Decision Support to Watershed Modelling Using Soft Computing

[1]P. V. Desai, [2]S. K. Ukarande

[1]Civil Engineering Department, Yadavrao Tasgaonkar College of Engineering & Management, University of Mumbai, Maharashtra, India

[2]K. J. Somaiya Institute of Engineering and Information Technology, University of Mumbai, Maharashtra, India

Abstract − In this study, Artificial intelligent(AI) technique i.e. Artificial Neural Network and Classical Multi linear regression method are used to predict one day ahead intermittent stream inflow i.e. runoff one of the important parameter of the watershed . To demonstrate the application of Artificial Neural Network, Ghataprabha basin which consist three watersheds i.e. Hiranyakeshi, Ghataprabha and Tambraparni in Maharashtra, India are selected as a part of case study. For the current study, the multilayer perceptron (MLP) neural network along with feed forward back propagation network is used. Based on the observed daily rainfall, antecedent rainfall, runoff, antecedent runoff with one, two and three day time lag and daily evaporation are used as inputs for model development. In this study runoff-runoff models, rainfall-runoff models and evaporation-rainfall- runoff models were developed to get one day ahead runoff. The model performance was evaluated using various performance criteria such as Correlation coefficient(R), Mean Square Error (MSE) and Root Mean Square Error (RMSE), from this result it is observed that ANN models are better than performance the MLR models but still both the methods had underestimated the peak flow. After the analysis of all the models it is observed that rainfall-runoff model gives the best performance. Total twenty years data for discharge (runoff), rainfall and evaporation, from the year 1990 to 2010, was used for developing the models to predict runoff, one day in advance.

Keywords: Artificial neural network, Multi Linear regression, feed forward back propagation

I. INTRODUCTION

Water and land are the two most important natural resources required for the survival of living and non-living thing, on earth next to air. The environment and economy encompasses the two vital resources. Therefore for the sustainable development of environment, economy and to provide life supporting system for human being and animal it is inevitable to inculcate efficient management practices of land, water and vegetation. A watershed is a natural, complete topographical and hydrological entity that assembles and converts all the rainwater falling on it to a common outlet. Watershed modeling deals with modeling of all of the hydrologic processes at the watershed level and combining them to determine the watershed response. The watershed scale in hydrology can be as small as a few square kilometers meters in area or as large as thousands of square kilometers. Because of the significant diversity between different sizes of watersheds, watershed models are classified as small watershed, medium watershed and large watershed models. Modeling of rainfall-runoff at watershed scale is important for water resource management as well as safe yield computation and most importantly for design of flood control structure. Frequent floods occurring worldwide had arises the need for good predictive models at the watershed scale. Watershed runoff rely on geomorphologic properties (such as vegetation, topology, soil type) of the watershed and other climatic factors (Precipitation, Evaporation, Temperation.etc

II. LITERATURE SURVEY

A. Literature Review

Assessment, Simulation and prediction of rainfall and corresponding runoff are essential for stakeholders and policy makers to plan or adort the required policies. There are various techniques available in the literature to assess, simulate and predict hydrological variables. However, the selection of various techniques normally depends on the objectives of the study, availability of required inputs data, some predefined assumptions and the quality of available models. According to Makridakis (1998) different methods are different in terms of accuracy, scope, time horizon and the cost. To facilitate an adequate level of accuracy, the developer has to be responsive to the characteristics of different methods is appropriate for the undertaken situation before the start of its usage in real application. As a result, the choice of a model is important among all factors that will influence the forecasting accuracy. In the following sections, literature review on the application of various hydrological and ANN models for rainfall and runoff assessment, simulation and prediction has been summarized.

B. Runoff simulation models

a. Based on Artificial neural network

A new magnitude has been added to the system-theoretic modelling approach through the adortion of the artificial neural network (ANN) technique for rainfall-runoff modelling. Presently many researchers are utilizing ANNs because these models had own desirable attributes of universal approximation, and the ability to learn from examples without the need for explicit physics. In 1890, William James published the first work about brain activity patterns. In 1943, McCulloch and Pitts composed a model of the neuron that is still used today in artificial neural networking. In 1949, Donald Hebb published ―The Organization of Behavior‖, which defined a law for synartic neuron learning. This law, later known as Hebbian Learning in honour of Donald Hebb is one of the simplest and most straight-forward learning rules for artificial neural networks. In 1951, Marvin Minsky built the first ANN while working at Princeton.

