Neural Networks Based Nonlinear Time Series Regression for Water Level Forecasting of Dungun River

Siti Hajar Arbain and Antoni Wibowo
Department of Modeling and Industrial Computing,
Faculty Computer Science and Information Systems,
University Teknologi Malaysia (UTM), Johor Bahru, Johor, 81310, Malaysia

Abstract: The Department of Irrigation and Drainage (DID) Malaysia and Meteorological Malaysia Department (MMD) has been measured the flood characteristics benchmark which included water level, area inundation, peak inundation, peak discharge, volume of flow and duration of flooding. In terms of water levels, DID have introduced three categories of critical level stages namely normal, alert and danger levels. One of the rivers detected by DID that had reached danger level is Sungai Dungun located at Dungun district, Terengganu. The aim of this study is to find suitable prediction model of water level with input variables monthly rainfall, rate of evaporation, temperature and relative humidity taken from the same catchment at Dungun River using Neural Networks based Nonlinear Time Series Regression methods which are Backpropagation Neural Network (BPNN) and nonlinear autoregressive models with exogenous inputs (NARX) networks. The variables selection criteria procedures are also developed to select a significant explanatory variable. In addition, the process of pre-processing data such as treatment of missing data has been made on the original data collected by DID and MMD. The methods are compared to obtain the best model for prediction water level in Dungun River. Based on the experiments, the NARX model with five predictor variables is the best model compared to BPNN. In addition, treatment of missing data using mean and OLR approach produced comparable results for this case study.

Keywords: Time series regression, missing data, BPNN, NARX, dungun river, terengganu, water level

INTRODUCTION

Overview problems: Malaysia is located at the tropics between 10°C N and 60°CN of the equator whereas Dungun is one of the seven districts in the Terengganu state. Dungun is located between 4°C 36'10N to 4°C 53'02N and 103°C 07'25E to 103°C 25'50E (Gasim *et al*., 2007). 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 situated at the coastal area between Dungun River (North) and Paka River (South). Dungun River is the longest river in Dungun which has flows about 110 kilometers long before reaching South China Sea and draining about 2507 kilometers of catchment area. In Dungun, flood occurs almost every year and during flood events, Dungun and its surrounding area was inundation by coastal flood water up to 1.50 m depth (Gasim *et al*., 2007). Recently, the Meteorological Malaysia Department (MMD) was warning on yellow stage of rainfall in some state included Dungun on November 2011. Yellow stage warned when

intermittent rain occasionally moderate is occurring over the state. In Dungun, flashfloods occurs almost every year when the level of Dungun River rises a few meters above the danger stage DID, 2002.

The recent developments in variables and feature selection methods have addressed the problems from the point of view of improving the performance of predictor's selection. Preprocessing data is one of the most important methods before we do the next step procedures where it will handle the imperfect characteristic data such as missing and inconsistent value of data. The original data that were collected from Department of Irrigation and Drainage (DID) and MMD involve some imperfect characteristics that need to undergo process of treatment missing data before going for the next method procedures. The missing data are part of the almost pre analysis research and we need to decide how to deal with it from time to time.

There are many analysis approaches have been made in order to overcome the hydrological problems where the neural networks is one of the methods that are widely used to solve them. Since Artificial Neural

Corresponding Author: Siti Hajar Arbain, Department of Modeling and Industrial Computing, Faculty Computer Science and Information Systems, University Teknologi Malaysia (UTM), Johor Bahru, Johor, 81310, Malaysia

Network (ANN) have ability to recognize time series patterns and nonlinear characteristics, which gives better accuracy over the others methods, it become most popular methods in making prediction (Vaziri, 1997; Sharda, 1994; Toriman *et al.*, 2009). A comparative study were made using ANN and conventional Auto-Regression (AR) model networks in forecasting the river flow for two well known River in USA and they found that ANN performed better than AR model. It has been reported that the results using Radial Basis Function (RBF) neural network is better than Back Propagation Neural Network (BPNN) in modeling a meteorological problems such as weather forecasting. In addition, there was analysis shows that pre-processing data analysis also can influenced the performance of prediction model (Zhang, 2002; Suguna and Thanuskodi, 2011). From previous work, the method of BPNN is better compared to SARIMA in obtaining water level prediction at Dungun River, Terengganu (Arbain and Wibowo, 2012).

