

# Forecasting River Water Level Using Artificial Neural Network



Norwaziah Mahmud, Nur Syuhada Muhammad Pazil, Izleen Ibrahim,  
Umi Hanim Mazlan, Siti Hafawati Jamaluddin  
and Nurul Zulaikha Othman

**Abstract** Flood is a natural phenomenon that occurs without warning. The occurrence of flood can cause many damages in terms of property, money, and even death. In order to prevent more loss, an early warning system of the increment of river water level needs to be forecasted in a more accurate and effective way to predict the best result. In this study, Artificial Neural Network (ANN) is implemented in order to forecast the river water level in Pelarit River, Perlis. The objectives of this study are to determine the efficiency of ANN in forecasting the river water level and to forecast the water level of Pelarit River for one week in advance. This study uses Multilayer Feed Forward Artificial Neural Network with two input neurons, seven hidden neurons, and one output neuron in order to forecast the water level. Three different types of algorithms, namely quick propagation, conjugate gradient descent, and Levenberg–Marquardt algorithms, are compared to get the most significant result in forecasting one week of river water level. Among all the algorithms used, conjugate gradient descent is proved to be the most reliable value to forecast the Pelarit River in terms of the lowest value of Root-Mean-Square Error and the highest value of correlation.

**Keywords** Flood · River water level · ANN · Forecasting

## 1 Introduction

Flood is natural disaster that usually happens in many countries around the world including Malaysia. In Malaysia, out of 189 river basins, 89 are located in Peninsular Malaysia, 78 in Sabah, and 22 in Sarawak with the main channel flows directly to the South China Sea. The types of flood in Malaysia can be categorized into flash flood and monsoon flood. In December 2005 and April 2006, the worst

---

N. Mahmud (✉) · N. S. Muhammad Pazil · I. Ibrahim · U. H. Mazlan  
S. H. Jamaluddin · N. Z. Othman  
Universiti Teknologi MARA Cawangan Perlis, Kampus Arau, Perlis, Malaysia  
e-mail: norwaziah@perlis.uitm.edu.my

flood occurred in Perlis after 30 years where two-thirds of Perlis were affected (Lawal et al. 2014). In addition, flood that occurred in Perlis in the year 2010 had caused a damage worth RM 34,747 224 (D/iya et al. 2014).

Every river in Perlis has different water level depending on their own depths taken from the sea level. The function of water level system is to produce a sign of immediate action that needs to be taken by authorized organizations when heavy rainfalls occurred in certain areas. It is important to analyze the accuracy of the river water level in order to take an immediate action to prevent more loss and damage. Therefore, the main purpose of this study is to apply a forecasting technique to predict the river water level in Pelarit River at Kaki Bukit in Perlis.

## 2 Artificial Neural Network (ANN) Modelling

Artificial Neural Network (ANN) is created based on the description of behavior on human's perceptions into a simplified computational system or mathematical model. ANN is commonly used in forecasting field as it is suitable to be implemented for problems that have complete data and observation but the solutions require specific information that is difficult to be identified (Suliman et al. 2013). It is also a good model in forecasting because of its ability in solving dynamic nonlinear time series problems (Bustami et al. 2007).

Basically, this model includes three types of layer, namely input layer, hidden layer, and output layer, in which all the layers are interconnected among the simple elements called artificial neurons (Alvisi et al. 2006). The input layer consists of number of neurons or nodes where all the neurons will work together in a parallel form to generate information for the final output layer (Kia et al. 2012). The parameters in input layer depend on the received data obtained from different sources, or it may come from the investigators' or researchers' opinion. The input layer may also contain the influences of the phenomenon that are being investigated (Sztobryn 2013). For the hidden layer, the number of neurons is processed as a trial and error procedure in order to search for the lowest number of neurons without disturbing the model efficiency. Each of the hidden neurons will respond according to the neuron connection via input layer. All the elements inside the hidden layer will transform into the nonlinear transfer function (Kia et al. 2012).

## 3 Methodology

### 3.1 Data Acquisition

This study will be using the river water level of Pelarit River that was recorded daily from January 1, 2012, to December 31, 2014. The recorded data were

obtained from the National Hydrological Network Management System (SPRHiN). The water level system is categorized into four types, namely normal level, alert level, warning level, and danger level. According to the Department of Irrigation and Drainage of Perlis, for Pelarit River, the normal level is 35.60 m, whereas the maximum water level which is also the danger level is 39.00 m. The water level reaches the alert level when the water level rises to 38.60 and 38.72 m is the warning level.

### 3.2 Data Preprocessing

Data preprocessing is implemented when there exists a complex data such as non-linearity and non-stationary data which cause difficulty in forecasting the case study using the chosen models. In ANN, preprocessing is a modification of the data before it can be used in the neural network, and it is a data transformation to develop a suitable neural network. Furthermore, data preprocessing also improves the quality of data because it involves filtering the outliers and approximating the missing values. The activities involved in data preprocessing are data cleaning, data reduction, outlier, and data normalization.

