Improving Rainfall Forecasting Efficiency using Wavelet Neural Network

R. Venkata Ramana and B. Krishna

Scientists, Deltaic Regional Center, National Institute of Hydrology, Siddartha Nagar, Kakinada-3, A.P.
E-mail: venkataramana_1973@yahoo.co.in

Abstract: *Rainfall is one of the most complicated effective hydrologic processes in runoff prediction and water management. Time series analysis requires mapping complex relationships between input(s) and output(s), since the forecasted values are mapped as a function of observed patterns in the past. In order to raise the forecasted precision, an alternative hybrid model called wavelet neural network model (WNN), which is the combination of wavelet analysis and ANN, has been proposed. The wavelet neural network (WNN) has been widely used for modeling different kinds of nonlinear systems including rainfall forecasting. 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 daily rainfall data series of IMD Patna rain gauge station. The observed time series is decomposed into sub-series using discrete wavelet transform and then appropriate sub-series is used as inputs to the neural network. The calibration and validation performance of the models is evaluated with appropriate statistical methods. The results showed that the model based WNN performed higher rainfall forecasting accuracy with low errors.*

Key words: *Rainfall, Training, Decomposition, Neural network and Wavelet*.

1. Introduction:

The most important variable affecting the hydraulic behavior of different construction and structure such as dams, spillways, culverts, bridges, irrigation ditches, urban and rural expansion based on accurate rainfall prediction, could decrease the effects of floods. In recent years, various new mathematical techniques have been developed to analyze the rainfall data. Moreover, the achieved results have some errors using those models. Undoubtedly, the accurate forecasting in rainfall signals could present useful information for water resource management, flood control and disaster relief. This paper introduces a new approach for forecast of rainfall series.

Several time series models have been proposed for modeling annual rainfall series such as the autoregressive model (AR) [28], the fractional Gaussian noise model [18], autoregressive moving-average models (ARMA) [8] and the disaggregation multivariate model [25]. 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 [17]; [22] ; [12]. The time-frequency character of long-term climatic data is investigated using the continuous wavelet transform technique [14]; [23]; [16] and wavelet analysis of wind wave measurements obtained from a coastal observation tower [12]. Combination of neural networks and wavelet methods to predict ground water levels [27]. 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 [4]. Wavelet transforms were also applied to time series prediction preprocessed for multistep prediction [24]. By coupling the wavelet method with the traditional AR model, the Wavelet-Autoregressive model (WARM) is developed for annual rainfall prediction. Conjunction model (wavelet-neuro-fuzzy) to forecast the Turkey daily precipitation to observe 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 [19].

In this paper, a Wavelet Neural Network (WNN) model, which is the combination of wavelet analysis and ANN, has been proposed for rainfall forecast modeling of IMD station, Patna.

2. Wavelet Analysis:

The wavelet analysis is an advanced tool in signal processing that has attracted much attention since its theoretical development [11]. 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 [1]. 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 φ (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 [5]. 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.

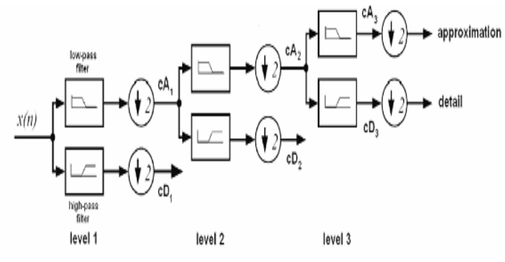

Figure 1. Diagram of multiresolution analysis of signal**.**

3. Artificial Neural Networks (ANN):

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 [20]. 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 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 [2]&[3]. 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

$$Y = f\left(\sum WX - \theta\right) \qquad (2)$$

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.

4. 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 [12]. The method also takes advantage of the internal recurrence to dynamically incorporate past experience in the training process [9].

The Levenberg-Marquardt Algorithm (LMA) is given by

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

where, X is the weights of neural network, J is 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. LM 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.

5. 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 by 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.

6. 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 Figure 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. 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 hidden layer 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.

