J. Indian Water Resour. Soc.,
Vol. 41, No. 1, Jan., 2021

APPLICATION OF WANN MODEL FOR GROUNDWATER LEVEL FORECASTING IN UR RIVER WATERSHED IN TIKAMGARH DISTRICT, INDIA

Ankur Kumar[1], Ravindra Vitthal Kale[2], Govind Pandey[3] and V. C. Goyal[4]

ABSTRACT

The use of Aquifers as a source of water supply is increasing on a global scale, leading to over-exploitation of available groundwater blocks. Thus, there is an increasing demand for checking the groundwater levels for better and sustainable management of groundwater resources. To acquire knowledge about the factors affecting the entire groundwater system, one should know the important variables and how they vary over time. It is well known that the groundwater head is considered to be one of the most essential hydrological variables and hence, it is monitored and predicted frequently at different locations and at frequent time intervals. Particularly, the groundwater prediction in hard rock areas is a complex task with the use of physically-based models as compared to the data-driven models. Therefore, in this study, an attempt has been made to verify the adequacy as well as the efficacy of the Artificial Neural Network model (ANN) and Wavelet-ANN conjunction (WANN) models in the prediction of groundwater levels in the Ur River watershed in Tikamgarh district of Madhya Pradesh, India. Although the Ur river basin having mainly granite type of aquifer, the obtained results reveal that the WANN and ANN models can be used to predict the groundwater levels in this watershed. The application of the ANN model in the groundwater prediction gives a higher estimate of the RMSE values during calibration and validation as compared to those obtained with the application of the WANN model for each one of the observation wells. Further, the WANN model is capable to provide groundwater level prediction with higher efficiency as reflected by higher R^2 values during calibration and validation as compared to the ANN model which indicates a substantial improvement in the model performance. Therefore, it can be concluded that the WANN model provides a significantly accurate prediction of groundwater levels as compared to the results of the ANN model. Besides, the comparison of the scatter plots of time series during calibration and validation indicates that the values of water level depth estimated by the WANN model are more precise than those estimated by the ANN. Thus, this paper reveals the significant features of ANN models for forecasting groundwater levels in hard rock aquifer and their performance enhancement with wavelet theory.

Keywords: *Modeling, Groundwater level depth, Artificial Neural Networks, Feedforward neural networks, Discrete Wavelet Transform, Mother Wavelet, Decomposition level, Hargreaves Temperature Model.*

INTRODUCTION

Groundwater is regarded to be one of the most reliable water supply sources for meeting the demands of water for various sectors in India including manufacturing, industries, agriculture, mining, and municipal water supply. It can also be regarded as a vital source in terms of clean drinking water in the country. However, on current trends, it is estimated that 60 percent of groundwater sources will be in a critical state of degradation within the next twenty years and thus have serious implications for the sustainability of agriculture, long-term food security, livelihoods, and economic growth. As groundwater resources are more intensively used, there is an increasing need for monitoring and modeling of groundwater systems. One of the most important and critical hydrological variables is the groundwater level, which is therefore monitored and predicted frequently at different locations and at frequent time intervals. Accurate prediction of the groundwater level helps a water engineer in developing better strategies to reduce the effects of various factors leading to the progressive decline of groundwater levels and thus it assists in the better sustainable management of groundwater sources. Thus, to assess the factors governing the rapid decline of groundwater levels, modeling groundwater resources can help a water engineer in achieving the objective in a better way.

Almost all the groundwater flow and available transport models solve the relevant partial differential equation by the finite element method. These modeling methods are very much data and labour intensive and costly. In the recent past, intelligent technique based ANN models is being applied intensively to gain insight into the hydrological processes due to their better performance over the traditional modeling techniques such as empirical models, statistical models (autoregressive, autoregressive moving average models), and as an alternative to the physical-based models.

ANNs also treated as Universal approximators are an alternative modeling and simulation tool, greatly suited to dynamic non-linear system modeling. Another attractive featureof ANNs is that they often do not require explicit characterization and quantification of physical assumptions like traditional physical-based numerical models. ANN models also include some drawbacks when handled with a non-stationary signal of a hydrologic process that involves seasonalities that vary from a single day to several decades.

1. *Assistant Professor, CE Dept, BIT, Gorakhpur, India*
   Email : ankur495@bit.ac.in
2. *Scientist 'D', WHRC, NIH, Jammu, India*
   Email: ravikale2610@gmail.com
3. *Professor, CE Dept, MMMUT, Gorakhpur, India*
   Email: pandey_govind@rediffmail.com
4. *Scientist 'F, NIH, Roorkee*
   Email : vcgoyal@yahoo.com

Manuscript No. 1549
Received 8 January 2021; Accepted 10 May 2021

Therefore, in such a situation pre-processing of time and space data may be an effective approach to overcome the drawbacks. The effectiveness of the wavelet transform is disintegrating non-stationary time series into sub-series at different scales (levels) is helpful for a betterunderstanding of the process. Therefore, the combination of ANN with wavelet transform as a hybrid wavelet-ANN (WANN) model that can explain simultaneously spectral and temporal information of the signal creates an effective implement for the prediction of hydrological processes (Vahid et al., 2013). The wavelet transform of a function is the improvedversion of the Fourier transform. Fourier transform is a powerful tool for analyzing the components of a stationary signal. But it is failed for analyzing the non-stationary signal whereaswavelet transform allows the components of a non-stationary signal to be analyzed hence any application using the Fourier transform can be formulated using wavelets to provide more accurately localized temporal and frequency information. The analyzing function in the wavelet transformation is called wavelet. It will adjust the time width of the frequencies in such a way that high frequency once will be broader. Wavelet transform is one of the best tools to determine low-frequency areas and high-frequency areas. Mainly there are two approaches available in the wavelet theory. The integral transform approaches continuous-time and the multi-resolution analysis MRA/ filter bank approach discrete time.

In this analysis, the ANN model with feed-forward structure has been developed to predict groundwater level with antecedent rainfall, evapotranspiration, and groundwater level as the input vector. The data of maximum and minimum temperature, extraterrestrial radiation from January 2004 to December 2013 have been collected from the Tikamgarh observation station used to compute the evapotranspiration which has been fed to ANN as an input. The different discrete wavelet transformation has been performed to decompose the input vector to forecast the groundwater level. The efficiency of the ANN model with wavelet transformed WANN is compared with the performance of the ANN model with a real input vector.

STUDY AREA AND DATA COLLECTION

Study Area

The performance of the proposed model is evaluated using the quarterly time series data which obtained using an average of daily rainfall and daily maximum, minimum, and mean temperature data; and quarterly measured average water table depth which is recorded for the Ur River watershed in the Tikmagarh district, situated in the northern part of the state of M.P., India. The Tikamgarh district encompasses a total area of 5048 km^2 bounded in the north and west by the Jhansi and Lalitpur of Uttar Pradesh. The entire district comes under the Betwa sub-basin of the Ganga basin. For the entire Tikamgarh district, altitude ranges from 200 to 400 m above mean sea level. The hill ranges are made up of hard compact and resistant granite masses intruded by quartz reefs. The normal annual rainfall during the monsoon season is recorded to be 1057.1 mm. The temperature of the study area varies widely between 7^o C and 41.8^o C. The outlet location of the district is 25^{o}40' N and 78^{o}26' E. Tikmagarh district is divided into six blocks as Tikamgarh, Baldeogarh, Jatara, Palera, Niwari, Prithipur. Out of which Jatara is the biggest with an area 1008.60 km^2 while Niwari is the smallest with an area of 606 km^2. The geographical area of the Ur river watershed is 990 km^2. Some of the observation wells falling outside the Ur river watershed are also considered in the analysis.

Geomorphological features are directly controlled by the geological formation and their structures. Soils of the entire study area fall into three categories Viz. black humus granitic and yellowish-grey colour with kankar soils derived due to disintegration and decomposition of the parent rock. The Location of the Ur river watershed is shown in Figure 1.

