NON-LINEAR WATER LEVEL FORECASTING OF DUNGUN RIVER USING HYBRIDIZATION OF BACKPROPAGATION NEURAL NETWORK AND GENETIC ALGORITHM

SITI HAJAR BINTI ARBAIN

A thesis submitted in fulfillment of the requirement for the award of the degree of Master of Science (Computer Science)

Faculty of Computing

Universiti Teknologi Malaysia

SEPTEMBER 2014

NON-LINEAR WATER LEVEL FORECASTING OF DUNGUN RIVER USING HYBRIDIZATION OF BACKPROPAGATION NEURAL NETWORK AND GENETIC ALGORITHM

SITI HAJAR BINTI ARBAIN

A thesis submitted in fulfillment of the requirement for the award of the degree of

Master of Science (Computer Science)

Faculty of Computing

Universiti Teknologi Malaysia

SEPTEMBER 2014

Thanks to God, Allah *'azza wa jalla*

Then to my beloved family and friends

ACKNOWLEDGEMENT

In preparing this thesis, I was in contact with many people, researchers, and lecturers. It is my greatest pleasure to take this opportunity to express my gratitude and thank you to all the people involve whether direct or indirect in my effort to successfully finish this project.

First of all, I would like to express my heartfelt thanks and deep sense of gratitude to my supervisor, Dr Antoni Wibowo for his guidance, patient, and encouragement as well as suggestions throughout the research. This project could never be successfully completed without his assistance and support. I am also very thankful to my co-supervisor, Prof. Madya Dr. Mohd Salihin Bin Ngadiman for his guidance and supportive in this research analysis. Without their continued support, this thesis would not be successfully finished.

I would like to extent my sincere appreciations to my lab research mates and friends for their help and support. The great comments and suggestions from all these individuals were invaluable to me. At it has been, and always will be, my greatest gratitude is to my family, who shared this burden fully and magnificently. Without their support, this project would have been difficult at best. Thank you

ABSTRACT

The Department of Irrigation and Drainage (DID) and Meteorological Malaysia Department (MMD) have identified that water level is one of the important indicators for flooding control. The aim of this study is to find the best regression model and to identify the dominant variables of water level in Dungun River. Autoregressive Integrated Moving Average (ARIMA),Seasonal ARIMA (SARIMA), Backpropagation Neural Network (BPNN) and Nonlinear Autoregressive Exogenous Model (NARX) are popular methods in time series forecasting. However, ARIMA and SARIMA produce linear models where the approximations of linear models for the complex real-world problems are not always satisfactory. Thus, Backpropagation Neural Network (BPNN) and Nonlinear Autoregressive Exogenous Model (NARX) can be implemented in the time series forescasting due to its nonlinear modelling capability. These four methods, however, cannot be used directly for water level prediction since the original data from DID and MMD contain missing data. In this thesis, two methods are employed to treat missing data which are pre-processing using Mean and preprocessing using Ordinary Linear Regression (OLR) substitutions. In addition, BPNN and NARX may be difficult to determine the optimal network architecture and weights design since the optimal weight are different in each learning process. Thus, it is difficult to get best model in prediction. Based on the limitation of BPNN and NARX, the hybridization of Single BPPN and Genetic Algorithms (S-BPNN-GA) and Multi BPNN and Genetic Algorithms (M-BPNN-GA) have been proposed in this study. Experiments indicate hybridization of M-BPNN-GA 5-6-1 using five predictor variables including monthly, rainfall, temperature, evaporation and humidity and give better results compared to the other methods.

