UNIVERSITI PUTRA MALAYSIA

TIME SERIES MODELING OF WATER LEVEL AT SULAIMAN STATION, KLANG RIVER, MALAYSIA

HADI GALAVI

TIME SERIES MODELING OF WATER LEVEL AT SULAIMAN STATION, KLANG RIVER, MALAYSIA

By

HADI GALAVI

Thesis submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the degree of Master of Science

November 2010

DEDICATION

This thesis is dedicated to my parents for their love, endless support and encouragement.

Abstract of thesis presented to the senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Master of Science

TIME SERIES MODELING OF WATER LEVEL AT SULAIMAN STATION, KLANG RIVER, MALAYSIA

By

HADI GALAVI

November 2010

Chairman: Prof Lee Teang Shui, PhD

Faculty: Engineering

The temporal and spatial flow in rivers represents the total response of a watershed and is affected by every hydrological event occurs in the watershed. The important point in planning and management of water projects is to achieve accurate estimation of river stage and flow. Among the two major approaches for modeling and forecasting hydrologic events, the empirical or black-box models have received more attention than the physical based models by hydrologists. Among the empirical models the Autoregressive Integrated Moving Average (ARIMA) models which have been conventionally used for hydrologic modeling and the Adaptive Neuro-Fuzzy Inference System (ANFIS) as a data driven model were used in this study to investigate their effectiveness in forecasting the river flows. The Klang River is one of the important urban rivers in Malaysia and is the subject of this research. A total of 3012 daily average water level measurements (2002 – 2010) of the river were used in this study. The number of inputs for both models is selected based on the Autocorrelation function (ACF) and the Partial Autocorrelation Function (PACF). The best ARIMA model is selected based on the minimum corrected Akaike Information Criterion (AICC) statistic

obtained among the several possible ARIMA models. An ARIMA(3,1,3) model was found to be the appropriate ARIMA model for the modelling of the selected case study. A Subtractive Clustering Method was used in the fuzzy inference system to determine the optimal number of Membership Functions (MF) and rules. Using the cross validation method the best training subset is selected to train the ANFIS model based on that dataset. The estimation of parameters of the model is accomplished using the hybrid learning algorithm consisting of standard neural network backpropagation algorithm and least squares method.

Among the several structures of ANFIS model examined, the model with four membership functions for the model inputs and output, Guassian membership function for the fuzzification, four rules for the fuzzy inference engine, and trained using hybrid algorithm was selected as the best constructed ANFIS model. The performance of the models, ANFIS and ARIMA models, was compared based on established statistical performance measures. Results show the superiority of the ANFIS model over the ARIMA model with interestingly better overall index of 1.514 against 1.108. However, the database of ARIMA model when became updated with each prediction step showed that results were dramatically better than the simple ARIMA model with 0.48 improvement in overall index. The updated ARIMA model showed very close results to those of the ANFIS model with overall index of 1.588 and 1.514, respectively. Overall, the updated ARIMA model in terms of overall index measure outperforms the ANFIS model for prediction of average water level, but not in the prediction of an exact time a day. Both models based on the obtained results (below one percent mean absolute percentage error) are appropriate models for the modelling and forecasting of Klang River water level. These models can be applied to the other case studies with model calibration in terms of the available dataset.

Abstrak tesis dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Master Sains

TIME SERIES MODELING OF WATER LEVEL AT SULAIMAN STATION, KLANG RIVER, MALAYSIA

Oleh

HADI GALAVI

November 2010

Pengerusi: Prof Lee Teang Shui, PhD

Fakulti: Kejuruteraan

Aliran di dalam sungai dari segi masa dan reruang membayangkan jumlah sambutan sebuah kawasan tadahan dan ia dipengaruhi setiap peristiwa hidrologi yang berlaku di dalam tadahan itu. Perkara terpenting di dalam perancangan dan pengurusan projek air adalah untuk mencapai penganggaran tetap paras air dan aliran. Di antara dua pendekatan utama bagi pemodelan dan peramalan peristiwa hidrologi, model empirik atau kotak-hitam telah mendapa lebih tumpuan daripada juru-hidrologi dibandingkan model berassakan fizik. Di antara model empirik Autoregressi Terkamir Purata Bergerak (ARIMA) yang biasa digunakan sebagai model hidrologi dan Sistem Taabir Kelam Saraf Suai (ANFIS) sebagai model berpandu data telah digunakan dalam kajian ini demi menyiasat keberkesanan untuk meramalkan aliran sungai. Sungai Klang ialah sebuah sungai perbandaran di Malaysia dan dipilih untuk kajian ini. Sebanyak 3012 ukuran harian paras air (2002 – 2010) sungai itu diguna dalam kajian.

