

## Applicability of Adaptive Neuro-Fuzzy Inference Systems in Daily Reservoir Inflow Forecasting

<sup>1</sup>S.H. Karimi-Googhari and <sup>2</sup>T.S. Lee

<sup>1</sup>Department of Water Engineering, Shahid-Bahonar University of Kerman, Kerman, Iran

<sup>2</sup>Department of Biological and Agricultural Engineering, Universiti Putra Malaysia,  
Serdang, Selangor, Malaysia

---

**Abstract:** Accurate prediction of reservoir inflows is crucial for optimizing the operations of managing water resources. With emerging new data driven modeling approaches, methods based on neuro-fuzzy are becoming established in academic and practical applications. This study investigates the ability of Adaptive Neuro-Fuzzy Inference System (ANFIS) method to improve the accuracy of daily reservoir inflow forecasting. The subtractive clustering method is used to find the best number of fuzzy rules. A comparison is made between the ANFIS model and the Artificial Neural Network (ANN) model. A wide range of statistics measures are used to evaluate the performance of the models. Based on comparisons, it was revealed that the ANFIS technique could not improve the accuracy of estimations in a small tropical catchment and the ANN performed better, especially in capturing peak inflows.

**Key words:** Reservoir inflow, fuzzy inference system, forecasting, peak inflows, clustering method, floods

---

### INTRODUCTION

The ability to predict the uncontrollable hydrological events such as dam reservoir inflows prior to making decisions on water release helps in better managing of water resources. It is necessary to estimate how much water enters into the dam reservoir in a specific time in order to provide the specific water release volume. The main aim of the short term prediction of reservoir inflow is to reduce the risk of floods through mitigating measures such as appropriate and timely release of flood waters. The non-linear relationship between input and output variables complicates effort to forecast reservoir inflow events. Various types of forecasting techniques have been developed and implemented in the past which basically there are two common types of solution that are either knowledge driven or data driven. In order to overcome difficulties of data availability, required for the knowledge driven technique, data driven techniques are more popular choice. Data-driven modeling approach is emerging as a relative new methodology and is based on identification of the relationships between input and output variables of a hydrological system using the historical data. In this context, different Artificial Intelligence (AI) techniques have become popular in modeling water resources events (Openshaw and Openshaw, 2006). Soft Computing (SC), an

innovative approach to constructing computationally intelligent systems has recently come into the limelight. It is realize that complex real world problems require intelligent systems that combine knowledge, techniques and methodologies from various sources (Jang *et al.*, 1997).

Artificial Neural Networks (ANN) and Fuzzy Inference Systems (FIS) are the most popular and successful approaches resulting from advances in a branch of nonlinear system theoretic modeling. ANN models have been used successfully to model complex nonlinear input-output time series relationships in a wide variety of water resources fields. FIS have been used in controller systems and model identification also. The synergism allows soft computing to incorporate human knowledge effectively to deal with imprecision and uncertainty and learn to adapt to unknown or changing environment for better performance (Jang *et al.*, 1997). Fuzzy systems and neural networks share their ability to improve the intelligence of systems working in uncertain, imprecise and noisy environments. Neuro-fuzzy networks combine the explicit knowledge representation of fuzzy logic with the learning power of neural networks (Babuska and Verbruggen, 2003). For this reason, recently interest in using the Adaptive Neuro-Fuzzy Inference System (ANFIS) is increasingly raised. ANFIS is a typical Neuro-fuzzy (Jang, 1993) which uses the Takagi-Sugeno

fuzzy inference engine model with a six layers feed-forward network. The ANFIS has been shown to be powerful in modeling numerous industrial processes such as motor fault detection and diagnosis (Altug *et al.*, 1999), power systems dynamic load (Oonsivilai and El-Hawary, 1999) and wind speed modeling (Sfetsos, 2000). Neuro-fuzzy approach has been employed successfully in some fields of water resources engineering such as rainfall runoff modeling (Vernieuwe *et al.*, 2005; Chang and Chen, 2001) and reservoir operation (Chang *et al.*, 2005; Chang and Chang, 2001; Ponnambalam *et al.*, 2003). ANFIS has been evaluated to improve the accuracy of daily evaporation estimation (Kisi, 2006), reservoir water level prediction (Chang and Chang, 2006) and river flow forecasting (Firat and Gungora, 2007).

The main objective of this study is to investigate the ability of adaptive neuro-fuzzy inference system to forecast daily reservoir inflow. The modeling procedure is demonstrated for an ungauged dam reservoir in Malaysia. Reservoir catchments in Malaysia are usually ungauged forest areas where collecting different types of data which are essential when knowledge driven models (Physical or Conceptual models) are used is very costly and impossible in most cases.

### ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM (ANFIS)

Adaptive Neuro-Fuzzy Inference System (ANFIS), first was proposed by Jang (1993) can achieve a highly nonlinear mapping and it is superior to common linear methods in producing nonlinear time series (Jang *et al.*, 1997). Throughout this research, it was considered the ANFIS architecture for the first order Sugeno fuzzy model (Sugeno and Kang, 1988). The ANFIS is a multilayer feed forward network which uses neural network learning algorithms and fuzzy reasoning to map an input space to an output space (Chang and Chang, 2006). Assuming the fuzzy inference system under consideration has two inputs,  $x$  and  $y$  and one output,  $f$  for a first-order Sugeno fuzzy model, a common rule set with two fuzzy if then rules can be expressed as:

$$\begin{aligned} \text{Rule 1: If } x \text{ is } A_1 \text{ and } y \text{ is } B_1 \text{ then} \\ f_1 = p_1 x + q_1 y + r_1 \end{aligned} \quad (1)$$

$$\begin{aligned} \text{Rule 2: If } x \text{ is } A_2 \text{ and } y \text{ is } B_2 \text{ then} \\ f_2 = p_2 x + q_2 y + r_2 \end{aligned} \quad (2)$$

where,  $A_1, A_2$  and  $B_1, B_2$  are the membership functions (mfs) for inputs  $x$  and  $y$ , respectively;  $p_i, q_i$  and  $r_i$  ( $i = 1$  or  $2$ ) are linear parameters in the consequent part of the first-order Sugeno fuzzy model. Figure 1a shows

