Monthly Rainfall Prediction Using Wavelet Neural Network Analysis

**R. Venkata Ramana · B. Krishna · S. R. Kumar ·
N. G. Pandey**

Abstract Rainfall is one of the most significant parameters in a hydrological model. Several models have been developed to analyze and predict the rainfall forecast. In recent years, wavelet techniques have been widely applied to various water resources research because of their time-frequency representation. In this paper an attempt has been made to find an alternative method for rainfall prediction by combining the wavelet technique with Artificial Neural Network (ANN). The wavelet and ANN models have been applied to monthly rainfall data of Darjeeling rain gauge station. The calibration and validation performance of the models is evaluated with appropriate statistical methods. The results of monthly rainfall series modeling indicate that the performances of wavelet neural network models are more effective than the ANN models.

1 Introduction

Rainfall is a complex atmospheric process, which is space and time dependent and it is not easy to predict. Due to the apparent random characteristics of rainfall series, they are often described by a stochastic process (Chinchorkar et al. 2012). For water resources planning purposes, a long-term rainfall series is required in hydrological and simulation models (Tantanee et al. 2005). There have been many attempts to find the most appropriate method for rainfall prediction for example, coupling physical, marine, and meteorological or satellite data with a forecasting model, or even

R. Venkata Ramana (✉) · S. R. Kumar · N. G. Pandey
Centre for Flood Management Studies, National Institute of Hydrology, Patna 801 505, Bihar, India
e-mail: venkataramana_1973@yahoo.co.in

S. R. Kumar
e-mail: srk9966@gmail.com

N. G. Pandey
e-mail: ngpandey@sify.com

B. Krishna
Deltaic Regional Center, National Institute of Hydrology, Kakinada 533 003, Andhra Pradesh, India
e-mail: krishna_nih@rediffmail.com

applying several techniques such as the artificial neural network or fuzzy logic as a forecasting approach (Hsu *et al*. 1995; Dawson and Wilby 2001; Hettiarachchi et al. 2005). In recent years, several numerical weather forecasts have been proposed for weather prediction but most of these models are limited to short period forecasts. This paper introduces a new approach for prediction of rainfall series.

Several time series models have been proposed for modeling monthly rainfall series (Bhakar et al. 2006) and annual rainfall series such as the autoregressive model (AR) (Yevjevich 1972), the fractional Guassian noise model (Matalas and Wallis 1971), autoregressive moving-average models (ARMA) (Carlson et al. 1970) and the disaggregation multivariate model (Valencia and Schaake 1973). Moustris et al. (2011) examine the possibility of long term precipitation forecast (four consecutive months) by the application of ANNs, using long monthly precipitation time series of four meteorological stations in Greece.

In the past decade, wavelet theory has been introduced to signal processing analysis. In recent years, the wavelet transform has been successfully applied to wave data analysis and other ocean engineering applications (Massel 2001; Teisseire et al. 2002; Huang 2004). The time-frequency character of long-term climatic data is investigated using the continuous wavelet transform technique (Lau and Weng 1995; Torrence and Compo 1997; Mallat 1998) and wavelet analysis of wind wave measurements obtained from a coastal observation tower (Huang 2004). Chou (2011) used wavelet denoising method in linear perturbation models (LPMs) and simple linear models (SLMs) for rainfall and runoff time series data. Wang and Li (2011) used a new wavelet transform method for developing the synthetic generation of daily stream flow sequences. Wu et al. (2004) used a combination of neural networks and wavelet methods to predict underground water levels.

Dynamical Recurrent Neural Network (DRNN) on each resolution scale of the sunspot time series resulting from the wavelet decomposed series with the Temporal Recurrent Back propagation (TRBP) algorithm (Aussem and Murtagh 1997). There are some appreciable studies of wavelet transform based neural network models (Wang and Ding 2003; Anctil and Tape 2004; Cannas et al. 2006; Kisi 2008; Wang et al. 2009). The wavelet transform is also integrated with multiple linear regression (Kucuk and Agiralioğlu 2006; Kisi 2009, 2010) and support vector machine approach (Kisi and Cimen 2011). Adamowski and Sun (2010) compared the relative performance of the coupled wavelet-neural network models (WA–ANN) and regular artificial neural networks (ANN) for flow forecasting at lead times of 1 and 3 days for two different non-perennial rivers in semiarid watersheds of Cyprus. Kisi (2011) investigated the performance of the wavelet regression (WR) technique in daily river stage forecasting and determined the WR model was improved combining two methods, discrete wavelet transform and a linear regression model. Sang (2013), developed a method for discrete wavelet decomposition and an improved wavelet modeling framework, WMF for short was proposed for hydrologic time series forecasting. By coupling the wavelet method with the traditional AR model, the Wavelet-Autoregressive model (WARM) is developed for annual rainfall prediction (Tantanee et al. 2005). Partal and Kisi (2007) used a conjunction model (wavelet-neuro-fuzzy) to forecast the Turkey daily precipitation. The observed daily precipitations are decomposed to some sub series by using Discrete Wavelet Transform (DWT) and then appropriate sub series are used as inputs to neuro-fuzzy models for forecasting of daily precipitations. Each of these studies showed that different black box models trained or calibrated with decomposed data resulted in higher accuracy than the single models that were calibrated with an undecomposed and noisy time series. In this paper, a Wavelet

Neural Network (WNN) model, which is the combination of wavelet analysis and ANN, has been proposed for rainfall forecast Darjeeling station, India.

2 Wavelet Analysis

The wavelet analysis is an advanced tool in signal processing that has attracted much attention since its theoretical development (Grossmann and Morlet 1984). Its use has increased rapidly in communications, image processing and optical engineering applications as an alternative to the Fourier transform in preserving local, non-periodic and multiscaled phenomena. The difference between wavelets and Fourier transforms is that wavelets can provide the exact locality of any changes in the dynamical patterns of the sequence, whereas the Fourier transforms concentrate mainly on their frequency. Moreover, Fourier transform assume infinite length signals, whereas wavelet transforms can be applied to any kind and any size of time series, even when these sequences are not homogeneously sampled in time (Antonios and Constantine 2003). In general, wavelet transforms can be used to explore, denoise and smoothen time series which aid in forecasting and other empirical analysis.

