APPLYING ADAPTIVE NEURO FUZZY INFERENCE SYSTEM APPROACH TO RIVER LEVEL FORECASTING

Imam Suprayogi[1], Joleha[2], Nurdin[3]
[1,2,3]Civil Engineering of Department, Faculty of Engineering
University of Riau, Telp : 081332 606360
Email : drisuprayogi @ yahoo.com

Abstract

River level forecasting is quite important for reservoir operation studies, flood planning and control, modeling and management water resources. In the last decade, the softcomputing model as a branch of the artificial intelligence science were introduction as a forecast tool beside knowledge based system, expert system, fuzzy logic, artificial neural network, and genetic algorithm. The method that used in this research was a combination between fuzzy logic and artificial neural network which usually called neuro fuzzy system of adaptive neuro fuzzy inference system (ANFIS) algorithm approach was used construct a river level forecasting system. The advantages of this method is that is use input- output data sets. In particular, the applicability of ANFIS as an estimation model for river flow was investigated. To illustrate the applicability and capability the ANFIS, the River Indragiri, located the Indragiri Hulu Residence and the most important water resources of Indragiri catchment's, was choosen as a case study area. To totally 1997-2008 annual data sets collected years were used to estimate the River level. The models having various input structures were constructed and the best structure was investigated. In addition four various training / testing data were constructed by cross validation methods and the best data set was investigated. The performance of the ANFIS models in training and testing sets were compared with the observation and also evaluated. The results indicated that the ANFIS can be applied successfully and provide high accuracy and reliability for River level estimation in Indragiri River.

***Keywords** : Forecasting, River level, Artificial Neural Network, Fuzzy Logic, Adaptive Neuro Fuzzy Inference System.*

1. Introduction

The modeling techniques used in hydrological processes are quite important to provide the accurate and sustainable use of the water resources. In modeling of the hydrological processes, the measurment of natural phenomena is firstly necessary. In order to estimate the hydrological processes such as precipitation, runoff and change of water level by using existing method, some parameters such as the physical properties of the watershed region and observed detail data are necessary. Softcomputing is one of the latest approaches for the development of systems that prossess computational intelligence. Softcomputing attempts to integrate several different computing paradigms including artificial neural networks, fuzzy logic and genetic algorithms. On their own, each of these techniques appears to be extremely effective at handling dynamic, nonlinear and noisy data, especially when the underlying physical relationships are not fully understood. However, when utilized together, the strengths of each technique can be exploited in a synergistic manner for the development of low cost, hybrid systems [13].

However, the integration of these different, softcomputing technologies to produce a single, hybrid solution for the enhancement of operational river level and flood forecasting systems still remains to be investigated [12] , [13] . Existing hydrological forecasting models are often highly data specific,and their operational performance depends upon both the specification of the model, which is based on existing hydrological knowledge, and the ability of themodel to respond to dynamic and rapidly changing events. Softcomputing, on the other hand, offers a more flexible, less assumption-dependent and potentially self adaptive approach to modelling flood processes, which by their nature are inherently complex, nonlinear and dynamic. Moreover, these techniques can be used for modelling

ISSN : 2085-9902

systems on a real-time basis. Other advantages include: the potential for improved performance, faster model development and execution times and, therefore, reduced costs, the capability to plug softcomputing components directly into conventional models, and the ability to provide a measure of prediction certainty via bootstrapping techniques. Moreover, it would even be possible to retrain the models online, and there ere by enhance the ability of the softcomputing components to adapt to rapidly changing future events an important feature for handling the unknown effects of future climatic changes or of storm damage.

