

FLOOD PREDICTION FOR KLANG RIVER USING MUSKINGUM AND ANN MODELS

(Date received: 09.10.2018/Date accepted: 18.05.2019)

Jer Lang Hong¹ and Kee An Hong¹

¹ Hong and Associates

Email: jerlang.hong@taylors.edu.my, keeahong@yahoo.co.uk

ABSTRACT

Temporal and spatial variations of a flood hydrograph moving through a river reach can be simulated using flood routing tools such as hydrodynamic, hydrological and the ANN (Artificial Neural Networks) models. The ANN models have emerged as viable tools in flood routing and are widely adopted for this purpose. The aim of this study is to make an objective comparison of these two flood routing models to evaluate their individual performance. Four flood events recorded for Klang river at Kuala Lumpur in the period October 1973 to December 1974 for stations at Leboh Pasar and Sulaiman Bridge which are 950m apart were used for this study. The statistical performance of the models is assessed using criteria such as peak flow, root mean square error, mean absolute error and Nash –sutcliffe coefficient. Results from calibration runs for the 02/05/1974 flood event show that the MAE, RMSE and NAE for ANN and Muskingum models are 0.75, 1.24, 0.9917 and 1.1, 1.3, 0.992 respectively. The performance of the two models was verified using three other different events. Results of simulation runs for the 10/12/1974 event gave 2.72, 3.24, 0.96 and 2.1, 3.1, 0.963 MAE, RMSE and NAE values for ANN and Muskingum. Graphical inspections and statistical tests show that the ANN and Muskingum methods performed equally well in flood prediction for this study, using the flood events of Klang river.

Keywords: ANN Muskingum, Calibration, Flood Routing NAE, RMSE, Validation

1.0 INTRODUCTION

Flood routing technique is important as they are applied in flood forecasting studies, reservoir and channel design, flood plain and watershed simulation. Understanding of the complex dynamics processes of flood routing has been made through the development of conventional hydraulic and hydrologic models, e.g. Muskingum method, HEC-HMS [1] and the data driven ANN approach e.g. Thirumalaiah and Deo [2], Sherestha *et al.*, [3] Razavi and Karamouz [4], Hong and Hong [5]. The fundamental theory and assumptions of the models are entirely different in model structure, data requirements and capabilities, it is therefore worth to perform an objective comparison for these models using locally available flow data.

2.0 MATERIALS AND METHODS

The following sections present the details of floods chosen for analysis and the methods used in river routing for the Klang river.

2.1 Flood Hydrographs of Klang River

Klang river has been gauged at several locations (Figure 1). The gauging stations at Leboh Pasar Kuala Lumpur and at Sulaiman Bridge in the centre of Kuala Lumpur city are within a distance of 950m. Concurrent discharge data are available from the automatic recorders from these stations from October 1973 to December 1974. Details of the gauging stations and records available are shown in Table 1. Examination of the flow records shows that several flood events in this period are available and can be adopted for river routing analysis. Selected flood events

used in this study are shown in Table 2. Figure 2 shows the selected flood hydrographs of 24/11/1973.

Table 1: Details of the Gauging Records

Station name	Station ID	Period of Record	Catchment Area km ²
Klang at Leboh Pasar Kuala Lumpur	3116432	Jul 1960 - Dec 1974	464
Klang at Sulaiman Bridge	3116430	Oct 1973 - to date	468