Nombor masukan dipilih berdasarkan Fungsi Autosekaitan (ACF) dan Fungsi Autosekaitan Separa (PACF) bagi kedua dua model. Model ARIMA terbaik dipilih berdasarkan statistik Kriterium Matlumat Akaike (AICC) minimum yang didapati di antara beberapa model ARIMA. Model ARIMA(3,1,3) didapati paling sesuai untuk memodelkan kes kajian terpilih itu. Kaedah Gugusan Penolakan telad digunakan di dalam sistem taabir kelam untuk menentukan nombor optima Fungsi Keahlian (MF) dan peraturan. Dengan menggunakan kaedah pengesahan silang subset latihan terbaik dipilih untuk melatih model ANFIS berdasarkan set data tersebut. Parameter penganggaran model dicapai mengunakan algoritma belajar hibrid yang mengandungi algoritma piawaian rangkaian saraf perambatan-balik dan kaedah kuasa dua terkecil.

Diantara beberapa struktur model ANFIS yang diuji, model dengan empat fungsi keahlian bagi masukan dan keluaran model, fungsi keahlian Gaussian untuk kekelaman, empat peraturan untuk injin taabir kelam dan terlatih dengan menggunakan algoritma hibrid dipilih sebagai model terbaik dibina. Prestasi model ANFIS dan ARIMA dibandingkan berdasarkan ukuran prestasi statistik tertubuh. Keputusan menunjukkan keunggulan model ANFIS dibandingkan model ARIMA. Akan tetapi, bila database model ARIMA dikemaskini bagi setiap langkah ramalan, keputusan menunjukkan bahawa ianya lebih baik dibandingkan dengan model ARIMA dahulu. Model ARIMA terkemaskini menunjukkan keputusan rapat dengan model ANFIS dan terbukti tertunjuk dengan beberapa ukuran statistik model ARIMA adalah lebih baik. Pada keseluruhan, model ARIMA terkemaskini didapati berukuran index keseluruhan lebih berprestasi. Berasaskan kepada keputusan

tercapai, kedua dua model (ARIMA dan ANFIS) adalah sesuai untuk memodelkan dan meramalkan paras air di dalam sungai Klang.

ACKNOWLEDGEMENTS

First of all, I would like to express my deepest appreciation to my supervisor Prof. Dr. Lee Teang Shui, for his valuable guidance and suggestions, kindly attitude, and for believing in me during this work. I have been honored to have the opportunity to be his student. I would also like to thank Prof. Dr. Mohammad Hamiruce B. Marhaban, my co-supervisor, for his valuable comments and discussions on this study.

I attribute a great deal of my success to the positive interactions I have had with my family. My mother's emotional support with endless pure love over the years has been immeasurable. My father gave me encouragement and support, without which my educational goals could not be realized. I have learnt from him to love the science, be patient, and work hard.

Additionally, I would like to thank Dr. Mahendran Shitan for always making time to answer my questions despite his busy schedule at department of mathematics.

Special thanks are due to the Department of Irrigation and Drainage (Kuala Lumpur) for providing very useful information in order to carry out this study.

I certify that a Thesis Examination Committee has met on 04 November 2010 to conduct the final examination of Hadi Galavi on his thesis entitled "Time Series Modelling of Water Level at Sulaiman Station, Klang River, Malaysia" in accordance with the Universities and University College Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(a) 106] 15 March 1998. The Committee recommends that the student be awarded the Master of Science.

Members of the Examination Committee were as follows:

Helmi Zulhaidi Mohd Shafri, PhD
Lecturer
Faculty of Engineering
Universiti Putra Malaysia
(Chairman)

Law Teik Hua, PhD
Lecturer
Faculty of Engineering
Universiti Putra Malaysia
(Internal Examiner)

Mohd Amin b Mohd Soom, PhD
Professor
Faculty of Engineering
Universiti Putra Malaysia
(Internal Examiner)

Othman A. Karim, PhD
Associate Professor
Information Technology Centre
Universiti Kebangsaan Malaysia
(External Examiner)

SHAMSUDDIN SULAIMAN, PhD
Professor and Deputy Dean
School Of Graduate Studies
Universiti Putra Malaysia

Date: 18 January 2011

This thesis was submitted to the senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Master of Science. The members of the supervisory committee were as follows:

Lee Teang Shui, PhD
Professor
Faculty of Engineering
Universiti Putra Malaysia
(Chairman)

Mohammad Hamiruce b. Marhaban, PhD
Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

HASSANAH MOHD GHAZALI, PhD
Professor and Dean
School Of Graduate Studies
Universiti Putra Malaysia

Date:

DECLARATION

I declare that the thesis is my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously and is not concurrently submitted for any other degree at Universiti Putra Malaysia or other institutions.

HADI GALAVI

Date:

TABLE OF CONTENTS