International Journal of the Physical Sciences Vol. 6(12), pp. 2875-2888, 18 June, 2011
Available online at http://www.academicjournals.org/IJPS
DOI: 10.5897/IJPS11.515
ISSN 1992 - 1950 ©2011 Academic Journals

Full Length Research Paper

Adaptive neuro-fuzzy inference system based model for rainfall forecasting in Klang River, Malaysia

A. El-Shafie*, O. Jaafer and A. Seyed

Civil and Structural Engineering Department, University Kebangsaan Malaysia, Malaysia.

Accepted 03 May, 2011

Runoff prediction still represents an extremely important issue in applied hydrology. On the other hand, rainfall is one of the most complicated effective hydrologic processes in runoff prediction. For a developing country such as Malaysia which is prone to flood disaster having such an expert model for runoff forecasting is a very vital matter. In this article, an adaptive neuro-fuzzy inference system (ANFIS) model is proposed to forecast the rainfall for Klang River in Malaysia on monthly basis. To be able to train and test the ANFIS and ANN models, the statistical data from 1997 to 2008, was obtained from Klang gates dam data. The optimum structure and optimum input pattern of model was determined through trial and error. Different combinations of rainfall were produced as inputs and five different criteria were used in order to evaluate the effectiveness of each network and its ability to make precise prediction. The performance of the ANFIS model is compared to artificial neural network (ANN) model. The five criteria are root mean square error (RMSE), Correlation Coefficient (R^2), and Nash Sutcliffe coefficient (NE), gamma coefficient (GC) Spearman correlation coefficient (SCC). The result indicate that the ANFIS model showed higher rainfall forecasting accuracy and low error compared to the ANN model. Furthermore, the rainfall estimated by this technique was closer to actual data than the other one.

Key word: Klang gate, ANFIS, forecasting model.

INTRODUCTION

Rainfall-runoff relationship is an essential component in the process of water resources evaluation and is considered as a central problem in hydrology. Recognition of relation between rainfall and runoff for a watershed is one of the most important issues which hydrologists and engineers have been busy with. Information about rainfall-runoff is needed for designation in hydrology engineering and management. This relation is complicated and non-linear. There are many parameters which effect runoff, such as: rainfall, soil primitive wetness, earth surface material, geomorphology watershed, evaporation, infiltration rainfall distribution, rainfall duration and so on.

Rainfall is one of the most complicated hydrologic processes. For many years, hydrologists have attempted to understand the transformation of precipitation to runoff, in order to forecast stream flow for purposes such as water supply, flood control, irrigation, drainage, water quality, power generation, recreation, and fish and wildlife propagation (Remesan and Shamim, 2009). Non-linearity characteristics of the traditional and historical information about the complexity of physical model have been the reasons, which led to more rational models like artificial intelligence tools.

In the last decade, the intelligence models such as neuro-fuzzy in rainfall-runoff modeling have been provided to simulate this relationship. A large number of parameters, including instability in watersheds, their particular characteristic and rainfall models make the issue more complicated. Hydrology has a large recorded rainfall - runoff prediction. With regard to the problems and weak points which exist in statistical and conceptual model, it seems necessary to have a model which can utilize input and output parameters. Over the years artificial intelligence tools has been added as a prediction tool.

*Corresponding author. Email: elshafie@vlsi.eng.ukm.my

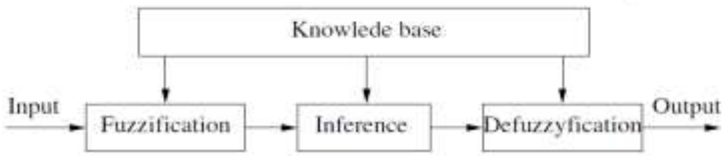

Figure 1. Fuzzy inference system.

Neural network (ANN) has been successfully applied in water resource. In hydrological forecasting contest, the artificial neural network is applied for rainfall-runoff modeling (Remesan and Shamim, 2009; Junsawang et al., 2007), light penetration estimation in reservoirs (Soyupak et al., 2007), streamflow prediction (Chang, 2006; Chen et al., 2006; Kisi, 2004a; Cigizoglu and Kisi, 2005); hydrological time series prediction (Zounemat-Kermani and Teshnehlab, 2008). Fuzzy based model could be used to model the process behaviors even with incomplete and imprecise or ambiguous information (Nourani et al., 2008).

Also, neuro-fuzzy has been used in rainfall-runoff modeling by (Nasr and Bruen, 2008; Remesan et al., 2009), hydrological time series prediction (Nayak, 2004; Firat, 2007, 2008) prediction of water level in reservoir (Chang, 2006); reconstructing missing flow data, suspended sediment (Kisi, 2005, 2009), evaporation (Moghadamnia et al., 2009), river flow estimation (Firat, 2007; El-Shafie, 2006). There are a few reports on employment of neuro fuzzy approaches in rainfall modelling. Aldrian et al. (2008) provided a multi variable adaptive neuro fuzzy inference system (ANFIS) in predicting daily rainfall using several surface weather parameters as predictors. The research showed that relative humidity is the best predictor with a stable performance slightly regardless of training data size and low error amount especially in comparison to those from other predictors. In Datorani et al. (2010) multi layer perceptron (MLP), generalized feed forward (GFF), modular neural network (MNN), principle component analysis (PGA), recurrent network (RN) and time lagged recurrent network (TLRN) and ANFIS were examined for annual rainfall prediction in Yazd meteorological station in central Iran. The model trained using monthly precipitation moving average; mean temperatures, relative humidity, mean wind speed, mean wind direction and evaporation. Comparison of the model's results indicated that the RN and TLRN and ANFIS had better performance in predicting the rainfall.

In recent years, ANNs have been used intensively for prediction and forecasting in a number of water-related areas, including water resource study (Najah et al., 2009; Ahmed et al., 2009; El-Shafie et al., 2007; El-Shafie et al., 2008; El-Shafie et al., 2009(a), El-Shafie et al., 2009(b); El-Shafie et al., 2010), prediction of evaporation (Sudheer et al., 2002), hydrograph simulator (Deka and Chandramouli, 2005; Lange, 1999), rainfall estimating (Hsu et al., 1999; Lin and Chen 2005; Luk et al., 2001).

The main purpose of this paper is to compare and analyze the performance of the adaptive inference system (ANFIS) and ANN to see their applicability in rainfall forecasting of Klang River basin in Malaysia.

NEURO-FUZZY APPROACH

The neuro-adaptive learning techniques provide a method for the fuzzy modeling procedure to gain information about a data set, in order to compute the membership function parameters that best allow the associated fuzzy inference system to track the given input/output data (Jang, 1993).

Each fuzzy system contains three main parts: fuzzification, inference, defuzzification. Basically a fuzzy inference system is composed of five functional blocks (Figure 1):

(i) Input characteristics to input membership functions.
(ii) Input membership function to rules.
(iii) Rules to a set of output characteristics.
(iv) Output characteristics to output membership functions.
(v) The output membership function to a single-valued output, or a decision associated with the output.