Kuok king kuok et.al(2007) has done artificial neural network for daily rainfall runoff modelling in sungai Bedup Basin, Sarawak as build using MLP, REC networks. Jain et al(2003) modeled an event based rainfall runoff process using unit hydrograph (UH)theory, statistical regression and the ANN.A comparative analysis of all of the modelling techniques revealed that the ANN is the most suitable technique and would be efficient in modelling an event based rainfall-runoff process for selecting peak discharge and its occurrence time very accurately. Jain et al(2004) presented a new class of model called gray box model that integrate deterministic and ANN technique, which was found to perform better than the purely black box type ANN rainfall-runoff model.

b. Based on Multi Linear Regression network

Since the basic features of the watershed remain unchanged in years, certain correlation between the input and output might exist, variable MLR is the simplest and well developed representation of a formal, time invariant relationship between time as an input function and corresponding output function. MLR models are considered as benchmark in reservoir inflow forecasting, Diskin(1970) viewed a LRM as a simple concertual model and made clear the physical meaning of the regression coefficients. Driver et.al (1989) used LRM for predicting urban storm-runoff quality and quantity. The use of different regression models in the data was explained in detail by Hirsch (1979) and Hirsch et al (1984). Chiew et al (1993) recommended minimization of the sum of the two aimed functions; these objective functions were minimized using the Levenberg-Marquardt (LM) non-linear least square algorithm. Loague and Freeze (1985) used unit hyderograph theory, regression analysis and quasi-physically based models for high ground Basins and concluded that regression models perform marginally better. Jagdeesh(2000) used sum of squares of errors(SSE) between assumed and observed values particularly useful to take low monthly flows into account. Jain and Prasad (2003) investigate two different types of regression models whiz are linear multi-regression and nonlinear multi-regression models to model an event based rainfall-runoff progression. Raman et al(1995) studied multiple regression models namely, runoff coefficient models, monthly linear regression model, single linear regression, monthly LR model, monthly LR with statistical probality descrirtion and double regression models. All these models were used to improve the monthly stream flow data at a site where the available historic rainfall and stream flow data are short for adequate system study. Jothiprakash et al (2007) developed multiple LR models based on different input function combinations namely, time-series, cause-effect and combined for modelling monthly rainfall-runoff relationship for Kanand watershed in Maharashtra, India. It was reported that multiple model combination approach was more efficient than a single forecast model. The above studies confirm that still MLR models are very much sought because of having the advantage of a relationship between input and output.

c. History of feed-forward Multi Linear ercertron-

The most widely used ANN in Water resource and a hydrological application is the feed forward MLP. (S.Rian and J.Mania,2004) had used multilayer feed forward back with propagation in their hydrological relevance of rainfall runoff modelling with artificial neural network. (Kumar Abhishek, Abhay Kumar,2012)had used feed forward network in their research work of model predicting rainfall using artificial neural network. (Sobri Harun and Nor Irwan) had compared the working of multilayer percertron (MLP)and radial basis function (RBF) algorithm to predict daily runoff as a work of daily rainfall for the Sungai Ketil catchment area. The three layer feed forward error back-propagation algorithm was used (Archana Sarkar, Rakesh Kumar, 2012)for runoff rainfall modelling in Ajay river. Feed-forward network was used in different papers in which speed was of importance (Shukla et.al. 1996;Yang et.al.1996).The feed-forward back propagation algorithm was used for evaporation modelling by(K.P.Sudheer, A.K.Gosain,2002).In a proportional study of artificial neural networks and neuro-fuzzy in continuous modelling of the hourly and daily behavior of runoff(Muhammad Aqil et al 2007)had constructed three different adartive techniques named, Levenberg-Marquardt feed forward neural network and Bayesian regularization feed forward neural network and neuro fuzzy and the performance was check between these three algorithm. (S.M.Chen et al)had done artificial neural network approach for modelling rainfall runoff for which feed forward back propagation network(FFBP)and conventional regression analysis(CRA)were used to study their performance, it was observed that the performance of FFBP exceeded that of regression analysis. The application of multilayer feed-forward ANN for rainfall runoff modelling of the Geer catchment was done by N.J.de.Vos et al.