ABSTRAK

Jabatan Pengairan dan Saliran (DID) dan Jabatan Meteorologi Malaysia (MMD) telah mengenal pasti bahawa paras air adalah salah satu petunjuk yang penting untuk pengawalan banjir. Tujuan kajian ini adalah untuk mencari model regresi yang terbaik dan untuk mengenal pasti pembolehubah dominan bagi paras air di Sungai Dungun. Autoregresi Bersepadu Purata Bergerak (ARIMA), ARIMA bermusim (SARIMA), *Backpropagation Neural Network* (BPNN) dan *Nonlinear Autoregressive Exogenous Model* (NARX) adalah kaedah yang popular dalam ramalan siri masa. Walau bagaimanapun, ARIMA dan SARIMA menghasilkan model linear di mana anggaran model linear kepada masalah dunia sebenar yang kompleks tidak sentiasa memuaskan. Oleh itu, kaedah BPNN dan NARX diaplikasikan dalam ramalan siri masa kerana keupayaan model tersebut untuk mengendalikan masalah tidak linear. Tetapi keempat-empat kaedah in tidak boleh digunakan secara langsung untuk meramal paras air menggunakan data dari DID dan MMD kerana mengandungi data yang tidak lengkap. Dalam tesis ini, dua kaedah digunakan untuk mengendalikan data yang tidak lengkap iaitu pra-pemprosesan menggunakan penggantian purata dan penggantian *Ordinary Linear Regression* (OLR). Di samping itu, BPNN dan NARX mungkin sukar untuk menentukan rangkaian seni bina dan reka bentuk pemberat yang optimum memandangkan setiap proses pembelajaran akan menghasilkan nilai yang berbeza. Akibatnya sukar untuk mendapatkan model terbaik dalam ramalan. Oleh kerana terdapat kekangan pada model BPNN dan NARX, maka model penghibridan BPPN dan Algoritma Genetik serta penghibridan Multi BPNN dan Algoritma Genetik dicadangkan dalam penyelidikan ini. Eksperimen menunjukkan penghibridan M-BPNN-GA 5-6-1 menggunakan lima pembolehubah peramal yang meliputi pembolehubah bulanan, kadar hujan, suhu, penyejatan dan kelembapan memberikan keputusan yang lebih baik berbanding dengan model yang lain.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Background of Research

Dungun is located between 4^{o}36’10’N to 4^{o}53’02’N and 103^{o}07’25’E to 103^{o}25’50’E, which is one of the seven districts in Terengganu state (Ekhwah *et al.*, 2009). Topographically the Dungun district comprises 35% lowland area, 20% swamp and other water bodies, and 45% forest reserves (Gasim *et al.* 2007). Dungun district (3 km width and 20 km long) is located at the coastal area and is bounded by Dungun River on the north and Paka River on the south. Dungun River is the longest river in Dungun which flows about 110 kilometers long before reaching South China Sea and draining about 2507 kilometers of catchment area.

