# Modeling of daily runoff from a small agricultural watershed using artificial neural network with resampling techniques

Gurjeet Singh, Rabindra K. Panda and Marc Lamers

## ABSTRACT

The reported study was undertaken in a small agricultural watershed, namely, Kapgari in Eastern India having a drainage area of 973 ha. The watershed was subdivided into three sub-watersheds on the basis of drainage network and land topography. An attempt was made to relate the continuously monitored runoff data from the sub-watersheds and the whole-watershed with the rainfall and temperature data using the artificial neural network (ANN) technique. The reported study also evaluated the bias in the prediction of daily runoff with shorter length of training data set using different resampling techniques with the ANN modeling. A 10-fold cross-validation (CV) technique was used to find the optimum number of hidden neurons in the hidden layer and to avoid neural network over-fitting during the training process for shorter length of data. The results illustrated that the ANN models developed with shorter length of training data set avoid neural network over-fitting during the training process, using a 10-fold CV method. Moreover, the biasness was investigated using the bootstrap resampling technique based ANN (BANN) for short length of training data set. In comparison with the 10-fold CV technique, the BANN is more efficient in solving the problems of the over-fitting and under-fitting during training of models for shorter length of data set.

**Key words** | 10-fold cross-validation, ANN, BANN, daily runoff, small agricultural watershed

**Gurjeet Singh**
**Rabindra K. Panda** (corresponding author)
School of Earth,
Ocean and Climate Sciences,
Indian Institute of Technology,
Bhubaneswar,
India
E-mail: rkpanda56@gmail.com

**Marc Lamers**
Institute of Soil Science and Land Evaluation,
University of Hohenheim,
Stuttgart,
Germany

## INTRODUCTION

Accurate simulation of the important hydrological processes such as runoff is essential for planning, design of structures, and management of water resources. Many physical factors such as infiltration, initial soil moisture, evaporation, land use/land cover, watershed geomorphology, and duration of the rainfall make hydrological processes extremely complex, non-linear, dynamic, and fragmented. The rainfall–runoff process is not only complex but also extremely difficult to simulate due to spatial–temporal variability and inter-relationships of the underlying climatic and physiographic variables (Zhang & Govindaraju 2003). Investigations in the past revealed that hydrological modeling is the best technique for satisfactory estimation of the runoff. Based on the degree of complexity, the hydrological models are categorized into empirical black-box models, lumped conceptual

models, and distributed physically based models (Dooge 1977). The lumped conceptual models are sometimes adopted due to limited data requirement, but these models have the limitation of lengthy calibration and parameterization requirement using rigorous optimization techniques (Duan *et al.* 1992). The physically based models consider the controlling physical processes; therefore these models are considered to be a better choice in a rigorous theoretical sense, even though their data requirements are higher. Due to limited availability of data, such models do not perform satisfactorily under field conditions. Generally, in intensively monitored watersheds, all the required data are not available. Thus, there is a need to look for alternative methods for estimation of the runoff using readily available data and information.

Recently, a data-driven modeling approach has received an important boost, due to its ability to overcome some of the limitations associated with the conceptual and physically based models. The data-driven modeling has been developed with contributions from artificial intelligence, data mining, computational intelligence, machine learning, intelligent data analysis, soft computing, pattern recognition, and knowledge discovery in databases. The last two decades have seen increasing popularity of the artificial neural network (ANN) based model for simulation of the hydrological processes. The ANN is one of the intelligence techniques which is flexible and robust and requires less variety of data. Ability of the ANN has been demonstrated while applying to complex systems that may be poorly described or understood using mathematical equations (Tokar & Johnson 1999). Despite the black-box nature of the ANN, it has the flexibility in inclusion of parameters and in capturing the non-linearity of rainfall–runoff processes, making it more attractive for modeling the hydrological processes (Hsu *et al.* 1995).

A preliminary study on modeling the rainfall–runoff process using ANN was initiated by Halff *et al.* (1993), who used a three-layered feed forward ANN for the prediction of hydrographs. Since then, many studies in the area of rainfall–runoff modeling using ANN have been carried out. Extensive review on the application of ANN models in hydrologic simulation and forecasting have been presented in the literature (Maier & Dandy 2000; ASCE 2000a, 2000b; Dawson & Wilby 2001). Hsu *et al.* (1995) found that the ANN model provides a better representation of the rainfall–runoff relationship than the ARMAX time series model or the conceptual SAC-SMA (Sacramento soil moisture accounting) model. Shamseldin (1997) investigated the capability of the ANN model for rainfall–runoff modeling by comparing the simulated discharge with that of the simple linear model, the seasonally based linear perturbation model, and the nearest neighbor linear perturbation model and concluded that the ANN model can simulate discharge more accurately than some of the traditional models. The capability of the ANN model in rainfall–runoff simulation has been demonstrated by many researchers (French *et al.* 1992; Hall & Minns 1993; Navone & Ceccatto 1994; Smith & Eli 1995; Mason *et al.* 1996; Dawson & Wilby 1998; Campolo *et al.* 1999; Sajikumar & Thandaveswara 1999;

Gautam *et al.* 2000; Chang & Chen 2001; Wu *et al.* 2005). The applicability of ANN was also investigated for modeling rainfall–runoff due to typhoon (Chen *et al.* 2013), who reported that a feed forward back-propagation based ANN model performs better than the conventional regression analysis method.

There are some other statistical techniques presented in the literature which have similar or better performance than ANN for rainfall–runoff modeling. There are approaches such as MARS (multiple adaptive regression splines) algorithm (Friedman 1991) and machine learning method M5 model tree (MTs) as given by Quinlan (1992), which use piece-wise linear approximations that are much easier to interpret and provide accuracy comparable to that of ANN. Solomatine & Dulal (2003) used MARS and MTs approaches in rainfall–runoff modeling and found that both techniques have almost similar performance. They concluded that even though the MT was slightly more accurate for 1-h ahead prediction of runoff, performance of ANN is slightly better than the MT for higher lead times. Toth & Brath (2007) used the ANN model and the ADM models for flood forecasting and concluded that the ADM may allow a significant improvement in prediction when focusing on the prediction of flood events and especially in the case of a limited training data availability condition. The fuzzy partitioning mechanisms are used in many hydrologic applications for creating a rule base to generate the output. An extensive review of the application of fuzzy inference system (FIS) for river flow forecasting has been presented by Jacquin & Shamseldin (2009). The FIS is based on the back-propagation algorithm which optimizes the fuzzy membership parameters to develop the best input–output relationship. The adaptive neuro-fuzzy inference system (ANFIS) based on the integration of FIS with the back-propagation algorithm was also applied by a good number of researchers for rainfall–runoff modelling and river flow forecasting (Nayak *et al.* 2005; Aqil *et al.* 2007; Mukerji *et al.* 2009; Pramanik & Panda 2009). Some of the investigators found that the ANFIS, ANN, and FIS models perform similarly in some cases, but the ANFIS predicts better than the ANN and FIS models in most of the cases. In spite of that, presently more and more researchers are utilizing the ANN because this model possesses desirable attributes of universal approximation, and the ability to learn from the examples.

Some previous studies concluded that the lack of physical concepts and relations has been a major limitation of ANN and the reason for the sceptical attitude towards this approach (ASCE 2000a, 2000b). Some hydrologists criticize the ANN model saying it does not reveal anything about the structure of the function that it represents. It is believed that the physics is locked up in the ANN model within the connection weights and threshold values, but these are not easily interpretable. The other limitation of the ANN model is its sensitivity towards the studied data, which means that the structure of ANN becomes totally different with the change of the training data set. A major drawback of the ANN model has been indicated as its incapability of predicting extreme values in river flow (Minns & Hall 1996; Campolo *et al.* 1999). There may be a number of reasons due to which the ANN models are incapable of predicting extreme values (Imrie *et al.* 2000). If the ANN models are trained using a data set that does not contain the maximum possible output value, trained networks may perform poorly for the encountering events containing previously unseen values. The requirement of a lengthy training data set is also a major issue for ANN modeling, because ANN learns from the examples without incorporating physical concepts. Toth & Brath (2007) investigated the impact of the amount of training data on the performance of the ANN model. They concluded that ANN provides an excellent performance for rainfall–runoff simulation of continuous periods, provided that an extensive set of hydro-meteorological data were available for calibration purposes. However, its performance was not satisfactory while focusing on the prediction of flood events, especially in the case of limited availability of training data. Some researchers have tried to address these limitations by applying various approaches while developing the ANN models (See & Openshaw 1999; Cigizoglu 2003; Hettiarachchi *et al.* 2005).