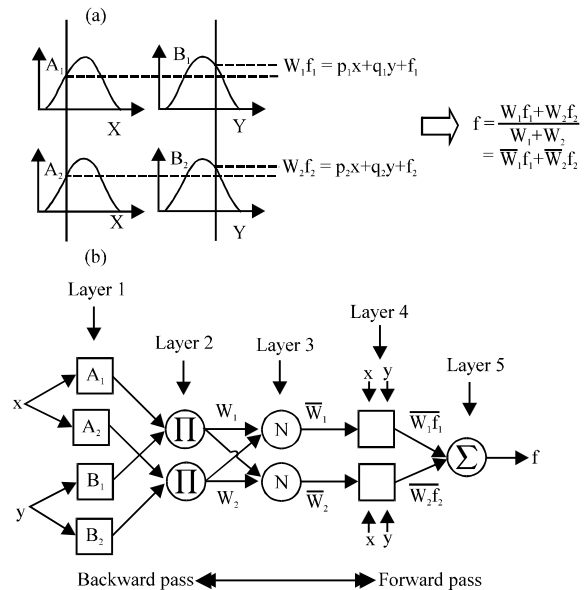


Fig. 1: (a) A two input first order Sugeno fuzzy model with two rules; (b) Equivalent ANFIS structure (Jang *et al.*, 1997)

the fuzzy reasoning mechanism for this Sugeno model to derive an output function ( $f$ ) from inputs  $x$  and  $y$ . The corresponding equivalent ANFIS architecture is shown in Fig. 1b where nodes of the same layer have similar functions. ANFIS consists of five layers.

**Layer 1:** Every node  $i$  in this layer is an adaptive node with a node function:

$$\begin{aligned} O_{1,i} = \mu_{A_i}(x), \quad \text{for } i = 1, 2 \text{ or} \\ O_{1,i} = \mu_{B_{i-2}}(y), \quad \text{for } i = 3, 4 \end{aligned} \quad (3)$$

where,  $x$  (or  $y$ ) is the input to node  $i$  and  $A_i$  (or  $B_i$ ) is the linguistic label (such as small, large, etc.) characterized by appropriate membership functions  $\mu_{A_i}, \mu_{B_{i-2}}$ , respectively.  $O_{1,i}$  is the membership grade of a fuzzy set  $A$  ( $= A_1, A_2, B_1, B_2$ ) and it specifies the degree to which the given input  $x$  (or  $y$ ) satisfies the quantifier  $A$ . Here the Gaussian membership function was used. Gaussian membership function is determined by two parameters  $\delta_i$  and  $\sigma_i$ :

$$\begin{aligned} \mu_{A_i} = e^{-\frac{(x-\delta_i)^2}{2\sigma_i^2}} \\ \mu_{B_{i-2}} = e^{-\frac{(y-\delta_i)^2}{2\sigma_i^2}} \end{aligned} \quad (4)$$

where,  $\{\delta_i, \sigma_i\}$  is the parameter set of the membership functions in the premise part of the fuzzy if then rules that

changes the shapes of the membership function. Parameters in this layer are referred to as the premise parameters.

**Layer 2:** Every node in this layer is a fixed node whose output is the product of all the incoming signals using AND operator as  $O_{2,k}$  which represents the firing strength:

$$O_{2,k} = w_k = \mu_{A_i}(x) \times \mu_{B_j}(y) \quad i = 1, 2; j = 1, 2; k = 1, \dots, 4 \quad (5)$$

**Layer 3:** Every node in this layer is a fixed node. It is to calculate the ratio of each  $i$ th rule's firing strength to the sum of all rules' firing strength. For convenience, outputs of this layer are called normalized firing strengths:

$$O_{3,i} = \bar{w}_i = \frac{w_i}{\sum_{k=1}^4 w_k}, i = 1, \dots, 4 \quad (6)$$

**Layer 4:** Every node  $i$  in this layer is an adaptive node with a node function:

$$O_{4,i} = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i + r_i), i = 1, \dots, 4 \quad (7)$$

where,  $\bar{w}_i$  is the normalized firing strength from layer 3 and  $\{p_i, q_i, r_i\}$  are the coefficients of this node and as parameters in this layer are referred to as consequent parameters.

**Layer 5:** The single node in this layer is affixed node which calculates overall output by summing all the incoming signals. Accordingly, the defuzzification process transforms each rule's fuzzy results into a crisp output in this layer:

$$O_{5,1} = \sum_{i=1}^4 \bar{w}_i f_i = \frac{\sum_{i=1}^4 w_i f_i}{\sum_{i=1}^4 w_i} \quad (8)$$

Premise parameters  $\{\delta_i, \sigma_i\}$  which identify the shape of the mfs and the consequent parameters  $\{p_i, q_i, r_i\}$  which determine the overall output of the system have to be optimized. It is possible to identify successfully parameters in an adaptive network using the back-propagation algorithm which is based on the gradient descent rule however, it is generally slow and likely to become trapped in local minima (Jang, 1993). The hybrid learning algorithm (Jang, 1993) is a faster learning algorithm which combines the gradient descent method and the Least Squares Estimate (LSE) to identify parameters.