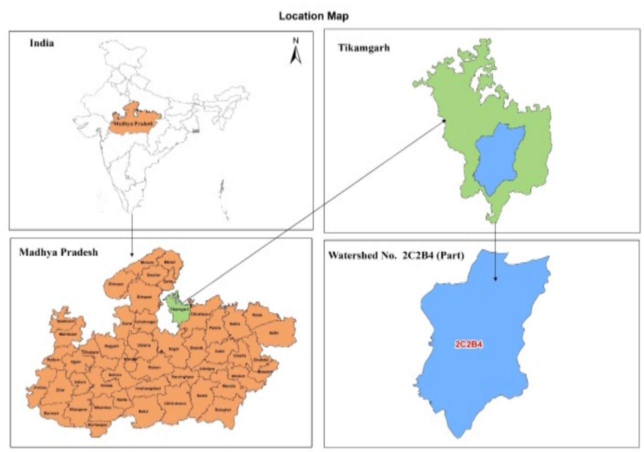

Fig. 1 : Location of Ur River watershed

Data collection

All the meteorological variables and groundwater level data used in this research work is respectively obtained from IMD and Central Ground Water Board (CGWB), Bhopal, Madhya Pradesh by NIH, Roorkee (Uttarakhand). The data consists of daily precipitation (mm), daily maximum, minimum, and mean air temperature (^{o}C) data, and quarterly average groundwater level (mbgl) data. This study uses the groundwater level data for eight observation wells observed for ten years (from January 2004 to December 2013). As shown in Figure 2, the selected observation wells are widely distributed over the study area.

The evapotranspiration data used in this research work is calculated using a Hargreaves Temperature Model. This model is applied for the present study area due to the non-availability of sunshine hour data. The Hargreaves temperature equation is one of the simplest and most accurate empirical equations used to estimate evapotranspiration (ET$_o$) in mm/day, which is expressed as.

$$ET_o = 0.0023\, R_a \left(T_{mean} + 17.8\right) \sqrt{T_{\max} - T_{\min}} \quad (1)$$

Here, ET_o = evapotranspiration in mm/day, T_{mean} , T_{max} , T_{min} = mean, maximum, and minimum air temperatures in (°C), respectively and R_a = extraterrestrial radiation in (mm/day) (Bhabagrahi et al. 2012).

Fig. 2: Geographical Locations of the observation wells

METHODOLOGY

Artificial neural networks (ANNs)

Artificial neural networks (ANNs) are inspired by the research processes that take place in biological systems. It can be regarded as an information-processing construct composed of several parallel interconnected processing elements known as simple nodes, analogous to neurons in the brain. An artificial neuron is a computational model inspired by natural neurons. Each node can be regarded as computational units, combines several inputs from the external environment, and process them (such as summing the inputs) to produce an output, which is then transmitted to many different locations, including other nodes (this node might contain another network). With the help of these processing nodes, a neural network can be used to predict future values of possibly noisy multivariate time-series based on past histories. The objective of the neural network is to transform the inputs into meaningful outputs. A neural network is characterized by its architecture that represents the pattern of connection between nodes, its method of determining the connection weights, and the activation function (Fausett, 1994). The prime goal beside the ANN model is to generalize a simple relationship of the form (Nayak et al., 2006):

$$Y^m = f(X^n) \tag{2}$$

where, X^n is an n-dimensional input vector consisting of variables $x_1 \dots x_i, \dots, x_n$; while Y^m is an m-dimensional output vector consisting of the resulting variables of interest $y_1 \dots y_i, \dots, y_m$. In ground-water level forecasting, x_i may represent precipitation, temperature, and groundwater level values at different antecedent time lags, and the value of y_i is generally the groundwater level for a subsequent period at a specific well (Nayak et al., 2006).

Feed forward neural networks (FNNs)

Feed forward neural networks (FNNs) means that all the interconnections of the nodes between the layer propagate forward to the next layer. In the current proposed study, the activation function used for calculation is a sigmoid logistic function, whose slope is confined in the interval range [-1, 1]. Each node is a simple processing element that responds to the weighted inputs it receives from other nodes. The receiving node sums the weighted signals from all nodes to which it is connected to the preceding layer. The net input x_j to node j is the weighted sum of all the incoming signals:

$$\text{Net\_input} = x_j = \sum w_{ij} y_j \tag{3}$$

Here, x_j =net input coming to node j.

w_{ij} = weight between node i and node j.

y_j = activation function at node j.

The activation function, y_j which is a nonlinear function of its net-input, is described by the sigmoid logistic function using the following equation.

$$y_j = \frac{1}{1+\exp(-x_j)} \tag{4}$$

Since the advent of the error backpropagation learning algorithm, Feedforward neural networks (FNNs) have been applied successfully in many different problems. This network architecture and the corresponding learning algorithm can be viewed as a generalization of the popular least-mean-square (LMS) algorithm (Haykin, 1999). In this type of network, the data flow through the network in one direction from the input layer to the output layer through the hidden layer(s). Each output value is based solely on the current sets of inputs. In most networks, the nodes of one layer are fully connected to the nodes in the next layer; however, this is not a requirement of feed-forward networks. Figure 3 shows a typical feed-forward neural network (Adamowski, 2007) with four input neurons, one hidden layer with four nodes, and one output node. The input signal propagates through the network in a forward direction, layer by layer. The advantage is that they are easy to handle, and can approximate any input/output map, as established by Hornik et al. (1989). The one key disadvantage is that they train slowly, and requires lots of training data.

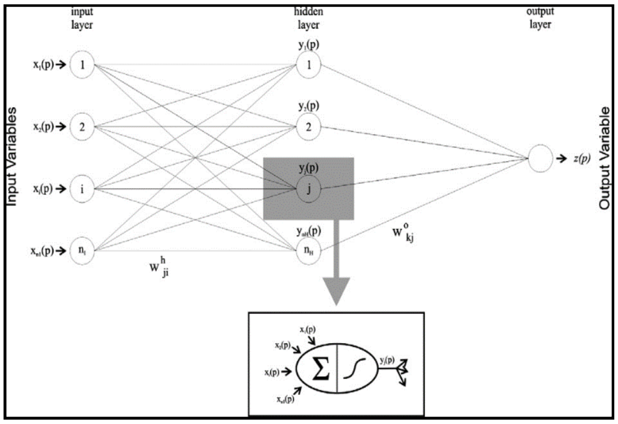

Fig. 3: A Feedforward Artificial Neural Network Architecture (Adamoswki, 2007).

WAVELET ANALYSIS

Wavelet transform (WT) analysis, developed during the last two decades in the mathematics community, appears to be a more effective tool than the Fourier transform (FT) and short-time Fourier transform in analyzing non-stationary time series. Wavelet analysis is the breaking up of a signal into shifted and scaled versions of the original (or mother) wavelet. Wavelet analysis can be used to decompose an observed time series (such as rainfall, evapotranspiration, and groundwater levels). In wavelet analysis, the use of a fully scalable modulated window solves the signal cutting problem. In this study, discrete wavelet transformation is used and it is described below.

I. *Discrete Wavelet Transform (DWT)*

The Discrete wavelet transform (DWT) allows one to reduce the computation time and it is considerably simpler to implement than continuous wavelet transform (CWT). In one-dimensional DWT the signal is split into two parts, usually the high frequency and low-frequency part. This splitting is called decomposition. The signal is passed through a series of high pass filters to analyze the high frequencies, and it is passed through a series of low pass filters to analyze the low frequencies.

$$y = (x * g)[n] = \sum_{k=-\infty}^{\infty} x[k] g[n-k] \tag{5}$$

The DWT of signal x is calculated by passing it through a series of filters. First, the samples are passed through a low pass filter with impulse response 'g' resulting in a convolution of the two.

$$y_{low}[n] = \sum_{k=-\infty}^{\infty} x[k] g[2n-k] \tag{6}$$

$$y_{high}[n] = \sum_{k=-\infty}^{\infty} x[k] h[2n-k] \tag{7}$$

The time series after wavelet decomposition allows one to have a look at the signal frequency at different scales. The discrete wavelet transformallows reducing computation time than CWT. High pass and low pass filters of different cutoff frequencies are used to separate the signal at different scales. The scale ischanged by upscaling and downscaling operations (Cannas et al. 2005). In this study, Haar wavelet and Daubechies wavelets are used to decompose the input signal time series.