Data pre-processing is an important statistical approach in ANN modeling, which leads to a reduction in the prediction error. The data pre-processing method is very helpful for good generalization ability in the case of inadequate data set. Different resampling techniques may be used for the data pre-processing to explore the hidden properties in the data sets that help in efficient input–output mapping during model training. Generally, hydrological modeling using ANN has adopted simple train-and-test (hold-out)

validation procedure to find the best ANN structure. Sometimes, due to high dimensionality, noise, and inadequate data set, ANN structure decided by the hold-out method can be affected by poor generalization ability and provides biased testing. A cross-validation (CV) procedure has been used for estimating the generalization performance for smaller length of data. Stone (1974) and Geisser (1975) used CV for selecting proper model parameters, as opposed to employing CV purely for selecting proper model performance. Nowadays, CV is broadly adopted in data mining and considered as a standard procedure for performance estimation and model selection. The basic form of CV is $K$-fold CV. In data mining and machine learning where the data set is relatively smaller, a 10-fold CV ($K = 10$) procedure is recommended to check the generalization ability of the model (Weiss & Kulikowski 1991). Molinaro *et al.* (2005) presented a comparison of the resampling methods to estimate the error in prediction. One of the goals of this study was to build the models using relatively smaller data sets to predict the outcome of future observations. Different resampling methods such as split-sample, $K$-fold (2-fold, 5-fold, and 10-fold) CV, leave-one-out CV (LOOCV), Monte Carlo cross-validation, and bootstrap were used to estimate error. Results show that simple split-sample estimates are seriously biased; LOOCV, 10-fold CV, and the bootstrap have the smallest bias. Additionally, the LOOCV, 10-fold CV, and bootstrap have the lowest mean square error. Also it was found that the bootstrap is quite biased in the case of small sample sizes with strong signal-to-noise ratios. More recently, Sharda *et al.* (2006) simulated runoff from middle Himalayan watersheds using the artificial intelligence technique with only 2 years' daily rainfall, runoff, base flow, and total flow data. The authors used a 10-fold CV procedure to check the generalization ability of the models. Akhtar *et al.* (2009) applied a 10-fold CV procedure to build the best ANN model for river flow forecasting. The use of bootstrap (Efron 1979) approach to build a neural network solution for application in hydrological modeling is on the rise. It is a computational procedure that uses intensive resampling with replacement, in order to reduce uncertainty (Efron & Tibshirani 1993). The neural bootstrap has been used to perform bootstrap aggregation, also known as multi-model ensembles, to produce averaged outputs and a more stable solution (Hsieh & Tang 1998). Abrahart (2003) employed

the bootstrap technique for rainfall–runoff modeling and reported that it offered marginal improvement in terms of greater accuracies and better global generalizations. Jeong & Kim (2005) used the bootstrap technique to simulate monthly rainfall–runoff. The authors concluded that the bootstrap based ANN model (BANN) is less sensitive to the input variable selection and the number of hidden nodes than the simple neural network. Shu & Ouarda (2007) applied ensemble ANN models for regional flood frequency analysis in the canonical physiographic space and achieved good results. Sharma & Tiwari (2009) developed BANN for better simulation of monthly rainfall–runoff relationships. Ensemble flood forecasting was made by averaging the output of member bootstrapped neural networks (Tiwari & Chatterjee 2010). The study showed that the ensemble prediction is more consistent, reproducible, and more stable as compared to the traditional ANN approach.

## Background and scope of the present study

The rainfall information alone is insufficient to compute the runoff from a river basin or catchment as the catchment characteristics (related to soil moisture) play an important role in determining the runoff rate variation (Minns & Hall 1996; Campolo *et al.* 1999). In order to represent the soil moisture status, a common approach adopted by most of the above-mentioned researchers, is that runoff or water level at previous time steps were used as inputs to develop the models. An ANN model has the limitation of being highly sensitive towards the inputs, which means that the prediction accuracy of a developed ANN model with specific number of inputs may reduce if a lesser number of inputs are used for prediction. The use of runoff at previous time steps as input may increase the accuracy of the simulation but these developed ANN models may be insufficient to provide good prediction accuracy if runoff at previous time steps is not available. However, it is very difficult to get continuously monitored runoff from small agricultural watersheds in developing countries like India. Therefore, prediction of the daily runoff from small agricultural watersheds requires development of ANN models, without using previous time step runoff as an input.

Towards this investigation, Zhang & Govindaraju (2000) used a feed forward ANN for the prediction of monthly runoff with average monthly rainfall of current and previous months and the average monthly temperature as inputs and utilized Bayesian concept in deriving the training algorithm. The performance of modular networks in predicting runoff over three medium-sized watersheds in Kansas, USA was examined. More recently, a geomorphology based ANN (GANN) model was used for prediction of the watershed runoff as demonstrated by Zhang & Govindaraju (2003). The study shows that GANNs have the potential scope for estimating the direct runoff. Sarangi *et al.* (2005) used geomorphological parameters as inputs with rainfall for prediction of surface runoff and concluded that selected geomorphological parameters with rainfall depth enhanced the accuracy of runoff rate predictions. Sharma & Tiwari (2009) incorporated geomorphologic parameters such as soil, topography, and vegetation information as inputs to the ANN for estimation of monthly runoff at the catchment scale. Raghuwanshi *et al.* (2006) developed ANN models for prediction of daily and weekly watershed runoff and sediment yield using the rainfall and temperature as inputs with 7 years' monsoon season data. As discussed earlier, long length of training data sets is also a limitation of ANN based rainfall–runoff modeling for good accuracy of prediction (Toth & Brath 2007). Some researchers found that a single ANN model is rigid in nature and not suitable for capturing a fragmented input–output mapping during short length of training data sets. Tiwari & Chatterjee (2010) found that short length of training data sets with appropriate representation can perform similarly to the ANN models with long length training data sets for flood forecasting. From the reported review it was found that different resampling techniques may be used for the data pre-processing to explore the hidden properties in the smaller length data set that help in efficient input–output mapping during model development under data scarcity conditions.

In view of the above-mentioned facts, the present study aimed at developing unbiased ANN models to predict daily runoff with limited quantum of continuously monitored runoff data from a small agricultural watershed, namely, Kapgari in Eastern India. The ANN models were developed using easily available climatic variables such as rainfall and temperature for predicting daily runoff from a small agricultural watershed. The 10-fold CV and bootstrapped resampling techniques were used to reduce the prediction

error under data scarcity condition. Effective inputs were selected based on correlation coefficients ($r$) and the ANN models were trained with resampling techniques using the time series of the selected inputs. Finally, the results obtained from the 10-fold CV and bootstrapped resampling technique based ANN models were compared.

## STUDY AREA AND DATA USED

A small agricultural watershed, namely, Kapgari watershed (KGW), situated in the Midnapore district of West Bengal state in Eastern India was selected for this study. This is mainly an agricultural watershed with unmanaged natural resources. The geographical boundary of the watershed lies between 86°50′ and 86°55′ E longitude and 22°30′ and 22°35′N latitude. An unlined main drain and three sub-drains exist in the whole watershed area of 973 ha. Further, on the basis of drainage network and land topography, the watershed was subdivided into three sub-watersheds (Figure 1). The areas of the sub-watersheds KGSW1, KGSW2, and KGSW3 are 280, 330, and 363 ha, respectively. The topography of the watershed is undulating. The slope ranges from 0.07 to 8.29% with an average slope of 3.03%. The major crop of the study area is paddy, which is usually cultivated during the rainy season. The major soil textural classes present are sandy loam, silt loam, clay loam, and loam. However, the predominant soil of the watershed is sandy loam soil. The climate is sub-humid sub-tropical with an average annual rainfall of 1,320 mm, of which about 80% occurs during the rainy months of June to October. Most of the rainwater drains from the watershed
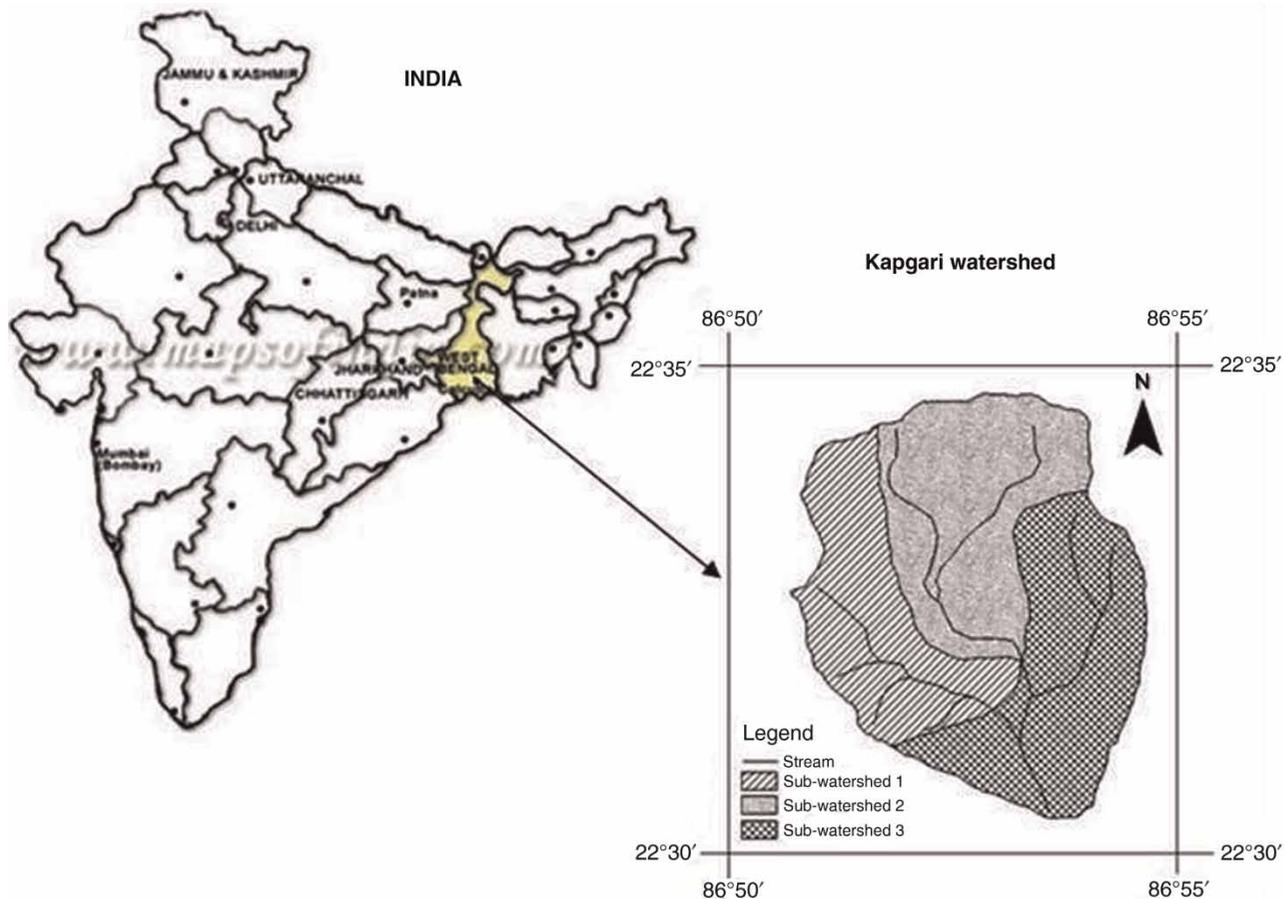


Figure 1 | Location map of the Kapgari watershed.

in the form of surface flow. The daily mean temperature ranges from a minimum of 24 °C to a maximum of 40 °C. The daily mean relative humidity varies from a minimum of 59.4% to a maximum of 94.3%. Average wind speed varies from 1.1 m/s to 2.1 m/s. The mean solar radiation varies from 5.0 kWh/m$^2$ to 7.4 kWh/m$^2$ and average evaporation varies from 2.1 mm/day to 5.7 mm/day.