For simplicity, a fuzzy inference system has two inputs x and y and one output is assumed. For a first-order Sugeno fuzzy model, a common rule set with two fuzzy if–then rules is defined as (Equations 1 and 2):

Rule 1: If x is x_1 and y is y_1, then $f_1 = p_1 x + q_1 y + r_1$ (1)

Rule 2: If x is x_2 and y is y_2, then $f_2 = p_2 x + q_2 y + r_2$ (2)

$$f_2 = p_2 x + q_2 y + r_2$$

Figure 1 illustrates the reasoning mechanism for this Sugeno model, the corresponding equivalent ANFIS architecture is shown in Figures 2a and b, where the

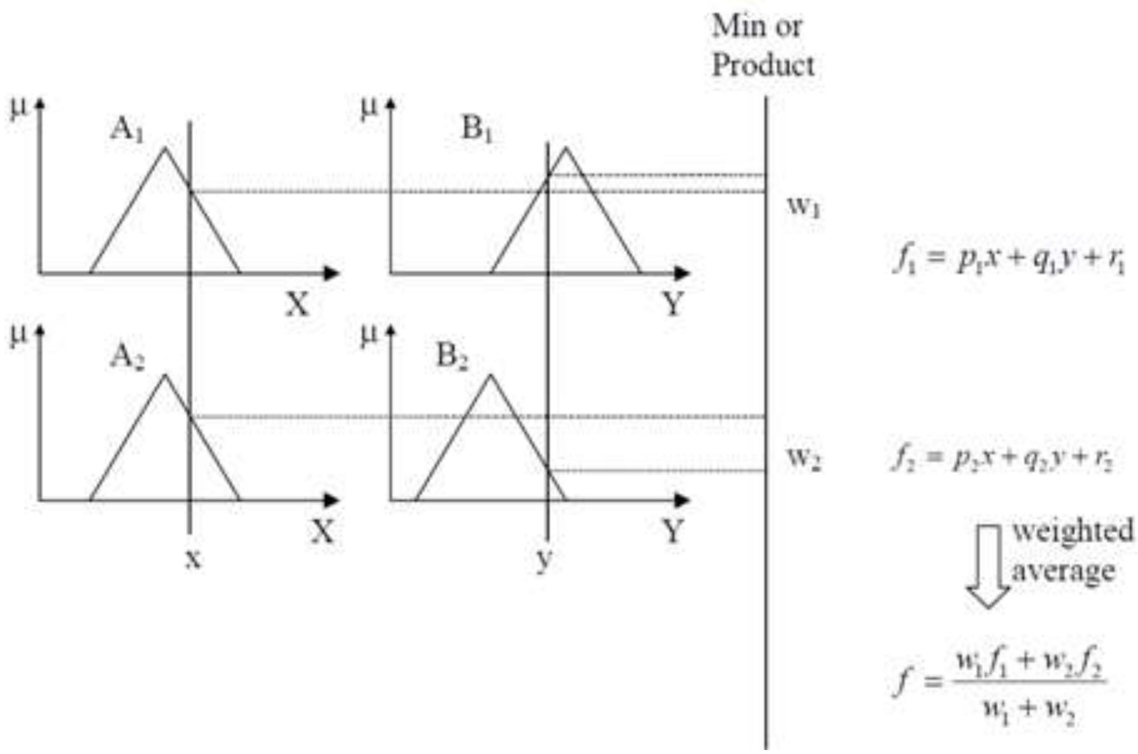

Figure 2a. Two-input first-order Sugeno fuzzy model with two rules.

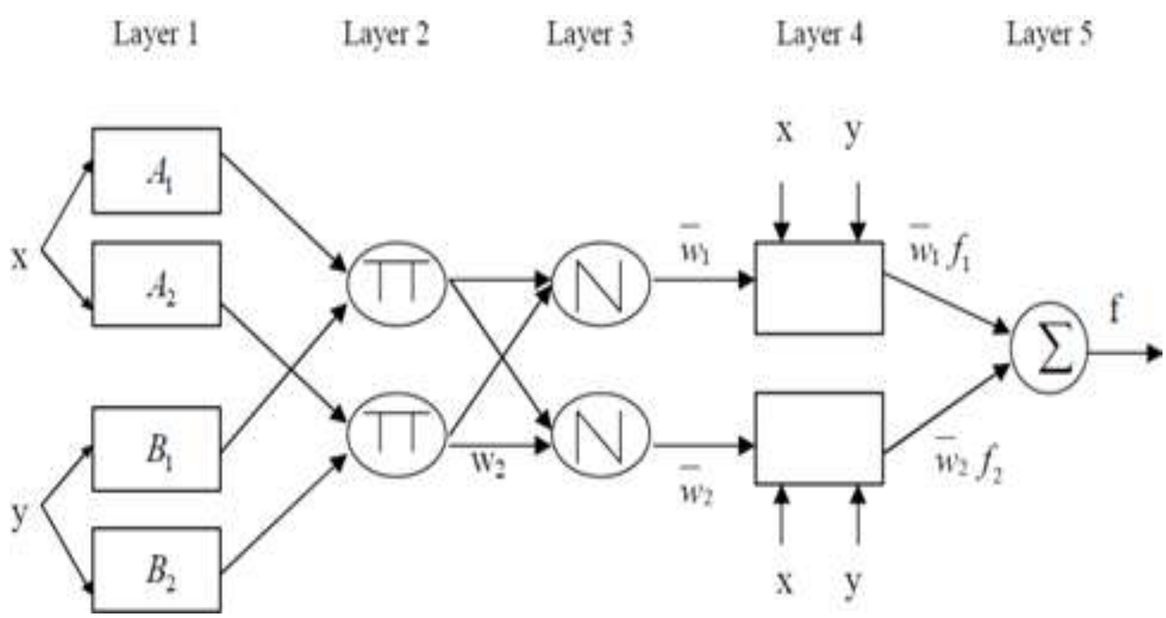

Figure 2b. Equivalent ANFIS structure.

nodes of the same layer have similar functions, as described:

Layer 1: Each node in this layer generates a membership grade of linguistic label. x_1, x_2, y_1 and y_2 are the linguistic labels, which used to define the membership functions. Every node in this layer is an adaptive node. Parameters in this layer are called premise parameters (Equations 3 and 4):

$Q_{l,i}$=μ_{A_i}(x) For i=1, 2, or (3)

$Q_{l,i}$=$\mu_{B_{i-2}}$(y) For i=3, 4 (4)

where x (or y) is the input to the i^{th} node and A_i(orB_{i-2}) is a linguistic label (such as "low" or "high") associated with this node. In words, $Q_{l,i}$ is the membership grade of fuzzy set A = (A_1, A_2, B_1 or B_2) and its specifies the degree to which the given input x (or y) satisfies the Quantifier A. The membership function for A and B are generally described by generalized bell function (Equation 5):

$$O_i^1 = \mu_{A_i}Q(t) = e^{\frac{-(Q(t)-c)^2}{2\sigma^2}} \quad (5)$$

where $\{a_i, b_i, c_i\}$ is the parameter set. As the values of these parameters change, the bell-shaped function varies accordingly, thus exhibiting various forms of membership functions on linguistic label A_i. Parameters in this layer are referred to as premise parameters. The outputs of this layer are membership values of the premise part.

Layer 2: Every node in this layer is a fixed node labelled Π, whose output is the product of all the incoming signals. Each node output represents the firing strength of a rule (Equation 6):

$Q_{2,i}$= w_i =μA_i(x) μB_i(y) i=1, 2 (6)

Each node output represents the firing strength of the rule.

