Prediction of Groundwater Quality Using Artificial Neural Network

Rita Maria Joseph*, Alna D Manjaly, Sreeram Unni, Able E C, Vinitha Sharon

Department of Civil Engineering, Christ College of Engineering, Thrissur, Kerala, India

*Corresponding author: ritamariajo21@gamil.com

doi: https://doi.org/10.21467/proceedings.112.26

ABSTRACT

Assessment and prediction of water quality is a vital tool for the management of water resources systems. It is necessitous for human existence, agriculture and industry. This project delves into the prediction of groundwater quality parameters and groundwater quality criterion based on the Artificial Neural Network Modelling with the study area as Kerala, a state of India. Two models were developed. The first model employs the water quality parameters such as pH, electrical conductivity, total hardness as the input parameters and calcium, magnesium, chloride, fluoride, nitrate concentration as the output parameters. The second model was designed by giving input as, input values and the predicted output values of the first model, and groundwater quality criterion corresponding to each location as the target values. The output qualitative parameters were estimated and compared with the measured values, to evaluate the influence of key input parameters. The number of neurons to be given in the hidden layer was decided by the trial-and-error method. Data of 506 water samples from all over Kerala were collected for modelling. The results show that the performance of the first model is having a regression value R^2=0.97 and the second model is having a regression value R^2=0.99, using backpropagation algorithms according to the best-chosen input parameters. MATLAB R2018a was used for developing the multi-layer perceptron ANN models and the performance function for calculating the model performance error statistics was the coefficient of determination (R^2). This project introduces a cost-effective and quick method for the estimation of groundwater quality.

Keywords: Artificial neural network (ANN), groundwater quality, Kerala.

1 Introduction

1.1 Background

Safe drinking water is one of the focal necessities of every human on earth for supporting life, because of which it should be accessible to all individuals in an adequate sum through protected and helpful methods of water supply (Kumar et al., 2012; Khudair et al., 2018). The expanded interest in clean drinking water causes to notice the superintendence of groundwater quality. The issues concerning the accessibility of clean water have become intense; particularly in developing countries (Akhil et al., 2013; Aswini et al., 2016; Lallahem et al., 2017). So the groundwater quality evaluation is important. Around, 33% of the total populace has been utilizing groundwater for drinking. In many countries situated in semiarid and dry areas, the management and security of groundwater assets are critical. Along these lines, protection notwithstanding the present moment and long haul planning of resources need to improve their viability (Kumar et al., 2014; Kheradpisheh et al., 2015). The nature of water has been contorted as of late by agriculture and modern exercises, and specialists have focused on quality boundaries like EC, Cl, F and NO_3 (Kheradpisheh et al., 2015; Thambi et al., 2015; Prasannakumari et al., 2019). The usage, management and protection of groundwater ought to be a fundamental guideline of country planning (Shigidi et al., 2003; Mirabbasi, 2015; Selvan et al., 2018).

Proceedings DOI: 10.21467/proceedings.112; Series: AIJR Proceedings; ISSN: 2582-3922; ISBN: 978-81-947843-3-3

ANN model in such a manner can be used as an effective apparatus by managers. A neural network model is an applied model and undoubtedly, is an improved framework of the mathematical model. The most noticeable issue observed by users and suppliers of mathematical models is the prerequisite for these models to address distinctive data (Kukreja et al., 2016). ANN which is inspired by the biological neural network can help in dealing with the before-referred issues. The networks are related to intelligent systems. A definitive point of Artificial Intelligence is to create an artificial human by planning a product program that can think like individuals. Such machines can settle on the best choices under significant circumstances by consolidating progressed scientific methods and the capacity of expert designers, government officials, and different researchers dependent on the enormous volume of data assets accessible (Kheradpisheh et al., 2015). The leading applications used in neural network modelling are Neuro Solution and MATLAB software. ANNs are associated with nervous systems and depend on our current comprehension of the brain. To form black-box representations of systems they use highly simplified models composed of many processing elements united by links of variable weights. Human nervous system performance is the foundation of information transferring in artificial neural networks. The neurons in the network are data processors and are responsible for reviewing input signals. ANN just computes output values from input values. The sum of transferred information is weighted. These weighted connections are adjusted during the learning process.

1.2 Objective

- To predict the water quality parameters.
- To predict the water quality criterion.
- To build ANN models using MATLAB.

