Supervised committee machine with artificial intelligence for prediction of fluoride concentration

Ata Allah Nadiri, Elham Fijani, Frank T.-C. Tsai and Asghar Asghari Moghaddam

ABSTRACT

The study introduces a supervised committee machine with artificial intelligence (SCMAI) method to predict fluoride in ground water of Maku, Iran. Ground water is the main source of drinking water for the area. Management of fluoride anomaly needs better prediction of fluoride concentration. However, the complex hydrogeological characteristics cause difficulties to accurately predict fluoride concentration in basaltic formation, non-basaltic formation, and mixing zone. SCMAI predicts fluoride by a nonlinear combination of individual AI models through an artificial intelligent system. Factor analysis is used to identify effective fluoride-correlated hydrochemical parameters as input to AI models. Four AI models, Sugeno fuzzy logic, Mamdani fuzzy logic, artificial neural network (ANN), and neuro-fuzzy are employed to predict fluoride concentration. The results show that all of these models have similar fitting to the fluoride data in the Maku area, and do not predict well for samples in the mixing zone. The SCMAI employs an ANN model to re-predict the fluoride concentration based on the four AI model predictions. The result shows improvement to the CMAI method, a committee machine with the linear combination of AI model predictions. The results also show significant fitting improvement to individual AI models, especially for fluoride prediction in the mixing zone.

Key words | artificial intelligence, artificial neural network, committee machine, fuzzy logic, ground water quality, neuro-fuzzy

Ata Allah Nadiri
Elham Fijani
Frank T.-C. Tsai (corresponding author)
Department of Civil and Environmental Engineering,
Louisiana State University,
3418G Patrick F. Taylor Hall,
Baton Rouge,
LA 70803,
USA
E-mail: *ftsai@lsu.edu*

Ata Allah Nadiri
Elham Fijani
Asghar Asghari Moghaddam
Department of Geology,
Faculty of Science,
University of Tabriz,
29 Bahman Boulevard,
Tabriz,
East Azarbaijan,
Iran

INTRODUCTION

Determination of ground water contamination is one of the major studies in hydrogeology. In this study, we focus on fluoride contamination that has drawn public attention due to its considerable impact on human life. Fluoride is an important constituent in drinking water. Although its trace amount in the diet is required for increasing the strength of bones and teeth, excessive amounts can also affect detrimentally, such as endemic fluorosis, on human health. The effects of fluoride on human health are well studied (Kharb & Susheela 1994). Different kinds of fluorosis were discussed by Jacks *et al.* (2005). Potential health effects associated with long-term ingestion of fluoride-bearing water were presented by Dissanayake (1991). Recently, the implication of fluoride ingestion for human health was discussed in detail (Rao 2003). Different permissible concentrations of fluoride in drinking water were reported by various agencies, such as the US Environmental Protection Agency (USEPA 2009). The World Health Organization recommends a guideline maximum fluoride value of 1.5 mg/L as a level at which fluorosis should be minimal (WHO 2006).

The occurrence of high fluoride concentration in ground water has been reported in different areas around the world (Gizaw 1996; Jacks *et al.* 2005; Guo *et al.* 2007). Naturally, the occurrence arises from geological settings that have fluoride bearing minerals (e.g. topaz, fluorite, fluorapatite, cryolite, amphiboles and mica); and its concentration in ground water is controlled by physicochemical parameters (Handa 1975; Bårdsen *et al.* 1996; Saxena & Ahmed 2003). Several studies

doi: 10.2166/hydro.2013.008

have shown the effect of water–rock interaction to fluoride occurrence in various aquifers with different geological settings (Saxena & Ahmed 2001; Gupta *et al.* 2005; Ozsvath 2006). In addition to the geological settings, the tectonic factor and the regional climate play an important role in affecting the fluoride distribution in ground water (Handa 1975). The complexity of the distribution of geogenic contamination of water resources such as fluoride contamination stems from geological condition complexity.

Because of the vital impact of fluoride on human health and complexity of its distribution, spatial prediction of fluoride is necessary for planning, risk assessment, decision-making, and management of ground water resources. For example, Zietsman (1991) applied statistical methods to study the spatial variation of fluorosis and fluoride content of water in an endemic area in Bophuthatswana and found that the fluoride content in the water did not adequately explain the differences in the occurrence of fluorosis. Hudak & Sanmanee (2003) used a geographic information system (GIS) to map and analyze nitrate, chloride, sulfate, and fluoride concentrations in 110 wells tapping the Woodbine Aquifer, Texas, USA. They deduced that volcanic ash deposits influenced fluoride concentrations.

Using physical-based simulation models to predict hydrochemical distribution in aquifer systems is rigorous, but is constrained on several limitations: (1) the models need sufficient spatial and temporal data, which may not be practical for initial site investigation; (2) the development of these models requires detailed characterization of the study area including the physical and chemical processes when such processes may not be fully known; and (3) the models usually utilize fine spatiotemporal discretization that requires substantial computational resources for multiple scenarios (McGrail 2001).

Recently, artificial intelligence (AI) models such as fuzzy logic (FL), artificial neural network (ANN), and neuro-fuzzy (NF) have been successfully utilized at different fields of hydrology, e.g. rainfall–runoff modeling (Nayak *et al.* 2004; Tayfur & Singh 2006), water levels forecasting in space and time (Chang & Chang 2006; Nadiri 2007; Nourani *et al.* 2008), parameter estimation (Garcia & Shigidi 2006; Merdun *et al.* 2006; Tutmez & Hatipoglu 2010; Anifowose & Abdulraheem 2011; Nadiri *et al.* 2013a), and water quality modeling (Govindaraju & Rao 2000; Gutiérrez-Estrada *et al.* 2004; Dixon 2005; Mark *et al.* 2010; Ranković *et al.* 2012). Some studies focus on the prediction of water contamination via AI models, which have high capability to predict hydrochemical anomalies (Sharma *et al.* 2003; Almasri & Kaluarachchi 2005; Shekofteh *et al.* 2013).

This study emphasizes the importance of using multiple AI methods for hydrochemical prediction rather than a single AI method. Each AI method has its own advantages. The fuzzy models tend to be robust to parameter changes, and are also tolerant to imprecision and uncertainty (Bárdossy & Disse 1993). Besides, the ANN model represents non-linear relationships and learns these relationships directly from the data being modeled (Palani *et al.* 2008). Obviously, the NF model takes advantage of the FL and ANN in modeling.

None of the AI methods can dominate other AI methods. Using multiple AI models for prediction may achieve the optimal performance, reap the benefits of all AI models, and avoid bias to a single AI model.

A combination of multiple AI models to estimate hydrological parameters was suggested using a committee machine with artificial intelligence (CMAI) method (Lim 2005; Chen & Lin 2006; Kadkhodaie-Ilkhchi *et al.* 2009). CMAI makes predictions by linearly combining the outputs of individual AI models through a set of weights. There are two methods to determine weights for CMAI (Chen & Lin 2006; Labani *et al.* 2010): simple averaging using equal weights and weighted averaging using optimized weights. Kadkhodaie-Ilkhchi *et al.* (2009) and Labani *et al.* (2010) used a genetic algorithm (GA) to optimize weights and found that weighted averaging performed superior to simple averaging for parameter estimation.

Instead of linearly combining AI models, this study introduces a supervised committee machine with an artificial intelligence (SCMAI) method that replaces linear combination with an ANN. In the SCMAI, the ANN receives individual model predictions as input and derives new predictions. The advantage of the SCMAI is the non-linear combination of the AI models under supervision such that the SCMAI may perform better for hydrochemical prediction in complex aquifer systems.

This study tests the SCMAI for predicting fluoride concentration in the Maku area, north of West Azarbaijan, northwest of Iran. The Maku area is mainly underlain by basaltic and non-basaltic aquifers. Previous studies indicated

a high concentration of fluoride in these complex aquifers in the study area (Asghari Moghaddam & Fijani 2008, 2009; Asghari Moghaddam *et al.* 2010). A better understanding on this anomaly is necessary in order to suitably manage water resources in this area. In this study, four artificial intelligence models, Sugeno fuzzy logic (SFL), Mamdani fuzzy logic (MFL), ANN, and NF, are adopted to test the SCMAI and CMAI methods. However, the SCMAI and CMAI can include any number of AI models for multimodel analysis.