For effective monitoring of surface runoff, culverts were constructed at the main outlet of the watershed and the outlets of the sub-watersheds, where gauging stations have been installed. Each gauging station consists of a stilling well of 1 m diameter, in which an automatic stage level recorder has been installed for the continuous monitoring of discharge from the sub-watersheds and the whole watershed for the monsoon season of the years 2003 to 2005. A current meter was used to measure the velocity of water passing through the culverts. The fluctuation of water level in the culverts was recorded in the chart papers of automatic stage level recorders. Other metrological data such as daily rainfall, daily maximum and minimum air temperature, were also collected for the same period from the meteorological observatory established in the study area.

## THEORETICAL CONSIDERATIONS

### Artificial neural network

An ANN is a massively parallel distributed processor with a natural propensity for storing knowledge and making it available for use (Haykin 1994). ANN generally consists of three layers: an input layer, an output layer, and one or more hidden layers. The information passes from input layer to output layer through neurons which are present in each layer. The network function is determined by connection between the nodes. The connection between the processing neurons is weighted by scalar weight and a bias, which is adapted during the model training. The number of input neurons, output neurons, and the neurons in the hidden layer depends upon the problem being studied.

A neural network can be trained to perform a particular function by adjusting the value of connection weights between the nodes (Rumelhart & McClelland 1986; Muller & Reinhardt 1991). At the beginning of the training, the

initial values of the weights can be assigned randomly or based on experience. In this process, the learning algorithm systematically changes the weights to correctly perform a desired input–output relationship. The process of training is said to be finished when the mean squared error between the ANN output $y_j(t)$ and the corresponding actual output $d_j(t)$ becomes closer to the preset error goal. The mean squared error $E(t)$ at any time $t$, is calculated over the entire data set using Equation (1).

$$E(t) = \frac{1}{2} \sum_{j=1}^{n} \left[ y_j(t) - d_j(t) \right]^2 \qquad (1)$$

Determination of appropriate ANN architecture is one of the most important tasks in the model-building process. From various past studies it was revealed that a multilayer feed forward network outperforms all the other networks. Although multilayer feed forward networks are one of the most fundamental models, they are also the most popular type of ANN structure suited for practical applications. The multilayer feed forward network has a layered arrangement of the nonlinear processing nodes (Figure 2). The architecture consists of hidden layers of neural network and hidden neurons in the hidden layers. Determination of the optimum number of hidden layers and nodes in the
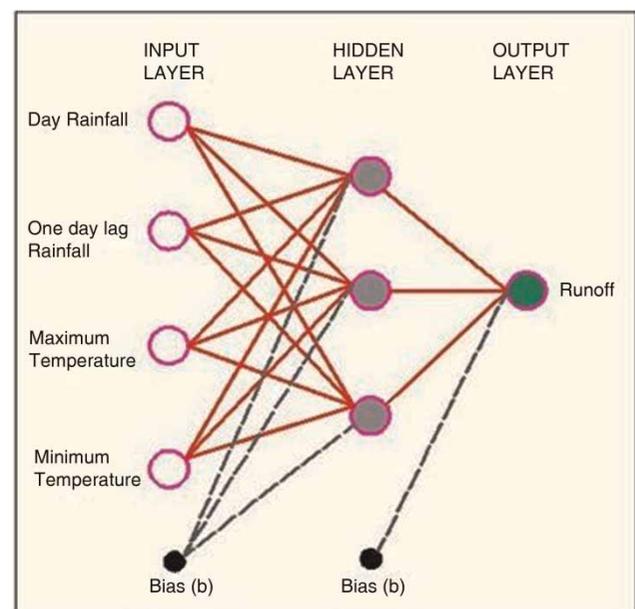


**Figure 2** │ ANN architecture with one hidden layer.

neural network architecture, to produce good results, is a trial-and-error procedure which depends on the type of problem and the availability of data. Hecht-Nielsen (1990) provided a proof that even one hidden layer of neurons can be sufficient to model any solution surface of practical interest. Wang *et al.* (2011) reported that one of the important issues in the development of an ANN model is the determination of the optimal number of neurons in the hidden layer that can satisfactorily capture the nonlinear relationship existing between the input and the output variables. If the number of neurons in the hidden layer is small, the network may not have sufficient degrees of freedom to learn the process correctly. The training takes a long time and the network may sometimes over-fit the data due to the number of neurons in the hidden layer being too large (Karunanithi *et al.* 1994). The learning algorithm to train the network for responding correctly to a given set of inputs also plays an important role in the process of model development. A properly trained network with Levenberg–Marquardt (LM) back propagation algorithm gives reasonable results when presented with new inputs during the validation (Hagan & Menhaj 1994).

In the present study, a multilayered feed forward network having one hidden layer and a sigmoid activation function in the neurons was trained with LM back propagation algorithm. A 10-fold CV method was adopted for selection of the optimal number of hidden neurons to avoid the neural network over-fitting during the training process. Finally, the trained architecture (having the optimized values of the connection weights) was tested using the unseen data sets to evaluate the model performance.

### Neural network training algorithm

A wide range of algorithms has been developed for training the ANN network to achieve the optimum model performance, while ensuring generalization and computational efficiency. A back-propagation algorithm is the most common neural network training algorithm (Fausett 1994; Patterson 1996). A standard second-order nonlinear least-square technique based on the back-propagation algorithm, such as the LM algorithm to increase the speed (Masters 1995) and efficiency of the training (Hagan & Menhaj 1994), was used for training the ANN models. Demuth &

Beale (1998) reported that the LM algorithm uses a second-order training mode without having to compute the Hessian matrix. When the performance function has the form of the sum of squares (as is typical in training feed-forward networks), the Hessian matrix can be approximated as $H = J^T J$ and the gradient can be computed as $g = J^T e$, where $J$ is Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases, and $e$ is a vector of network errors. The LM algorithm uses the above approximation to the Hessian matrix in the Newton-like weights update as Equation (2):

$$\vec{w}(n + 1) = \vec{w}(n) - \left[ \vec{J^T}(n) \, \vec{J}(n) + \mu I \right]^{-1} \vec{J^T}(n) \, \vec{e}(n) \qquad (2)$$

where $w$ indicates the weight of the neural network, and $\mu$ is a non-negative scalar that controls the learning process. When the parameter $\mu$ is large, the above expression approximates gradient descent with a small step size; while for a small $\mu$ the algorithm approximates the Newton method. The LM can overcome between its two extremes, the gradient descent and Newton's algorithm. The LM method is a standard method for minimization of the mean square error criterion, due to its rapid convergence properties and robustness (Demuth & Beale 1998).

### Ten-fold CV technique

In the data mining and machine learning, where the length of data for the model development is inadequate, a 10-fold CV is the most common procedure recommended to check the generalization ability of the model (Weiss & Kulikowski 1991). This is performed to avoid neural network over-fitting during the training process. To perform the 10-fold CV procedure, data for the model development are first partitioned into 10 equally (or nearly equal) sized segments or folds. After making 10-fold, 10 iterations of the training and validation are performed such that in each iteration, a different fold of the data is held-out for the validation while the remaining nine folds are used for training. Subsequently, the trained models are used for predictions of the data in the validation fold. Thus, each time a model is constructed and tested with an 'unseen' data set. The predetermined performance functions can be used to track the performance of each

learning algorithm on each fold. The samples obtained from the results of performance of each learning algorithm on each fold can be used in a statistical hypothesis test or an aggregate measure of these samples obtained by the different methodologies such as averaging. The magic of this procedure is that, during each fold, 90% of the data is available for training, yet the final performance is based on 100% of data. The advantage of this method is that it performs reliable unbiased testing on the smaller data set, because this process requires much more computational effort than simple trained-and-tested (hold-out) approach.

## Bootstrapping

Bootstrapping is a data-driven simulation method that uses intensive resampling with replacement. Efron & Tibshirani (1993) developed the bootstrapping method which manipulates the training data sets in order to generate different models and uses them to obtain an aggregated predictor. The bootstrap procedure involves resampling with replacement, to reduce the uncertainties. The aim of resampling is to mimic the random component of a process and to reduce variance through averaging over numerous different partitions of the data. The neural network simulations generally converge at the local minima. This results in slightly different predictions every time neural networks are trained, due to the random initializations of the weight matrix. The neural bootstraps have been used to perform the bootstrap aggregation (bagging) of multi-model ensembles which

produce averaged outputs and a more stable solution. This is done by repeated sampling with replacement of the original data set of size N, to obtain B bootstrap data sets, each with a size of N. Each bootstrap data set contains different data, resulting in B neural network models, all of which may differ slightly. Efron & Tibshirani (1993) suggested B to be between 50 and 200.