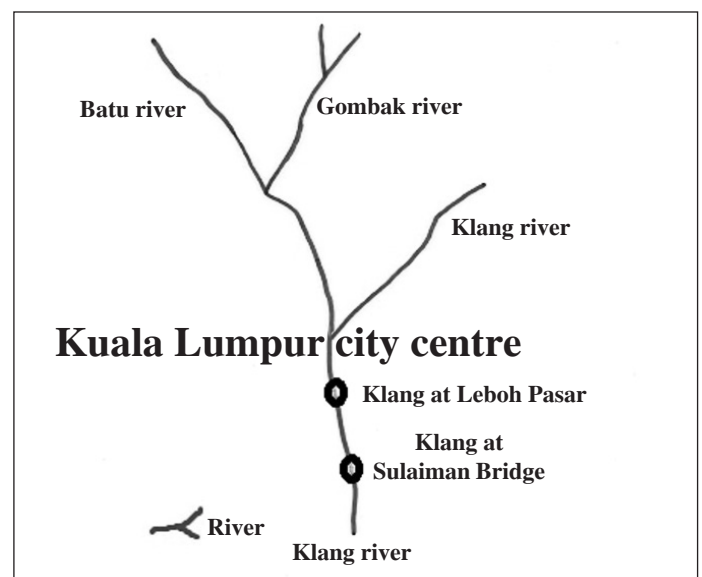


Figure 1: Gauging stations of Klang river

Table 2: Selected floods used for routing analysis

Flood Event	Period	Peak flow m ³ /s	
		At Leboh Pasar	At Sulaiman Bridge
1	22/12/1973 10:00 to 24/12/1973 20:30	75.5	75.7
2	24/11/1973 13:00 to 27/11/1973 16:00	110.9	93.6
3	02/05/1974 15:00 to 03/05/1974 19:00	63.5	63.8
4	10/12/1974 10:00 to 13/12/1974 07:00	69.3	70.8

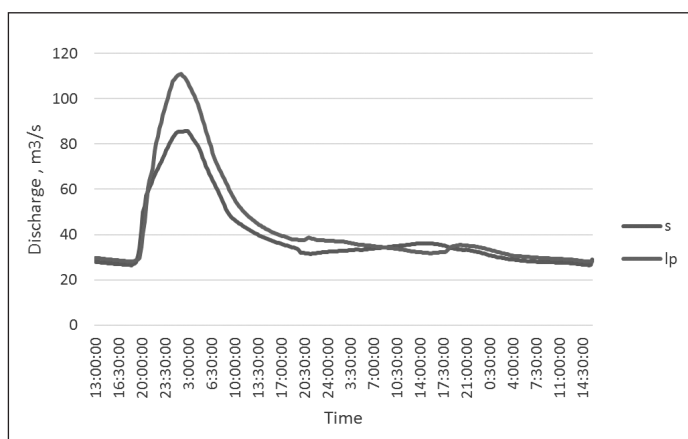


Figure 2: Flood hydrographs of 24/11/1973

2.2 Muskingum Method

The Muskingum method was developed to directly accommodate the looped relationship between storage and outflow that exists in rivers. With the Muskingum method, storage within a reach is visualized in two parts: prism storage and wedge storage. Prism storage is essentially the storage under the steady flow water surface profile. Wedge storage is the additional storage under the actual water surface profile. As shown in Figure 3, during the rising stages of the flood wave the wedge storage is positive and added to the prism storage. During the falling stages of a flood wave, the wedge storage is negative and subtracted from the prism storage.

Prism storage is computed as the outflow (O) times the travel time through the reach (K). Wedge storage is computed as the difference between inflow and outflow ($I-O$) times a weighting coefficient x and the travel time K . The coefficient K corresponds to the travel time of the floodwave through the reach. The parameter x is a dimensionless value expressing a

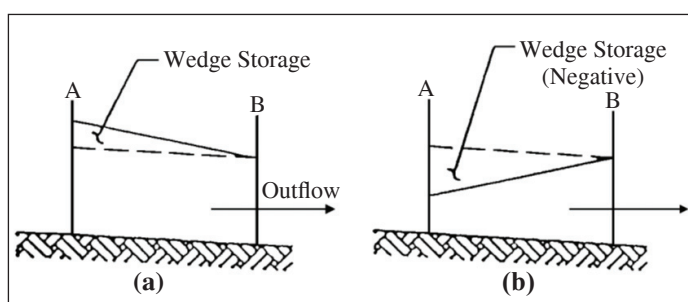


Figure 3: Muskingum prism and wedge storage concept

weighting of the relative effects of inflow and outflow on the storage (S) within the reach. Thus, the Muskingum method defines the storage in the reach as a linear function of weighted inflow and outflow:

$S = \text{prism storage} + \text{wedge storage}$

$$S = KO + Kx(I - O)$$

$$S = K [xI + (1 - x)O] \tag{1}$$

where

S = total storage in the routing reach

O = rate of outflow from the routing reach

I = rate of inflow to the routing reach

K = travel time of the floodwave through the reach

x = dimensionless weighting factor, ranging from 0.0 to 0.5

The quantity in the brackets of Equation 1 is considered an expression of weighted discharge. When $x = 0.0$, the equation reduces to $S = KO$, indicating that storage is only a function of outflow, which is equivalent to level-pool reservoir routing with storage as a linear function of outflow. When $x = 0.5$, equal weight is given to inflow and outflow, and the condition is equivalent to a uniformly progressive wave that does not attenuate. Thus, “0.0” and “0.5” are limits on the value of x , and within this range the value of x determines the degree of attenuation of the floodwave as it passes through the routing reach. A value of “0.0” produces maximum attenuation, and “0.5” produces pure translation with no attenuation.

The routing period T of the Muskingum method should be rounded off to an even interval for ease of calculation, and should be selected to conform generally with the following conditions:

$$T \leq 0.25T_R \tag{2}$$

Where T_R is the time of rise of the inflow hydrograph

$$T \leq K \tag{3}$$

$$T \geq 2Kx \tag{4}$$

2.3 The ANN Model

Artificial Neural Networks (ANN) are nonlinear and flexible massively parallel distributed information processing systems. For a number of non linear processing units, it is possible to train the neural networks to learn from experience and compute the complex functional relationships with accuracy. A number of neural networks has been proposed in the literature but the most commonly method used in hydrology for flood forecasting is the feed forward Multi Layer Perceptron (MLP) model. As an example, a typical three layer feed forward MLP model is shown in Figure 4. For most flood forecasting studies carried out in the past two decades, the three layer MLP model was used with the input nodes consist the lagged rainfall and streamflow or water level values and the output is the forecasted future value. In this particular MLP model, hidden nodes are used to process the information transmitted from the input nodes with a particular non linear transfer function. For this study, the model considered is a single output MLP model. The network is processed through training, testing and validation stages in order to forecast the flood level using the input upstream data. Back Propagation (BP) algorithm [6] is usually used to correct the weights of the interconnecting neurons. Back Propagation (BP) uses the steepest gradient descent method to correct the weight of the interconnecting neurons. This method (BP) solves

the interconnection of the processing of processing elements by adding hidden layers. For the learning process in the back propagation method, the interconnection weights are adjusted using the error convergence method to obtain a desired output from an input. The BP algorithm propagates the error at the output to the input layer through the hidden nodes to obtain the final output. The gradient technique is used to calculate the weight of the network and adjust the weight of the interconnections to minimize the output error.

BP uses the following equation [7] to correct the weighting factor:

$$\Delta w_{ij}(n) = \alpha \Delta w_{ij}(n-1) - \varepsilon \left(\frac{\partial E}{\partial w_{ij}} \right) \quad [5]$$

Where $\Delta w_{ij}(n)$ and $\Delta w_{ij}(n-1)$

are weights interconnecting nodes i and j during the n th and $(n-1)$ th steps.

α is the momentum factor used to speed up training in flat regions of the error surface and helps to prevent oscillations in the weights.

A learning rate ε is used to increase the chance of avoiding the training process being trapped in a local minima instead of a global minima.

The number of neurons in the input and output layers are problem dependent and decided by the number of input and output variables in the MLP model. The size of hidden neurons is an important factor in solving the problems using MLP. There are no fixed rules in determining the number of hidden neurons required for the model and trial and error experiments are normally adopted to determine the hidden node that gives the model the best performance. However, empirical relationships between optimum hidden neurons and number of input and output elements were given by some authors e. g. Mishra *et al.*, [8] used $2n+1$ for estimating the number of hidden neurons. Where n is the number of input neurons