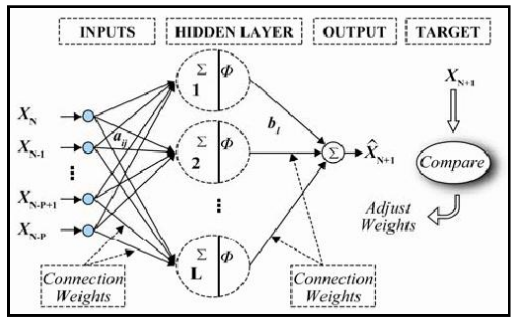

Figure 2. Signal data based multilayer perceptron (MLP) neural network structure.

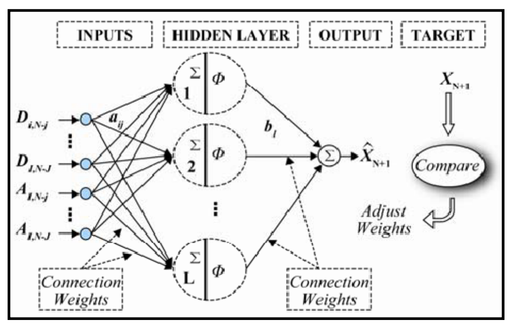

Figure 3. Wavelet based multilayer perceptron (MLP) neural network structure.

7. 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).

8. Study Area:

Patna, capital of Bihar state is situated on the bank of river Ganga with Latitude 25^0 $37^{'}$ and longitude 85^0 $10^{'}$ and has a mean elevation of 49.68m above the MSL. Patna and its upland is sandwiched between the high Himalayan ranges in the far north and the high tracts of Chhotanagpur in the south. Due to its location with relation to latitude and other features, Patna has a humid subtropical climate with hot summers from late March to early June, the monsoon season from late June to late October and a mild winter from November to February. Highest temperature ever recorded is 46.6^0C (in 1966) lowest ever is 2.3^0C (in 2003) and highest daily rainfall was 250.8 mm (in 1987). Patna has a temperate climate, suitable for urban living. Patna is a linear city and is about 30 Km long from east to west and 5-7 Km from north to south. The city is situated between the river Ganga in the north, river Punpun in the south and river Sone in the west (Fig. 4).

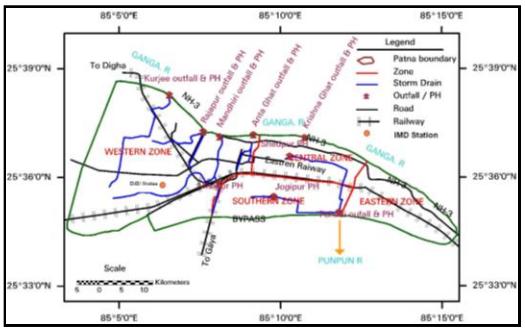

Figure 4. Study area and location of IMD station of Patna.

9. Development of Wavelet Neural Network (WNN) Model:

The original time series was decomposed into Details and Approximations to certain number of sub-time series $\{D_1, D_2, \ldots D_p, A_p\}$ by wavelet transform algorithm. These play different role in the original time series and the behavior of each sub-time series is distinct [26]. 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. 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 are the antecedent values of the time series and were presented in Table 1. 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 deter-mined by trial and error procedure.

Table 1. Model Inputs

Model I	X (t) = f (x [t-1])
Model II	X (t) = f (x [t-1], x [t-2])
Model III	X (t) = f (x [t-1], x [t-2], x [t-3])
Model IV	X (t) = f (x [t-1], x [t-2], x [t-3], x [t-4])
Model V	X (t) = f (x [t-1], x [t-2], x [t-3], x [t-4], x [t-5])
Model VI	X (t) = f (x [t-1], x [t-2], x [t-3], x [t-4], x [t-5] x [t-6])

10. Results and Discussion:

To forecast the rainfall at Patna gauging station of IMD Patna (Fig. 4), the daily rainfall data of 29 years from 1975 to 2004 was used. The first twenty years (1975-95) data were used for training of the model, and the remaining nine years (1995-2004) data were used for validation. The model inputs (Table 1) 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 was determined as six by trial and error procedure. The performance of various models estimated to forecast the rainfall was presented in Table 2.