By calculating averages and standard deviations of B testing results, one obtains robust values of the predicted variable and associated uncertainty estimates for independent data. A model $f_{\text{ANN}}(x)$ fitted to each of the generated bootstrap data sets and bootstrapping estimate $f_{\text{bootstrap}}(x)$ is calculated as the mean of each model (Equation (3)):

$$f_{\text{bootstrap}}(x) = \frac{1}{B} \sum_{b=1}^{B} f_{\text{ANN}}(x) \tag{3}$$

## METHODOLOGY

### Selection of input variables

Determination of significant input variables is one of the most important steps in the ANN hydrologic model development process (Bowden *et al*. 2005a, b). The reported study used a correlation coefficient (significant at 0.01 levels, two-tailed) of the hydro-climatic data to select the effective inputs. The correlation coefficients between variables are presented in Table 1, which shows that there is a very

**Table 1** | Correlation coefficients ($r$) of the hydro-climatic data for the Kapgari watershed

| | $P_t$ | $P_{t-1}$ | $P_{t-2}$ | $T_{\max}$ | $T_{\max-1}$ | $T_{\min}$ | $T_{\min-1}$ | $Q_t$ |
|---|---|---|---|---|---|---|---|---|
| $P_t$ | 1 | 0.232 | 0.084 | *0.331* | *0.068* | *0.288* | *0.102* | 0.856 |
| $P_{t-1}$ | | 1 | 0.232 | *0.353* | *0.342* | *0.271* | *0.299* | 0.486 |
| $P_{t-2}$ | | | 1 | *0.279* | *0.333* | *0.195* | *0.253* | 0.191 |
| $T_{t,\,\max}$ | | | | 1 | 0.753 | 0.689 | 0.588 | *0.356* |
| $T_{t-1,\,\max}$ | | | | | 1 | 0.556 | 0.692 | *0.118* |
| $T_{t,\,\min}$ | | | | | | 1 | 0.772 | *0.305* |
| $T_{t-1,\,\min}$ | | | | | | | 1 | *0.123* |
| $Q_t$ | | | | | | | | 1 |

Italic numbers represent negative correlations. $P_t$ = current day rainfall; $P_{t-1}$ = 1-day lag rainfall; $P_{t-2}$ = 2-day lag rainfall; $Q_t$ = current day runoff; $T_{t,\,\max}$ = current day maximum temperature; $T_{t,\,\min}$ = current day minimum temperature; $T_{t-1,\,\max}$ = 1-day lag maximum temperature; $T_{t-1,\,\min}$ = 1-day lag minimum temperature.

good correlation of the present day runoff with present day rainfall and significant correlation with the 1-day lag rainfall. The correlation coefficient between the present day runoff and the present day maximum and minimum temperatures was also significant. It is also observed that the correlation among daily runoff, 2-day lag rainfall, and 1-day lag maximum and minimum temperature was quite low. Therefore, these combinations were not considered for this study. The description of different ANN models developed for simulating daily runoff is presented in Table 2. A schematic diagram of architecture of developed neural networks with a single hidden layer is shown in Figure 2.

## Data preparation

ANN has the ability to handle nonlinear, noisy, and nonstationary data. However, with suitable data preparation beforehand, it is possible to improve the modeling performance (Maier & Dandy 2000; Bray & Han 2004). The data

preparation involves a number of processes such as data collection, data division, and data pre-processing. In the present study, daily runoff as well as daily rainfall, maximum and minimum temperatures during monsoon season were collected for a period of 3 years (2003–2005) from the observatory installed in the study area. The results of statistical analysis of available hydro-metrological data for the present study are presented in Table 3. The statistical analysis showed that the year 2005 had the highest peak value followed by the years 2004 and 2003. The trend of variation of all the variables is similar during all 3 years. However, due to some high intensity rainfall events, the runoff peaks were higher in 2004 and 2005. Normally, a major portion of the data set is used for the training and a relatively smaller portion is used for testing the ANN models. Thus, 2 years' data were selected for the development of models and the remaining 1 year's data were used for testing the developed models. However, a major limitation of the ANN models is that they are not usually able to extrapolate. The training pattern therefore should go at least to the edge of the problem domain in all dimensions. To use a higher range of available data in the model training, data sets of the years 2004 and 2005 which cover all the possible scenarios were used for the model development and the developed models were tested using the unseen data sets of the year 2003 to evaluate the model performance.

It is also essential for the ANN model to have training pattern evenly distributed within the region. If the pattern

**Table 2** | Selected input combinations to develop ANN models for prediction of daily runoff

| Model no | Input layer neuron combinations | Output layer neuron |
|----------|--------------------------------|---------------------|
| M1 | $P_t$ | $Q_t$ |
| M2 | $P_t$ and $P_{t-1}$ | $Q_t$ |
| M3 | $P_t$, $P_{t-1}$, $T_{t, max}$ and $T_{t, min}$ | $Q_t$ |

**Table 3** | Basic statistical parameters of the observed daily hydro-meteorological data of KGW and its sub-watersheds

| Year | Statistics | Meteorological data | | | Runoff (mm) | | | |
|------|-----------|----------|-----------|-----------|---------|-------|-------|-------|
| | | Rainfall | $T_{max}$ | $T_{min}$ | Kapgari | KGSW1 | KGSW2 | KGSW3 |
| 2003 | Mean | 7.55 | 33.64 | 29.09 | 1.12 | 1.31 | 0.32 | 1.70 |
| | STD | 12.50 | 2.53 | 1.78 | 1.90 | 2.26 | 0.54 | 2.89 |
| | Maximum | 63.00 | 39.50 | 36.00 | 8.18 | 9.88 | 2.49 | 12.41 |
| 2004 | Mean | 8.22 | 33.93 | 28.65 | 0.93 | 1.01 | 0.72 | 1.27 |
| | STD | 13.91 | 2.64 | 1.89 | 1.86 | 2.02 | 1.43 | 2.53 |
| | Maximum | 68.40 | 39.00 | 33.00 | 10.64 | 11.35 | 7.95 | 14.28 |
| 2005 | Mean | 5.80 | 34.78 | 29.58 | 0.63 | 0.70 | 0.50 | 0.86 |
| | STD | 16.57 | 2.38 | 1.89 | 1.93 | 2.18 | 1.55 | 2.62 |
| | Maximum | 110.40 | 39.50 | 34.00 | 13.83 | 15.89 | 10.98 | 18.65 |

STD represents the standard deviation.

is not evenly distributed, the trained network performs poorly for the region where density of the training data set is less. Therefore, to perform a 10-fold CV procedure, data for the model development are stratified prior to being split into 10 folds. Stratification is a data rearrangement process to ensure that each fold is a good representative of the whole. The sigmoid activation function used for training the network has lower and upper limits of 0 and 1, respectively. Therefore, in order to suit the consistency of the model, the data used in input and output layers were normalized in the range of 0 to 1 and then returned to original values after the simulation. Equation (4) was adopted to normalize the data set:

$$X_{i,\text{norm}} = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \qquad (4)$$

where $X_i$ is the original values of different selected inputs, $X_{i,\text{norm}}$ is the normalized value, $X_{\max}$ and $X_{\min}$ are the maximum and minimum values respectively.

The normalization of the data, ANN model development, and analysis of the results were performed by using MATLAB software. MATLAB is an interactive software tool for engineering analysis and mathematical computations entirely written in the C language. The command window of MATLAB (http://www.mathworks.in/help/nnet/index.html) was used for the ANN model development.

### Ten-fold CV based ANN models development

The goal of the reported study was to develop the ANN models for predicting the daily runoff using climatic variables with smaller length of data set. Therefore, in this study, different combinations of the rainfall and temperature were considered as input for the ANN models' development. To develop ANN models using shorter length of the training data set, such as weather data for 2 years (2004–2005), a 10-fold CV procedure was used to check the generalization ability of the model. The main purpose of using a 10-fold CV technique was to find the optimum number of hidden neurons in the hidden layer to avoid neural network over-fitting during the training process. A procedure was followed as mentioned earlier. The performance of each learning algorithm on each fold was tracked using root mean square error

(RMSE) and mean absolute error (MAE). After completing 10 iterations of the training and validation, average RMSE and MAE were used to obtain an aggregate measure from these samples. The ANN structure, which showed the least CV error, was selected for the daily runoff prediction. After selection of the optimum number of inputs and neurons in the hidden layer, the model was trained with the whole data set of the years 2004 and 2005 and evaluated with the observed data of the year 2003.

### Bootstrap based ANN models development

The bootstrap resampling method was used to generate different realizations of the data sets to create a set of bootstrap samples using intensive resampling with replacement. These data sets provided a better understanding of the average and variability of the original unknown distribution or process. The bootstrap BANN is less sensitive to the input variable selection and the number of hidden nodes than the simple neural network (Jeong & Kim 2005). Therefore, the optimized number of inputs and neurons in the hidden layer were used to develop the BANN. Finally, results obtained from the BANN models were compared with the simple ANN models.

### Performance evaluation of models

The performance of the trained network was evaluated using the criteria proposed by previous researchers (Nash & Sutcliffe 1970; WMO 1975; ASCE Task Committee on Definition of Criteria for Evaluation of Watershed Models 1993). In the reported study, the standard goodness-of-fit such as coefficient of simulation efficiency (COE), coefficient of determination ($R^2$), RMSE, MAE, and percent deviation ($D_v$) were used for model evaluation. The COE provides a measure of the ability of a model to predict the values that are different from the mean. The value of COE close to 1 indicates good agreement between the observed and predicted values. The $R^2$ value provides a measure of how well future outcomes are likely to be predicted by the model. High $R^2$ value indicates a close relationship between the observed and model predicted values. The MAE is a linear score which indicates that all the individual differences are weighted equally in the average but the RMSE is

a quadratic scoring rule that provides a relatively high weight to the larger errors. This means that the *RMSE* is most significant when large errors are particularly undesirable. Both the *MAE* and *RMSE* can range from 0 to ∞. *MAE* and *RMSE* equal to zero indicate a perfect match between the observed and predicted values. These coefficients are independent of the scale of the data used and are useful in assessing the goodness of fit of the model (Dawson & Wilby 1998). The magnitude of the percent deviation ($D_v$) indicates the magnitude of difference between the measured value and the simulated value. Bingner et al. (1989) reported that over- and underprediction limits for model simulation within 20% from the measured values are considered as the acceptable levels of accuracy for the simulations. The mathematical expressions of the goodness-of-fit indices (GFI) used in the study are presented in Table 4.