Layer 3: Every node in this layer is a fixed node labelled N. The i^{th} node calculates the ratio of the i^{th} rule's firing strength (w_1' and w_2'). Thus the outputs of this layer are called normalized firing strengths (Equation 7):

$$O_i^3 = \bar{w}_i = \frac{w_i}{\sum_i w_i} \quad i = 1,2 \quad (7)$$

Layer 4: Every node in this layer is an adaptive node:

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i p_i Q(t) + q_i Q(t-1) + r_i \quad i = 1,2 \quad (8)$$

where w_i is the output of layer 3, and $\{p_i, q_i, r_i\}$ is the parameter set. Parameters of this layer are referred to as consequence or output parameters (Equation 8).

Layer 5: The single node in this layer is a fixed node labelled Σ, which computes the overall output as the summation of all incoming signals (Equation 9):

$$Q_i^5 = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (9)$$

ARTIFICIAL NEURAL NETWORK

The artificial neural network model (ANN) is proposed to uncover the non-linear relationship between rainfall and runoff.

Feed forward back propagation neural network (FFNN)

This class of networks consists of multiple layers of computational units, usually interconnected in a feed-forward way. Each neuron in one layer has directed connections to the neurons of the subsequent layer. In many applications the units of these networks apply a sigmoid function as basically the MLP consists of three layers: the input layer, where the data are introduced to the network; the hidden layer, where the data are processed (that can be one or more) and the output layer, where the results for given inputs are produced (Junsawang et al., 2007). Each layer is made up of several nodes, and layers are inter-connected by sets of correlation weights. Each input node in input layer broadcasts the activation function (Figure 3). Multi-layer networks use a variety of learning techniques, Back-propagation is the most commonly used supervised training algorithm in the multilayer feed-forward networks. The objective of a back propagation network is to find the weight that approximate target values of output with a selected accuracy. The network weights are modified by minimizing the error between a target and computed outputs. The error between the outputs of the network and the target outputs are computed at the end of each forward pass. If an error is higher than a selected value, the procedure continuous with a reverse pass, otherwise, training is stopped (Kisi, 2003).

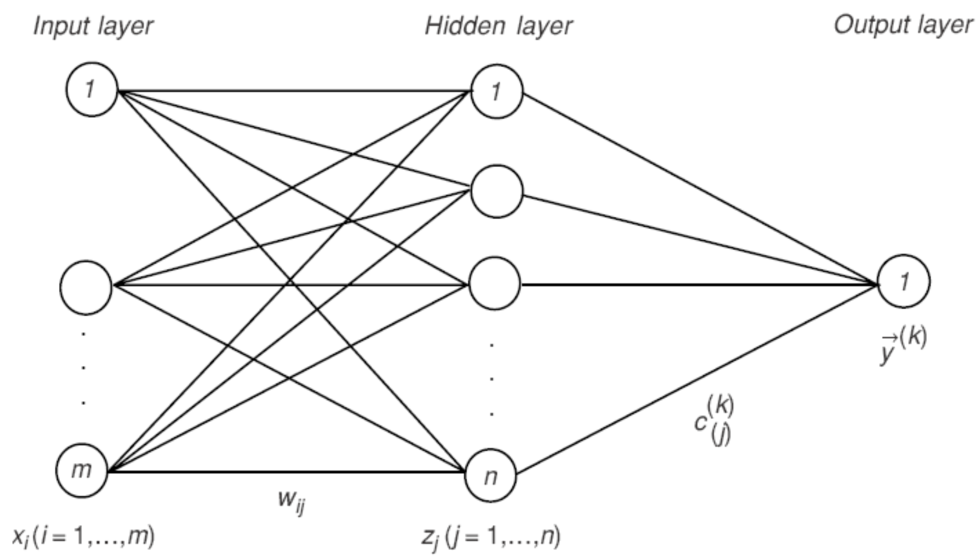

Figure 3. Three layer feed forward perceptron.

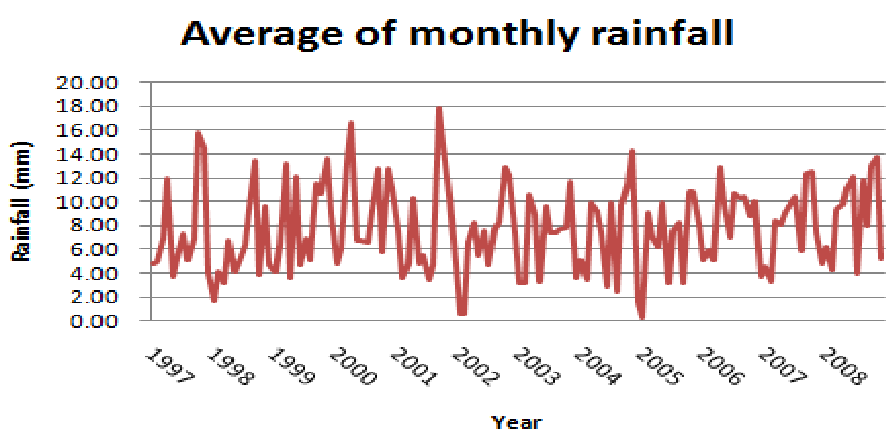

Figure 4. Rainfall time series.

With respect to the theory of artificial neural network, having statistical information based on data is useful in data management. With respect to artificial neural network, the data were classified into different groups and so their harmony with activation function and parameter design is very important. Having a pervasive outlook on data of artificial neural network is needed. Figure 3 shows the rainfall time series.

Case study

The Klang River basin is located on the west coast of Peninsular Malaysia and encompasses the Federal Territory of Kuala Lumpur, parts of Gombak, Hulu Langat, Klang, and Petaling districts in Selangore Stats, and the municipal areas of Ampang Jaya, Petaling Jaya, and Shah Alam. Klang is geographically located at latitude(3.233°) 3° 13' 58" North of the equator and longitude (101.75°) 101° 45' 0" East of the Prime Meridian on the Map of Kuala Lumpur. With an estimated population of over 3.6 million, (about 21% of the national population) and with a growth rate of almost 5% per year, the basin has experienced the highest economic growth in the country.

It is also characterized by uniform high temperature, high relative humidity, heavy rainfall and little wind. The average annual rainfall depth in the study area is about 2400 mm. The highest rainfall occurs in the month of April and November with a mean of 280 mm. The lowest rainfall occurs in the month of June with a mean of 115 mm. The maximum and minimum rainfall chart is shown in Figure 4 and 5 respectively.

The Klang River originates in the mountainous area

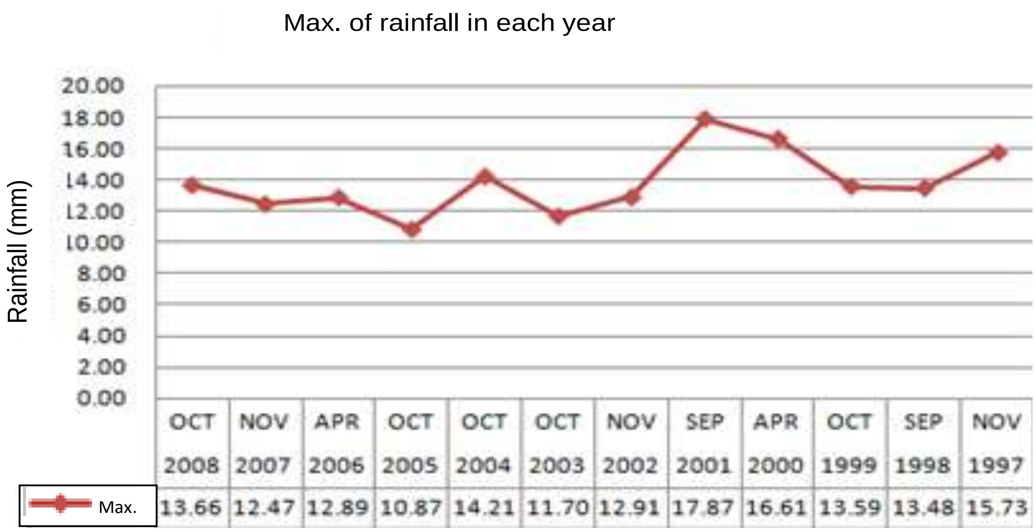

Figure 5. Maximum rainfall.