#### 3.2.1 Data Cleaning

In order to obtain unbiased result, missing value must be overcome first. To solve the problem of missing value in the data, this study will be using average method. The average of the daily river data is calculated separately for every month from 2012 to 2014 using Waikato Environment for Knowledge Analysis (WEKA) to replace the missing value.

#### 3.2.2 Data Reduction

The total of 1096 of daily water level data will be reduced to the average of 156 weeks in order to forecast the average of one week of river water level. The following formula will be applied for the data reduction:

$$\text{Weekly Data}_k = \sum_{n=0}^6 \text{Raw Water Level Day}_{n+1} / 7, \quad (1)$$

where  $k = 1, 2, 3, \dots, 156$  and  $n = 0, 1, 2, \dots, 1095$ .

### 3.2.3 Outlier

After the data reduction, the presence of the outliers has been detected at week 144 and week 155 where the values of the average water level are the highest. For this study, the numeric outlier is substituted by median since it will not be affecting the forecasting of one week of river water level.

### 3.2.4 Data Normalization

Data normalization is used to scale the data into (0, 1) and (-1, 1). The input of the average water level data at time  $t$  and the lag average water level at time  $t-1$  will be normalized using the formula in Eq. (2) (Gazzaz et al. 2012), while the output of the next water level will be using logistic sigmoid activation function as in Eq. (3).

$$X_s = [(b - a) * (X_0 - X_{\min}) / (X_{\max} - X_{\min})] + a, \quad (2)$$

where  $X_s$  and  $X_0$  express the normalized and raw observations of parameter  $X$ , respectively, while  $a$  and  $b$  represent the lower and upper limits of the normalization range.  $X_{\min}$  and  $X_{\max}$  are minimum and maximum values of parameter  $X$ .

$$y = \frac{1}{(1 + e^{-kx})}, \quad (3)$$

where  $y$  is the sigmoid value,  $k$  is the sigmoid steepness coefficient, and  $x$  represents the data value, namely the total number of the input and weight values.

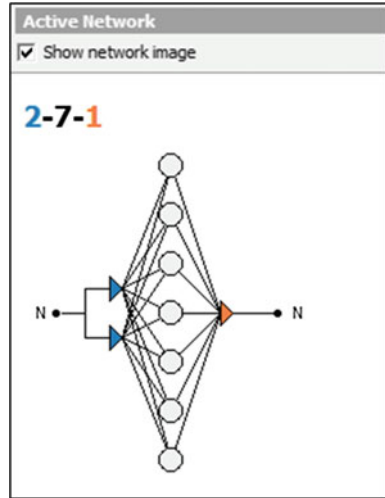
## 3.3 Network Design

Five possible network architectures are verified using the Artificial Neuro-Intelligence software. Later on, the architecture design is chosen based on the highest number of correlation coefficient,  $r$ , where the closer the value of  $r$  to 1, the better the network is. Figure 1 shows the best architecture design for this study that consists of two neurons of input layer, seven neurons in hidden layer, and one neuron for output layer, (2-7-1).

## 3.4 Training an ANN Using Algorithms

This study chooses three different algorithms, namely quick propagation, conjugate gradient descent, and Levenberg–Marquardt, in order to identify the best algorithms that can produce better results. For this study, the number of iteration that represents

**Fig. 1** Best architecture design for Pelarit River



a complete presentation of the training set to the network training is 500. Furthermore, the value of 0.001 is chosen as the mean square error (MSE) as the stopping condition for the over-fitting and minimum error.

### 3.5 Testing the Algorithms

After the training process is completed, the forecasted values and the actual values are verified. Based on the acquired results, it shows that the conjugate gradient descent algorithm is the most efficient algorithm in forecasting the water level for Pelarit River.

### 3.6 Performance Evaluation

The performance of the actual water level and the corresponding neural network is evaluated by using different measurements such as Root-Mean-Square Error, RMSE (Eq. 4), and Nash–Sutcliffe, NS (Eq. 5). RMSE describes the average magnitude of error of the observed and predicted values, while NS coefficient of efficiency is widely used to describe the forecasting accuracy (Sulaiman et al. 2011).

$$NS = 1 - \frac{\sum_{i=1}^N (O_i - F_i)^2}{\sum_{i=1}^N (O_i - \bar{O})^2}, \tag{4}$$

$$RMSE = \sqrt{\frac{\sum_{i=0}^N (O_i - F_i)^2}{N}}, \tag{5}$$

where  $O_i$  is represented as an actual value,  $F_i$  is represented as a predicted value,  $\bar{O}$  is mean of the actual value, and  $N$  is the number of data being evaluated. Error is calculated from the differences between the actual and predicted values. The formula is given as the following:

$$\text{Error} = (O_i - F_i), \tag{6}$$

where  $O_i$  is an actual value and  $F_i$  is a predicted value. The values obtained indicate the accuracy of the forecasting results for the developed model.