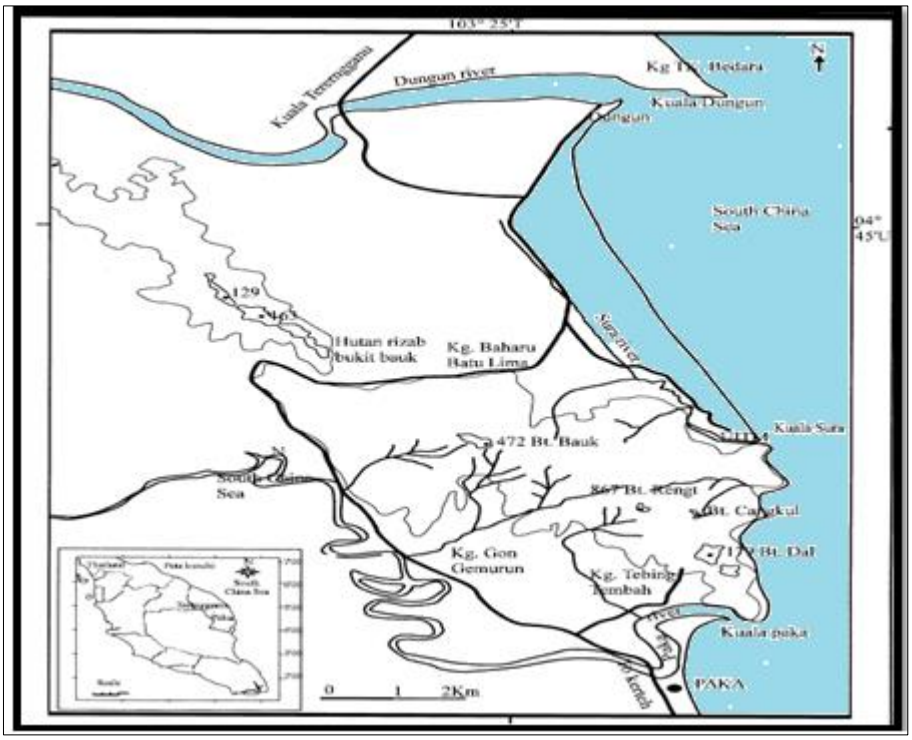

Figure 1.0 Location of the Dungun District

Figure 1.0 shows the location of the Dungun District between two main rivers which are the Dungun River and Sura River. In reports of flooding in Dungun District, Department of Irrigation and Drainage (DID) stated that there are two types of flooding which are flash floods and river flood. Flash flood usually occurs in urban areas where it is usually caused by short, intense localized thunderstorm rains, where it is usually experienced during the evening (DID, 2009). Besides flash flood, there is also river flood usually happens when the flow in a river exceeds its conveyance capacity, the water in the river rises above its bank level and overspills into adjacent low-lying areas, causing river floods.

In the beginning of the year 2000 and December 2001, an unexpected destruction happened to the livelihood and the local populations in Dungun, due to the continuous

heavy rain, resulting in massive flooding in the southern state of the peninsular (DID, 2002). It was known to cause major destruction to the residents of that particular area. According to the flood report by DID, there is a series of floods occurred in Dungun district-Terengganu from the year 2000 till 2004. Dungun was hit by a larger flood in the third week of November 2000. From March 2011 till April 2011, Terengganu was affected by a heavy rainfall which exceeded 200mm per day and resulted in a massive flood.

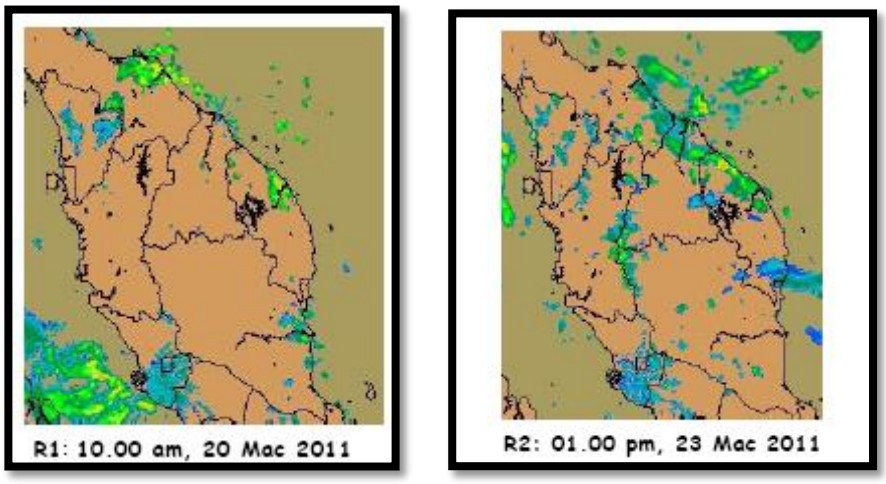

IMAGE RADAR ON 20-23 MARCH 2011

Figure 1.1 Image Radar of R1 and R2 on 20-23 Mac, the Widespread of Rainfall

Sources: *Report on Heavy Rainfall That Cause Flash Flood in Terengganu During The Period Prepared by MMD*

It was stated that heavy rainfall was brought by the strong winds from the South China Sea and western part of the Pacific Ocean, and the radar can be seen in Figure 1.1. From Figure 1.1, Image R1 shows the scattered rain (green color) which occurred over

coastal areas of southern Thailand, Kelantan, Terengganu and Pahang, morning of 20th March 2011.

However, there is slightly different on image R4 in Figure 1.2 where the widespread of torrential rain can be clearly observed over Kelantan and north of Terengganu. Whereas, the widespread rain (green colour) still continues since 23rd March over other part of Terengganu and Pahang.