II. *Harr Wavelets*

The Haar wavelet operates on data by calculating the sums and differences of adjacent elements. The Haar wavelet operates first on adjacent horizontal elements and then on adjacent vertical elements. After each transform is performed the size of the square which contains the most important information is reduced by a factor of 4. The next step in the image compression process is quantization. In Haar wavelet, the basics functions are scaled and translated versions of a "mother wavelet" $\varphi(t)$. The Haar wavelet transform has many advantages such as it is conceptually simple, fast, memory-efficient since it can be calculated in place without a temporary array and it is exactly reversible without the edge effects that are pro intimate connections with the theory of fractals. The haar wavelets used in this study are shown in Figure 4.

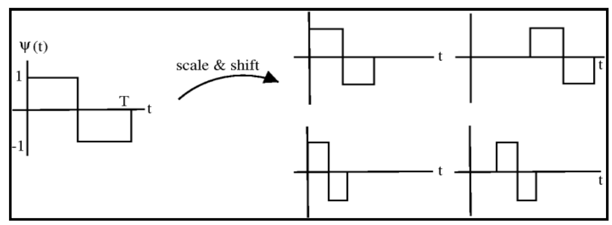

Fig. 4: Type of Harr wavelet used in this study.

III. *Daubechies Wavelets*

The Daubechies wavelets consist of a family of orthogonal wavelets defining a discrete wavelet transform and characterized by a maximal number of vanishing moments for some given support. With each wavelet type of this class, there is a scaling function (also called father wavelet) that generates an orthogonal multi-resolution analysis. The Daubechies wavelets have surprising features-suchas intimate connections with the theory of fractals. This wavelet type has balanced frequency responses but non-linear phase responses. Figure 5 shows the different types of Daubechies wavelets used for this study. Decomposed signals using these wavelets are used as inputs for the WANN model. In this research, input data are split using haar and Daubechies wavelets to analyze non-stationary signals.

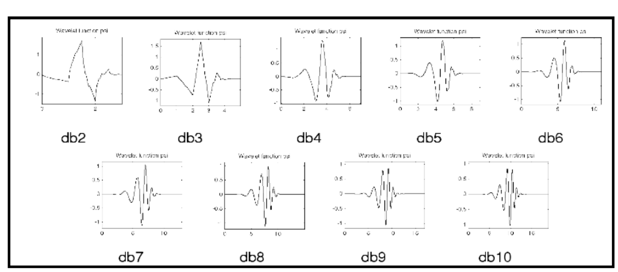

Fig. 5: Type of Daubechies wavelet used in this study.

Training with Bayesian Regularization Algorithm

The next step was to determine the best values of all the weights, called training the ANN. In the so-called supervised learning mode, the actual output of a neural network was compared to the desired output. Weights, which were randomly set to begin with, were then adjusted so that the next iteration produces a closer match between the desired and the actual output. The main objective of training/calibrating a neural network was to produce an output vector Y= ($y_1, y_2,\ldots, y_p$) close as possible to the target vector X= ($x_1, x_2, \ldots, x_p$) before feeding to the ANN. In this process, weight matrices W and bias vectors V were determined by minimizing a predetermined error function as explained as follows:

$$E=\sum_P \sum_P \left(y_i - t_i\right)^2 \quad (8)$$

Here, t_i is a component of the desired output T; y_i is the corresponding ANN output; p is the number of output nodes; and P is the number of training patterns. The training phase can consume a lot of time. It is considered complete when the artificial network reaches a user-defined performance level. At this level the network has achieved the desired statistical accuracy as it produces the required outputs for a given sequence of inputs. When no further learning is judged necessary, the resulting weights are typically fixed for the application.

Bayesian regularization algorithm is used in this study in order to train the given network more efficiently. The Bayesian regularization is an algorithm that automatically sets optimum values for the parameters of the objective function. In the approach used, the weights and biases of the network are assumed to be random variables with specified distributions. In order to estimate regularization parameters, which are related to the unknown variances, statistical techniques are being used. The advantage of this algorithm is that whatever the size of the network, the function won't be over-fitted. Bayesian regularization has been effectively used in literature (Anctil et al.; 2004; Coulibaly et al., 2001a, Porter et al., 2000). The network geometry is generally highly problem oriented in order to get optimal network geometry trial and error procedure is adopted. The numbers of nodes in the input layer were decided based on the inputs to the model. The number of hidden neurons in the network, which is responsible for capturing the dynamic and complex relationship between various input and output variables, was identified by various trials. For each set of hidden neurons, the network was trained with input datasets in batch mode to minimize the mean square error at the output layer.

Coupled wavelet and Artificial Neural Network (WANN)

WANN models are ANN models that use, as inputs, sub-series components (DWs), which are derived from the DWTs of the original time series data. The DWT was performed in this study because it requires less computational effort than the CWT. One of the advantages of the WANN method compared to the ANN method is its ability to identify data components in a time series such as irregular components with multi-level wavelet decomposition (Adamowski and Sun, 2010).

Criteria for model performance evaluation of ANN model

The whole data series was divided into two subsets based on the statistical properties of the time series such as mean and standard deviation. One of the subsets is used during calibration of the model and another subset is used for validation of the model outputs. Various performance criteria are used to assess model performance. Numerous researchers proposed that a perfect evaluation of model performance should include at least one 'goodness-of-fit' or relative error measure (e.g. coefficient of determination (R^2)) and at least one absolute error measure (e.g. Root mean Square Error (RMSE) or Mean Absolute Error (MAE)) (Rajaee, 2011). In this proposed study various performance criteria viz. Coefficient of Efficiency (CE), Root Mean Square Error (RMSE), Explained Variance (EV), and Coefficient of Determination (R^2) are used to assess model performance and its ability to make the precise prediction. Based on the standardization of residual variance with initial variance, the coefficient of efficiency can be used to compare the relative performance of the two approaches effectively. It is expressed as.

$$CE=\left\{1-\frac{residual\ variance}{initial\ variance}\right\}=\left\{1-\frac{\sum_{j=1}^{n}\left(y_j-x_j\right)^2}{\sum_{j=1}^{n}\left(y_j-\overline{y}\right)^2}\right\} \quad (9)$$

Chiew et al. (1993) classified the coefficient of efficiency into three categories viz. perfectly acceptable simulation (C.E. > 0.90), acceptable simulation (C.E. between 0.60 and 0.90), and unacceptable simulation (C.E. < 0.60).

RMSE indicates the discrepancy between the observed and calculated values. The lowest the RMSE, the more accurate the prediction is. It is expressed using the following equation.

$$RMSE=\sqrt{\frac{residual\ variance}{n}}=\sqrt{\sum_{j=1}^{n}\frac{\left(y_j-x_j\right)^2}{n}} \quad (10)$$

Explained Variance measures the proportion to which a mathematical model accounts for the variation (dispersion) of a given data set. It is given by equation

$$EV=\sqrt{\left(\frac{\sum\left(x_j-\overline{y_j}\right)^2}{\sum\left(y_j-\overline{x_j}\right)^2}\right)} \quad (11)$$

The Coefficient of Determination $\left(R^2\right)$ given by equation

$$R^2 = \left(1 - \frac{SSE}{SST}\right)$$

$$\text{SST} = \sum_{i=1}^{N}\left(y_j - \overline{y_j}\right)^2$$

$$\text{SSE} = \sum_{i=1}^{N}\left(y_j - x_j\right)^2 \tag{12}$$

Here, y_j = Observed water table depth, x_j = Predicted water table depth, N= Number of observations, $\overline{y_j}$ = Mean of observed water table depth, $\overline{x_j}$ = Mean of predicted water table depth. SSR = Sum of regression, SSE = Sum of Square Error, and SST = Sum of the square total. The best fit between observed and calculated data, which is unlikely to occur, would have RMSE = 0 and R^2 = 1.

