ANN-based sediment yield models for Vamsadhara river basin (India)

Avinash Agarwal*, RD Singh, SK Mishra, and PK Bhunya
Western Himalayan Regional Centre, National Institute of Hydrology, Jammu Cantt., Jammu, J & K, India

Abstract

Most universally accepted feed-forward error back-propagation artificial neural network models, supported by batch- and pattern-learning, daily, weekly, ten-daily and monthly sediment yield were developed for the Vamsadhara River basin of India. The fast gradient descent optimisation technique improved with variable learning rate (α) and momentum term (β) was used for optimisation. In the process of optimisation and updating of weights, criteria adopted to terminate the process of learning was selected as a per-decided high number of iteration and the other is the generalisation of model through cross-validation. In all cases of model formulation, the data were normalised with the maximum value of the variable of the series individually. The pattern-learned models were found superior to batch-learned models. High numbers of iterations adopted for model development were found to reduce the value of the objective function, but with model's over-learning and that is reflected? Unclear what is meant by an increase and decrease of the performance in calibration and cross-validation, respectively. The generalised pattern- learned models for different time scales were compared with linear transfer function models and it was found that the pattern-learned models developed with generalisation through cross-validation were superior in general, except weekly for the study area.

Keywords: back propagation artificial neural network, sediment yield modelling, generalised modelling.

Introduction

Since the 1930s, numerous linear and non-linear hydrological models have been developed to simulate and forecast various hydrological processes and variables. The suggested models have continuously been improved by introducing new tools to effectively simulate the processes. The developed models in the area of hydrology can broadly be classified into regression models, stochastic models, conceptual or parametric models and system models.

Regression models are either regression or correlation based, and correlate the input(s) and output(s) of a process in the form of linear or non-linear relationship, which estimate the constants of regression models. Few universally acceptable hydrological models that have been developed in this category are the USLE, MUSLE and Elwell models. The stochastic models normally extract the statistical properties of time series and propagate these properties during prediction. Such models normally require a long time series and their extrapolation properties are poor. Some of the widely used stochastic models in hydrological studies are auto-regressive, auto-regressive moving average, auto-regressive integrated moving average, seasonal auto-regressive integrated moving average, etc. The conceptual models are designed to approximate within their structure the general internal physical subprocesses. The models usually incorporate simplified forms of physical laws or a series of physical laws to represent the transformation of input to output. These laws are generally linear or non-linear, time variant or time invariant, lumped or distributed, casual or non-casual and dynamic or memory less. Among the most widely used conceptual models in the field of hydrology are the Sacramento Soil Moisture Accounting (SAS-

SMA) model of the U.S. National Weather Service, HEC of U.S. Army Corps of Engineers and the Stanford Watershed Model and System Hydrologic European (SHE) model.

Artificial neural network (ANN) is a new soft computing technique composed of densely interconnected processing nodes which has the ability to extract and store the information from the few patterns (data) in training through learning. ANN architecture parallels in processing with that designed to process the information in neuro-computing (Vemuri, 1992). The model is easy to develop; yields satisfactory results when applied to complex systems poorly defined or implicitly understood; and more tolerant to variable, incomplete or ambiguous input data. Hydrologic applications of ANN include the modelling of daily rainfall-runoff-sediment yield process, snow-rainfall process, assessment of stream's ecological and hydrological responses to climate change, rainfall-runoff forecasting, ground water quality prediction and ground water remediation. ASCE (2000a;b), Jagadeesh (2000), Tokar (2000), Rajurkar (2002) among others provided a good overview of the ANN application to rainfall-runoff simulation and prediction. Imrie (2000) improved the generalisation by adding a guidance system to the cascade correlation learning architecture and extrapolation properties using an activation function. Wilby (2003) was able to interpret the internal behaviour of an ANN-based rainfall-runoff model. To this end, he deleted all the nodes other than the hidden nodes and compared with the state variables and internal fluxes. Danh (1999) and Elshorbagy (2000) proposed feed-forward error back- propagation artificial neural network (BPANN) models for runoff forecasting using fixed stopping criterion and independent variables, respectively, and compared them for performance with the available conceptual models. The works of Thirumalaiah (2000), Xu (2002), Birikundavyi (2002), Shivakumar (2002), Cigizoglu (2003), and Xiong (2002) among others are notable for real time forecasting of runoff. Other ANN applications include derivation of unit hydrograph (Lange, 1998). Only a few studies (Tayfur, 2002; Nagy, 2002; Cigizoglu

2004) focused on ANN-event based sediment yield modelling and sediment concentration.