Several studies have been carried out using fuzzy logic in hydrology and water resources planning [1], [4], [7], [8], [9], [11]. Recently ANFIS which consists of of the ANN and fuzzy logic methods, have been used for many application such as, data base management, system design and planning/forecasting of the water resources [1], [2], [3], [11]. The main purpose of this study is to the present novel ANFIS to the real-time river level and flood forecasting. The methodology for combining each of the techniques into a single forecasting solution is outlined. To verify the application approach, the Indragiri River Catchment's location in Indragiri Hulu District is chosen as the case study area. Indragiri Catchment's is one the most important water resources in Indragiri Hulu District, Riau Province. In the region Indragiri River has a quite significant effect of water drinking,demand use and recreation facilities. In the face of these impacts, forecasts of future river level can be of help in making efficient operating decision of water demand for a wise and sustainable use of the river. To exemplify its applicability and demonstrate that the adaptive network fuzzy inference system has the ability to deal with hhuman activities, we developed ANFIS models for time series forecasting with having various input structures. In the following, firstly the main network structure of the ANFIS model and the parameters estimating algorithms are defined and a description of the study area, available data and the model construction are describe. Finally, the results of all ANFIS modes are discussed.

Adaptive Neuro Fuzzy Inference System

Basic fuzzy terminology

The fuzzy logic approach is based of the linguistic uncertain expression rather than numerical uncertainy. It is an artificial intelligence technique that has been used currently in hydrological processes. Since Zadeh proposed the fuzzy logic approach to describe complicated systems, it has become popular and been successfully used in various engineering problem, especially on control processes [2], [3], [7], [9], [10], [11]. Fuzzy set theory was first developed by Zadeh. It is primarily used to deal with uncertain and imprecise knowledge. It canbe thought of as an extension of the traditional crisp set theory. Let A be a crisp set. A individual x from a universal set X is determined either to member of A or a non member of A. This can be expressed by

$$\mu_A(x) \; : \; X \;\; \text{for} \; \{0,1\}$$

Crisp set theory using Boolean logic cannot be used to represent vague concepts such as the terms large, medium and small. This be overcome by using fuzzy logic which extends the range of true values to all real numbers in the interval between 0 and 1. Fuzzy logic can be best understood using set membership where the membership values represent the degrees with which each object is associated with the properties that are distinctive to the collection of objects with membership values between 0 (complete exclusion) and 1 (complete membership). Membership grade of each element in X is determined through a membership function μ_A which map the elements of an universe of discourseX to the unit interval [0,1], that is

$$\mu_A \; : \; X \;\; \text{for} \; [0,1]$$

By using approximate reasoning, a fuzzy logic description can be used to effectively model the uncertainy and non linearity of a system.

Adaptive Neuro Fuzzy Inference System Architecture

The proposed adaptive neuro fuzzy inference system model is multilayer artificial neural networks based fuzzy system. A typical architecture of an ANFIS, in which a circle indicates an adaptive node, is shown in Fig.1. In this connectionist structure, the input and output nodes and the hidden layers, there are nodes functioning as membership functions (MFs) and rules. This eliminates the disadvantage of a normal feedforward multilayer network, which is difficult for an observer to understand or to modify [5]. For simplicity, we assume that the examined fuzzy inference system (FIS) has two inputs and one output. For a first order Sugeno fuzzy model, a typical rule set with to two fuzzy “if - then” rules can be expressed as follows:

Rule 1 : If x is A_1 and y is B_1 then $f_1 = p_1x + q_1y + r_1$ (1)

Rule 2 : If x is A_2 and y is B_2 then $f_2 = p_2x + q_2y\ r_2$ (2)

Where x and y are the two crisps inputs, and A_1 and B_1 are the linguistic label associated with the node function. As indicated in Fig.1, the system has a total of five layers. The functioning of each layer is described as follows.

