Available online at www.sciencedirect.com
SCIENCE DIRECT®
Applied Soft Computing 6 (2006) 272–282
Applied Soft Computing
www.elsevier.com/locate/asoc

An evaluation of artificial neural network technique for the determination of infiltration model parameters

Ashu Jain *, Amit Kumar

Department of Civil Engineering, Indian Institute of Technology, Kanpur 208016, India

Received 7 June 2004; received in revised form 3 December 2004; accepted 13 December 2004

Abstract

Infiltration is a key component in the rainfall runoff models employed for runoff prediction. Conventionally, the hydrologists have relied on classical optimization techniques for obtaining the parameters of various infiltration equations. Recently, artificial neural networks (ANNs) have been proposed as efficient tools for modelling and forecasting. This paper proposes the use of ANNs for calibrating infiltration equations. The ANN consists of rainfall and runoff as the inputs and the infiltration parameters as the outputs. Classical optimization techniques were also employed to determine flow hydrographs for comparison purposes. The performances of both the approaches were evaluated using a variety of standard statistical measures in terms of their ability to predict runoff. The results obtained in this study indicate that the ANN technique can be successfully employed for the purpose of calibration of infiltration equations. The regenerated and predicted storms indicate that the ANN models performed better than the classical techniques. It has been found that the ANNs are capable of performing very well in situations of limited data availability since the differences in the performances of the ANNs trained on partial information and the ANNs trained on the complete information was only marginal and the ANN trained on partial information consisted of a more compact architecture. A wide variety of standard statistical performance evaluation measures are needed to properly evaluate the performances of various ANN models rather than relying on a few global error statistics (such as RMSE and correlation coefficient) normally employed.

Keywords: Artificial neural networks; Model calibration; Optimization; Infiltration; Horton's infiltration model; Green-Ampt infiltration model

1. Introduction

When the rain water falls on the surface of the earth, some of it gets intercepted on the obstructions such as buildings, vegetation, etc., some of it gets trapped in the depressions on the surface of the earth, some of it evaporates back into the atmosphere, some of it seeps into the soil, and the remaining portion runs off towards oceans via streams and rivers in the form of what is known as runoff. The runoff in a river is measured and represented graphically by a curve showing volumetric runoff discharge (m^3/s) passing

* Corresponding author. Tel.: +91 512 2597411; fax: +91 512 2597395.
E-mail address: ashujain@iitk.ac.in (A. Jain).

1568-4946/$ – see front matter © 2005 Published by Elsevier B.V.
doi:10.1016/j.asoc.2004.12.007

through a cross section in a river as a function of time. This type of graph between direct runoff in a river as a function of time is called a direct runoff hydrograph (DRH). The downward seepage of water through the soil surface is known as 'infiltration'. The infiltrated water percolates deep and joins what is called the groundwater. The groundwater can move horizontally for long distances over long periods of time to appear on the earth in the rivers in the form of base flow. The water can evaporate from the vegetation, soil moisture, and the surface of the earth in the form of evapotranspiration. This whole cycle of movement of water from the atmosphere to the oceans via different parts of the earth is called the rainfall-runoff process. The rainfall-runoff process is an extremely complex, non-linear, dynamic, and fragmented process that is affected by many physical factors. The involvement of many often-interrelated physiographic and climatic factors makes the rainfall-runoff process not only very complex to understand but also extremely difficult to model [1]. Researchers have devoted considerable attention in developing the mathematical models of the complex rainfall-runoff process using either deterministic/conceptual techniques or the systems theoretic techniques. The mathematical models of the rainfall-runoff process attempt to capture the characteristics of the underlying physical processes through the use of equations of mass, momentum, and energy in case of the deterministic models and their simplified forms in case of the conceptual models. The systems theoretic models do not consider the underlying physics of the rainfall-runoff process.

The mathematical models of the rainfall-runoff process are essential components in the planning, design, and operation of many water resources projects. For example, in order to plan for the distribution and allocation of the available water resources for different uses (such as drinking, irrigation, industrial, hydro-power, etc.) in a region, accurate estimates of the runoff forecasts in the area are needed. The design of major hydraulic structures such as dams, bridges etc. requires the knowledge of the rainfall-runoff process under extreme conditions. Runoff forecasts are also needed for the routine operation and management of various municipal water supply systems, and for the floods and drought management, etc. Many mathematical rainfall-runoff models of varying degree of sophistication and complexity have been proposed by various researchers and hydrologists in the past.