III. STUDY AREA

In the present study runoff –runoff model, rainfall-runoff model and rainfall-runoff-evaporation model are developed to predict runoff one day in advance at five rain gauge stations namely, Ajara Ramtirth, Kadal, Tarewadi, Jambre Umgaon and Nadgadwadi in Ghataprabha basin, which are along the stream Hiranyakeshi, Ghataprabha and Tambraparni.Ajara Ramtirth station is in Ajara taluka (District-Kolhapur). It is on Ghataprabha, tributary of Krishna basin on Hiranyakeshi stream (exact location detail: 6 km away from Ajara on Ajara-Ramtirth Road). Kadal station is in Gadhinglaj taluka (District-Kolhapur). It is on Ghataprabha, tributary of Krishna basin on Hiranyakeshi stream (exact location details: 6 km away from Sankeshwar on Sankeshwar-Halkarni (via Nool) road).Tarewadi station is in Gadhinglaj taluka (District-Kolhapur). It is on Ghataprabha, tributary of Krishna basin on Ghatprabha stream (exact location details: 3 km away from Nesari on Chandgad Kolhapur road). Jambre Umgaon station is in Chandgad taluka (District-Kolhapur). It is on Ghataprabha, tributary of Krishna basin on Tambraparni stream (exact location-15 km away from Chandgad on Chandgad-Jambre project Approach road). Nadgadwadi station is in Chandgad taluka (District-Kolhapur). It is on Ghatprabha, tributary of Krishna basin on Tambraparni stream (exact location-7 km away from kowad on Ajara-Belagam road)Ghataprabha Basin comprises of 12 watersheds extending over an area of 2000 sq.km.which includes the stream like Ghataprabha and its important tributaries, Hiranyakeshi and Tamraparni which drain through the southern part of the Kolhapur district. The origins of Ghataprabha and Hiranyakeshi rivers are located in Choukul and Amboli respectively. The river has developed a narrow valley with dendritic to sub-parallel drainage pattern having general flow towards the east. The river passes out of the district near village Saroli after a heavy meandering course. The annual rainfall in the area varies between 6500 mm to 500 mm. The basin is partially covered in three talukas viz. Gadhinglaj, Ajara and Chandgad which comes under the Kolhapur District of Maharashtra state.

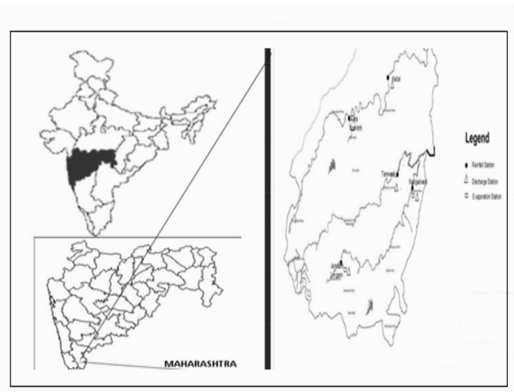

Figure 1 Index Map Of Ghataprabha Sub-basin (Source-AutoCad)

IV. METHODOLOGY

A. General

Basically, an artificial neuron is a system of three elements: a) first A set of weights and biases, responsible for ANN learning; b) second Summation units computing the linear combination of the inputs; and c) third a transfer function, which determines the response. The most frequently used functions are the linear and sigmoid functions (Haykin, 1994). Usually, the values are normalized, thus transforming the real data into a scale compatible with the characteristics of the transfer function. Although other ANN methods are reported in the literature, the ANN employed in this study uses the feed-forward back-propagation method. In this method, the connection of several neurons is distributed across layers.