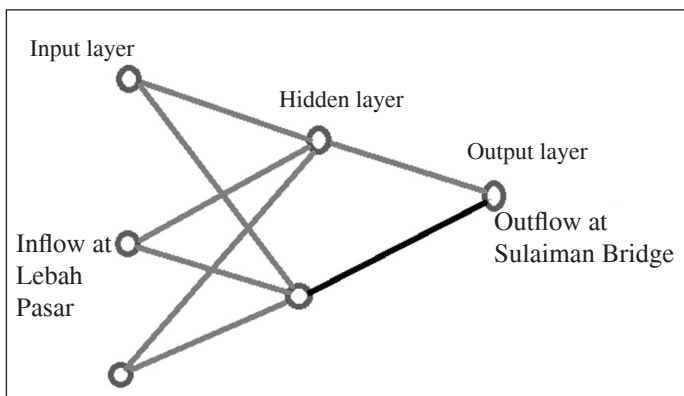


Figure 4: A three layer MLP neural network method

The neural network add-in version 1.5 software developed by the University of Adelaide [9] was used for the training of the ANN model. The activation function used is the logical sigmoid function.

For training purposes, data are normalised using the scaling method. Input variables are selected using the partial mutual information selection option and data are split randomly with 60%, 20%, 20% for training, testing and validation purposes using the program.

The data is trained using learning rate of 0.01 and momentum coefficient of 0.9.

2.4 Performance of The Routing and Data Driven Models

The performance of the ANN model and Muskingum routing model in predicting the flood flows at downstream Sulaiman Bridge station are assessed using:

$$NAE = 1 - \frac{\sum_1^n (Q_o - Q_p)^2}{\sum_1^n (Q_o - \bar{Q}_o)^2} \quad [6]$$

$$RMSE = \sqrt{\frac{\sum_1^n (Q_p - Q_o)^2}{n}} \quad [7]$$

$$MAE = 1/n \sum_1^n |Q_o - Q_p| \quad [8]$$

Where Q_o and Q_p represent the observed and predicted flows, n = total number of observations
 \bar{Q}_o is the mean of observed flows

MAE is the mean absolute error

RMSE is root mean square error

NAE is the Nash-Sutcliffe efficiency coefficient

2.5 Input Data Preparation

Of the four events selected, the 02 May 1974 flood event is used for calibration and the 24/11/1973, 10/12/1974 and the 22/12/1973 events are used for testing the performance of both the Muskingum and ANN models.

To use the data of Leboh Pasar as input to forecast the flow at Sulaiman Bridge downstream using the Muskingum routing method and the ANN data driven model, it is necessary to:

- Adjust the downstream flow volume so that the volume of inflow is equal to the outflow volume
- The travel time of flood wave from Leboh Pasar to Sulaiman Bridge
- Suitable time interval for accurate estimation of flow at Sulaiman Bridge

As the area between the gauging stations is small and no tributaries enter the reach between the stations, the runoff volume of inflow hydrograph should equal that of outflow hydrograph. Therefore, within the analysis period, if the inflow runoff volume is found to be higher than the outflow volume, the outflow hydrograph was scaled proportionally so that the volume of flow at upstream side is equal to that of downstream as there are no records available for local inflow. The start and end times for the analysis are selected so that average discharges in the reach were approximately equal in the start and end time of the hydrographs.

Flow data are available for various time intervals and to obtain more accurate flood forecasting, a time interval of 15 minutes was adopted.

The flood wave travel time from Leboh Pasar to Sulaiman Bridge is found to be 0.8 hour (48 minutes) from Muskingum

routing analysis. Considering a computation time interval of 15 minutes, a 45 minutes flood wave travel time is adopted for ANN model.

2.6 Calibration and Testing of Muskingum Method

The Muskingum model was calibrated using the HEC-HMS computer software of Hydrologic Engineering Centre [1] for the 02/05/1974 event. The parameters obtained are then used to assess the performance of the 3 testing events, namely 24/11/1973, 10/12/1974 and 22/12/1973.

2.7 ANN Mode Architecture

The goal of the ANN is to generalize a relationship in the form

$$Y^m = f(X^n) \tag{9}$$

Where X^n is the n dimensional input vector consisting of variables $X_1, \dots, X_i, \dots, X_n$; Y^m is an m-dimensional output vector consisting of the resulting variables $Y_1, \dots, Y_i, \dots, Y_m$. In this study, value of X_i is the upstream station flow with a lag of $3 \Delta t$ with Δt at 15 minutes, (a total lag of 45 minutes) and Y_i is the flow at Sulaiman Bridge three time steps ahead of the upstream station flow at 15 minutes increment.

