Int.J.Curr.Microbiol.App.Sci (2021) *10*(01): 3327-3334

International Journal of Current Microbiology and Applied Sciences
ISSN: 2319-7706 **Volume 10 Number 01 (2021)**
Journal homepage: http://www.ijcmas.com

Original Research Article https://doi.org/10.20546/ijcmas.2021.1001.390

Forecasting Groundwater Level Fluctuation of Veppanthattai Block using Artificial Neural Network

A. Saranya[1*], U. Arulanandu[2], M. Kalpana[2], P. Sujatha[3] and R. Parimalarangan[4]

[1]Agricultural Statistics, [4]Department of Agricultural Economics, Tamil Nadu Agricultural University, Coimbatore, India
[2]Department of Social Sciences, Anbil Dharmalingam Agricultural College and Research Institute, Thiruchirappalli, India
[3]Agricultural Engineering College and research institute, Thiruchirappalli, India

**Corresponding author*

ABSTRACT

Keywords

Artificial Neural Network, Feed Forward Neural Network, Back Propagation, Levenberg-Marquardt algorithm

Article Info

Accepted:
20 December 2020
Available Online:
10 January 2021

Groundwater is an essential source of water for the domestic, agricultural, and industrial sectors. Due to over-extraction, the trend of groundwater levels declining continues steadily. So, there is a need to monitor the behavior of fluctuations and the prediction of groundwater levels for making effective policies and management practices that support sustainable groundwater usage. In this study, fluctuations in the groundwater level of the Veppanthattai block observation wells were forecasted using Artificial Neural Networks (ANN). Multilayer Feed Forward Neural Network (FFNN) was selected for the network architecture and Levenberg- Marquardt (LM)algorithm was used for training the data. The performance of the network was evaluated by Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE) and Theil's U and the optimal networks of observation wells were discussed.

Introduction

Groundwater makes up about twenty percent of the world's freshwater supply, which is about 0.61per cent of the entire world's water. Besides, it is an essential source of water for the domestic, agricultural, and industrial sectors. Groundwater is highly susceptible to meteorological drought where the water level is limited. Due to monsoon failure accompanied by overexploitation, a drastic reduction has been reported in Perambalur district (5.5m), according to monthly assessment by State Ground and Surface Water Resources Data Centre.

A clear understanding of the groundwater level dynamics is required for making effective management practices to avoid stress and shocks. Several methods have been deployed to study the dynamics of the groundwater level. Among them, some traditional techniques used to model and forecast the groundwater levels rely on the profound information of the observed system dynamics and they require spatial, geological, and hydrological properties of the aquifer.

Artificial Neural Network (ANN) has been successfully employed in many hydrological modelling techniques with relatively higher accuracy than traditional methods. ANN is a soft computing tool inspired by the activities and learning mechanism of the human brain. It is treated as universal approximators, suited well for non-linear dynamic system modelling. ANN learns from past data through the neurons similar to that of the brain. Neurons are highly interconnective structures that update the weights to give the desired output. ANN models are the black-box models, which behave like the human brain's biological neuron system (Haykin *1994*). ANN was first introduced by McCulloch and Pitts (1943) but became more popular after Rumelhart *et al.,* (1986) when the error back propagation method for ANN training was developed. Since then, ANNs are widely accepted by cross disciplines and suitable for applying hydrological processes due to their informative processing characteristics, such as nonlinearity, parallelism, noise tolerance, and learning and generalization capabilities in precipitation groundwater estimation. Some studies in the literature compared the results of data-driven ANN models with any other numerical prediction methods, wherein ANN results showed that the performance of ANN is better than other numerical methods.

The study aims to develop a neural network architecture to model the groundwater levels of the Veppanthatttai block of the Perambalur district and evaluate their prediction capabilities. The rest of the paper is structured as follows. Section II briefly explains the materials and methodology. Section III presents the results and discussions of the developed ANN. Finally, the conclusions drawn from this work are summarized in Section IV.