The fast gradient descent optimisation is probably the most widely used supervised algorithm applied in multilayer feed-forward artificial neural network to optimise the error function and therefore it is adopted in the present study. In model building, the convergence of error is normally linked with a pre-decided tolerance value, such as a minimum error, minimum error gradient or a high level of iteration. The use of a pre-decided tolerance value to stop convergence may under- or over-learn the model (Fu, 1996). The cross-validation which improves the generalisation of network and obviates the under- or over-learning, however, requires intensive computations and demands large data set and computer memory. Therefore, it is in order to develop the generalised batch- and pattern-learned BPANN-based sediment yield models considering high level of iteration and cross-validation as criteria to terminate the process of learning, which forms the objective of this paper. These models are simultaneously subjected to network pruning to achieve parsimony, and compared with linear transfer function (LTF) models using daily, weekly, ten-daily, and monthly data of Vamsadhara River basin in India.

Linear transfer function model

The linear transfer functions (LTFs) are the time-dependent regression models with simple mathematics, requiring minimum input, little computation, and yielding the results of desired accuracy (Johnston, 1972). For a lumped linear system, the two or more time-dependent observations are linked as:

$$S_t = \sum_{j=1}^{p} a_j \cdot S_{t-j} + \sum_{j=1}^{q} b_j \cdot R_{t-j+1} + \sum_{j=1}^{r} c_j \cdot Q_{t-j+1} \qquad (1)$$

where:
  S_t is a dependent observation
  R_t and Q_t are independent observations
  p, q and r are the time responses
  a_j, b_j and c_j are parameters associated with the jth variable.

The least square method can be used to solve the set of t linear equations for parameters. Representing $[a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q, c_1, c_2, \ldots, c_r]^T$ as $\hat{H}$, the variance of the time response $\{var.(\hat{H})\}$ is given as (Johnston, 1972):

$$Var(\hat{H}) = [A^T A]^{-1} \sigma^2 \qquad (2)$$

where:
  A is the input matrix and σ^2 is the variance of error term (e_t) expressed as:

$$\sigma^2 = (t - p - q - r)^{-1} \sum_{j=1}^{t} e_j^2 \qquad (3)$$

The $Var(\hat{H})$ helps avoid the parsimonious selection of time responses by comparing the respective standard error associated with each parameter as follows:

$$\hat{H} \geq \sqrt{Var(\hat{H})} \qquad (4)$$

For parsimonious selection of time responses, the initially selected time response value is increased one by one and if the resulting $\hat{H}$ is less than the respective standard error $\{\sqrt{Var(\hat{H})}\}$, the parameter value is decreased. The obtained time response value represents the number of successive past events of the variable affecting output (Johnston, 1972).

Artificial neural network

In a multilayer back propagation artificial neural network (BPANN), the nodes of input layer receive the input data, process it and pass the output to the nodes of subsequent hidden layer(s), and from last hidden layer to the output layer. The structure of the system with nodes in input, hidden and output layers can be represented by j, i, and k, where j varies from 1 to j, i from 1 to i, and k from 1 to k. A particular structure for j=2, i=4 and k=1 is shown in Fig. 1.
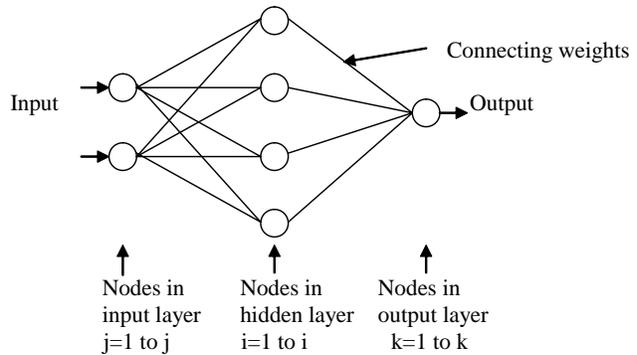
Figure 1
Structure (2, 4,1) and notations of a multilayer BPANN