A key component of any mathematical model of the complex rainfall-runoff process is modelling of the infiltration process. Many models of the infiltration process are available such as Overton's model, Green-Ampt model, Horton's model, Holtan's model, and Kostikov method, etc. [2]. The Horton's and Green-Ampt's infiltration equations are the most commonly used methods, which provide estimates of the infiltration capacities as a function of time. The Horton's infiltration equation is a simplified version of the Richard's equation under simplified assumptions. The Richard's equation is the basic governing differential equation for the movement of water through unsaturated soil under unsteady conditions that is based on the laws of conservation of mass and momentum [3]. While the Horton's equation is developed from approximate solution of the Richard's equation, the Green-Ampt equation is based on a more approximate physical theory that has an exact analytical solution. Regardless of the choice for the infiltration model to be adopted in a rainfall-runoff model, the first step in its use is the determination of the parameters of the infiltration model. The infiltration parameters are normally determined through model calibration or field measurements. In using model calibration, classical non-linear optimization techniques can be adopted to determine the optimal set of infiltration parameters using known rainfall and runoff data. The performance of the rainfall-runoff models using infiltration parameters determined using classical optimization techniques are only reasonable. Recently, the soft computing techniques have become very popular especially in the last couple of decades. Artificial neural networks (ANNs) have been used as efficient tools of modelling and forecasting in all disciplines. The ANNs are inspired by the workings of a human brain, and have the capability to generalize from facts or the information presented to them. The ANNs have been used in a wide variety of areas including modelling of the complex rainfall-runoff process [4–12] but the efforts of using the ANNs for model calibration have been limited. It may be possible to improve the performance of the rainfall-runoff models by the use of infiltration parameters determined using ANN technique but it needs to be explored.

The objectives of the study presented in this paper are to (a) employ the technique of ANN to determine a set of optimal infiltration parameters using known rainfall and runoff data, (b) employ a classical optimization technique to determine the infiltration parameters using the same data set for comparison purposes, and (c) evaluate the performance of the two methodologies of determining the optimal set of infiltration parameters in terms of certain standard statistical performance evaluation criteria in predicting runoff. A major problem in using the ANNs for rainfall-runoff modelling is the involvement of a large number of input variables that represent the input layer of the ANN model. More so while using the ANNs for the calibration of an infiltration model, as there would be a large number of runoff hydrograph ordinates that need to be presented to the ANN as inputs. The large number of inputs to an ANN increases its complexity, which is not desirable and may also limit the use of the developed model. Therefore, another objective of the study presented in this paper is to evaluate the impacts of presenting partial information to the input layer of the ANN on the overall quality of the model calibration and runoff prediction. This paper begins with a brief description of the technique of ANN followed by the description of the two infiltration models employed in this study. The details of both classical optimization and ANN models developed in this study for the purpose of model calibration are provided next followed by a description of the various standard statistical performance evaluation criteria before discussing the results and making concluding remarks.

2. Artificial neural networks

An artificial neural network is a highly interconnected network of many simple processing units called neurons, which are analogous to the biological neurons in the human brain. Neurons having similar characteristics in an ANN are arranged in groups called layers. The neurons in one layer are connected to those in the adjacent layers, but not to those in the same layer. The strength of connection between the two neurons in adjacent layers is represented by what is known as a 'connection strength' or 'weight'. An ANN normally consists of three layers, an input layer, a hidden layer, and an output layer. In a feed-forward network, the weighted connections feed activations only in the forward direction from an input layer to the output layer. On the other hand, in a recurrent network additional weighted connections are used to feed previous activations back into the network. The structure of a feed-forward ANN is shown in Fig. 1. In the Fig. 1, the circles represent neurons; the lines joining the neurons represent weights; the inputs are represented by X's; Y represents the output; V_{ji} and W_{kj} represent the weights between input and hidden and hidden and output layers, respectively.

An important step in developing an ANN model is the training of its weight matrix. The weights are initialized randomly between a suitable range, and then updated using certain training mechanism. There are primarily two types of training mechanisms, supervised and unsupervised. A supervised training algorithm requires an external teacher to guide the training process. This typically involves a large number of examples (or patterns) of inputs and outputs for training. The inputs in an ANN are the

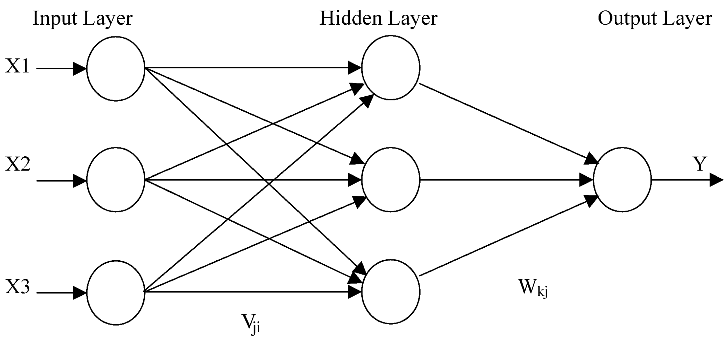

Fig. 1. Structure of a feed-forward ANN.

cause variables and outputs are the effect variables of the physical system being modelled. The primary goal of training is to minimize the error function by searching for a set of connection strengths that cause the ANN to produce outputs that are equal to or closer to the targets. A supervised training mechanism called back-propagation training algorithm [13] is normally adopted in most of the engineering applications. In the back-propagation training mechanism, the input data are presented at the input layer, the information is processed in the forward direction, and the output is calculated at the output layer. The target values are known at the output layer, so that the error can be estimated. The total error at the output layer is distributed back to the ANN and the connection weights are adjusted. This process of feed-forward mechanism and back propagation of errors and weight adjustment is repeated iteratively until convergence in terms of an acceptable level of error is achieved. This whole process is called the training of the ANN. The trained ANN is then validated on the testing data set, which it has not seen before. Once an ANN has been trained and tested, it can be used for prediction or modelling the physical system for which it is has been designed.