From the ANFIS architecture shows in Fig. 1, it is observed that given the values of the premise parameters,

the overall output can be expressed as linear combinations of consequent parameters. More precisely, the output  $f$  can be rewritten as:

$$f = \bar{w}_1 f_1 + \bar{w}_2 f_2 = (\bar{w}_1 x) p_1 + (\bar{w}_1 y) q_1 + (\bar{w}_1) r_1 + (\bar{w}_2 x) p_2 + (\bar{w}_2 y) q_2 + (\bar{w}_2) r_2 \quad (9)$$

This is linear in the consequent parameters ( $p_i, q_i, r_i, p_2, q_2$  and  $r_2$ ). As a result, the total number of parameters in an ANFIS can be divided into a set of premise parameters and a set of consequent parameters. Consequently the hybrid-learning algorithm which combines the back propagation gradient descent and least squares method can be used as an effective search for the optimal parameters of the ANFIS. More specifically in the forward pass of the hybrid learning algorithm, the node output goes forward until layer 4 and the consequent parameters are identified by the least squares method. In the backward pass, the error signal propagates backwards and the premise parameters are updated by gradient descent. As mentioned earlier, the consequent parameters thus identified are optimal under the condition that the premise parameters are fixed. Accordingly, the hybrid approach converges much faster since, it reduces the dimension of the search space of the original back propagation method. More details of hybrid learning algorithm and its mathematical equations can be found in Jang *et al.* (1997) and Jang (1993).

## MODEL DEVELOPMENT

**Study area and data:** The Sembrong dam flood mitigation program was considered for this study and it is located in the West of Malaysia. The climate at the reservoir's catchment is equatorial through the year with relatively uniform temperatures, high humidity and heavy rainfalls. The catchment has no single principal stream but a series of short water courses converging into the reservoir. The available data were collected and reservoir inflows were evaluated using the water balance equation for the reservoir as below:

$$I_t = \pm \Delta S_t + (O_t + Se_t + E_t + Ws_t - P_t) \quad (10)$$

Where:

- $\pm \Delta S_t$  = The storage change
- $I_t$  = The inflow
- $P_t$  = The precipitation in dam lake
- $O_t$  = The release water from dam gates and penstocks
- $Se_t$  = The dam seepage
- $E_t$  = The evaporation
- $Ws_t$  = The water supply withdrawal at time  $t$

Table 1: Statistics of raw and transformed reservoir inflow data

Data	Types of sets	No. of data	Mean m <sup>3</sup> day <sup>-1</sup>	Standard deviation m <sup>3</sup> day <sup>-1</sup>	Coefficient of skewness
Raw data	Training and validation set	2188	319389.00	592623.0	6.20
	Testing set	304	160251.00	326862.0	3.87
Transformed data	Training and validation set	2188	11.74	1.5	0.65
	Testing set	304	11.07	1.2	0.74

Average rainfall across the catchment was calculated by constructing the Thiessen polygon. The 6 years daily data, 1995-1998 and 2002-2004 were considered as the training validation set and the other 10 months daily data, March 2005 to end of December 2005 as the testing set.

**Data pre-processing and model input selection:** In most conventional statistical models, the data should be normally distributed before the model parameters can be estimated efficiently. In practice, for most applications it is however advantageous first to transform the data into some new representation before model optimization. Similarly, the outputs of the network are often post-processed to give the required output values. Therefore, the models have been developed with transformed data to normal distributions. In this study, the reservoir inflow and rainfall data were tested for different probability distributions by using Probability-Probability (P-P) curves. The best fitted distribution was the Log-Normal distribution for reservoir inflow data. A summary of the data result is shown in Table 1. The statistics of the transformed training-validation and testing data sets are similar and the coefficient of skewness has been reduced significantly. The average rainfall values followed the Gamma distribution and the root square transform was considered.

Familiarity with the hydrologic system under consideration plays an important role in the successful implementation of ANFIS. This helps in avoiding missing some useful model inputs or preventing the complex model being assigned insignificant inputs. In the daily reservoir inflow modeling, values of inputs include average daily rainfall across the catchment and reservoir inflow in suitable lags and the value of output is the reservoir inflow during the immediate first following period, Q (t+1). However, the number of antecedent values to be included in the input vector was determined using the Auto-Correlation Function (ACF), Partial Auto-Correlation Function (PACF) and Cross-Correlation Function (CCF) as Q (t), Q (t-1), Q (t-2), Q (t-3), Q (t-4), R (t), R (t-1), R (t-2), R (t-3) where Q (t) and R (t) are daily reservoir inflow and average rainfall data at time step t.