In a feed-forward BPANN scheme, nodes of the input layer receive the normalised data set (input). The weighted sum corresponding to each node of the next layer is calculated and passed to the next layer usually through a sigmoid activation function. The error (E) calculated at the output is propagated back to hidden layer(s) and finally to input layer by updating the weights of interconnection. The error (E) is defined as:

$$E = \frac{1}{2} \sum_{1}^{k} [d_{(k)} - O_{(k)}]^2 \qquad (5)$$

where:
  $d_{(k)}$ is the observed output at the kth node of the output layer
  $O_{(k)}$ is the estimated output at the kth node of the output layer

The updating of weights in all iterations is carried out using the following equation:

$$W_{(ij)\,n+1} = W_{(ij)\,n} + \Delta W_{(ij)\,n} \qquad (6)$$

The speed of convergence is increased normally by introducing a momentum term β and the effect of previous weight change as:

$$W_{(ij)\,n+1} = W_{(ij)\,n} + \Delta W_{(ij)\,n} + \beta[W_{(ij)\,n} - W_{(ij)\,n-1}] \qquad (7)$$

The change in weights (ΔW) in the direction of negative gradient is given by:

$$\Delta W_{(ij)} = -\alpha \frac{\partial E}{\partial W_{(ij)}} \qquad (8)$$

where:
  α is the learning rate such that $0 < \alpha < 1$ and governs the rate of change of weights.

The model parsimony can be achieved through an A information criterion (AIC) (Akaike, 1974), B information criterion (BIC) (Rissanen, 1978) or by the pruning of network (Karnin, 1990). AIC

Available on website **http://www.wrc.org.za**

Figure 2
Index map of Vamsadhara River basin showing hydrological details
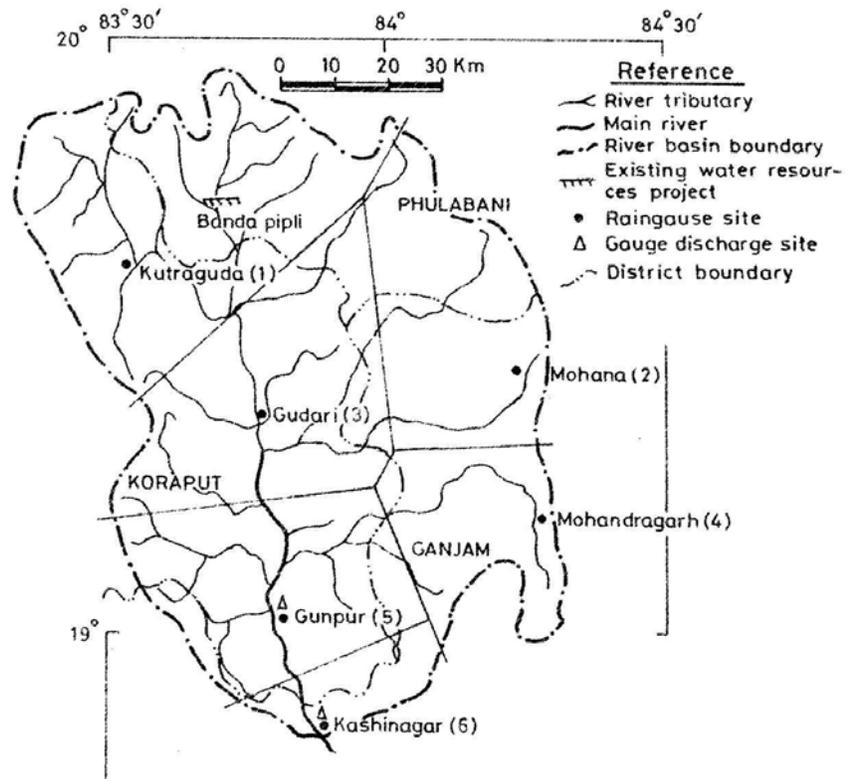
and BIC criteria utilises root mean square error (RMSE) statistic which is penalised for having more numbers of free parameters for deciding the number of free parameters (Xu, 2002). In pruning, Karnin (1990) suggested to use the sensitivity of error $Se_{(ij)}$ with respect to weight $W_{(ij)}$ for elimination of the respective weight without excessive calculations. $Se_{(ij)}$ is defined as:

$$Se_{(ij)} = -\frac{\partial E}{\partial W_{(ij)}} \int_{W_{(ij)}^{f}}^{W_{(ij)}^{i}=0} W_{(ij)}^{f} \qquad (9)$$

Finally the sensitivity of the error with respect to weight reduces to:

$$Se_{(ij)} = \sum_{1}^{it} \frac{\Delta W_{(ij)}^{2}}{\alpha} \cdot \frac{W^{f}_{(ij)}}{W^{f}_{(ij)} - W^{i}_{(ij)}} \qquad (10)$$

where:

i and f indicate the initial and final values of weights

Since the weight update is available for each iteration during learning, the determination of the summation of squared of change in weight $\Delta W_{(ij)}^{2}$ to estimate the sensitivity of error function is the only extra computation required. A low sensitivity suggests the respective weight to be in sensitive and pruning of the corresponding node?unclear. Suggest: "insensitive and pruning of the corresponding node is recommended". The generalisation of model can be checked using the available statistical evaluation criteria: root mean square error (RMSE), correlation coefficient (CC), and coefficient of efficiency (CE) (Nash and Sutcliffe, 1970). Here, it is noted that in cross-validation, the model is trained on training data set and for every iteration the model is simultaneously verified on another data set through statistical evaluation criteria. The training is continued until the performance evaluation criteria show improvement in both training and verification. The method monitors the generalised performance and stops the process of learning when there is no more

improvement in training and its performance in first verification. Since in this method the performance of the developed model is checked in all iterations, the level of accuracy is not fixed in the beginning of model formulation.

Study area and data

The selected Vamsadhara River basin up to Kashinagar (area = 7 820 km²) (Fig. 2) is located between Mahanadi and Godavari River basins of India. The basin is narrow and highly undulated. The daily rainfall data (mm), runoff (m³/s), and suspended sediment yield (kg/s) of the active period (June 1 to October 31) for years 1984 to 89 and 1992 to 95 were available. The weighted average rainfall is computed with the rainfall values of 1 to 6 rain-gauge stations (Fig. 2) with Theissen weights as 0.2640, 0.1835, 0.2696, 0.1096, 0.1509 and 0.0224, respectively.

The collected daily rainfall (mm), runoff (m³/s) and sediment yield (kg/s) of the active period for all the years were subjected to pre-analysis and formulation of the database for different time units, i.e. weekly, ten-daily and monthly. The yearly weighted average rainfall, runoff, sediment yield, runoff-rainfall ratios and the sediment-runoff ratios (sediment concentration) for the active monsoon period is shown in Table 1.

It can be seen that the runoff-rainfall ratio in the first four years varied in between 0.10 to 0.23 and for subsequent six years it varied in between 0.20 to 0.43. It indicates that the catchment behaviour to infiltration and other losses has changed and runoff corresponding to rainfall has increased in years 1988 to 1995. The weighted average rainfall and runoff indicated no specific trend and correlation between them.

The sediment concentration was found to be decreasing with time (Table 1). The change in sediment concentration indicates that the catchment is improving with time in respect of soil conservation. The main reason for this could be the regular *Jhum* cultivation practices being adopted by the tribes of these areas. On the other hand, it also reflected that the data are not exactly from one