3. Infiltration models

Two different infiltration models have been employed in this study, Horton's infiltration model and the Green-Ampt infiltration model. The Horton's infiltration model has been employed for the determination of its optimal parameters while the Green-Ampt model was used to generate synthetic runoff data from the synthetic rainfall data. The details of these two infiltration models are provided below.

3.1. Horton's infiltration model

The Horton's equation is the most commonly used infiltration equation that provides estimates of potential infiltration rates as a function of time during a rainfall storm. The Horton's infiltration equation is a simplified version of the Richard's equation under simplified assumptions. The Richard's equation is the basic governing differential equation for the movement of water through unsaturated soil under unsteady conditions that is based on the laws of conservation of mass and momentum [3]. The Horton's infiltration equation can be represented by the following mathematical expression [2,14]:

$$f_t = f_c + (f_0 - f_c)e^{-kt} \tag{1}$$

where f_t is the potential infiltration rate at time t (mm/h), f_c is the constant steady-state infiltration rate after sufficient time has elapsed (mm/h), f_0 is the initial infiltration rate at the beginning of a rainfall event (mm/h), k is an exponential decay constant (h^{-1}), and t is time. Clearly, the three parameters involved in the Horton's equation that need to be determined using the known set of rainfall and runoff data are f_0, f_c, and k. The parameters of the Horton's model can be determined using the rainfall and runoff data.

3.2. Green-Ampt infiltration model

Green-Ampt infiltration model was employed in this study to compute infiltration for the purpose of synthetic data generation (explained later). The Green-Ampt infiltration model was selected because it not only provides an exact analytical solution to the simplified governing continuity and momentum equations for the movement of water through the soil beneath the earth but also provides estimates of the potential incremental infiltration at each time step based on the updated soil moisture storage continuously. The Green-Ampt model of computing infiltration can be mathematically represented by the following equations [3]:

$$f_t = K\left(\frac{\psi\Delta\theta}{F_t} + 1\right) \tag{2}$$

$$F_t = Kt + \psi\Delta\theta \ln\left[\frac{F_t + \Psi\Delta\theta}{\Psi\Delta\theta}\right] \tag{3}$$

$$\Delta F_t = K\Delta t + \psi\Delta\theta \ln\left[\frac{F_t + \Psi\Delta\theta}{F_{t-1} + \Psi\Delta\theta}\right] \tag{4}$$

$$\Delta\theta = (1 - S_e)\theta_e \tag{5}$$

$$S_e = \frac{\theta - \theta_r}{\eta - \theta_r} \tag{6}$$

where f_t is the potential infiltration rate at time t (mm/h), F_t is the potential cumulative infiltration into the soil at time t (mm), ΔF_t is the potential incremental infiltration during time interval Δt (mm), K is the saturated hydraulic conductivity of the soil (mm/h), ψ is the soil suction head (mm), S_e is the effective saturation of soil at time t (varies between 0 and 1), θ is the moisture content in the soil at any time t (varies between 0 and η), θ_r is the residual moisture content, and η is the porosity of the soil. It is clear from the Eqs. (2) through (6) that the Green-Ampt model of infiltration is a non-linear model and is capable of providing accurate estimates of infiltration through different types of soils. Also, the calculation of infiltration using Eqs. (2) through (6) is not a straightforward process but involves iterative solution techniques. The details of the solution procedure are not included here and interested readers are referred to [3]. The Green-Ampt infiltration parameters can be physically measured in the laboratory from spatially distributed samples. The parameters can also be calibrated using the rainfall and runoff data. In this study, since the Green-Ampt model has been used to generate synthetic runoff data, its parameters are taken from [3] corresponding to clayey soil.

4. Model development

Two types of approaches have been investigated in this study for the purpose of determination of an optimal set of infiltration parameters of the Horton's model. The first approach uses the classical optimization technique to estimate the parameters of the Horton's infiltration equation using known rainfall and runoff data. The second approach employs the ANN technique to estimate the Horton's infiltration parameters using the same data set. In the first approach, an optimization program is formulated in which the objective function consists of the sum of the squares of the differences between the observed and modelled DRH ordinates. This model is referred to as the runoff optimization model (ROM) in this study. The second approach uses multi-layer feed forward ANNs with back-propagation training mechanism to determine infiltration parameters. Two different ANN models have been developed, the first ANN model uses all the input data and the second ANN model uses portion of the data set as input in order to test the robustness of the ANN technique in situations where only partial information is available. All the models were developed using the synthetic rainfall and runoff data generated in this study; however, the methodologies proposed here can be easily extended to existing catchments.