In this study, we are focused on two types of ANN which are Backpropagation Neural Network (BPNN) and nonlinear autoregressive models with exogenous inputs (NARX) network to obtain prediction of Water Level (WL) monitoring stations in the Dungun River. To apply these two methods, data which contain some imperfect characteristics collected from DID and MMD need to undergo preprocessing analysis before going for the next methods procedure.

The remain of this study is organized as follow. The first part will briefly discussed about theories and methods of BPNN and NARX. After that we will discussed about the experiments of case study of all methods proposed and followed by model selection. Finally the study is closed with the conclusion of the case study.

MATERIALS AND METHODS

Backpropagation Neural Network (BPNN): Backpropagation is the most widely used as learning algorithm and was a popular technique because it is easy to implement. It does require a data for conditioning the network before using it for predicting the output. A backpropagation network includes one or more hidden layers. The network is considered as feed forward approach, since there are no interconnections between the output of a processing element and the input of node on the same layer or on the preceding layer. Externally provided correct patterns are compared with the ANN output during training and feedback is used to adjust the weights until all training patterns are correctly categorized by the network.
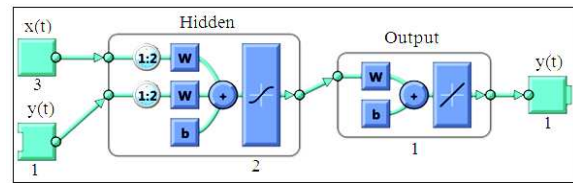

Fig. 1: The Snapshot of NARX network

Backpropagation was first described by Paul Werbos in 1974, then in 1986, through the study of David E. Rumelhart, Geoffrey E. Hinton and Ronald J.Williams, it gained recognition and led to a "renaissance" in the field of ANN research.

In BPNN network, there are four important considerations are comprised in network designs which are the network architecture determination, hidden neuron number determination, activation function optimization and training algorithm optimization (Hagan *et al.*, 1996).

Nonlinear Autoregressive Network with Exogenous Inputs (NARX): The NARX models are commonly used in the system of identification area (Xie *et al.*, 2009). All the specific dynamic networks discussed so far have either been focused networks, with the dynamics only at the input layer, or feedforward networks. The nonlinear autoregressive network with exogenous inputs (NARX) is a recurrent dynamic network, with feedback connections enclosing several layers of the network. The NARX model is based on the linear ARX model, which is commonly used in time-series modelling.

Figure 1 illustrates the standard NARX network. The standard NARX network used here is a two-layer feedforward network, with a sigmoid transfer function in the hidden layer and a linear transfer function in the output layer. This network also uses tapped delay lines (d) to store previous values of the input, x(t) and output, y(t) sequences. First, load the training data and use tapped delay lines with two delays for both the input and the output, so training begins with the third data point. There are two inputs to the series-parallel network, the x (t) sequence and the y(t) sequence. Notice that the y(t) sequence is considered a feedback signal, which is an input that is also an output (target). The model can be shown as in Eq. 1 as follow:

$$y(t) = f(y(t-1),...,y(t-d), x(t-1),...,(t-d)) \qquad (1)$$

where, y(t) is the output of the NARX network and also feedback to the input of the network and tapped delay lines (d) that store the previous values of x(t) and y(t) sequences. It also has been reported that gradient-descent learning can be more effective in NARX networks than in other recurrent architecture (Horne and Giles, 1995).

Case study: In this case study, 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 the total number of observation which is 75 data covered from year 2006 until 2012. The following part will discuss for each procedure that have been used in this case study.

