Application of Optimized Neuro-Fuzzy Models for Estimation of Water Quality Index in Karun River

Mohammad Lotfi Dashbolagh[1] and Mohammad Najafzadeh[2*]

[1] Master of Water Resources Management, Water Engineering Department, Faculty of Civil and Surveying Engineering Graduate University of Advanced Technology

[2*] Assistant Professor, Water Engineering Department, Faculty of Civil and Surveying Engineering, Graduate University of Advanced Technology

ABSTRACT

Management of water quality is inextricably bound up with making good management decisions and this typical management is at the mercy of predicting Water Quality Index (WQI). The use of board range of artificial intelligence models for conducting analysis of surface water quality is one of the most efficient techniques to predict water quality parameters and WQI. In the current research, at the first, datasets accumulated from nine hydrometry stations, located in Karun River, were included those of 13 water quality parameters (i.e., Dissolved Oxygen, Chemical Oxygen Demand, Biochemical Oxygen Demand, Electrical Conductivity, Nitrate, Nitrite, Phosphate, Turbidity, pH, Calcium, Magnesium, Sodium, and water temperature) which was used to estimate WQI. So, in order to obtain optimal selection of ANFIS model-feeding-input variables, Gamma Test (GT), Forward Selection (FS), and Principle Component Analysis (PCA) were applied. Ultimately, constant coefficients of membership function used in ANFIS model were computed by using evolutionary techniques including Genetic Algorithm (GA), Ant Colony Optimization (ACO), and Particle Swarm Optimization (PSO) were used to train the structure of ANFIS model. Results of statistical assessments indicated that GT-ANFIS-PSO model with Correlation Coefficient of 0.952, Mean Absolute Error of 1.68, and, Root Mean Square Error of 3.05 had satisfying performance for prediction of WQI compared with other optimized ANFIS models. Moreover, values of WQI ranged from 30 to 58.4 which were indicative of being relatively poor to good water quality of Karun River.

Keywords: **Water Quality Index; Adaptive NeuroFuzzy Inference System; Sensitivity Analysis; Heuristic Algorithms; Karun River**

* Corresponding Author: Email: **m.najafzadeh@kgut.ac.ir**

1. Introduction

Prediction of Water Quality Index (WQI) plays a key role in monitoring of natural water bodies such as rivers, lakes, estuarine, and ocean environments. There are a variety of conventional methodologies for estimation of WQI in natural rivers [1-3]. The most well-known WQI model is associated with National Sanitation Foundation (NFS) which has been frequently applied for various real world problem in water quality monitoring [1-3]. The current WQI methodology may occasionally experience a number of restrictions into water quality applications. The major drawback of NSF model is related to the lack of available Water Quality Parameters (WQPs) for calculation of WQI. Nowadays, Artificial Intelligence (AI) models were used to predict WQI and WQPs with satisfying accuracy level [1-3]. In accordance with this, the most frequently-used AI models into water quality monitoring are as Artificial Neural Networks (ANNs), Adaptive Neuro-Fuzzy Inference System (ANFIS), Gene-Expression Programming (GEP), M5 Model Tree (MT), and Evolutionary Polynomial Regression (EPR) [1-3]. Hence, this study investigates WQI for Karun River by 200 recorded WQPs series data over a 16-years period beginning at May 1995. In this study, 13 WQPs are used to predict WQI on the basis of NSF instructions. Furthermore, the optimal number of WQPs in order to estimate WQI value are obtained by using Principle Component Analysis (PCA), Gamma Test (GT), and Forward Selection (FS). Afterwards, ANFIS model are improved in the training stages by three evolutionary algorithms, as Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), and Genetic Algorithm (GA). Ultimately, the performance of various optimized ANFIS model are evaluated by different statistical benchmarks and additionally water quality of Karun River is classified.

2. Methodology

2.1. Case study

In the current study, 13 WQPs including Dissolved Oxygen (DO=3.1-29.4mg/l), Biochemical Oxygen Demand (BOD=3.7-40mg/l), Chemical Oxygen Demand (COD=1.06-34.2mg/l), Potential of Hydrogen (pH=1.5-8.71), Turbidity (Tur=1-25 NTU), Sodium (Na^+=1.42-40 mg/l), Magnesium (Mg^{2+}=2.1-60mg/l), Electrical Conductivity (EC=1.7-2.2 dS/m), Nitrite (NO_2^-=0.08-1.2mg/l), Nitrate (NO_3^-=0.34-2.7mg/l), Phosphate (PO_4^{3-}=0.13-3.21mg/l), Calcium (Ca^{2+}=1-35mg/l), and Temperature (T=9.3-30.3 ^{0}C) were used to predict monthly WQI based on NSF guidelines. All the water quality parameters were accumulated from nine hydrometry stations (i.e., 5th Bridge, Khoramshahr, Zergan, Kutamir, Deirfam, Marun, Mollasani, Darkhuien, and Nahr-e-Ghasbeh) in Karun River. All the observational datasets have been ranged in May 1995- Jan 2012. According to the NSF guidelines, WQI values are obtained by recorded WQPs and there are seven classifications for WQI values: very poor (WQI<15), poor (WQI=15-29.9), relative poor (WQI=30-44.9), moderate (WQI=45-55), relative moderate (WQI=55.1-70), good (WQI=70.1-85), and very good (WQI>85). Furthermore, the following relationship is generally used to develop ANFIS,

$$WQI=\varphi\,(DO, BOD, COD, pH, Tu, EC, Ca^{2+}, Na^{+}, Mg^{2+}, NO_2^-, NO_3^-, PO_4^{3-}, T) \quad (1)$$