Layer 1 : All the nodes in the first layer are adaptive nodes which means that the outputs of the nodes depend on the parameters pertaining to these nodes. Each node corespondends to a linguistic label which has a membership function μ_{Ai} and μ_{Bi}.. The output of a node in this layer specifies the degree to which the given input satisfies the membership function. The node function of a node i can be expressed by :

$O_i = \mu_{Ai}(x)$ for i=1,2

$O_i = \mu_{Bi}(x)$ for i=3,4

(3)

In this paper, membership functions μ_{Ai} and μ_{Bi} are chosen to be Gaussian shaped with maximum equal to 1 and minimum equal to 0. Parameters in this layer are referred to as premise parameters

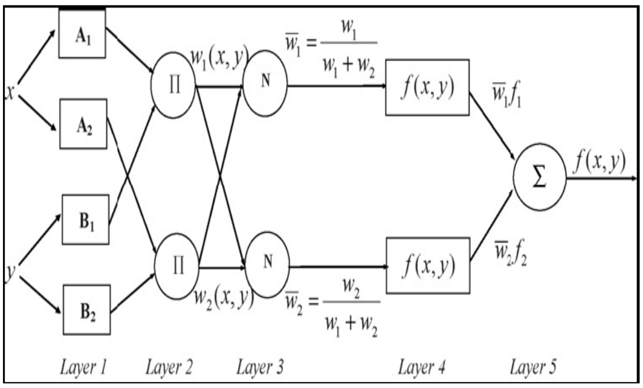


Fig 1 Architecture of the ANFIS

Layer 2 : Nodes in the first layer are labeled Π, whose output represents a firing strength of rule. The nodes generates the the outputs (firing strength) by cross multiplying all the incoming signals:

$O_i = W_i = \mu_{Ai}(x) \times \mu_{Bi}(y)$ for i = 1,2 (4)

Layer 3 : Every nodes in this layer is fixed node labeled N. The i th node calculates the ratio between the i th rule's firing strength to the sum of all rule's firing strengths:

$$O_i = \overline{w}_i = \frac{w_i}{w_1 + w_2}, i = 1, 2 \tag{5}$$

The outputs of layer are called normalized firing strengths.

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

$$O_i^4 = \overline{w} f_i = \overline{w_i} (p_i x + q_i y + r_i) \tag{6}$$

Where $\{p_i, q_i, r_i)$ is the parameter set of this node. These parameters are reffered to as consequent parameters.

Layer 5 : The single node in this layer is a fixed node labeled E, which computes to overall output by summing all incoming signal:

$$O_i^5 \; overalloutput = \sum_i \overline{w_i} \; f_i = \frac{\sum_i w_i \; f_i}{\sum_i w_i} \tag{7}$$

There are two major phases for implementing the ANFIS for specific application : the structure identification phase and parameter identification phase. The structure identification phase involves finding a suitable number of fuzzy rules and fuzzy sets and a proper partition feature space. The parameter identification phase involves the adjustment of the premis and consequence parameters of system. More detailed descriptions of the two phase are provided in the following two sections.

Parameter identification using hybrid learning algorithm

During the learning process,the premise parameters in the layer 1, {c, σ}, and the consequent parameters in the layer 4, {p, q, r}, are tuned until desired response of the FIS is achieved. The two frequently used training methods are the backpropagation (BP) algorithm [1] and the hybrid learning algorithm [5]. In this paper, the hybrid learning algorithm, which combines the least squares method (LSM) and the BP algorithm, is used rapidly train and adapt the FIS. This algorithm convergens much faster since it reduces the dimension of the search space of the original BP algorithm [5].

When the premise parameters are fixed, the overall output can be expressed as a linear combination of the consequent parameters. The output f can then be written as:

$$f = \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 \; w_2} f_2$$

$$f = \overline{w}_1 (p_1 x + q_1 y + r_1) + \overline{w}_2 (p_2 x + q_2 y + r_2)$$

$$f = (\overline{w}_1 x) \, p_1 + (\overline{w}_1 y) \, q_1 + (\overline{w}_1) \, r_1 + (\overline{w}_2 x) \, p_2 + (\overline{w}_2 y) \, q_2 + (\overline{w}_2) \, r_2 \tag{8}$$

Equation (3) is linear the consequent parameters p_1, q_1, r_1, p_2, q_2 and r_2.