Preprocessing data: The original data that were collected from DID and MMD involve some imperfect characteristics that need to undergo process of treatment data before going for the next method procedures. There are numbers of alternative ways of dealing with missing data. In this study, we used the method of finding mean and Ordinary Linear Regression (OLR) substitution in order to treat a missing data. There is a lot of research have been used these two methods (Westin, 2004; Little and Rubin, 1987; Outhwaite and Turner, 2007). Missing data will be conducted by finding the mean and using OLR for each month throughout 7 years and the Non Value (NA) are replaced by the corresponding values.

Treatment of missing data: There are several reasons why the data may be missing. There are maybe missing because of the malfunctioned equipment, the weather was terrible, people who are responsible to collect the data got sick or maybe the data were entered incorrectly. Missing data is very important to be treating first in every starting analysis because it is not only about the issues of 'missingness" but the value of observation is the most important in time series data. Table 1 shows the snapshot of raw data from 1st until 12th month which is in January 2006 until December 2006. From this figure, we can see NA values shows that there are missing values of rainfall in January until September 2006.

There are a lot of approaches to deal with missing data. In some cases, deletion or elimination the missing variable is the default method for most procedures (Suguna and Thanuskodi, 2011). However in time series regression, this approach seems like not the best

methods to be used. In the next part, we will discuss about two methods selected from previous study to treat missing data which is using mean and OLR substitutions (Outhwaite and Turner, 2007).

Using mean approach: One approach that is sometimes used on an independent variable data missing is to substitute for the missing scores with the mean of that value for all nonmissing cases. Substitute the missing data is one of the methods substitution where the remaining missing values are estimated using a measure of central tendency either using mean or median. In this study, we find a mean to replace all missing data observations and Eq. 2 was used to calculate the missing value. Table 2 shows the snapshot of replacement values of missing data using mean calculations:

$$\bar{X} = \frac{\sum X}{N} \qquad (2)$$

where,
$\bar{X}$ = The average of non missing data in certain month and used to replace the missing values
X = The values for all non missing cases in certain month
N = Total number of non missing value in certain month

Using Ordinary Linear Regression (OLR) approach: Using OLR approach, we will predict the value of missing data using regression model for each variables. The predicted value obtained from regression model will replace the missing data. For an input variable with missing data, the replacement values are estimated by treating this input as target and using the remaining input variable as predictor in regression model. In this study, all input variables which are rainfall, rate of evaporation, temperature and relative humidity will use month (t) as predictor variable. The Eq. 3-6 are the regression model for each variable whereas Table 3 shows the snapshot of replacement values of missing data using OLR calculations.

Table 1: The snapshot of raw missing data

Months	t	Rainfall	Evaporation	Temperature	Relative humidity	Water level
Jan	1	NA	3.8548	26.2419	78.6513	14.7200
Feb	2	NA	3.9194	28.8107	79.1893	14.8300
Mar	3	NA	4.8387	27.2452	78.1774	13.9600
Apr	4	NA	5.2484	27.9567	77.7867	13.8100
May	5	NA	4.9032	27.6323	79.0806	13.9500
Jun	6	NA	3.8548	27.5033	79.1600	14.1300
Jul	7	NA	3.4194	28.0839	77.5581	13.7500
Aug	8	NA	4.0161	27.3903	78.5613	13.7500
Sep	9	NA	3.7258	26.9500	79.7400	13.9200
Oct	10	3.5161	4.0323	27.2323	80.2355	13.7800
Nov	11	11.2258	3.6129	26.6300	83.7767	14.1400
Dec	12	20.5484	3.1452	26.7194	81.1548	14.2500