STUDY AREA

The study area is shown in Figure 1. The Maku area is located in the north of West Azarbaijan province, northwest of Iran, on the foothills of Ararat Mountain. It is bounded on the west by Turkey and on the east by the Aras River. The area is approximately 1,600 km^2, where up to 400 km^2 is covered by basaltic lavas. Maku, Poldasht and Bazargan are the main cities in the Maku area. Two main rivers, Sari Su and Zangmar, flow across the study. The climate in the area is arid with a yearly mean temperature 10.85 °C (1996–2006). The annual mean precipitation is about 291.7 mm (1992–2006). The maximum and minimum precipitations occur in May and September, respectively. The annual evaporation is 1,500 mm, five times more than annual precipitations. In the Maku area, ground water is the main water source to be used for various purposes such as drinking, agriculture and industry. Ground water discharges at 12 large-scale springs ranging from 20 to

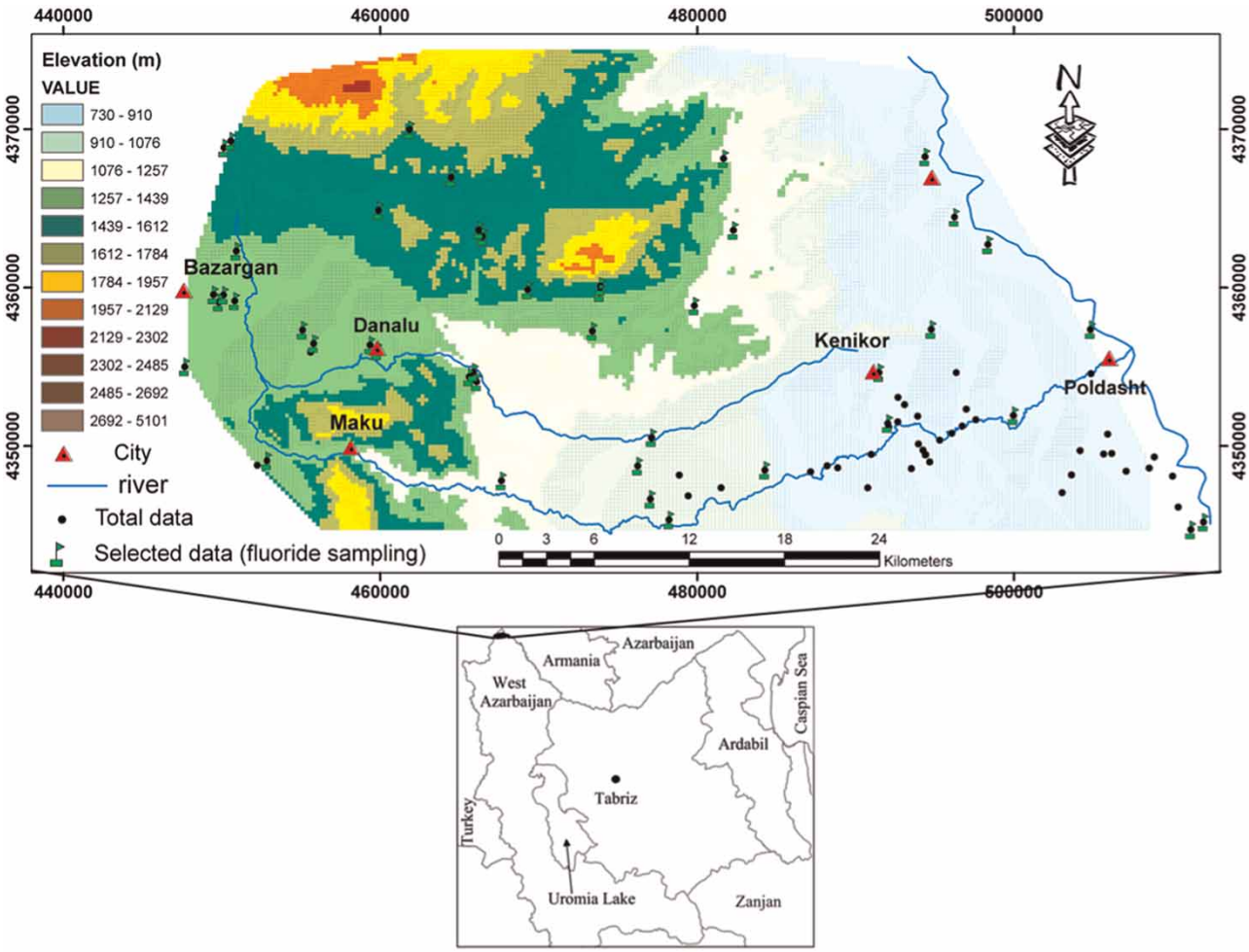

Figure 1 | Topographical map of the study area and sample locations.

4,000 L s^{-1} and at several withdrawal wells (Asghari Moghaddam & Fijani 2009).

The geological formations are shown in Figure 2. Hydrogeologically, two major formations dominantly cover the study area: (a) the basaltic area, where the aquifer mainly consists of basaltic rocks and covered by alluvium (basalt-alluvium aquifer), and (b) the non-basaltic area, where the aquifer consists of limestones of Qom formation (Oligo-Miocene age) and massive limestone and dolomites of Ruteh formation (Permian). Other formations, such as shale, marl and conglomerate, are presented in small areas. According to the prior studies (Asghari Moghaddam & Fijani 2008, 2009; Asghari Moghaddam *et al.* 2010), ion concentration, electrical conductivity (EC), total dissolved solids (TDS) and alkalinity in the basaltic aquifer are higher than those in the non-basaltic aquifer.

From a geoelectrical surveying in the Bazargan Plain with two profiles including 13 soundings, the thickness of the basalt-alluvium aquifer in this area is estimated to be about 150 m (Asghari Moghaddam & Fijani 2009). The basaltic aquifer is intensively karstified in some places. High fluoride ground water in the study area is mostly concentrated in basaltic rocks. Mixing of non-basaltic and basaltic ground waters causes a variation of fluoride concentration in some places (which are called the mixing zone). Some of these zones are located in the basaltic aquifer and some in the non-basaltic aquifer. Habitants who consume water from basaltic springs and wells suffer from dental fluorosis. Asghari Moghaddam & Fijani (2008) found that low ground water flow rate, rock geochemistry, prolonged water–rock interactions, and relatively high pH, HCO_3^- and Na^+ are the main factors controlling fluoride concentration in the ground water beneath the study area. This study focuses on predicting fluoride concentration using fluoride-related water quality parameters via the SCMAI model.

Factor analysis to determine input data

To determine important input data of water quality parameters in the AI models to estimate the fluoride concentration, we collected water samples from 131 water sources (wells, springs, ganats, etc.) in a five-year period from 2004 to 2008. The sampling locations are indicated in Figure 1. The sampling time and data size of the seven data sets are listed in Table 1. We had a largest data set (38 samples) in August 2006 and a smallest data set (eight samples) in July 2004.

The samples were analyzed in the hydrogeological laboratory in the University of Tabriz for fluoride ion (F^-), major ions (Ca^{2+}, Mg^{2+}, Na^+, K^+, HCO_3^-, CO_3^{2-}, SO_4^{2-}, and Cl^-), pH, electrical conductivity (EC), and silicon (SiO_2), following the standard procedure of APHA (1998). Asghari Moghaddam & Fijani (2008) summarized the methods used for analysis of fluoride. The observed precision of ion analysis was within the range of acceptability (±5%) used in most of chemical analysis. The statistical characteristics of the fluoride and other 11 hydrochemical parameters are presented in Table 2. The correlation matrix of all water quality parameters is presented in Table 3. Na^+, K^+, HCO_3^-, pH, EC and SiO_2 are positively correlated with F^-; Ca^{2+}, Mg^{2+}, CO_3^{2-} and SO_4^{2-} are negatively correlated with F^-; and Cl^- shows extreme low correlation with F^-. Input selection and data reduction to construct an AI model were discussed by Nourani & Sayyah Fard (2012) and Nourani *et al.* (2012).

In this study, we introduce factor analysis (FA) (Davis 1973; Grande *et al.* 1996) to select proper hydrochemical parameters for fluoride concentration prediction and reduce the dimensionality of our large data sets. Prior to factor analysis, the raw data for the input vector are standardized such that they have zero mean and unit variance in order to deal with different units in the data set.

The starting point of the FA is to generate new groups as factors of variables via linear combination of original variables. The Kaiser criterion is applied to determining a number of factors (Kaiser 1958; Davis 1973). The factors which better describe the variance of the analyzed data (eigenvalue > 1) and can be reasonably interpreted are accepted for further analysis.