TABLE 1 Summary of annual hydrological data of Vamsadhara River basin along with statistical summary of data						
Year	Weighted average rainfall	Runoff	Sediment yield	Runoff	Runoff-rainfall ratio	Sediment yield runoff ratio (sediment concentration)
	(mm)	(m³/s)	(kg/s)	(mm)		(kg/m³)
1984	965.2	13 245.2	28 917.2	146.3	0.15	2.18
1985	1 079.6	17 195.9	29 802.3	190.0	0.18	1.73
1986	971.7	20 519.6	36 233.0	226.7	0.23	1.77
1987	715.8	6 506.6	12 230.4	71.9	0.10	1.88
1988	1 053.7	25 454.9	42 742.8	281.2	0.27	1.68
1989	1 074.2	19 789.9	30 699.6	218.7	0.20	1.55
1992	1 028.7	40 017.2	62 668.2	442.1	0.43	1.57
1993	773.3	15 005.5	20 177.2	165.8	0.21	1.34
1994	1 142.3	31 896.4	33 675.8	352.4	0.31	1.06
1995	1 010.6	34 237.4	30 291.5	378.3	0.37	0.88
Maximum	1 142.3	40 017.2	62 668.2	442.1	0.43	2.18
Minimum	715.8	6506.6	12 230.4	71.9	0.10	0.88
Mean	981.5	22 386.9	32 743.8	247.3	0.25	1.56
Standard deviation	136.1	10 427.1	13 419.3	115.2	0.10	0.39
Skewness	-1.2	0.3	1.0	0.3	0.57	-0.44

TABLE 2 Performance evaluation of LTF sediment yield models for Vamsadhara River basin for different time scales									
Model and structure	Calibration period, (1984 to 1987)			Verification period, (1988 to 1989)			Verification period, (1992 to 1995)		
	RMSE (Abs.)	CC (%)	CE (%)	RMSE (Abs.)	CC (%)	CE (%)	RMSE (Abs.)	CC (%)	CE (%)
Daily (2,1,1)	199	87.1	71.2	228	78.8	53.2	308	85.0	57.4
Weekly (1,1,0)	766	90.8	77.7	1018	74.2	45.3	1433	87.4	63.6
Ten-daily (1,1,0)	833	93.5	83.0	1315	76.9	54.4	1828	87.1	63.8
Monthly (1,1,0)	1404	94.0	85.9	2800	68.6	45.3	3988	83.4	51.3

homogenous population.

Since the sediment concentration is found to be fast decreasing over the years, a sediment model developed with the data of 1984, 1985, 1986, 1987 and verified on years 1988, 1989, 1992, 1993 and 1994 to 1995 may therefore over-estimate for these years.

The daily data of the active period of monsoon (June 1st to October 31st) of water years were used to determine weekly, ten-daily and monthly data. The weekly data consisted of weekly summation starting from June 1st to October 31st. The last six days of October were not used in the weekly data. The ten-daily data were the summation of days 1 to 10, 11 to 20 and 21 to 30 or 29/31 days of each month. The monthly data were the summation of data of all the days of the month. Thus the data for four different time bases (daily, weekly, ten-daily and monthly) comprise 153 d or 21 weeks or 15 ten-dailies or 5 months and were used for the development of respective simulation and forecasting models.

Model development

The development of daily, ten-daily, and monthly rainfall-runoff-sediment yield (RQS) models based on the above described LTF and multilayer BPANN approaches utilises the above data of 1984 to 87, and their generalisation/verification on the data of 1988 to 89 and 1992 to 95, as follows.

LTF models

Linear transfer function (LTF) models were developed with an initial selection of time response for the input variables as twice time units. The response of the respective variable was increased by one, if estimated $\hat{H}$ was greater than the respective value of the standard error of the variance of $\hat{H}$ as $\{ \sqrt{Var(\hat{H})} \}$, otherwise the time

