

Prediction of River Water Quality by Adaptive Neuro Fuzzy Inference System (ANFIS)

Archive of SID

Safavi, H.R.

Assoc. Prof., Department of Civil Engineering, Isfahan University of Technology, Isfahan- Iran.

Received: Dec., 2008

Accepted: Nov., 2009

Extended Abstract

Limitations on freshwater resources have caused researchers and water resources managers to focus an increasing attention over the past few decades on water quality protection. Surface water quality management in such resources as rivers, seas, lakes, and estuaries is of a greater importance than other water resources and a greater number of studies have been conducted on them as they are more accessible and, therefore, more directly exposed to a variety of contaminants and pollutants. Over the past two decades, industrial development and population growth with the ensuing escalation in waste disposal from rural, industrial, and agricultural areas have led to an increasing release of pollutants into receiving surface waters, which happen in most cases to be the main sources of drinking water, thereby jeopardizing public health. The response of water resources to such varied types of pollutants requires proper investigation and water quality modeling based on water resources quality standards in order to estimate the capacity of the receiving waters for handling the incoming pollution. Application of appropriate and efficient mathematical models for river water quality simulation is essential for the formulation of comprehensive guidelines used in evaluating measures that are employed for river pollution control and management. The non-linear equations dominating pollutant transfer phenomena in rivers, the complexity of their simultaneous solution, and the multiplicity of kinetic constants and coefficients have made it difficult, or at times impossible, to use physically-based models and methods for this purpose. Therefore, most of these models can only be applied to simplified cases or to situations where the models are strictly calibrated and validated, with no adequate accuracy when applied to unrestricted conditions. The uncertainties in water quality problems have made fuzzy inference systems, especially as combined with adaptive neural networks, to be used as a novel approach. The main objective of the present study is to exploit the capabilities of the adaptive neuro-fuzzy inference system (ANFIS) for river quality predictions with emphasis on DO and BOD. For this purpose, the Zayandehroud River in Isfahan is selected as a case study for which adequate hydraulic and quality data are available. A fuzzy system is one which is based on the logical rules of premises and conclusions that cannot be analyzed with the classical probability theories. The starting point in constructing a fuzzy system is the development of a series of fuzzy "if-then" rules. An efficient tool to this end is a method that can transform the available data into the required fuzzy rules. On the other hand, artificial neural networks are capable of creating relevant relations between input and output variables through their learning capabilities based on various training patterns. A combined system of fuzzy inference and artificial neural network capable of using numerical data for predicting outputs can create a powerful tool that has come to be called the adaptive neural-fuzzy inference system. In this combined system, the fuzzy part establishes relations between input and output variables while the fuzzy membership functions are determined by its neural network part.

The Zayandehroud River is the most important river in central Iran which stretches over a length of 360 km, originating in the Zardkough Mountain in Bakhtiari and ending in the Gavkhooni Marshland after passing through the city of Isfahan.

www.SID.ir

At a distance of 110 km west of Isfahan, the origins of the river flow into the Zayandehroud storage dam with a storage capacity of 1450 MCM. In addition to the natural flows from various springs in the region, transferred flows from neighboring watersheds are also diverted into this storage dam through the Kouhrang and Cheshmeh Lengan tunnels. The river's natural flows amount to an annual average of around 950 MCM, and the flows from the first, second, and the Cheshmeh Lengan tunnels account for an additional 850 MCM. The annual natural flows downstream the dam reservoir make up a total flow of 200 MCM.

Due to the numerous cities, towns, agricultural sites, and industrial developments around and along the river, the Zayandehroud River receives a variety of pollutants causing its water quality to severely downgrade as it distances away from the dam site.

Industrial effluents account for 7.2 MCM, urban effluents for 47.5 MCM, and farm drainage for 17 MCM of the pollution released into the river annually. In water

quality predictions for the Zayandehroud River, DO and BOD have been taken as the main indexes. As already mentioned, the river's water loses its quality after it passes through the city of Isfahan.