We use factor loadings to show the significance of the correlation between factors and variables. Strong positive or negative correlation between a variable and a factor is indicated by high factor loadings close to 1 or –1, respectively (Nadiri *et al.* 2013b). Among different rotation techniques of factor loading matrix, the varimax rotation technique (Kaiser 1958) is adopted in the study. The varimax rotation technique is able to distinguish variables with high factor loadings (close to 1 or –1) from variables with low factor loadings (close to 0).

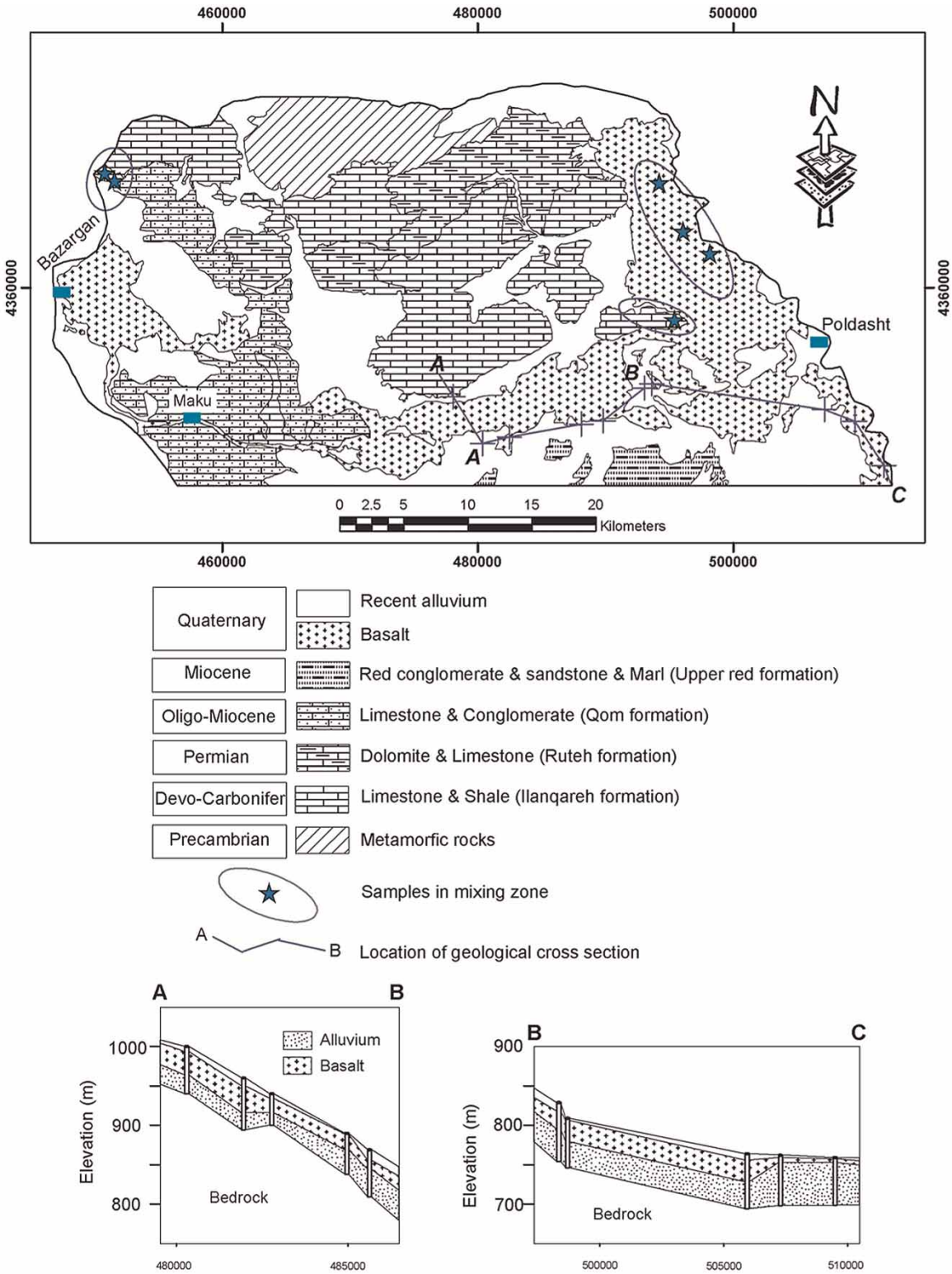

Figure 2 | Geological formations and cross sections in the study area. (Modified from Asghari Moghaddam & Fijani 2009.)

Table 1 | Sample data sets

Data set	Year	Month	No. of samples
1	2004	April	10
2	2004	July	8
3	2004	December	10
4	2006	August	38
5	2006	September	34
6	2007	December	12
7	2008	March	20

Table 2 | Statistical characteristics of hydrochemical parameters

Parameter	Unit	Minimum	Maximum	Mean	Std. deviation
F^-	mg/L	0.05	7.91	2.76	1.73
Ca^{2+}	mg/L	4.81	166.81	73.9	29.01
Mg^{2+}	mg/L	< 0.1	174.96	37.31	22.96
Cl^-	mg/L	< 0.1	432.49	45.24	53.98
CO_3^{2-}	mg/L	< 0.1	240	28.369	42.98
HCO_3^-	mg/L	17.08	902.8	568.91	223.93
Na^+	mg/L	0.7	335.67	147.39	91.5
K^+	mg/L	0.4	18.561	8.6	4.56
SiO_2	mg/L	< 0.1	124.9	50.42	24.62
SO_4^{2-}	mg/L	< 0.1	700.98	108.68	97.34
pH[a]	–	6.5	8.99	7.89	0.52
EC[a]	μS/cm	295	4,500	1,184	507.46

[a]Measured at 25 °C.

According to Dalton & Upchurch (1978), a set of correlated physico-chemical parameters in a particular factor can be explained by a special chemical process (contamination, anthropogenic or genetic), which can lead to variability observed in the data set. In this research, three factors were used to interpret hydrochemical process as they explained 73.1% of data set variance. The results of FA are presented in Table 4. Factors 1 and 3 present the largest and smallest variances, respectively. Factor 1 shows high factor loadings for F^-, Na^+ (positive), K^+ (positive), Ca^{2+} (negative) and HCO_3^-, (positive) and low factor loadings for the remaining parameters. High factor loadings in Factor 2 are CO_3^{2-} (positive), SO_4^{2-} (positive), pH (positive) and SiO_2 (negatively), and in Factor 3 are Na^+ (positive), Cl^- (positive) and EC (positive). The high correlated parameters to F^- in Factor 1 are consistent with the information in the correlation matrix in Table 3. Based on the prior study (Dragon 2006), Factor 1 is a fluoride concentration factor. This factor clearly describes effective hydrochemical parameters that relate to an increase of fluoride concentration in ground water (Asghari Moghaddam & Fijani 2008). From hydrological investigations and the correlation matrix, Factors 2 and 3 show the anthropogenic and genetic (occurring with the dissolution of geological setting as usual) processes, respectively, in the study area. Therefore, the four highly correlated parameters (Na^+, K^+, HCO_3^-, and Ca^{2+}) to fluoride concentration are chosen as input data of the AI models.

Table 3 | Correlation matrix of hydrochemical parameters

Parameters	F^-	Ca^{2+}	Mg^{2+}	Cl^-	CO_3^{2-}	HCO_3^-	Na^+	K^+	SiO_2	SO_4^{2-}	pH	EC
F^-	1											
Ca^{2+}	– 0.68	1										
Mg^{2+}	– 0.35	0.24	1									
Cl^-	0	0.14	0.73	1								
CO_3^{2-}	– 0.2	0.08	– 0.07	– 0.39	1							
HCO_3^-	0.29	– 0.3	0.2	0.39	– 0.8	1						
Na^+	0.64	– 0.77	– 0.14	0.01	– 0.07	0.51	1					
K^+	0.41	– 0.61	– 0.14	– 0.28	0.03	0.3	0.61	1				
SiO_2	0.2	0.05	– 0.07	0.23	– 0.52	0.39	– 0.02	0.06	1			
SO_4^{2-}	– 0.21	0.11	0.18	– 0.03	0.39	– 0.14	0.19	0.13	– 0.25	1		
pH	0.5	– 0.22	0.36	0.68	– 0.29	– 0.41	0.41	0.12	0.2	– 0.02	1	
EC	0.01	– 0.29	0.01	– 0.28	0.56	– 0.09	0.53	0.4	– 0.49	0.53	– 0.13	1

Table 4 | Factor loadings using varimax rotation. Loadings larger than 0.6 is marked by bold face

Parameter	Factor 1	Factor 2	Factor 3
F^-	**0.77**	− 0.29	− 0.03
Na^+	**0.93**	0.17	**0.84**
Ca^{2+}	− **0.86**	0.01	0.08
K^+	**0.78**	0.16	− 0.09
Mg^{2+}	− 0.29	0.22	0.18
CO_3^{2-}	− 0.15	**0.80**	− 0.36
SO_4^{2-}	0.03	**0.74**	0.19
HCO_3^-	**0.60**	− 0.45	0.52
Cl^-	− 0.14	− 0.20	**0.92**
SiO_2	0.08	− **0.67**	0.13
pH	0.36	**0.70**	− 0.18
EC	0.40	− 0.06	**0.82**

For the AI models discussed below in this research, 100 samples were used in the training step and 32 samples were used in the test step. The samples were selected to cover the entire region. Similar statistical characteristics of the training data and test data sets are presented in Table 5. Different types of the input data were normalized to have a range between 0 and 1.