Table 2: The snapshot of substitution missing values using mean approach

Months level	t	Rainfall	Evaporation	Temperature	Relative humidity	Water
Jan	1	17.0672	3.8548	26.2419	78.5613	14.7200
Feb	2	4.2204	3.9194	26.8107	79.1893	14.8300
Mar	3	12.5565	4.8387	27.2452	78.1774	13.9600
Apr	4	6.4097	5.2484	27.9567	77.7867	13.8100
May	5	8.2258	4.9032	27.6323	79.0806	13.9500
Jun	6	7.1129	3.8548	27.5033	79.1600	14.1300
Jul	7	7.1774	3.4194	28.0839	77.5581	13.7500
Aug	8	9.2710	4.0161	27.3903	78.5613	13.7500
Sep	9	12.1387	3.7258	26.9500	79.7400	13.9200
Oct	10	3.5161	4.0323	27.2323	80.2355	13.7800
Nov	11	11.2258	3.6129	26.6300	83.7767	14.1400
Dec	12	20.5484	3.1452	26.7194	81.1548	14.2500

Table 3: The snapshot of substitution missing values using OLR approach

Months	t	Rf	Eva	Temp	Humid	WL
Jan	1	9.4567	3.8548	26.2419	78.5613	14.7200
Feb	2	9.5241	3.9194	26.8107	79.1893	14.8300
Mar	3	9.5915	4.8387	27.2452	78.1774	13.9600
Apr	4	9.6588	5.2484	27.9567	77.7867	13.8100
May	5	9.7262	4.9032	27.6323	79.0806	13.9500
Jun	6	9.7936	3.8548	27.5033	79.1600	14.1300
Jul	7	9.8609	3.4194	28.0839	77.5581	13.7500
Aug	8	9.9283	4.0161	27.3903	78.5613	13.7500
Sep	9	9.9957	3.7258	26.9500	79.7400	13.9200
Oct	10	3.5161	4.0323	27.2323	80.2355	13.7800
Nov	11	11.2258	3.6129	26.3600	83.7767	14.1400
Dec	12	20.5484	3.1452	26.7194	81.1548	14.2500

Table 4: The snapshot of normalized treatment data

Treat-ment	Months	t	Rf	Eva	Temp	Humid	WL
OLR	Jan	1	0.2490	0.7343	0.8999	0.9034	0.8669
	Feb	2	0.2507	0.7465	0.9194	0.9106	0.8734
	Mar	3	0.2525	0.9217	0.9343	0.8990	0.8221
	Apr	4	0.2543	0.9997	0.9587	0.8945	0.8133
	May	5	0.2561	0.9339	0.9476	0.9094	0.8216
	Jun	6	0.2578	0.7343	0.9432	0.9103	0.8322
	Jul	7	0.2596	0.6513	0.9631	0.8919	0.8098
	Aug	8	0.2614	0.7650	0.9393	0.9034	0.8098
	Sep	9	0.2632	0.7097	0.9242	0.9170	1.8198
	Oct	10	0.926	0.7680	0.9339	0.9227	0.8115
	Nov	11	0.2955	0.6882	0.9132	0.9634	0.8327
	Dec	12	0.5410	0.5991	0.9163	0.9332	0.8392
Mean	Jan	1	0.4494	0.7343	0.8999	0.9034	0.8669
	Feb	2	0.1111	0.7465	0.9194	0.9106	0.8734
	Mar	3	0.3306	0.9217	0.9343	0.8990	0.8221
	Apr	4	0.1688	0.9997	0.9587	0.8945	0.8133
	May	5	0.2166	0.9339	0.9476	0.9094	0.8216
	Jun	6	0.1873	0.7343	0.9432	0.9103	0.8322
	Jul	7	0.1890	0.6513	0.9631	0.8919	0.8098
	Aug	8	0.2441	0.7650	0.9393	0.9034	0.8098
	Sep	9	0.3196	0.7097	0.9242	0.9170	0.8198
	Oct	10	0.0926	0.7680	0.9339	0.9227	0.8115
	Nov	11	0.2956	0.6882	0.9132	0.9634	0.8327
	Dec	12	0.5410	0.5991	0.9163	0.9332	0.8392