**Construction of fuzzy rule base using subtractive fuzzy clustering:** In the ANFIS model, each input variable might be clustered into several class values in layer 1 to build up fuzzy rules and each fuzzy rule would be constructed through several parameters of membership function in

layer 2. The number of membership functions (mfs) and rules are mutually related; more membership functions implies more rules and determine the level of details of the model. For a successful network training and estimation of parameters, good initialization is important. Fuzzy clustering, subtractive clustering method was used in this study to identify the antecedent mfs. Subtractive clustering was introduced by Chiu (1994). For this method which is described in (Vernieuwe *et al.*, 2005), data points have to be rescaled to [0, 1] in each dimension. Each data point z<sub>j</sub> = (x<sub>j</sub>, y<sub>j</sub>) is assigned a potential P<sub>j</sub>, according to its location to all other data points:

$$P_j = \sum_{k=1}^N e^{-4/r_a^2 \|z_j - z_k\|^2} \quad (11)$$

Where:

- N = The number of data points
- ||| = The Euclidean distance
- r<sub>a</sub> = A positive constant called cluster radius

The potential of a data point to be a cluster center is higher when more data points are closer. The data point with the highest potential, denoted by P<sub>1</sub>\* is considered as the first cluster center c<sub>1</sub> = (d<sub>1</sub>, e<sub>1</sub>). The potential is then recalculated for all other points excluding the influence of the first cluster center according to;

$$P_j^{new} = P_j^{old} - P_1^* e^{-4/r_a^2 \|z_j - c_1\|^2} \quad (12)$$

with r<sub>b</sub> = η × r<sub>a</sub> the radius defining the neighborhood that will have measurable reductions in potential and η a positive constant, called the quash factor. Generally η = 1.5 as suggested by Chiu (1994). Again, the data point with the highest potential is considered to be the next cluster center c<sub>2</sub> if;

$$P_2^* > \xi P_1^* \quad (13)$$

Where, ξ is the accept ratio. If this is not the case but the following condition holds:

$$\frac{d_{min}}{r_a} + \frac{P_2^*}{P_1^*} \geq 1 \quad (14)$$

where, d<sub>min</sub> is the minimal distance between c<sub>2</sub> and all previously found cluster centers, the data point is still accepted as the next cluster center c<sub>2</sub>. Further iterations

can then be performed to obtain new cluster centers  $c_i$ . If a possible cluster center does not fulfill the above described conditions, it is rejected as a cluster center and its potential is set to 0. The data point with the next highest potential  $P_k^*$  is selected as the new possible cluster center and re-tested. The clustering ends if the following condition is fulfilled:

$$P_k^* - \langle \varepsilon P_1^* \rangle \quad (15)$$

where,  $\varepsilon$  is the reject ratio. Indicative parameters values for  $r_{\infty}$ ,  $\eta$ ,  $\varepsilon$  and  $\underline{\varepsilon}$  have been suggested by Chiu (1994). Each cluster center is considered as a fuzzy rule that describes the system behavior (Chiu, 1994). The degree to which a rule  $i$  is fulfilled is defined in terms of the distance to the defined cluster centers:

$$w_i(x) = e^{-(4/r_i^2)\|x-d_i\|} \quad (16)$$

In this study, first order Takagi-Sugeno (TS) fuzzy inference technique (Sugeno and Kang, 1988) with a rule base of the following type is developed:

$$R_i: \text{If } (Q(t), Q(t-1), Q(t-2), Q(t-3), \\ Q(t-4), R(t), R(t-1), R(t-2), R(t-3))$$

Belong to cluster  $C_i$  then:

$$Q_i(t+1) = p_i Q(t) + q_i Q(t-1) + r_i Q(t-2) + s_i Q(t-3) + \\ k_i Q(t-4) + l_i R(t) + g_i R(t-1) + h_i R(t-2) + j_i R(t-3) + d_i \quad (17)$$

In the second part of each fuzzy rule ( $R_i$ ), there are 10 parameters ( $p_i, q_i, r_i, s_i, k_i, l_i, g_i, h_i, j_i, d_i$ ) that have to be determined in the training process besides premise parameters which belong to membership functions. The Gaussian membership function was considered for modeling as it is more popular and simple. The sensitivity of selected models for using another practical membership functions was evaluated.

The fuzzy rule-base for the TS model based on Subtractive Clustering (SC) was determined using the Fuzzy Logic Toolbox of MATLAB. In this case, command `anfisedit` was used to run the Graphical User Interface (GUI) of ANFIS in MATLAB. In order to find the optimal model, the parameters of the SC algorithm were 1, 1.25 and 1.5 for squash factor ( $\eta$ ) and 0.2-0.8 for range of influence ( $r_a$ ) with steps of 0.05. The value of accept ratio and reject ratio were fixed based on MATLAB defaults (0.5 and 0.15, respectively). For each combination of these parameters, a model was built and trained. Through optimization of the performance RMSE obtained on the identification data, the optimal parameter combination was sought.

**Neural networks model:** Artificial Neural Networks (ANNs) are essentially semi-parametric regression estimators and are suitable for simulate the behavior of complicate physical phenomena. A significant advantage of the ANN approach in system modeling is that there is no need to have a well defined physical relationship for systematically converting an input to an output. It is needed for most networks provide a collection of representative examples (input-output pairs) of the desired mapping. The ANN then adapts itself to reproduce the desired output when presented with training sample inputs. Network architecture determines the number of connection weights (free parameters) and the way information flows through the network. Determination of appropriate network architecture is one of the most important but also one of the most difficult, tasks in the model building process.

