UNIVERSITI TEKNOLOGI MARA

LONG MEMORY PROCESS IN HYDROLOGICAL TIME SERIES BASED ON THREE CATCHMENTS IN MALAYSIA

NORAZYANI OMAR

Thesis submitted in fulfillment of the requirements
for the degree of
Master of Science

Faculty of Civil Engineering

October 2008

ABSTRACT

The long memory process in a hydrological time series refers to considerable correlation and dependence between all data in a long time span of observations. Studies of long memory have been attempted in various fields of application such as in financial time series, meteorological time series and hydrological time series.

This research is carried out to investigate the presence of long memory process in hydrological time series, and to estimate the parameter *d* that is the long memory parameter (or fractional differencing parameter), in order to generate the synthetic hydrological time series. The presence of long memory in hydrologic data as detected by Hurst (1951) has enhanced the various estimations and procedures of developing models until today. The presence of long memory process can be characterized in several ways: from autocorrelation structure plots (ACF) and the classical rescaled range (R/S) analysis (also known as the heuristic method); or by statistical tests of long memory such as Lo's modified R/S Statistic and GPH (Geweke and Porter-Hudak) Test (also known as the statistical method). The long memory parameter *d* can be estimated by using R/S analysis and Periodogram method. The analysis of long memory process includes detrending, normalization, deseasonalizing of time series and transformation of data.

In this study, some fundamental properties of long memory that are present in the hydrological time series are presented and have been applied to the daily streamflow time series using Autoregressive Fractional Integrated Moving Average models (ARFIMA) for the purpose of modelling. The class of ARFIMA models introduced by Granger and Joyeux (1980) and Hosking (1981) provides a convenient model for modelling long-term time series data. The synthetic hydrological time series are then generated by incorporating fractional differencing *d*. The comparison between the original and synthetic data is made where the statistical characteristics of the series are observed and compared.

The streamflow series of Sungai Selangor, Sungai Linggi and Sungai Johor were chosen for the research. Several important aspects of stochasticity are discussed namely, trend analysis, normality, seasonality, autocorrelation and lastly long memory process. From R/S analysis and Periodogram method, the parameter of *d* obtained are 0.3601, 0.3597, 0.3605 and 0.3983, 0.3753, 0.3157 for each Sungai Selangor, Sungai Linggi and Sungai Johor. The parameter of *d* obtained is in the range of $0 < d < 0.5$ for all streamflow and this shows that the long memory process is present. This statement is supported by the statistical evidence where the test statistics for each streamflow was significant at 1% level of confidence to reject null hypothesis of no long memory.

The significance of the existence of long memory in the hydrological time series cannot be neglected. The model of long memory process in hydrology is useful for simulating a synthetic series of observations in order to study water resources management procedures and allows one to plan and test the water resources management with respect to many different hydrological scenarios.

ACKNOWLEDGEMENTS

Alhamdulillah, praise to God Almighty Allah, the Most Gracious and the Most Merciful, for with His security and guidance, I am able to complete my studies within the timeline given. It is a pleasure to thank the many people who have made this thesis possible.

It is difficult to overstate my gratitude to my supervisor, Prof. Madya Dr. Ismail Bin Atan for his enthusiasm, inspiration, and great efforts to explain things clearly and simply to me. Throughout my thesis-writing period, he provided encouragement, advice, good teaching, and lots of good ideas. He showed me different ways to approach a research problem and the need to be persistent to accomplish any goal. Without his encouragement and constant guidance I could not possibly have finished this study on time.

I am also grateful to the staff at the Drainage and Irrigation Department, Jalan Ampang, for their assistance and guidance in providing me with the research data and material necessary for my study.

I am also very indebted to my friends for providing me with a fun environment to learn and grow. Tremendous support was given until the completion of my study. I wish to thank my husband, Amir Khomeiny Ruslan, for his support and understanding.

Finally, I wish to thank my beloved family, my father, my mother, my brothers and my little sisters, for the tender, loving and caring environment they have given to me. I dedicate this work to them.

TABLE OF CONTENTS

CHAPTER ONE

INTRODUCTION

1.1 Background of Research

Studies related to the streamflow process are important part of hydrological studies. This study is not just associated to the global environmental issues that arise nowadays but is included in the branch of streamflow modelling. Streamflow modelling has been extensively studied by the researcher since a decade ago for its potential in providing conceptual justifications for streamflow and predicting the future behavior or properties of streamflow.

The study of the streamflow process has been improved by the discovery of an important part of analysis known as the long memory (or persistent) process. The long memory process is a term that is not only used in the hydrological time series but also in financial and other fields of application. In the long memory process, the term known as the degree of differencing '*d* 'is the interesting coefficient to be discovered. The reason is that the degree of differencing in the class of the autoregressive integrated moving average (ARIMA) time series plays an important role in the mobilization of the time series model. In the modelling process, *d* is used to formulate an ARFIMA model.

The significance of *d* in the time series, in fact, plays an important role in terms of accuracy, compatibility, and mobility especially when used for modelling purposes