Rainfall (RF) OLR model:

$$Rf(t) = 9.38935 + 0.06737(t) \quad (3)$$

Evaporation (Eva) OLR model:

$$Eva(t) = 4.09290 - 0.00228(t) \quad (4)$$

Temperature (Temp) OLR model:

$$Temp(t) = 27.162 + 0.00047(t) \quad (5)$$

Relative Humidity (Humid) OLR model:

$$\text{Humid}(t) = 78.9441 + 0.0088(t) \qquad (6)$$

Normalized data: The treatment data were transformed into normalized data by using Eq. 7 as follow:

$$\text{Normalized data} = \qquad (7)$$

In which the normalized treatment data are presented in Table 4. The data should be normalized in the range of 0-1 and after the data has been transformed to the equivalent normalized value, to change the data back to its original state, the original value after have been forecast can be determine by using back this formula. According to previous researches, it is important to make normalized the data because normalization of data is procedure for scaling the numbers in a data set to upswing the accuracy of the consequent numeric computations (Birbir *et al*., 2007; Wang, 1993).

Model evaluation performance: The evaluation of performance is essential with the purpose of finding the best neural network architecture, which gives the most reliable and accurate predictions. Based on previous researches, there are some performance function can be used to control the performance of network (Muhammad, 2009). Some might prefer the performance tool of the back-propagation algorithm is Mean Square Error (MSE) of training and testing (Hagan *et al*., 1996). Moreover, the selection of MSE is supported by MATLAB software, which also the default indicator in training the network. The neural network model with the smallest MSE value is considered to be the best neural network model.

Methods: The prediction models for water level in Dungun River were developed based two non linear time series regression which are BPNN and NARX. Since the neural network is a nonlinear procedure and the network parameters will affect each other, the adjustment of each parameter to optimize the whole network is not an easy task (Lee, 2006). In addition, these two methods will try to find the smallest variation in detecting appropriate predictors among the variables uses. By carry out this variable selection analysis, it will become one of the efforts to enhance the accuracy model in predicting the water level in Dungun River, Terengganu. The following discussions present the results between these two methods in variables selection methods and also the comparison between using normalized treatment using mean and OLR approaches.

Variables selection: The selection of appropriate predictors is one of the most important steps in enhancement of accuracy in prediction the water level of Dungun River. The predictors are chosen based on the smallest Mean Square Error (MSE) in two methods BPNN and NARX. The results will then be compared between two types of treatment data which is using mean and OLR. Next, we will discuss the results of variable selection analysis using BPNN and NARX methods respectively.

RESULTS

BPNN method: BPNN is performed in this experiment to build the model for water level in Dungun River. In these study, two types of sigmoid activation functions are selected for several numbers of hidden, k_1 and one output layer, k_2 which are logarithmic sigmoid function (logsig) and hyperbolic tangent sigmoid function (tansig). According to Zhang, both sigmoid functions are often used in hidden layer due to the ability of authoritative nonlinear approach (Zhang and Wu, 2009). The training algorithm used here is trainlm function that modified bias and weight values based on Lavenberg-Marquardt optimization. This part presents the results of the experiment using BPNN (k_1, k_2, k_3) where k_1 are the numbers of inputs or variables selected, k_2 are the numbers of hidden layer and k_3 are the numbers of output. From Table 5, we can see that for using BPNN method, BPNN 5-10-1 and BPNN 5-8-1 have their stable value of MSE in mean and OLR treatment respectively. In this case, five predictors namely months, rainfall, evaporation, temperature and relative humidity will be consider as the variables selected to model the water level in Dungun River. Otherwise, for using both mean and OLR treatment, it also have been detected that using four predictors namely months, rainfall, evaporation and relative humidity (BPNN 4-10-1) have stable MSE for its training and testing. In addition, three predictors which are months, rainfall and evaporation (BPNN 3-6-1 and BPNN 3-8-1) also give smallest and stable values of MSE in training and testing for mean and OLR treatment respectively. For two predictor using months and rainfall, there are also have stable MSE for mean treatment (BPNN 2-6-1 and BPNN 2-4-1).