The ANN model consists of flow input data from upstream stations (Leboh Pasar) with a lag of three 15 minutes intervals, for the 4 selected flood events. Details are shown in Table 3.

Table 3: The ANN model for Klang River

Type of Node	Site	Variable
Input	Leboh Pasar flow	$X_{t-3\Delta t}$
Output	Sulaiman Bridge	Y_t

Note : Time increment 15 minutes

$X_{t-3\Delta t}$ Leboh Pasar flow at time $t-3\Delta t$

Y_t Sulaiman Bridge flow at time t

For this study, the input variable adopted is the upstream flow at Leboh Pasar which occurred at a time of $3\Delta t$ (45 minutes) before the Sulaiman flow since there are no significant lateral inflows entering the river reach between the two gauging stations.

After inputs were determined, the ANN models were optimized to obtain the best prediction model. The number of nodes was changed in the hidden layer to determine the optimum number for the ANN model.

3.0 RESULTS AND DISCUSSION

3.1 General Set-up of the Flood Prediction Computations

The Muskingum and ANN routing methods have been applied to the 4 flood events of Klang river (Table 2) mentioned earlier and the results obtained were compared to the observed hydrographs at the outflow location. This includes the processes of calibration

and verification using computer program HEC-HMS and ANN model. The calibration process is to optimise the model parameters using inflow and outflow hydrographs recorded at the gauging stations. After successful calibration, the model parameters were used in the validation (verification) process using selected floods of different events. Results are presented in the sections as follows:

a. Calibration Results

This is a summary of the calibration results on optimised parameters of srlected models and the goodness of fit indices for the observed and predicted hydrographs. Model parameters ascertained are used in validation runs.

b. Verification Results

This is a summary of the simulation runs using the model parameters determined from calibration runs to obtain hydrographs for the selected flood events and the goodness of fit statistical parameters.

c. Comparison of Predicted and Observed Hydrographs For Verification Runs

Hydrographs from verification runs are used to compare with observed hydrographs to see the goodness of fit of the predicted and observed hydrographs visually.

Flow data of Leboh Pasar at 15 minutes interval and at a lag of 45 minutes were used as inputs for the Muskingum routing and ANN models to predict the flow at Sulaiman Bridge of Klang river. Results of the calibration and testing runs for the four selected flood events are presented and discussed as follows:

3.2 Calibration Results

Calibration results for the 02/05/1974 event are presented in Table 4 and Table 5 for the adopted models. The best ANN model is the 1 input node, 5 hidden node and 1 output node model structure.

Table 4: Calibration results of Muskingum Method

Flood Occurred On	K	No. of subreaches	x	Peak flow m3/s		MAE	RMSE	NAE
				Qo	Qs			
02/05/1974	0.8	1	0.142	63.8	62.7	1.1	1.3	0.992

Table 5: Calibration Results of ANN Model

Flood Occurred On	Peak flow m3/s		MAE	RMSE	NAE	Model Structure
	Qo	Qs				
02/05/1974	63.8	62.9	0.75	1.24	0.9917	1-5-1*

*Denotes a network with 1 input node 5 hidden node and 1 output node.

Results show that the ANN produced lower MAE and RMSE values but also lower NAE. Best fit is achieved if MAE or RMSE approached 0 and NAE approached 1.

3.3 Test Results

The Muskingum parameters obtained from the 02/05/1974 event and the ANN model obtained from calibration for the same event are used to predict the flows for the test events. Test results for the selected events are presented in Table 6 and Table 7 for the adopted models.

Table 6: Test results of Muskingum method for selected flood events

Flood Occurred On	Peak flow m ³ /s		MAE	RMSE	NAE
	Q _o	Q _s			
24/11/1973	93.6	109.8	4.2	6.7	0.852
10/12/1974	70.8	68.1	2.1	3.1	0.963
22/12/1973	75.7	75	1.7	4	0.934

Table 7: Test results of ANN method for selected flood events

Flood Occurred On	Peak flow m ³ /s		MAE	RMSE	NAE
	Q _o	Q _s			
24/11/1973	93.6	110.8	4.16	6.83	0.844
10/12/1974	70.8	69.6	2.72	3.24	0.960
22/12/1973	75.7	75.8	1.81	3.204	0.957

The results indicate that both the Muskingum and ANN models are able to produce good results. The results are almost similar for the two models compared to observed flows.