PREDICTIONS BY AI MODELS AND SCMAI

Fuzzy logic

Fuzzy logic is based on the fuzzy set theory which was first introduced by Zadeh (1965). Fuzzy sets, in contrast to their crisp counterparts, include partial membership ranging between 0 and 1. Fuzzy sets have ambiguous boundaries and gradual transitions between the defined sets which render them amenable to overcome the inherent uncertainty in the system and to control the human errors (Kadkhodaie-Ilkhchi & Amini 2009; Pulido-Calvo & Gutiérrez-Estrada 2009; Grande *et al.* 2010). Therefore, fuzzy sets are suitable to describe hydrogeological and hydrochemical parameters (such as fluoride concentration) that are inherently imprecise.

The FL method is an effective tool for hydrochemical parameter prediction, e.g. traces, minor and major ion concentration prediction. Each fuzzy set is represented by a membership function (MF). The membership functions may have different shapes, such as Gaussian, triangular, trapezoidal, sigmoid, etc. An FL method consists of three main parts, fuzzification, inference engine (fuzzy rule base), and defuzzification, shown in Figure 3. In fuzzification, the four crisp inputs change to fuzzy set and the fuzzified results are obtained through the inference engine. The fuzzified results are aggregated and defuzzified to obtain the final results. Based on the type of the output membership function, FL may be constructed by the methods proposed by Mamdani and Sugeno. In the MFL method, the output membership functions are fuzzy sets. After the aggregation process, there is a fuzzy set for each output variable that needs defuzzification (Mamdani & Assilian 1975; Mamdani 1976, 1977). The most appropriate clustering method compatible with the MFL is the fuzzy c-means (FCM) clustering method (Bezdec 1981; Newton *et al.* 1992; Lee 2004). The FCM method categorizes data points that populate multidimensional spaces into a specific number of clusters. The FCM clustering starts with an initial guess for the cluster centers, which are intended to mark the mean location of each cluster. The initial guess for these cluster centers is most likely

Table 5 | Statistical characteristics of hydrochemical parameters in the training and test steps

		Training				Test			
Param.	Unit	Minimum	Maximum	Mean	Std. deviation	Minimum	Maximum	Mean	Std. deviation
Ca^{2+}	meq/L	0.56	6.17	3.61	1.48	0.28	6.41	3.37	1.44
HCO_3^-	meq/L	1.64	15.6	9.63	3.36	2.92	14.35	8.22	3.95
Na^+	meq/L	0.06	20.45	6.72	4.07	0.03	14.59	5.94	4.69
K^+	meq/L	0.03	0.47	0.23	0.11	0.02	0.45	0.2	0.14
F^-	meq/L	0.003	0.37	0.15	0.09	0.013	0.34	0.13	0.09

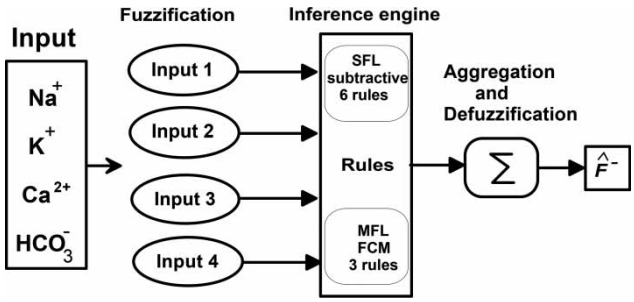

Figure 3 | Schematic diagram of the FL model.

incorrect. Additionally, FCM assigns every data point a membership grade for each cluster. By iteratively updating the cluster centers and the membership grades for each data point, FCM iteratively moves the cluster centers to the correct location within a data set. During this iteration, an objective function is minimized and shows the distance from any given data point to a cluster center weighted by the membership grade of that data point. The FCM output is a list of cluster centers and several membership grades for each data point. Therefore, the MFL fuzzy rules are extracted through the FCM. In this fashion, the model matrices of data are passed through the FCM algorithm and the cluster centers are calculated. In the FCM algorithm, the number of clusters is defined by the user. Choosing the optimum number of clusters is accomplished by measuring the performance of the model through systematically changing the number of the clusters from 1 to the number of the model data points. A Gaussian membership function $\mu_i(x)$ for data x is fitted to the extracted ith input and output cluster as follows (Kadkhodaie-Ilkhchi *et al.* 2009):

$$\mu_i(x) = \exp\left(-\frac{(c_i - x)^2}{2\sigma_i^2}\right) \tag{1}$$

where c_i and σ_i are the mean and the standard deviation for ith cluster, respectively. For F^- prediction, a fuzzy if–then rule i can be also expressed as:

$$\begin{aligned}&\text{Rule } i\text{: If}\left(\mathrm{Na}^+ \text{ belongs to } \mathrm{MF}^i_{\mathrm{Na}^+}\right) \text{ and } \left(\mathrm{K}^+ \text{ belongs to } \mathrm{MF}^i_{\mathrm{K}^+}\right)\\ &\text{and } \left(\mathrm{Ca}^{2+} \text{ belongs to } \mathrm{MF}^i_{\mathrm{Ca}^{2+}}\right), \text{and}\\ &\left(\mathrm{HCO}_3^- \text{ belongs to } \mathrm{MF}^i_{\mathrm{HCO}_3^-}\right), \text{ then } \left(\hat{\mathrm{F}}^- \text{ belongs to } \mathrm{MF}^i_{\mathrm{F}^-}\right)\end{aligned} \tag{2}$$

where $\hat{F}^-$ is the predicted fluoride concentration, $MF^i_{F^-}$ is the membership function for the output of rule i, $MF^i_{Na^+}$ is the membership function of the ith cluster of input Na^+, $MF^i_{K^+}$ is the membership function of the ith cluster of input K^+, and so forth. The operator among the input membership function is 'and' (minimize) operator and the outputs from the rules are aggregated via the 'or' (maximize) operator. The most popular defuzzification method, centroid calculation, was applied to produce the crisp output.

The SFL method is similar to the MFL method in many aspects. The major difference between them is that the output membership functions are either linear or constant for SFL (Sugeno 1985; Takagi & Sugeno 1985).

The most capable clustering algorithm in the SFL method is the subtractive clustering method, which can be used for extraction of the clusters and fuzzy if–then rules (Chiu 1994). The details of the subtractive clustering can be found in the work of Chiu (1994), Chen & Wang (1999), and Li *et al.* (2000). The important parameter in subtractive clustering, which controls the number of clusters and the fuzzy if–then rules, is the cluster radius. Decreasing the cluster radius will increase the number of clusters and lead to smaller clusters. This will create more rules and complicate the system behavior and may lead to a low performance of the model. In contrast, a large cluster radius produces large clusters in the data and results in few rules (Chiu 1994), which may not be sufficient to cover the entire domain.

Searching for the optimal cluster radius can be accomplished by systematically varying cluster radius value from 0 to 1 until minimal root mean squared error (RMSE) is met.