Within the ANN, the data flow in a single direction (feed-forward) i.e. the input data are propagated through the ANN, layer by layer, in the forward direction. The inputs in the input layer are multiplied by the weights of the respective connections. In mathematical terms, the output from a three layer ANN is represented by equation (1)

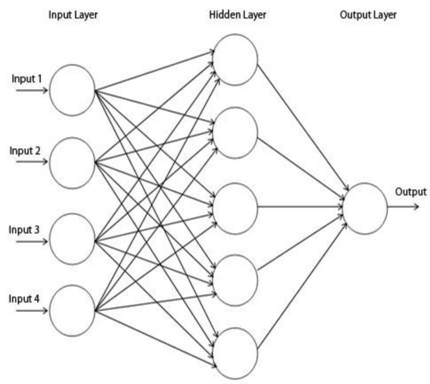

Figure 2 Multi layer Architecture with single hidden layer

Fig 2. Multi layer Architecture with single hidden layer

$$y_k = \varphi\left(\sum_{j=1}^{q} W_{kj}\, \varphi\left(\sum_{i=1}^{p} w_{ji}\, x_i + b_j\right) + b_k\right) \ldots\ldots (1)$$

Where x are the input elements, w are the weights between the connections, b the biases, p the number of neurons in the middle layers, is the transfer function, y is the ANN output and i,j, and k are neurons of the input, middle and output layers, respectively.

B. Regression Analysis:

a. Multiple Linear Regression (MLR):

MLR is the expansion of Simple linear Regression to the case of multiple explanatory variables. MLR is the

procedure of establishing relationship between a dependent variables "y" and set of independent variables

$X_1, X_2, X_3 \ldots X_n$, governing a phenomenon. In these work, one day ahead runoff variable is taken as dependent variable where as runoff at current time step, previous time step similarly rainfall at current time step, previous time step and Evaporation at current time step are taken as independent variable.

A well-know spread-sheet EXCEL was used for MLR model. Regression analysis ortion available in data analysis tools ortions was explored.

C. Data transformation

Luk et al.(2000) and Aqil et al(2007) revealed that networks trained on transformed data achieve better performance. A logarithmic transformation has been pre-owned to bring the observed data to near normal distribution. Transformation is performed on every input output variable independently using the following equation.

$$Z_{p.t.} = a\log_{10}\left(P_{obs\,(t)} + b\right)$$

$$Z_{q.t.} = a\log_{10}\left(Q_{obs\,(t)} + b\right)$$

$$Z_{e.t.} = a\log_{10}\left(E_{obs\,(t)} + b\right)$$

The predicted results were then back-transformed using the following equation

$$Q_{pred\,(t)} = 10^{Z_{qt}/a} - b$$

Where $Z_{p.t.}, Z_{q.t.}, Z_{e.t.}$ are the transformed values of the rainfall, runoff and evaporation during time period‘t’, a and b are arbitrary constant assumed as 0.5 and 1, respectively.

D. Evaluation criteria for ANN prediction-

1) The correlation coefficient (R-value) has been widely used to evaluate the goodness-of-fit of hydrology and hydrodynamic models (Evaluating the use of goodness-of-fit measures in hydrologic and hydro-climatic model validation, Legates, D.R. and G.J.McCabe;1999).This is obtained by performing a linear regression among the ANN predicted values and the targets and is computed by

$$R = \frac{\sum_{i=1}^{N} t_i p_i}{\sqrt{\sum_{i=1}^{n} t_i^2 \sqrt{\sum_{i=1}^{n} p_i^2}}}$$

Where R is correlation coefficient; N is the number of samples; $t_i = T_i - \overline{T}$; $p_i = P_i - \overline{P}$ and T_i

And P_i are the target and predicted values for i=1..., N, and $\overline{T}$ and $\overline{P}$ are the mean values of the target and predicted data set, respectively.