One of the most popular ANN architectures is Multi Layer Perceptron (MLP). A typical MLP has neurons arranged in a distinct layered topology. The input layer simply sends the values of the input variables into the hidden layer. The hidden and output layer neurons are fully connected to all of the units in the preceding layer. Each hidden neuron in an ANN receives a number of inputs from original data or other layer nodes. Each input comes via a connection that has a strength (or weight) attached. The weighted sum of the inputs is formed, to compose the activation of the neuron. The activation signal is passed through an activation (transfer) function produce the output of the neuron. A feed-forward MLP network where nodes in one layer are only connected to nodes in the next layer was used for modeling. Network geometry determines the number of connection weights and how these are arranged. This is generally done by fixing the number of hidden layers and choosing the number of nodes in each of these. It has been shown that ANNs with one hidden layer can approximate any function. The number of nodes in the input layer is fixed by the number of model inputs whereas the number of nodes in the output layer is equal to the number of model outputs. In this study, there was only one output (reservoir inflow). For selecting the final structure of the ANN model, it being a trial and error procedure, started with a minimum number of nodes in the hidden layer and the network was trained until a minimum mean square error will be attained. The number of nodes in a hidden layer has been increased gradually until such an increase did not significantly improve the performance of the neural network.

The process of optimizing the connection weights is known as training or learning. Here, the Levenberg-Marquardt Backpropagation training (LMBP) has been used for train a Feed-forward Neural Network (FNN)

where nodes in one layer are only connected to nodes in the next layer. The sigmoidal transfer functions that are most common was used.

**Models evaluation:** The performances of the models which were developed in this study were evaluated by using a variety of standard statistical performance evaluation measures. Specifically, five different statistical performance indices have been employed: Average Absolute Relative Error (AARE), Pearson’s correlation coefficient (R), Nash-Sutcliff Efficiency (CE), Normalized Mean Bias Error (NMBE) and Normalized Root Mean Square Error (NRMSE). These statistical parameters can be calculated using the total observed and estimated inflow data from the ANFIS and ANN models. On view of the need to consider these different performances criteria, the decision to select the best model is not so easy. Therefore, in this study an Overall Index (OI) comprising of the different performances criteria is proposed as below:

$$OI = \frac{(R^2 + E) \times 100}{((|NMBE| + RMSE) \times AARE)} \quad (18)$$

In order to optimize this performance index, it is desirable to have the maximum value of R and E and minimum value of NMBE, NRMSE and AARE.

The mean error in estimating peak inflow (MF%) and persistence coefficient (Eper) of models were calculated as follows:

$$MF(\%) = \frac{Q_{max} - Q_{omax}}{Q_{omax}} \times 100 \quad (19)$$

$$E_{per} = \frac{EP - E_2}{EP} \quad (20)$$

Where:

$$EP = \sum_{t=1}^n (Q_o(t) - Q_o(t-1))^2 \quad (21)$$

$$E_2 = \sum_{t=1}^n (Q(t) - Q_o(t))^2 \quad (22)$$

Where:

- $Q_o(t)$  = The observed inflow at time t
- $Q(t)$  = The estimated inflow at time t
- n = The total number of runoff data points estimated from an ANFIS model
- $Q_{max}$  = The maximum estimated inflow
- $Q_{omax}$  = The maximum observed inflow
- $Q_o(t-1)$  = The observed inflow at time (t-1)

The MF% statistic quantifies the error in predicting the peak magnitude of inflow. The Eper statistic gives a measure of the goodness of a model in comparison to a persistence model. The values of Eper statistic larger than zero, ensure that the model forecasts are better than those from a simple persistence model.

### APPLICATION AND RESULTS

Different models were built using the transformed and normalized data. The Gaussian membership function was considered. Subtractive Clustering (SC) was used for determining number of membership functions and rules with different value of influence range and squash factor (Table 2). For the case of influence range and squash factor of 0.2 and 1, respectively, it resulted in 61 rules. In this case, the number of model parameters is more than training data pairs (>1600 parameters) and it could not be trained. It is apparent from Table 2 that increasing the influence range and squash factor decreases the rules number. ANFIS training needs at least 2 rules. Thus, in cases of one rule, it could not be trained. The total of 34 models were obtained for different values of influence range and squash factor and they are named by their influence range and squash factor. For

Table 2: Values of overall index for different FISs during training

FIS's name	No. of mfs	Overall index	FIS's name	No. of mfs	Overall index
FIS20125	26	0.634	FIS50100	5	0.335
FIS20150	15	0.435	FIS50125	4	0.297
FIS25100	32	0.662	FIS50150	2	0.261
FIS25125	12	0.489	FIS55100	6	0.345
FIS25150	6	0.577	FIS55125	4	0.345
FIS30100	19	0.510	FIS55150	2	0.259
FIS30125	9	0.397	FIS60100	5	0.303
FIS30150	4	0.305	FIS60125	2	0.256
FIS35100	13	0.414	FIS60150	2	0.231
FIS35125	8	0.359	FIS65100	3	0.231
FIS35150	4	0.281	FIS65125	2	0.250
FIS40100	11	0.359	FIS70100	3	0.296
FIS40125	6	0.297	FIS70125	2	0.228
FIS40150	4	0.255	FIS75100	2	0.251
FIS45100	8	0.420	FIS75125	2	0.252
FIS45125	4	0.289	FIS80100	3	0.278
FIS45150	3	0.258	FIS80125	2	0.212

Table 3: Number of mfs and statistical indices value in selected trained ANFIS models during testing

FIS's name	No. of mfs	R	CE	AARE	NMBE	NRMSE	Overall index
FIS20125	26	0.720	0.387	0.856	-12.948	132.170	0.729
FIS25100	32	0.783	0.520	0.974	-17.470	116.947	0.935
FIS25125	12	0.714	0.388	0.828	-21.065	132.115	0.707
FIS25150	6	0.856	0.699	0.733	-24.985	92.655	1.659
FIS30100	19	0.685	-0.054	0.922	-10.582	173.345	0.245