The hybrid learning algorithms of ANFIS consist of the following two parts [1] : (a) the learning of the premise parameters by backpropagation and (b) the learning of the consequent parameters by least square estimation. In the forward pass of the hybrid learning algorithm, functionsignals go forward until layer 4 to calculate each node output. The non linear or premise parameters in the layers 2 remaind fixed in this pass. The consequent parameter are identified by the least squares estimate. In the backward pass, the error propagate backward from the output end towards the input end, and the premise parameters are updated by the gradient descent. Provided the detailed description and the mathematical background of the hybrid learning algorithm [6].

2. Research Methods

Study Area and Data

The applicability of ANFIS as a water level forecasting model is investigated. To illustrate the validity and capability of ANFIS method for time series forecasting and modeling, Indragiri River of Indragiri Hulu Residence is chosen. It has been operated for drinking water, domestic use and recreation facilities. There is one River level staff gauging station, equipped which manual daily River level records on Indragiri River as shown in Figure 2.

Figure 2. The river staff gauge station in Pekan Heran Bridge, Indragiri River

One of the most important steps in developing a satisfactory forecasting model is the selection of the input variable. Because, these variable determine the structure of the ANFIS model and affect the wighted coefficient and the results of the model. Different combination of the antecedent level of the river staff gauge station in Pekan Heran Bridge was construce the appropriate input structure in the time series forecasting model. The general structure of time series model forecasting model is given in Equation (9).

$$H(t+1) = f\left[(H_t)\right] \quad (9)$$

Where H_t represent the river level at time (t), H_{t-1} is the river level at time (t-1). It is evident that the training data sets should cover all the characters of the problem in order to get effective estimation. For this aim, data set were devided two training/testing subsets, where are m_3 and m_2 by cross validation method as a systematic process to get effective and sensitive modeling. The structures of data set are given in Table 1, Table 2 and Figure 1.

Table 1. The Cross Validation Data Sets for River Level Estimation

Cross Validation	Dates of the data	Numbers of Data
m_3	01.1.2007- 31.1.2009	1095

m_2	01.6.2002- 30.6.2005	1095

Table 2. The Structure of the Training and Testing Data Sets

Model ANFIS	Dates of the training set	Numbers of training Data	Dates of the testing set	Numbers of testing Data
M-1	1998-2006	3285	2007-2009	1095
M-2	1998-2002 2005-2009	1642 1643	2003-2005	1095