In this study, an attempt has been made to predict the groundwater level of the different observation wells located in the Ur river watershed in the Tikmagarh District, Madhya Pradesh India. An ANN model was developed with the historical data of rainfall, maximum, minimum, and mean temperature, and water table depth of different observation wells in this study area. The best trained ANN model with the input data is derived based on statistical analysis. The original time seriesis decomposed into sub-series using discrete wavelet transform and these decomposed signals were used as input to the WANN model. The input values of quarterly rainfall and evapotranspiration used in the ANN model development were kept the same for each one of the observation well, the prediction of groundwater levels for next month ahead is done for each one of the observation wells using past historic groundwater level data for the particular well. The best prediction model is selected by using the above mentioned various performance criteria parameters.

MODEL DEVELOPMENT

ANN Models

Primarily, ANN models for groundwater level forecasting for each one of the observation well was developed using the computer program in MATLAB 2009b software. The significant input variables for the neural network were selected based on cross-correlation, autocorrelation, and partial autocorrelation techniques to avoid noisy and non-correlated input variables for the modeling system. Three techniques applied in this research work have been listed below:

- The auto-correlation coefficient (Salas et al., 1980) is expressed using an equation.

$$r_k = \frac{\sum_{t=1}^{N-k}\left(x_t - \overline{x_t}\right)\left(x_{t+k} - \overline{x_{t+k}}\right)}{\left[\sum_{t=1}^{N-k}\left(x_t - \overline{x_t}\right)^2 \sum_{t=1}^{N-k}\left(x_{t+k} - \overline{x_{t+k}}\right)^2\right]^{1/2}} \tag{13}$$

Here, r_k is called the lag-k correlation coefficient, the serial correlation coefficient or the autocorrelation function (ACF), x_t is the time series for t = 1,.., N, x_{t+k} is the lagged time series for t = 1,….., N-k, $\overline{x}_t$ is the sample mean for t = 1, … , N, $\overline{x}_{t+k}$ is the sample mean for t = 1,…., N-k, N is the sample size.

- The partial autocorrelation coefficient (Salas et al., 1980) may be obtained recursively by the Durbins relations as expressed using an equation.

$$\phi_k(1) = \rho_1, \quad \phi_1(2) = \frac{\rho_1(1-\rho_2)}{\left(1-\rho_1^2\right)}, \quad \phi_2(2) = \frac{\rho_2 - \rho_1^2}{\left(1-\rho_1^2\right)}$$

$$\phi_k(k) = \frac{\rho_k - \sum_{j=1}^{k-1}\phi_j(k-1)\rho_{k-j}}{1-\sum_{j=1}^{k-1}\phi_j(k-1)\rho_j} \quad \phi_j(k) = \phi_j(k-1) - \phi_k(k)\phi_{x-j}(k-1) \tag{14}$$

To determine the partial auto-correlation function from a sample series x_1 ,…, x_N , in equation (11) the values of the autoregression coefficient, ρ are replaced by the values of the autocorrelation coefficient, r .

- The cross-correlation coefficient (Salas et al., 1980) is expressed using an equation.

$$r_k^{ij} = \frac{\sum_{t=1}^{N-k}\left(x_t^{(i)} - \overline{x_t^{(i)}}\right)\left(x_{t+k}^{(j)} - \overline{x_{t+k}^{(j)}}\right)}{\left[\sum_{t=1}^{N-k}\left(x_t^{(i)} - \overline{x_t^{(i)}}\right)^2 \sum_{t=1}^{N-k}\left(x_{t+k}^{(j)} - \overline{x_{t+k}^{(j)}}\right)^2\right]^2} \tag{15}$$

Here, r_k^{ij} is the lag-k cross-correlation coefficient, $x_t^{(i)}$ is the time series i, $x_t^{(j)}$ is the time series values of series j, $\overline{x_t^{(i)}}$ is the mean of the first N-k values of series i, and $\overline{x_{t+k}^{(j)}}$ is the mean of the last N-k values of series j.

After selecting the significant input variables, it was required to normalize the time series data between the range 0 and 1 before passing into neural networks. Since the sigmoidal function slope is considerably based on their proportion in the interval [-1, 1]. To prevent the saturation of the network, it is required to normalize them using a suitable technique before finalizing them for a neural network. In this study, empirical equation 16 is used to normalize the given data set:

$$N_i = \frac{R_i - Min_i}{Max_i - Min_i} \tag{16}$$

Here, R_i is the real value applied to neuron, i; N_i is the subsequent normalized value calculated for neuron i; Min_i is the minimum value of all values applied to neuron i; Max_i is the maximum value of all values applied to neuron i.

The next step was to determine the best values of all the weights, called training the ANN. The ANN models have been trained using the Bayesian regularization algorithm. For each one of the inputs from different observation wells, the whole time series data set is divided into sets for the training and validation of the ANN model to prevent the complexity of the network. The quarterly groundwater level data from January 2004 to December 2013 (40 months) have been considered for the development of the model. Out of 40 months datasets, the 32 month data set is available for analysis due to consideration of 8 month time lag for rainfall series. In 32 month datasets, 25 sets (78.125%) of data were used for calibration (training), 7 sets(21.875%) of data were used for validation. These data sets were selected by the trial and error method.

WANN Models

Interpretation of time series simultaneously in both spectral and temporal terms helps ANN for better interpretation of the process. The wavelet transform was used to decompose rainfall and evapotranspiration time series using haar and Daubechies wavelets at different decomposition levels. The selection of mother wavelet for different input time-series signals was based on the performance of the model. The optimum level was obtained through a trial-error procedure. In order to have a comprehensive overview of the decomposition level, initially, the following formula was employed which gives a maximum level of decomposition (Nourani, 2008).

$$L = \text{int}\,[\log (N)] \qquad (17)$$

Here, L is the decomposition level and N is the number of time-series data. In this study N=32, thus L=3. This experimental equation was derived for fully autoregressive signals and only considering time-series length without paying attention to seasonal signatures of a hydrologic process (Nourani et al., 2008). Therefore, the other decomposition levels (i.e., 2, 3, and 4) were also examined, the decomposition level which led to better results via the presented modeling and was selected at the optimum decomposing level.

Due to the proportional relationship between the amount of rainfall and evapotranspiration, these signals were supposed to have the same seasonality level and both time series were decomposed at the same level (i.e. level 2). Daubechies-2 (db2) mother wavelets were applied to decompose both rainfall and evapotranspiration time series for this station because the db2 wavelet is similar to theevapotranspiration and rainfall signal so that it could capture the signal features, especially peak points, efficiently and led to comparatively good results.

5. RESULTS

The performance of the best ANN and best WANN models for the prediction of water level depth at Tikamgarh town during calibration and validation is presented in Figure 6 to 9. The scatter plots demonstrate the potential ability of the developed ANN and WANN models for the prediction of water level depth. The performance of the best ANN and best WANN model in terms of observed and computed monthly water table depths of observation well during validation are presented in Figures 10 and 11 respectively. The results of the calibration and validation of the best ANN and best WANN models in terms of various statistical indices are presented in Table 1.

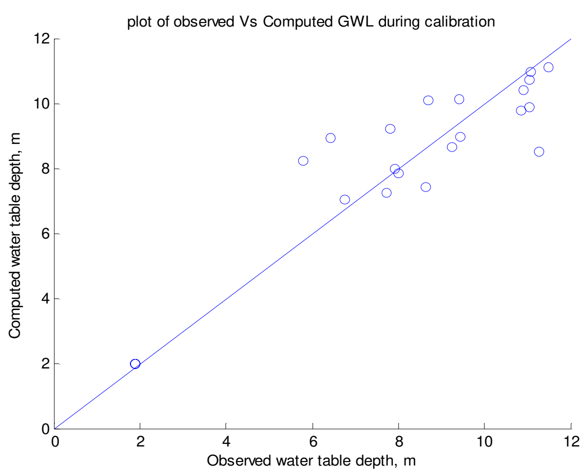

Fig. 6: Scatter plot for the result of the best ANN model for the Tikamgarh Town station during Calibration.