4.1. Model data

The data needed to test the proposed methodologies include rainfall and runoff data from a catchment. The methodologies proposed in this paper for the calibration of infiltration models have been tested using synthetic data generated in a hypothetical catchment. This can be achieved by randomly generating a series of rainfall values for a specified duration and then transforming this rainfall sequence through a rainfall-runoff model. A simple unit hydrograph (UH) for the hypothetical catchment was adopted as the rainfall-runoff model in this study to compute runoff. The Green-Ampt infiltration model was employed to determine the infiltration losses. Rainfall and runoff data were generated at 10 min time-interval for 12 different rainfall storm events. The procedure of generating the rainfall and runoff data is described in a step-wise procedure below:

1. For a given storm, six values of total rainfalls (in mm) were randomly generated. The six values correspond to six intervals of 10 min each (total duration is $10 \times 6 = 60$ min).
2. The Green-Ampt infiltration model was then used to calculate infiltration at each time step and subtracted from total rainfall to obtain 'effective rainfall' at each time step. The sum of six 'effective rainfall' values gives the 'total effective rainfall' during 1 h event.
3. An assumed 1 h UH for the hypothetical catchment was then used to compute the direct runoff hydrograph ordinates resulting from the computed 'total effective rainfall' using the method of superposition [3].
4. The steps 1 through 3 above were repeated 12 times for each storm in order to generate rainfall runoff data for 12 different storms.

The hypothetical catchment considered in this study consisted of clayey soil having certain infiltration characteristics. The values of the Green Ampt infiltration parameters for the hypothetical catchment were $\eta = 0.475$, $\Psi = 31.63$ cm, $K = 0.03$ cm/h and $\theta_r = 0.0$ corresponding to clay soil taken from [3]. In order to simulate the observation and instrumentation errors, random error component was added to the DRH ordinates computed above, to obtain the observed DRH ordinates to be used in the two approaches investigated in this study. This was achieved by adding/subtracting a fraction of the generated runoff value to itself. The fraction used was randomly generated but was chosen such that the maximum perturbation was not more than 10% of the original value. The rainfall and runoff data corresponding to the first six storm events were used for training and the data corresponding to the remaining six storms were used for testing purposes. It should be noted that although the proposed methodology has been tested using synthetic data in this study, it can be easily extended to real-world problems by using the rainfall and runoff data derived from existing catchments.

4.2. Runoff optimization model

The runoff optimization model involves estimation of the Horton's infiltration parameters by minimizing the sum of squares of differences between the observed and modelled DRH ordinates using the classical least squares optimization technique for unconstrained case.

Let P be the total rainfall during a storm, and ΔF be the cumulative infiltration during the storm, then ΔF can be computed by integrating the Horton's model represented by Eq. (1) during the time limits of t_1 and t_2.

$$\Delta F = \left[f_c(t_2 - t_1) - \left(\frac{f_0 - f_c}{k}\right)(e^{-kt_2} - e^{-kt_1}) \right] \tag{7}$$

where t_1 is the starting time and t_2 is the ending time of the rainfall storm event, and other variables have same meaning as explained earlier. Once the cumulative infiltration during the storm is known, effective rainfall can be calculated by subtracting actual incremental cumulative infiltration ΔF from the total rainfall (P) for the storm. Let $U(i)$ be the UH ordinate at time step i, then the estimated DRH ordinate at time i, QE(i), can be computed as follows:

$$\mathrm{QE}(i) = [P - \Delta F] \times U(i); \qquad i = 1, 2, \ldots, N \tag{8}$$

where N is the total number of DRH ordinates, and the other variable are same as described earlier. Eq. (8) represents a set of N equations in three unknown parameters of the Horton's infiltration equation (f_0, f_c, and k). Because the set of equations represented by Eq. (8) is over-determined, an optimization technique is needed to find the best solution in terms of the three unknowns. The least squares principle can be employed to determine the best solution for the three unknown infiltration parameters. An objective function in terms of the total error equal to the sum of squares of the differences between the observed and estimated DRH ordinates can be represented by the following equation:

$$\mathrm{Min}\, E(f_0, f_c, k) = \sum_{i=1}^{N} [\mathrm{QO}(i) - \mathrm{QE}(i)]^2 \tag{9}$$

where $E(f_0, f_c, k)$ is the objective function to be minimized, which is a function of the three unknown infiltration parameters, QO(i) is the ordinate of the observed DRH at time step i, and other notations have the same meaning as explained earlier. This objective function is non-linear in nature and needs to be solved using a non-linear optimization solution procedure. In this study, the mathematical software MATLAB was employed to determine the optimal set of Horton's infiltration parameters. First, the data from all the 12 storm events were used to determine 12 different optimal sets of Horton's infiltration parameters. These calibrated values of the Horton's infiltration parameters were then used to compute the modelled estimates of the 12 DRH ordinates. The performance of the Horton's infiltration parameters obtained from the classical optimization technique was evaluated by computing various standard statistical performance evaluation criteria (explained later) using both calibration and validation storms.