In summary, the number of neurons in the input layer, k_1 was changed in several times and we used different numbers of hidden layer, k_2 neurons (between 4 and 10) to select the best variables in model water level in Dungun River. After several trials, the method of BPNN models gives smallest MSE values on the network BPNN 5-10-1, BPNN 5-8-1, BPNN 4-10-1, BPNN 3-6-1, BPNN 2-6-1 and BPNN 2-4-1 for using mean treatment approach. Otherwise, using OLR approach the smallest MSE values occur when the networks are BPNN 5-8-1, BPNN 4-10-1 and BPNN 3-8-1.

Table 5: The variables selection using BPNN

Method	Network	Mean Training	Mean Testing	OLR Training	OLR Testing
BPNN	5-10-1	0.0007	0.0015	0.0058	0.0039
	5-8-1	0.0006	0.0016	0.0008	0.0019
	5-6-1	0.0219	0.0071	0.0009	0.0015
	5-4-1	0.0009	0.0023	0.0009	0.0016
	4-10-1	0.0009	0.0017	0.0009	0.0019
	4-8-1	0.0010	0.0019	0.0010	0.0022
	4-6-1	0.0010	0.0015	0.0010	0.0025
	4-4-1	0.0009	0.0024	0.0010	0.0080
	3-10-1	0.0007	0.0018	0.0010	0.0016
	3-8-1	0.0010	0.0015	0.0003	0.0017
	3-6-1	0.0010	0.0013	0.0010	0.0018
	3-4-1	0.0009	0.0026	0.0009	0.0020
	2-10-1	0.0008	0.0020	0.0008	0.0025
	2-8-1	0.0009	0.0012	0.0010	0.0015
	2-6-1	0.0009	0.0011	0.0010	0.0019
	2-4-1	0.0009	0.0011	0.0010	0.0020

Table 6: The variables selection using NARX

Method NARX	Network Testing	Mean Training	Mean Testing	OLR Training	Testing
With d= 2	5-10-1	0.0003	0.0025	0.0004	0.0015
	5-8-1	0.0004	0.0034	0.0010	0.0056
	5-6-1	0.0003	0.0047	0.0009	0.0045
	5-4-1	0.0009	0.0020	0.0008	0.0015
	4-10-1	0.0007	0.0070	0.0007	0.0042
	4-8-1	0.0006	0.0035	0.0005	0.0058
	4-6-1	0.0008	0.0011	0.0010	0.0048
	4-4-1	0.0006	0.0020	0.0012	0.0056
	3-10-1	0.0006	0.0048	0.0011	0.0014
	3-8-1	0.0005	0.0092	0.0008	0.0025
	3-6-1	0.0008	0.0026	0.0012	0.0045
	3-4-1	0.0004	0.0091	0.0018	0.0004
	2-10-1	0.0004	0.0039	0.0026	0.0005
	2-8-1	0.0006	0.0015	0.0021	0.0098
	2-6-1	0.0009	0.0021	0.0007	0.0075
	2-4-1	0.0007	0.0015	0.0006	0.0016

NARX method: In this study, nonlinear autoregressive with exogenous (external) input or NARX will be used in order to make future prediction values of a time series, y (t), from past values of that time series and past values of a second time series, x (t). In these experiments, we also conducted NARX with variety numbers of hidden layers, numbers of tapped delay lines (d) and one output neuron with two layer feed forward networks (hyperbolic tangent transfer function in the hidden layer and linear transfer function in the output layer) were used in these experiments. The standard Lavenberg-marquardt backpropagation algorithm is used to train the network with learning rate equal to 0.001. The method regularization has been used which consist of 1000 epoch and regularization parameter used is 1.00e-05.