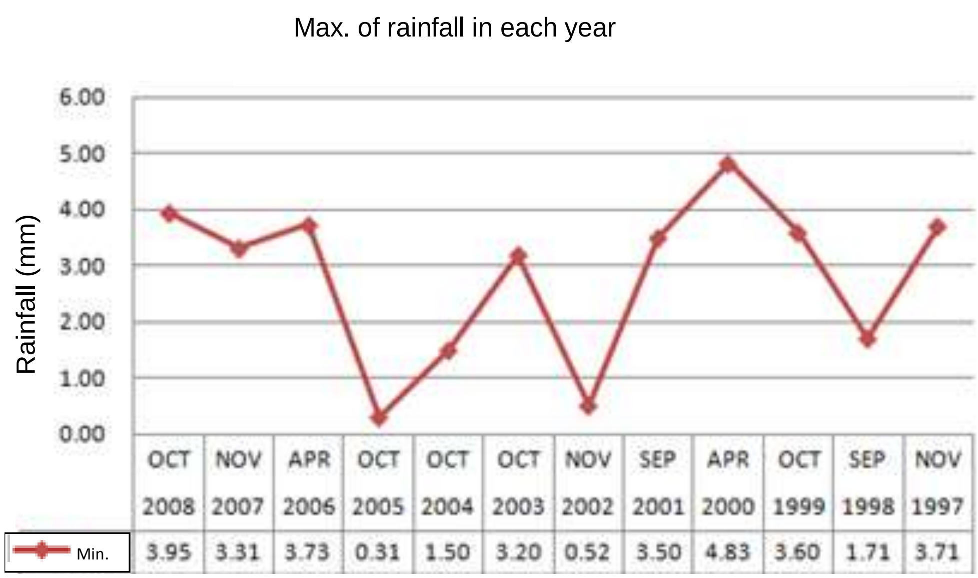

Figure 6. Minimum rainfall.

about 25 km northeast of Kuala Lumpur. It is joined by 11 major tributaries, while passing through the Federal Territory and the area downstream of Kuala Lumpur, before joining the Straits of Malacca at port Klang. The Klang river has a total lengh of about 120 km. The basin is 1290 km^2, about 35% of which has been developed for residental, commercial, industrial, and instituionsl use. The upper catchments of the Klang River and its tributaries – the Gombak and Batu Rivers- are covered with well maintained forests. However, the lower reaches of the basin, with extensive urban land development activities, are major contributors of sediment load and flood peaks.The study area location map has been shown in Figure 6.

RAINFALL FORECASTING:

Data collection

All information and data that are available about Klang River were based on Klang gates dam data. For this study, the data used were from year 1997 to 2008 .The available data for catchment (144 patterns) is divided into two groups: training set (calibration) and a testing set (validation), checking set (checking). The data for training and testing patterns are selected to 100 and 44 and 30 input patterns respectively. The rainfall data statistic for training and testing sets are given in Table 1, which contains the maximum, minimum, mean, standard deviation, variance, skew standard.

The skewness value can be positive or negative, or even undefined. Qualitatively, a negative skew indicates that the tail on the left side of probability density function is longer than the right

Table 1. The statistical parameter for data set.

Statistical parameter	Year											
	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997
Max	13.66	12.47	12.89	10.87	14.21	11.70	12.91	17.87	16.61	13.59	13.48	15.73
Min	3.95	3.31	3.73	0.31	1.50	3.20	0.52	3.50	4.83	3.60	1.71	3.71
Mean	9.03	8.04	8.31	7.03	7.26	7.04	6.84	7.89	9.41	8.40	6.01	7.59
Variance	11.87	8.99	8.11	10.83	16.78	9.11	14.19	21.17	14.58	13.94	11.19	17.22
St. deviation	3.45	3.00	2.85	3.29	4.10	3.02	3.77	4.60	3.82	3.73	3.35	4.15
Skew St.	3.45	3.00	2.85	3.29	4.10	3.02	3.77	4.60	3.82	3.73	3.35	4.15

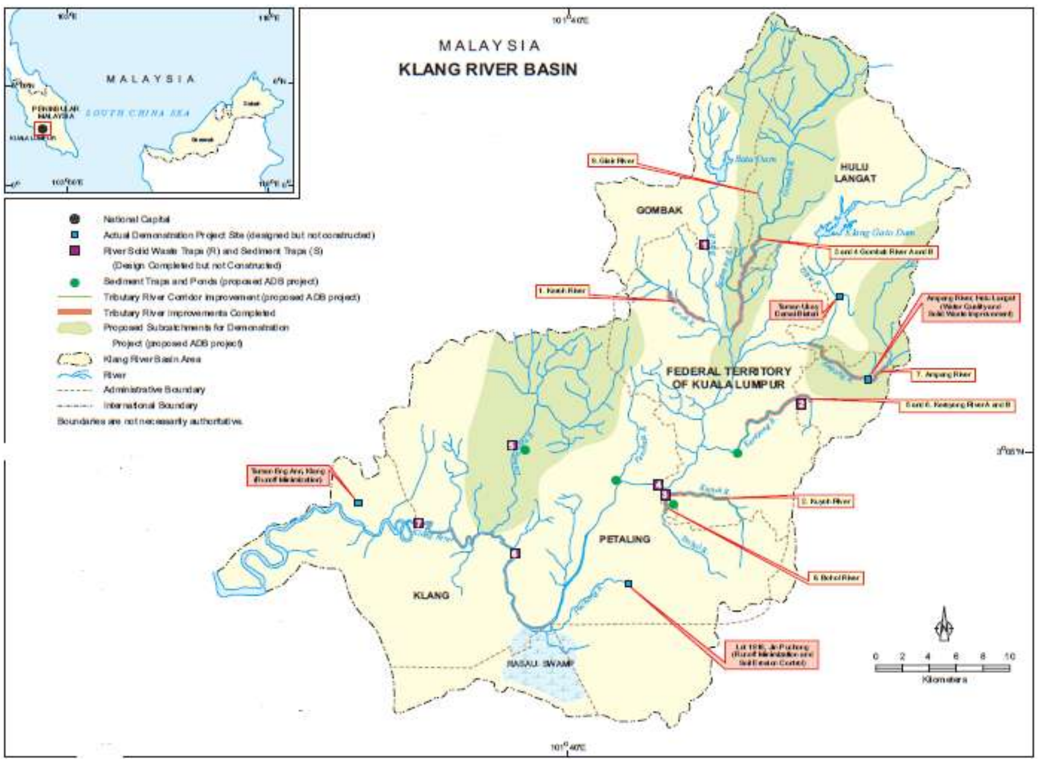

Figure 7. Local authorities within Klang River basin.

side and the bulk of the values (including the median) lie to the right of the mean. A positive skew indicates that the tail on the right side is longer than the left side and the bulk of the values lie to the left of the mean. A zero value indicates that the values are relatively evenly distributed on both sides of the mean, typically but not necessarily implying a symmetric distribution.