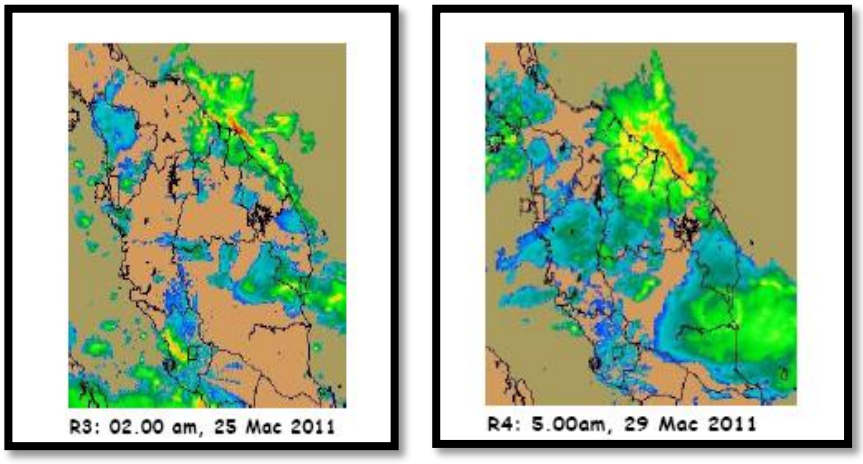

Figure 1.2 Image Radar of R3 and R4 on 25-29 Mac, the Widespread of Rainfall

Sources: *Report on Heavy Rainfall That Causes Flash Flood in Terengganu During The Period Prepared by MMD*

The DID also had been measuring some of the characteristics of the flood which included water level, inundation area, peak inundation, peak discharge, volume of flow

and duration. In terms of water level, DID have introduced three categories of critical level stages namely normal, alert and danger levels (Gasim *et al.* 2007). One of the rivers spotted by DID that had reached the danger level was Sungai Dungun located at Dungun district.

There is a case study which identified the factors that are related to the occurrence of flood phenomenon in Dungun, Terengganu. The occurrence of flood in Dungun district was due to a combination of physical factors such as elevation, its close proximity to the sea as well as the heavy rainfall experienced during the monsoon period (Gasim *et al.*, 2007).

The stages of water level are designed to make the local authority aware of the level of danger posed by the rising water level so that a necessary emergency arrangement could be initiated for the welfare of the local community affected by the river. As water level forecasting could reduce damage from impact of flooding in agriculture, public uses, avoid life losing and economic loses, it is therefore important to predict its appearance.

Prediction the pattern of water level is one of the benchmark points in flood forecasting analysis. It has been one of the most important issues in hydrological research. Water level is an essential component in the process of forecasting flood resources evaluation and is considered as a central problem in hydrology (Ekhwah *et al.*, 2009). Flood forecasting relies on real time rainfall and water level data obtained from DID's network of telemetric stations.

In preparation for flood during the monsoon season, the performance of the telemetric stations will be checked, repaired and forecasting models will be tested and parameters are adjusted if necessary including all the variables such as water level and

rainfall data (DID, 2009). In Terengganu, there are 94 and 28 stations set up for rainfall and water level inventory stations, respectively, prepared by DID so that the data can be collected in real-time. For flood forecasting to be effective and sufficient, the lead time must be available. Lead time is the time between the detection or the forecasting of flood and the time of actual occurrence of the event.

Water level forecasting is one of the forecasting problems that have attracted the interest of scientists (Amir *et al.*, 1999). Estimating the water level can have significant impact, as this can help in agricultural water management and in protection from water shortages and possible flood damages. In statistics analysis, there are a lot of ways in doing the predicting. Time series forecasting is highly used in predicting economic and business trends.

Recently there are many researchers who have been successfully make prediction in hydrology like rainfall forecasting, flooding, water quality, water level, water consumption and etc (Faruk, 2010). In this study, the methods of Autoregressive Integrated Moving Average (ARIMA) or seasonal ARIMA (SARIMA), Backpropagation Neural Network (BPNN), Nonlinear Autoregressive Models with exogenous inputs (NARX), have been applied as a model of water level prediction allocation at the selected water level, rainfall, evaporation, temperature and relative humidity at monitoring stations prepared by DID and MMD in the Dungun River-Terengganu.

To deal with the issue of linearity, we can use Back Propagation Neural Network (BPNN) and Nonlinear Autoregressive Exogenous Model (NARX). However, BPNN and NARX may be difficult to determine the optimal network of architecture and weights design. Therefore, we propose hybrid techniques to overcome the limitations of the above techniques. Hybrid Neural Network based on genetic algorithm (GA) then will

be compared to the traditional methods such as ARIMA, BPNN and NARX standalone methods to evaluate the predictive performance between those methods. Since the weights in BPNN model are hard to estimate due to the large number of possible solutions, therefore GA is implemented to improve the performance of artificial neural networks. GA is adopted in searching optimal design based on the weight and biased in BPNN which proposed hybrid sense.