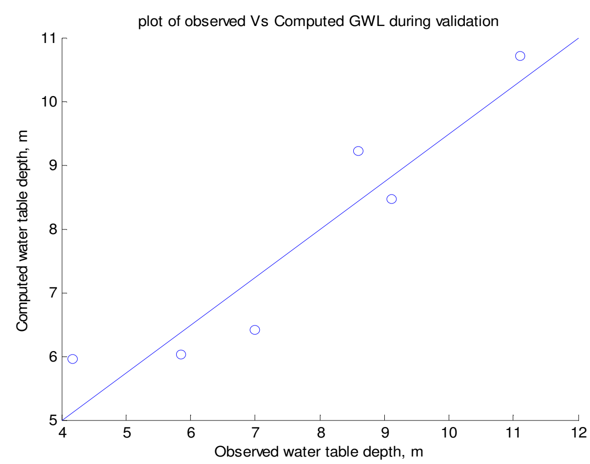

Fig. 7: Scatter Plot for the result of the best ANN model for the Tikamgarh Town station during Validation.

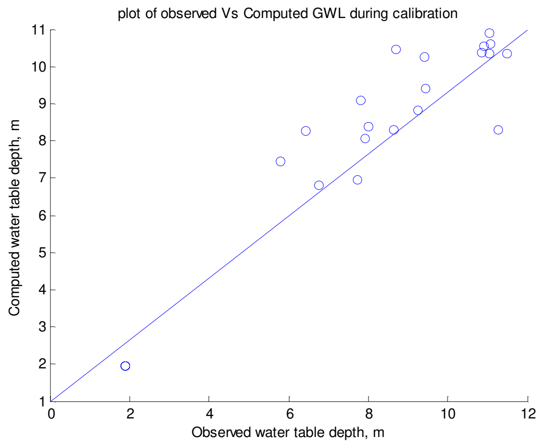

Fig. 8: Scatter Plot for the result of the best WANN model for the Tikamgarh Town station during Calibration.

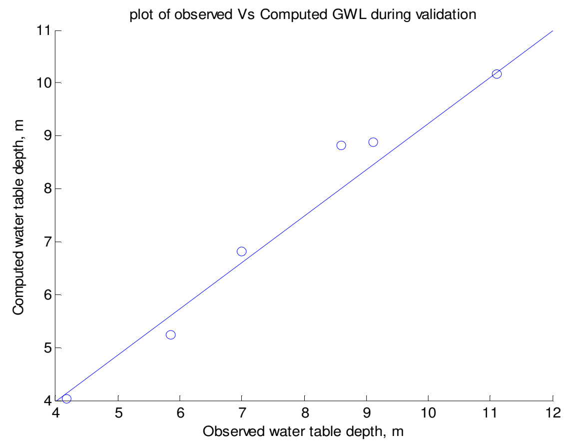

Fig. 9: Scatter Plot for the result of the best WANN model for the Tikamgarh Town station during Validation.

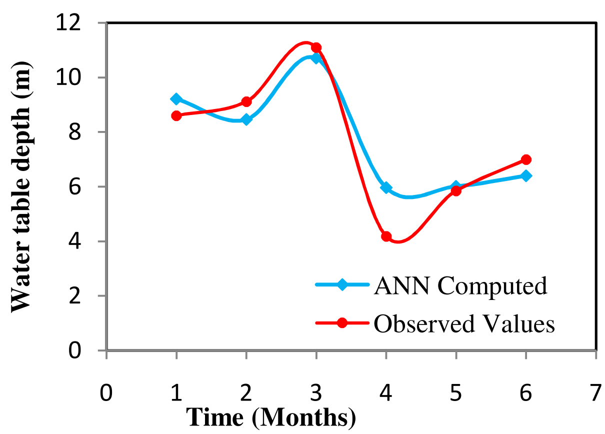

Fig. 10: Observed and computed monthly water table depths by the best WANN model for the Tikamgarh Town station during Calibration.

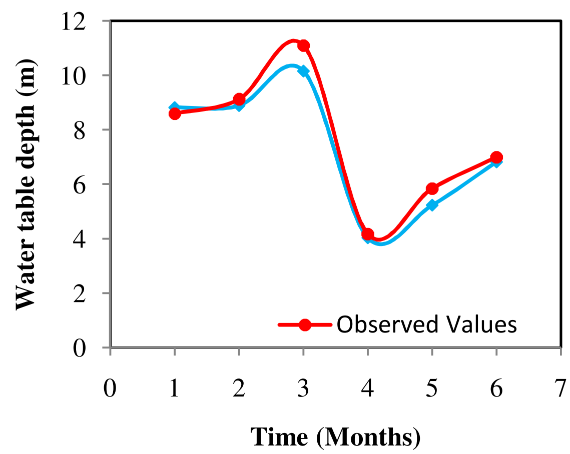

Fig. 11: Observed and computed monthly water table depths by the best WANN model for the Tikamgarh Town station during Validation.

Table 1.Comparison of results between ANN and WANN models for the Tikamgarh Town station.

Station Name	Models	Calibration			Validation		
		CORR	RMSE (m)	EFF (%)	CORR	RMSE (m)	EFF (%)
Tikamgarh Town	ANNGWL11 (8-11-1)	0.95	1.09	**0.90**	0.94	0.87	**0.85**
	WANNGWL11 (8-11-1)	0.96	0.98	**0.92**	0.99	0.49	**0.95**

From Table 1, it can be revealed that the coefficients of correlation with the use of the WANN model are higher during calibration as well as validation as compared to those with the ANN model. The RMSE values obtained with the ANN model, during calibration and validation are higher than those with the use of the WANN model. The model efficiency of the ANN model deteriorates during the validation. whereas the model efficiency of the WANN model is improved for both during validation. From these results, it appears that the WANN model performance is significantly better than the ANN model even though the times series data length is short. Also, the scatter plots of the best ANN and WANN for the predicted and observed groundwater levels during calibration and validation at Tikamgarh town as shown in Figs. 6-9 indicates that the ANN models underestimate the groundwater levels during calibration as well as during validation as compared to the observed levels. However, the WANN model is capable of predicting the groundwater level close to that of the observed values. This observation can also be verified by the examination of Figs. 10 and 11 which shows that the WANN model predicts the groundwater levels at the Tikamgarh town more satisfactorily than the ANN model. Thus, the WANN model can be considered as a more suitable model in predicting the fluctuations of the groundwater level as compared to the ANN model.

Further to check the applicability of these two models at other observation wells, these models applied for the groundwater levels prediction at three different observation wells at Majna, Palera, and Nengawan. Note that the observation well at the Palera is having a shallow aquifer. The ANN and WANN models were developed for each one of the observation wells as per the procedures used for the development of models for the Tikamgarh town observation well. The input variables such as precipitation and evapotranspiration were kept the same for these three observation wells. The observed values of water level fluctuations in these observation wells and their comparison with the values computed by the ANN and WANN models during calibration and validation are shown in Fig. 16 to Fig.24. The scatter plots showing the observed groundwater levels and ANN and WANN models predicted levels at Majna, Palera (Shallow), and Nengawan during calibration and validation are shown in Figs 12-15, Figs. 17-20 and Figs. 22-23, respectively. Further, the performance criteria for these models are shown in Table 2, Table 3, and Table 4 respectively.

From these results, it can be concluded that the WANN model has a better performance in the prediction of groundwater levels at Majna, Palera, and Nengawan stations during calibration and validation as compared to the results by the ANN model.

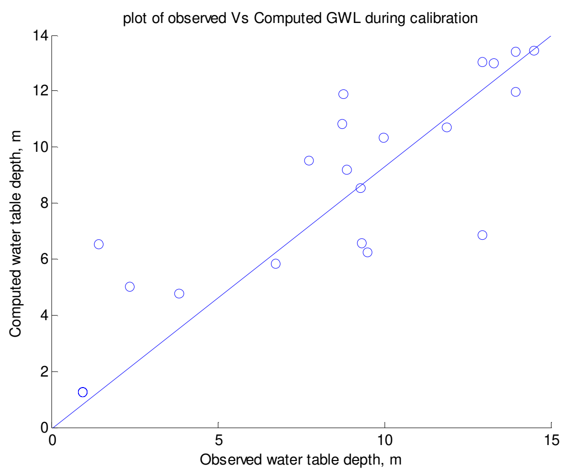

Fig. 12: Scatter plot for the result of the best ANN model for the Majna station during Calibration.