Reconstruction of model

Before making a structure of neural network, data pre-processing such as missing data, filling, ill data elimination and normalization are fulfilled.

Data normalization

One of the steps of data pre-processing is data normalization. The need to make harmony and balance between network data range and activation function used causes the data to be normal in activation function range. Sigmoid logarithm function is used for all layers. By considering Sigmoid, it can be seen that the range is between 0 and 1, so data must be normalized between 0, 1. (Equation 10) The following formula was used:

$$x_n = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{10}$$

Where x is actual data and $x_{\min}$ is minimum value of original series and $x_{\max}$ is maximum value of original series.

Activation function used in this research is Logsig function (Figure 7). It means that the input and target values of rainfall, is normalized into an appropriate scale by the maximum and minimum values so that all are combined in the unit interval [0, 1] (Equation 11):

$$y = \frac{1}{1+e^{-x}} \tag{11}$$

The procedure of neural network making is such that the network is made by a simple structure and then the performance of each model is evaluated using the root mean square error (RMSE), correlation coefficient (R^2), and Nash Sutcliffe coefficient (NE), gamma test (G) Spearman correlation coefficient (SCC).

The selection of training data is very important in the application of artificial neural network. The data duration should be large enough to represent the characteristics of the watershed and to accommodate the requirement of the ANN architecture. The objective of training is to minimize a global error that measures the difference between the model output and the actual values that are defined as:

$$E = \frac{1}{p}\sum_{p=1}^{p} E_p \tag{12}$$

where p total number of training is pattern and E_p is error for training pattern, p which is given by:

$$E_p = \frac{1}{2}\sum_{i=1}^{n}(Q_i - \hat{Q}_i)^2 \tag{13}$$

where n is total number of output nodes Q_i is network output at the i^{th} output node and $\hat{Q}_i$ is target output at the i^{th} output node.

Criteria of performance

Five different criteria are used in order to evaluate the effectiveness of each network and its ability to make precise predictions. The five criteria are root mean square error (RMSE), correlation coefficient (R^2), and Nash Sutcliffe coefficient (NE), gamma test (G) Spearman correlation coefficient (SCC). The performance is mathematically expressed as:

(1) Nash- Sutcliffe: $E_f = 1 - \frac{\Sigma(Q_D - Q_Y)^2}{\Sigma(Q_Y - \bar{Q})^2}$ (14)

(2) Spearman rank correlation: $\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$ (15)

(3) Correlation coefficient:

$$\mathrm{CORR} = \frac{\sum_{i=1}^{N}(Q_D - \overline{Q_D})(Q_Y - \overline{Q_Y})}{\sqrt{\sum_{i=1}^{N}(Q_D - \overline{Q_D})^2(Q_Y - \overline{Q_Y})^2}} \tag{16}$$

(4) Root mean square error:

$$\mathrm{RMSE} = \left[\sum_{i=1}^{N}\frac{(Q_D - Q_Y)^2}{N}\right]^{0.5} \tag{17}$$

(5) Gamma test: Basically, the gamma test is able to provide the best mean square error that can possibly be achieved using any nonlinear smooth models:

$$\mathrm{Gamma} = \frac{Ns - Nd}{Ns + Nd} \tag{18}$$

where N_s, the number of pairs of cases ranked in the same order on both variables; N_d, the number of pairs of cases ranked differently on the variables, di is the difference in the ranks given to the two variable values for each item of data.

RESULTS AND DISCUSSION

ANN and ANFIS models are introduced into rainfall science as a powerful, flexible, and statistical modelling technique to address complex pattern recognition problems. One of the most important tasks in developing a satisfactory ANN and ANFIS forecasting model is the selection of the input variables, but it is not clear how long the time-delay should be continued to achieve the proper structure. Therefore, the correlation coefficient between initial data should be considered first (Chen et al., 2006). In this study, different combinations of input data were explored to assess their influence on the rainfall estimation modeling. Table 2 demonstrates the linear relationship between input and target rainfall. It shows the persistence effect of current rainfall versus several previous water levels (t-1, t-2, and t- 3) at the river flow gauging station of the Klang River.

As we move the rainfall forward to 3 months, the pattern and peak of the rainfall seems well matched. Consequently, several previous average rainfalls of the watershed (R (t-4), R (t-5), and R (t-6)) are used as input but as shown in the Table 3, the correlation coefficient is varying and not stable. However, if we only have a limited number of data sets, the models built might only be good for the training sets due to overtraining and could not be used for future (Chen et al., 2006). events.

The available data for catchment (144 patterns) is divided into three groups: training set (calibration) and a testing set (validation), checking set (checking). The data for training and testing patterns are randomly selected to 100 and 44 and 30 input patterns respectively. Different combinations of the rainfall are used to construct the appropriate input structure. The general structures of the rainfall forecasting models are given in Equations 19, 20, 21 and 22 for the rainfall models, respectively.

Network Inputs with rainfall data in current and previous months (P (t), P (t-1) (19)

Table 2. Linear relationship between input and target rainfall.

Parameter	Testing data	Model					
		R+1	R+2	R+3	R+4	R+5	R+6
Correlation coefficient (%)	2004-2008	17.5	22.3	29.1	29.4	30.8	31.6
NASH (%)	2004-2008	23.16	28.37	36.29	36.57	-	-

Table 3. Model structures for estimation of rainfall.

Model	Input structure numbers of the variable	Output
1	P (t-1)	1
2	P (t-1), P (t-2)	1
3	P (t-1), P (t-2), P (t-3)	1
4	P (t-1), P (t-2), P (t-3), P (t-4)	1

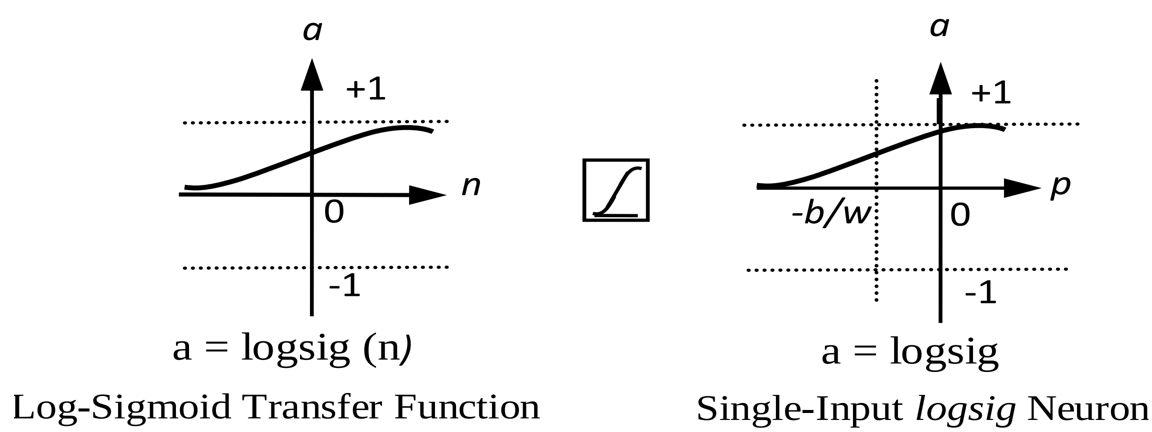

Figure 8. Log Sigmoid transfer function.