2) The ability of the ANN-predicted values to match measured data is calculated by the Mean Square Error (MSE). It is defined

$$MSE = \left(\frac{1}{N}\sum_{i=1}^{N}(T_i - P_i)^2\right)$$

3) The ability of the ANN-predicted values to match measured data is calculated by the Root Mean Square Error (RMSE). It is defined

$$RMSE = \sqrt{\left(\frac{1}{N}\sum_{i=1}^{N}(T_i - P_i)^2\right)}$$

E. Model Development

Overall the ANN results are more precise if R, MSE and RMSE are found to be close to 1, 0 and 0, respectively. Following Table(4.5.1) is sample of nine different models for each monsoon month one station one month, these models are same for all five station

Table 1 sample of nine different models

Sr. No	Station Name	Model Name	Input	Output
1	Kadal	Kajune1	Qt-1,Qt	Qt+1
2	Kadal	Kajune2	Qt-2,Qt-1,Qt	Qt+1
3	Kadal	Kajune3	Qt-3,Qt-2, Qt-1, Qt	Qt+1
4	Kadal	Kajune4	Qt-1, Qt, Pt	Qt+1
5	Kadal	Kajune5	Qt-1, Qt, Pt-1, Pt	Qt+1
6	Kadal	Kajune6	Qt-2, Qt-1, Qt, Pt	Qt+1
7	Kadal	Kajune7	Qt-2, Qt-1, Qt, Pt-1, Pt	Qt+1
8	Kadal	Kajune8	Qt-3, Qt-2, Qt-1, Qt, Pt	Qt+1
9	Kadal	Kajune9	Qt-3, Qt-2, Qt-1, Qt, Pt-1, Pt	Qt+1

Table 2 Month-july performance criteria

Model name	R		MSE		RMSE	
Kaaug4	ANN	MLR	ANN	MLR	ANN	MLR
	0.9715	0.9022	0.0043	0.0031	0.0656	0.0560

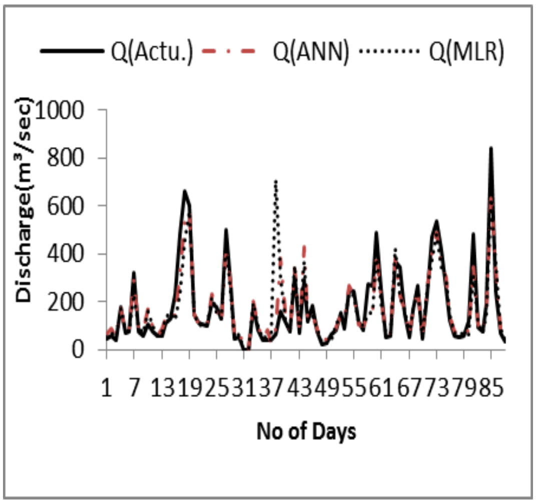

Figure 3 Hydrograph for Kaaug4 Model to compare ANN and MLR

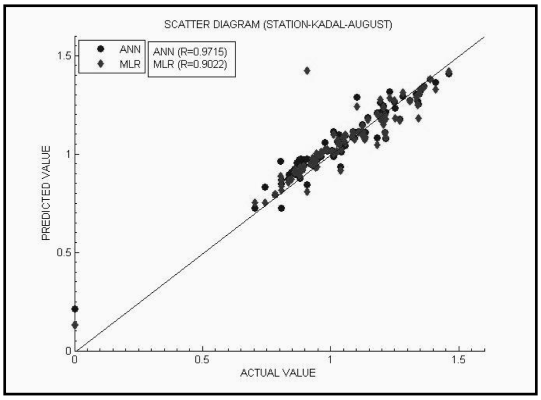

Figure 4 Scatter Plot forKaaug4 Model to compare ANN and MLR

During the model preparation of month of August,9 model were used among which model no 4 with input [Qt-1, Qt, Pt]& Output [Q(t+1)] was selected the best model by considering the performance criteria i.e. Regression (R)=0.9715 for ANN and R=0.9022 for MLR. The maximum peak point 836.91m³/sec was predicted by ANN as 652.13m³/sec and MLR as 632.02m³/sec. The minimum peak point 0.00m³/sec was predicted by ANN as 1.64m³/sec and MLR as 0.79m³/sec.The ANN value had reached very close to the highest peak value as compared to MLR, which shows ANN is better than MLR.