TABLE 3 Performance evaluation of BPANN sediment yield models for Vamsadhara River basin for different time scales													
Model and structure		No. of iteration	Value of objective function	Calibration period, (1984 to 1987)			Cross-validation period (1988 to 1989)			Verification period, (1992 to 1995)		
				RMSE (Abs.)	CC (%)	CE (%)	RMSE (Abs.)	CC (%)	CE (%)	RMSE (Abs.)	CC (%)	CE (%)
Batch-learning process												
1. Generalisation with cross validation												
Daily	(4,8,1)	1 707	0.032	190	86.8	74.8	310	72.5	15.9	463	70.0	3.9
Weekly	(2,5,1)	501	0.067	760	90.7	78.2	1 300	75.0	11.9	2105	68.6	21.5
Ten-daily	(2,4,1)	123	0.049	840	91.9	82.6	1 300	96.4	54.1	2459	65.1	34.5
Monthly	(2,4,1)	1 480	0.022	1 100	95.7	90.8	1 600	91.2	82.9	3923	72.9	52.9
2. Generalisation with high level of iteration												
Daily	(4,8,1)	5 000	0.032	250	82.3	56.5	330	75.2	6.1	492	68.6	-8.2
Weekly	(2,5,1)	5 000	0.051	1 000	87.2	61.3	1 600	69.8	-37.6	1990	19.4	29.8
Ten-daily	(2,4,1)	5 000	0.025	900	93.4	80.1	1 500	78.9	40.0	2048	75.7	54.6
Monthly	(2,4,1)	5 000	0.014	1 500	95.7	83.4	2 000	91.6	72.7	4065	72.5	49.4
Pattern-learning process												
1. Generalisation with cross validation												
Daily	(4,8,1)	19	0.094	200	86.7	72.2	200	79.3	62.8	267	83.2	68.0
Weekly	(2,5,1)	21	0.150	760	91.6	78.1	820	80.2	64.1	1648	75.1	51.8
Ten-daily	(2,4,1)	135	0.080	830	91.6	83.3	1 100	82.7	68.2	2342	65.0	40.6
Monthly	(2,4,1)	2 260	0.019	1 100	96.0	92.0	1 700	89.4	79.1	3889	74.1	53.7
2. Generalisation with high level of iteration												
Daily	(4,8,1)	5 000	0.041	130	93.6	87.4	210	83.2	59.5	380	69.0	35.2
Weekly	(2,5,1)	5 000	0.106	620	92.7	85.6	1 000	77.3	46.6	1598	76.0	54.7
Ten-daily	(2,4,1)	5 000	0.039	630	95.2	90.2	1 200	80.9	63.1	2026	75.0	35.5
Monthly	(2,4,1)	5 000	0.024	1 000	96.2	92.5	1 800	88.8	77.3	3931	73.0	52.7

response was decreased. The model was re-developed after each change in time response and finally the chosen LTF models satisfying the criterion were reported through Eqs. (11) to (14).

$$S_t = 8.629*R_t + 3.636*R_{t-1} + 1.617*Q_t + 0.053*S_{t-1} \qquad \text{— Daily} \qquad (11)$$
$$S_t = 12.800*R_t + 1.371*Q_t \qquad \text{— Weekly} \qquad (12)$$
$$S_t = 13.240*R_t + 1.265*Q_t \qquad \text{— Ten-daily} \qquad (13)$$
$$S_t = 12.280*R_t + 1.158*Q_t \qquad \text{— Monthly} \qquad (14)$$

Here, it can be observed that the sediment yield from the daily model depends on its successive past sediment yield value, whereas for weekly, ten-daily and monthly models, it is independent of their past values. The performance evaluation results of all the developed simulation modes are given in Table 2. It can be seen that the model performance in calibration generally improves with an increase in the time scale, but it does not hold in verification. The performance in second verification period (1992 to 95) is better than the first verification period (1988 to 89).

ANN models

In BPANN the same number of input variables was taken as those obtained while developing LTF models. All the input-output pairs of data set were first normalised considering the maximum value of the series and thus reducing the individual variables in the range of 0 to 1 to avoid any saturation effect that may arise from the use of sigmoid activation function. In both batch- and pattern-learning, the initial values of interconnecting weight were randomly selected between –0.5 and +0.5 (Dawson and Wilby, 1998) against the recommended ranges of (–1.0, +1.0) or (–0.5, +0.5) or (–0.1, +0.1) (Lorrai and Sechi, 1995; Dawson and Wilby, 1998). The sigmoid activation function was considered in model development. The values for both learning rate (α) and momentum term (β) were initially considered as 0.5, which, however, decreased in successive iterations. All the interconnecting weights were updated using the error of input-output pairs. In batch-learning, these weights were adjusted only after processing all data sets for error, and the data set indicating the highest error was used for leaning? and updating of weights. This type of learning is fast, but requires more computation for it is governed by the data set indicating the highest error (Minns and Hall, 1996). The pattern-learning is governed by the error of each data set; the interconnecting weights were simultaneously adjusted. The processing in pattern-learning was slow for it continuously improved the weights for each data set.