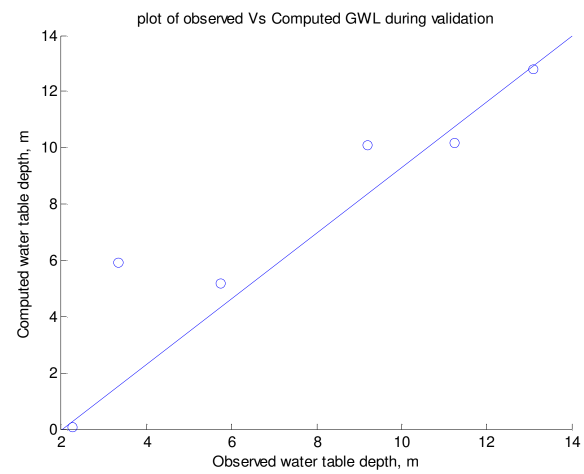

Fig. 13: Scatter plot for the result of the best ANN model for the Majna station during Validation.

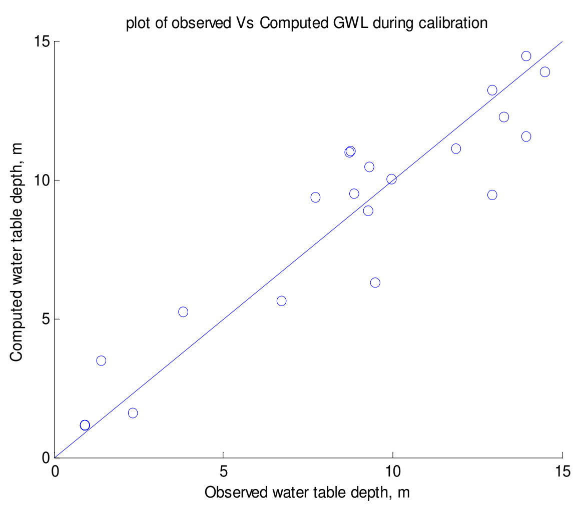

Fig 14. Scatter plot for the result of the best WANN model for the Majna station during Calibration

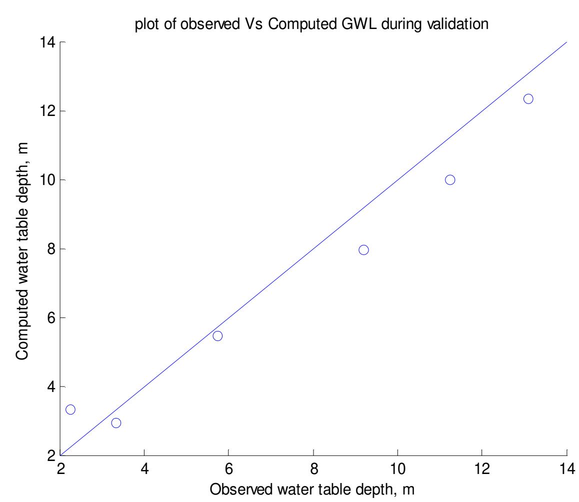

Fig 15. Scatter plot for the result of the best WANN model for the Majna station during Validation

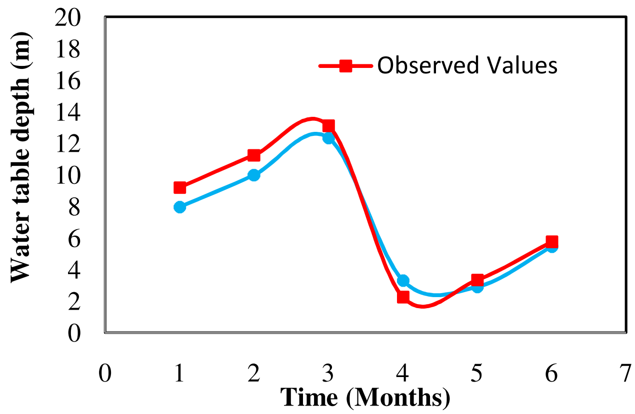

Fig. 16: Observed and computed monthly water table by the best WANN model for the Majna station during Validation.

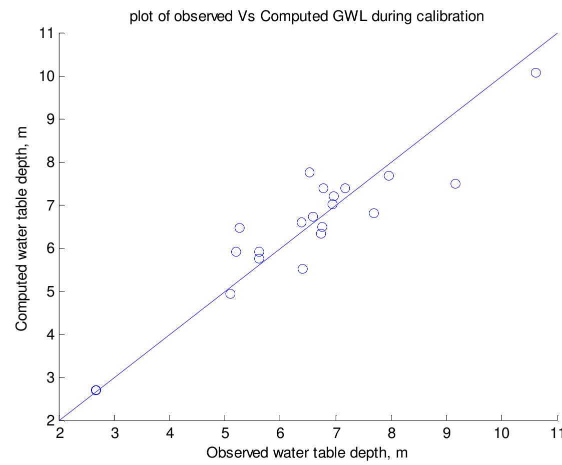

Fig. 17: Scatter plot for the result of the best ANN model for the Palera station during Calibration.

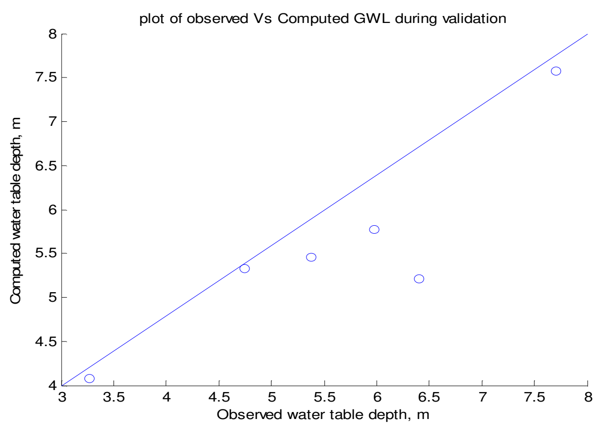

Fig 18.Scatter plot for the result of the best ANN model for the Palera station during Validation.

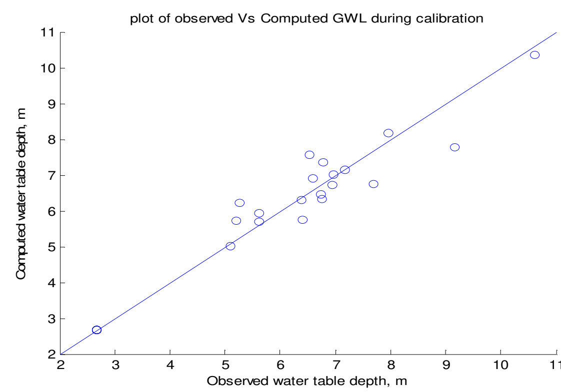

Fig. 19 : Scatter plot for the result of the best WANN model for the Majna station during Calibration.

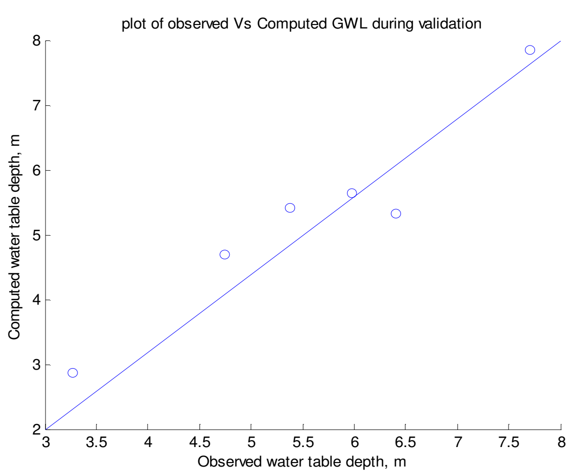

Fig. 20: Scatter plot for the result of the best WANN model for the Majna station during Validation station during Validation.