Table 3 The Correlation Coefficiant for all the selected models using ANN and MLR methods.

Sr. No.	Model Name	R(ANN)	R(MLR)
1	Ajajune6	0.9349	0.8413
2	Ajajuly7	0.9641	0.9352
3	Ajaaug9	0.9689	0.9276
4	Ajasep8	0.9676	0.8190
5	Ajaoct4	0.9891	0.9275
6	Kajune4	0.9584	0.8368
7	Kajuly6	0.9737	0.9162
8	Kaaug4	0.9715	0.9022
9	Kasep7	0.9768	0.8801
10	Kaoct8	0.9720	0.9493
11	Tajune7	0.9393	0.8281
12	Tajuly6	0.9675	0.9229
13	Taaug5	0.9689	0.9196
14	Tasep6	0.9651	0.9097
15	Taoct3	0.9814	0.9293
16	Jamjune1	0.9692	0.8202
17	Jamjuly4	0.9475	0.9069
18	Jamaug7	0.9734	0.9129
19	Jamsep4	0.9798	0.8905
20	Jamoct4	0.9782	0.9132
21	Najune3	0.9731	0.8680
22	Najuly8	0.9768	0.9333
23	Naaug6	0.9730	0.8411
24	Nasep5	0.9849	0.9201
25	Naoct4	0.9865	0.9396
26	Jajune10	0.9092	0.8168
27	Jajuly10	0.9458	0.9114
28	Jaaug10	0.9243	0.9079
29	Jasep10	0.9649	0.8928
30	Jaoct10	0.9598	0.9182
31	Najune10	0.9252	0.8706
32	Najuly10	0.9679	0.9487
33	Naaug10	0.9456	0.9168
34	Nasep10	0.9403	0.9295
35	Naoct10	0.9741	0.9393

V. RESULT AND CONCLUSION

The above developed models for monsoon month June, July, August, September and Octomber were developed from whch last ten models were tested using evaporation as an additional parameter to the rainfall and runoff,it is being observed that models developed by the combination of rainfall and runoff performed better than models of rainfall, runoff and evaporation parameter. The models were also tested for their performance using Hyderograph and Scatter plot from which it is concluded that ANN perform better than MLR model.

VI. ACKNOWLEDGEMENT

I am also grateful to Hydrology Project, Nashik for providing the necessary data of Ghataprabha Sub-basin, which made this work possible. I would like to thank Kolhapur Irrigation Department for providing the valuable documents of Ghataprabha basin.

REFERENCES

[1] M.H.Diskin(1970), “Definition and Uses of the Linear Regression Model”. Water Resource Research, Vol 6,No 6,1668-1673.

[2] Robert Hirsch (1979), “An Evaluation of Some Record Reconstruction Techniques”. Water Resources Reasearch, Vol 15 No 6,1781-1790.

[3] Robert Hirsch, Edward J.Gilroy(1984), “Methods of fitting a straight line to data: Examples in Water Resources”. Water Resource Bulletin, Vol.20,No.5,705-711.

[4] Keith Loague, Allan Freeze (1985), “A Comparison of Rainfall-Runoff Modeling Techniques on Small Upland Catchments”.Water Resource Research, Vol 21,No.2, 229-248.

[5] Driver N.E. and Troutman B.M. (1989), “Regression model for estimating urban storm runoff quality and quantity in the United States”. Journal of Hydrology, 109 (1989) 221-236.

[6] F.H.S. Chew, M.J. Stewardson, T.A. McMahon(1993), “Comparison of six rainfall runoff modelling approaches”. Journal of hydrology 147 (1993),1-36.

[7] Donna M Rizzo, David E. Dougherty (1994), “Characterization of aquifer properties using artificial neural networks: Neural kriging”. Water Resource research, Vol 30, No 2,483-497.

[8] H. Raman and N. Sunilkumar (1995), “Multivariate modelling of water resources time series using artificial neural networks”. Hydrolog. Sci. J,145-163.