Model ANFIS	Training	Training	Testing
M-1	m_1	m_2	**m_3**

Model ANFIS	Training	Testing	Training
M-2	m_1	**m_2**	m_3

Figure 3. The training and testing data sets of River level estimation

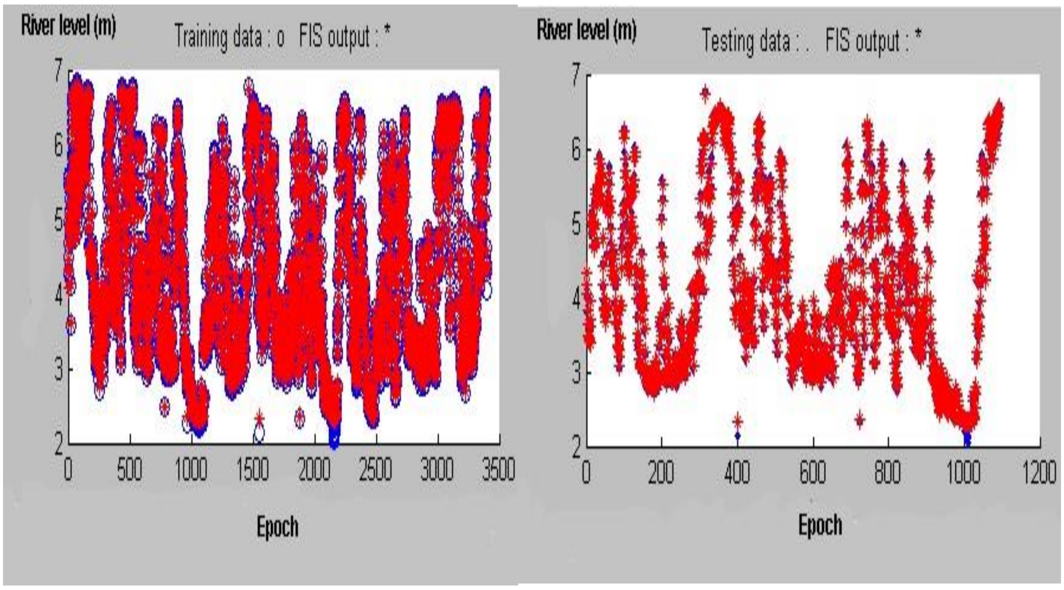
In each model every input variable might be clustered into several class values in layer 1 to build up fuzzy rules. Each fuzzy rule is constructed throught several parameter of membership function in layer 2. If the number of parameters that needs to be determined increases with the fuzzy rule increment, the model structure will becomes more complicated. In this study, the subtractive fuzzy clustering function is used to establish the fuzzy rule based on the relationship between the input output variables. In order to determine the nonlinear input and linear output parameters, the hybrid algorithm is used. The learning procedure and the construction the rule are provided by this algorithm. The performances of the ANFIS model both training/testing data were evaluated and the best training/testing data selected according to statisical parameter such as, mean square error (MSE). The MSE formulation is expressed by :

$$MSE = \frac{\sum_{t=p+1}^{N}(H_t - \hat{H}_t)^2}{N} \qquad (10)$$

Where X_t represent time series data the river level at time period t, $\hat{X}_t$ represent time series prediction value the river level at time period t for t =1, 2, 3 ...., N and p represent the longest period of time which influence the forecast.

3. Result and Analysis

The ANFIS is simulated using Toolbox with MATLAB version 7.0 of river level forecasting in Indragiri River as an aid program. The results of the ANFIS models and observation are compared with the observed levels in order to evaluate the performance of the training/testing of the real time river level model. Figure 4 show the estimated values of the training/testing of the ANFIS models and observation values.



Sources : Running Program ANFIS GUI using Toolbox with MATLAB 7.02

Figure 4. Training and Testing of ANFIS (M-1) Model and Measured

The Figure 4 nicely demonstrates that ANFIS (M-1) model performance, in general, accurate and good, where all data points are quite near the line of agreement. The results of the ANFIS model demonstrate that the ANFIS (M-1) can be successfully applied to establish accurate and realiable River level estimation models (see Table 3).

Table 3. The training and testing parameter of ANFIS (M-1) Model

Number of Rules	24
Epoch	35
Range of Influence	0.05
Number of Membership function	Gauss 8
AND method	prod
Defuzzification method	wtaver

Sources : Running Program ANFIS GUI using Toolbox with MATLAB 7.02

The comparison of the resuts of the ANFIS (M-1), ANFIS (M-2) and Naive models is given in Table 4.

Table 4. The performance of the ANFIS (M-1) model

Models	Training set	Testing set	Cheking set
	MSE	MSE	MSE
ANFIS (M-1)	0.27	0.30	0.007
ANFIS (M-2)	0.28	0.33	0.009
Naïve	0.35	0.45	0.130

4. Conclusion

In this study, the applicability and capability of ANFIS method was investigated for river level forecasting. To verify application this approach, the Indragiri River cathcment's is chosen as the case study area. The data set include 4380 daily water level data for the period at 1997-2008 years and data set was divided two subsets by cross validation data sets method. The model having various input structure were trained and tested with all cross validation data sets. ANFIS model is compared on their performance in training and testing sets. It may be noted that the model is trained using non-transformed data. Based on the finding of this study following conclusion can be drawn:

- The performance of ANFIS model and observation are compared and evaluated. It appears that the ANFIS models are more accurate and reliable the value MSE is smaller than conventional methods (Naive model). The performance ANFIS model was found to be satisfactory on the basis of performance evaluation of model.
- The model that use Toolbox with Matlab version 7.0 as an aid program software were very sensitive to changes in range of influence parameter and were also have an accurate forecast range for a day ahead (H_{t+1}).
- ANFIS (M-1) model gives best results among ANFIS (M-2) model and Naïve Model.

