J. Indian Soc. Coastal Agric. Res. 36(2): 122-127 (2018)

# Prediction of Water Quality Parameters for Irrigation in Konkan Region using Artificial Neural Network Technique

H. N. BHANGE[1]*, D. M. MAHALE[1], P. M. INGLE[1] and P. K. SINGH[2]

[1]Department of Soil and Water Engineering, College of Agricultural Engineering and Technology, Dr. B. S. Konkan Krishi Vidyapeeth, Dapoli - 415 712, Maharashtra

[2]Department of Soil and Water Engineering, College of Technology and Engineering, Maharana Pratap University of Agriculture & Technology, Udaipur - 313 001, Rajasthan

Received: 11.07.2018 Accepted: 31.10.2018

**In the present study, artificial neural networks (ANN) were used to derive and develop models for prediction Kelly's ratio (KR ), percent sodium (Percent Na), permeability index (PI ), residual sodium carbonate (RSC), sodium absorption ratio (SAR) and soluble sodium percentage (SSP) as groundwater quality parameters of Ratnagiri district by using post monsoon season groundwater quality parameters (Na, Mg, K, $CaCO_3$ and $HCO_3$) collected for time period 1999-2014 from Groundwater Surveys and Development Agency, Navi Mumbai as input variables. The ANN model was developed with multilayer feed forward back propagation (MLFBP) with sigmoid transfer function. While developing ANN model for different input parameters, three steps were followed as identification of model structures, to evaluate the performance and adopting model for forecasting. The model development data set was further divided into three subsets; training set, cross validation set and testing set in 70:15:15 proportions. The ANN models were developed for prediction KR, Percent Na, PI, RSC, SAR and SSP using neurosolutions. Performance of model was evaluated by statistical criteria which included correlation coefficient, root mean square error, index of agreement and mean bias error. The analysis revealed that the selected ANN based models had correlation coefficientr>0.95, RMSE<0.3438, IA>0.91and MBE<-0.2450 for Model 3-2-1 to predict KR, Model 4-2-1 to predict percent Na, Model 4-6-1 to predict PI, Model 4-4-1 to predict RSC and Model 3-6-1 to predict SAR and SSP during post monsoon season. The results confirmed that developed ANN models were found suitable for prediction of water quality indicators used for irrigation purpose. It was recommended that two numbers of nodes in hidden layer can be used for modelling of KR and percent Na, four number of nodes for RSC and six numbers of nodes can be used for SAR, SSP and PI under limited data availability.**

(***Key words:*** *Artificial neural network, Water quality parameters, Irrigation Water Quality*)

The Ratnagiri district of Maharashtra is outlined by coastline and creeks and agricultural land is reported to be saline due to sea water ingress along the coast and creeks. Ground water quality has become saline therefore rendering it unsuitable for irrigation. Population has almost doubled within a span of two decades in this district directly resulting in a rise in irrigation and micro-scale industries. Water required for all these activities is sourced from the groundwater reserves, thereby deteriorating the quality due to manmade tampering. Besides anthropogenic activities, natural phenomena such as weathering of rocks and dissolution of minerals and climate changes also lead to release of certain elements into water. Therefore, assessment of the water quality and its prediction will help to suggest the type of amendment to be adopted or suitable measure to be adopted in that area. Routine applications of fertilizers on crop fields cause contamination of groundwater as well as accumulation of the nutrients in groundwater. Several researchers evaluated the suitability of groundwater for irrigation quality *viz*., Jeevanandam *et al*. (2006), Laluraj and Gopinath (2006), Nagarajan *et al*. (2009), Prasanna, *et al*. (2010) and Tyagi *et al*. (2009). Total dissolved solids (TDS) values are also considered as an important parameter in determining the usage of water, and groundwater with high TDS values are not suitable for both irrigation and drinking purposes as reported by Freeze and Cherry (1979).

Artificial neural network is a massively parallel distributed information processing system that has

**Corresponding author:* E-mail: harshalbhange@gmail.com

certain performance characteristics resembling biological neural networks of the human brain (Junsawang *et al*., 2007). The basic building block of an ANN is neuron. They receive an input and produce an output, which is passed to other neurons in other layers. Neurons having similar characteristics in an ANN are arranged in groups called layers. The neurons in one layer are connected to those in the adjacent layers, but not to those in same layer. The strength of connection between two neurons in adjacent layers is represented by connection strength or weight (s). Khandelwal and Singh (2005),Yesilnacar *et al*. (2008) and Ehteshami *et al*. (2016) observed that the predicted results of chemical parameters of groundwater by ANN was very satisfactory and acceptable and seen to be a good alternative for prediction of groundwater quality parameters. The present study focuses on ascertaining the irrigational suitability of groundwater in the study area and intends to highlight such issues, if any, for future implementation of preventive measures and to model and predict the groundwater quality parameters using artificial neural network technique.