From Table 2, it is found that low RMSE values (3.0236mm to 8.9116mm), correlation coefficient (0.9705 to 0.6823) and coefficient of efficiency (46.2979% to 94.1760%) 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 205.4 mm). Further, it is observed that the WNN model having four antecedent values of the time series, estimated minimum RMSE (3.4725mm), high correlation coefficient (0.9597) and coefficient of efficiency (>91%) during the validation period. The model IV of WNN was selected as the best-fit model to forecast the rainfall one-day in advance.

Table 2. Goodness of fit statistics of the calibration and validation for the forecasted rainfall.

Model	Calibration			Validation		
	RMSE	R	COE(%)	RMSE	R	COE(%)
WNN						
Model I	7.9231	0.7751	60.0084	8.9116	0.6823	46.2979
Model II	4.0116	0.9474	89.7479	4.8875	0.9158	83.8518
Model III	3.6296	0.9572	91.6077	4.3767	0.9331	87.0545
Model IV	3.2088	0.9667	93.4406	3.9893	0.9471	89.2476
Model V	**3.0236**	**0.9705**	**94.1760**	**3.4725**	**0.9597**	**91.8555**
Model VI	2.5381	0.9693	93.8963	4.5157	0.9314	86.2307
ANN						
Model I	11.6497	0.368	13.5430	11.6048	0.2871	8.9351
Model II	11.5549	0.3867	14.9432	11.8686	0.2751	4.7755
Model III	11.4854	0.3996	15.9634	11.8111	0.2768	5.7219
Model IV	11.4312	0.4094	16.7550	11.7725	0.2862	6.3653
Model V	11.396	0.4156	17.2672	11.7837	0.2878	6.2137
Model VI	11.3885	0.4149	17.3764	11.8008	0.2829	5.9682
AR						
	11.9537	0.3379	8.9721	12.3344	0.1796	-2.9345

Figure 5 and 6, shows the observed and modeled graphs for ANN and WNN models 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.

An analysis to assess the potential of each of the model to preserve the statistical properties of the observed rainfall series was carried out and reveals that the rainfall series computed by the WNN model reproduces the first three statistical properties (*i.e.* mean and standard deviation) better than that computed by the ANN model. The values of the first three properties for the observed and modeled rainfall series for the validation period were presented in Table 3 for comparison. In Table 3, AR model performance was not presented because of its low efficiency compared to other models.

The distribution of error along the magnitude of rainfall computed by WNN and ANN models during the validation period has been presented in Figure 7. From Figure 7, it was observed that the estimation of rainfall was very good as the error is minimum when compared with ANN model.

Figure 8 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 degrees line. From this analysis, it was worth to mention that the performance of WNN was much better than ANN and AR models in forecasting the rainfall in one-day advance.

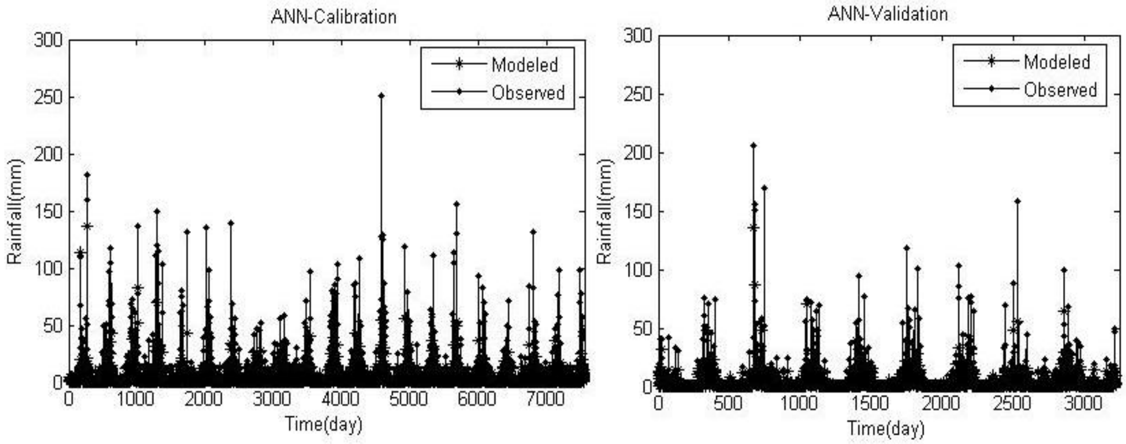

Figure 5. Observed and modeled rainfall for ANN model during calibration and validation.