Wavelet analysis is the breaking up of a signal into shifted and scaled versions of the original (or mother) wavelet. In wavelet analysis, the use of a fully scalable modulated window solves the signal-cutting problem. The window is shifted along the signal and for every position the spectrum is calculated. Then this process is repeated many times with a slightly shorter (or longer) window for every new cycle. In the end, the result will be a collection of time-frequency representations of the signal, all with different resolutions. Because of this collection of representations we can speak of a multiresolution analysis. By decomposing a time series into time-frequency-space, one is able to determine both the dominant modes of variability and how those modes vary in time. Wavelets have proven to be a powerful tool for the analysis and synthesis of data from long memory processes. Wavelets are strongly connected to such processes in that the same shapes repeat at different orders of magnitude. The ability of the wavelets to simultaneously localize a process in time and scale domain results in representing many dense matrices in a sparse form.

2.1 Discrete Wavelet Transform (DWT)

The basic aim of wavelet analysis is to determine the frequency (or scale) content of a signal and then it assess and determine the temporal variation of this frequency content. This property is in complete contrast to the Fourier analysis, which allows for the determination of the frequency content of a signal but fails to determine frequency-time dependence. Therefore, the wavelet transform is the tool of choice when signals are characterized by localized high frequency events or when signals are characterized by a large numbers of scale variable processes. Because of its localization properties in both time and scale, the wavelet transform allows for tracking the time evolution of processes at different scales in the signal.

The wavelet transform of a time series $f(t)$ is defined as

$$f(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t)\varphi\left(\frac{t-b}{a}\right)dt \qquad (1)$$

where $\varphi(t)$ is the basic wavelet with effective length (t) that is usually much shorter than the target time series $f(t)$. The variables 'a' is the scale or dilation factor that determines the characteristic

frequency so that its variation gives rise to a spectrum and 'b' is the translation in time so that its variation represents the 'sliding' of the wavelet over $f(t)$. The wavelet spectrum is thus customarily displayed in time-frequency domain. For low scales i.e. when $|a|<<1$, the wavelet function is highly concentrated (shrunken compressed) with frequency contents mostly in the higher frequency bands. Inversely, when $|a|>>1$, the wavelet is stretched and contains mostly low frequencies. For small scales, we obtain thus a more detailed view of the signal (also known as a "higher resolution") whereas for larger scales we obtain a more general view of the signal structure.

The original signal $X(n)$ (Fig. 1) passes through two complementary filters (low pass and high pass filters) and emerges as two signals as Approximations (A) and Details (D). The approximations are the high-scale, low frequency components of the signal. The details are the low-scale, high frequency components. Normally, the low frequency content of the signal (approximation, A) is the most important part. It demonstrates the signal identity. The high-frequency component (detail, D) is nuance. The decomposition process can be iterated, with successive approximations being decomposed in turn, so that one signal is broken down into many lower resolution components (Fig. 1).

2.2 Mother Wavelet

The choice of the mother wavelet depends on the data to be analyzed. The Daubechies and Morlet wavelet transforms are the commonly used "Mother" wavelets. Daubechies wavelets exhibit good trade-off between parsimony and information richness, it produces the identical events across the observed time series and appears in so many different fashions that most prediction models are unable to recognize them well (Benaouda et al. 2006). Morlet wavelets, on the other hand, have a more consistent response to similar events but have the weakness of generating many more inputs than the Daubechies wavelets for the prediction models.

3 Artificial Neural Networks

An ANN, can be defined as a system or mathematical model consisting of many nonlinear artificial neurons running in parallel, which can be generated, as one or multiple layered. Although the concept of artificial neurons was first introduced by McCulloch and Pitts, the major applications of ANN's have arisen only since the development of the back-propagation method of training by Rumelhart (Rumelhart et al. 1986). Following this development, ANN research has resulted in the successful solution of some complicated problems not easily solved by traditional modeling methods when the quality/quantity of data is very limited. ANN models are 'black box' models with particular properties, which are greatly suited to dynamic nonlinear
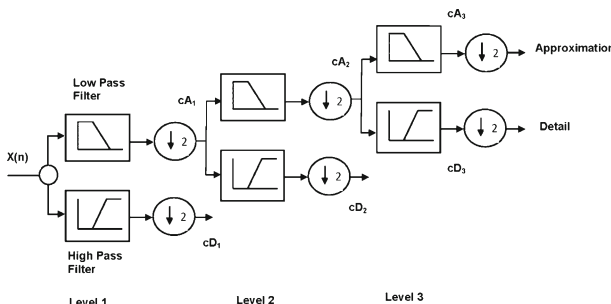

Fig. 1 Diagram of multiresolution analysis of signal

system modeling. The main advantage of this approach over traditional methods is that it does not require the complex nature of the underlying process under consideration to be explicitly described in mathematical form. ANN applications in hydrology vary, from real time to event based modeling.

The most popular ANN architecture in hydrologic modeling is the multilayer perceptron (MLP) trained with BP algorithm (ASCE(2000a, b)). A multilayer perceptron network consists of an input layer, one or more hidden layers of computation nodes, and an output layer. The number of input and output nodes is determined by the nature of the actual input and output variables. The number of hidden nodes, however, depends on the complexity of the mathematical nature of the problem, and is determined by the modeler, often by trial and error. The input signal propagates through the network in a forward direction, layer by layer. Each hidden and output node processes its input by multiplying each of its input values by a weight, summing the product and then passing the sum through a nonlinear transfer function to produce a result. For the training process, where weights are selected, the neural network uses the gradient descent method to modify the randomly selected weights of the nodes in response to the errors between the actual output values and the target values. This process is referred to as training or learning. It stops when the errors are minimized or another stopping criterion is met. The Back Propagation Neural Network (BPNN) can be expressed as

$$Y = f\left(\sum WX - \theta\right) \tag{2}$$

where X = input or hidden node value; Y = output value of the hidden or output node; $f(.)$ = transfer function; W = weights connecting the input to hidden, or hidden to output nodes; and θ = bias (or threshold) for each node.