## MATERIALS AND METHODS

An ANN normally consists of three layers, an input layer, a hidden layer and an output layer. Input layer usually receives input signal values. Neurons in output layer produce output signal. ANN has been found useful for modeling and prediction of uncertain and complex phenomena. A neural network trained from previous data to forecast future events, without accurately understanding physical parameters which influence present and future events. In training, network was adjusted based on comparison of output and target, until network output matches the target. The number of nodes in input layer depends on number of parameters used in estimating such as residual sodium carbonate (RSC), sodium adsorption ratio (SAR), soluble sodium percentage (SSP), permeability index (PI), percent sodium (% Na) and Kelly's ratio (KR). The input parameters and number of nodes in hidden layer consider for different class models is given in Table 1.

A total of 2236 samples data sets collected were divided into a model development set and evaluation data. According to Rao and Rao (1995) and Kumar (2002, 2008) the data was then normalized in order to train network.

The model development data set was further divided into three subsets; training set, cross validation set and testing set in 70:15:15 proportion. The first subset was used for computing and updating the network weights and biases where as second subset was used for cross validation. The error in validation set was monitored during the training process. When validation error increases for a specified number of iterations, training was stopped, and weights and biases at the minimum of the validation error were returned. The most common architecture composed of the input layer, where data are introduced, the hidden layer where data were processed and output layer where results of given inputs were obtained was used with learning functions namely Levenberg- Maquardt learning algorithm (Zanetti *et al*., 2007). Single hidden layer with sigmoid transfer function was used. The network was trained for 5000 epochs and with goal for mean square error of 0.001. The number of neurons in input and output layers were fixed as per given combinations. The numbers of neurons in hidden layer were varied as per input parameters in input layer *i.e.* maximum twice input parameters (2n) in order to get best performance as given in Table 1.

**Table1.** Number of nodes in hidden layer

| Sr. No. | Input parameters | No. of nodes in hidden layer | Output parameters |
|---|---|---|---|
| 1 | $CO_3^{2-}$ , $HCO_3^-$, $Ca^{2+}$, $Mg^2$ | 2, 4, 6, 8 | Residual sodium carbonate (RSC) |
| 2 | $Na^+$, $Ca^{2+}$, $Mg^{2+}$ | 2, 4, 6 | Sodium adsorption ratio (SAR) |
| 3 | $Na^+$, $Ca^{2+}$, $Mg^{2+}$ | 2, 4, 6 | Soluble sodium percentage (SSP) |
| 4 | $HCO_3^-$ , $Na^+$, $Ca^{2+}$, $Mg^{2+}$ | 2, 4, 6, 8 | Permeability index (PI) |
| 5 | $Na^+$, $Ca^{2+}$, $Mg^{2+}$ | 2, 4, 6 | Kelly's ratio (KR) |
| 6 | $Na^+$, $Ca^{2+}$, $Mg^{2+}$, K | 2, 4, 6, 8 | Percent Sodium (Percent Na) |

**Performance evaluation of ANN model**

The evaluation of performance of different ANN models was done using statistical analysis. The artificial neural network technique was used for evaluation of groundwater quality parameters during post monsoon seasons for irrigation purpose.

The ANN models were tested using different statistical indicator *viz.*, MBE, RMSE and IA. (Jacovides and Kontoyiannis, 1995, Stockle *et al.*, 2004; Vazquez and Feyan, 2003, Willmott, 1982) as given below. To evaluate the effectiveness of each network and its ability to make precise prediction. The Root Mean Square Error (RMSE) can be used to determine how well the network output fits the desired output.The smaller values of RMSE ensure better performance. It is defined as follows:

$$\text{RMSE} = \left[\frac{1}{N}\sum_{i=1}^{N}\left(P_i - O_i\right)^2\right]^{0.5} \quad \ldots(1)$$

Where,

$P_i$ = Predicted for $i^{th}$ observation

$O_i$ = Observed value for $i^{th}$ observation

N = Number of observations.