4.3. Artificial neural network models

The first step in developing an ANN model involves identifying input and output variables and normalizing the data between 0 and 1. The method of 'channelized normalization' was used to normalize the input and output data, wherein the data representing separate physical variables/parameters are normalized separately. Then the best ANN architecture is determined by finding the optimal number of hidden neurons through training of the various architectures using a trial and error method. Once the best ANN architecture is trained, it is validated using the testing data set. As mentioned earlier, two different types of ANN models were developed. The first ANN model (called ANN-1 Model) used all the input data in terms of rainfall and DRH ordinates; whereas, the second ANN model (called ANN-2 Model) employed only a portion of the data from the available data set.

4.3.1. ANN-1 Model

The ANN-1 Model developed in the present study consisted of three layers, one input layer, one hidden layer, and an output layer. The input data of the ANN-1 Model consisted of the total storm rainfall, and the DRH ordinates. The output data of the ANN-1 Model consisted of the three Horton's infiltration parameters. The number of ordinates in the DRH was 27, therefore, the total number of neurons in the input layer was 28, and with three Horton's infiltration parameters representing the output layer, a structure of 28-N-3 was investigated. The number of hidden neurons was varied from 1 to 20 and an ANN architecture giving the minimum training error and the principle of parsimony was used to select the final ANN-1 Model. The rainfall and runoff data along with the 12 sets of infiltration parameters constitute the training and testing patterns for the ANN model. The data from the first six 1 h storms were used for training and those from the remaining six storms were used for testing the ANN model architecture. The ANN architecture consisting of five hidden neurons (i.e. 28-5-3 ANN) was found suitable for the data set considered. Once trained, the 28-5-3 ANN model structure was used to compute 12 different sets of Horton's infiltration parameters using the rainfall and runoff data from the 12 storms. These estimates were then used to regenerate the first six DRHs and predict the remaining six DRHs. The performance of the Horton's infiltration parameters obtained from the ANN technique was then evaluated by computing various standard statistical performance evaluation criteria using both training and testing storms.

4.3.2. ANN-2 Model

The ANN-2 Model was developed to test the performance of the ANNs when presented with partial information. The ANN-2 Model also consisted of three layers, one input layer, one hidden layer, and an output layer. The input data of the ANN-2 Model consisted of the total storm rainfall, and the DRH ordinates at reduced number of time intervals. The output data of the ANN-2 Model consisted of the three Horton's infiltration parameters. The total number of neurons in the input layer of the ANN-2 Model was 13 (1 total rainfall and 12 DRH ordinates), and with three Horton's infiltration parameters representing the output layer, a structure of 13-N-3 was investigated. The number of hidden neurons was varied from 1 to 20 and an ANN architecture giving the minimum training error and the principle of parsimony was used to select the final ANN-2 Model. The data from the first six 1 h storm patterns were used for training and those from the remaining six storms were used for testing the ANN-2 Model architecture. For this category also, an ANN architecture consisting of five hidden neurons (i.e. 13-5-3 ANN) was found suitable for the data set considered. Once trained, the 13-5-3 ANN model structure was used to compute 12 different sets of Horton's infiltration parameters using the rainfall and runoff data from the 12 storms. These estimates were then used to regenerate the first six DRHs and predict the remaining six DRHs. The performance of the Horton's infiltration parameters obtained from the ANN technique was then evaluated by computing various standard statistical performance evaluation criteria using both training and testing storms.

5. Performance evaluation criteria

Three different types of standard statistical performance evaluation criteria were employed to evaluate the performance of the Horton's infiltration parameters computed using both classical optimization and ANN techniques in this study. These are average absolute relative error (AARE), threshold statistics for an absolute relative error (ARE) level of x% (TS_x), and the correlation coefficient (R). The three performance evaluation criteria used in the current study can be calculated using the following equations:

$$\mathrm{AARE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\mathrm{QO}(t) - \mathrm{QE}(t)}{\mathrm{QO}(t)}\right| \times 100\% \qquad (10)$$

$$\mathrm{TS}_x = \frac{n_x}{N} \times 100\% \qquad (11)$$

$$R = \frac{\sum_{t=1}^{N}(\mathrm{QO}(t) - \overline{\mathrm{QO}})(\mathrm{QE}(t) - \overline{\mathrm{QE}})}{\sqrt{\sum_{t=1}^{N}(\mathrm{QO}(t) - \overline{\mathrm{QO}})^2(\mathrm{QE}(t) - \overline{\mathrm{QE}})^2}} \qquad (12)$$

where n_x is the number of data points for which the ARE (expression within absolute in equation 10) is less than x%, N is the total number of data points computed, $\overline{\mathrm{QO}}$ is the mean of observed runoff series, $\overline{\mathrm{QE}}$ is the mean of predicted/estimated runoff series, and other variables have the same meaning as explained earlier. Threshold statistics were computed for ARE levels of 5%, 10%, 20%, and 50% in this study. Clearly, lower AARE values and higher TS_x values would indicate good model performance. Correlation coefficient values close to 1.0 indicate good model performance.

