Universiti Teknologi MARA

Outliers in Bilinear Time Series Model

Ibrahim bin Mohamed

Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy
Faculty of
Information Technology and Quantitative Sciences

March 2005

Acknowledgements

The research recorded in this thesis was conducted under the excellent supervision of Dr. Azami Zaharim and Assoc. Prof. Dr. Mohd. Sahar Yahya (between December 2000 and March 2005) and Prof. Dr. Mohammad Said Zainol (between December 2003 and March 2005) at the Faculty of Information Technology and Quantitative Sciences (FTMSK), MARA University of Technology (UiTM), Shah Alam, Malaysia. I would like to express my deepest appreciation to my supervisors for their guidance, continuous encouragement, patience and help throughout this PhD program which contributed to the completion of this thesis.

My heartfelt appreciation goes to my parents, my wife, my brother Haji Yusof and my family for their continuous love, inspiration, motivation, support and prayers for my success.

I would like to record my appreciation to the staff of the FTMSK, notably Assoc. Prof. Dr. Arsmah Ibrahim, Assoc. Prof. Dr. Adnan Ahmad, Prof. Dato' Dr. Mohamad Ali Hassan, Assoc. Prof. Dr. Mohamad Alias Lazim, Prof. Dr. Ishak Ab. Ghani, Assoc. Prof. Dr. Daud Mohamad and Puan Norizam Daud; staff of the ISM and PASUM, Universiti Malaya, especially Assoc. Prof. Dr. Nor Aishah Hamzah, Assoc. Prof. Dr. Daud Yahya, Assoc. Prof. Dr. Abdul Ghapor Hussin, Ng Lee Leng, Yusniza Abdul Razak, Abdul Rahim Sukiman and Budiyah Yeop; and staff of IPS, UiTM, particularly Assoc. Prof. Dr. Noorlaila Ahmad and En. Amad Hamdan Jemiran, for their assistance during my course of study.

I am most grateful to Universiti Malaya for giving me the opportunity to fulfil my ambition.

TABLE OF CONTENTS

		Page
ACKNOWLEGEMENT		iii
LIST OF TABLES		x
LIST OF FIGURES		xii
LIST OF SYMBOLS AND ABBREVIATIONS		xvi
ABSTRACT		xix
1.	**INTRODUCTION**	
1.1	Background	1
1.2	Approaches to time series analysis	2
1.3	Nonlinearity tests	3
1.4	Nonlinear time series models	4
1.5	Various time series models	6
	1.5.1 Linear model	7
	1.5.2 Bilinear model	8
	1.5.3 Exponential autoregressive model	9
	1.5.4 Threshold autoregressive model	9
1.6	Outliers	10
1.7	Statement of problem	11
1.8	Objectives of study	12
1.9	Significance of study	12
1.10	Organization of thesis	13

ABSTRACT

This study has two main objectives: Model building and detection of outliers in BL(1,1,1,1) models in time-domain framework.

In model building, the Box-Jenkins approach was closely followed. In the identification stage, time-domain based nonlinearity tests were considered to distinguish nonlinearity data set from linear time series data. In general, identifying the order of bilinear model including BL(1,1,1,1) model is not possible yet due to the complexity of form taken by the moments of bilinear model. In the estimation stage, the nonlinear least squares method were used to estimate the parameters of BL(1,1,1,1) models. In the diagnostic stage, the residuals were examined to check the adequacy of model. In addition, the Akaike's information criteria, the Akaike's Bayesian information criteria and Schwarz's criterion were used for model comparison purposes.

Outliers exist due to many possibilities such as misrecording, disaster or changes of nature. The occurrence of four types of outliers; the additive outlier (AO), innovational outlier (IO), level change (LC) and temporary change (TC), in data from BL(1,1,1,1) models were considered in the study. Their features were studied so that different patterns caused by each type of outliers were distinguishable. Further, the measures of outlier effect for AO, IO, LC and TC were derived using the least squares method. Their variances were obtained using bootstrap method. The detection of outliers was carried out by examining the maximum value of the standardized statistics of the outlier effects. The outlier detection procedure for identifying the type of outlier at time point *t* was proposed. Simulation studies were carried out to study the performance of the procedure in BL(1,1,1,1) models. It was found out that, in general, the proposed procedure performed well in detecting outliers. As for illustration, the proposed procedure was applied on three hydrological data.

CHAPTER ONE
INTRODUCTION

1.1 BACKGROUND

In most statistical studies, the development of statistical theories and applications involves linear forms. In the classical theory of time series analysis, one assumes that the structure of series can be described by linear models. Examples include the autoregressive (AR) model, moving average (MA) model, autoregressive moving average (ARMA) model and autoregressive integrated moving average (ARIMA) model. A good account of the theory is available, for example in Box and Jenkins [1976], Fuller [1976] and Chatfield [1996]. Tong [1983] highlighted three significant achievements of ARMA models. He pointed out that the availability of the theory of linear difference equation and Gaussian sequences made works on ARMA model simpler. The required computation time had been reduced with the availability of ARMA modeling in most ready-made statistical packages. These models had reasonably been successfully applied in forecasting and control.

However, Tong [1983] also pointed out that ARMA models had some limitations. For example, stationary Gaussian ARMA models were not ideally suited for data exhibiting strong asymmetry, sudden bursts of large amplitude at irregular time points, strong cyclicity or time-irreversibility. Further, not all linear models were adequate for the time series data. The famous Wolfer's sunspot data (Box and Jenkins [1976]) was a good example of a nonlinear data set. By considering the time series plot of the sunspot data, one could observe that "there exist a systematic periodic cycles with the downturn is faster than the upturn. This pattern will never be explained fully by any linear model". Granger and Andersen [1978a] and Rao [1981] among others had shown that the sunspot data was better fitted by nonlinear time series model if compared to linear model. Other nonlinear data sets that had been studied in the literature include the Canadian lynx data (Tong [1983]), the blowfly data (Brillinger *et al.* [1980]), the