The Mean bias error (MBE) is given as

$$\text{MBE} = \frac{1}{N}\sum_{i=1}^{N}\left(P_i - O_i\right) \quad \ldots(2)$$

Index of agreement (IA) provides a relative measure of error and allowing cross comparison of the model.

$$d = 1 - \left[\frac{\sum_{i=1}^{N}(P_i - O_i)^2}{\sum_{i=1}^{N}\left(\left(\left|P_i'\right| + \left|O_i'\right|\right)^2\right)}\right] \quad \ldots(3)$$

$$P_i' = P_i - \overline{O} \quad O_i' = O_i - \overline{O}$$

Where,

d = index of agreement (IA)

N = number of observations,

The correlation coefficient (r) is often used to evaluate the linear relationship between the predicted and measured values.

$$r = \frac{\sum PO - \frac{\sum P \sum O}{N}}{\sqrt{\left(\sum P^2 - \frac{(\sum P)^2}{N}\right)\left(\sum O^2 - \frac{(\sum O)^2}{N}\right)}} \quad \ldots(4)$$

## RESULTS AND DISCUSSIONS

The ANN models developed were used for prediction of KR, Percent Na, PI, RSC, SAR and SSP using neurosolutions.

**Performance of ANN for different neurons in hidden layer for KR**

Kelly's Ratio (KR) is measured considering sodium ion concentration against calcium and magnesium ion concentrations. Water sample with a KR value <1 are considered suitable for irrigation, while those with greater ratios are rendered unsuitable. During post monsoon KR values varied between 0.10 - 0.91. According to Kelly's ratio, water analyzed is suitable for irrigation during post monsoon. Table 2 showed that, correlation coefficient (r) of model 2 was highest among all three models but the IA was high in model 1. Mean bias error and RMSE for model 1 was lowest among all. MBE values with negative sign means it underestimated the values of KR. Therefore, model 1 (3-2-1) was better for predicting KR values during post monsoon season in Ratnagiri district.

**Performance of ANN for different neurons in hidden layer for percent Na**

It is observed from Table 2 that, r and IA values were found to be constant for training, testing and cross validation in all cases. But values of MBE and RMSE changes comparatively more in training, testing and cross validation to r and IA values. It may be due to the fact that MBE was found to be nearer to zero and therefore changes were more visible on small scale. MBE was lowest in model 1. MBE values with negative sign means it underestimated the values of percent Na. RMSE was lowest among all models. Performance of model 1 was better as compared to model 2 and model 3. Therefore, for post monsoon season model 1 (4-2-1 model) was best suited for prediction of percent Na values during post monsoon season for Ratnagiri district.

### Performance of ANN for different neurons in hidden layer for PI

Another modified criterion has evolved based on the solubility of salts and the reaction occurring in the soil solution from cation exchange for estimating the quality of irrigative water. Soil permeability is affected by long-term use of irrigation water and is influenced by total dissolved solids, sodium contents and bicarbonate content, etc. It is observed that, correlation coefficient (r) for training and testing and cross validation was 1.00, 0.98 and 0.99 respectively. There was very close index of agreement for training, testing and cross validation being 1.00, 0.99 and 0.99, respectively for model 3. Mean bias error was close to zero in negative direction. RMSE was observed lowest among all the models for training, testing and cross validation in post monsoon season for Ratnagiri district. In Ratnagiri district, for post monsoon season model 3 (4-6-1) model was best suited to predict PI.

### Performance of ANN for different neurons in hidden layer for RSC

The residual sodium carbonate (RSC) index of water/soil signifies the alkalinity hazard posed by it and it finds the suitability of water for irrigation in case of clay soils. Residual sodium carbonate values should be preferably less than 1.25 to be rendered suitable for irrigational purposes. Data pertaining to performance indices *viz*., correlation coefficient (r), index of agreement (IA), mean bias error (MBE) and root mean square error (RMSE) of different models for changing number of neurons is given in Table 2. It was observed that, r values in training, testing and cross validation was nearly equal for all considered models. In the case of IA change in the number of neurons in hidden layer improved performance of ANN models. But values of MBE and RMSE changed comparatively more in training, testing and validation as compared to the r values. The value of correlation coefficient and IA was equal in model 2, 3 and 4. MBE and RMSE was low as in model 2 compared to other three models. Model 2 (4-4-1) was found best suited

