Mukesh Kumar Mehla *et al*, International Journal of Advances in Agricultural Science and Technology,
Vol.7 Issue.6, June-2020, pg. 19-24 **ISSN: 2348-1358**
Impact Factor: 6.057
NAAS Rating: 3.77

USE OF MACHINE LEARNING IN GROUNDWATER LEVEL FORECASTING

Mukesh Kumar Mehla*; Mahesh Kothari; Harsh Upadhyay; Priyanka Jagtap
Department of Soil and Water Engineering, College of Technology and Engineering, MPUAT, Udaipur, Rajasthan, India-313001
* mukeshmehla310@gmail.com

Abstract
Knowledge about the current groundwater levels of an area play a very important role in proper utilization and management of groundwater supply. Use of machine learning for groundwater levels prediction can drastically change the way forecasting, monitoring and management of groundwater is carried out traditionally, reducing the need of costly surveys and reduce dependence on manual monitoring of wells. Artificial Neural Networks (ANNs) have been used progressively in recent years for various hydrological applications because of their ability to truthfully model the complicated non-linear relationships. Availability of proper database and frequency of data collection for creating accurate and usable dataset is one of many constraints faced when using machine learning algorithms for groundwater level prediction. Use of ANN, GP, SVM and ELM for groundwater levels prediction has many advantages over traditional methods; Performance of SVM can be enhanced by using SVM hybrid models such as SVM-QPSO and SVM-RBF. Adaptive neural fuzzy inference system (ANFIS) and genetic programming (GP) are the artificial intelligence tools that help to predict and simulate groundwater levels. GP simulated equations decrease the computational effort by using common simulation packages that can yield results with acceptable accuracy. Statistical indicators RMSE, r, R^2, MAE and MAPE can be used for the comparison of these models and find best suited model under a given condition.

***Keywords**:* Groundwater Level, Machine Learning, ANN, SVM and Soft computing

1. Introduction
Authors are advised to provide an introduction for their article. Introduction can be considered as the first Groundwater is a very precious resource on earth and is available to us in a very limited amount. It is mainly utilized in agriculture primarily for irrigation which accounts for 80% of its use. Alongside agriculture, it is also used for drinking water supply, domestic and industrial uses, etc. For proper utilization and management of groundwater resource it is mandatory to predict the groundwater level. Groundwater is subjected to variations resulting from differences in the recharge and draft of groundwater, streamflow variations, tidal effects, meteorological impacts and also global climatic changes. Forecasting groundwater level is a very important aspect of hydrological studies. Groundwater level varies spatially and temporally therefore it is important to consider both aspects for the forecasting of groundwater level. There are various models used for that ranges from mathematical and conceptual models to black box models like artificial neural networks (ANN) and other models based on artificial intelligence. There are number of studies on all these models that analysed either the spatial or the temporal aspects of groundwater level variation but the that studies that consider both spatial or the variation are few and far between (Nourani *et al.* 2008, Jha and Sahoo 2015). Also, due inadequate availability of both the spatial and temporal data poses a great challenge in predicting groundwater. Use of models can help in predicting in advance groundwater levels for a sparsely monitored region and also analyse spatial-temporal variability. Hybrid models which are combinations of traditional models and soft computing techniques. For spatial-temporal predictions, kriging has been used in combination with neural networks to develop neural-geostatistic models for groundwater level prediction and forecasting (Nourani *et al.* 2008). ANNs has ability to identify complex relationships from given patterns has made them useful in solving complex hydrological problems. It has been used for groundwater level prediction by different researchers using different input parameters (Daliakopoulos *et al.* 2005, Nayak *et al.* 2006, Dash *et al.*