In this research, we used monthly rainfall, rate of evaporation, rate of temperature, relative humidity and water level which are collected from Department of Irrigation (DID) and Drainage and Meteorological Department (MMD) for Dungun district-Terengganu, comprises a total of 75 data observations covered from year 2006 until 2012.

1.2 Statement of Problem

After the flood in 2000 and 2001, the Dungun River might be over saturated hence, the rise in the water level. When the water level starts to rise above the river banks or dams, the water starts overflowing which often causes flood. The phenomenon of the flood could result in a lot of destructions to the livelihood and the local populations.

The recent developments in variables selection methods have addressed the problems from the point of view in improving the performance of the selection of predictors. It is notice that the original data that was collected from DID and MMD have some imperfect characteristics that need to undergo the process of cleaning and treating missing data before proceeding to the next method procedures. Hence, pre-processing of data is one of the most important methods before we move on to other procedures

where it will handle the imperfect characteristics of the data such as missing and inconsistent value of data.

Due to the fact that water level forecasting involves a complex nonlinear data pattern; there are a lot of forecasting methods to improve the forecasting accuracy. It was reported that DID used general regression and linear transfer function model for their hydrological modelling (DID, 2012). It means that DID perform linear models for forecasting hydrological modelling. Commonly, the linear models when applying to complex real-world problem are not always satisfactory. To deal with the issues of linearity of model, hybrid nonlinear model will be proposed.

There are a number of alternative ways to deal with missing data. In some cases, deletion or elimination of the missing variable is the default method for most procedures (Howell, 2007). However in time series regression, this approach does not seem to be a favour. Treatment of missing data is very important in time series data because it is not only about the issues of 'missingness" but the value of observation is the most important one. In this research, we used two types of treatments in dealing with missing data which are using Ordinary Linear Regression (OLR) and Mean substitution. Therefore, we need to identify how good these two types of methods can handle the missing data, in order to get better prediction model.

There were many analyses about forecasting time series approaches had been made in hydrological problems. The choice of the forecasting model is one of the most important factors on how to improve some application in forecasting accuracy. Autoregressive Integrated Moving Average (ARIMA), Seasonal ARIMA (SARIMA), Back Propagation Neural Network (BPNN) and Nonlinear Autoregressive Exogenous Model (NARX) are popular methods in time series forecasting. Time Series ARIMA are generated from linear processes, which may be inappropriate if the underlying

mechanism is nonlinear. However, we need to conduct variables selection to find the dominant variables which influence the water level predicition. The reducing of the number of variables can decrease the learning process time in water level prediction.

In addition, BPNN and NARX may be difficult to determine the optimal network of architecture and weights design since the optimal weights are different in each learning process, implying that, it is difficult to get the best model in prediction. Therefore, we considered hybrid BPNN and another metaheuristic method to improve the quality of water leter prediction.

1.3 Problem Questions

Based on statement of problem above, we summarized the problem questions as follows:

i) What is the suitable method for pre-processing treatment of missing data?
ii) What are the dominant variables in water level prediction of Dungun River?
iii) How to improve the performance of Neural Network in water level prediction?
iv) Can the proposed methods enhance the water level prediction compared to the existing methods?

1.4 Objective of Research

Based on the problems in this research, the objectives of the research are:

i) To treat missing data using OLR and mean substitution in order to achieve complete time series data.

ii) To find best variables using the common time series methods such as ARIMA/ SARIMA, BPNN and NARX.

iii) To hybrid single and multi-classifiers BPPN with GAs for water level prediction in Dungun District.

iv) To compare the performance of two types of hybrid BPNN-GA methods with the common time series methods which is standalone BPNN, NARX and ARIMA in prediction of water level.

1.5 Scope of Research

Based on the objective, the scope f the research are:

i) For data analysis, data months, monthly rainfall, rate of evaporation, rate of temperature, relative humidity and water level were used, which were collected from DID and MMD for Dungun district-Terengganu. The total number of observation is 75 which covered from January year 2006 until March 2012.

ii) Only two methods are considered for treatment missing data which is OLR and Mean Substitution.

iii) Hybridization process included Single BPNN-GA and Multi BPNN-GA.

iv) A comparative study is carried out between ARIMA/ SARIMA, BPNN, NARX, S-BPNN-GA and M-BPNN-GA.