**Table 2.** Performance of the various groundwater parameters

| Parameter | Architecture | Data set | Correlation | MBE | RMSE | IA |
|---|---|---|---|---|---|---|
| KR | 3-2-1 | Training | 0.92 | -0.0478 | 0.1674 | 0.94 |
| | | Testing | 0.94 | -0.0211 | 0.1005 | 0.94 |
| | | Cross validation | 0.95 | -0.0759 | 0.1451 | 0.95 |
| percent Na | 4-2-1 | Training | 0.98 | -0.4682 | 3.2246 | 0.99 |
| | | Testing | 0.99 | -0.4081 | 2.0632 | 0.99 |
| | | Cross validation | 0.99 | -1.0383 | 2.6365 | 0.99 |
| PI | 4-6-1 | Training | 1.00 | -0.0270 | 0.0719 | 1.00 |
| | | Testing | 0.98 | -0.0173 | 0.0725 | 0.99 |
| | | Cross validation | 0.99 | -0.0144 | 0.0629 | 0.99 |
| RSC | 4-4-1 | Training | 1.00 | 0.0379 | 0.0713 | 1.00 |
| | | Testing | 0.99 | 0.0225 | 0.0579 | 1.00 |
| | | Cross validation | 1.00 | 0.0335 | 0.0599 | 1.00 |
| SAR | 3-6-1 | Training | 0.98 | -0.2176 | 0.3698 | 0.97 |
| | | Testing | 0.97 | -0.1348 | 0.2482 | 0.90 |
| | | Cross validation | 0.97 | -0.2137 | 0.3438 | 0.91 |
| SSP | 3-6-1 | Training | 0.99 | 0.0440 | 1.7385 | 1.00 |
| | | Testing | 0.99 | -0.4797 | 1.4172 | 1.00 |
| | | Cross validation | 0.99 | -0.2450 | 1.7354 | 1.00 |

to predict RSC in Ratnagiri district for post monsoon season among four models.

**Performance of ANN for different neurons in hidden layer for SAR**

Sodium adsorption ratio is a measure of the sodicity of the soil and is determined through quantitative chemical analysis of water in contact with it. An excess of $HCO_3^-$ and $CO_3^{2-}$ ions in water react with $Na^+$ in soil, resulting in a sodium hazard. Based on the SAR values all samples have low sodium hazard and hence can be considered suitable for irrigation. High correlation coefficient (r) for training (0.98), testing and cross validation (0.97) were observed. A high index of agreement for training (0.97), testing (0.90) and cross validation (0.91) was observed. Mean bias error was found to be close to zero. The biasness which was indicated by Mean Bias Error (MBE) represents underestimation as it was negative. RMSE was close to zero for training, testing and cross validation during post monsoon season for Ratnagiri district. Therefore, Model 3(3-6-1) was best suited to predict RSC in Ratnagiri district for post monsoon season among three models is given in Table 2.

**Performance of ANN for different neurons in hidden layer for SSP**

High sodium ion concentration in soil can take a toll on internal drainage patterns in soil as release of calcium and magnesium ions are facilitated due to absorption of sodium by clay particles. The SSP values range from 5.02 - 42.01 in post monsoon. The SSP values was suitable for irrigation during post monsoon. It is observed from Table 2 that the correlation coefficient (r) for training, testing and cross validation was 0.99 and index of agreement for training, testing and cross validation was 1.00 for model 3. Mean bias error was close to zero and lowest among all models. The biasness as indicated by MBE represents underestimation when it is negative. RMSE was also lowest among the entire model for training, testing and cross validation during post monsoon season for Ratnagiri district. Therefore, Model 3 (3-6-1) was best suited to predict SSP in Ratnagiri district for post monsoon season among three models.

The different ANN models were developed for prediction of water quality indicators such as KR, % Na, PI, RSC, SAR and SSP to test the suitability of groundwater for irrigation purpose for post monsoon periods. The output of predicted water quality indicators were compared with observed parameters using statistical performance indicators. The analysis revealed that selected ANN based modelshad correlation coefficient more than 0.98, index of agreement more than 0.98 with less bias error and mean error in terms of MBE and RMSE for post monsoon season. The analysis revealed that for the selected ANN based models had correlation coefficient (r>0.95), RMSE (<0.3438), IA (>0.91) and MBE (<-0.2450) for Model 3-2-1 to predict KR, Model 4-2-1 to predict percent Na, Model 4-6-1 to predict PI, Model 4-4-1 to predict RSC and Model 3-6-1 to predict SAR and SSP during post monsoon season. The results confirmed that the developed ANN models were suitable for prediction of water quality indicators used for irrigation purpose. It was recommended that two numbers of nodes in hidden layer can be used for modelling of KR and percent Na. Four number of nodes for RSC and six numbers of nodes can be used for SAR, SSP and PI under limited data conditions. The present study revealed that the quality parameters for more than 94.60% of the total number of samples were within the permissible limits for irrigation during post monsoon season, with a very few isolated exceptions as per irrigation water quality standards.Thus, it could be concluded that the groundwater of the study area is in general suitable for irrigation purpose.