3.1 Method of Network Training

Levenberg-Marquardt method (LM) was used for training of the given network. It is a modification of the classic Newton algorithm for finding an optimum solution to a minimization problem. In practice, LM is faster and finds better optima for a variety of problems than most other methods (Hagan and Menhaj 1994). The method also takes advantage of the internal recurrence to dynamically incorporate past experience in the training process (Coulibaly et al. 2000).

The Levenberg-Marquardt Algorithm (LMA) is given by

$$X_{k+1} = X_k - \left(J^T J + \mu I\right)^{-1} J^t e \tag{3}$$

where, X is the weights of neural network, J are the Jacobian matrix of the performance criteria to be minimized, 'μ' is a learning rate that controls the learning process and 'e' is residual error vector. If scalar μ is very large, the above expression approximates gradient descent with a small step size, while if it is very small; the above expression becomes Gauss-Newton method using the approximate Hessian matrix. The Gauss-Newton method is faster and more accurate near an error minimum. Hence we decrease μ after each successful step and increase only when a step increases the error. LMA has great computational and memory requirements, and thus it can only be used in small networks. It is faster and less easily trapped in local minima than other optimization algorithms.

3.2 Selection of Network Architecture

Increasing the number of training patterns provide more information about the shape of the solution surface, and thus increases the potential level of accuracy that can be achieved by the network. A large training pattern set, however can sometimes overwhelm certain training algorithms, thereby increasing the likelihood of an algorithm becoming stuck in a local error minimum. Consequently, there is no guarantee that adding more training patterns leads to improve solution. Moreover, there is a limit to the amount of information that can be modeled by a network that comprises a fixed number of hidden neurons. The time required to train a network increases with the number of patterns in the training set. The critical aspect is the choice of the number of nodes in the hidden layer and hence the number of connection weights.

Based on the physical knowledge of the problem and statistical analysis, different combinations of antecedent values of the time series were considered as input nodes. The output node is the time series data to be predicted in one step ahead. Time series data was standardized for zero mean and unit variation, and then normalized into 0 to 1. The activation function used for the hidden and output layer was logarithmic sigmoidal and pure linear function respectively. For deciding the optimal hidden neurons, a trial and error procedure started with two hidden neurons initially, and the number of hidden neurons was increased up to 10 with a step size of 1 in each trial. For each set of hidden neurons, the network was trained in batch mode to minimize the mean square error at the output layer. In order to check any over fitting during training, a cross validation was performed by keeping track of the efficiency of the fitted model. The training was stopped when there was no significant improvement in the efficiency, and the model was then tested for its generalization properties. Figure 2 shows the multilayer perceptron (MLP) neural network architecture when the original signal taken as input of the neural network architecture.

3.3 Method of Combining Wavelet Analysis with ANN

The decomposed details (D) and approximation (A) were taken as inputs to neural network structure as shown in Fig. 3, where 'i' is the level of decomposition varying from 1 to I and j is the number of antecedent values varying from 0 to J and N is the length of the time series.
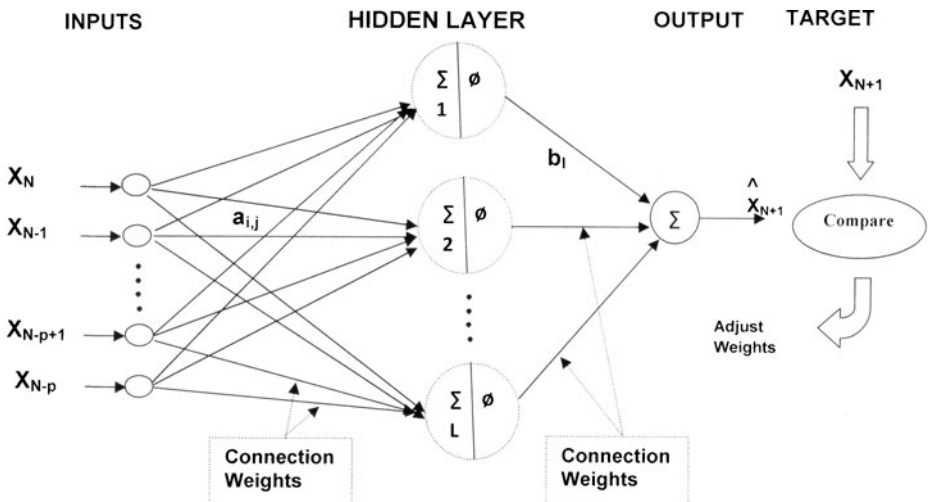

Fig. 2 Signal data based multilayer perceptron (MLP) neural network structure

To obtain the optimal weights (parameters) of the neural network structure, LM back propagation algorithm was used to train the network. A standard MLP with a logarithmic sigmoidal transfer function for the hiddenlayer and linear transfer function for the output layer were used in the analysis. The number of hidden nodes was determined by trial and error procedure. The output node will be the original value at one step ahead.