Table 7: The performance of optimization NARX 5-10-1 model using several delay lines (d)

value d	MEAN Training	MEAN Testing	R	OLR Training	OLR Testing	R
2	0.0003	0.0025	0.55	0.0003	0.0031	0.75
3	0.0002	0.0050	0.64	0.0002	0.0013	0.89
4	0.0001	0.0008	0.86	0.0001	0.0018	0.82

Table 8: The performance between BPNN and NARX for mean and OLR treatment data process

[Methods	Mean Training	Mean Testing	OLR Training	OLR Testing
BPNN 5-10-1	0.0007	0.0015		
BPNN 5-8-1	0.0006	0.0016		
BPNN 5-4-1	0.0009	0.0023		
BPNN 4-10-1	0.0009	0.0017		
BPNN 5-8-1			0.0008	0.0019
BPNN 4-10-1			0.0009	0.0019
BPNN 3-8-1			0.0003	0.0017
NARX 5-10-1 (d=4)	0.0001	0.0008		
NARX 5-10-1 (d=3)			0.0002	0.0013

Training automatically stops when generalization stops improving, as indicated by an increase in the Mean Square Error (MSE) of the validation samples.

From Table 6 it can be seen clearly that five predictors using in NARX methods (NARX 5-10-1) with tapped delay lines, d = 2, give smallest MSE in both treatment methods which are consists variables of months, rainfall, evaporation, temperature and relative humidity. In addition, model NARX 4-6-1 also give stable MSE values when using mean treatment analysis. Otherwise, for OLR treatment, model NARX 5-4-1 and NARX 3-10-1 give stable MSE for both training and testing value.

However, we assume that these NARX model still not in optimum analysis since we just used one number of tapped delay line (d) which is 2. Therefore, to optimize the NARX model, we were try another delay lines which are 2, 3 and 4 for NARX 5-10-1in both treatment analysis since this model already chosen as the best and stable among others NARX models. Table 7 shows the results of that experiment. From Table 7, we can see that the NARX 5-10-1 become more stable when the delay lines (d) are increase. We can see here that using mean treatment, NARX 5-10-1 with d = 4 have more stable MSE compared to others. Otherwise NARX 5-10-1 with d = 3 also have stable MSE when using OLR treatment. Therefore, several models will be used in water level prediction in Dungun River.

Model selection: For this analysis, we want to select the best model in order to predict water level in Dungun River based on performance of BPNN and NARX methods. The good performance is based on accuracy measurement using Mean Square Error (MSE) of both methods.
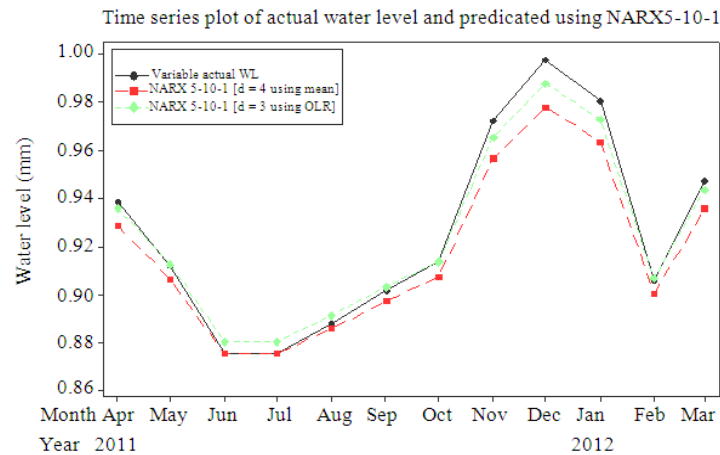
Fig. 2: The time series plot between actual and prediction monthly water level for Dungun River with test data in 2011 and 2012

From the previous discussion, we already see the performance of BPNN and NARX using two types of treatment analysis which is mean and OLR respectively. The type of treatment data either using the mean or OLR clearly have give influences on performance of prediction water level in Dungun River. Table 8 is the summarization of the results between two methods with their own treatment analysis. From Table 8, if we compare between two method BPNN and NARX results, we can see that the performance of NARX 5-10-1 for both treatment approach is more stable than BPNN.

DISCUSSION

Since NARX 5-10-1 is the best among the other NARX model, we will use NARX with 10 hidden layers with d = 4 for mean data treatment and d = 3 for OLR data treatment to predict the water level of Dungun River for the rest of twelve months. After that, the results of predicted water level will then be compared with the actual value of water level to test the performance of proposed models. Figure 2 shows the comparison between predicted and actual water level NARX 5-10-1(with d = 3 and d = 4). From this figure, we can see that the predicted value with NARX 5-10-1 are closer to actual value of water level in Dungun River.