2.2. Optimal Selection of WQPs

In this research, three well-known sensitivity analysis were used to choose optimal number of WQPs. In the first place, FS techniques select 10 water quality parameters (i.e., BOD, COD, DO, PO_4^{3-}, EC, pH, NO_3^-, Mg^{2+},Ca^{2+},Tu) among 13 WQPs, as the most contributory factors on prediction of WQI values. Additionally, results of Gamma Test (GT) indicated that COD and Mg^{2+} have the lowest level of contribution on estimation of WQI. Hence, the rest of WQPs (11 input variables) can be used to develop ANFIS technique. Finally, PCA technique converted 13 WQPs into four components as 4 input variables.

3. Development of ANFIS model

In this section, Fuzzy C Mean (FCM) was used to generate the structure of ANFIS model. Number of clusters in ANFIS model is in close connection with number of input variables. In this way, there are 4, 10, and 11 clusters (or fuzzy rules) for application of preprocessing-data-strategies: PCA, FS, and GT, respectively. Moreover, weighting coefficients associated with Gaussian Membership Functions were optimized by GA, PSO, and ACO.

4. Results and discussion

To assess performance of ANFIS models underlying evolutionary algorithms, three statistical benchmarks such as correlation of Coefficient (R), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) were employed. In accordance with applying PCA, ANFIS-PSO had the best performance (R=0.88, RMSE=3.45, and MAE=2.54) in the training stage compared with the ANFIS-ACO and ANFIS-GA techniques. Also, in the testing phase, ANFIS-PSO with R of 0.89, RMSE of 3.41, and MAE of 2.69 predicted the

WQI with most accuracy level than ANFIS-GA (R=0.84, RMSE=3.7, and MAE=2.84) and ANFIS-ACO (R=0.83, RMSE= 3.51, and MAE=2.95).

In the case of FS application, performance of training phase demonstrated that ANFIS-GA (R=0.942, RMSE=3.16, and MAE=1.82) had the most accurate estimation for the WQI in comparison with ANFIS-ACO and ANFIS-GA techniques. From testing results, it was found that application of PSO into structure of ANFIS model (R=0.943, RMSE=3.13, and MAE=1.87) provided compromising estimation of WQI rather than ANFIS-GA (R=0.88, RMSE=3.5, and MAE=2.1) and ANFIS-ACO (R=0.889, RMSE=3.49, and MAE=2.21). As GT was applied as a preprocessing-data-technique, statistical results of training phase showed that ANFIS-PSO (R=0.88, RMSE=3.45, and MAE=2.54) had superiority to other optimized ANFIS models. Furthermore, it was inferred from testing results that ANFIS-PSO model (R=0.89, RMSE=3.41, and MAE=2.69) produced the most convincing efficiency for WQI estimation when compared to ANFIS-GA (R=0.84, RMSE=3.7, and MAE=2.84) and ANFIS-ACO (R=0.83, RMSE=2.95, and MAE=3.51). Overall, statistical benchmarks given by training and testing stages demonstrated that GT-ANFIS-PSO had the most successful performance in the WQI estimation than other ANFIS models fed by PCA and FS.

5. Conclusions

In this research, the WQI-NSF values of Karun River were predicted by using optimized ANFIS techniques. Results of statistical assessments indicated that FS and GT strategies had the highest impact on reducing number of WQPs, leading to optimal selection of input variables for feeding ANFIS technique. Furthermore, application of PSO in order to train GMFs utilized in ANFIS had the most accurate performance when compared with ANFIS-GA and ANFIS-ACO. In addition to this, Moreover, WQI for Karun River were between 30 and 58.4 which have been classified into relatively poor to good quality. With regard to the best ANFIS model, it can be used to monitor and management the water quality of Karun River as a cost-effective method.

6. References

[1] Z.M. Yaseen, M.M. Ramal, L. Diop, O. Jaafar, V. Demir, O. Kisi, Hybrid adaptive neuro-fuzzy models for water quality index estimation, Water Resources Management, 32(7) (2018) 2227-2245.

[2] M. Najafzadeh, A. Ghaemi, S. Emamgholizadeh, Prediction of water quality parameters using evolutionary computing-based formulations, International Journal of Environmental Science and Technology, 16(10) (2019) 6377-6396.

[3] M. Najafzadeh, A. Ghaemi, Prediction of the five-day biochemical oxygen demand and chemical oxygen demand in natural streams using machine learning methods, Environmental monitoring and assessment, 191(6) (2019) 380.