. Bacanli, U.G., Firat, M. and Dikbas, E.F. Adaptive Neuro-Fuzzy Inference System For Drought Forecasting. *Stoch. Environ. Res. Risk Assess*. 2009. 23, 1143–1154.
4. Cooray T. M. J. A. *Applied Time Series Analysis And Forecasting*. Oxford U.K.: Alpha Science International Ltd. 2008.
5. Durdu, O. F. Application Of Linear Stochastic Models For Drought Forecasting in the Buyuk Menderes River Basin, Western Turkey. *Stoch Environ Res Risk Assess*. 2010. 24: 1145-1162.
6. Fathabadi, A., Gholami, H., Salajeghe, A., Azanivand, H. and Khosravi, H. Drought Forecasting Using Neural Network And Stochastic Models. *Advances in Natural and Applied Sciences.* 2009. 3(2): 137-146.
7. Fernandez, C., Vega, J. A., Fonturbel, T. and Jimenez, E. Streamflow Drought Time Series Forecasting: A Case Study In A Small Watershed In North West Spain. *Stoch. Environ. Res. Risk Assess*. 2009. 23, 1063–1070.
8. Frechtling, D. C. *Forecasting Tourism Demand Methods And Strategies*. Great Britain: MPG Books Ltd, Bodmin, Cornwall. 2001.
9. Guttman, N. B. Accepting The Standardized Precipitation Index: A Calculation Algorithm. *J. Am. Water Resour. Assoc*. 1999. 35 (2), 311–322.
10. Han, P., Wang, P.X., Zhang, S.Y. and Zhu, D. H. Drought Forecasting Based On The Remote Sensing Data Using. ARIMA Models. *Mathematical and Computer Modelling*. 2010. 51 (11–12), 1398–1403.

11. Hughes, B. L and Saunders, M.A. Drought Climatology For Europe. *Int J Climatol*. 2002. 22:1571–1592
12. Leilah, A. A. and Al-Khateeb, S. A. Statistical Analysis Of Wheat Yield Under Drought Conditions. *J. Arid Environ*. 2005. 61, 483–496.
13. Liang, Y.-H. Analyzing And Forecasting The Reliability For Repairable Systems Using The Time Series Decomposition Method. *International Journal of Quality & Reliability Management*. 2011. *28*(3), 317–327.
14. Makridakis, S. A Survey Of Time Series. *International Statistical Review.* 1976. Vol. 44, No.1. 29-70.
15. McKee, T.B., Doesken, N. J. and Kliest., J. The Relationship Of Drought Frequency And Duration To Time Scales. In: *Proceedings of the 8th conference on applied climatology 17–22 January, Anaheim, CA. American Meteorological Society, Boston, MA*. 1993. pp 179–184
16. Mishra, A. K. and Desai, V. R. Drought Forecasting Using Feed Forward Recursive Neural Network. *Int. J. Ecol. Model*. 2006. 198, 127–138.
17. Mishra, A. K. and Desai, V. R. Drought Forecasting Using Stochastic Models. *Stoch Environ Res Risk Assess*. 2005. 19:326–339.
18. Mishra, A. K. and Singh, V. P. A Review Of Drought Concepts. *Journal of Hydrology*. 2010. 391: 202-216.
19. Mishra, A. K. and Singh, V. P. Drought Modeling- A Review. *Journal of Hydrology*. 2011. 403: 157-175.
20. Mishra, A. K., Desai, V. R. and Singh, V. P. Drought Forecasting Using A Hybrid Stochastic And Neural Network Model. *J. Hydrol. Eng., ASCE*. 2007. 12 (6), 626–638.
21. Modarres, R. Streamflow Drought Time Series Forecasting. *Stoch. Environ. Res. Risk Assess.* 2007. 15 (21), 223–233.
22. Mohamed, N., Ahmad, M. H., Ismail, Z. and Suhartono. Double Seasonal ARIMA Model For Forecasting Load Demand. *METEMATIKA*. 2010. 26(2): 217-231.
23. Moreira, E. E., Paulo, A. A., Pereira, L. S. and Mexia, J. T. Analysis Of SPI Drought Class Transitions Using Loglinear Models. *J. Hydrol*. 2006. 331, 349–359.

24. Randsan, N. and Titida, N. ARIMA Model For Forecasting Oil Palm Price. *Proceedings of the 2nd IMT-GT Regional Conference on Mathematics, Statistics and Application Universiti Sains Malaysia*. 2006.
25. Rossi. et. al. *Methods And Tools For Drought Analysis And Management*. Springer. 2007. 29-48.
26. Samah, M. and Abo, E. H. Modeling Sulphur Dioxide Time Series Data Using Parametric And Nonparametric Classical Decomposition Methods. *Journal of Environmental Statistics*. 2011. Vol. 2.
27. Sheffield, J., Goteti, G., Wen, F. and Wood,E.F. A Simulated Soil Moisture Based On Drought Analysis For The United States. *Journal of the Geophysical Research*. 2004. Vol. 109.
28. Temraz, H. K. et. al. Application Of The Decomposition Technique For Forecasting The Load Of A Large Electric Power Network. *IEE*. 1996. Vol. 143. No.1
29. Thom, H. C. S. A Note On Gamma Distribution. *Monthly Weather Rev*. 1958. 86:117–122.
30. Wilhite, D. A. *Drought As A Natural Hazard: Concepts And Definitions*. In: Wilhite DA (ed) Drought : A Global Assesment. Routledge Publishers, London,UK. 2000. Vol. 1, pp3-18
31. Zhang, G. P. Time Series Forecasting Using A Hybrid ARIMA And Neural Network Model. *Neurocomputing*. 2003. *50*, 159–175.