2 Methodology

2.1 Selection of the study area

Kerala (38,863 km^2), a state in India was taken as the study area based on the availability of data required and for the easiness of the experiment conduct. Kerala is in the southern part of India, situated between north latitudes 8^0 18' and 12^0 48' and east longitudes 74^0 52' and 77^0 22' (Jayasankar et al., 2015). Out of 14 districts in Kerala, 11 districts (Alapuzha, Ernakulam, Idukki, Kannur, Kasargod, Kollam, Kozhikode, Palakkad, Pathanamthitta, Thrissur and Trivandrum) were selected with the data as shown in Fig. 1. It was necessary to gather data for training purposes, therefore 506 data of samples from different locations were collected.

Proceedings DOI: 10.21467/proceedings.112
ISBN: 978-81-947843-3-3
Series: AIJR Proceedings
ISSN: 2582-3922

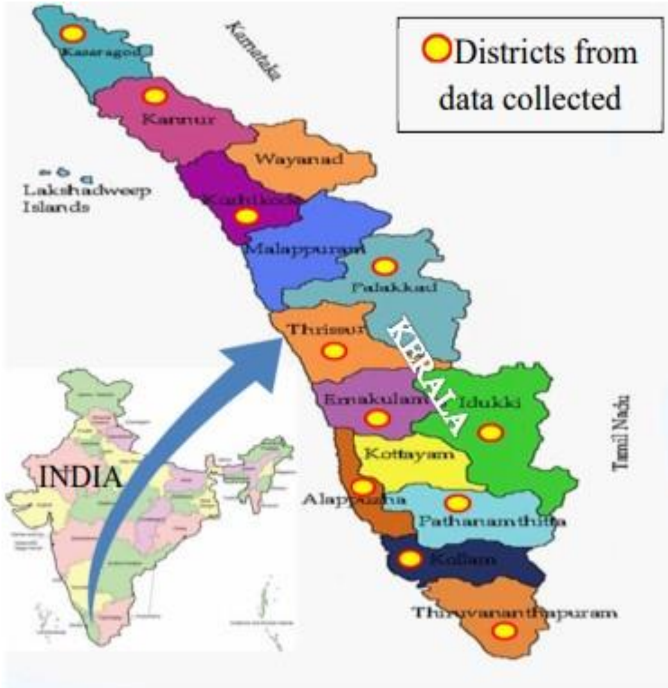

Figure. 1. Study area: Kerala, India.

2.2 Designing ANN models using MATLAB

Designing ANN models follows several systemic procedures (Al Shamisi et al., 2011). Basic steps include;

1) Data collection: Collection and preparation of sample data for designing ANN models are done in this step. Data comprises of water quality parameters such as pH, electrical conductivity, total hardness, calcium content, magnesium content, chlorine content, fluoride content and nitrate content from different locations of Kerala. For the first model, that is, the model for the prediction of water quality parameters, the input parameters are pH, electrical conductivity and total hardness, and target parameters are calcium, magnesium, chloride, fluoride and nitrate concentration. The second model is designed by giving input as, input values and the predicted output values of the first model, and groundwater quality criterion corresponding to each location as the target values; by evaluating every data and categorizing its quality under three category as 1,0 and -1 which corresponds to good, moderate and poor respectively based on the acceptable and permissible limits of the parameters.
2) Data pre-processing: After data collection, to train the ANNs more efficiently, three data preprocessing procedures are conducted. These techniques are to take care of the issue of missing data, normalize data and randomize data. On the off chance that any data is missing, they are supplanted by the average of adjoining values. Method of normalization prior to conceding the information to the network is typically a decent practice. Incorporating variables with enormous extents and little sizes will confound the learning algorithm on the significance of every factor and may compel it to refuse the variable with the smaller size.
3) Building the network: Artificial neural network models are developed using the MATLAB R2018a software. In the network building stage, the designer defines the number of hidden layers, neurons in each layer, transfer function in each layer, training function, weight/bias learning function, and performance function. Here in this project, for both the models a multilayer perceptron (MLP) ANN architecture and feed-forward backpropagation network is used. In these models number of hidden layers

are 2 with 10 neurons in each layer. TRAINLM is the training function and the performance function is MSE. The weight/bias learning function LEARNGDM is used in these network models.