## RESULTS AND DISCUSSION

### Development of the ANN models for Kapgari watershed

The ANN model structure was determined using a 10-fold CV procedure for unbiased testing. The hidden neurons in

a single hidden layer were varied from 1 to 5. For each hidden neuron, 10 iterations of the training and validation were performed and performance of each learning algorithm on each fold was tracked using the *RMSE* and *MAE* performance evaluation functions. Averaging was used to obtain an aggregate measure of all 10 iterations. The average performance of all hidden neurons is presented in Table 5. It is observed from Table 5 that as we went on adding hidden neurons the *RMSE* and *MAE* decreased for both training as well as CV but after a certain number of hidden neurons the *RMSE* and *MAE* decreased for the training and increased for the CV. Hence, in this study, the ANN structure was decided after considering the least *RMSE* and *MAE* from the CV results. Using only a present day rainfall as input, the Model KGM1 (1-3-1) gave a minimum *RMSE* and *MAE* values of 0.726 mm and 0.382 mm, respectively, during the CV. Adding 1-day lag of the daily rainfall as input improved the prediction performance and the Model KGM2 (2-3-1) gave minimum *RMSE* and *MAE* values of 0.638 and 0.339 mm, respectively, for prediction of the daily runoff during CV. Therefore, Model KGM2 (2-3-1) was considered the best when only daily rainfall data are available. Similarly, addition of the daily maximum and minimum temperature data along with the rainfall data also improved the performance of prediction and the

**Table 4** | Mathematical expressions of the GFI

| Goodness-of-fit indices | Abbreviation | Formula |
|---|---|---|
| Coefficient of determination | $R^2$ | $\left( \dfrac{\sum_{i=0}^{n} (O_i - \overline{O_i})(P_i - \overline{P_i})}{\sqrt{\sum (O_i - \overline{O_i})^2 \sum_{i=1}^{n} (P_i - \overline{P_i})^2}} \right)^2$ |
| Coefficient of simulation efficiency | COE | $1 - \dfrac{\sum_{i=1}^{n} (O_i - P_i)^2}{\sum_{i=1}^{n} (O_i - \overline{O_i})^2}$ |
| Root mean square error | RMSE | $\sqrt{\dfrac{1}{n} \sum_{i=1}^{n} (O_i - P_i)^2}$ |
| Mean absolute error | MAE | $\dfrac{1}{n} \sum_{i=1}^{n} |O_i - P_i|$ |
| Percent deviation | $D_v$ (%) | $\dfrac{V - V'}{V} \times 100$ |

$O_i$ and $P_i$ are the observed and predicted values; $\overline{O}$ is the mean of observed values; $V$ is the measured value of variable and $V'$ is the model predicted value.

**Table 5** | Performance results of the 10-fold CV technique for prediction of the daily runoff from Kapgari watershed

| Model no. | ANN structure | Training error RMSE (mm) | Training error MAE (mm) | CV error RMSE (mm) | CV error MAE (mm) |
|---|---|---|---|---|---|
| KGM1 | 1-1-1 | 0.588 | 0.395 | 0.851 | 0.433 |
| | 1-2-1 | 0.557 | 0.376 | 0.803 | 0.416 |
| | 1-3-1 | 0.498 | 0.347 | 0.726 | 0.382 |
| | 1-4-1 | 0.452 | 0.285 | 0.734 | 0.406 |
| | 1-5-1 | 0.424 | 0.186 | 0.738 | 0.413 |
| KGM2 | 2-1-1 | 0.536 | 0.363 | 0.711 | 0.396 |
| | 2-2-1 | 0.482 | 0.320 | 0.682 | 0.362 |
| | 2-3-1 | 0.411 | 0.273 | 0.638 | 0.339 |
| | 2-4-1 | 0.371 | 0.268 | 0.703 | 0.348 |
| | 2-5-1 | 0.366 | 0.261 | 0.730 | 0.383 |
| KGM3 | 4-1-1 | 0.486 | 0.337 | 0.624 | 0.376 |
| | 4-2-1 | 0.423 | 0.271 | 0.586 | 0.348 |
| | 4-3-1 | 0.400 | 0.260 | 0.573 | 0.334 |
| | 4-4-1 | 0.384 | 0.241 | 0.665 | 0.368 |
| | 4-5-1 | 0.379 | 0.238 | 0.668 | 0.375 |

Model KGM3 (4-3-1) gave the least *RMSE* and *MAE* of 0.573 and 0.334 mm, respectively, for prediction of the daily runoff during CV. Hence, for this study, the Model KGM3 (4-3-1) was considered the best for prediction of daily runoff and was used for the further analysis.

The selected ANN structure KGM3 (4-3-1) was trained using the whole data set of the years 2004 and 2005. The weights for this trained structure were saved and the network was evaluated with testing data set of the year 2003. The scatter plots of the observed and KGM3 (4-3-1) model predicted daily runoff during the training and testing period are shown along with 1:1 line in Figures 3(a) and 3(b) respectively. It is observed from the figures that during the training period, the major portion of the scatter plot is well distributed about the 1:1 line. The results for the testing period showed an overestimation of the very low runoff values which may be due to the fact that the minimum value of the output training vector was higher than the corresponding value in the testing data set. It was found from the training and testing period results that the predictions of peak or higher runoff values are better than low and medium runoff values. This may be due to the fact that the model is trained well with the peak or higher values and is capable of predicting peak values with considerable accuracy. A statistical analysis was performed to compare the predicted daily runoff during both training and testing periods and the results are presented in Table 6. The value of coefficient of determination ($R^2$) of 0.95 and 0.90 indicates a close match between the measured and predicted daily runoff values during the training and

testing periods respectively. The *COE* of 0.94 for the training and 0.88 for the testing shows a good agreement between the measured and predicted daily runoff.

## Development of the ANN models for Kapgari sub-watersheds

All the networks for the sub-watersheds were developed using the same procedure that was used for model development for the whole Kapgari watershed. Generalized ANN structures were selected using the 10-fold CV results. On the basis of minimum error in CV, the best ANN structures were selected for each input combination (M1, M2, and M3). The performance of the selected best model for each input combination, in terms of *RMSE* and *MAE* is presented in Table 7. It is observed from Table 7 that the ANN structure with input combination M3 has better generalization ability than the combination of M1 and M2 for all the sub-watersheds. Therefore, the ANN structures KGSW1M3 (4-3-1), KGSW2M3 (4-4-1), and KGSW3M3 (4-3-1) were selected for prediction of the daily runoff from Kapgari sub-watersheds 1, 2, and 3, respectively.

The selected ANN models were trained using the whole data sets of the years 2004 and 2005. The weights for all these trained models were saved and these networks were evaluated using the testing data set of the year 2003. The scatter plot of the observed and predicted daily runoff during the training and testing periods for sub-watersheds 1, 2, and 3 are shown along with 1:1 line in Figure 4. The daily runoff values were slightly overpredicted for
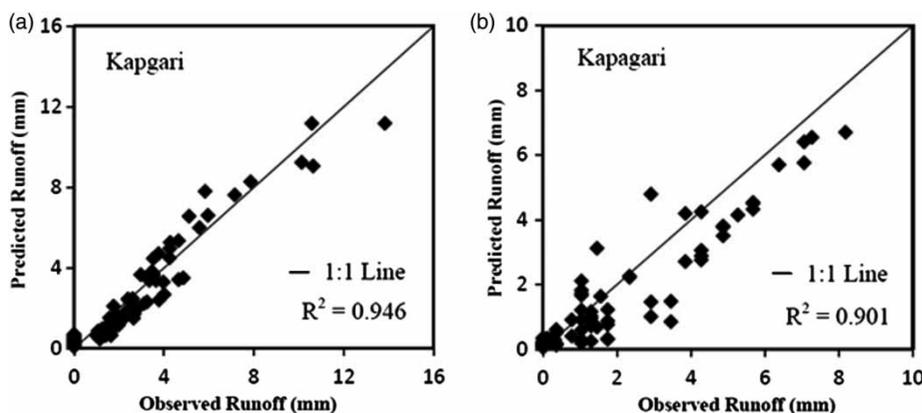


**Figure 3** │ Scatter plots of observed versus ANN model predicted daily runoff from KGW during (a) training period (2004–05) and (b) testing period (2003).

**Table 6** | Goodness-of-fit statistics of the observed and predicted daily runoff from Kapgari watershed

| Statistical parameters | Runoff (mm) | | | |
| | Training period (2004–05) | | Testing period (2003) | |
| | Observed | Simulated | Observed | Simulated |
|---|---|---|---|---|
| Mean | 0.78 | 0.87 | 1.12 | 0.96 |
| Standard deviation | 1.89 | 1.78 | 1.90 | 1.57 |
| Maximum peak | 13.83 | 11.18 | 8.18 | 6.70 |
| Total | 212.38 | 237.85 | 153.61 | 131.98 |
| COE | 0.94 | | 0.88 | |
| $R^2$ | 0.95 | | 0.90 | |
| RMSE | 0.44 | | 0.67 | |
| MAE | 0.32 | | 0.41 | |
| $D_v$ % | 11.99 | | 14.09 | |

**Table 7** | Performance results of the best models selected by a 10-fold CV technique for prediction of the daily runoff from Kapgari sub-watersheds

| Model no. | ANN structure | Training error | | CV error | |
| | | RMSE (mm) | MAE (mm) | RMSE (mm) | MAE (mm) |
|---|---|---|---|---|---|
| KGSW1M1 | 1-3-1 | 0.475 | 0.319 | 0.808 | 0.428 |
| KGSW1M2 | 2-3-1 | 0.472 | 0.318 | 0.758 | 0.413 |
| KGSW1M3 | 4-3-1 | 0.468 | 0.317 | 0.705 | 0.388 |
| KGSW2M1 | 1-3-1 | 0.318 | 0.233 | 0.473 | 0.272 |
| KGSW2M2 | 2-4-1 | 0.315 | 0.231 | 0.469 | 0.269 |
| KGSW2M3 | 4-4-1 | 0.313 | 0.208 | 0.414 | 0.226 |
| KGSW3M1 | 1-3-1 | 0.556 | 0.411 | 1.007 | 0.556 |
| KGSW3M2 | 2-3-1 | 0.542 | 0.398 | 0.983 | 0.532 |
| KGSW3M3 | 4-3-1 | 0.536 | 0.385 | 0.852 | 0.474 |

the medium flow and underpredicted for the peak and low flow conditions. The training period scatter plot shows the data points to be very well distributed around the 1:1 line. However, during the testing periods, mostly under-prediction was observed, even though the models were capable of predicting the low and high runoff values with considerable accuracy. Statistical analysis was performed to compare the predicted daily runoff for both training and testing periods and is presented in Table 8. The values of the coefficients of determination ($R^2$) were found to be 0.944, 0.951, and 0.956 during the training and 0.902,

0.755, and 0.873 during the testing for sub-watersheds 1, 2, and 3, respectively. These high $R^2$ values indicate a close relationship between the observed and predicted daily runoff. The COE of 0.941, 0.947, and 0.954 during training and 0.875, 0.732, and 0.859 during the testing period for sub-watersheds 1, 2, and 3, respectively, show a good agreement between the daily observed runoff and predicted runoff.