ISSN: 2348-1358
Impact Factor: 6.057
NAAS Rating: 3.77

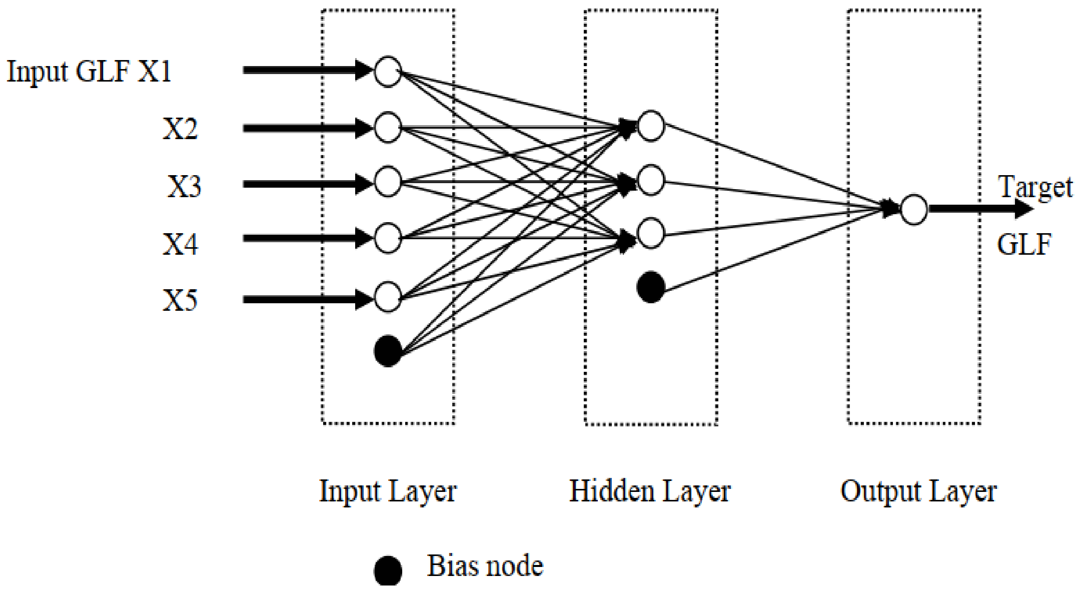
2010, Jha and Sahoo 2015). A limitation of most ANN models developed is the requirement for a large dataset for achieving proper training of ANN model. By combining ANN and FIS into the adaptive neuro-fuzzy inference system (ANFIS) and use of the same for modelling has many advantages in a computational framework. The learning capability of ANN can be utilized Chang and Chang (2006) used it to construct a water level forecasting system for flood periods. Alvisi *et al.* (2006) predicted water level using fuzzy logic and ANN. Vapnik (1995) developed the basis for the support vector machines (SVM), solutions provided by SVM are unique and global as its implementation requires the solution of a convex quadratic constrained optimization problem (Scholkoff B, Smola AJ 2002) Although SVM has been successful in prediction, the output of SVM depends on the choice of the suitable kernel function and its parameters that are being adopted. Extreme learning machine (ELM) is a novel data-driven algorithm for a single layer feed-forward neural network. (Huang *et al*,). It reduces the computational time required for training a neural network. Moreover, since ELM simplifies the entire learning process, it yields faster learning with good generalization performance. Genetic programming (GP) is an evolutionary algorithm similar to the genetic algorithm and it uses the concept of natural selection and genetics on evolutionary computation (Sreekanth and Datta 2010). This paper discuses major soft computing tools that are used for the groundwater level prediction and forecasting namely ANN, ANFIS, SVM, ELM, and GP.

1. Artificial neural network

An artificial neural network (ANN) is very much different from a conventional system such as an analytical or statistical model, it is used determine the complex nonlinear relationships between a set of inputs and desired outputs without giving any information about the actual processes involved (Haykin 1994). ANN is a network consisting of an arbitrary number of very simple elements called nodes. Each node is a simple processing element that responds to the weighted inputs it receives from other nodes (Lee *et al.*, 2004). The arrangement of the nodes is referred to as the network architecture. There are several types of ANN architecture and algorithm that are used successfully in groundwater level forecasting (Daliakopoulos, 2005). The feed-forward neural network with Levenberg-Marquardt algorithm (LM) and radial basis function (RBF) will be presented. They are more reliable and faster to convergence. They can be used successfully used for time series modelling and forecasting of groundwater level. ANN is being used for prediction in a wide range of fields as a multilayer feed-forward network with backpropagation learning algorithm.

1.1. Feed-forward neural network

A common type of feed-forward neural network (FFNN) consists of three layers (*i.e.*,an input layer is connected to a hidden layer, which is connected to an output layer) as seen in Figure 1.