Materials and Methods

Study area: The geographical location of Veppanthattai block is 575 km², located in Perambalur District. It lies between 10°13’44” to 10°26’47” N latitude and 77°53’8” to 78°01’24” E longitude and situated at an altitude of 149 m above mean sea level. The district receives the rainfall under the influence of both the southwest and northeast monsoon. Veppaanthattai block enjoys a typical semi-arid climate with hot summers and moderately cool winters. The hottest season is from March to May. During the period, the maximum temperature often exceeds 40°C. The winter season is spread over two months, viz. January and February. The district generally has high humidity. The district experiences strong winds during the southwest monsoon season. The wind speed from June to August is more than 25 km/hr. After that, there is a gradual decrease in wind speed, reaching the lowest 7.7 km/hr value. The area is a highly active region for agriculture activities. The chief irrigation sources in the area are tanks, dug wells and tube wells, and canals. The block has a total irrigated area of about 14181ha (Dept. of Economics & Statistics, Tamil Nadu). There are seven observation wells of the Veppanthattai block in the study area namely, Iraiyur, Veppanthattai, Vengalam, Pasumbalur, Periyavadakarai, Arumbavur, and Mettupalayam were taken to represent the area. Black cotton soil and red loamy soil are the major types of soil present in the region. The major agricultural crops cultivated in the area include paddy, maize, millets, groundnut, sugarcane, cotton, and oilseeds and the horticultural crops grown are small onion, turmeric, tapioca, coriander, chilies, and some medicinal and aromatic crops.

Data preparation and challenges in data preparation: The monthly water table depth of groundwater of wells Veppanthattai block from the year 2000 to 2019 is collected from the Public Works Department, Chennai. Seven wells are chosen for the study as they better represent the study area. Rainfall data collected from the Department of Disaster Management have continuous patches of missing observations for all the rain gauge stations. So, performing imputation may violate the results. Despite more missing observations found in the real rainfall data collected from the Department of Disaster Management, gridded month total of rainfall data is acquired from the NASA Power from 2000 to 2019.

Missing observation is a common problem observed in groundwater level data. Missing values are non-ignorable; it may reduce network performance. So missing values should be replaced by a proper imputation algorithm. Depending on the pattern of missing values, an appropriate algorithm of imputation can be used.

For this study, many imputation algorithms have been analyzed for the imputation. On CRAN, *imputeTS* is the only package solely dedicated to univariate time series imputation. Finally, the *imputeTS* package in R is adopted for the imputation. From the imputeTS package, the na_interpolation function is used for the imputation. It has a function called spline; from that, the cubic spline is used.

Spline interpolation: In general, to find the missing values of the data set in a graph, a combination of polynomials that gives the best fit can be used. But to find the missing values with some local irregularities, then to fit the non-linear sub-region of the data with different polynomial can be used. Splines are polynomial, used to mathematically reproduce flexible shapes. Knots are placed at several places within the data range, to identify the points where adjacent functional pieces join each other. For the study, cubic spline polynomial is used to interpolate missing values in the series with the non-linear discontinuities. The cubic spline method is a popular method among spline techniques. Cubic spline interpolation, which fits the given points by a piecewise polynomial function $s(x)$, known as the spline, a composite function formed by n low-degree polynomials $P_i(x)$ each fitting $f(x)$ in the interval between x_{i-1} and $x_i, (i = 1, \ldots n)$. The graph of cubic spline has a continuous slope and curvature throughout the region. From the nature of the sub-region, the coefficients of the equation a, b, c, and d are found and then the missing values are interpolated. The plot of the imputed values of observation well located at Iraiyur is shown in Fig. I.

$$s(x) = \begin{cases} P_1(x) & x_0 \le x \le x_1 \\ \vdots & \vdots \\ P_i(x) & x_{i-1} \le x < x_i \\ \vdots & \vdots \\ P(x) & x_{n-1} \le x \le x_n \end{cases}$$

Where,
$P_i = a_i + b_i x + c_i x^2 + d_i x^3 (i = 1, \ldots n)$

Input selection: The selection of inputs for the particular well is based on the lag of the well and the lags of nearby wells. The inputs of a site-specific well are selected through cross-correlation analysis. Lags with greater significant value were taken as the inputs for the modelling. The inputs selected for the site-specific well is given in Table I.

Selection of a Suitable Neural Network Model: Multilayer FFNN is selected to design the network. Fig. III shows the diagrammatic representation of the developed ANN. The input vectors and the output to the model were normalized using Minimax scaling to introduce the non-linearity then, it was reverse-backed to the original units. The training and testing data were allocated as 85 percent and 15 per cent respectively.

To consider the efficiency of every algorithm and reach the best-desired conditions, several parameters, and variables such as number of hidden layers, number of neurons in hidden layers, activation function, learning rate and number of repeating epochs were varied. Finally, the results obtained for each well were compared for accuracy by using MAE, RMSE, MAPE and Theil's U. Then, the network which showed higher accuracy for training and testing subsets were selected and tabulated (Table II and Table III).