These results clearly indicate that the ANN models KGM3 (4-3-1), KGSW1M3 (4-3-1), KGSW2M3 (4-4-1), and KGSW3M3 (4-3-1) can be used to predict the daily runoff using daily rainfall and temperature data as input. This also illustrates the application of ANN in modeling highly complex and nonlinear phenomenon of the rainfall–runoff process. Thus, the ANN models with easily available information can be used to predict the daily runoff.

## Development of the BANN models for Kapgari watershed and its sub-watersheds

The best input combination M3 was chosen to develop the BANN for KGW and its sub-watersheds. The BANN models were used to perform the bootstrap aggregation of multi-model ensembles to get a more stable solution. Chernick & LaBudde (2010) suggested that for a small sample size, the lower order bootstrap works better than the higher order and suggested that 20 to 100 bootstrap samples be used for small sample sizes. Therefore, Bootstrap.xla an Excel Add-In (Barreto & Howland 2006) was used to develop 50 bootstraps' resample for this study. Each bootstrapped resample was used to develop an ANN model, and likewise, 50 ANN models were developed and then combined (using mean) to approximate the relationship between model inputs and outputs.

Ensemble prediction of the daily runoff was done by 50 sets of weights instead of one set of weights for the testing year 2003. Comparison between the observed and the predicted daily runoff by the ANN and BANN models during the testing period is presented graphically in Figure 5 for the whole KGW and its sub-watersheds KGSW1, KGSW2, and KGSW3 respectively. The scatter plots show that BANN models estimate better peaks or higher runoff values than the 10-fold CV based ANN models for KGW and its sub-watersheds. It is observed from these plots that
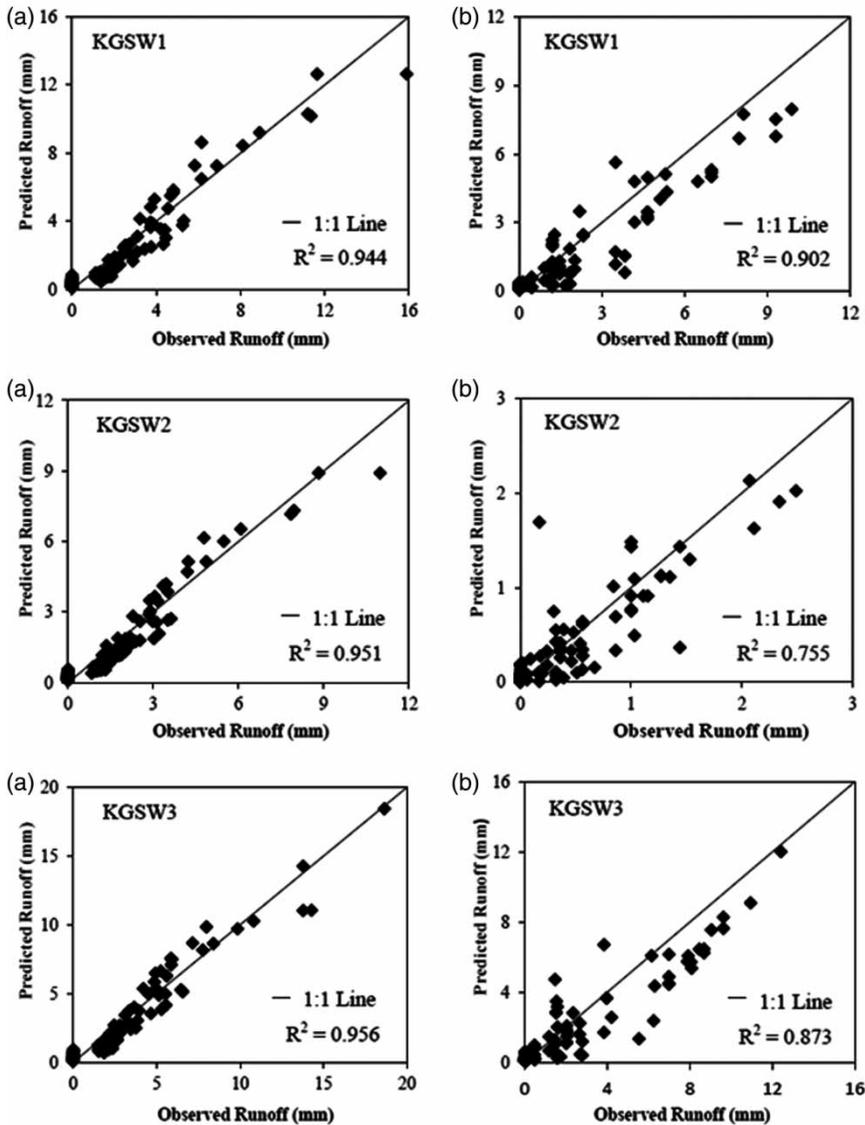
**Figure 4** | Scatter plots of observed versus ANN model predicted daily runoff during (a) training period (2004–05) and (b) testing period (2003) from Kapagari sub-watersheds 1, 2, and 3, respectively.

the BANN models had both underprediction and overprediction for the medium runoff values which are very well distributed about the 1:1 line than ANN predicted values. Statistical analysis was also performed for the BANN models' predicted runoff for the testing year 2003. The results of the BANN models and the corresponding ANN models in terms of *COE*, *RMSE*, coefficient of determination ($R^2$), and percent deviation (% $D_v$) for KGW and its sub-watersheds during the testing period are presented in Table 9. The higher *COE* of 0.902, 0.894, 0.771, and 0.907

by the BANN models as compared to the ANN models for Kapagari watershed and its sub-watersheds 1, 2, and 3, respectively, show a very good agreement between the daily observed runoff and the predicted runoff. The percent deviation (% $D_v$) values of 3.51, 7.09, 2.87, and 6.15 in the case of the BANN models for the Kapagari watershed and its sub-watersheds 1, 2, and 3, respectively, indicate that the magnitude of difference between the observed values and the predicted values is much less (less than 10%) as compared to that of the ANN models. The higher $R^2$ and

**Table 8** | Goodness-of-fit statistics of the observed and predicted daily runoff from Kapgari sub-watersheds

| Watershed | Statistical parameters | Runoff (mm) | | | |
| | | Training period (2004–2005) | | Testing period (2003) | |
| | | Observed | Simulated | Observed | Simulated |
|---|---|---|---|---|---|
| KGSW1 | Mean | 0.856 | 0.957 | 1.306 | 1.105 |
| | Standard deviation | 2.105 | 1.976 | 2.262 | 1.839 |
| | Maximum peak | 15.89 | 12.65 | 9.88 | 7.96 |
| | Total | 234.57 | 262.19 | 178.83 | 151.46 |
| | $COE$ | 0.941 | | 0.875 | |
| | $R^2$ | 0.944 | | 0.902 | |
| | $RMSE$ | 0.510 | | 0.796 | |
| | $MAE$ | 0.354 | | 0.489 | |
| | $D_v$ % | 11.77 | | 15.32 | |
| KGSW2 | Mean | 0.609 | 0.684 | 0.323 | 0.292 |
| | Standard deviation | 1.490 | 1.398 | 0.545 | 0.464 |
| | Maximum peak | 10.98 | 8.92 | 2.49 | 2.13 |
| | Total | 166.96 | 187.33 | 44.37 | 39.96 |
| | $COE$ | 0.947 | | 0.732 | |
| | $R^2$ | 0.951 | | 0.755 | |
| | $RMSE$ | 0.344 | | 0.281 | |
| | $MAE$ | 0.255 | | 0.139 | |
| | $D_v$ % | 12.20 | | 9.94 | |
| KGSW3 | Mean | 1.062 | 1.161 | 1.705 | 1.455 |
| | Standard deviation | 2.575 | 2.459 | 2.893 | 2.399 |
| | Maximum peak | 18.65 | 18.45 | 12.41 | 12.41 |
| | Total | 291.03 | 318.16 | 233.53 | 199.38 |
| | $COE$ | 0.954 | | 0.859 | |
| | $R^2$ | 0.956 | | 0.873 | |
| | $RMSE$ | 0.550 | | 1.082 | |
| | $MAE$ | 0.368 | | 0.641 | |
| | $D_v$ % | 9.32 | | 14.62 | |