As we use the subtractive clustering method, the number of rules is the same as the number of clusters. For F^- prediction in this study, a fuzzy if–then rule i can be expressed as:

$$\begin{aligned}&\text{Rule } i\text{: If}\left(\mathrm{Na}^+ \text{ belongs to } \mathrm{MF}^i_{\mathrm{Na}^+}\right) \text{ and } \left(\mathrm{K}^+ \text{ belongs to } \mathrm{MF}^i_{\mathrm{K}^+}\right)\\ &\text{and } \left(\mathrm{Ca}^{2+} \text{ belongs to } \mathrm{MF}^i_{\mathrm{Ca}^{2+}}\right),\\ &\text{and } \left(\mathrm{HCO}_3^- \text{ belongs to } \mathrm{MF}^i_{\mathrm{HCO}_3^-}\right),\\ &\text{then } \hat{\mathrm{F}}^- = m_i\mathrm{Na}^+ + n_i\mathrm{K}^+ + p_i\mathrm{Ca}^{2+} + q_i\mathrm{HCO}_3^- + c_i\end{aligned} \tag{3}$$

where m_i, p_i, q_i and c_i are coefficients. The final output is the weighted average of all outputs (aggregation) as

follows:

$$\hat{F}^{-} = \frac{\sum_i w_i \hat{F}_i^{-}}{\sum_i w_i} \tag{4}$$

where w_i is the firing strength of rule i, which is obtained via 'and' (minimize) operator.

Artificial neural network

Artificial neural networks imitate the human brain by using mathematical method. This method has proven to be extremely beneficial for prediction and forecasting of water quality modeling, contaminant ion concentration such as nitrate and ammonia in the hydrological system (Sharma *et al.* 2003; Gutiérrez-Estrada *et al.* 2004; Almasri & Kaluarachchi 2005; Dixon 2005). The most widely used artificial neural network method is the multi-layer perceptron (MLP) (Hornik *et al.* 1989; Haykin 1998; Daliakopoulos *et al.* 2005; Nourani *et al.* 2008) that consists of one input layer, hidden layers, and one output layer. This study considers one hidden layer for the MLP network shown in Figure 4. To apply the ANN to the Maku case study, there are four neurons in the input layer corresponding to the input data $I_i = \left(Na^+, K^+, Ca^{2+}, HCO_3^-\right)$, three neurons in the hidden layer which was found sufficient after trying different numbers, and one neuron in the output layer. The normalized input signal propagates through the network in a forward direction via connections between neurons. Incoming signals are linearly combined and converted to outgoing signals. The signal conversion is carried out by assigned activation functions. The mathematical expression of a three-layer feedforward ANN is

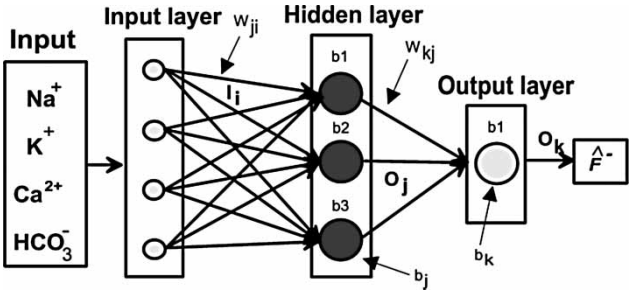

Figure 4 | Feedforward ANN model.

(ASCE 2000a, b):

$$O_j = f_1\left(b_j + \sum_i W_{ji} I_i\right) \tag{5}$$

$$O_k = \hat{F}^{-} = f_2\left(b_k + \sum_j W_{kj} O_j\right) \tag{6}$$

where f_1 and f_2 are the activation functions for the hidden layer and output layer, respectively, I_i is the ith input, O_j is the jth output, W_{ji} and W_{kj} are the weights that control the strength of connections between two layers, and the biases b_j and b_k are used to adjust the mean value for input layer and hidden layer, respectively. The activation function for the hidden layer is typically a continuous and bounded nonlinear transfer function such as sigmoid and log sigmoid functions. The activation function for the output layer is usually a linear function. In this study, we select hyperbolic tangent sigmoid (Tansig) for f_1 and linear (Purelin) for f_2. The output O_k of the ANN is predicted fluoride concentration $\hat{F}^{-}$. In the ANN training step, we use the Levenberg–Marquardt (LM) algorithm as a supervised learning algorithm to estimate the weights (W_{ij} and W_{kj}) and the biases (ASCE 2000a, b). Several previous research studies have found the performance of the LM method to be superior to other training algorithms such as the conjugate gradient (CG), Bayesian regularization (BR), and gradient descent with momentum (GDX) algorithms (Daliakopoulos *et al.* 2005; Nourani *et al.* 2008).

Neuro-fuzzy

The NF modeling is a combination technique for describing the behavior of a system using fuzzy inference rules within a neural network structure. The NF inference system consists of a given input/output data set and an SFL model, for which the membership function parameters are optimized using a hybrid algorithm (Kaynak *et al.* 1998; Jacks *et al.* 2005). For this study, an NF architecture of a five-layer MLP network is considered for fluoride concentration prediction, shown in Figure 5. This MLP network optimizes the aforementioned SFL model. In each layer, the NF operations are as follows.

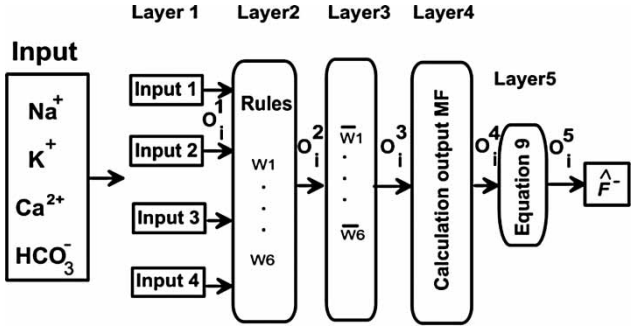

Figure 5 | Schematic architecture of the NF model.

Layer 1: Generate membership function of input data. The output of neuron i is defined by:

$$O_i^1 = \mu_{ji}(X) \tag{7}$$

where j is the number of input and i is the membership function index. $X = \left\{\mathrm{Na}^+, \mathrm{K}^+, \mathrm{Ca}^{2+}, \mathrm{HCO}_3^-\right\}$ is a set of input. $\mu_{ji}(X)$ is a fuzzy set associated with neuron i given a membership function. Again, we use a generalized Gaussian function to develop membership functions.

Layer 2: Calculate firing strength w_i for the ith rule via multiplication:

$$O_i^2 = w_i = \mu_{1i}(X)\,\mu_{2i}(X)\,\mu_{3i}(X)\,\mu_{4i}(X) \tag{8}$$

Layer 3: Compute the normalized firing strengths for the ith neuron:

$$O_i^3 = \overline{w}_i = \frac{w_i}{\sum_i w_i} \tag{9}$$

Layer 4: Compute the contribution of the ith rule in the model output based on the first-order SFL method:

$$O_i^4 = \overline{w}_i \hat{\mathrm{F}}_i^- = \overline{w}_i\left(m_i\mathrm{Na}^+ + n_i\mathrm{K}^+ + p_i\mathrm{Ca}^{2+} + q_i\mathrm{HCO}_3^- + c_i\right) \tag{10}$$

Layer 5: Calculate the final output as the weighted average of all rule outputs (aggregation):

$$O^5 = \hat{\mathrm{F}}^- = \sum_i \overline{w}_i \hat{\mathrm{F}}_i^- \tag{11}$$

The NF parameters in Equation (10) and membership function parameters are estimated using a hybrid algorithm in this study, which is a combination of the gradient descent and least-squares methods (Aqil *et al.* 2007).

SCMAI model

The committee machine with artificial intelligence (CMAI) approach combines the aforementioned artificial intelligence model outputs to reap advantage of all AI models to produce the final output. Previous works recommend two methods for construction of a CMAI (Chen & Lin 2006; Kadkhodaie-Ilkhchi *et al.* 2009; Labani *et al.* 2010): simple averaging and weighted averaging.