3.4 Verification the Flood Routing Models with the Klang Flood Events

In order to assess the ability of the models in reproducing the observed floods, the floods simulated in the verification runs of Klang river at the outflow gauging site were used to compare with the observed floods. Details are presented as follows:

a. The Klang 24/11/1973 Flood

The corresponding hydrographs for the flood event are shown in Figure 5. The observed and calculated discharges at the reach outlet at Sulaiman Bridge for the two models are shown. From visual inspection of the hydrographs, the two routing models appear to be equally good. However, the goodness of fit test show that the Muskingum overall fits better than the ANN model. Sharp rise events such as this flood are more difficult to predict using the kinematic wave type of models such as Muskingum and Lag routing methods.

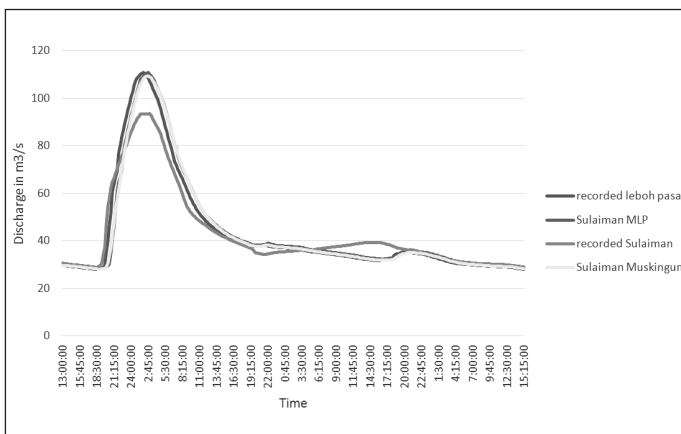


Figure 5: Observed, Muskingum and ANN results for the 24/11/1973 event

(b) The Klang 10/12/1974 Event

The second verification run of the routing models is based on the 10/12/1974 flood. The results are plotted in Figure 6. From graphical inspections, the hydrographs predicted from the two models are similar. The agreement with the observed hydrograph is also good. However, for this particular case, in terms of goodness of fit, the ANN model performs better.

(c) The Klang 22/12/1973 Flood

The third verification run of the routing models is based on the 22/12/1973 flood. The results are plotted in Figure 7. From graphical inspections, the hydrographs predicted from the two models are similar. The agreement with the observed hydrograph is also good.