The number of input nodes in the input layer was taken equal to the number of input variables. Since no guideline is yet available on the number of hidden nodes in the hidden layer(s) (Vemuri, 1992), these were initially taken equal to twice of input nodes (Hipel et al., 1994), and increased one at a time considering the improved generalisation and the above pruning criteria. However, corresponding to one output, only one node was taken in the output layer. Thus,

a three-layer network structure with varying numbers of hidden nodes in the hidden layer was tried, and the finally selected ANN structure along with the performance of developed daily, weekly, ten-daily, and monthly models for both pattern- and batch-learning is listed in Table 3.

Table 3 shows a comparative performance of pattern and batch-learned daily, weekly, ten-daily, and monthly BPANN models for generalisation:

- With maximum number of iterations restricted to 5 000
- Using cross-validation (by least errors both in calibration and cross validation).

Apparently the batch-learned models for both the cases of generalisation yield a relatively low value of the objective function compared to the similar pattern-learned models, but perform poorly in both cross-validation and verification. It suggests avoiding the use of fast optimisation methods in model generalisation with such data. Furthermore, in calibration, the performance of pattern-learned BPANN models generalised with cross-validation is generally superior to those generalised using high level of iteration. It suggests preference of best-fit models to ANN models requiring high numbers of iterations, for it over-learns and memorises the data, and performs poorly in cross-validation and simultaneously in verification (French et al., 1992; Hus et al., 1995). The generalised pattern-learned sediment yield models are superior in both calibration and verification based on RMSE, CC and CE criteria (Table 3) and, therefore, are preferred to the other models based on the batch-learning and LTF concept. Similar to the model based on the LTF concept, the performance of the best pattern-learned BPANN model also generally improves with increase in time scale in calibration only.

Conclusion

Based on the selected performance evaluation criteria, viz., RMSE, CC, and CE, the developed BPANN and LTF sediment yield models for the Vamsadhara River basin exhibit an improvement with increase in time scale in model calibration. The pattern-learned BPANN models perform better than batch-learned models irrespective of their high convergence. On the other hand, the pattern-learned BPANN models generalised with cross-validation perform better than those generalised with a high level of iteration and LTF models. This study suggests that a fast and high convergence is not essential in generalised model development.

References

AKAIKE H (1974) A new look at the statistical model identification. *IEEE Trans. Autom. Control*. **AC-19** 716-723.

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology (2000a) Artificial neural networks in hydrology, I: Preliminary concepts. *J. Hydrol. Eng*. *ASCE* **5** (2) 124-137.

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology (2000b) Artificial neural networks in hydrology, II: Hydrologic application. *J. Hydrol. Eng. ASCE* **5** (2) 115-123.

BIRIKUNDAVYI S, LABIB R, TRUN T and ROUSSELLE J (2002) Performance of neural networks in daily stream flow forecasting. *J. Hydrol. Eng.* **7** (5) 392-398.

CIGIZOGLU HK (2003) Estimation, forecasting and extrapolation of river flows by artificial neural networks. *Hydrol. Sci. J.* **48** (3) 349-361.

CIGIZOGLU HK (2004) Estimation and forecasting of daily suspended sediment data by multi layer perceptrons. *Advances Water Res.* **27** 185-195.

DANH NT, PHIEN HN and GUPTA AD (1999) Neural network models for river flow forecasting. *Water SA* **25** (1) 33-39.

DAWSON CW and WILBY R (1998) An artificial neural network approach to rainfall-runoff modelling. *Hydro. Sci. J.* **43** (1) 47-66.

ELSHORBAGY A, SIMONOVIC SP and PANU US (2000) Performance evaluation of artificial neural networks for runoff prediction. *J. Hydrol. Eng.* **5** (4) 424-427.