Table 4: Parameters values of each rule for FIS25150 model

Parameters	Rule number (R <sub>i</sub> )					
	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	R <sub>6</sub>
p	0.26850	0.84940	0.71200	0.09407	0.30370	0.07005
q	0.10670	-0.03653	0.21460	0.53950	0.25310	0.01347
r	0.09951	0.14530	0.00118	0.07349	0.08870	-0.12340
s	0.05382	-0.08029	0.10730	0.04987	-0.00712	0.19980
k	0.17010	0.21630	0.06256	0.10480	0.03099	0.13880
l	-0.04967	0.11730	-0.03665	-0.06978	0.01103	-0.04006
g	0.10760	-0.11850	-0.04445	0.01002	-0.03788	0.12900
h	0.98250	0.06110	0.04976	0.26530	0.09612	0.11810
j	-0.07096	0.03009	0.07283	0.14960	0.13360	0.00707
d	0.19550	-0.09805	-0.12370	0.05067	0.20240	0.45130

instance, FIS30125 is given to a fuzzy inference system that is obtained by subtractive clustering method with influence range 0.3 and squash factor 1.25.

The different fuzzy inference systems were trained by hybrid method and all performance indices were computed. Validation data was considered to avoid over training. The higher values of overall index, belonged to models with more rules (Table 2). Most of the models with ≤4 rules give similar performance indices. The more complex models which consist of more rules caused higher number of model parameters and in most cases over training.

Therefore, to adjust this problem, the five best models based on overall index value were selected and tested with the test data set. The selected models were FIS20125, FIS25100, FIS25125, FIS25150 and FIS30100. The results are shown in Table 3. The overall index results during testing show that the best performance and generalization has been obtained with the FIS25150 model with a reasonable difference against the other 4 models. The values of NMBE show that all models have some under prediction.

The minimum values of NRMSE and AARE indices and maximum values of R and CE show the superiority FIS25150 clearly. Comparison of the ability of models to predicting peak inflows is shown in Fig. 2 using the MF% index. This figure reveals that the least value of error in capturing the peak inflows comes from the FIS25150 model. Figure 3 shows that the FIS25150 could keep its persistency with the testing data just like with the training data. Finally the FIS25150 was selected as the best model. This FIS constructed with 9 inputs and one output. Each input has 6 Gaussian membership functions

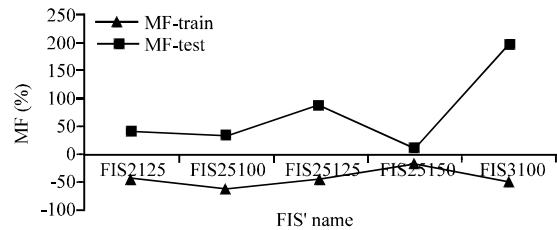


Fig. 2: MF% index for selected models during training and testing

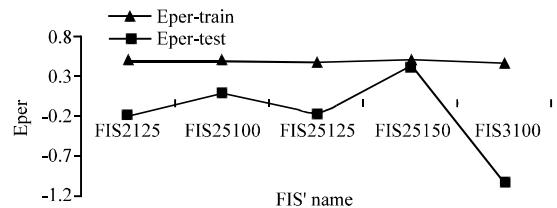


Fig. 3: E-per index for selected models during training and testing

and there are 6 rules. The output has also 6 member functions. The parameters of rules for FIS25150 (consequent parameters) are shown in Table 4. The FIS25150 was applied to the training and testing data and the results are shown in Fig. 4 and 5. The model simulates the base flows in testing as well as training but it is not able to capture peak inflows well. The peak inflows under-prediction is seen in training data but in testing it is apparently under and over predicting peak values.

To evaluate the impacts of the choosing different membership functions on the simulated results, the sensitivity analysis was carried out. For construction of