ISSN: 2348-1358
Impact Factor: 6.057
NAAS Rating: 3.77

Figure 1: A three-layered feed forward Artificial Neural Network

During this operation, each node *j* receives incoming signals from every node *i* in the previous layer. Each incoming signal (y_i) associates with a weight (w_{ji}). The net input, x_j, to node *j* is a sum of the incoming signal times the weight, as:

$$x_j = \sum_i y_i w_{ji}$$

The output signal is given as $f(x_j)$, which is a non-linear function, is produced by a transfer function of its summed input. The most commonly used transfer or activation function is the logistic sigmoid and hyperbolic tangent functions. The nonlinear nature of this logistic transfer function plays an important role in the performance of the ANN. Other functions can be used as long as they are continuous and possess a derivative at all points.

$$f(x_j) = \frac{1}{1 + e^{x_j}}$$

The backward pass which is concerned with error computation and weight update and algorithm normally used for same used is a backpropagation algorithm. In this algorithm, the difference between the calculated output of the output layer and the desired output is backpropagated to the previous layer and the weights are adjusted. Other method generally used for weight and bias values update is Levenberg-Marquardt (LM) algorithm. This process continuously proceeds until the criterion achieved. The LM algorithm is widely applied to many different domains. It works extremely well in practice and is considered the most efficient algorithm. Like Quasi-Newton methods, the LM algorithm was designed to approach second-order training speed without having to compute Hessian matrix.

1.2. The Radial basis function network

The radial basis function (RBF) network also consists of three layers, namely an input layer, a hidden layer or radial basis layer, and an output layer or linear layer. The input layer collects the input information. The hidden layer consists of a set of basis functions performing nonlinear transformations of the inputs. Commonly used transformation is Gauss function as the nonlinearity of the hidden nodes. The response of the *j*-th hidden node to x_i is given by

$$h_{ij}(x) = exp\left(\alpha \left\| x_i + c_j \right\|^2\right)$$

where, $\|..\|$ is Euclidean norm, c_j is the center of the basis function and α is a positive constant that determines the width of the symmetric response of the hidden node.

The output values of the network are computed as linear combination of these basis functions (hidden nodes),

$$\hat{y}_i = \sum_{j=1}^{K} w_j h_{ji}(x)$$

Where, w_j is the network connection weights and K is the number of hidden nodes.

Assume that N samples of the signal are available for training. The center, c_j, $1 \le j \le K$, can be selected from the network training input x_i, $1 \le i \le N$. The weights can then be solved using the least squares method. They have been widely used for nonlinear system identification because of their simple topological structure and their ability to reveal explicitly the learning process (Lin & Chen, 2004). RBF networks have increasingly attracted interest for engineering applications due to their advantages over traditional multilayer perceptrons, namely faster convergence, smaller extrapolation errors, and higher reliability (Moradkhani *et al.*, 2004). The architecture and training algorithms for radial basis function networks (RBF) are simple and clear.

2. Adaptive neuro-fuzzy inference system (ANFIS)

Concept of Fuzzy logic is about mapping an input space to an output space, and the primary mechanism of this mapping is a list of if-then statements called rules (Zadeh, 1965). The fuzzy inference system (FIS) is based on the concept of fuzzy set theory and fuzzy reasoning. It is a method that interprets the values in the input vector

ISSN: 2348-1358
Impact Factor: 6.057
NAAS Rating: 3.77

and, based on some set of rules, assigns values to the output vector. There are two types of fuzzy rule system being widely used, and these two were proposed by (Namdani, 1974) and (Takagi and Sugeno, 1985). ANFIS was originally proposed by Jang (1993). ANFIS is a fuzzy system trained by an algorithm derived from neural network theory. The algorithm is a hybrid training algorithm based on backpropagation and the least squares approach. In this algorithm, the parameters defining the shape of the membership functions are identified by a backpropagation algorithm, while the consequent parameters are identified by the least squares method. An ANFIS can be viewed as a special three-layer feedforward neural network. The first layer represents input variables, the hidden layer represents fuzzy rules, and the third layer is an output.

3. Genetic programming (GP)

Genetic programming is very much similar to genetic algorithm (GA), an evolutionary algorithm based on the Darwin's theory of natural selection and survival of the fittest. But GP is different in the sense that it operates on parse trees, rather than on bit strings as in GA, to approximate the equation that best describes how the output is related to the input. The algorithm considers an initial population of random generated equations, derived from a random combination of input variables, random numbers and functions. The selection of the function has to be done appropriately to ensure the comprehension of the process. The population of the possible solutions is dependent on an evolutionary process, and then, the 'fitness' of the evolved problems is evaluated. Fitness is a measure of how well they solve the problem. The programs that best fit the data are selected from the initial population.