In this study, we introduce a SCMAI model that employs an artificial neural network model as a supervised combiner of all AI models to replace simple averaging or weighted averaging. The SCMAI model consists of four AI models, shown in Figure 6, and includes two major steps. In the first step, fluoride concentration is predicted from hydrochemical data using aforementioned AI models (MFL, SFL, ANN, and NF). In the second step, a supervised artificial neural network is constructed as a nonlinear, supervised combiner. The mathematical expression of the SCMAI model is:

$$\hat{\mathrm{F}}_i^- = \mathrm{AI}_i\left(\mathrm{Na}^+, \mathrm{K}^+, \mathrm{Ca}^{2+}, \mathrm{HCO}_3^-\right) \tag{12}$$

$$O_j = f_1\left(b_j + \sum_i W_{ji}\hat{\mathrm{F}}_i^-\right) \tag{13}$$

$$O_k = \hat{\mathrm{F}}_{\mathrm{SCMAI}}^- = f_2\left(b_k + \sum_j W_{kj}O_j\right) \tag{14}$$

where $\hat{\mathrm{F}}_i^-$ is the output of the each AI model which has been used as ith input, f_1 and f_2 are activation functions for the hidden layer and output layer, respectively, O_j is the jth output of nodes in hidden layer, W_{ji} and W_{kj} are weights that

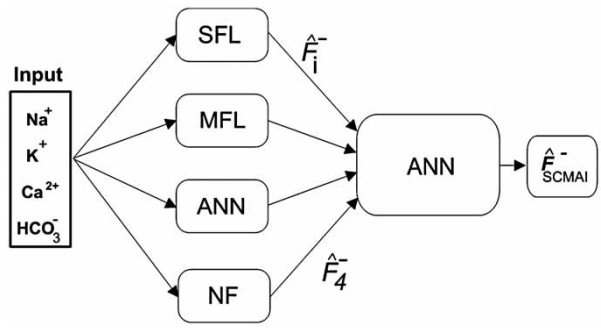

Figure 6 | The framework of the SCMAI model.

control the strength of connections between two layers, and the biases b_j and b_k are used to adjust the mean value for hidden layer and output layer, respectively. The activation functions are hyperbolic tangent sigmoid (Tansig) for f_1 and linear (Purelin) for f_2. The output O_k of the SCMAI model is $\hat{F}^-_{SCMAI}$. In the ANN training step, the LM algorithm was adopted as a learning algorithm to estimate the weights (W_{ji} and W_{kj}) and the biases (ASCE 2000a, b).

Effectiveness evaluation

Two different criteria are used to evaluate the effectiveness of the ANN and NF models and their ability to make precise predictions. One criterion is the root mean square error (RMSE) (Hyndman & Koehler 2006):

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(F_i^- - \hat{F}_i^-\right)^2} \tag{15}$$

where F_i^- and $\hat{F}_i^-$ are the measured and predicted fluoride concentration, respectively. n is the number of data. RMSE measures the discrepancy between the data and prediction. The other criterion is the coefficient of determination (R^2) (Nash & Sutcliffe 1970):

$$R^2 = 1 - \frac{\sum_i^n \left(F_i^- - \hat{F}_i^-\right)^2}{\sum_i^n \left(F_i^- - \overline{F}_i^-\right)^2} \tag{16}$$

where $\overline{F}^-$ is the mean of the data. R^2 measures the closeness between the data and prediction. Perfect fitting will result in RMSE = 0 and $R^2 = 1$.

RESULTS AND DISCUSSION

Artificial intelligence models

FL model

In this study, two different FL models (MFL and SFL) are adopted for predicting the fluoride concentration. The FCM is used to develop fuzzy rules for the MFL model and the subtractive clustering method is used to develop fuzzy rules for the SFL model (Li *et al.* 2000).

For the SFL model, the cluster radius is determined by minimizing the RMSE. Through a systematic search, the cluster radius is 0.6 and the minimal RMSE is 0.040 meq/L. This cluster radius generates six fuzzy if–then rules and establishes six Gaussian membership functions. In each rule, the parameters m_i, n_i, p_i, q_i, z_i and c_i in the output membership function were estimated by minimizing the RMSE. The results show that the coefficient of determination (R^2) and RMSE in the training step were 0.78 and 0.04 (meq/L), respectively. The R^2 and RMSE in the test step were 0.76 and 0.04 (meq/L), respectively. Figure 7(a) shows a scatter plot of measured F^- versus predicted F^- for the training data. The closer the data to the 45° solid line, the better the accuracy of the model. The scatter plot for test data is shown in Figure 7(b). We note that the model error for predicting fluoride concentration at the mixing zone (filled squares) is significant due to its complex hydrogeological characteristics.

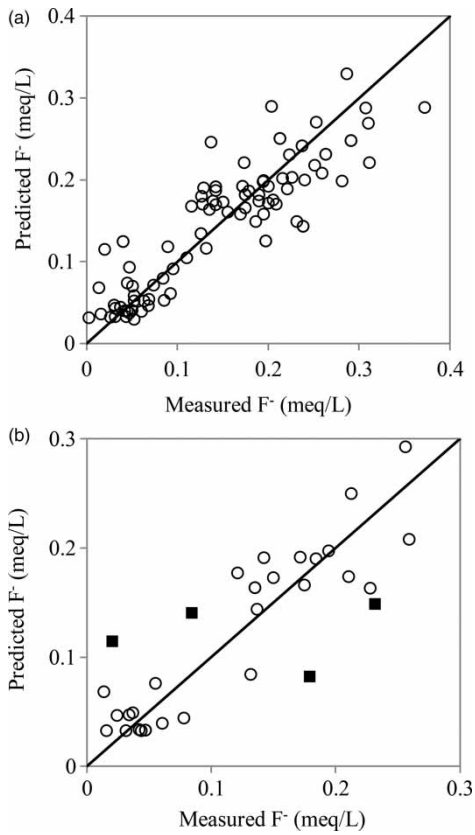

Figure 7 | Comparison of the measured fluoride concentration to predicted fluoride concentration by SFL method in (a) training step and (b) test step. The filled squares are the samples in the mixing zone.

Using the FCM clustering method for the MFL model, the minimal RMSE is 0.05 meq/L with three clusters in the training step. The R^2 is 0.70. Therefore, our MFL model generates three rules to link input to output via 'and' operator. The output from the rules are aggregated via the 'or' (maximize) operator. Then, we used a defuzzification method, centroid calculation, to produce crisp output.

The fitting performance of the test step is RMSE = 0.05 and $R^2 = 0.66$, which are slightly worse than the SFL model. Figure 8 shows the comparison between the measured and predicted fluoride concentrations for the training and test data sets. Again, the MFL model does not predict well for the fluoride samples in the mixing zone.

Artificial neural network

An MLP network is adopted to predict fluoride concentration. The MLP structure is shown in Figure 4. The transfer function for the hidden layer is Tansig and for the output layer is Purelin. The ANN model is trained by minimizing RMSE in the training step. The LM optimization method was applied to optimizing ANN weights and biases values.

After 85 epochs of training, the minimum RMSE = 0.02 meq/L was reached. The R^2 was 0.88. In the testing step, the RMSE and R^2 were 0.03 meq/L and 0.82, respectively, which show better performance than the FL models. Comparisons between the measured and predicted F^- values in the training and test steps are shown in Figure 9. The ANN model predicts the fluoride concentration at the mixing zone better than the FL models.

Neuro-fuzzy

In this study, an NF model is developed for prediction of fluoride concentration. The input and output data of NF model

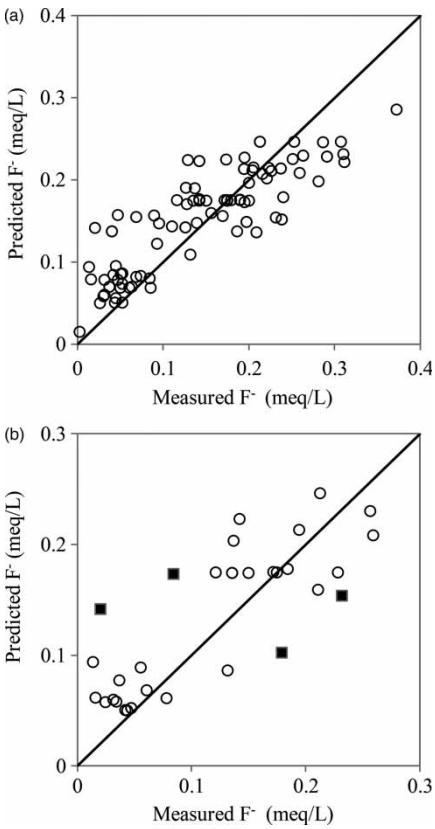

Figure 8 | Comparison of the measured fluoride concentration to predicted fluoride concentration by MFL method in (a) training step and (b) test step. The filled squares are the sample in the mixing zone.