## 4 Results and Discussions

### 4.1 Analysis of Training Algorithms

Based on the training process from quick propagation algorithm, conjugate gradient descent algorithm, and Levenberg–Marquardt algorithm as shown in Table 1, the best result for the training process in this study is conjugate gradient descent

**Table 1** Results of the training process for quick propagation, conjugate gradient descent, and Levenberg–Marquardt algorithms

Algorithm						
Parameter	Quick propagation		Conjugate gradient descent		Levenberg–Marquardt	
	Training	Validation	Training	Validation	Training	Validation
Absolute error	0.1040	0.1131	0.0944	0.1522	0.1686	0.1503
Network error	0.0085	0	0.0077	0	0.0152	0
Error improvement	0.000006		0.000003		0.000321	
Correlation coefficient, $r$	0.8482	0.8307	0.8487	0.8463	0.8408	0.8412
R-squared	0.5994	0.6430	0.5791	0.6782	0.2651	0.3397

algorithm with the lowest value of absolute error and the highest value of correlation that indicate that there is a strong positive relationship between the actual values and the network outputs. Apart from that, conjugate gradient descent also has the smallest network error compared to others.

Based on the result, the value of network error of conjugate gradient descent algorithm in the training set is 0.0077 which is higher than the network error in the validation set, 0, and this proves that the overtraining is being controlled during the iterations. Overtraining usually occurs when the value of network error in validation set is increasing, while the value of the network error in training set is decreasing. Moreover, the lowest value of the error improvement for conjugate gradient descent algorithm also indicates the ability of the neural network to improve the forecast value.

### 4.2 Analysis of Performance Algorithm

The performance of the actual river water level and the forecasted is evaluated by using the Root-Mean-Square Error (RMSE) and Nash–Sutcliffe (NS) as shown in Table 2. Based on the table, it shows that the value of NS for each of the algorithms reached the optimal value. However, RMSE of conjugate gradient descent algorithm has the minimum value compared to others.

### 4.3 Analysis of River Water Level Forecasting

Table 3 shows the forecast value of water level for Pelarit River in week 157 for conjugate gradient descent algorithm. Based on the result, the average one week ahead shows a decreasing pattern from the previous week.

**Table 2** RMSE and NS for quick propagation, conjugate gradient descent, and Levenberg–Marquardt algorithms

	Quick Propagation	Conjugate gradient descent	Levenberg–Marquardt
RMSE	0.1862	0.1860	0.2098
NS	1.000	1.000	1.000

**Table 3** Forecast value for the average of one week ahead

Week	Average water level (m)
156	36.13
157	35.98

## 5 Conclusion

In conclusion, the conjugate gradient descent algorithm is proved to be the most reliable value to forecast the water level of Pelarit River. The forecast value for week 157 is 35.98 m which is lower than 36.13 m in week 156. Therefore, the forecast value indicates that the water level is at a normal level. It is hoped that this study can be applied to forecast river water level in various areas that have risk of flood. In future, it is also recommended to consider other variables such as weather, river water level from the nearest river, and time taken in collecting the data.

## References

- Alvisi, S., Mascellani, G., Franchini, M., & Bárdossy, A. (2006). Water level forecasting through fuzzy logic and neural network approaches. *Hydrology and Earth System Sciences*, *10*(1), 1–17.
- Bustami, R., Bessaih, N., Bong, C., & Suhaili, S. (2007). Artificial neural network for precipitation and water level predictions of Bedup River. *IAENG International Journal of Computer Science*, *34*(2), 228–233.
- D/iya, S. G., Gasim, M. B., Toriman, M. E., & Abdullah, M. G. (2014). Floods in Malaysia historical: Reviews, causes, effects and mitigations approach. *International Journal of Interdisciplinary Research and Innovations*, *2*(4), 59–65.
- Gazzaz, N. M., Yusoff, M. K., Aris, A. Z., Juahir, H., & Ramli, M. F. (2012). Artificial neural network modelling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors. *Marine Pollutin Bulletin*, *64*(11), 2409–2420.
- Kia, M. B., Pirasteh, S., Pradhan, B., Mahmud, A. R., Sulaiman, W. N. A., & Moradi, A. (2012). An Artificial Neural Network model for flood simulation using GIS: Johor River Basin, Malaysia. *Environmental Earth Sciences*, *67*, 251–264.
- Lawal, D. U., Matori, A. N., Yusuf, K. W., Hashim, A. M., & Balogun, A. L. (2014). Analysis of the flood extent extraction model and the natural flood influencing factors: A GIS-based and remote sensing analysis. *IOP Conference Series: Earth and Environmental Sciences*, *18*, 1–6.
- Sulaiman, M., El-Shafie, A., Karim, O., & Basri, H. (2011). Improved water level forecasting performance by using optimal steepness coefficients in an artificial neural network. *Water Resources Management*, *25*, 2525–2541.
- Suliman, A., Nazri, N., Othman, M., Malek, M. A., & Ku-Mahamud, K. R. (2013). Artificial neural network and support vector machine in flood forecasting: A review. In *Proceeding of the 4th international conference on computing and informatics*, pp. 237–332.
- Sztobryn, M. (2013). Application of Artificial Neural Network into the water level modelling and forecast. *The International Journal on Marine Navigation and Safety of Sea Transportation*, *7* (2), 219–223.