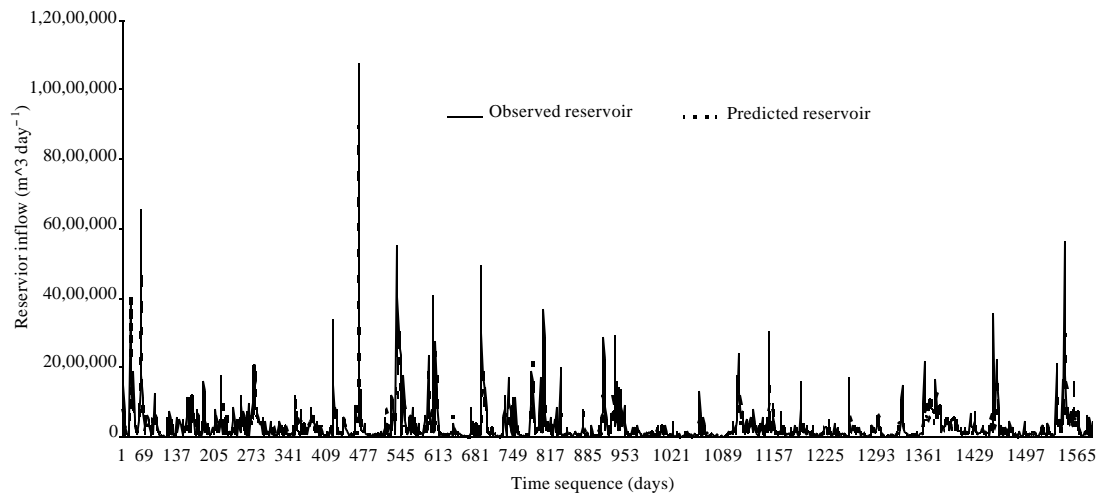


Fig. 4: Simulations of training inflow data by FIS25150 model

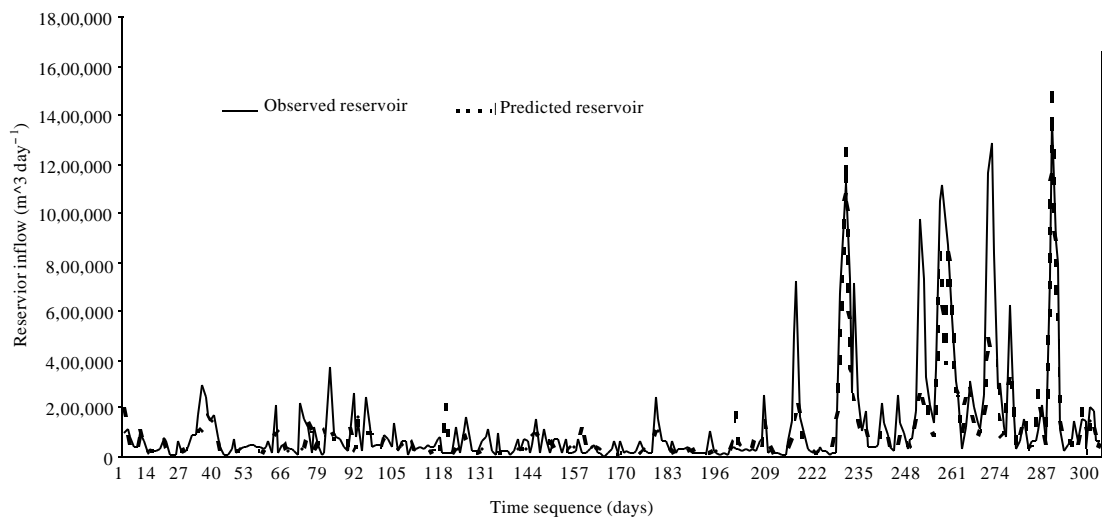


Fig. 5: Simulations of testing inflow data by FIS25150 model

fuzzy inference systems in all ANFIS models, the Gaussian membership function was used. For the final selected model, FIS25150, the FIS was built by Combined Gaussian, Generalized Bell and Difference Sigmoidal membership functions of the Fuzzy Toolbox of MATLAB. The equivalent ANFIS models were trained and tested with the ANFIS editor. The best result in training was obtained by the Gaussian membership function and its overall index was reasonably more than the others (Fig. 6). Figure 7 shows that using Gaussian membership function lead to lesser value of under-prediction of peak flows and higher persistency of model in training period. The best ANN model to forecast inflow to reservoir of Sembrong dam was obtained from the same data and model inputs. Its structure was 9-13-1 with maximum overall

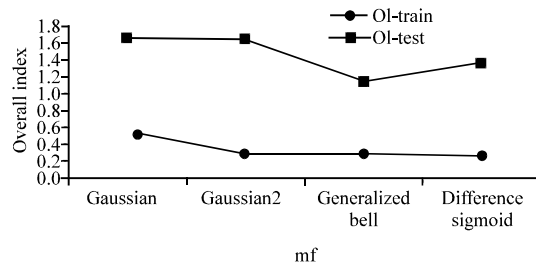


Fig. 6: Overall index for FIS25150 model with different types of membership functions

performances and best value of R and CE. Figure 8 shows this model's architecture. The statistic measures for selected ANN and ANFIS models for validation and testing data are shown in Fig. 9 and 10. With validation all



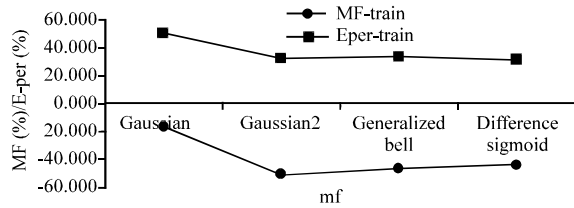


Fig. 7: MF% and E per indices for FIS25150 model with different types of membership functions

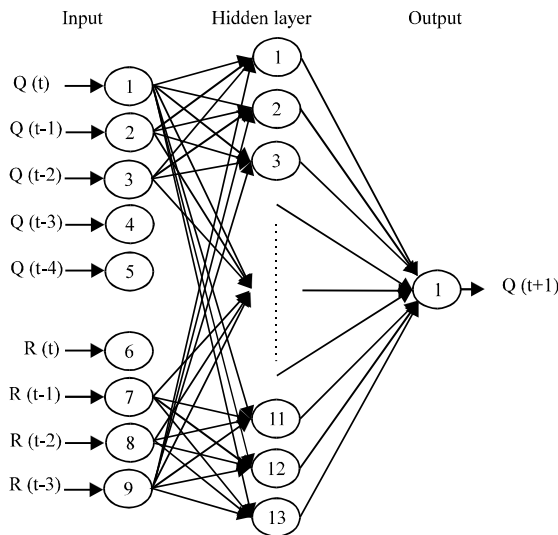


Fig. 8: Structure of selected ANN model

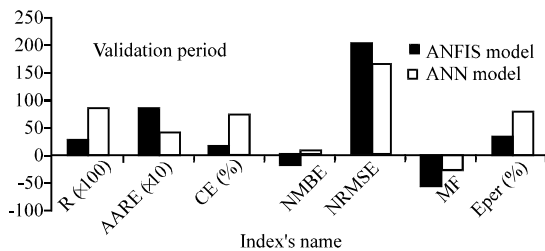


Fig. 9: Statistics measures of final ANFIS and ANN models in validation

performance indices of selected ANFIS model were worse than ANN model. The differences between indices were obviously high in this phase. Figure 9 and 10 show the superiority of the selected ANN model to the ANFIS model. The selected ANFIS model has only six rules and it seems that these linear rules are not sufficient to cover the entire complex relationships between rainfall and discharge. It is noted that in these catchments, there are many water paths behaving like small inflows channels into reservoirs. In other words, the density of overland drainage system is high. Apart from this, the soil