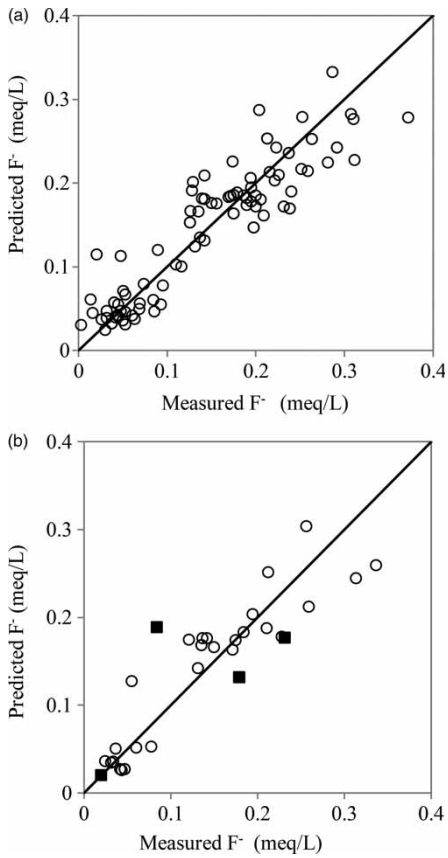

Figure 9 | Comparison of the measured fluoride concentration to predicted fluoride concentration by the ANN model in (a) training step and (b) test step. The filled squares are the samples in the mixing zone.

are classified similar to the aforementioned SFL model. The same clusters of the input and output, and the same number of rules are used for NF construction. A hybrid algorithm is applied for optimizing parameters in the Gaussian membership functions and the coefficients in the output linear equations (Zounemat-Kermani & Teshnehlab 2008).

After three epochs of training, the minimum RMSE = 0.03 meq/L is reached. The R^2 is 0.83. In the test step, the trained NF model obtains RMSE = 0.02 meq/L and R^2 = 0.81. The scatter plots for the measured fluoride concentration and predicted value are shown in Figure 10. By taking advantage of SFL and ANN with more model parameters, NF performed slightly better fitting than the FL models in both the training and test steps. Again, significant model prediction error exists in the mixing zone.

The performance results using individual AI models are listed in Table 6. It concludes that the ANN model is the best model and the MFL model is the worst with respect to the other three AI models for the study area. However, all four AI models are applicable for prediction of fluoride concentration since their performance difference is not significant. Their fluoride predictions at the mixing zone are equally poor. To achieve the optimal performance and reap the benefits of all work, in the following, we conduct the SCMAI and CMAI that utilize the advantages of each AI model.

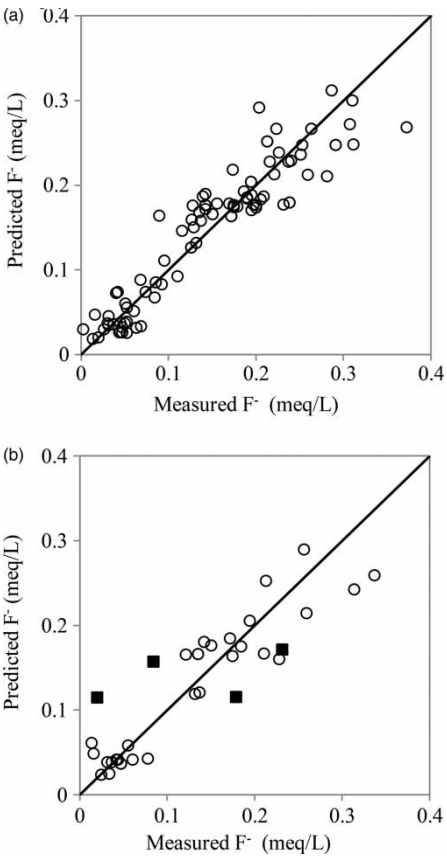

Figure 10 | Comparison of the measured fluoride concentration to predicted fluoride concentration by NF model in (a) training step and (b) test step. The filled squares are the samples in the mixing zone.

Table 6 | Fitting performance of the AI models in the training and test steps

		Model			
Step	**Criteria**	**SFL**	**MFL**	**ANN**	**NF**
Training	RMSE (meq/L)	0.04	0.05	0.02	0.03
	R^2	0.78	0.70	0.88	0.83
Test	RMSE (meq/L)	0.04	0.05	0.03	0.04
	R^2	0.76	0.66	0.82	0.81

SCMAI model

The SCMAI method shown in Figure 6 adopts a simple ANN method to re-estimate fluoride concentration obtained by SFL, MFL, ANN, and NF in the training step (100 sample data). The MLP-LM structure based on Equations (13) and (14) is employed in the ANN model to be used in the SCMAI. This ANN model has four neurons in the input layer, three neurons in the hidden layer, and one neuron in the output layer. The transfer function for the hidden layer is Tansig and for the output layer is Purelin.

The LM algorithm was used to optimize 15 weights and four biases in the ANN. After 52 epochs, the RMSE = 0.005 meq/L is obtained. Then, the SCMAI model is tested by the 31 data points in the test step. The RMSE and R^2 for SCMAI predictions are 0.01 (meq/L) and 0.97, respectively. Comparing to Table 6, SCMAI outperforms individual AI models with much smaller fitting errors. The improved fitting, especially for predictions in the mixing zone (filled squared), is shown in Figure 11 and Table 7. These results support the importance of a nonlinear combination of AI models and the effectiveness of using the SCMAI for predicting the fluoride concentration value in the highly heterogeneous unconfined aquifer underneath the Maku area.

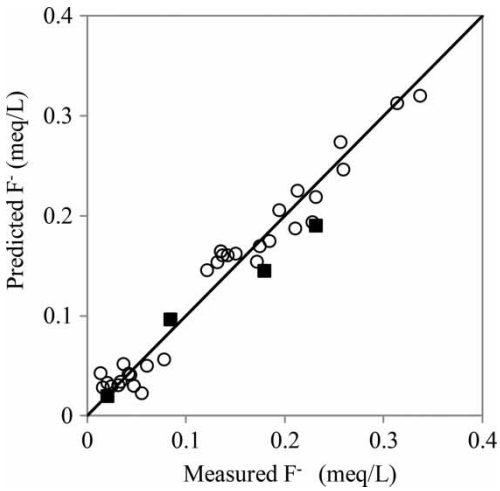

Figure 11 | Comparison of the measured fluoride concentration to predicted fluoride concentration by the SCMAI model in the test step. The filled squares are the samples in the mixing zone.

Comparison of SCMAI to CMAI for predicting the fluoride concentration

Here, we compare the SCMAI model to the CMAI model. For the simple averaging method, the CMAI estimates the fluoride concentration by averaging the predictions from four AI models (SFL, MFL, ANN, and NF) with equal weights:

$$\hat{F}^{-}_{CMAI} = 0.25\hat{F}^{-}_{SFL} + 0.25\hat{F}^{-}_{MFL} + 0.25\hat{F}^{-}_{ANN} + 025\hat{F}^{-}_{NF} \quad (17)$$

For the weighted averaging method, the optimized weights, w_i are determined by minimizing the mean squared error (MSE):

$$\mathrm{MSE} = \frac{1}{m}\sum_{i=1}^{m}\left(w_1\hat{F}^{-}_{i,SFL} + w_2\hat{F}^{-}_{i,MFL} + w_3\hat{F}^{-}_{i,ANN} + w_4\hat{F}^{-}_{i,NF} - F^{-}_{i}\right)^2 \quad (18)$$

where m is the number of training data (100 samples). The weights, w_i, range between 1 and 0. Sum of weights is unity, $\sum w_i = 1$.

A GA optimizer in MATLAB® is used to minimize the MSE. The initial population size is set to 30. The maximum number of generations is set to 140. The probability for crossover operation is set to 0.8. The mutation function is Gaussian. The parameters that control the mutation are specified as the scale value of 1 and shrink value of 1. After the optimal weights are obtained, the CMAI model predicts the fluoride concentration by:

$$\hat{F}^{-}_{CMAI} = 0.22\hat{F}^{-}_{SFL} + 0.19\hat{F}^{-}_{MFL} + 0.31\hat{F}^{-}_{ANN} + 0.28\hat{F}^{-}_{NF} \quad (19)$$

The order of the weights in Equation (19) is consistent with the ranking of the AI models in Table 6. The ANN model receives the highest weight, 0.31, and the MFL model receives the lowest weight, 0.19. However, the best model does not have a dominant weight, which supports the incentive of developing the CMAI.

The performance result of the SCMAI and CMAI is shown in Table 8 for fluoride data in the test step. The R^2 and RMSE for CMAI with simple averaging are 0.84 and 0.03, respectively, and for CMAI with weighed averaging are 0.93 and 0.02, respectively. Table 8 concludes that CMAI and SCMAI perform better than individual AI models. CMAI with weighted averaging performs better than CMAI with simple averaging, which agrees with Kadkhodaie-Ilkhchi *et al.* (2009) and Labani *et al.* (2010). The SCMAI model which takes advantage of the nonlinear combination of the AI models performs better than the linear combination in the CMAI for prediction of fluoride concentration in the Maku area.