The TS and AARE statistics measure the "effectiveness" of a model in terms of its ability to accurately predict data from a calibrated model and have been used in literature [15,16,8,17,10,11]. The other statistic, correlation coefficient R, quantifies the "efficiency" of a model in capturing the complex, dynamic, and non-linear rainfall-runoff process. The global error statistics such as correlation coefficient R and RMSE, tend to give higher weightage to the high magnitude flows due to the involvement of square of the difference between observed and predicted flows, or equivalent expressions. Therefore, the errors in estimating flows are dominated by the errors in estimating high magnitude flows in such global statistics. The error statistics based on percentage error in prediction with respect to observed value (such as TS_x and AARE) are better for performance evaluations as they give appropriate weightage to all magnitude flows (low, medium, or high). This aspect of relative errors, such as AARE and TS_x, has been found to give more appropriate assessment and comparison of various models by some researchers [18] including the authors.

6. Results and discussions

The results in terms of the averaged Horton's infiltration parameters obtained from classical optimization technique and the two ANN models are presented in Table 1. The results in terms of the various standard statistical performance evaluation criteria from the three models are presented in Table 2. It can be noted from the Table 1 that the values of the three Horton's infiltration parameters determined using the rainfall and runoff data by all the three models are comparable, and those from the two ANN models are identical. The average values of the initial infiltration rate (f_0), the final steady-state infiltration rate (f_c), and the Horton's decay exponent value (k) from the classical optimization and the two ANN models were 37.46 mm/h, 6.735 mm/h, 2.2627 h^{-1}; and 36.75 mm/h, 7.683 mm/h, 2.3285 h^{-1}, respectively. The comparable values of the Horton's infiltration parameters suggest that the ANNs can be successfully employed for model calibration purposes as they are able to recognize the inherent relationships among the parameters and the rainfall runoff data from a catchment.

The performance of the Horton's infiltration parameters determined using the ANN technique was found to be better than those determined using classical optimization techniques in terms of various standard statistical performance evaluation criteria for predicting runoff. It can be noted from Table 2 that the ANN-1 Model outperformed the ROM Model in terms of the threshold statistics and AARE during both training and testing; whereas, the performance in terms of the correlation coefficient was identical from all the three models during both training and testing. The best AARE of only 7.6% was achieved from the ANN-1 Model during testing. Further, 91.67% of the predicted DRH ordinates had an ARE of less than 50% (see TS50 in Table 2) from the two ANN Models during both training and testing. The correlation coefficient of 0.9932 was achieved from both the models during

Table 1
Horton's infiltration parameters

Model	f_0 (mm/h)	f_c (mm/h)	K (h^{-1})
ROM	37.46	6.735	2.2627
ANN-1 Model	36.75	7.683	2.3285
ANN-2 Model	36.75	7.683	2.3285

Table 2
Statistical performance evaluation criteria

Model	AARE	*R*	TS5	TS10	TS20	TS50
During training/calibration						
ROM Model	11.54	0.9932	19.86	42.77	58.06	91.67
ANN-1 Model	10.13	0.9932	25.00	48.61	66.67	91.67
ANN-2 Model	10.13	0.9932	22.14	43.09	66.67	91.67
During testing validation						
ROM Model	15.00	0.9932	15.97	34.03	61.11	90.97
ANN-1 Model	7.60	0.9932	37.50	66.67	81.25	91.67
ANN-2 Model	11.33	0.9932	29.17	52.78	67.36	91.67

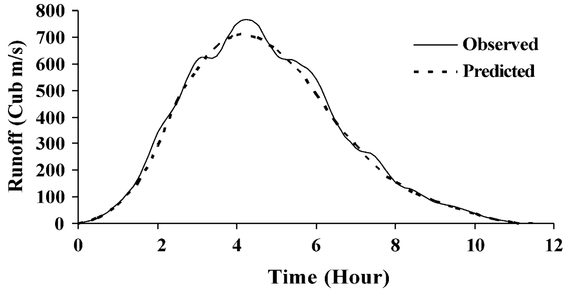

Fig. 3. Observed and predicted DRH by ANN-1 Model for a testing storm.