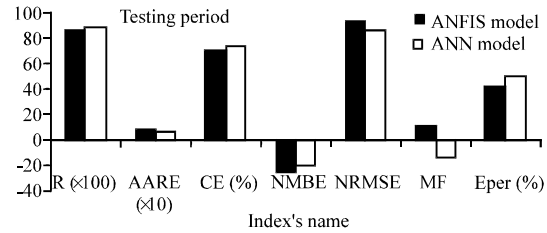


Fig. 10: Statistics measures of final ANFIS and ANN models in testing

antecedent moisture content and the tropical rainfall patterns increases the complexity and non-linearity of reservoir inflow forecasting. However, it may be noted that the ANFIS simplifies the model building process.

### CONCLUSION

The potential of a particular neuro-fuzzy model, Adaptive Neuro-Fuzzy Inference System (ANFIS) for forecasting the daily reservoir inflow was investigated for the Sembrong dam catchment in Malaysia. Daily reservoir inflow was computed using a water balance equation. Average daily rainfall across the catchment was obtained by constructing Thiessen polygons. The appropriate model inputs for modeling the next day daily reservoir inflow were extracted using auto-correlation, cross-correlation and partial auto-correlation functions. The historical data was divided into three independent sets for training, verifying and testing the constructed models. Data transformation to normality reduced the skewness of different data sets.

Different fuzzy inference systems were constructed by subtractive clustering methods and they were trained with hybrid learning algorithm. The models were evaluated for an independent data set by a wide range of performance indices. In general, simulation results of best selected model showed good generalization for base flows and medium peaks.

However, the ANFIS model was not able to capture peak flows. A comparison was made between the simulate results provided by the ANFIS model and the Artificial Neural Networks (ANNs) developed model using same data for study area. The final ANN model performed better comparing to the selected ANFIS model.

### REFERENCES

Altug, S., M.Y. Chow and H.J. Trussell, 1999. Fuzzy inference systems implemented on neural architectures for motor fault detection and diagnosis. IEEE Trans. Ind. Electron., 46: 1069-1079.

- Babuska, R. and H. Verbruggen, 2003. Neuro-fuzzy method for nonlinear system identification. *Ann. Rev. Control*, 27: 73-85.
- Chang, F.J. and Y.C. Chen, 2001. A counterpropagation fuzzy-neural network modeling approach to real time stream-flow prediction. *J. Hydrol.*, 245: 153-164.
- Chang, F.J. and Y.T. Chang, 2006. Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. *Adv. Water Resour.*, 29: 1-10.
- Chang, L.C. and F.J. Chang, 2001. Intelligent control of modeling of real time reservoir operation. *Hydrol. Process.*, 15: 1621-1634.
- Chang, Y.T., L.C. Chang and F.J. Chang, 2005. Intelligent control for modeling of real-time reservoir operation-part II: ANN with operating rule curves. *Hydrol. Process.*, 19: 1431-1444.
- Chiu, S.L., 1994. Fuzzy model identification based on cluster estimation. *J. Intell. Fuzzy Syst.*, 2: 267-278.
- Firat, M. and M. Gungora, 2007. River flow estimation using adaptive neuro fuzzy inference system. *Math. Comput. Simul.*, 75: 87-96.
- Jang, J., 1993. ANFIS: Adaptive network based fuzzy inference systems. *IEEE Trans. Syst. Man Cybernetics*, 23: 665-685.
- Jang, J.S.R., C.T. Sun and E. Mizutani, 1997. *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. Prentice-Hall Inc., New York, USA., ISBN-13: 9780132874670, pp: 614.
- Kisi, O., 2006. Daily pan evaporation modeling using a neuro-fuzzy computing technique. *J. Hydrol.*, 329: 639-646.
- Oonsivilai, A. and M.E. El-Hawary, 1999. Power system dynamic load modeling using adaptive-network-based fuzzy inference system. *Proc. IEEE Can. Conf. Electr. Comput. Eng.*, 3: 1217-1222.
- Openshaw, S. and C. Openshaw, 1997. *Artificial Intelligence in Geography*. John Wiley and Sons Ltd., Chichester, UK., ISBN-13: 9780471969914, pp: 329.
- Ponnambalam, K., F. Karry and S.J. Musavi, 2003. Minimizing variance of reservoir systems operations benefits using soft computing tools. *Fuzzy Sets Syst.*, 139: 451-461.
- Sfetsos, A., 2000. A comparison of various forecasting techniques applied to mean hourly wind speed time series. *Renewable Energy*, 21: 23-35.
- Sugeno, M. and G.T. Kang, 1988. Structure identification of fuzzy model. *Fuzzy Sets Syst.*, 28: 15-33.
- Vernieuwe, H., O. Georgieva, B. de Baets, V.R.N. Pauwels, N.E.C. Verhoest and F.P. de Troch, 2005. Comparison of data-driven Takagi-Sugeno models of rainfall-discharge dynamics. *J. Hydrol.*, 302: 173-186.