4) Training the network: In this training process, to make the actual outputs (predicated) close to the target (measured) outputs of the network the weights are adjusted. MATLAB gives built-in transfer functions that are used for this job.
5) Testing the network: To test the performance of the developed model is the next step. In this project, for testing the first ANN model by stimulating in MATLAB, the unseen data from Thrissur district of Kerala is used. For the second ANN model, the predicted data is compared with the collected data. To judge the performance of the developed ANN models quantitatively, and check whether there is an underlying trend in the performance of ANN models, statistical analysis of the coefficient of determination (R^2) is conducted.

3 Results and Discussions

The models were trained with these ranges of values of parameters from different locations of Kerala as shown in Table 1 (Balakrishnan, 2013; Chand, 2013; Chandran, 2013; Joji, 2013; Murugan, 2013; Rani, 2013; Ravi, 2013; Saritha, 2013; Shyam, 2013; Singadurai, 2013; Subramani, 2013).

In this project work, for creating the two models, Multi-Layer Perceptron (MLP) ANN architecture and feed-forward backpropagation network were used. The first model employs the input parameters like pH, EC, and TH and output parameters as Ca, Mg, Cl, F and NO_3 concentration. The input parameters and predicted output parameters of the first model were given as input for the second model and the water quality criterion corresponding to each location was given as target. Artificial neural network modelling was developed using the MATLAB R2018a software package. ANN modelling was done using the following functions as shown in Table 2 for both models. These were chosen based on the trial and error method.

Table 1. Ranges of values of parameters.

Parameter	pH	EC µs/cm	TH mg/l	Ca mg/l	Mg mg/l	Cl mg/l	F mg/l	NO_3 mg/l
Range	4.37-10.86	10-5260	1.6-675	1-108	0-336	2.8-1178	0-7.3	0-149
Mean ± SD	7.77 ± 0.69	271.79 ± 363.57	56.34 ± 63.49	13.85 ± 13.43	6.04 ±17.69	41.60 ±98.39	0.26 ± 0.41	11.64 ± 0.41

Table 2. Functions used for ANN modelling.

Functional elements of modelling	Optimum parameters
Number of hidden layers	2
Number of neurons in each layer	10
Training function	TRAINLM
Performance function	Mean Square Error (MSE)
Weight/bias learning function	LEARNGDM
Transfer function	TANSIG

Proceedings DOI: 10.21467/proceedings.112 Series: AIJR Proceedings

ISBN: 978-81-947843-3-3 ISSN: 2582-3922

3.1 Results of the First Model

Fig. 2 illustrates the architecture of model 1. The best results are obtained with a combination of input parameters as pH, EC and TH and output parameters as Ca, Mg, Cl, F and NO_3. The number of hidden layers is 2 with 10 neurons in each layer.

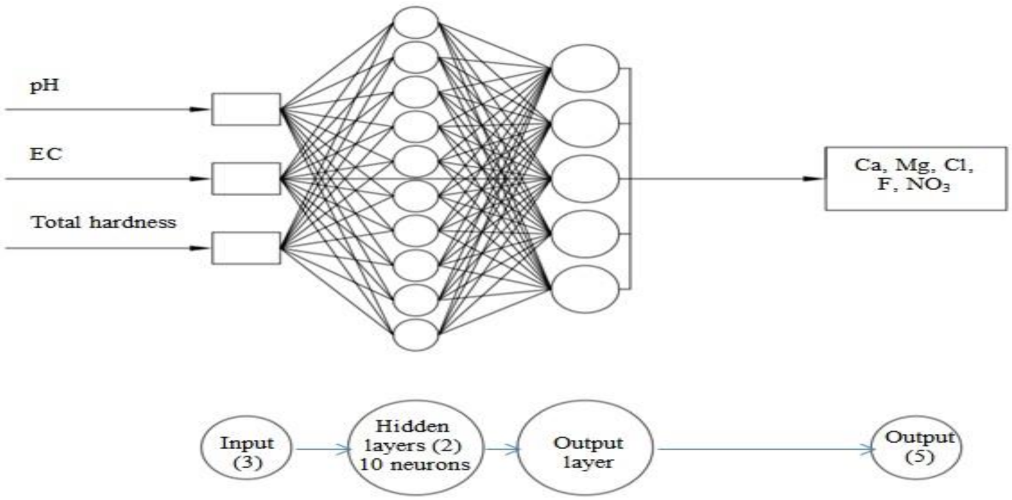

Figure. 2. Multi-layered perceptron (MLP) network of Model 1.

The performance of the ANN model 1; for the prediction of water quality parameters is shown in Table 3.