4 Linear Auto-Regressive (AR) Modeling

A common approach for modeling univariate time series is the autoregressive (AR) model:

$$X_t = \delta + \phi_1 X_{t-1} + \phi_2 X_{2-1} + \cdots\cdots + \phi_2 X_{2-1} + A_t \tag{4}$$

where X_t is the time series, A_t is white noise, and

$$\delta = \left(1 - \sum_{i-1}^{p} \phi_i\right)\mu \tag{5}$$

where 'μ' is the mean of the time series. An autoregressive model is simply a linear regression of the current value of the series against one or more prior values. AR models can be analyzed with linear least squares technique. They also have a straightforward interpretation. The determination of the model order can be estimated by examining the plots of Auto Correlation Function (ACF) and Partial Auto Correlation Function (PACF). The number of non-zero terms (i.e. outside confidence bands) in PACF suggests the order of the AR model. An AR (k) model will be implied by a sample PACF with k non-zero terms, and the terms in the sample ACF will decay slowly towards zero. From ACF and PACF analysis for rainfall, the order of the AR model is selected as 1.
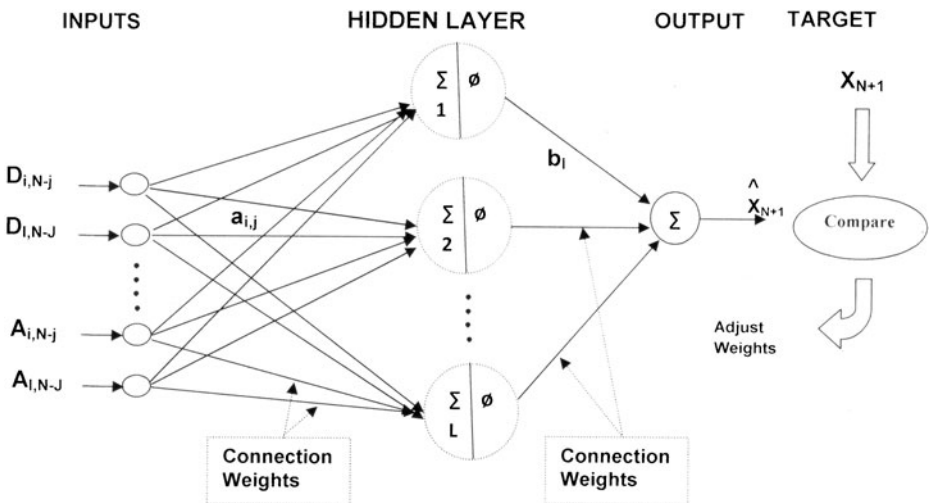

Fig. 3 Wavelet based multilayer perceptron (MLP) neural network structure

5 Performance Criteria

The performance of various models during calibration and validation were evaluated by using the statistical indices: the Root Mean Squared Error (RMSE), Correlation Coefficient (R) and Coefficient of Efficiency (COE).

The definitions of different statistical indices are presented below:

$$\text{RMSE} = \sqrt{\frac{\left[\sum_{t=1}^{N}\left(y_t^o - y_t^c\right)^2\right]}{N}} \tag{6}$$

$$R = \frac{\sum_{i=1}^{N}\left(y_t^o - \overline{y^o}\right)\left(y_t^c - \overline{y^c}\right)}{\sqrt{\left[\sum_{t=1}^{N}\left(y_t^o - \overline{y^o}\right)^2\right]\left[\sum_{t=1}^{N}\left(y_t^c - \overline{y^c}\right)^2\right]}} \tag{7}$$

$$\text{COE} = 1 - \frac{\left[\sum_{t=1}^{N}\left(y_t^o - y_t^c\right)^2\right]}{\left[\sum_{i=1}^{N}\left(y_t^o - \overline{y^o}\right)^2\right]} \tag{8}$$

where y_t^o and y_t^c are the observed and calculated values at time t respectively, $\overline{y^o}$ and $\overline{y^c}$ are the mean of the observed and calculated values.

6 Study Area and Data Collection

Darjeeling is a small town in the Himalayan foot hills (Fig. 4), lying at an altitude of 2,130 m above mean sea level and known as the Queen of the Hills. It is located in Darjeeling District in the extreme north of West Bengal State in the east of India. The district extends from tropical plains at about 91 m above sea level to an altitude of 3,658 m on the Sandakfu-Phalut ridge. The hills of Darjeeling are part of the Mahabharat Range or Lesser Himalaya. Kanchenjunga, the world's third-highest peak, 8,598 m (28,209 ft) high, is the most prominent mountain visible from Darjeeling.

To forecast the rainfall at Darjeeling rain gauge station (Fig. 4), monthly rainfall, minimum and maximum temperature data of 74 years from 01 January 1901 to 01 September 1975 was used. The first 44 years (60 % data) from 01 January 1901 to 01 November 1945 was used for calibration of the model, and the remaining 26 years (40 % data) from 01 December 1945 to 01 September 1975 data were used for validation. The annual mean maximum temperature is 14.9 °C while the mean minimum temperature is 8.9 °C, with monthly mean temperatures range from 5–17 °C. The lowest temperature recorded was −3.7 °C in the month of February 1905. The average annual precipitation is 3,092 mm, with an average of 126 days of rain in a year. The highest rainfall occurs in July.
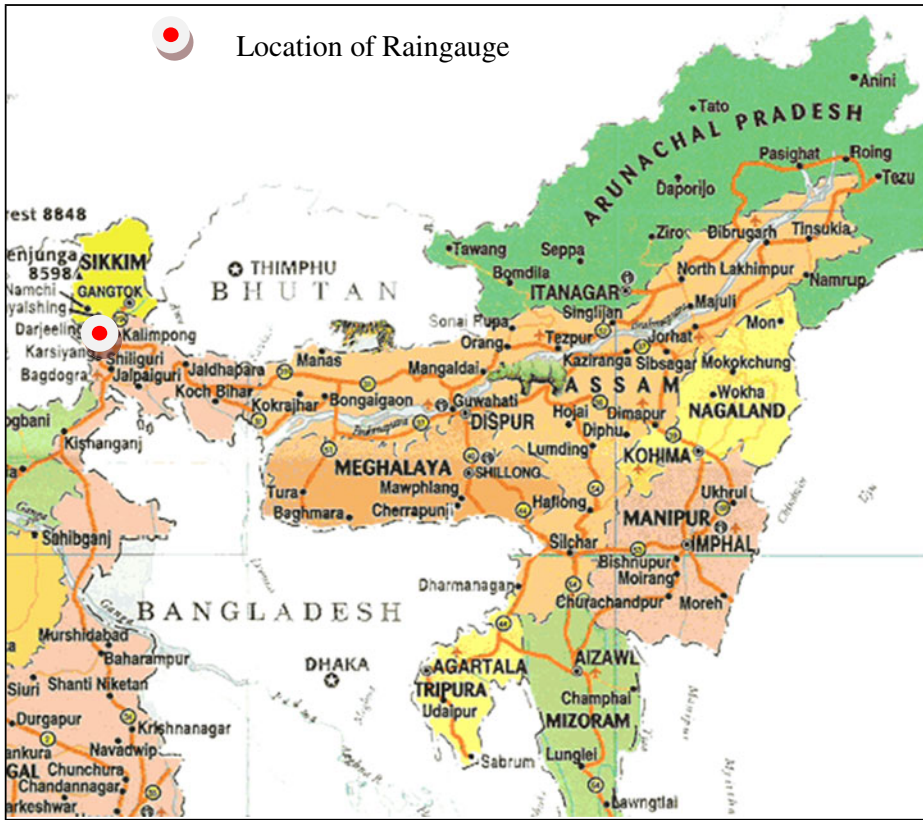
Fig. 4 Location of the study area