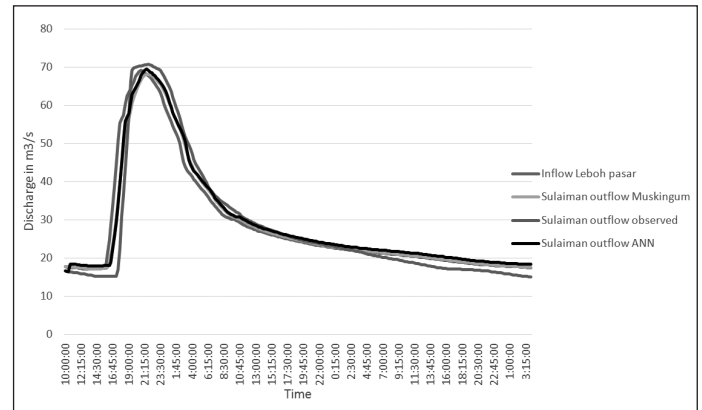


Figure 6: Observed, Muskingum and ANN results for the 10/12/1974 event

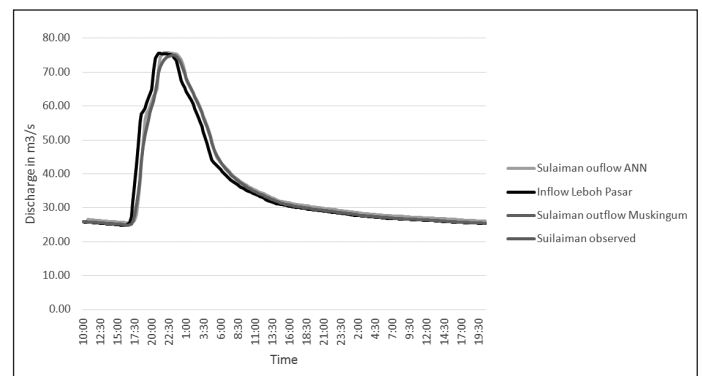


Figure 7: Observed, Muskingum and ANN results for the 22/12/1973 event

4.0 CONCLUSION

This study made an objective comparison using the Muskingum and ANN models in predicting downstream flood flows of Klang river at two gauging stations 950m apart. Generally, the Muskingum and ANN models both are able to produce good results. The results are almost similar for the two models compared to observed flows. These are seen when the modeled and observed hydrographs were compared graphically. However, in terms of model performance, the Muskingum model is slightly superior as is shown in the goodness of fit indices such as NAE, MAE and RMSE. As results obtained from ANN are almost congruent with those of Muskingum model, it is considered that both the methods are suitable for flood predictions for the study area.

5.0 ACKNOWLEDGEMENT

The permission of Drainage and Irrigation Department to use the data for this study is gratefully acknowledged. ■

6.0 REFERENCES

- [1] Hydrologic Engineering Centre (2016) Hydrologic Engineering Systems, User Manual US Army Corps of Engineers.
- [2] Thirumailah K. , Deo M.C. (2000) Hydrological forecasting using neural network Journal of hydrologic engineering.
- [3] Shrestha R. R. and Nestmann F. (2005) River Water Level Prediction Using Physically Based and Data Driven Models Allen Institute.
- [4] Razavi S. and Karamouz M. (2007) Adaptive Neural Networks for Flood Routing in River Systems, Water International, vol. 32.
- [5] Hong J. L. and Hong K. A. (2016), Flood Forecasting for Klang River at Kuala Lumpur using Artificial Neural Network International Journal of Hybrid Information Technology Vol. 9, No. 3.
- [6] Rumelhart D. E., Hilton G.e., William R. J. (1986). Learning representations by back propagation errors, Nature 313, 533-536.
- [7] ASCE Task Committee (2000) Artificial neural networks in hydrology 1 Preliminary concepts Jour. Of hydrologic engineering 5(2) 115-123.
- [8] Mishra A. K., Desain V.R. (2006) Drought forecasting using feed forward recursive neural network, Ecological modeling, Vol. 198, No. -2 pg 117-138.
- [9] School of Engineering, Adelaide University (2014) Neural networks add in, version 1.5.

PROFILES



DAVID HONG JER LANG is currently a lecturer of Taylor's University, Malaysia. He received his B. Sc degree in computer science from Nottingham University in 2005 and the Ph. D. degree from Monash University in 2010. He is author and co-author of several Association for Computing Machinery/Institution of Electrical and Electronic Engineers (ACM/IEEE) conference papers and is also the author of several papers in automatic wrapper design published in computer science journals and is the co-author of the book "Hydrological Analysis and Design: An approach based on Malaysian Rainfall and Runoff". David Hong is listed in Who's Who in the World.
Email address: jerlang.hong@taylors.edu.my



HONG KEE AN is presently a consultant in hydrology and water resources and he worked as a hydrologist, water resources planner and modeler for various water supply projects in Malaysia. He received his B. Sc. degree in Agricultural Engineering from National Taiwan University in 1971 and the M. Eng. degree in Water Science from Asian Institute of Technology in 1973. He is the author of several papers published in the journal and bulletins of the Institution of Engineers, Malaysia and Malaysian Water Work Association and the books "A Guide to Flood and Yield Estimation" and "Hydrological Analysis and Design: An approach based on Malaysian Rainfall and Runoff". He is a Fellow of the Institution of Engineers, Malaysia and is listed in Who's Who in Engineering published by the American Association of Engineering Societies.
Email address: keeahong@yahoo.co.uk