Table 3. Performance of the ANN model.

Model		Coefficient of determination (R^2)			
		Training	Testing	Validation	All
1	Prediction of water quality parameters	0.97	0.97	0.96	0.97

The graph below in Fig. 3 shows the comparison between measured values and predicted values of chloride which have the highest performance among other parameters (Cl: (2.8, 1178) mg/l).

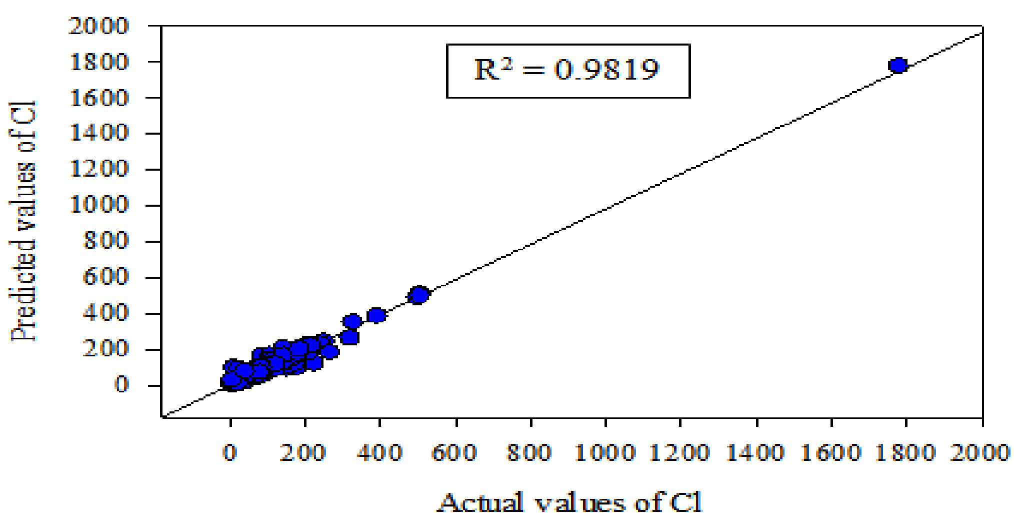

Figure. 3. Comparison between measured values and predicted values of Cl in Model 1.

3.2 Results of the Second Model

The figure below illustrates the architecture of model 2 with eight inputs and one output. For model 2 the input parameters and predicted output parameters of the first model were given as input and the water quality criterion corresponding to each location is given as target. This model also employs 2 hidden layers with 10 neurons in each layer as shown in Fig. 4.

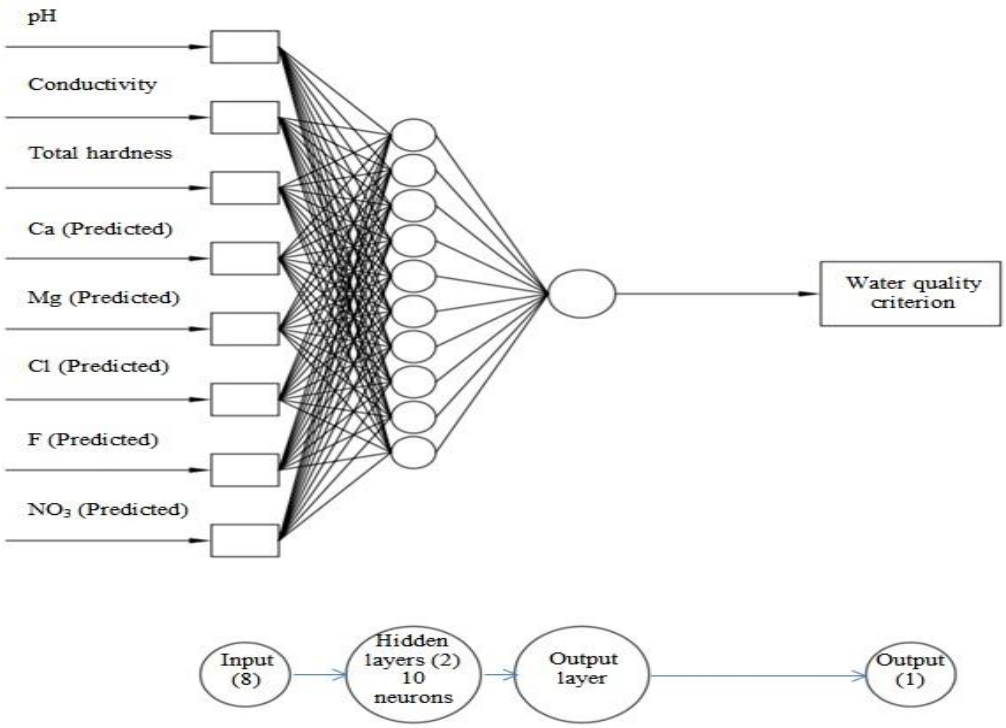

Figure. 4. Multi-layered perceptron (MLP) network of Model 2.