Network Inputs with Rainfall data in Current and previous months (P (t), P (t-1), P (t-2) (20)

Network Inputs with Rainfall data in Current and previous months (P (t), P (t-1), P (t-2), P(t-3) (21)

Network Inputs with Rainfall data in Current and previous months (P (t), P (t-1), P (t-2), P (t-3), P (t-4) (22)

The structure of models for Klang River is shown in Tables 3.

ANN model

In this study a model based on a feed forward neural network with a single hidden layer is used for monthly rainfall. The back propagation (BP) algorithm and sigmoid activation function are used to train the network. In the training and testing of ANN the same data set is used and performance of models are also evaluated and compared based on mentioned criteria. The qualities of the results produced by ANN are shown in Figures 8 and 9, as well as Table 2. According to the results, it can also be said that ANN is an efficient method to predict the rainfall. As seen from the fit line equation and R^2 value in scatter plot, the estimates of neural network model are closer to the exact fit line(y = x) line (Figure 10).

In brief, the models' predictions are optimum if Nash and RMSE are found to be close to 1 and 0 respectively. Referring to the relative performance of ANN methods, it can be observed that the highest and very close to unity Nash as well as the lowest RMSE between the observed and simulated results in all models is three previous rainfall values and it is selected as the best-fit model for describing the rainfall prediction in Klang River basin (Table 4). It also shows that the value of RSME is 0.074 and the highest value is 0.082. In addition, the values of the Nash-Sutcliff, correlation coefficient, and Spearman rank correlation and gamma coefficient in three previous rainfall values are higher than other models. Accordingly,

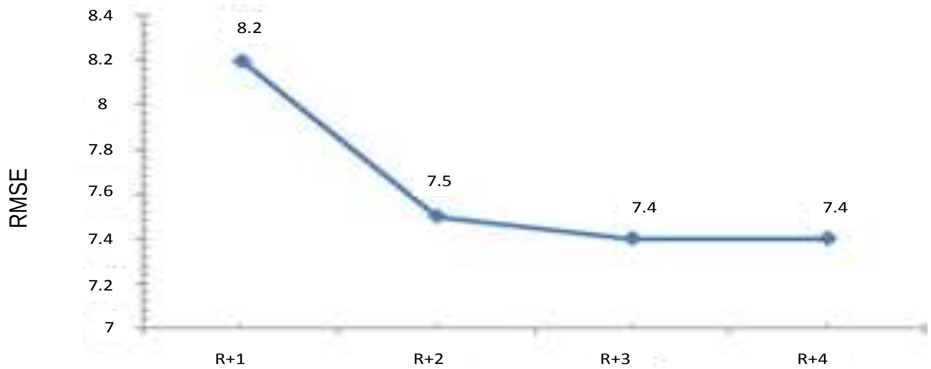

Figure 9. RMSE of ANN models-Testing set.

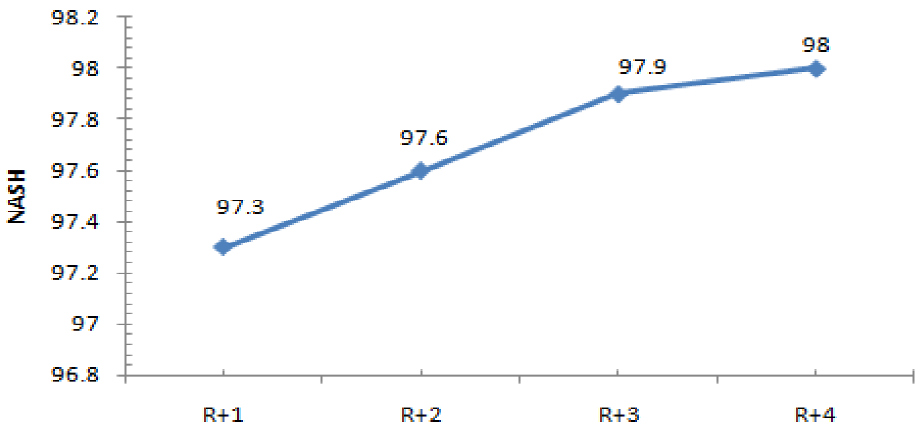

Figure 10. Performance of AN models-Testing set.

Table 4. Performance of ANN.

Parameter	Testing data	Models			
		1	2	3	4
RSME	2004-2008	0.082	0.075	0.074	0.74
NASH	2004-2008	0.973	0.976	0.979	0.980
R^2	2004-2008	0.735	0.768	0.785	0.788
Spearman	2004-2008	0.727	0.766	0.788	0.785
Gamma	2004-2008	0.556	0.604	0.624	0.626

it can be said that the ANN models whose input are the P (t-1), P (t-2), P (t-3) performs the best than the other models.

ANFIS model

The goal of ANFIS is to find a model, which will simulate correctly the inputs with the outputs (Keskin, 2006). In this paper, secondly, the various input variables are trained and tested by ANFIS method and the performances of models for rainfall forecasting models are compared and evaluated based on testing performances. In order to determine the nonlinear input and linear output parameters, the hybrid algorithm was used. The learning procedure and the construction of the rules were provided by this algorithm. The best fit model structure is determined according to criteria of performance

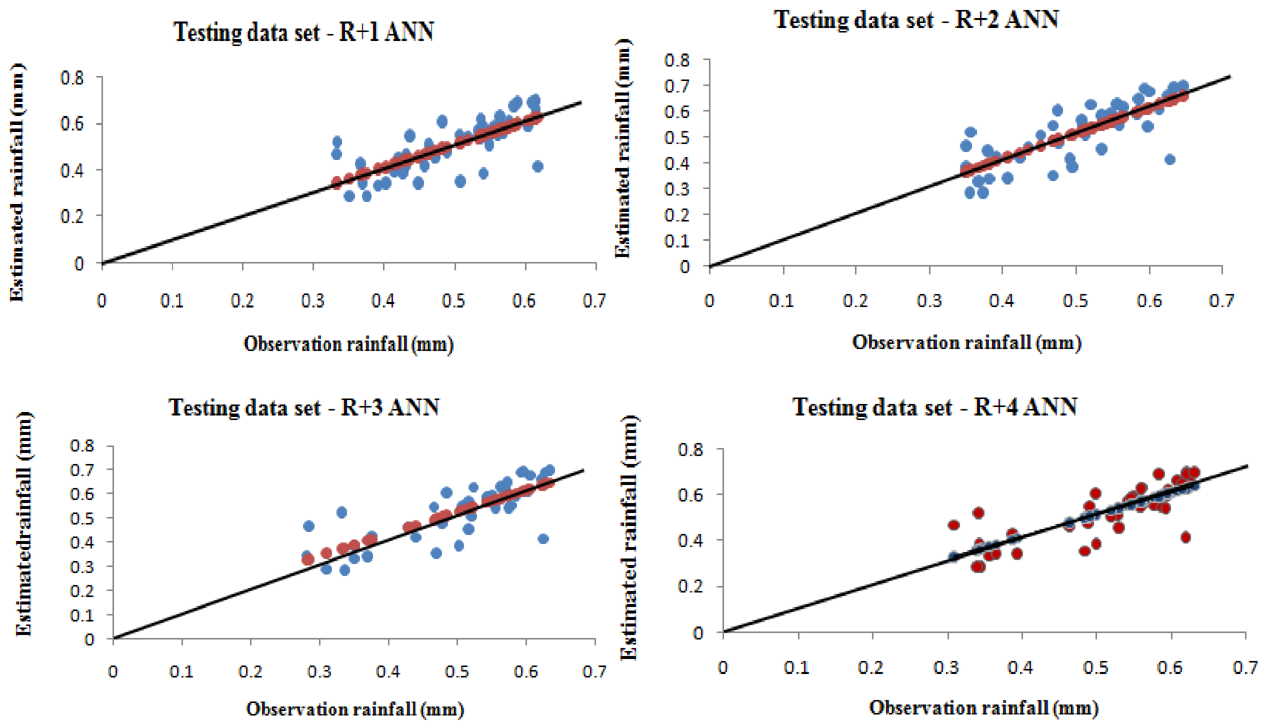

Figure 11. Scatter of observed and computed rainfall.