1.6 Significance of Research

The domain of this research is forecasting of water level prediction in Dungun River, Terengganu. In this research, we used time series regression method which included linear and nonlinear methods. Both methods have good forecasting model but there are some limitation problems that need to be considered. Therefore, hybrid nonlinear method is proposed to overcome the limitation of traditional methods. It is hoped that this research will be one of the beneficial findings that will be used by the government as one of the solutions to predict water level for further investigation. Although this is caused by the natural phenomenon, but hopefully this research will be one of the efforts to analyze the pattern of water level, rainfall, rate of evaporation, temperature and relative humidity to provide a better flood warning service and flood forecasting service in DID and MMD.

1.7 Thesis Organization

This report is organized in six chapters. Chapter 1 illustrates an introduction and brief overview of researches including background of research, formulation problems, objectives, scopes, significance of study, and report organizations. Chapter 2 is about literature reviews that discuss the related theories to be used as foundation of this study. Chapter 3 covers on methodology and operational framework of studies. Chapter 4 will briefly discuss the processes of experimental studies. Chapter 5 will discuss about the simulation analysis and results. Chapter 6 gives the conclusion and discussion for future development of this study.

REFERENCES

Amir, F. A., Suzan, M.E., Samir, I.S., and Mohamed, S.E., (1999). A Comparison Between Neural Network Forecasting techniques- Case Study: River Flow Forecasting. *IEEE Transactions on Neural Networks,* Vol. 10, No. 2.

Areekul, P., Senjyu, T., Toyama H., and Yona, A., (2010). A Hybrid ARIMA and Neural Network Model for Short-Term Price Forecasting in Deregulated Market. Japan : Department of Electric. & Electron, Engineer University of the Ryukyus, Nishihara.

Aryal, D.R. and Wang, Y.W., (2003). Neural Network Forecasting of the Production level of Chinese Construction Industry. *Journal of Comparative International Management*, Vol. 6, No. 2, 45-64.

Barrett, C.B., and others, (1985). Hydrology Subcommittee of the interagency Advisor Committee on Water Data, Guidelines on Community Local Flood Warning and Response System. Springfield.

Billa, L. , Mansor, S.B. and Mahamod, A.R., (2000). Quantitative Precipitation Forecasting using Cloud-Based Technique on AVHRR data. *Journal of Disaster Research*, 11(1): 1-2.

Bowerman, B.L. and Connell, R.T., (1987). Time Series Forecasting, *2nd* Ed, USA: Duxbury Press.

Box, G.E.P. and Jenkins, G.M., (1976). Time Series analysis: Forecasting and control, Holden-Day, San Francisco.

Box, G.E.P. and Jenkins, G.M., (1994). Time Series Analysis, Forecasting and Control, Prentice Hall, Englewood Cliffs.

Brown, L.D., (1993). Earning Forecasting Research: Its Important Implications for Capital Markets Research. *International Journal of Forecasting*, 9: 295-320.

Burn, D.H. and McBean, E.A., (1985). River Flow Forecasting Model For Sturgeon River. *Journal of Hydraulic Engineering*, Vol 111(2), 316-333.

Chen, T.L., (2011). Forecasting the Taiwan Stock Market with a Novel Momentum-based Fuzzy Time Series. *Academic Research Centre of Canada*, *Review Economics &Finance*, 1923-8401.

Cohen, J. and Cohen, P.(1983). Missing data. Applied Multiple Regression: Correlation Analysis for the Behavioral Ssciences ,pp. 275-300.

Connor, J., and Atlas, L., (1993).Recurrent Neural Networks and Time Series Prediction. *Proceedings of* the International Joint Conference on Neural Networks 1:301-306.

Dorigo, M. and Caro, G.D., (1999). Ant Algorithm for Discrete Optimization, *Artificial Life*, Vol(5), pp. 137-172.

Ekhwah, M.T., Juahir, H., Mokhtar, M., Gazim, M.B., Abdullah, S.M.S. and Jaafar, O., (2009). Predicting for Discharge characteristics in Langat River, Malaysia using Network Network Application Model. *Research journal of Earth Sciences*, 191:15-21.

El-Fandy, M.G., Ashour, Z.H., Taiel, S.M.M., (1994). Time Series Models Adopted for Forecasting Nile Floods and Ethiopian Rainfalls, *Bulletin of the American Meteorological Society*, 75(1):83-94.