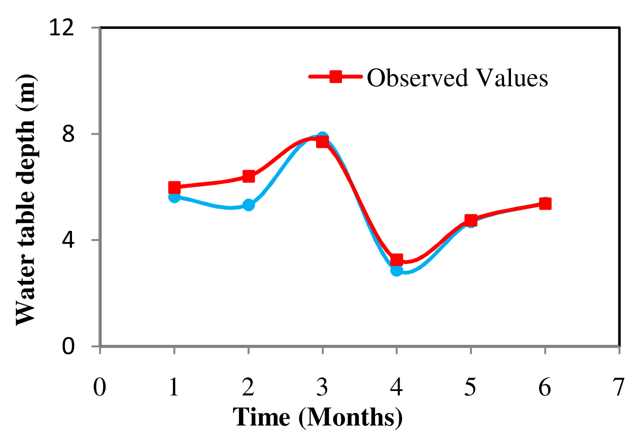

Fig. 21: Observed and computed monthly water table depths by the best WANN model for the Palera

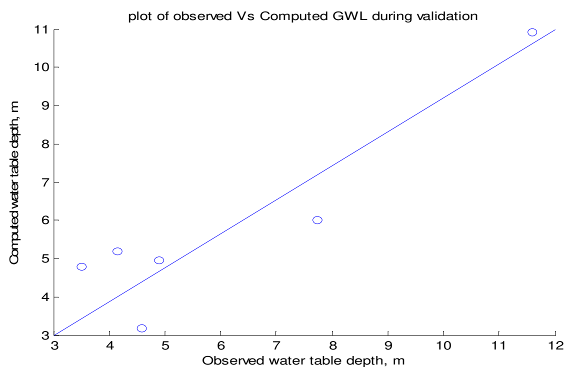

Fig. 22: Scatter plot for the result of the best ANN model for the Nengawan station during Validation

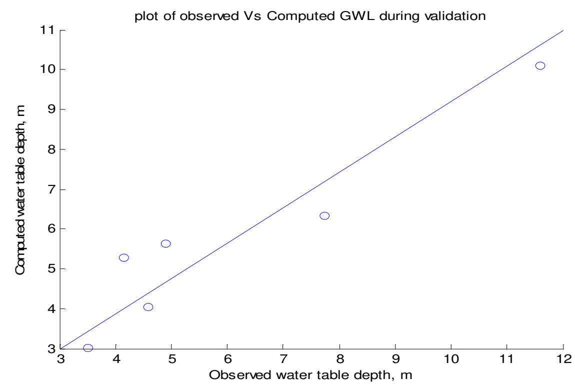

Fig. 23: Scatter plot for the result of the best WANN model for the Nengawan station during Validation.

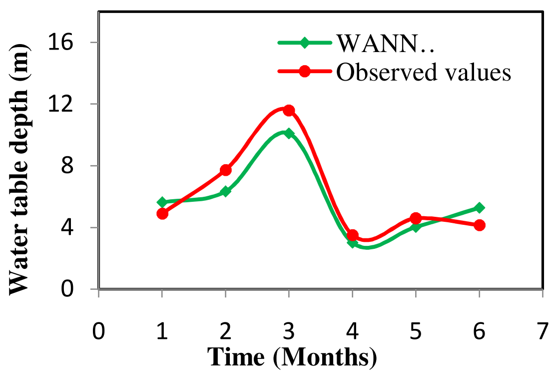

Fig. 24: Observed and Computed monthly water table depths by the Best WANN model for the Nengawan station during Validation.

CONCLUSIONS

ANN model based WANN models have been presented for predicting the groundwater levels in the Ur river watershed in the Tikamgarh district of Madhya Pradesh. This watershed area mainly consists of the granite type of poor aquifers and hence the application of a physically-based groundwater level prediction model seems to be a challenging task. Therefore, this study intended to investigate the applicability of the ANN and WANN model in the prediction of groundwater levels in the Ur river watershed. In the development of these models, the daily rainfall and daily maximum, minimum, and mean temperature data are converted into the quarterly data. Further, the quarterly rainfall, temperature (mean, maximum and minimum), and quarterly measured groundwater level data recorded from January 2004 to December 2013 for 10 years were used for the development

Table 2: Comparison of results between ANN and WANN model for the Majna station.

Station Name	Models	Calibration			Validation		
		CORR	RMSE (m)	EFF (%)	CORR	RMSE (m)	EFF (%)
Majna	ANNGWL3 (8-3-1)	0.90	2.16	**0.81**	0.93	1.50	**0.86**
	WANNGWL3 (8-3-1)	0.95	1.47	**0.91**	0.99	0.92	**0.95**

Table 3: Comparison of results between ANN and WANN models For the Palera Station

Station Name	Models	Calibration			Validation		
		CORR	RMSE (m)	EFF (%)	CORR	RMSE (m)	EFF (%)
Palera	ANNGWL5 (11-5-1)	0.96	0.60	**0.92**	0.90	0.64	**0.78**
	WANNGWL5 (1157-1)	0.97	0.51	**0.94**	0.96	0.49	**0.87**

Table 4: Comparison of results between ANN and WANN models for the Nengawan station.

Station Name	Models	Calibration			Validation		
		CORR	RMSE (m)	EFF (%)	CORR	RMSE (m)	EFF (%)
Nengawan	ANNGWL7 (9-7-1)	0.93	1.32	**0.87**	0.91	1.16	**0.82**
	WANNGWL7 (9-7-1)	0.93	1.31	**0.87**	0.94	1.04	**0.86**

of both of these models. The evapotranspiration is calculated using the Hargreaves temperature model. The statistical parameters such as ACF, PACF, and CCF have been used for the selection of input vector. The feed-forward neural network architecture was trained with a Bayesian regularization algorithm having input, hidden, and output nodes. The number of neurons in the hidden layer is optimized to 15 by trial and error, network parameters are also optimized by the trial and error method. Out of 40 months datasets, 32 month data set is available for analysis due to consideration of 8 month time lag for rainfall series. For each one of the inputs, out of 32 months datasets, 25 sets (78.125%) of data is used for training, 7 sets (21.875%) of data is used for validation.

Similarly, the WANN model is developed by using decomposed signals of rainfall and evapotranspiration time series by DWT as input data in the ANN model. The optimum level of decomposition was decided based on model performance. The selection of mother wavelet type was also based on the model performance which was investigated using different types of mother wavelets viz. db2, db3, db4, haar3, haar4, etc. However, previous studies were done by B. Krishna et al. and Jan Adamowski et al. [23] who argued that to develop the WANN model, the decomposition level was selected based on the empirical formula given by Eq. (17) were not found appropriate. The performance criteria such as coefficient of correlation, root mean squared error (RMSE), and model efficiency have been used to evaluate the performance of both the model.

The analysis of the predicted groundwater levels during calibration and validation by the ANN and WANN models for three ground observation wells indicates that the ANN and WANN models are applicable for the prediction of groundwater levels in the hard rock area particularly in the Ur River watershed. Further, the use of the WANN model resulted in significant improvement in groundwater levels prediction closely with the observed groundwater levels data as compared to the ANN models at all the four considered groundwater observation wells. Based on this study it can be envisaged that the WANN model is most suitable for prediction in the hard rock areas.