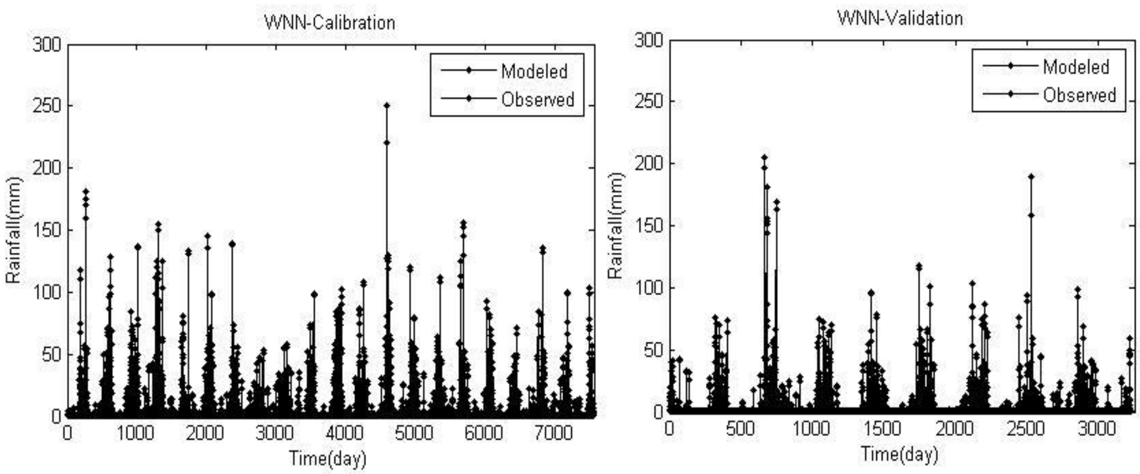

Figure 6. Observed and modeled rainfall for WNN model during calibration and validation.

Figure 8 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 degrees line. From this analysis, it was worth to mention that the performance of WNN was much better than ANN and AR models in forecasting the rainfall in one-day advance.

Table 3. Statistical properties of observed and model of rainfall series during validation period

Statistical properties	**Observed**	**WNN**	**ANN**
Minimum	0	0	0.5538
Maximum	205.4	197	135.7
Mean	3.226	3.472	3.4830
Standard deviation	12.17	12.12	5.563

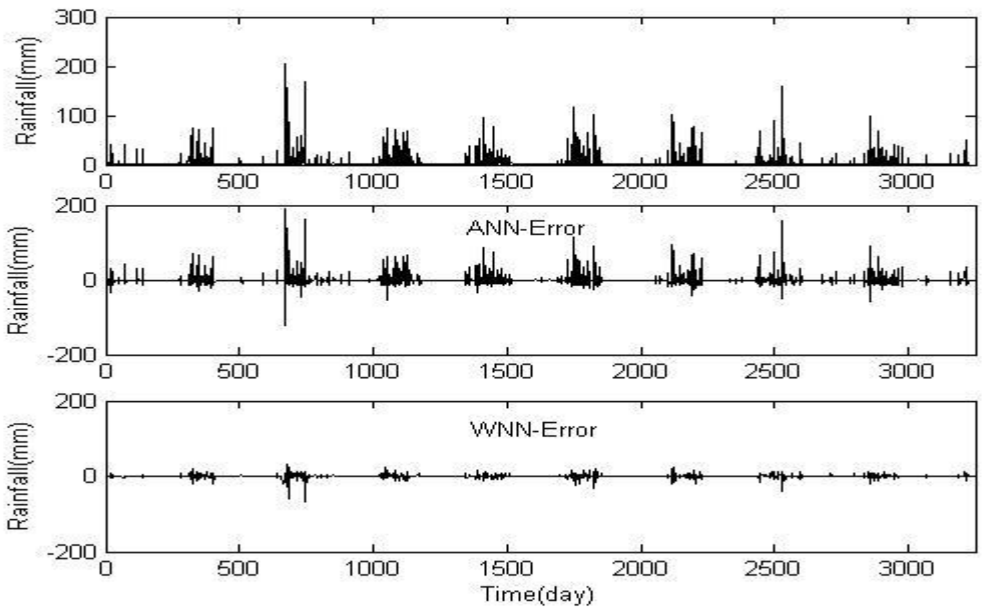

Figure 7. Distribution of error plots along the magnitude of rainfall for (a) ANN model and (b) WNN model during validation period.