The performance of the ANN model 2; for the prediction of water quality criterion is shown in Table 4.

Table 4 Performance of the ANN model.

Model		Coefficient of determination (R^2)			
		Training	Testing	Validation	All
2	Prediction of water quality criterion	0.99	0.99	0.96	0.99

Model 2 is based on the outputs obtained from model 1. It helps in categorizing the quality of the water sample as good, moderate and poor. The regression value of model 2 for training, testing, validation and all-purpose are 0.99, 0.99, 0.96 and 0.99. The results of model 2 indicate higher performance in regression values obtained from Kannur and Kasargod and lower performance in regression values obtained from Palakkad and Ernakulam. The lower values show the evidence for saline water intrusion in coastal areas and anthropogenic activities mainly due to the overutilization of fertilizers. Out of 506 data of samples of different locations, 424 were of good quality water as $R^2 \geq 0.93$, 49 were of moderate quality as $R^2 = 0.92$ to 0.82 and 33 were of poor quality as $R^2 < 0.82$. From these results, the average quality of water in each district was determined as shown in Fig. 5.

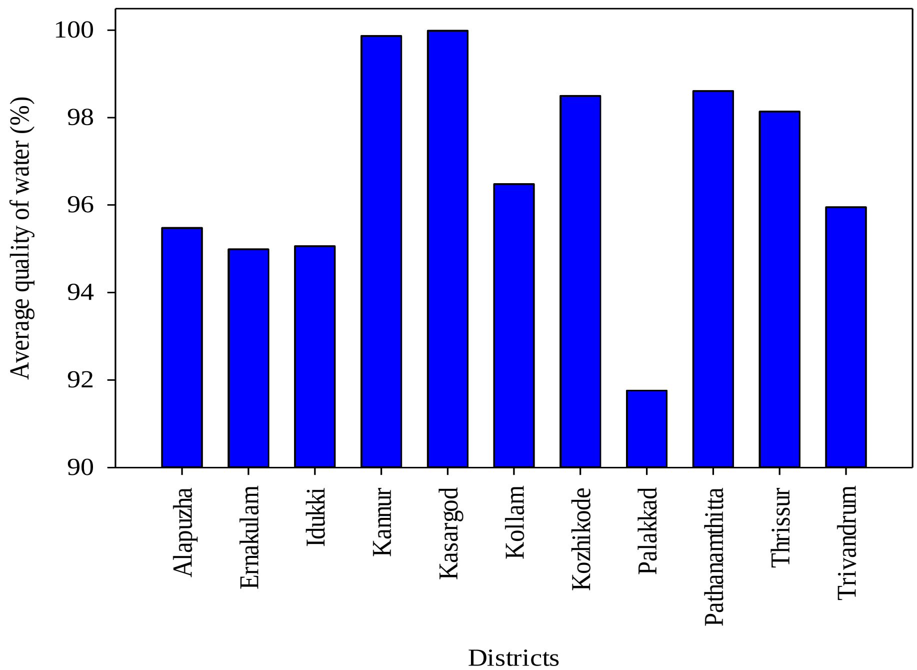

Figure. 5. Average quality of water in each district in Kerala.