training and testing, which is considered excellent. Looking at the TS10 statistics from the Table 2 during testing, it can be noted that only 34.03% of the predicted cases from the ROM model had ARE values less than 10% (TS10 = 34.03% in Table 2) while the same statistics was almost double (TS10 = 66.67%) from the ANN-1 Model, which clearly highlights the superior predictive capability of the ANN-1 Model. The performance of the ANN models was consistently superior to that of the ROM model in terms of the other TS statistics also. Among the two ANN models, it can be noted that the performance of the ANN-1 Model was better than that of the ANN-2 Model. This is expected as the ANN-2 Model was developed on the partial information, and the loss of information would have an impact on the quality of the model. However, it is encouraging to note that the performance of the ANN-2 Model does not deteriorate significantly and the drop in its predictive capability is within the acceptable limits for practical considerations. It can be noted from the Table 2 that the differences in the various statistics from the two ANN models during training is minimal. Further analyzing the results during testing, it can be noted that the AARE of 7.60% from the ANN-1 Model increased to 11.33% from the ANN-2 Model; and the TS5 value of 37.50% from ANN-1 Model reduced to 29.17% only from the ANN-2 Model. This clearly indicates that the reduction in performance of the ANN-2 Model was only minimal. This is an important finding from the view point of the ANN model development of the complex, dynamic, and nonlinear rainfall-runoff process. The marginal reduction in the performance of the ANN-2 Model developed on partial information is more than compensated by its compact architecture (13-5-3) involving fewer parameters (weights) to be determined through training as compared to the massive architecture (28-5-3) of the ANN-1 Model developed on complete information. The graphical results in terms of the observed and predicted DRH from the ROM, ANN-1, and ANN-2 Models are presented in Figs. 2–4, respectively. It is clear from the graphical results that all the models perform very well.

It has been observed in this study that even though the performance from any two models can be comparable in terms certain global statistical measures, such as correlation coefficient *R*, the two models can be quite different in making accurate predictions of the variable being modelled. This is apparent from

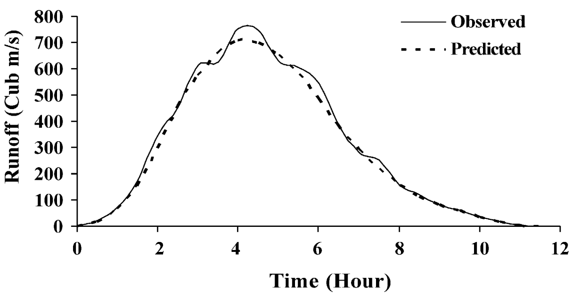

Fig. 2. Observed and predicted DRH by ROM for a testing storm.

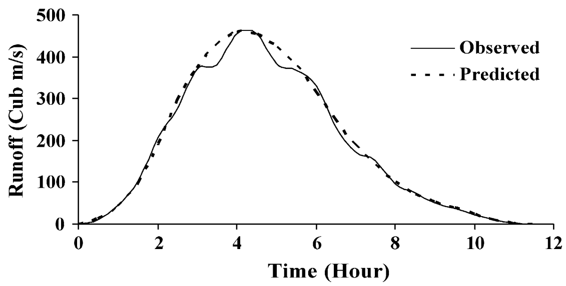

Fig. 4. Observed and predicted DRH by ANN-2 Model for a testing storm.

the identical R values from different models developed in this study but varying TS and AARE statistics from the same models. This highlights the need of evaluating the performance of various ANN models using a wide variety of standard statistical measures rather than relying on a few global error statistics such as correlation coefficient that are similar in nature to the global error being minimized at the output layer of an ANN. The global error statistics, such as RMSE and R, etc., can be biased towards high magnitudes due to the square of the differences between observed and predicted values. The TS and AARE statistics are not biased towards any magnitude of the flow and give unbiased error estimate for the purpose of evaluating the predictive capability of an ANN model.

7. Summary and conclusions

This study presents the findings of an investigation of use of ANN technique for the purpose of determination of infiltration model parameters using observed rainfall and runoff data. In addition to the ANN technique, the classical optimization technique was also explored for comparison purposes. The data in terms of both storm rainfall and runoff were synthetically generated using a hypothetical catchment having known soil properties and a 1 h unit hydrograph. Green-Ampt infiltration model was used to generate the synthetic runoff data from the known rainfall fields, which were generated randomly. Horton's infiltration model was selected for the purpose of the determination of optimal set parameters using known rainfall and runoff data. A total of 12 storms were generated out of which data from the first six storms were used for training and the data from the remaining six storms were employed for validation purposes. The performance of the two techniques was evaluated using certain standard statistical performance evaluation criteria.

The results obtained in this study indicate that the technique of ANN can be suitable for the purpose of determination of infiltration parameters. The performance of the Horton's infiltration parameters determined using the ANN technique was found to be better than those obtained using the classical optimization technique in terms of various standard statistical performance evaluation criteria in predicting runoff. The study of the evaluation of the robustness of the ANN technique in determining the unknown relationships inherent in the input output data when presented with partial information revealed that though the performance of the ANN trained on partial information deteriorated slightly, the reduction in the predictive capability was within acceptable limits for practical considerations. Also, the compact architecture of the ANN model trained on partial information more than compensates for the reduction in its performance as compared to the complex ANN model trained on complete information. Moreover, it has been found that the performances of various ANN models need to be evaluated using a wide variety of standard statistical performance evaluation measures (such as TS_x and AARE) rather than relying on a few global error statistics, such as correlation coefficient and RMSE, normally employed that are similar in nature to the global error minimized at the output layer of an ANN.