## REFERENCES

Ehteshami, M., Dolatabadi Farahani N. and Tavassoli, S. (2016). Simulation of nitrate contamination in groundwater using artificial neural networks. *Modeling Earth Systems and Environment* **2**:28, DOI 10.1007/s40808-016-0080-3.

Freeze, R. A. and Cherry, J. A. (1979). *Groundwater*, Prentice Hall, Englewood Cliffs, 604 p.

Jacovides, C. P. and Kontoyiannis, H. (1995). Statistical procedures for the evaluation of evapotranspiration computing models. *Agricultural Water Management* **74**: 87-97.

Jeevanandam, M., Kannan, R., Srinivasalu, S. and Rammohan, V. (2006). Hydrogeochemistry and groundwater quality assessment of lower part of the Ponnaiyar river basin, Cuddalore District, South India. *Environmental Monitoring and Assessment* **132**(1-3): 263-274.

Junsawang, P., Asavanant, J. and Lursinsap, C. (2007). Artificial neural network model for rainfall-runoff relationship, Proceeding 2nd Asian Simulation and Modelling (ASSIMOD2007), January 9-11, 2007, Chaing Mai, Thailand. pp 267-274.

Khandelwal, M. and Singh, T. N. (2005). Prediction of mine water quality by physical parameters. *Journal of Scientific and Industrial Research* **64**: 564-570.

Kumar, M., Bandyopadhyay, A., Rahguwanshi, N.S. and Singh, R. (2008). Comparative study of conventional and artificial neural network based ETS estimation models. *Irrigation Science* **26**: 531–545.

Kumar, M., Raghuwanshi, N. S., Singh, R., Wallender, W. W. and Pruitt, W. O. (2002). Estimating evapotranspiration using artificial neural network. *Journal of Irrigation and Drainage Engineering*, ASCE **128**: 224–233.

Laluraj, C. M. and Gopinath, G. (2006) Assessment on seasonal variation of groundwater quality of phreatic aquifers - a river basin system. *Environment Monitoring and Assessment* **117**: 45-47.

Nagarajan, R., Rajmohan. N., Mahendran, N., Senthamilkumar, S. (2009). Evaluation of groundwater quality and its suitability for drinking and agricultural use in Thanjavur city, Tamil Nadu, India. *Environment Monitoring and Assessment* **171**(1-4): 289-308. doi: 10.1007/s10661-009-1279-9.

Prasanna, M. V., Chidambaram, S., Gireesh, T. V. and Jabir Ali, T. V. (2010). A study on hydrochemical characteristics of surface and sub-surface water in and around Perumal Lake, Cuddalore district, Tamil Nadu, South India. *Environmental Earth Sciences* **63**(1): 31-47.

Rao, V. B. and Rao, H. (1995). *C++ Neural Networks and Fuzzy Logic*, Second Edition, BPB Publications. 551 p.

Stockle, Claudio O., Kjelgaard, J. and Bellocchi, G. (2004). Evaluation of estimated weather data for calculating Penman-Monteith reference crop evapotranspiration. *Irrigation Science* **23**: 39-46.

Tyagi, S. K., Datta, P. S. and Pruthi, N. K. (2009). Hydrochemical appraisal of groundwater and its suitability in the intensive agricultural area of Muzaffarnagar district, Uttar Pradesh, India. *Environmental Geology* **56**(5): 901-912.

Vazquez, R. F. and Feyen, J. (2003). Effect of potential evapotranspiration estimates on effective parameters and performance of the MIKESHE-code applied to a medium-size catchment. *Journal of Hydrology* **270**: 309-327.

Willmott, C. J. (1982). Some comments on the evaluation of model performance. *American Meteorological Society* **63**: 1309-1313.

Yesilnacar, M. I., Sahinkaya, E., Naz, M. and Ozkaya, B. (2008). Neural network prediction of nitrate in groundwater of Harran Plain, Turkey. *Environmental Geology* **56**: 19-25.

Zanetti, S. S., Sousa, E. F., Oliveira, V. P. S., Almeida, F. T. and Bernardo, S. (2007). Estimating evapotranspiration using artificial neural network and minimum climatological data. *Journal of Irrigation and Drainage Engineering*, ASCE. **133**: 83-89.