Pre-requisites - division, pre-processing and post-processing of data: The whole data set must be split into training and testing sub-sets to train and test the network. In our study, the training and testing subsets are divided into 85 per cent and 15 per cent, respectively. Generally, the data set should be scaled to both the input vectors and the target vectors in the data set. In this way, the network output always falls into a normalized range. Then the output vector can be reverse backed to get the original unit of the data. Minimax scale is used in the study to normalize the data.

Development of the Neural Network Model: FFNN often has one or more hidden layers of sigmoid neurons followed by an output layer of linear neurons. Multiple layers of neurons with nonlinear transfer functions allow the network to learn nonlinear relationships between input and output vectors. Sigmoid and tanh activation functions are used in the neurons of hidden layers to introduce the data's non-linearity.

In this study, the whole process of neural networking is performed by using the *neural net* package from R programming. Theoptimal number of hidden layers and the number of neurons in the hidden layer was determined by testing all reasonable possible values for each well until the best match between observed and predicted groundwater level changes. The maximum number of neurons in the hidden layer was set at 15 and the maximum number of the hidden layer is three. The LM algorithm is considered to be one of the most efficient learning algorithms [Maier and Dandy, 2000], was selected for back-propagation and updating the connection weights.

Evaluation criteria: The network performance of NN in the training and the testing set are assessed by using Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root mean square error (RMSE) and Theil's U.

$$MAE = \frac{1}{n}\sum |Y_i - \hat{Y}_i|$$

$$MAPE = \frac{\left(\sum_{i=1}^{n}\left(\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right|\right) \times 100\right)}{n}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2}{n}}$$

$$U = \sqrt{\frac{\sum_{t=1}^{n-1}(\hat{Y}_{t+1} - Y_{t+1})^2}{\sum_{t=1}^{n-1}\left(\frac{Y_{t+1} - Y_t}{Y_t}\right)^2}}$$

Results and Discussion

Multilayer FFNN is selected to design the network. Fig. III shows the diagrammatic representation of the developed ANN. The input vectors and the model's output were normalized using Minimax scaling to introduce the non-linearity; then, it was reverse-backed to the original units. The training and testing data were allocated as 85 percent and 15 per cent, respectively. To consider every algorithm's efficiency and reach the best-desired conditions, several parameters and variables such as number of hidden layers, number of neurons in hidden layers, activation function, learning rate and number of repeating epochs were varied. Finally, the results obtained for each well were compared for accuracy by using MAE, RMSE, MAPE and Theil's U. Then, the network which showed higher accuracy for training and testing subsets were selected and tabulated (Table II and Table III).

Table.1 Inputs selected for the site-specific wells

Well no.	Lag between inputs and output	Inputs							
		Water table depth of wells							
		Iraiyur	Veppanthattai	Vengalam	Pasambalur	Periyavadakarai	Arumbavur	Mettupalayam	Rainfa
Iraiyur	One month	0.830**	0.230**	0.488**	0.415**	0.232**	0.348**	0.618**	-0.226
	Two months	0.597**	0.086**	0.370**	0.338**	0.155**	0.256**	0.484**	-0.376
Veppanthattai	One month	0.382**	0.796**	0.451**	0.487**	0.625**	0.319**	0.505**	-0.084
	Two months	0.346**	0.588**	0.397**	0.457**	0.561**	0.228**	0.471**	-0.198
Vengalam	One month	0.572**	0.393**	0.837**	0.523**	0.289**	0.301**	0.707**	-0.180
	Two months	0.519**	0.309**	0.678**	0.501**	0.250**	0.239**	0.655**	-0.283
Pasambalur	One month	0.412**	0.389**	0.472**	0.943**	0.619**	0.331**	0.581**	-0.290
	Two months	0.329**	0.305**	0.404**	0.874**	0.563**	0.296**	0.520**	-0.335
Periyavadakarai	One month	0.294**	0.550**	0.297**	0.665**	0.928**	-0.379**	0.372**	-0.165
	Two months	0.282**	0.458**	0.269**	0.651**	0.852**	-0.353**	0.357**	-0.216
Arumbavur	One month	0.430**	0.335**	0.317**	0.342**	0.385**	0.854**	0.296**	-0.107
	Two months	0.382**	0.247**	0.270**	0.306**	0.325**	0.692**	0.281**	-0.196
Mettupalayam	One month	0.645**	0.332**	0.634**	0.593**	0.299**	0.211**	0.915**	-0.309
	Two months	0.535**	0.189*	0.536**	0.543**	0.221*	0.135*	0.785**	-0.419