Based on physically-based models and the equation governing the quality-quantity processes in surface waters, the main hydrodynamic and quality factors affecting BOD levels at each section of the river include: river flow (Q) at the section, water temperature (T), water depth (d), velocity (V), and BOD level measured at one and two upstream stations. The factors affecting DO level include: river flow (Q) at the section, water temperature (T), water depth (d), surface width (W), velocity (V), and DO levels measured at one and two upstream stations. The 16-year data available were used in this study which included 10-year data used for training, 3-year data for calibration, and 3-year data for validation and control. Minimum, maximum, mean, and standard deviation values for the DO and BOD parameters are reported in Table (1). The hydraulic characteristics of this section of the river including flow rate, velocity, depth, and surface width were extracted from the hydrometric station at Isfahan Regional Water Company.

Table 1: BOD and DO statistical figures measured at Pol-e-Ziar on the Zayandehroud River

Parameter	Min. (mg/l)	Max (mg/l)	Mean (mg/l)	SD (mg/l)
BOD	16	45	31.2	6.41
DO	1.5	5.1	2.6	0.82

For each parameter, 112 runs of the adaptive neuro-fuzzy inference system using the networking segregation method showed that the best results in the calibration stage for BOD predictions at the Pol-e-Ziar station could be achieved by those models in which the vectors of input data include river flow at the station and one and two upstream stations, temperature, and BOD at one and two upstream stations. It must be mentioned that in the MATLAB toolbox (MATLAB 7.4) both methods of networking segregation and partial clustering can be used to employ the adaptive neuro-fuzzy inference system. In the present study, we used the networking segregation in which type and number of membership functions of the input data are determined by the user. Table(2) presents the criteria for error evaluation of three different ANFIS types with different data sets for calibration and validation stages for BOD quality parameter.

Table 2: Evaluation of the performance of three ANFIS types according to different BOD

Model	Input variables	R ²		RSME		NSME		MAPE	
		calibration	validation	calibration	validation	calibration	validation	calibration	validation
NF1	Q, T, BOD ₋₂ , BOD ₋₁	0.953	0.931	1.182	1.376	0.432	0.512	0.043	0.068
NF2	BOD ₋₂ , BOD ₋₁ , V, d, Q, T	0.849	0.768	2.141	2.437	0.386	0.418	0.036	0.051
NF3	Q, T, V, d	0.796	0.729	2.685	2.915	0.616	0.721	0.056	0.072

Clearly, all the three models show a satisfactory performance in predicting BOD values. This indicates the capability of ANFIS models in making predictions on the basis of input data sets. The first model (NF1), however, whose data set includes flow, temperature, and BOD measurements at one and two upstream stations exhibits a better performance compared to the other two. It can, therefore, be concluded that flow velocity and depth play a minor role as input data in predicting BOD values and that flow and temperature have a higher sensitivity.

Table (3) presents the criteria for error evaluation of three ANFIS models with different data sets for calibration and validation for DO.

Table 3: Evaluation of the performance of three ANFIS types according to different DO

Model	Input variables	R ²		RSME		NSME		MAPE	
		calibration	validation	calibration	validation	calibration	validation	calibration	validation
NF4	V, d, T, W, DO ₋₁	0.921	0.904	1.496	1.654	0.548	.0593	0.058	0.069
NF5	T, Q, DO ₋₂ , DO ₋₁ , d	0.793	0.726	1.743	1.978	0.621	0.768	0.066	0.071
NF6	Q, DO ₋₁ , W, d	0.726	0.679	2.234	2.431	0.751	0.726	0.073	0.079

In the case study carried out on the Zayandehroud River, BOD predictions were obtained by the proposed system with a correlation coefficient of 0.953 in the calibration stage and 0.931 in the validation stage and DO predictions were obtained with a correlation coefficient of 0.921 in the calibration stage and 0.904 in the validation stage. Comparison of the results obtained from the six different types (three models for BOD and three for DO) revealed that increasing the number of input parameters does not necessarily enhance model predictive accuracy but that identification of the effect of each input parameter and sensitivity analysis of the parameters involved play a more important role. Comparison of the results provided by the adaptive neuro-fuzzy inference system and the measured values reveals the high accuracy level of the proposed model. It must always be remembered that although data-based models can do well without physically-based equations which are typically complex and non-linear, understanding the effects of the processes involved and input parameters on the output is of crucial importance, without which no satisfactory results can be expected.

Key words

Water quality, Fuzzy logic, Neural network, Zayandehroud River, Water resources management