lesser $RMSE$ values for the BANN models as compared to the ANN models also indicates that the use of ANN with bootstrapped resampling provides more accurate prediction of daily runoff. These results illustrate that the bootstrap resampling technique based artificial neural network (BANN) is more capable of solving the problems of overfitting and underfitting than 10-fold CV technique based ANN models during the training of the models with shorter length of data set. The BANN models were better than the 10-fold CV technique based ANN models because these models were developed with 50 resamples, out of which some models were trained with the higher peaks, some were trained with the lower peaks, and some were trained with both higher and lower peaks. The average prediction from all these 50 models increased the prediction efficiency. On the other hand, 10-fold CV technique based ANN models were trained using only one data set and simulation was made according to peaks presented in that data set.
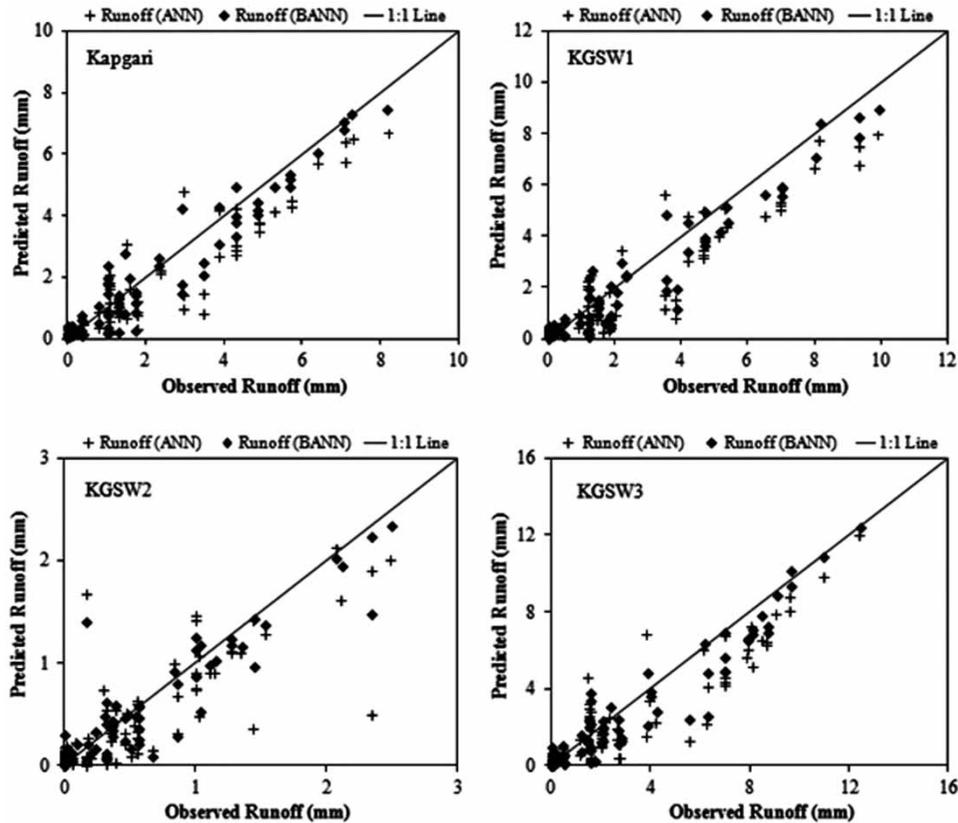
**Figure 5** │ Comparison of observed and predicted daily runoff of ANN and BANN models during testing period (2003) from KGW and its sub-watersheds 1, 2, and 3, respectively.

**Table 9** │ Comparison of the ANN and BANN models for prediction of the runoff from KGW and its sub-watersheds during the testing period 2003

| Watershed | COE | | RMSE (mm) | | $R^2$ | | % $D_v$ | |
|---|---|---|---|---|---|---|---|---|
| | ANN | BANN | ANN | BANN | ANN | BANN | ANN | BANN |
| KGW | 0.877 | 0.902 | 0.665 | 0.544 | 0.905 | 0.928 | 14.09 | 3.51 |
| KGSW1 | 0.875 | 0.894 | 0.796 | 0.697 | 0.902 | 0.912 | 15.32 | 7.09 |
| KGSW2 | 0.732 | 0.771 | 0.281 | 0.260 | 0.755 | 0.772 | 9.94 | 2.87 |
| KGSW3 | 0.859 | 0.907 | 1.082 | 0.877 | 0.873 | 0.913 | 14.62 | 6.15 |

## SUMMARY AND CONCLUSIONS

The major goal of the present study was to develop unbiased ANN models using readily available inputs and shorter length of training data sets for the prediction of daily runoff from a small agricultural watershed and its sub-watersheds in Eastern India. Only 3 years' (2003–2005) rainy season hydro-metrological data were available for the analysis. Two years' data were used for development of the ANN models and 1 year's data were used to test the developed models. In the reported study, ANN models were developed using the readily available climatic variables such as rainfall and temperature. To deal with the limitation of the ANN modeling with shorter length of training data set, a 10-fold CV method was used to avoid the neural network overfitting. It was done to minimize the bias while

testing with shorter length of data set using simple train-and-test (hold-out) technique. Several ANN models were developed for the watershed and its sub-watersheds. Out of the developed models, four ANN models KGM3 (4-3-1), KGSW1M3 (4-3-1), KGSW2M3 (4-4-1), and KGSW3M3 (4-3-1) considering both rainfall and temperature as input, were selected to predict the daily runoff for KGW and its sub-watersheds 1, 2, and 3, respectively. Biasness in the daily runoff prediction associated with the shorter length of training data set was also investigated using the bootstrap resampling technique BANN. Optimized numbers of inputs were used to develop 50 bootstrapped resamples and each resample was used to develop an ANN model to perform bootstrap aggregation of multi-model ensembles which produced averaged outputs. The approaches used in this study are of an objective nature and can be easily applied to other small agricultural watersheds under data scarcity conditions. The results of this study also have practical importance and wider applicability. The following specific conclusions were drawn from the study:

1. It was found that climatic variables such as daily rainfall and temperature, which are readily available, can be used for accurate prediction of the daily runoff from small agricultural watersheds using ANN modeling.

2. The ANN models constructed from climatic variables only will have the potential of filling missing data in a daily runoff time series and for predicting the influence of climatic change on runoff in data sparse regions.

3. It was revealed that the 10-fold CV technique can help in selecting the ANN structure, which provides unbiased estimation in the case of shorter length of training data set.

4. The bootstrapped resampling technique based ANN models can provide a more stable solution as compared to the 10-fold CV technique based ANN models and improve the prediction accuracy for the shorter length of training data set.

5. The peaks in the time series of the daily runoff data can be estimated more accurately using the bootstrap resampling technique based ANN models than that of 10-fold CV technique based ANN models for the shorter length of training data set.

## REFERENCES

Abrahart, R. J. 2003 Neural network rainfall-runoff forecasting based on continuous resampling. *J. Hydroinform.* 5, 51–61.

Akhtar, M. K., Corzo, G. A., Andell, S. J. & Jonoski, A. 2009 River flow forecasting with artificial neural networks using satellite observed precipitation pre-processed with flow length and travel time information: case study of the Ganges river basin. *Hydrol. Earth Syst. Sci.* 13, 1607–1618.

Aqil, M., Kita, I., Yano, A. & Nishiyama, S. 2007 A comparative study of artificial neural networks and neuro-fuzzy in continuous modeling of the daily and hourly behaviour of runoff. *J. Hydrol.* 337, 22–34.

ASCE Task Committee 1993 Criteria for evaluation of watershed models. *J. Irrig. Drain. Eng.* 119 (3), 429–442.

ASCE Task Committee 2000a Artificial neural networks in hydrology I: preliminary concepts. *J. Hydrol. Eng.* 5 (2), 115–123.

ASCE Task Committee 2000b Artificial neural networks in hydrology II: hydrologic applications. *J. Hydrol. Eng.* 5 (2), 124–137.

Barreto, H. & Howland, F. M. 2006 *Introductory Econometrics: Using Monte Carlo Simulation with Microsoft Excel.* Cambridge University Press, New York.

Bingner, R. L., Murphee, C. E. & Mutchler, C. K. 1989 Comparison of sediment yield models on various watersheds in Mississippi. *Am. Soc. Agric. Biol. Eng.* 32 (2), 529–534.

Bowden, G. J., Dandy, G. C. & Maier, H. R. 2005a Input determination for neural network models in water resources applications. Part 1 – Background and methodology. *J. Hydrol.* 301, 75–92.

Bowden, G. J., Dandy, G. C. & Maier, H. R. 2005b Input determination for neural network models in water resources applications. Part 2 – Forecasting salinity in a river. *J. Hydrol.* 301, 93–107.

Bray, M. & Han, D. 2004 Identification of support vector machines for runoff modeling. *J. Hydroinform.* 6 (4), 265–280.

Campolo, M., Andrreussi, P. & Soldati, A. 1999 River flood forecasting with a neural network model. *Water Resour. Res.* 35 (4), 1191–1197.

Chang, F. J. & Chen, Y. C. 2001 A counter-propagation fuzzy-neural network modeling approach to real time stream flow prediction. *J. Hydrol.* 245, 153–164.

Chen, S. M., Wang, Y. M. & Tsou, I. 2013 Using artificial neural network approach for modeling rainfall–runoff due to typhoon. *J. Earth Syst. Sci.* 122 (2), 399–405.

Chernick, M. R. & LaBudde, R. 2010 Revisiting qualms about bootstrap confidence intervals. *Am. J. Math. Manag. Sci.* 29, 437–456.

Cigizoglu, H. K. 2003 Estimation, forecasting and extrapolation of river flows by artificial neural networks. *Hydrolog. Sci.* 48 (3), 349–361.

Dawson, C. W. & Wilby, R. 1998 An artificial neural network approach to rainfall modeling. *Hydrol. Sci. J.* **43** (1), 47–66.

Dawson, C. W. & Wilby, R. L. 2001 Hydrological modeling using artificial neural network. *Prog. Phys. Geogr.* **25** (1), 80–108.

Demuth, H. & Beale, M. 1998 *Neural Network Toolbox for Use with MATLAB, Users Guide*. The MathWorks Inc., Natick, MA, USA.