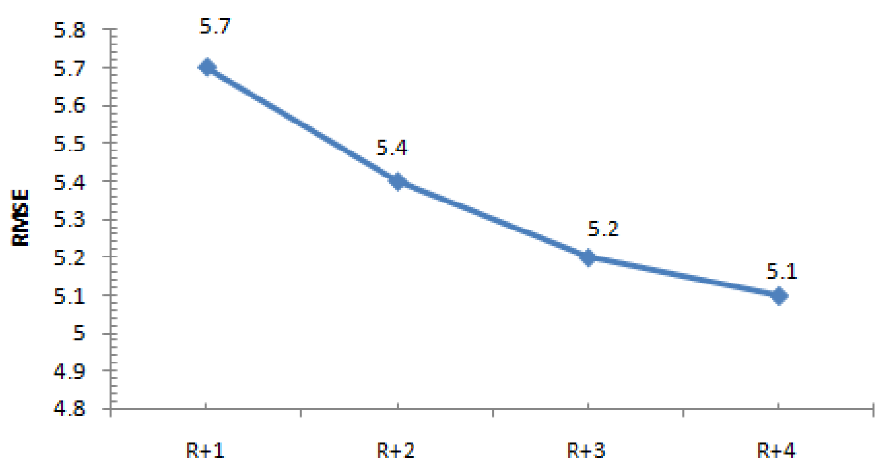

Figure 12. RMSE of ANFIS model- Testing set.

evaluation. The performances of the ANFIS models are shown in Figures 11 and 12 and Table 5.

Firstly, for Klang River, it shows that the lowest value of the RMSE and the highest values of the Nash is R+3 ANFIS model. R+3 ANFIS model, which consists of three antecedent rainfall in input, has shown the highest Nash, correlation and the minimum RMSE and was selected as the best-fit model for modelling of rainfall flow in the Klang River.

Figure 13 shows the scatter diagrams of the estimated values of the testing of R+3 ANFIS models and observed values. Figure 14 shows the scatter diagrams of the estimated values of the testing of R+3 ANFIS models and observed values shows the estimates of neural network model are closer to the exact fit line (y = x) line. Figures demonstrate that ANFIS could be applied successfully to establish accurate and reliable monthly total rainfall forecasting models. The structure parameter, which were

Table 5. Performance of ANFIS.

Parameter	Testing data	Models			
		1	2	3	4
RSME	2004-2008	0.057	0.054	0.052	0.051
NASH	2004-2008	0.985	0.991	0.990	0.991
R^2	2004-2008	0.876	0.886	0.90	0.906
Spearman	2004-2008	0.884	0.886	0.885	0.896
Gamma	2004-2008	0.715	0.707	0.729	0.722

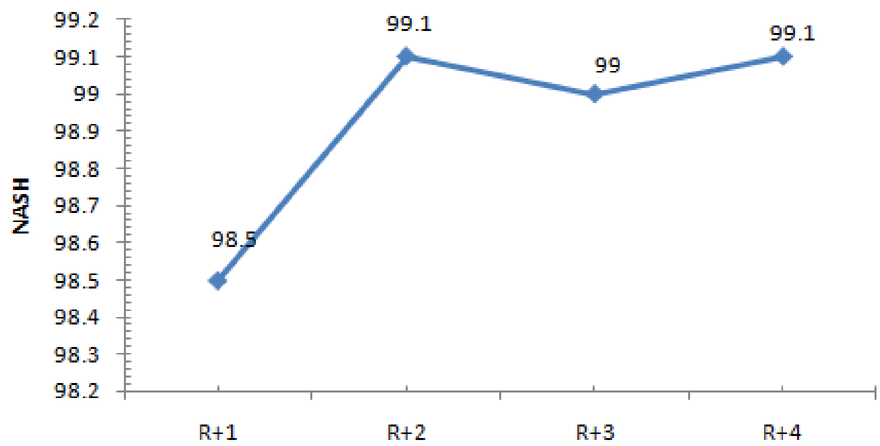

Figure 13. Performance of ANFIS model- Testing set.

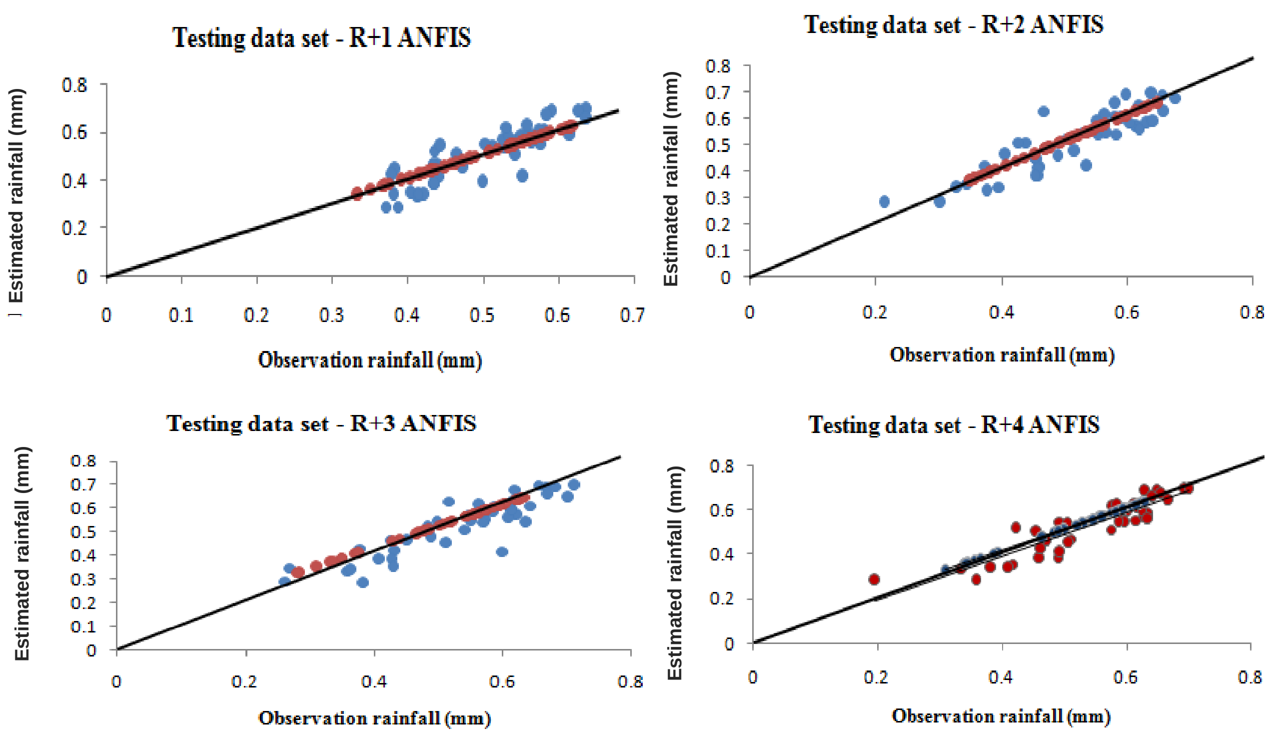

Figure14. Scatter of observed and computed rainfall using ANFIS.