Table.2 Training accuracy the optimized networks of site-specific observation wells

Location of the observation well	No of the NN	Activation function	MAE	MAPE	RMSE	Theil's U
Iraiyur	4,2,2	tanh	0.714	0.192	0.932	0.173
Veppanthattai	3,1,1	tanh	1.792	0.181	2.562	0.214
Vengalam	3,3,3	tanh	0.920	0.074	1.286	0.0975
Pasumbalur	1,1	logistic	0.910	0.1802	0.160	0.175
Periyavadakarai	3,2,2	tanh	1.024	0.214	0.178	0.177
Arumbavur	6,2	tanh	1.031	0.081	0.0786	0.097
Mettupalayam	5,4,2	tanh	0.437	0.114	0.107	0.113

Table.3 Testing accuracy

Location of the observation well	No of the NN	Activation function	MAE	MAPE	RMSE	Theil's U
Iraiyur	4,2,2	tanh	1.037	0.191	1.651	0.253
Veppanthattai	3,1,1	tanh	1.960	0.127	2.509	0.149
Vengalam	3,3,3	tanh	0.057	0.058	1.148	0.067
Pasumbalur	1,1	logistic	1.46	0.104	1.977	0.139
Periyavadakarai	3,2,2	tanh	0.963	0.073	1.580	0.118
Arumbavur	6,2	tanh	2.236	0.142	2.499	0.136
Mettupalayam	5,4,2	tanh	0.893	0.109	1.130	0.121

Fig.1 Imputed values of Iraiyur well

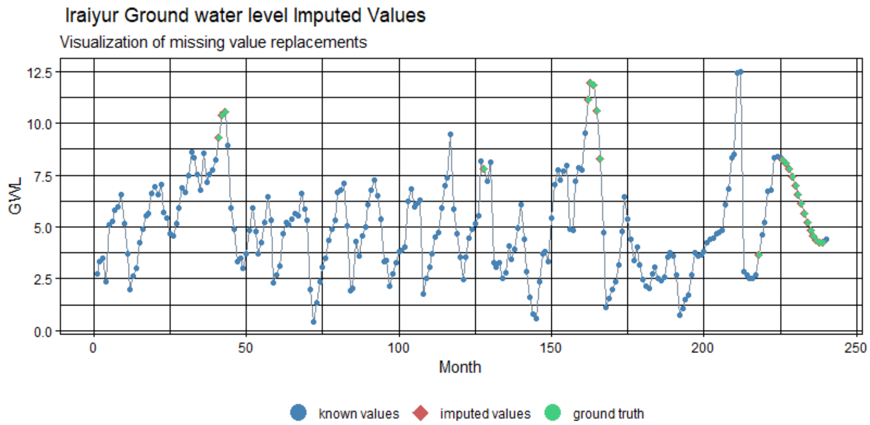


Fig.2 FFNN Model

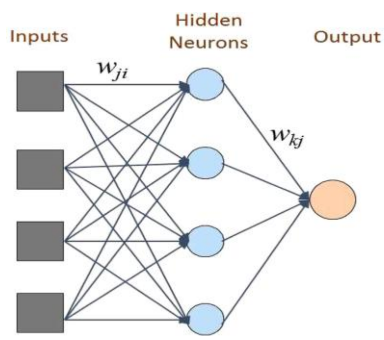
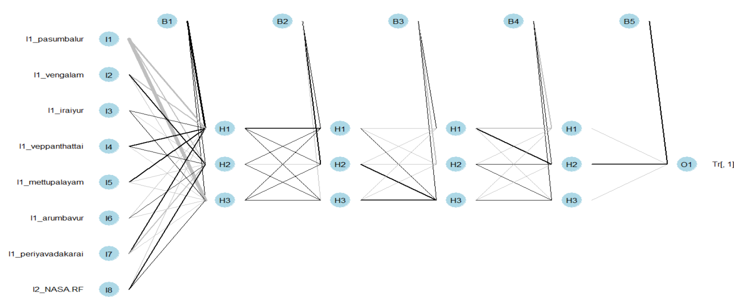