Dooge, J. C. I. 1977 Problems and methods of rainfall–runoff modeling. In: *Mathematical Models for Surface Water Hydrology* (T. A. Ciriani, U. Maione & J. R. Wallis, eds). Wiley, London, UK, pp. 71–108.

Duan, Q., Sorooshian, S. & Gupta, V. K. 1992 Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resour. Res.* **28**, 1015–1031.

Efron, B. 1979 Bootstrap methods: another look at the jackknife. *Ann. Stat.* **7**, 1–26.

Efron, B. & Tibshirani, R. J. 1993 *An Introduction to the Bootstrap*. Chapman and Hall, London, UK.

Fausett, L. 1994 *Fundamentals of Neural Networks Architectures, Algorithms, and Applications*. Prentice-Hall, Upper Saddle River, NJ, USA.

French, M., Krajewski, W. F. & Cuykendall, R. R. 1992 Rainfall forecasting in space and time using a neural network. *J. Hydrol.* **137**, 1–31.

Friedman, J. H. 1991 Multivariate adaptive regression splines. *Ann. Stat.* **19**, 1–141.

Gautam, M. R., Watanabe, K. & Saegusa, H. 2000 Runoff analysis in humid forest catchment with artificial neural network. *J. Hydrol.* **235**, 117–136.

Geisser, S. 1975 The predictive sample reuses method with applications. *J. Am. Stat. Assoc.* **70** (350), 320–328.

Hagan, M. T. & Menhaj, M. B. 1994 Training feed forward networks with the Marquardt algorithm. *IEEE Trans. Neural Netw.* **5** (6), 989–993.

Halff, A. H., Halff, H. M. & Azmoodeh, M. 1993 Predicting runoff from rainfall using neural networks. In: *Proceedings of the Symposium sponsored by the Hydraulics Division of ASCE, Engineering Hydrology*, July 25–30, San Francisco, CA, pp. 768–775.

Hall, M. J. & Minns, A. W. 1993 Rainfall-runoff modeling as a problem in artificial intelligence: experience with neural network. In: *Proceedings of the 4th British Hydrological Society Symposium*, Cardiff, UK 5, pp. 51–57.

Haykin, S. 1994 *Neural Networks: A Comprehensive Foundation*. Macmillan College Publishing, New York, USA.

Hecht-Nielsen, R. 1990 *Neurocomputing*. Addison-Wesley, Reading, MA, USA, pp. 115–117.

Hettiarachchi, P., Hall, M. J. & Minns, A. W. 2005 The extrapolation of artificial neural networks for the modelling of rainfall–runoff relationships. *J. Hydroinform.* **7**, 291–296.

Hsieh, W. W. & Tang, B. 1998 Applying neural network models to prediction and data analysis in Meteorology and Oceanography. *Bull. Amer. Meteorol. Soc.* **79**, 1855–1870.

Hsu, K. V., Gupta, H. & Sorooshian, S. 1995 Artificial neural network modeling of the rainfall-runoff process. *Water Resour. Res.* **31** (10), 2517–2530.

Imrie, C. E., Durucan, S. & Korre, A. 2000 River flow prediction using artificial neural networks: Generalization beyond the calibration range. *J. Hydrol.* **33**, 138–153.

Jacquin, A. P. & Shamseldin, A. Y. 2009 Review of the application of fuzzy logic in river flow forecasting. *J. Hydroinform.* **11**, 202–210.

Jeong, D. & Kim, Y. O. 2005 Rainfall-runoff models using artificial neural networks for ensemble streamflow prediction. *Hydrol. Process.* **19**, 3819–3835.

Karunanithi, N., Grenney, W. J., Whitley, D. & Bovee, K. 1994 Neural networks for river flow prediction. *J. Comput. Civil Eng. ASCE* **8** (2), 201–220.

Maier, H. & Dandy, G. C. 2000 Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications. *Environ. Model. Softw.* **15**, 101–124.

Mason, J. C., Price, R. K. & Temme, A. 1996 A neural network model of rainfall-runoff using radial basis functions. *J. Hydraul. Res.* **34** (4), 537–548.

Masters, T. 1995 *Advanced Algorithms for Neural Networks: a C++ Sourcebook*. Wiley, New York, USA.

Minns, A. W. & Hall, M. J. 1996 Artificial neural networks as rainfall runoff models. *Hydrol. Sci. J.* **41** (3), 399–417.

Molinaro, A. M., Simon, R. & Pfeiffer, R. M. 2005 Prediction error estimation: a comparison of resampling methods. *Bioinformatics* **21** (15), 3301–3307.

Mukerji, A., Chatterjee, C. & Raghuwanshi, N. S. 2009 Flood forcasting using ANN, Neuro-Fuzzy and Neuro-GA models. *J. Hydrol. Eng.* **14** (6), 647–652.

Muller, B. & Reinhardt, J. 1991 *Neural Networks: An Introduction*. Springer-Verlag, Berlin, Germany.

Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models. Part 1 – A: discussion principles. *J. Hydrol.* **10**, 282–290.

Navone, H. D. & Ceccatto, H. A. 1994 Predicting Indian monsoon rainfall: a neural network approach. *Clim. Dyn.* **10**, 305–312.

Nayak, P. C., Sudheer, K. P., Rangan, D. M. & Ramasastri, K. S. 2005 Short-term flood forecasting with a neurofuzzy model. *Water Resour. Res.* **41**, WR003562.

Patterson, D. 1996 *Artificial Neural Networks*. Prentice-Hall, Singapore.

Pramanik, N. & Panda, R. K. 2009 Application of neural network and adaptive neuro-fuzzy inference systems for river flow prediction. *Hydrolog. Sci. J.* **54** (2), 247–260.

Quinlan, J. R. 1992 Learning with continuous classes. In: *Proceedings AI'92 World Scientific*, Singapore, pp. 343–348.

Raghuwanshi, N. S., Singh, R. & Reddy, L. S. 2006 Runoff and sediment yield modeling using artificial neural networks: Upper Siwane River, India. *J. Hydrol. Eng. ASCE* **11** (1), 71–79.

Rumelhart, D. E. & McClelland, J. L. 1986 *Parallel Distributed Processing*. MIT Press, Cambridge, MA, USA.

Sajikumar, N. & Thandaveswara, B. S. 1999 A nonlinear rainfall runoff model using artificial neural network. *J. Hydrol.* **216**, 32–55.

Sarangi, A., Madramootoo, C. A., Enright, P., Prasher, S. O. & Patel, R. M. 2005 Performance evaluation of ANN and geomorphology-based models for runoff and sediment yield prediction for a Canadian watershed. *Current Sci.* **89** (12), 2022–2033.

See, L. & Openshaw, S. 1999 Applying soft computing approaches to river level forecasting. *Hydrolog. Sci. J.* **44** (5), 763–779.

Shamseldin, A. Y. 1997 Application of a neural network technique to rainfall-runoff modeling. *J. Hydrol.* **199**, 272–294.

Sharda, V. N., Patel, R. M., Prasher, S. O., Ojasvi, P. R. & Chandra, P. 2006 Modeling runoff from middle Himalayan watersheds employing artificial intelligence techniques. *Agric. Water Manage.* **83**, 233–242.

Sharma, S. K. & Tiwari, K. N. 2009 Bootstrap based artificial neural network (BANN) analysis for hierarchical prediction of monthly runoff in Upper Damodar Valley catchment. *J. Hydrol.* **374**, 209–222.

Shu, C. & Ouarda, T. B. M. J. 2007 Flood frequency analysis at ungauged sites using artificial neural networks in canonical correlation analysis physiographic space. *Water Resour. Res.* **43** (7), W07438.

Smith, J. & Eli, N. 1995 Neural network models of rainfall-runoff process. *J. Water Resour. Plan. Manage.* **121** (6), 499–508.

Solomatine, D. P. & Dulal, K. N. 2003 Model trees as an alternative to neural networks in rainfall-runoff modeling. *Hydrolog. Sci. J.* **48** (3), 399–411.

Stone, M. 1974 Cross-validatory choice and assessment of statistical predictions. *J. Roy. Stat. Soc.* **36** (2), 111–147.

Tiwari, M. K. & Chatterjee, C. 2010 Uncertainty assessment and ensemble flood forecasting using bootstrap based artificial neural networks (BANNs). *J. Hydrol.* **382**, 20–33.

Tokar, A. S. & Johnson, P. A. 1999 Rainfall-runoff modeling using artificial neural networks. *J. Hydrol. Eng. ASCE* **4** (3), 232–239.

Toth, E. & Brath, A. 2007 Multistep ahead stream flow forecasting: Role of calibration data in conceptual and neural network modeling. *Water Resour. Res.* **43** (11), W11405.

Wang, Y. M., Traore, S., Kerh, T. & Leu, J. M. 2011 Modeling reference evapotranspiration using feed forward back propagation algorithm in arid regions of Africa. *Irrig. Drain.* **60** (3), 404–417.

Weiss, S. M. & Kulikowski, C. A. 1991 *Computer Systems That Learn.* Kaufmann Publishers, San Mateo, CA, USA.

World Meteorological Organization (WMO) 1975 Intercomparison of Conceptual Models used in Operational Hydrological Forecasting. Technical Report No. 429. Geneva, Switzerland.

Wu, J. M., Han, J., Shastri, A. & Scott, B. 2005 Artificial neural networks for forecasting watershed runoff and streamflows. *J. Hydrol. Eng. ASCE* **10** (3), 216–222.

Zhang, B. & Govindaraju, R. S. 2000 Prediction of watershed runoff using Bayesian concepts and modular neural networks. *Water Resour. Res.* **36** (3), 753–762.

Zhang, B. & Govindaraju, R. S. 2003 Geomorphology-based artificial neural networks (GANNs) for estimation of direction runoff over watersheds. *J. Hydrol.* **273**, 18–34.