Table 1 shows the statistics of the calibration, validation and total data for monthly rainfall, minimum temperature and maximum temperature. The recorded monthly maximum rainfall is 1,401 mm, while the mean monthly rainfall was 2,350 mm. The observed monthly rainfall shows low positive skewness (1.2) and indicates that the data has less scattered distribution. The minimum of the maximum temperature is 5.5 °C, while the maximum of the maximum temperature is 21.5 °C.

Table 1 Statistical properties of rainfall, minimum and maximum temperatures

Variable	Data	Minimum	Maximum	Mean	Standard Deviation	Skewness
Rainfall (mm)	Total	0.00	1401.8	235.0	284.42	1.20
	Calibration	0.00	1231.3	240.1	286.61	1.19
	Validation	0.00	1401.8	226.5	276.10	1.23
Minimum temperature (°C)	Total data	−3.70	15.8	8.9	4.67	−0.29
	Calibration	−1.70	15.4	8.9	4.66	−0.27
	Validation	−3.70	15.8	9.0	4.68	−0.33
Maximum temperature (°C)	Total	5.50	21.5	15.6	3.95	−0.65
	Calibration	5.50	21.5	15.3	4.04	−0.57
	Validation	6.10	21.4	16.2	3.74	−0.75

6.1 Development of Wavelet Neural Network Model

The original time series was decomposed into Details and Approximations to certain number of sub-time series $\{D_1, D_{2....} D_p, A_p\}$ by wavelet transform algorithm. These play different role in the original time series and the behavior of each sub-time series is distinct (Wang and Ding 2003). So the contribution to original time series varies from each successive approximations being decomposed in turn, so that one signal is broken down into many lower resolution components, tested using different scales from 1 to 10 with different sliding window amplitudes. In this context, dealing with a very irregular signal shape, an irregular wavelet, the Daubechies wavelet of order 5 (DB5), has been used at level 3. Consequently, D_1, D_2, D_3 were detail time series, and A_3 was the approximation time series. The decomposed sub series of details and approximations along with original series for rainfall was shown in Fig. 5.

An ANN was constructed in which the sub-series $\{D_1, D_2, D_3, A_3\}$ at time t are input of ANN and the original time series at $t+T$ time are output of ANN, where T is the length of time to forecast. The input nodes for the WNN are the decomposed subsets of antecedent values of the rainfall, minimum and maximum temperatures and were presented in Table 2. The Wavelet Neural Network model (WNN) was formed in which the weights are learned with Feed forward neural network with Back Propagation algorithm. The number of hidden neurons for BPNN was determined by trial and error procedure.

7 Results and Discussion

To forecast the rainfall at Darjeeling rain gauge station, the monthly rainfall, is minimum temperature and maximum temperature data of 74 years was used. The first 44 years data were used for calibration of the model, and the remaining 26 years data were used for validation. The
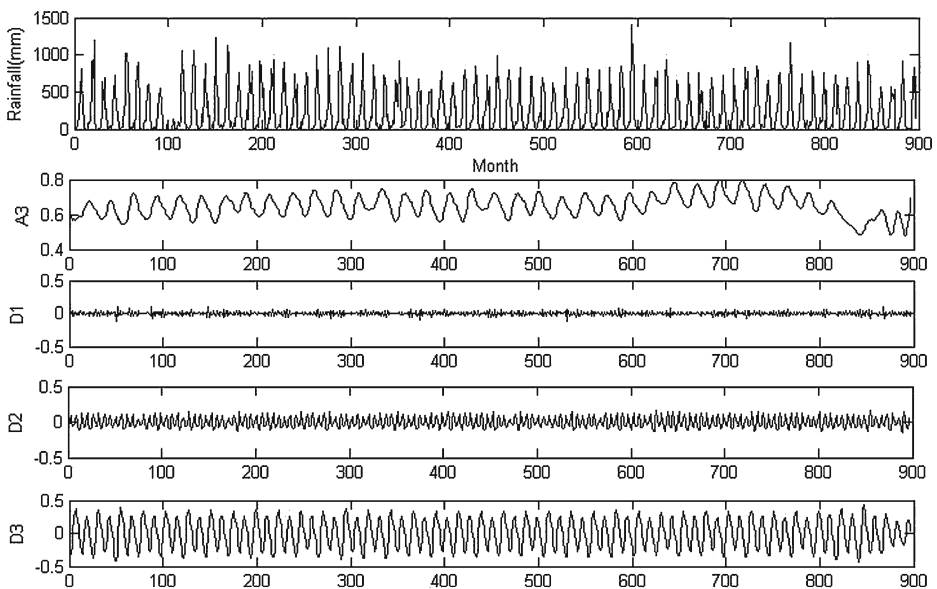