Eugen, (2008). The use of NARX Neural Networks to Predict Chaotic Time Series. *WSEAS Transactions of Computer Research*. Issue 3, Volume 3.

Faruk, D. O., (2010). A Hybrid Neural Network and ARIMA Model for Water Quality Time Series Prediction. *Journal of Engineering Applications of Artificial Intelligence* 23, 586-594.

Ganji, H. and Wang, L., (2011). Hybrid Neural Network Models for Hydrologic Time Series Forecasting Based on Genetic Algorithm. Fourth International joint Conference on Computational Sciences and Optimization.

Gasim, M. B., Adam, J. H., Toriman, M. E., Rahim, S. A. and Juahir, H., (2007). Coastal Flood Phenomenon in Terengganu, Malaysia: Special Reference to Dungun. *Research Journal of Environmental Sciences*, 1: 102-109.

Hagan, M.T., H.B. Demuth and M.H. Beale, 1996. Neural Network Design. 1st Edition, PWS Publishing Company, Boston, ISBN-10: 0534943322, pp: 736.

Hanke, J.E. and Wichern, D., (2009). Business Forecasting, Pearson International Edition.

Holger, K. and Thomas, S., (2004). Non Linear Time Series Analysis, Cambridge Second Edition.

Holland, J. H. (1975/1992). Adaptation In Natural And Arti¯Cial Systems. Cambridge,MA: MIT Press. Second Edition (1992). (First Edition, University Of Michigan Press, 1975).

Horne, B.G. and Giles, C.L., (1995). An Experimental Comparison of Recurrent Neural Network, in Advances in Neural Information Processing Systems 7, MA:MIT press, pp. 697-704.

Howell, D.C., (2007). The Analysis of Missing Data. In Outhwaite, W. & Turner, S. Handbook of Social Science Methodology, London.

Huang, G. and Wang, L., (2011). Hybrid Neural Network Models for Hydrologic Time Series Forecasting Based on Genetic Algorithm. Fourth International Joint Conference on Computational Sciences and Optimization.

Jan, S., and Katarina, H., (2010). The Implementation of Hybrid ARIMA-Neural network Prediction Model for Aggregate Water Consumption Prediction. *Journal of Applied Mathematics*, Vol (3),No. 3.

Jonathan, P.R, and Montas, H.J., (2006). A Comparison of Artificial Neural Networks and Statistical Regression with Biological Resources Applications. Thesis Master of Science, 2006.

Kaiser, J.(1990). The robustness of regression and Substitution by Mean Methods in Handling Missing Values. Paper Presented at the second Islamic Countries Conference on Statistical Science, Malaysia.

Kirkpatrik, S., Gelatt, C.D. and Vecchi, M.P., (1983). Optimization by Simulated Annealing, *Sciences* 220, pp. 671-680.

Kost, B. (1995). Evolution Strategies in Structural Topology Optimization of Trusses. International Conference on Computer in Civil and Building Engineering.

Kuan, C. M. and White, H. (1994) Artificial Neural Networks: An Econometric Perspective. Econometric Reviews, 13, 1-91.

Kumar, G., (2005). Artificial Neural Network and Its Applications, New Delhi.

Liang, Y.H., (2009). Combining seasonal time series ARIMA method and neural networks with genetic algorithms for predicting the production value of the mechanical industry in Taiwan. Neural Computing & Applications, Vol. 18 Issue 7, p833.

Lisboa, P.J.G., Edisbury, B.and Vellido, A., (2000). Business Applications of Neural Networks, the State of the Art of Real World Application, Springer Publications.

Little, R.J.A. and Rubin, D. B. (1987). Statistical Analysis with Missing Data. John Wiley & Sons.

Maasoumi E., Khotanzad,, A., and Abaye, A., (1996). Artificial neural networks for some macroeconomic series: a first report. Econometric Reviews, 13 (1), 105-122.

Nair, B.B., Sai, S.G., Naveen, A.N., Lakshmi, A., Venkatesh, G.S., and Mohandas, V.P., (2011). A GA-Artificial Neural Network Hybrid System for Financial Time Series Forecasting. Springer-Verlag Berlin Heidelberg, pp. 499-506.

Pham, D.T. and Liu, X., (1995). Neural Network for Identification, Prediction and Control, Springer Publications.