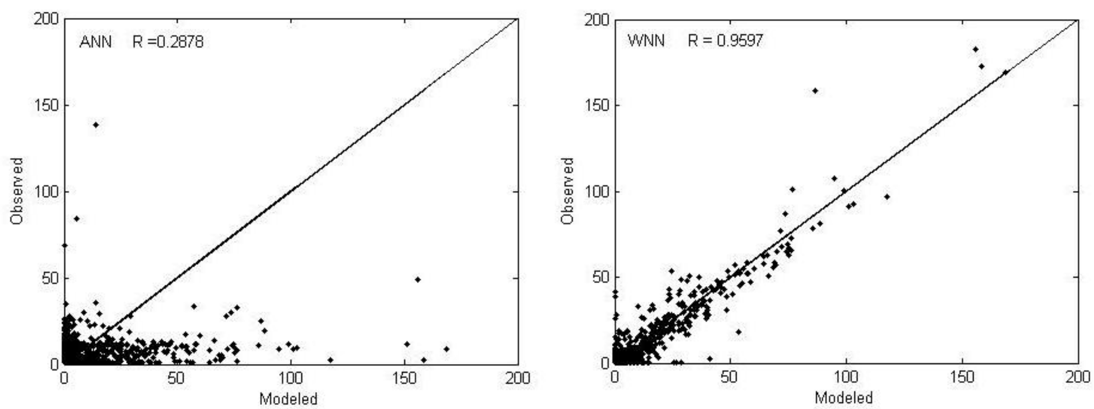

Figure 8. Scatter plot between observed and modeled rainfall during validation period ANN and WNN.

11. Conclusions:

This paper reports a hybrid model called wavelet based neural network model for time series modeling of rainfall. The proposed model is a combination of wavelet analysis and artificial neural network (WNN). Wavelet decomposes the time series into multilevels of details and it can adopt multiresolution analysis and effectively diagnose the main frequency component of the signal and abstract local information of the time series. The proposed WNN model has been applied to daily rainfall of Patna gauging station, IMD Patna India. The time series data of rainfall was decomposed into sub series by DWT. Each of the sub-series plays distinct role in original time series. Appropriate sub-series of the variable used as inputs to the ANN model and original time series of the variable as output. From the current study it is found that the proposed wavelet neural network model is better in forecasting rainfall of Patna gauging station. In the analysis, original signals are represented in different resolution by discrete wavelet transformation; therefore, the WNN forecasts are more accurate than that obtained directly by original signals. Where rainfall prediction is needed in several hydrological models, WNN can be applied to those models in order to make irrigation or water resources planning management more effective.

12. References:

[1]. Antonios, A. and Constantine, E. V. “Wavelet Exploratory Analysis of the FTSE ALL SHARE Index,” Preprint submitted to Economics Letters University of Durham, Durham, 2003.

[2]. ASCE Task Committee, “Artificial Neural Networks in hydrology-I: Preliminary Concepts,” *Journal of Hydrologic Engineering*, Vol. 5, No. 2, 2000(a), pp. 115-123.

[3]. ASCE Task Committee, “Artificial Neural Networks in Hydrology-II: Hydrologic Applications,” *Journal of Hydrologic Engineering*, Vol. 5, No. 2, 2000(b), pp. 124- 137.

[4]. Aussem, A. and Murtagh, F. “Combining Neural Network Forecasts on Wavelet Transformed Series,” *Connection Science*, Vol. 9, No. 1, pp. 113-121. 1997

[5]. Benaouda, D., Murtagh, F,.. Starck, J. L. and Renaud, O. “Wavelet-Based Nonlinear Multiscale Decomposition Model for Electricity Load Forecasting,” *Neurocomputing*, Vol. 70, No. 1-3, 2006, pp. 139-154.

[6]. Box, G.E., and G. Jenkins. 1970. Time series analysis, forecasting and control, Holden-Day Inc., San Francisco.