Fig. 5 Decomposed wavelet sub-time series of rainfall from 1901 to 1974

Table 2 Model Inputs for ANN

Model	Input Variables
I	R(t)=f (R [t-1])
II	R(t)=f (R [t-1], R [t-2])
III	R(t)=f (R [t-1], R [t-2], TM [t-1])
IV	R(t)=f (R [t-1], R [t-2], TM [t-1], TX [t-1])
V	R(t)=f (R [t-1], R [t-2], TM [t-1], TM [t-2], TX [t-1])
VI	R(t)=f (R [t-1], R [t-2], TM [t-1], TM [t-2], TX [t-1], TX [t-2])

R rainfall, *TM* minimum temperature, *TX* maximum temperature

model inputs (Table 2) were decomposed by wavelets and decomposed sub-series were taken as input to ANN. ANN was trained using back propagation with LM algorithm. The optimal number of hidden neurons for WNN and ANN was determined as six by trial and error procedure. The final structure of the WNN model is the number of inputs used in ANN Multiplied by 4 sub-series for every input neuron as obtained by wavelet decomposition, 6 hidden neurons and 1 output neuron. Whereas, the final structure of the ANN model is the number of inputs used in ANN, 6 hidden neurons and 1 output neuron. The performance of various models estimated to forecast the rainfall was presented in Table 3.

From Table 3, it is found that low RMSE values (63.01 to 117.39 mm) for WNN models when compared to ANN and AR models. It has been observed that WNN models estimated the peak values of rainfall to a reasonable accuracy (peak rainfall in the data series is 1401.80 mm). Further, it is observed that the WNN model having two antecedent values of the rainfall, one

Table 3 Goodness of fit statistics of the calibration and validation for the forecasted rainfall

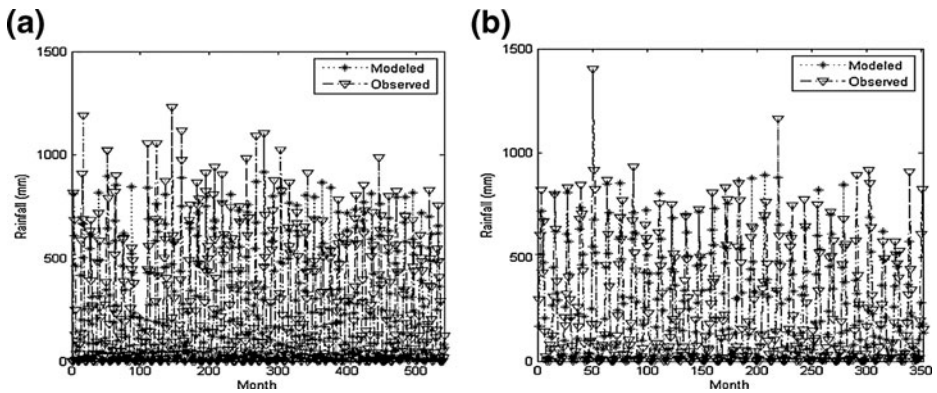| Model	Calibration			Validation		
	RMSE	R	COE(%)	RMSE	R	COE(%)
WNN						
I	102.26	0.934	87.19	117.39	0.906	81.78
II	52.78	0.983	96.58	63.92	0.973	94.60
III	40.97	0.989	97.94	81.51	0.955	91.24
IV	**35.12**	**0.992**	**98.48**	**63.01**	**0.974**	**94.78**
V	34.74	0.993	98.51	84.79	0.953	90.57
VI	35.75	0.992	98.43	82.24	0.955	91.10
ANN						
I	205.96	0.695	48.38	203.98	0.672	45.01
II	176.56	0.788	62.05	171.15	0.784	61.35
III	145.43	0.861	74.24	178.31	0.783	58.10
IV	123.23	0.902	81.49	163.79	0.807	64.73
V	116.75	0.913	83.36	178.08	0.786	58.42
VI	114.66	0.916	83.98	176.84	0.784	58.89
AR						
	226.00	0.659	37.70	221.82	0.642	34.91

Fig. 6 Plot of observed and modeled rainfall for ANN model during (**a**) calibration and (**b**) validation

minimum temperature and maximum temperature estimated minimum RMSE (63.01 mm), high correlation coefficient (0.9736) and highest efficiency (>94 %) during the validation period. The model IV (Bold) of WNN was selected as the best fit model to forecast the rainfall in 1 month advance.

Figures 6 and 7, shows the observed and modeled graphs for ANN and WNN models during calibration and validation respectively. It is found that values modeled from WNN model properly matched with the observed values, whereas ANN model underestimated the observed values. From this analysis, it is evident that the performance of WNN was much better than ANN and AR models in forecasting the rainfall.

The distribution of error along the magnitude of rainfall computed by WNN and ANN models during the validation period has been presented in Fig. 8. From Fig. 8, it was observed that the estimation of peak rainfall was very good as the error is minimum when compared with ANN model. Figure 9 shows the scatter plot between the observed and modeled rainfall by WNN and ANN. It was observed that the rainfall forecasted by WNN models were very much close to the 45° line. From this analysis, it was worth to mention that the performance of WNN was much better than the standard ANN and AR models in forecasting the rainfall in 1-month advance.
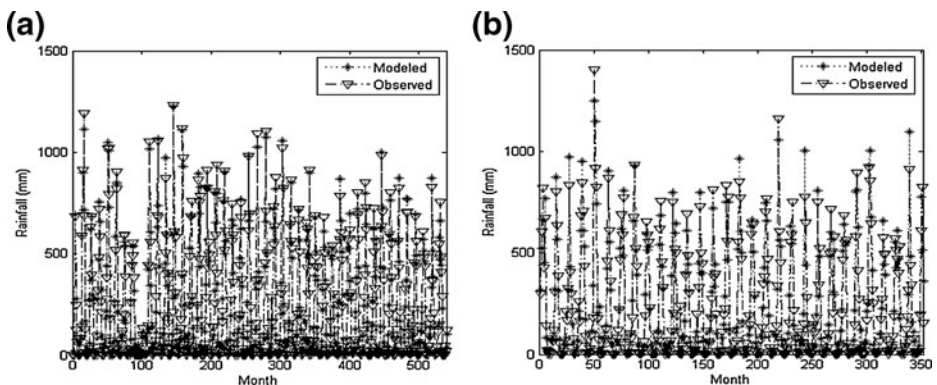