Rajasekaran, S. and Vijayalakshmi, G.A., (2007). Neural Networks, Fuzzy Logic, and Genetic Algorithms: Synthesis and Applications. Prentice Hall of India Private Limited, New Delhi.

Roth, P. (1994). Missing data: A conceptual review for applied psychologists. Personnel Psychology, 47, 537-560.

Schafer, J. L., and Graham, J.W., (2002). Missing Data : Our View of the State of The Art. American Psychological Association. Inc, Vol. 7, No. 2, 147-177.

Sharda, R. and Patil, K., (1994). Neural Networks for the MS/OR Analysis. *International Journal of Economics,* 24: 116-130.

Simon H., (1999). Neural Networks, Second Edition, Pearson Education.

Suguna, N. and K.G. Thanuskodi, 2011. Predicting Missing Attribute Values using K-means Clustering. J. Comp. Sci., 7: 216-224. DOI: 10.3844/jcssp.2011.216.224

Tsay, R., (2010). Estimation of Univariate ARIMA Models. Lecture 8 Time Series Analysis.

Tseng, F.M., Yu, H.C., Tzeng, G.H., (2002). Combining Neural Network Model with Seasonal Time Series ARIMA model. *Journal of Technological Forecasting and Social Change*, 69 : 71-87.

Vaziri, M., (1997). Predicting Caspian Sea Surface Water Level by ANN and ARIMA models. *Journal of Waterway, Port, Coastal, and Ocean Engineering*, Vol (123) No. 4.

Walczak, S. (2001). An Empirical Analysis of Data Requirements for Financial Forecasting with Neural Networks. *Journal of Management Information Systems, Spring*, 17, (4), 203-22.

Wang L., Tang F., and Wu H., (2005). Hybrid genetic algorithm based on quantum computing for numerical optimization and parameter estimation. *Applied Mathematics and Computation*, 171(2): 1141-1156.

Wang, D., Lum, K.Y., Yang, G., (2002). Parameter Estimation of ARX/NARX Model: A Neural Network Based Method. *Proceedings of the 9th International Conference on Neural Information Processing*, Vol. 3.

Weigend, A., Mangeas, M. and Srivastava, A., (1995). Nonlinear gated experts for time series: Discovering regimes and avoiding overfitting. *Int. J. Neural Syst*., vol. 6, pp. 373–399.

Westin, L.K., (2002). Missing Data and The Preprocessing Perceptron. Department of Computing Sciences.

White, H (1988). Economic Prediction using Neural Networks: The Case of IBM Daily Stock Returns. *Proceeding of the IEEE International Conference on Neural Networks II*, 451-58.

Wolfram, S., (1994). Cellular Automata and Complexity. First edition, Addison-Wesley.

Xie, H., H. Tang and Y.H. Liao, 2009. Time series prediction based on NARX neural networks: An advanced approach. Proceeding of the International Conference on Machine Learning and Cybernetics, Jul. 12-15, IEEE Xplore Press, Baoding, pp: 1275-1279. DOI: 10.1109/ICMLC.2009.5212326.

Zhang, G., Patuwo, E.B. and Hu, M.Y., (1998). Forecasting with artificial neural networks: the state of the art. *International .Journal of Forecasting*, 14, 35-62.

Zhang, G.P., (2002). Time Series Forecasting Using Hybrid ARIMA and ANN Model. *World Applied Sciences Journal,* 50:159-175.

Zhang, G.P., M.Qi, (2005). Neural Network Forecasting for Seasonal and Trend Time Series. *European Journal of Operational Research,* 160: 501-514.

—, (2010). World Meteorological Organizational (WMO). Guide to Agriculture Meteorological Practices, Edition 2010, No. 124.

__, (2001). Department of Irrigation and Drainage (DID). Laporan Banjir 2000/2001. Unit Hidrologi Jabatan Pengairan dan Saliran Negeri Terengganu.

__, (2002). Department of Irrigation and Drainage (DID). Laporan Banjir 2001/2002. Unit Hidrologi Jabatan Pengairan dan Saliran Negeri Terengganu.

__, (2009). Department of Irrigation and Drainage (DID). Flood Forecasting and Warning system Report. Unit Hidrologi Jabatan Pengairan dan Saliran Malaysia.