4 Conclusion

Current progress in the management of water resources has expanded the demands for modelling techniques that can provide safe, effective and exact nonlinear dynamics of groundwater quality representation. Restricted water quality information and the significant expense of water quality checking frequently cause major issues for process-based modelling strategies. In this project work, MATLAB software was used to predict the groundwater quality parameters and groundwater quality criterion using ANN. Data were collected from different locations of Kerala for training the neural network. By giving the parameters like pH, EC, TH as input, the predicted values show a very close resemblance to the reference value and high R^2 value were also obtained. Thus, the model is exhibiting high prediction efficiency. Using the Artificial Neural Network modelling, we were able to classify whether the quality of water was good, moderate, or poor and also to predict the output parameters with very high accuracy. Results show that the water in Palakkad and Ernakulam district has slightly low quality. This may be due to the intrusion of seawater in the coastal areas and also due to anthropogenic activities such as the over-utilization of fertilizers (Boominathan et al., 2012). Also, ANN models have in normally great execution regardless of whether one or more input parameters are missing. The back-propagation algorithm was used in training the neural network. Supervised learning methods and feed-forward architecture are used by a back-propagation neural network (BPNN). As indicated by the profoundly costly and tedious tests, this is a fast and financially savvy technique for the management rehearses, particularly in crises. It is highly beneficial in regions where the groundwater is on the verge of deterioration. Indeed, ANN modelling requires unique thoughtfulness regarding the training process. Utilizing an ineffectively trained network is unequipped for accurately anticipating the response. ANN modelling demands a large number of input data for the training purpose. ANN modelling experiences the issue when the limited information record will result in over-fitting of the model to the training information and bringing about the weak speculation of the model for the information that is out of the training information range (Sakizadeh, 2015). But increasing the number of data beyond a limit for training will have a significant impact on the computational time and cost.

How to Cite this Article:

Joseph, R. M., Manjaly, A. D., Unni, S., Able, E. C., & Sharon, V. (2021). Prediction of Groundwater Quality Using Artificial Neural Network. *AIJR Proceedings*, 213-221.

References

Akhil, P. S., Sumangala, K. N. and Sujatha, C. H. (2013). *Groundwater hydrochemistry in the crystalline laterites/ forest loams of Kasargod district, Kerala, India*. Water Science and Technology, 13(2), 486-498. 10.2166/ws.2013.049.

Al Shamisi, M. H., Assi, A. H. and Hajase, H. A. N.: Using MATLAB to Develop Artificial Neural Network Models for Predicting Global Solar Radiation in Al Ain City – UAE. IntechOpen, University Campus STeP Ri Slavka Krautzeka 83/A 51000 Rijeka, Croatia (2011).

Aswini, P. V., Mary, C., Anusmitha, A. and Mohanan, A. (2016). *Groundwater Quality Analysis*. International Journal of Engineering Research & Technology (IJERT), (ISSN), 4(13), 2278-0181.

Balakrishnan, K.: *Ground Water Information Booklet Of Kasargod District, Kerala State*, Government Of India Ministry Of Water Resources Central Ground Water Board Kerala Region, Technical Reports: Series 'D', Kedaram Kesavadasapuram Thiruvananthapuram Kerala (2013).

Boominathan, M., Karthick, B., Ali, S. and Ramachandra, T. V. (2012). *Spatial assessment of groundwater quality in Kerala India*. IUP Journal of Soil & Water Sciences, 5(1), 52-71.

Chand, A. D. A.: *Ground Water Information Booklet Of Palakkad District, Kerala State*, Government Of India Ministry Of Water Resources Central Ground Water Board Kerala Region, Technical Reports: Series 'D', Kedaram Kesavadasapuram Thiruvananthapuram Kerala (2013).

Chandran, M.: *Ground Water Information Booklet Of Kollam District, Kerala State*, Government Of India Ministry Of Water Resources Central Ground Water Board Kerala Region, Technical Reports: Series 'D', Kedaram Kesavadasapuram Thiruvananthapuram Kerala (2013).

Jayasankar, P. and Babu, M. N. S. (2017). *An Assessment of Ground Water Potential for state of Kerala, India: A Case Study*. International Conference of Water-From Pollution to Purification, 5(2), 2348-6732.

Joji, V. S.: *Ground Water Information Booklet Of Thrissur District, Kerala State*, Government Of India Ministry Of Water Resources Central Ground Water Board Kerala Region, Technical Reports: Series 'D', Kedaram Kesavadasapuram Thiruvananthapuram Kerala (2013).

Kheradpisheh, Z., Talebi, A., Rafati, L., Ghaneian, M. T. and Ehrampoush, M. H. (2015). *Groundwater quality assessment using artificial neural network: A*

case study of Bahabad plain, Yazd, Iran. Desert, 20(1), 65-7.

Khudair, B. H., Jasim, M. M. and Alsaqqar, S. A. (2018). *Artificial Neural Network Model for the Prediction of Groundwater Quality*. Civil Engineering Journal, 4(12), 2959-2970. 10.28991/cej-03091212.