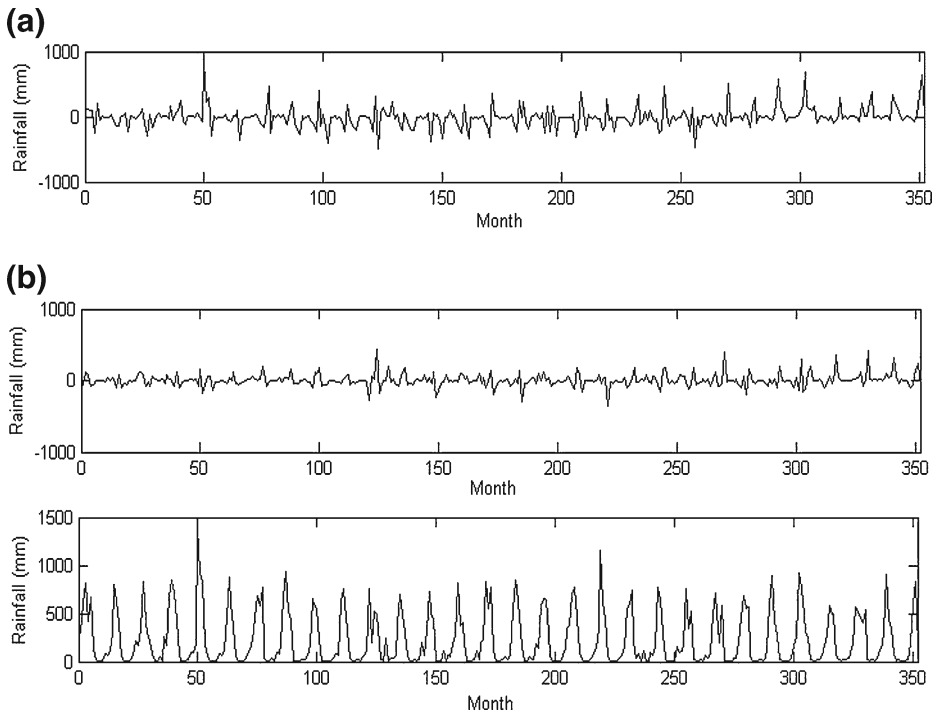**Fig. 7** Plot of observed and modeled rainfall for WNN model during (**a**) calibration, and (**b**) validation

(a)

(b)

Fig. 8 Distribution of error plots along the magnitude of monthly rainfall for (**a**) ANN model and (**b**) WNN model during validation period

8 Conclusions

This paper reports a hybrid model called wavelet neural network model (WNN) for time series modeling of monthly rainfall. The proposed model is a combination of wavelet analysis and artificial neural network. Wavelet decomposes the time series into multilevel of details and it can adopt multiresolution analysis and effectively diagnose the main frequency component of
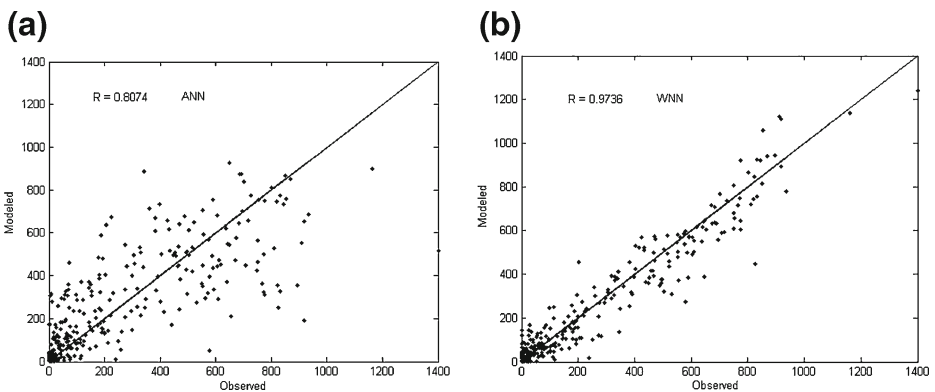
(a) **(b)**

Fig. 9 Scatter plot between observed and modeled monthly rainfall for (**a**) ANN model and (**b**) WNN model during validation period

the signal and abstract local information of the time series. The proposed WNN model has been applied to monthly rainfall of Darjeeling rain gauge station. The time series data of rainfall was decomposed into sub series by DWT. Appropriate sub-series of the variable used as inputs to the ANN model and original time series of the variable as output. Model parameters are calibrated using 44 years of data and rest of the data is used for model validation. From this analysis, it was found that efficiency index is more than 94 % for WNN models whereas it is 64 % for ANN models respectively.

Overall analysis indicates that the performance of ANN are relatively lower compared to that of WNN models; this may be plausibly due to the variation in nonlinear dynamics of temperature which plays a predominant role in hilly areas included in rainfall process that mapped effectively by Wavelet based models. It may be noted that hydrological data used in the WNN model has been decomposed in details and approximation, which may lead to better capturing the rainfall processes. The study only used data from one rain gauge station and further studies using more rain gauges data from various areas may be required to strengthen these conclusions.