No study is complete in itself and there is always scope for further improvements. The inferences drawn in this study are based on synthetically generated data, and the validity of the developed methodology needs to be verified using the real data from existing catchments. The feed-forward ANN with back-propagation training algorithm was employed in this study. Many researchers have reported about the problems associated with the back-propagation method while developing ANN rainfall-runoff models [12]. Recently, some researchers have explored the use of real-coded genetic algorithms (GAs) for training of the ANN rainfall-runoff models, which overcomes some of the problems associated with the back-propagation training algorithm [12,19]. It may be possible to improve the predictive capability of rainfall-runoff models using infiltration parameters determined using ANN models trained using other training methods such as real-coded GAs. It is hoped that future research efforts will focus in these directions to take advantage of the relatively new and emerging soft computing techniques for better planning, design, operation, and management of the water resources systems.

References

[1] B. Zhang, S. Govindaraju, Prediction of watershed runoff using bayesian concepts and modular neural networks, Water Resource Res. 36 (3) (2000) 753–762.

[2] V.P. Singh, Elementary Hydrology, Prentice Hall of India Pvt. Ltd, New Delhi, 1994.
[3] V.T. Chow, D.R. Maidment, L.W. Mays, Applied Hydrology, McGraw-Hill International Editions, McGraw Hill Book Company, New York, 1988.
[4] H. Raman, N. Sunilkumar, Multivariate modeling of water resources time series using artificial neural networks, Hydrol. Sci. J. 40 (2) (1995) 145–163.
[5] D.W. Dawson, R. Wilby, An artificial neural network approach to rainfall-runoff modeling, Hydrol. Sci. J. 43 (1) (1998) 47–65.
[6] A.S. Tokar, M. Markus, Precipitation runoff modeling using artificial neural network and conceptual models, J. Hydrol. Eng. ASCE 5 (2) (2000) 156–161.
[7] S. Birikundavyi, R. Labib, H.T. Trung, J. Rousselle, Performance of neural networks in daily streamflow forecasting, J. Hydrol. Eng. ASCE 7 (5) (2002) 392–398.
[8] A. Jain, S.K.V.P. Indurthy, Comparative analysis of event based rainfall-runoff modeling techniques-deterministic, statistical, and artificial neural networks, J. Hydrol. Eng. ASCE 8 (2) (2003) 93–98.
[9] R.L. Wilby, R.J. Abrahart, C.W. Dawson, Detection of conceptual model rainfall-runoff processes inside an artificial neural network, Hydrol. Sci. J. 48 (2) (2003) 163–181.
[10] A. Jain, S. Srinivasulu, R. Bhattacharjya, Determination of an optimal unit pulse response function using *real-coded* genetic algorithm, J. Hydrol. 303 (1–4) (2005) 199–214.
[11] A. Jain, K.P. Sudheer, S. Srinivasulu, Identification of physical processes inherent in artificial neural network rainfall runoff models, Hydrol. Process. 118 (3) (2004) 571–581.
[12] A. Jain, S. Srinivasulu, Development of effective and efficient rainfall-runoff models using integration of deterministic, real-coded genetic algorithms, and artificial neural network techniques, Water Resource Res. 40 (4) (2004), W04302, doi: 10.1029/2003WR002355..
[13] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, Nature 323 (1986) 533–536.
[14] K. Subramanya, Engineering Hydrology, second ed., Tata McGraw-Hill Publishing Company Limited, New Delhi, India, 1997.
[15] A. Jain, A.K. Varshney, U.C. Joshi, Short-term water demand forecast modeling at IIT Kanpur using artificial neural networks, Water Res. Manage. 15 (5) (2001) 299–321.
[16] A. Jain, L.E. Ormsbee, Evaluation of short-term water demand forecast modeling techniques: conventional v/s artificial intelligence, J. Am. Water Works Assoc. 94 (7) (2002) 64–72.
[17] A. Jain, L.E. Ormsbee, An evaluation of the available techniques for estimating missing fecal coliform data, J. Am. Water Res. Assoc. 40 (6) (2004) 1617–1630.
[18] A. Kumar, K. Minocha, Discussion on rainfall runoff modeling using artificial neural networks, J. Hydrol. Eng. ASCE 6 (2) (2001) 176–177.
[19] S. Srinivasulu, A. Jain, Comparative analysis of training algorithms for ANN rainfall-runoff models, in: Proceedings of the IICAI2003: First Indian International Conference on Artificial Intelligence, December 18–20, 2003, Hyderabad, India, 2003.