FRENCH MN, KRAJESWSKI WF and CUYKENDALL RR (1992) Rainfall forecasting in space and time using a neural network. *J. Hydrol.* **137** 1-31.

FU LIMIN (1996) *Neural Networks in Computer Intelligence*. McGraw-Hill, Inc., Newyork.

HIPEL KW, IAN A, PANN US and SINGH VP (1994) *Stochastic And Statistical Methods in Hydrology and Environmental Engineering*. Vol. **3**. *The Series Analysis in Hydrology and Environmental Engineering*. The Netherlands. Kluwer Academic Publishers.

HSU, KUO-LIN, GUPTA HV and SOROOSH SOROOSHIAN (1995) Artificial neural network modelling of the rainfall-runoff process. *Water Resour. Res.* **31** (10) 2517-2530.

IMRIE CE, DURUCAN S and KORRE A (2000) River flow prediction using artificial neural networks: generalization beyond the calibration range. *J. Hydrol.* **233** 138-153.

JAGADEESH A, ZHANG B and GOVINDARAJU RS (2000) Comparison of ANNs and empirical approaches for predicting watershed runoff. *J. Water Resour. Plann. Manag.* **126** (3) 156-166.

JOHNSTON J (1972) *Econometric Methods*. McGraw Hill, Tokyo.

KARNIN ED (1990) A simple procedure for pruning back propagation trained neural networks. *Inst. Elect. Electron. Eng. Trans. Neural Networks* **1** (20) 239-242.

LANGE N (1998) Advantage of unit hydrograph derivation by neural networks. In: Babovie V and Larsen CL (eds.) *Hydroinformatics*. *Proc. 3rd Int. Conf. on Hydroinformatics*. Copenhagen, Denmark. **2** 783-789 AA Balkema, Rotterdam, Netherlands.

LORRAI M and SECHI GM (1995) Neural nets for modelling rainfall-runoff transformations. *Water Resour. Manage.* **9** 299-313.

MINNS AW and HALL MJ (1996) Artificial neural networks as rainfall-runoff models. *Hydrol. Sci. J.* **41** (3) 399-417.

NAGY HM, WATANABE KAND and HIRANO M (2002) Prediction of sediment load concentration in rivers using artificial neural network model. *J. Hydraul. Eng.* **128** (6) 588-595.

NASH JE and SUTELIFFE JV (1970) River flow forecasting through conceptual models. *J. Hydrol.* **10** 282-290.

RAJURKAR MP, KOTHYARI UC and CHAUBE UC (2002) Artificial neural networks for daily rainfall-runoff modelling. *Hydrol. Sci. J.* **47** (6) 865-877.

RISSANEN J (1978) Modelling by short data description. *Automation* **14** 465-471.

SHIVAKUMAR B, JAYAWARDENA AW and FERNANDO TMKG (2002) River flow forecasting: use of phase-space reconstruction and artificial neural networks approaches. *J. Hydrol.* **256** 225-245.

TAYFUR, GOKMEN (2002) Artificial neural networks for sheet sediment transport. *Hydrol. Sci. J.* **47** (6) 879-892.

THIRUMALAIAH T and MAKARAND C DEO (2000) Hydrological forecasting using neural networks. *J. Hydrol. Eng.* **5** (2) 180-189.

TOKAR AS and MARKUS M (2000) Precipitation-runoff modelling using artificial neural networks and conceptual models. *J. Hydrol. Eng.* **5** (2) 256-161.

VEMURI V RAO (1992) Artificial neural networks: Concepts and control application. Los Alamitos, California. *IEEE Computer Society Press*. 509 pp.

WILBY RL, ABRAHART RJ and DAWSON CW (2003) Detection of conceptual model rainfall-runoff processes inside an artificial neural network. *Hydrol. Sci. J.* **48** (2) 163-181.

XIONG L, KIERAN M and O'CONNOR (2002) Comparison of four updating models for real-time river flow forecasting. *Hydrol. Sci. J.* **47** (4) 621-639.

XU ZX and LI JY (2002) Short term inflow forecasting using an artificial neural network model. *Hydrol. Proc.* **16** 2423-2439.

Available on website **http://www.wrc.org.za**