References

Adamowski J, Sun K (2010) Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds. J Hydrol 390(1–2):85–91

Anctil F, Tape DG (2004) An exploration of artificial neural network rainfall-runoff forecasting combined with wavelet decomposition. J Environ Eng Sci 3:121–128

Antonios A, Constantine EV (2003) Wavelet Exploratory Analysis of the FTSE ALL SHARE Index. In proceedings of the 2nd WSEAS international conference on non-linear analysis. Non-linear systems and Chaos, Athens

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

Aussem A, Murtagh F (1997) Combining neural network forecasts on wavelet transformed series. Connect Sci 9(1):113–121

Benaouda D, Murtagh F, Starck JL, Renaud O (2006) Wavelet-based nonlinear multiscale decomposition model for electricity load forecasting. Neurocomputing 70(1–3):139–154

Bhakar SR, Singh RV, Neeraj C, Bansal AK (2006) Stochastic modeling of monthly rainfall at kota region. ARPN J Eng Appl Sci 1(3):36–44

Cannas B, Fanni A, See L, Sias G (2006) Data preprocessing for river flow forecasting using neural networks: wavelet transforms and data partitioning. Phys Chem Earth A/B/C 31(18):1164–1171

Carlson RF, MacCormick AJA, Watts DG (1970) Application of linear models to four annual streamflowseries. Water Resour Res 6(4):1070–1078

Chinchorkar SS, Patel GR, Sayyad FG (2012) Development of monsoon model for long range forecast rainfall explored for Anand (Gujarat-India). Int J Water Resour Environ Eng 4(11):322–326

Chou C (2011) A threshold based wavelet denoising method for hydrological data modelling. Water Resour Manag 25:1809–1830

Coulibaly P, Anctil F, Rasmussen P, Bobee B (2000) A recurrent neural networks approach using indices of Low-frequency climatic variability to forecast regional annual runoff. Hydrol Process 14(15):2755–2777

Dawson DW, Wilby R (2001) Hydrological modeling using artificial neural networks. Prog Phys Geogr 25(1):80–108

Grossmann A, Morlet J (1984) Decomposition of hardy functions into square integrable wavelets of constant shape. SIAM J Math Anal 15(4):723–736

Hagan MT, Menhaj MB (1994) Training feed forward networks with Marquardt algorithm. IEEE Trans Neural Netw 5:989–993

Hettiarachchi P, Hall MJ, Minns AW (2005) The extrapolation of artificial neural networks for the modeling of rainfall-runoff relationships. J Hydroinformatics 7(4):291–296

Hsu KL, Gupta HV, Sorooshian S (1995) Artificial neural network modeling of rainfall-runoff process. Water Resour Res 31(10):2517–2530

Huang MC (2004) Wave parameters and functions in wavelet analysis. Ocean Eng 31(1):111–125

Kisi O (2008) Stream flow forecasting using neuro-wavelet technique. Hydrol Process 22(20):4142–4152

Kisi O (2009) Neural networks and wavelet conjunction model for intermittent stream flow forecasting. J Hydrol Eng 14(8):773–782

Kisi O (2010) Wavelet regression model for short-term stream flow forecasting. J Hydrol 389:344–353

Kisi O (2011) Wavelet regression model as an alternative to neural networks for river stage forecasting. Water Resour Manag 25(2):579–600

Kisi O, Cimen M (2011) A wavelet-support vector machine conjunction model for monthly stream flow forecasting. J. Hydrol 399:132–140

Kucuk M, Agiralioğlu N (2006) Wavelet regression technique for stream flow prediction. J Appl Stat 33(9):943–960

Lau KM, Weng HY (1995) Climate signal detection using wavelet transform: how to make a time seriessing. Bull Am Meteorol Soc 76:2391–2402

Mallat S (1998) A wavelet tour of signal processing. Academic, San Diego

Massel SR (2001) Wavelet analysis for processing of ocean surface wave records. Ocean Eng 28:957–987

Matalas NC, Wallis JR (1971) Statistical properties of multivariate fractional noise process. Water Resour Res 7:1460–1468

Moustris KP, Ioanna K, Larissi (2011) Precipitation forecast using artificial neural networks in specific regions of Greece. Water Resour Manag 25:1979–1993

Partal T, Kisi O (2007) Wavelet and neuro-fuzzy conjunction model for precipitation forecasting. J Hydrol 342:199–212

Rumelhart DE, Hinton GE, Williams RJ (1986) Learning Representations by Back-Propagating Errors. Nature 323(9):533–536

Sang Y (2013) Improved wavelet modeling framework for hydrologic time series forecasting. Water Resour Manag 27(8):2807–2821

Tantanee S, Patamatammakul S, Oki T, Sriboonlue V, Prempree T (2005) Coupled wavelet-autoregressive model for annual rainfall prediction. J Environ Hydrol 13(18):1–8

Teisseire LM, Delafoy MG, Jordan DA, Miksad RW, Weggel DC (2002) Measurement of the instantaneous characteristics of natural response modes of a spar platform subjected to irregular wave loading. J Offshore Polar Eng 12(1):16–24

Torrence C, Compo GP (1997) A practical guide to wavelet analysis. Bull Am Meteorol Soc 79(1):61–78

Valencia DR, Schaake JC Jr (1973) Disaggregation processes in stochastic hydrology. Water Resour Res 9(3):580–585

Wang D, Ding J (2003) Wavelet network model and its application to the prediction of hydrology. Nat Sci 1(1):67–71

Wang W, Li Y (2011) Wavelet transform method for synthetic generation of daily stream flow. Water Resour Manag 25:41–57

Wang W, Jin J, Li Y (2009) Prediction of inflow at three gorges dam in Yangtze River with wavelet network model. Water Resour Manag 23(13):2791–2803

Wu D, Wang J, Teng Y (2004) Prediction of under-ground water levels using wavelet decompositions and transforms. J Hydrol Eng 5:34–39

Yevjevich V (1972) Stochastic processes in hydrology. Water Resour Pub, Colorado