CONCLUSION

The preprocessing data analysis using mean and OLR have been done to treat the missing data collected from DID and MMD. These two types of treatment data have been used to develop a model based two

nonlinear regressions which are BPNN and NARX method. After the experiments, it was that the treatment either using the mean or OLR approaches have influence the performance of prediction water level in Dungun River. The approaches from BPNN and NARX models were compared to find the best model to predict water level in Dungun River. After we optimized the NARX model using several tapped delay lines (d), we can clearly see that NARX model is better than BPNN. Therefore, in this study we can conclude that the suitable model to predict the water level at Dungun River are NARX 5-10-1 with tapped delay lines (d = 4) using mean and tapped delay lined (d = 3) using OLR treatment data process since it has best performance compared to BPNN and others approaches.

ACKNOWLEDGEMENT

REFERENCES

Arbain, S. and A. Wibowo, 2012. Time series methods for water level forecasting of Dungun river in Terengganu Malaysia. Int. J. Eng. Sci. Technol., 4: 1803-1811.

Birbir, Y., H.S. Nogay and V. Topuz, 2007. Estimation of total harmonic distortion in short chorded induction motors using artificial neural network. Proceedings of the 6th Conference on Applications of Electrical Engineering, (AEE'07), ACM Press, USA, pp: 206-210.

Gasim, M.B., J.H. Adam, M.E. Toriman, S.A. Rahim and H. Juahir, 2007. Coastal flood phenomenon in terengganu, Malaysia: Special reference to dungun. Res. J. Environ. Sci., 1: 102-109.

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

Horne, B.G. and C.L. Giles, 1995. An Experimental Comparison of Recurrent Neural Network. 7th Edn., MIT Press, pp: 697. clgiles.ist.psu.edu/papers/NIPS94.rnn.comp.pdf

Lee, T.L., 2006. Neural network prediction of a storm surge. Ocean Eng., 33: 483-494. DOI: 10.1016/j.oceaneng.2005.04.012

Little, R.J.A. and D.B. Rubin, 1987. Statistical Analysis with Missing Data. 1st Edn., John Wiley and Sons, New York, ISBN-10: 0471802549, pp: 278.

Muhammad, Z.M., 2009. Solid Model reconstruction using neural network and the mathematical representation. Master Thesis, UTM Skudai, Malaysia.

Outhwaite, W. and S. Turner, 2007. Handbook of Social Science Methodology. SAGE Publications London, ISBN-10: 1412901197, pp: 640.

Sharda, R., 1994. Neural networks for the MS/OR analyst: An application bibliography. Interfaces, 24: 116-130.

Suguna, N. and K.G. Thanuskodi, 2011. Predicting missing attribute values using K-means clustering. J. Comput. Sci., 7: 216-224. DOI: 10.3844/jcssp.2011.216.224

Toriman, M.E., H. Juahir, M. Mokhtar, G.M. Barzani and S.A.S. Mastura *et al*., 2009. Predicting for discharge characteristics in Langat River, Malaysia using neural network application model. Res. J. Earth Sci., 1: 15-21.

Vaziri, M., 1997. Predicting caspian sea surface water level by ANN and ARIMA models. J. Waterway, Port, Coastal Ocean Eng., 123: 158-162.

Wang, S., 1993. Neural network approach to generating the learning curve. Inform. Syst. Operat. Res., 31: 136-150.

Westin, L.K., 2004. Missing Data and the Preprocessing Perceptron. 1st Edn., Umea, pp: 26.

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.P., 2002. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50: 159-175. DOI: 10.1016/S0925-2312(01)00702-0

Zhang, Y. and L. Wu, 2009. Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network. Expert Syst. Appli., 36: 8849-8854. DOI: 10.1016/j.eswa.2008.11.028