Genetic programming (GP) uses a tree structure to present each expression. In this structure, different variables, functions and operators are located in the nodes, which are related by several branches. Two different sets are located in the nodes: (1) a terminal set that includes numerical and non-numerical variables and (2) a function set that involves arithmetic operators (±, ×, ÷), mathematical functions (e.g. sin, cos), Boolean operators (e.g. and, or), logical expressions (e.g. if-then-else) and other user-defined expressions (Sreekanth and Datta, 2010). A GP mechanism starts with a randomly generated set of trees. The maximum number of generations and population size used in the present GP model are 8 and 10, respectively. The error criterion of each tree is calculated. The trees with least error values are selected using methods such as the roulette wheel or tournament method. The selected trees are prepared for the next iteration by two genetic operators: crossover and mutation. These two genetic operators are applied repeatedly till the termination.

4. Support vector machine (SVM)

The SVM equations are formulated as per Vapnik's theory (Vapnik, 1995) that if $\{(I_1, T_1), \dots (I_N, T_N)\}$ are assumed as the given training data sets, where I_k [R^n refers to the space of input variable, T_k [R refers to the space of the target value, and N represents the length of the training data. The linear regression of SVM is estimated by solving the equation,

4.1 SVM-RBF

The Gaussian RBF kernel is the most successful kernel and has been widely used in many problems. It uses the Euclidean distance between two points in the original space to find the correlation in the augmented space

$$K(x,y) = exp(-\gamma\|x-y\|^2)$$

4.1 SVM-QPSO

In quantum physics, the state of a particle with momentum and energy can be depicted by its wave function w (x,t). As per QPSO theory, each particle is in a quantum state and is formulated by its wave function w (x,t) instead of the position and velocity described by PSO. According to the statistical significance of the wave function, the probability of a particle's appearing in a certain position can be obtained from the probability density function $|\psi(x.t)^2|$. The probability distribution function of the particle's position can be calculated through the probability density function. The PSO algorithm with position update equation is called as the quantum delta-potential-well-based PSO (QDPSO) algorithm. Keeping in view the vital position of L for convergence rate and performance of the algorithm, an improvement was proposed to evaluate parameters L. As per this algorithm, the mean best position (mbest) is defined as the centre of best positions of the swarm. The PSO algorithm with is called as the quantum-behaved particle swarm optimization (QPSO). The most commonly used control strategy of b is to initially setting it to 1 and reducing it linearly to 0.5.

5. Extreme learning machine (ELM)

ELM is a three-layered structure algorithm (Huang *et al.*, 2011). In the ELM structure, the input weight (connection between the input and hidden) and the bias values (in the hidden layer) are randomly generated. ELM analytically calculates the output weight matrix between hidden layers and output layers through a simple generalized inverse operation of the hidden layer output matrix. ELM can be formulated as a function with L hidden nodes and N training sample. In the ELM algorithm, the input weight and bias are randomly chosen at the initial stage. If the ELM model with L hidden nodes is able to learn these N training samples with no residuals, then w can be predicted.

2. Conclusion

Groundwater level prediction is obligatory to understand and resolve various water management issues. Use of ANN, ANFIS, GP, SVM and ELM for groundwater levels prediction has many advantages over traditional methods, Performance of SVM can be enhanced by using SVM hybrid models such as SVM-QPSO and SVM-RBF. Adaptive neural fuzzy inference system (ANFIS) and genetic programming (GP) are the artificial intelligence tools that help to predict and simulate groundwater levels. GP simulated equations decrease the computational effort by using common simulation packages that can yield results with acceptable accuracy Availability of proper database and frequency of data collection for creating accurate and usable dataset is one of many constraints faced when using machine learning algorithms for groundwater level prediction.