REFERENCES

1. Anctil, F., Perrin, C., Andreassian, V., 2004. “Impact of the length of observed records on the performance of ANN and of conceptual parsimonious rainfall-runoff forecasting models”. Environ. Modeling Software., 19(4): 357-368.
2. ASCE Task Committee on Application of Artificial Neural Network in Hydrology. 2000a. “Artificial neural networks in hydrology-I: Preliminary concepts”. J. Hydrologic Engineering, 5(2): 115-123.
3. ASCE Task Committee on Application of Artificial Neural Network in Hydrology. 2000b. “Artificial neural networks in hydrology-II: Preliminary concepts”. J. Hydrologic Engineering, 5(2): 124-137.
4. Aussem, A., Campbell, J., and Murtagh, F., 1998. “Wavelet-based feature extraction and decomposition strategies for financial forecasting”. J. Comput. Intell. Financ., 6: 5-12.
5. Azhar K., Affandia, Kunio Watanabe, and Haryadi Tirthomihardjo, 2007. “Application of an artificial neural network to estimate groundwater level fluctuation”. J. Spatial Hydrol., 7(2): 23-46.
6. Bhabagrahi Saho, Imtisnla Walling, Bidyut C. Deka, and Bhagwati P. Bhatt., 2012. “Standardization of Reference Evapotranspiration Models for a Subhumid Valley Rangeland in the Eastern Himalayas”. J. Irrig. Drain Eng. 138: 880-895.
7. Cannas, B., Fanni, A., See, L., and Sias, G., 2006. “Data pre-processing for river flow forecasting using neural networks: wavelet transforms and data partitioning”. Phys. Chem. Earth., 31: 164-1171.
8. Chiew, F. H. S., Stewardson, M. J., and McMahon, T. A., 1993. “Comparison of six rainfall-runoff modeling approaches”. J. Hydrology, 147: 1-36.
9. Coppola, E., Rana, A., Poulton, M., Szidarovszky F., and Uhl, V., 2005. “A neural network model for predicting water table elevations”. Ground Water, 43(2): 231-241.
10. Coppola, E., Szidarovszky, F., Poulton, M., and Charles, E., 2003. “Artificial neural network approach for predicting transient water levels in a multilayered groundwater system under variable state, pumping, and climate conditions”. J. Hydrol. Eng., 8(6): 348-360.
11. Coulibaly, P., Anctil, F., Aravena, R., and Bobe’e, B., 2001a. “Artificial neural network modeling of water table depth fluctuations”. Water Resour. Res. 37(4): 885-896.
12. Daliakopulos, I. N., Coulibaly, P. and Tsanis, I.K., 2005. “Groundwater level forecasting using artificial neural networks”. J. Hydrol., 309: 229-240.
13. Dawson, C.W., and Wilby, R., 1998. “An artificial neural network approach to rainfall-runoff modeling”. Hydrol. Sci. J., 43(1): 47-66.
14. Deb Roy Aditi., Tushaar Shah., 2002. “The Socio-Ecology of groundwater in India”. Water Policy Briefing, Issue 4 IWMI-TATA. FAI.
15. Fausett, L. 1994. “Fundamentals of Neural Networks”. Pentice Hall, Englewood Cliffs, N.J.
16. French, M. N., Krajewski, W. F., and Kuykendall, R. R., 1992. “Rainfall forecasting in space and time using neural network”. J. Hydrol., 137: 1-31.
17. Govindaraju, R. S., Ramachandra Rao, A., 2000. “Artificial Neural Networks in Hydrology”. Kluwer Academic Publishing, The Netherlands.
18. Hagan, M.T., H.B. Demuth, and Beale, M., 1996. “Neural Network design”. Boston, Mass., PWS.

19. Haykin, S., 1999. "Neural Networks, A Comprehensive Foundation", second ed. Prentice-Hall, Englewood Cliffs, NJ.
20. Hertz, J., A. Krogh, and Palmer. R.G., 1991. "Introduction to the theory of neural computation reading", Mass., Addison Wesley.
21. Hopfield, J. J., 1982. "Neural networks and physical systems with emergent collective computational abilities". Proc., Nat. Academy of Scientists. 79: 2554-2558.
22. Hornik, k., Stinchcombe, M., White, M., 1989. "Multilayer feedforward networks are universal approximators". Neural Networks 2, 359-366.
23. http://www.engr.scu.edu/~emaurer/tools/calc_solar_cgi.pl.
24. Jan Adamowski, Hiu Fung Chan, 2011. "A wavelet neural network conjunction model for groundwater level forecasting". J. Hydrol., 407, 28-40.
25. Krishna, B., Satyaji Rao, Y. R., and Vijaya, T., 2008. "Modelling groundwater levels in an urban coastal aquifer using artificial neural networks". Hydrol. Process., 22: 1180-1188.
26. MATLAB 2009b (The Mathworks, Inc., 2001)
27. Maier, H. R., and Dandy, G. C., 2000. "Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications", Environ. Modell Softw., 15: 101-124.
28. McCulloch, W. S., and Pitts, W., 1943. "A logic calculus of the ideas immanent in nervous activity". Bulletin of Math's Biophysics. 5: 115-133.
29. Mohanty, S., Jha, M. K. Kumar, A., Srivastava, R. C. and James, B. K., 2007, "Groundwater level forecasting in Kathajodi River basin along with artificial neural network". Third International Groundwater Conference (IGC-2007), held at TNAU, Coimbatore, India, TS-5-12.
30. Nayak, P. C., Satyaji Rao, Y. R., and Sudheer, K. P., 2006. "Groundwater level forecasting in a shallow aquifer using artificial neural network approach". Water Resour. Manag., 20: 77-90.
31. Nourani, V., Kumasi, M., and Mano, A. 2009a. "A multivariate ANNWavelet approach for rainfall-runoff modeling". Water Resour. Manage., 23: 2877-2894.
32. Deka, P. C., Haque, L., and Bamhatti, A. G., 2012. "Discrete Wavelet-ANN Approach in Time series Flow Forecasting- A Case study of Brahmaputra River", International Journal of Earth Science and Engineering. 5(4): 673-685.
33. Porter, D.W., Gibbs, P.G., Jones, W.F., Huyakorn, P.S., Hamm, L.L., Flach, G.P., 2000. "Data fusion modeling for groundwater systems". J. Contaminant Hydrol., 42: 303-335.
34. Principe, J.C., Euliano, N.R., and Lefebvre, W.C., 2000. "Neural and adaptive systems: fundamental through simulation". New York, John Wiley.
35. Rajeev R Sahay, 2008. 'Coupled wavelet-autoregression models for predicting monsoon flows for the Kosi River (India)". Journal of Mechanical and Civil Engineering. 2320(334X), 43-53.
36. Salas, J. D., Delleur, J. W., Yevjevich, V., and Lane, W. L. 1980. "Applied Modeling of Hydrological Time Series", Water Resource Publications, Denver.
37. Satyaji Rao, Y. R., Krishna, B., and Venkatesh, B., 2014. "Wavelet-Based Neural Networks for Daily Stream Flow Forecasting". International Journal of Emerging Technology and Advanced Engineering, 4(1): 307-317.
38. Vahid Nourani and Masoumeh Parhizkar, 2013. "Conjunction of SOM-based feature extraction method and hybrid wavelet-ANN approach for rainfall-runoff modeling". J. Hydroinform., 15(3): 829-848.
39. Venkata Ramana. R., Krishna. B., Kumar. S. R., and Pandey, N.G., 2013. "Monthly Rainfall Prediction Using Wavelet Neural Network Analysis". Water Resource Manage. 27: 3697-3711.
40. Vijay Shankar, P. S., Himanshu Kulkarni, Sunderrajan Krishnan, 2011. "India's Groundwater Challenge and the Way Forward". Economic & Political Weekly. 2: 37-45.
41. Singh, B., 1998. "Water resources of India". J. Indian Water Works Association. Oct-Dec., pp. 229-231.
42. Vijay Shankar, P. S., Himanshu Kulkarni, Sunderrajan Krishnan, 2011. "India's Groundwater Challenge and the Way Forward". Economic & Political Weekly. 2: 37-45.
43. Wanakule, N., and Aly, A., 2005. "Artificial neural networks for forecasting groundwater Levels, In: Computing in Civil Engineering Proceedings of the 2005 ASCE International Conference on Computing in Civil Engineering", Edited by Lucio Soibelman and Feniosky Pena-Mora, July 12-15, 2005. Cancun, Mexico.
44. Wasserman, P.D. 1989, "Neural computing: theory and practice". Van Nostrand Reinhold, New York.
45. Wyrwoll, P., 2012. 'India's groundwater crisis, GWF Discussion Paper 1228, Global Water Forum, Canberra, Australia.