[7]. Burrus, C.S., R.A. Gopinath, and H. Gao. Introduction to wavelet and wavelet transform, Prentice-Hall International, Inc. 1998.

[8]. Carlson, R.F., A.J.A. MacCormick, and D.G. Watts. Application of linear models to four annual streamflow series, Water Resour. Res., 6,4, pp 1070-1078, 1970.

[9]. Coulibaly, P., Anctil, F.., Rasmussen, P.and Bobee, B. “A Recurrent Neural Networks Approach Using Indices of Low-Frequency Climatic Variability to Forecast Regional Annual Runoff,” *Hydrological Processes*, Vol. 14, No. 15, 2000, pp. 2755-2777.

[10]. Feiring, M.B., and B.B. Jackson. Synthetic hydrology monograph No.1, American geophysical Union,Washington D.C. 1971.

[11]. Grossmann, A. and Morlet, J. “Decomposition of Hardy Functions into Square Integrable Wavelets of Constant shape,” *SIAM Journal on Mathematical Analysis*, Vol. 15, No. 4, 1984, pp. 723-736.

[12]. Huang, M.C. 2004. Wave parameters and functions in wavelet analysis, Ocean Engineering 31:1, pp 111-125

[13]. Jury, M.R., and J.L. Melice. 2000. Analysis of Durban rainfall and Nile river flow 1871-1999, Theor. Appl.Climatol. 67, pp 161-169

[14]. Lau, K.M., and H.Y. Weng. Climate signal detection using wavelet transform: how to make a time seriessing, Bull. Amer. Meteor. Soc. 76, pp 2391-2402, 1995.

[15]. M. T. Hagan and M. B. Menhaj, “Training Feed forward Networks with Marquardt Algorithm,” *IEEE Transac-tions on Neural Networks*, Vol. 5, No. 6, 1994, pp. 989-993.

[16]. Mallat, S. 1998. A wavelet tour of signal processing, San Diego, CA: Academic Press.

[17]. Massel, S.R. 2001. Wavelet analysis for processing of ocean surface wave records, Ocean Engineering 28, pp 957-987

[18]. Matalas, N.C., and J.R. Wallis. Statistical properties of multivariate fractional noise process, Water Resour.Res., 7, pp 1460-1468. 1971.

[19]. Partal. T and Kisi. O. “Wavelet and Neuro-Fuzzy Con-junction Model for Precipitation Forecasting,” *Journal of Hydrology*, Vol. 342, No. 1-2, 2007, pp. 199-212.

[20]. Rumelhart, D. E., Hinton G. E. and Williams, R. J. “Lear- ning Representations by Back-Propagating Errors,” *Nature*, Vol. 323, No. 9, 1986, pp. 533-536

[21]. Salas, J.D., J.W. Delleur, V. Yevjevich, and W.L. Lane. Applied modeling of hydrological time series, Book Crafters, Inc. Michigan. 1985

[22]. Teisseire, L.M., M.G. Delafoy, D.A. Jordan, R.W. Miksad, and D.C. Weggel. Measurement of the instantaneous characteristics of natural response modes of a spar platform subjected to irregular wave loading, International Journal of Offshore and Polar Engineering 12:1, pp 16-24, 2002.

[23]. Torrence, C., and G.P. Compo. A practical guide to wavelet analysis, Bull. Amer. Meteor. Soc. 79:1, pp 61-78. 1997.

[24]. Tsui, F.C., C.C. Li, M. Sun, and R.J. Sclabassi. 1997. A comparative study of two bio-orthogonal wavelet

[25]. Valencia, D.R., and J.C. Schaake, Jr. Disaggregation processes in stochastic hydrology, Water Resour. Res.9:3, pp 580-585, 1973.

[26]. Wang, D. and Ding, J. "Wavelet Network Model and Its Application to the Prediction of Hydrology," *Nature and Science*, Vol. 1, No. 1, 2003, pp. 67-71.

[27]. Wu, D. J., Wang, J and Teng, Y. "Prediction of Under-ground Water Levels Using Wavelet Decompositions and Transforms," *Journal of Hydro-Engineering*, Vol. 5, 2004, pp. 34-39.

[28]. Yevjevich, V. Stochastic processes in hydrology, Water Resources Publications, Colorado.1972.