Table 6. The training parameters of the ANFIS model.

Epoch	50
AND method	Prod
Imp. Method	Minimum
Aggr. Method	Maximum
Defuzzification method	Wtaver
Step size	0.21
Step size decrease rate	0.8
Step size increase rate	1.20

Table 7. The comparing of the performances of the ANFIS and ANN models.

Models	Testing data				
	RSME	**NASH**	R^2	**Spearman**	**Gamma**
ANN	0.074	0.979	0.785	0.788	0.624
ANFIS	0.052	0.990	0.90	0.885	0.729

selected by trial and error method during the training process is shown Table 6.

Conclusion

In this article, an adaptive neuro-fuzzy inference system (ANFIS) model and ANN is proposed to forecast the rainfall for Klang River in Malaysia on monthly basis. Different combinations of rainfall were produced as inputs and five different criteria were used in order to evaluate the effectiveness of each network and its ability to make precise predictions. R+3 ANFIS model and ANN models, which consists of three antecedent rainfall in input, has shown the highest Nash, correlation and the minimum RMSE and selected as the best-fit model for modelling of rainfall flow in the Klang river.

Table 7 summarizes the results obtained by the fuzzy model and the ANN models, during the calibration and the verification period. Comparing the performances of the ANFIS and ANN models, RMSE values of ANFIS model is lower than provided by ANN model. In addition, values of NASH, Spearman and CORR of ANFIS model are also higher than those achieved utilizing of ANN models. Thus, the performance of ANFIS method is better than ANN method according to criteria. The results demonstrate that ANFIS method is superior to the ANN method in forecasting monthly rainfall.

ACKNOWLEDGEMENTS

This research was supported by the research grant for the first author from University Kebangsaan Malaysia UKM-GUP-PLW-08-13-308 and (01-01-02-SF0581) grant from the Ministry of Science, Technology and Innovation (MOSTI) of Malaysia.

REFERENCES

Ahmed AN, Elshafie A, Karim O, Jaffar O (2009). Evaluation the Efficiency Of Radial Basis Function Neural Network For Prediction Of Water Quality Parameters. Eng., Intell. Syst., 17(4): 221-231.

Chen SH, Lin YH (2006). The strategy of building a flood forecast model by neuro-fuzzy network. Hydro.Proc., 20(7): 1525-1540.

Cigizoglu HK, Kisi O (2005). Flow prediction by three back propagation tech., using k-fold partitioning of neural network training data. Nordic Hydrol., 36(1): 49-64.

Deka P, Chandramouli V (2005). Fuzzy neural network model for hydrologic flow routing. J. Hydrol. Eng., 10(4): 302-314.

El-Shafie A , Najah AA , Karim O (2009)b. Application of neural network for scour and air entrainment prediction. Inter. Conference on Computer Techno. Dev., Art. no. 5360151: 273-277

El-Shafie A, Alaa EA, Noureldin A, Mohd RT (2009a). Enhancing Inflow Forecasting Model at Aswan High Dam Utilizing Radial Basis Neural Network and Upstream Monitoring Stations Measurements. Water Resour. Manage., 23(11): 2289-2315.

El-Shafie A, Noureldin A (2010). Generalized versus non-generalized neural network model for multi-lead inflow forecasting at Aswan High Dam. Hydro., Earth Syst. Sci. Discuss., 7(5): 7957-7993.

El-Shafie A, Noureldin AE, Taha MR, Basri H (2008). Neural network model for Nile River inflow forecasting based on correlation analysis of historical inflow data. J. Appl. Sci., 8(24): 4487-4499

El-Shafie A, Reda TM, Noureldin A (2007). A Neuro-Fuzzy Model for Inflow Forecasting of the Nile River at Aswan High Dam. Water Resources Manage., 21(3): 533-556.

Firat M, Güngör M (2007). River flow estimation using adaptive neuro fuzzy inference syst. Math. Comput. Simul., 75(3-4): 87-96.

Firat M, Güngör M (2008). Hydro., time-series modelling using an adaptive neuro-fuzzy inference syst. Hydrol. Proc., 22(13): 2122-2132.

Hsu KL, Gupta HV, Gao X, Sorooshian S (1999). Estimation of physical variables from multichannel remotely sensed imagery using a neural network: application to rainfall estimation. Water Resour. Res., 35(5):1605-1618.

Jang, SR (1993). ANFIS: Adaptive-network-based fuzzy inference system. IEEE Trans. Syst. Man Cybern., 23(3): 665-685.

Kisi O (2003). Daily River Flow Forecasting using Artificial Neural Networks and Auto Regression Model. Turkish J. Eng. Environ. Sci., 29: 9-20.

Kisi O (2004a). River flow modeling using artificial neural networks. J. Hydrol. Eng. ASCE, 9(1): 60-63.

Kisi O (2005). Suspended sediment estimation using neuro-fuzzy and neural network approaches/Estimation des matières en suspension par des approches neurofloues et à base de réseau de neurones. Hydro. Sci. J., 50(4): 683-696.

Kisi O, Haktanir T (2009). Adaptive neuro-fuzzy computing tech., for suspended sediment estimation. Adv. Eng. Softw., 40(6): 438-444.

Lange NT (1999). New mathematical approaches in hydro. modeling an application of artificial neural networks. Phys. Chem. Earth, (B); 24(1–2):31–5.

Lin GF, Chen LH (2005). Application of an artificial neural network to typhoon rainfall forecasting. Hydro. Proc., 19:1825-37.

Luk KC, Ball JE, Sharma A (2001). An application of artificial Neural networks for rainfall forecasting. Math. Comput. Model., 33: 683-693.

Nasr A, Bruen M (2008). Development of neuro-fuzzy models to account for temporal and spatial variations in a lumped rainfall-runoff model. J. Hydro., 349(3-4): 277-290.

Nayak P C, Sudheer K P (2004). A neuro-fuzzy computing technique for modeling hydro. time series. J. Hydro., 291(1-2): 52-66.

Nourani V, Mogaddam AA, Nadiri AO (2008a). An ANN-based model for spatiotempral groundwater level forecasting. Hydro. Proc., 22: 5054-5066.

Remesan R, Shamim MA (2009). Runoff prediction using an integrated hybrid modelling scheme. J. Hydro., 372(1-4): 48-60.

Soyupak S, Karaer F (2007). Adaptive neuro-fuzzy inference technique for estimation of light penetration in reservoirs. Limnology, 8(2): 103-112.

Sudheer KP, Gosain AK, Rangan DM, Saheb SM (2002). Modelling evaporation using an artificial neural network algorithm, Hydro. Process. lfi, pp. 3189-3202.

Zhang G, Patuwo E, Hu MY (1998). Forecasting With Artificial Neural Networks", the State of the Art. Int. J. Forecast., 14: 35-62.

Zounemat-Kermani M, Teshnehlab M (2008). Using adaptive neuro-fuzzy inference system for hydro., time series prediction. Appl. Soft. Comput., 8(2): 928-936.