Fig.3 Diagrammatic representation of developed FFNNof observation well



In conclusion the study aimed to construct the Feed Forward Neural Network (FFNN) model and to test the ability to predict groundwater level fluctuation in Veppanthattai block, Perambalur District of Tamil Nadu. Month total of rainfall and the monthly average of groundwater level are the parameters used to forecast the groundwater level of the study area. The available data to simulate groundwater were the monthly total rainfall and the monthly average of groundwater level during 20 years period. The inputs are selected based on the cross-correlation analysis of each well with the lags of the well, the lags of the nearby wells and the lags of the rainfall data. The whole data were divided into two sub-sets namely training(85 per cent) and testing (15 per cent) sub-set. The network for each well was created for all possible combinations of the number of hidden layers, the number of neurons in the hidden layers and the activation function. The best-fitted network of each well was selected based on the performance evaluation criteria. The developed ANN model provided a good prediction of groundwater levels at all 7 sites, with considerably lower values of accuracy metrics. In general, it is concluded that the ANN-based algorithms were a better choice for the prediction of water table depths. The ANNmodeling approach as presented in this study can provide decision-makers, scientists, and water managers to know the dynamics of groundwater level fluctuations and to make effective management strategies.

References

McCulloch, W. S., and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. The bulletin of mathematical biophysics. *5*(4), 115-133.

Haykin, S., 1999. Neural networks, a comprehensive foundation second edition by prentice-hall, Englewood Cliffs, NJ.

Moritz, S. and Bartz-Beielstein, T., 2017. imputeTS: time series missing value imputation in R. Contributed research article. 9 (1), p.207.

Govindaraju, S., 2000. Artificial neural networks in hydrology.Journal of Hydrologic Engineering, Vol. 5, No. 2, 19090.

Perperoglou, A., Sauerbrei, W., Abrahamowicz, M. and Schmid, M., 2019. A review of spline function procedures in R. BMC medical research methodology, 19 (1), p.46.

Nikolos, I.K., Stergiadi, M., Papadopoulou, M.P. and Karatzas, G.P., 2008. Artificial neural networks as an alternative approach to groundwater numerical modelling and environmental design. Hydrological Processes:An International Journal, *22*(17), pp.3337-3348.

Dadaser-Celik, F. and Cengiz, E., 2013. A neural network model for simulation of water levels at the Sultan Marshes wetland in Turkey. Wetlands ecology and management, *21*(5), pp.297-306.

Kaya, Y.Z., Üneş, F., Demirci, M., Taşar, B. and Varçin, H., 2018. Groundwater level prediction using artificial neural network and M5 tree models. Aerul si Apa. Componente ale Mediului, pp.195-201.

Voyant, C., Tamas, W., Nivet, M.L., Notton, G., Paoli, C., Balu, A. and Muselli, M., 2015. Meteorological time series forecasting with pruned multi-layer perceptron and two-stage Levenberg-Marquardt method. International Journal of Modelling, Identification and Control, *23*(3), pp.287-294.

Kavehkar, S., Ghorbani, M.A., Khokhlov, V., Ashrafzadeh, A. and Darbandi, S., 2011. Exploiting two intelligent models to predict water level: a field study of Urmia lake, Iran. International Journal of Environmental and Ecological Engineering, *5*(3), pp.162-166.

Chitsazan, M., Rahmani, G. and Neyamadpour, A., 2015. Forecasting groundwater level by artificial neural networks as an alternative approach to groundwater modeling. Journal of the Geological Society of India, *85*(1), pp.98-106.

Husna, N.E.A., Bari, S.H., Hussain, M.M., Ur-Rahman, M.T. and Rahman, M., 2016. Ground water level prediction using artificial neural network. International Journal of Hydrology Science and Technology, *6*(4), pp.371-381.

Lepot, M., Aubin, J.B. and Clemens, F.H.,

2017. Interpolation in time series: An introductive overview of existing methods, their performance criteria and uncertainty assessment. Water, 9 (10), p.796.
Sun, Y., Wendi, D., Kim, D.E. and Liong, S.Y., 2016. Application of artificial neural networks in groundwater table forecasting-a case study in a Singapore swamp forest. Hydrology & Earth System Sciences, 20 (4).

How to cite this article:

Saranya, A., U. Arulanandu, M. Kalpana, P. Sujatha and Parimalarangan, R. 2021. Forecasting Groundwater Level Fluctuation of Veppanthattai Block using Artificial Neural Network. *Int.J.Curr.Microbiol.App.Sci.* 10(01): 3327-3334.
doi: https://doi.org/10.20546/ijcmas.2021.1001.390