Kukreja, H., Bharath, N., Siddesh, C. S. and Kuldeep, S. (2016). *An introduction to artificial neural network*. International Journal Of Advance Research And Innovative Ideas In Education, 1(5), 2395-4396.

Kumar, P. J. S., Babu, P.T. and Delson, P. D. (2012*). A preliminary assessment of groundwater quality in Thrissur District, Kerala, India*. Elixir Pollution 51, 10763-10765.

Kumar, V. S., Amarender, B., Dhakate, R., Sankaran, S. and Kumar, K. R. (2014). *Assessment of groundwater quality for drinking and irrigation use in shallow hard rock aquifer of Pudunagaram, Palakkad District Kerala*. Appl Water Sci, 6(2), 10.1007/s13201-014-0214-6.

Lallahem, S. and Hani, A. (2017). *Artificial neural networks for defining the water quality determinants of groundwater abstraction in coastal aquifer*. AIP Conference Proceedings, 1814(1). 10.1063/1.4976232.

Mirabbasi, R.: *Application of Artificial Intelligence Methods in Geosciences and Hydrology*. OMICS Group eBooks, 731 Gull Ave, Foster City, CA 94404, USA (2015).

Murugan, S. S.: *Ground Water Information Booklet Of Pathanamthitta District, Kerala State*, Government Of India Ministry Of Water Resources Central Ground Water Board Kerala Region, Technical Reports: Series 'D', Kedaram Kesavadasapuram Thiruvananthapuram Kerala (2013).

Prasannakumari, P., Pradeep, B., Jessy, M. D, Ulaganathan, A. and Jacob, J. (2019). *Water quality in major land use systems of Western Ghats in Kottayam and Idukki districts of Kerala*. Rubber Science, 32(2), 137-149.

Rani, V. R.: *Ground Water Information Booklet Of Trivandrum District, Kerala State*, Government Of India Ministry Of Water Resources Central Ground Water Board Kerala Region, Technical Reports: Series 'D', Kedaram Kesavadasapuram Thiruvananthapuram Kerala (2013).

Ravi, A.: *Ground Water Information Booklet Of Kozhikode District, Kerala State*, Government Of India Ministry Of Water Resources Central Ground Water Board Kerala Region, Technical Reports: Series 'D', Kedaram Kesavadasapuram Thiruvananthapuram Kerala (2013).

Sakizadeh, M. (2015). *Artificial intelligence for the prediction of water quality index in groundwater systems*. Modeling Earth Systems and

Environment, 2(8), 10.1007/s40808-015-0063-9.

Saritha, S.: *Ground Water Information Booklet Of Alappuzha District, Kerala State*, Government Of India Ministry Of Water Resources Central Ground Water Board Kerala Region, Technical Reports: Series 'D', Kedaram Kesavadasapuram Thiruvananthapuram Kerala (2013).

Selvan, J., Raj, D. P. and Murugan. (2018). *Natural and anthropogenic contaminants in dug wells along Pampa river bank*. International Journal of Scientific & Engineering Research, 9(9).

Shigidi, A. and Garcia, L. A. (2003). *Parameter Estimation in Groundwater Hydrology Using Artificial Neural Networks*. Journal of computing in civil engineering, 17(4). 10.1061/(ASCE)0887-3801(2003)17:4(281).

Shyam, T. S. A.: *Ground Water Information Booklet Of Ernakulam District, Kerala State*, Government Of India Ministry Of Water Resources Central Ground Water Board Kerala Region, Technical Reports: Series 'D', Kedaram Kesavadasapuram Thiruvananthapuram Kerala (2013).

Singadurai, S.: *Ground Water Information Booklet Of Idukki District, Kerala State*, Government Of India Ministry Of Water Resources Central Ground Water Board Kerala Region, Technical Reports: Series 'D', Kedaram Kesavadasapuram Thiruvananthapuram Kerala (2013).

Subramani, M. S.: *Ground Water Information Booklet Of Kannur District, Kerala State*, Government Of India Ministry Of Water Resources Central Ground Water Board Kerala Region, Technical Reports: Series 'D', Kedaram Kesavadasapuram Thiruvananthapuram Kerala (2013).

Thambi, M., Cherian, T., Jose, A. T. and Jamal, S. (2015). *Physico-chemical Characteristics Of Groundwater Samples From Different Areas Of Thrissur District, Kerala State, India*. International Journal of Environment, 4(2), 10.3126/ije.v4i2.12636.