[9] Jason Smith, Robert Eli (1995), “Neural Network Models of Rainfall-Runoff Process”. J. Water Resour. Plann. Manage. 1995.121:499-508.

[10] D. Achela K. Fernando, A. W. Jayawardena,(1998), “ Runoff forecasting using RBF networks with OLS algorithm”, J. Hydrol. Eng. 1998.3:203-209.

[11] Tokar, A. and Johnson, P. (1999). "Rainfall-Runoff Modeling Using Artificial Neural Networks." J. Hydrol. Eng.1999.4, 232-239.

[12] Kuok King Kuok, Nabil Bessaih(2000). “Artificial Neural Networks (ANNS) for daily rainfall runoff modelling.” The Institution of Engineers, Malaysia.

[13] Tokar, A. and Markus, M. (2000). "Precipitation-Runoff Modeling Using Artificial Neural Networks and Conceptual Models." J. Hydrol. Eng., 2000.5, 156-161.

[14] Jagadeesh Anmala, Bin Zhang and Rao Govinda Raju(2000), “Comparison of ANN and Empirical approaches for predicting watershed runoff”. J. Water Resour. Plann. Manage. 2000.126:156-166.

[15] Meuser Valenca, Teresa Ludermir (2000), “Monthly streamflow forecasting using an Neural Fuzzy Network Model”. 1522-4899/00 $10.00 0 2000 IEEE, 117-119.

[16] Sobri Harun, Nor Irwin Ahmat Nor (2002). “Artificial Neural Network Model for Rainfall–Runoff Relationship” Jurnal Teknologi 37(1) December 2002.

[17] Atsu S.S. Dorvlo, Joseph A. Jervase(2002), “Solar radiation estimation using artificial neural networks”. Applied Energy 71,307–319.

[18] S. Birikundavyi, R. Labib (2002), “Performance of Neural Networks in Daily Streamflow Forecasting”. J. Hydrol. Eng. 2002.7:392-398.

[19] K.P.Sudheer, A.K.Gosain (2002), “Modelling evaporation using an artificial neural network algorithm”. Hydrol. Process. 16, 3189–3202 (2002).

[20] Shie-Yu Liong and Chandrasekaran Sivapragasam (2002), “Flood Stage Forecasting With Support Vector Machines”. Journal of the American Water Resource Association, Vol. 38, no. 1,173-186.

[21] M.P. Rajurkar, U.C. Kothyari, U.C. Chaube (2003), “Modelling of the daily rainfall-runoff relationship with artificial neural network”.J.Hydrol.(2004),96-113.

[22] S.Riad, J.Mania (2003), “Rainfall-Runoff Model Using an Artificial Neural Network Approach”. Mathematical and Computer Modelling (40), 839-846.

[23] Bunchingiv Bazartseren, Gerald Hildebrandt (2003), “Short-term water level prediction using neural networks andneuro-fuzzy approach”. Neurocomputing 55 (2003) 439 – 450.

[24] A Jain, Indurthy(2003), “Comparative Analysis of Event-based Rainfall-runoff Modeling Techniques—Deterministic, Statistical, and Artificial Neural Networks”. J. Hydrol. Eng. 2003.8:93-98.

[25] Ashu Jain, Sanaga Srinivasulu (2004), “Deveopment of effective and efficient rainfall runoff models using integration of deterministic, real coded GA, and ANN techniques”. Water Resource Research 40(4), W04302.

[26] B.Bhattacharya, D.P. Solomatine (2004), “Neural networks and M5 model trees in modelling water level–discharge relationship”. Neurocomputing 63 (2005) 381–396.

[27] P.C. Nayak, K.P. Sudheer, D.M.Rangan (2004), “A neuro-fuzzy computing technique for modelling hydrological time series”. Journal of Hydrology 291 (2004) 52–66.

[28] F.Anctil and N.Lauzon(2004), “Generalisation for neural networks through data sampling and training procedures, with applications to streamflow prediction”. Hydrology and Earth system Sciences, 8(5), 940-958

◈◈◈