References

[1]. Alvisi, S., Mascellani, G., Franchini, M. And Bardossy, A. (2006) Water level forecasting through fuzzy logic and artificial neural network approaches. Hydrology and Earth System Sciences 10: 1-17.
[2]. Chang, F.-J. And Chang, Y.-T. (2006) Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. Advances in Water Resources 29: 1-10.
[3]. Daliakopoulos, I.N., Coulibaly, P., and Tsanis, I.K., (2005). Groundwater level forecasting using artificial neural networks. *Journal of Hydrology*, 309, 229–240. doi:10.1016/j.jhydrol.2004.12.001.
[4]. Huang GB, Zhu QY, Siew CK (2004) Extreme learning machine: a new learning scheme of feedforward neural networks. In: Proceedings of the international joint conference on neural net- works, 25–29 July 2004, Budapest, Hungary, vol 2, pp 985–990.
[5]. Huang GB, Wang DH, Lan Y (2011) Extreme learning machines: a survey. Int J Mach Learn Cybern 2(2):107–122.
[6]. Haykin, S., (1994). *Neural networks*. New York: Macmillan College Publishing Company.
[7]. Jang, J.S.R. (1993) ANFIS: Adaptive network based fuzzy inference system, IEEE Trans. On System, Man, and Cybernetics 23(3): 665-685
[8]. Jha, M.K. and Sahoo, S., (2015). Efficacy of neural network and genetic algorithm techniques in simulating spatio-temporal fluctuations of groundwater. *Hydrological Processes*, 29, 671–691. doi:10.1002/hyp.v29.5.
[9]. Lee, S., Ryu, J.H., Won, J.S. & Park, H.J. (2004) Determination and application of the weights for landslide susceptibility mapping using an artificial neural network. Engineering Geology 71: 289-302.
[10]. Moradkhani, H., Hsu, K.-L., Gupta, H.V. and Sorroshian, S. (2004) Improved streamflow forecasting using self-organizing radial basis function artificial neural networks. Journal of hydrology 295: 246-262.
[11]. Namdani, E.H. (1974) Application of fuzzy algorithm for control of simple dynamic plant, Proc. IEEE 121:1585-1889.
[12]. Nash, J.E. and Sutcliffe, J.V., (1970). River flow forecasting through conceptual models part I - A discussion of principles. *Journal of Hydrology*, 10 (3), 282–290. doi:10.1016/0022-1694(70)90255-6.
[13]. Nayak, P.C., Rao, Y.R.S., and Sudheer, K.P., (2006). Groundwater level forecasting in a shallow aquifer using artificial neural network approach. *Water Resources Management*, 20, 77–90. doi:10.1007/s11269-006-4007-z.
[14]. Nourani, V., Mogaddam, A.A., and Nadiri, A.O., (2008). An ANN-based model for spatiotemporal groundwater level forecasting. *Hydrological Processes*, 22, 5054–5066. doi:10.1002/hyp.v22:26.
[15]. Scholkoff B, Smola AJ (2002) Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT Press, Cambridge, p 648.
[16]. Sreekanth, J. and Datta, B., (2010). Multi-objective management of saltwater intrusion in coastal aquifers using genetic programming and modular neural network based surrogate models. *Journal of Hydrology*, 343, 245–256. doi:10.1016/j.jhydrol.2010.08.023.

Vol.7 Issue.6, June-2020, pg. 19-24 ISSN: 2348-1358
Impact Factor: 6.057
NAAS Rating: 3.77

[17]. Takagi, T. and Sugeno M. (1985) Fuzzy identification of systems its application to modeling and control, IEEE Trans. On System, Man, and Cybernetics 15(1): 116-132.
[18]. Vapnik VN (1998) Statistical learning theory. Wiley, New York.
[19]. Zadeh, L.A.,(1965). Fuzzy sets. Information and control 8(3): 338-353.

A Brief Author Biography

Mr. Mukesh Kumar Mehla is PhD Scholar in the Department of Soil and Water Engineering, College of Technology and Engineering, Maharana Pratap University of Agriculture and Technology, Udaipur. He has qualified GATE 2017, AIEEA-PG 2017 and AICE-JRF/SRF(PGS) 2019. He has published research papers in various journals and 8 abstracts in conference proceedings.

Dr. Mahesh Kothari is Professor and Head, Department of Soil and Water Engineering, College of Technology and Engineering, Maharana Pratap University of Agriculture and Technology, Udaipur. He has numerous publications in the form of research papers, articles and books.

Mr. Harsh Upadhyay is PhD Scholar of the Department of Soil and Water Engineering, College of Technology and Engineering, Maharana Pratap University of Agriculture and Technology, Udaipur. He was awarded with Jain Irrigation Medal for merit in M. Tech. in year 2019. He has published, abstracts and articles in various journals and proceedings.

Ms, Priyanka Jagtap is PhD Scholar of the Department of Soil and Water Engineering, College of Technology and Engineering, Maharana Pratap University of Agriculture and Technology, Udaipur. He was awarded with Departmental Gold Medal for merit in M. Tech. in year 2019. He has published, abstracts and articles in various journals and proceedings.