Acknowledgement

The author is grateful for Prof.Ir.H.Nizam, MSc, PhD (Civil Engineering Department, Faculty of Civil Engineering and Environmental, Gadjah Mada of University, Jogyakarta), Dr. Ir. Edijatno, CES, DEA (Civil Engineering Department, Faculty of Civil Engineering and Planning, Sepuluh Nopember Institute of Technology,Surabaya), Prof. Dr. Techn. M Isa Irawan, MT (Mathematic Department, Faculty of Matemathic and Natural Science, Sepuluh Nopember Institute of Technology, Surabaya) and Dr. Hj.Sri Kusumadewi,SSi, MT (Informatic Engineering Department, Islamic Indonesia of University, Jogyakarta) for their hepful and support of this paper.

References

[1] [Chang F.J, Chang Y.C. Counter Propagation Fuzzy-Neural Network for River Flow Reconstruction, Hydrolic Processes , 2001; 15, 219-232.
[2] Chang F.J, Chang Y.T. Adaptive Neuro Fuzzy Inference System for Prediction Water Level in Reservoir, Advances in Water Resources 29, 2006,1-10.
[3] Cheng S.H, Lin Y.H, Chang L.C, Chang F.J., The Strategy of Building a Flood Forecast Model by Neuro Fuzzy Network, Hydrolic Process, 2006; 20, 525-1540.
[4] Ertunga C.O, Duckstein L. Fuzzy Conceptual Rainfall-Runoff Models. Journal of Hydrology. 2001; 253: 41-68.
[5] Jang J.S.R. ANFIS : Adaptive Network Based Fuzzy Inference System. Journal of IEEE Transactions System, Man and Cybernetics. 1993; 23 (3), 655-685.
[6] Jang J.S.R, Sun C.T, Mizutani E. Neuro-Fuzzy and Softcomputing. New York : Springer Englewood Cliffs, 1997: 607-620.
[7] Liong S.Y, Lim W.H, Kojiri T, Hori T. Advance Flood Forecasting for Flood Stricken Bangladesh with Fuzzy Reason Method. Hydrolic Processes , 2000; 14, 431- 448.
[8] Mahabir P.C., Hick F.E., Fayek A.R . Application of Fuzzy Logic to the Seasonal Runoff, Hydrolic Processes , 2000; 17, 3749-3762.
[9] Mitra B, Scott F.E, Fayek A.R . Application of Fuzzy Logic to the Prediction of Soil Erosion in a Large Watershed, Geoderma, 1998; 86, 183-209.
[10] Nayak P.C., Sudheer K.P, Ramasastri., Fuzzy Computing Based Rainfall- Runoff Model for Real Time Flood Forecasting, Hydrolic Process, 2004 a; 17, 3749-3762.
[11] Nayak P.C, Sudheer, K.P, Ramasastri, A Neuro Fuzzy Computing Technique for Modelling Hydrological Time Series, Journal of Hydrology, 2004b; 29, 52-56.
[12] See L, Shaw O. Using Softcomputing Techniques to Enhance Flood Forecasting on The River Ouse. In: Proceding Hydroinformatics (ed. by V. Babovic & L. C. Larsen), Rotterdam. 1998; 98 : 819-824.
[13] See L, Shaw O. Applying Artificial Neural Network Approach to River Level Forecasting. Journal de Sciences Hydrologiques. 1999; 44(5) : 763-778.