Table 7 | Comparison of predicted fluoride concentration in the test step by different AI models and SCMAI model against the measured data in the mixing zone

Location (UTM, meters)			Predicted fluoride concentration (meq/L)				
X	Y	Measured fluoride concentration (meq/L)	SFL	MFL	ANN	NF	SCMAI
494383	4368257	0.0202	0.1149	0.1419	0.0206	0.1149	0.0199
496248	4364458	0.0842	0.1409	0.1738	0.1891	0.1573	0.0963
494770	4357357	0.1789	0.0824	0.1023	0.1320	0.1155	0.1452
449683	4369665	0.2316	0.1491	0.1540	0.1771	0.1720	0.1904

Table 8 | Fitting performance comparison of SCMAI with CMAI in the test step

		Model		
Step	Criteria	SCMAI	CMAI with simple averaging	CMAI with weighted averaging
Training	RMSE (meq/L)	0.005	0.023	0.013
	R^2	0.99	0.92	0.98
Test	RMSE (meq/L)	0.011	0.032	0.023
	R^2	0.97	0.84	0.93

CONCLUSIONS

The study presents a supervised committee machine with artificial intelligence (SCMAI) method to better predict the fluoride concentration in the basaltic and non-basaltic aquifers underneath the Maku area, Iran. For this purpose, four artificial intelligence models, Sugeno fuzzy logic (SFL), Mamdani fuzzy logic (MFL), artificial neural network (ANN), and neuro-fuzzy (NF), are adopted to construct the SCMAI model. The findings and future work are summarized as follows:

1. The four AI models (SFL, MFL, ANN, and NF) are applicable for prediction of fluoride concentration in the complex, heterogeneous and unconfined aquifer underlying the Maku area. The ANN and NF model show better predictions with respect to the measured fluoride concentration than the FL models. However, because of complex hydrogeological characteristics in the mixing zone, all AI models do not predict well for the fluoride concentration.
2. The study shows that the prediction of fluoride concentration prediction can be improved by linearly combining the outputs of the AI models. The CMAI model predicts better than individual AI models using either equal weights or optimized weights. This study also shows that using optimized weights performs better than equal weights in the CMAI method.
3. The SCMAI method improves the CMAI model by replacing the linear model output combination with a nonlinear model output combination through an artificial intelligence system. Nonlinear combination is particularly imperative for predicting the fluoride concentration in complex aquifer environments. This study shows that the SCMAI has a better prediction result than the CMAI and outperforms individual AI models, especially in the mixing zone.
4. Since most real aquifer systems are heterogeneous and complex, the presented SCMAI method has a potential to estimate other hydrochemical or hydrogeological parameters. Moreover, uncertainty analysis using the SCMAI model is a crucial subject for future research.

ACKNOWLEDGEMENTS

Ata Allah Nadiri and Elham Fijani were supported by the Iran Ministry of Science, Research and Technology and Iran's National Elites Foundation, to conduct research at Louisiana State University. Frank Tsai was supported by Grant/Cooperative Agreement Number G10AP00136 from the United States Geological Survey to conduct multimodel study. The contents in the publication are solely the responsibility of the authors and do not necessarily represent the official views of the USGS.

REFERENCES

Almasri, M. N. & Kaluarachchi, J. J. 2005 Modular neural networks to predict the nitrate distribution in ground water using the on-ground nitrogen loading and recharge data. *Environ. Model. Softw.* **20** (7), 851–871.

Anifowose, F. & Abdulraheem, A. 2011 Fuzzy logic-driven and SVM-driven hybrid computational intelligence models applied to oil and gas reservoir characterization. *J. Nat. Gas Sci. Eng.* **3** (3), 505–517.

APHA 1998 *Standard Method for the Examination of Water and Wastewater*. 17th edn, American Public Health Association, Washington, DC.

Aqil, M., Kita, I., Yano, A. & Nishiyama, S. 2007 Analysis and prediction of flow from local source in a river basin using a neuro-fuzzy modeling tool. *J. Environ. Manage.* **85** (1), 215–223.

ASCE 2000a Task Committee on Application of Artificial Neural Networks in Hydrology, Artificial neural network in hydrology. I: preliminary concepts. *J. Hydro. Eng.* **5** (2), 115–123.

ASCE 2000b Task Committee on Application of Artificial Neural Networks in Hydrology, Artificial neural network in hydrology. II: Hydrologic applications. *J. Hydro. Eng.* **5** (2), 124–137.

Asghari Moghaddam, A. & Fijani, E. 2008 Distribution of fluoride in groundwater of Maku area, northwest of Iran. *Environ. Geol.* **56** (2), 281–287.

Asghari Moghaddam, A. & Fijani, E. 2009 Hydrogeologic framework of the Maku area basalts, northwestern Iran. *Hydrogeol. J.* **17** (4), 949–959.

Asghari Moghaddam, A., Nadiri, A. A. & Fijani, E. 2010 Spatial prediction of fluoride concentration using artificial neural networks and geostatistics models. *Water Soil Sci. (Agric. Sci.)* **19.1** (2), 129–145.

Bárdossy, A. & Disse, M. 1993 Fuzzy rule-based models for infiltration. *Water Resour. Res.* **29** (2), 373–382.

Bårdsen, A., Bjorvatn, K. & Selvig, K. A. 1996 Variability in fluoride content of subsurface water reservoirs. *Acta Odontol. Scand.* **54** (6), 343–347.

Bezdec, J. C. 1981 *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York.

Chang, F.-J. & Chang, Y.-T. 2006 Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. *Adv. Water Resour.* **29** (1), 1–10.

Chen, C.-H. & Lin, Z.-S. 2006 A committee machine with empirical formulas for permeability prediction. *Comput. Geosci.* **32** (4), 485–496.

Chen, M.-S. & Wang, S.-W. 1999 Fuzzy clustering analysis for optimizing fuzzy membership functions. *Fuzzy Sets Syst.* **103** (2), 239–254.

Chiu, S. L. 1994 Fuzzy model identification based on cluster estimation. *J. Intell. Fuzzy Syst.* **2**, 267–278.

Daliakopoulos, I. N., Coulibaly, P. & Tsanis, I. K. 2005 Groundwater level forecasting using artificial neural networks. *J. Hydrol.* **309** (1–4), 229–240.

Dalton, M. G. & Upchurch, S. B. 1978 Interpretation of hydrochemical facies by factor analysis. *Ground Water* **16** (4), 228–233.

Davis, J. C. 1973 *Statistics and Data Analysis in Geology*. John Wiley and Sons, New York, 550 pp.

Dissanayake, C. B. 1991 The fluoride problem in the groundwater of Sri Lanka – environmental management and health. *Int. J. Environ. Stud.* **38** (2–3), 137–155.

Dixon, B. 2005 Applicability of neuro-fuzzy techniques in predicting ground-water vulnerability: a GIS-based sensitivity analysis. *J. Hydrol.* **309** (1–4), 17–38.

Dragon, K. 2006 Application of factor analysis to study contamination of a semi-confined aquifer (Wielkopolska Buried Valley aquifer, Poland). *J. Hydrol.* **331** (1–2), 272–279.

Garcia, L. A. & Shigidi, A. 2006 Using neural networks for parameter estimation in ground water. *J. Hydrol.* **318** (1–4), 215–231.

Gizaw, B. 1996 The origin of high bicarbonate and fluoride concentrations in waters of the Main Ethiopian Rift Valley, East African Rift system. *J. Afr. Earth Sci.* **22** (4), 391–402.

Govindaraju, R. S. & Rao, A. R. 2000 *Artificial Neural Networks in Hydrology*. Springer Publication, p. 384.

Grande, J. A., Andújar, J. M., Aroba, J., Beltrán, R., de la Torre, M. L., Cerón, J. C. & Gómez, T. 2010 Fuzzy modeling of the spatial evolution of the chemistry in the Tinto River (SW Spain). *Water Resour. Manage.* **24** (12), 3219–3235.

Grande, J. A., González, A., Beltrán, R. & Sánchez-Rodas, D. 1996 Application of factor analysis to the study of contamination in the aquifer system of Ayamonte-Huelva (Spain). *Ground Water* **34** (1), 155–161.

Guo, Q., Wang, Y., Ma, T. & Ma, R. 2007 Geochemical processes controlling the elevated fluoride concentrations in groundwaters of the Taiyuan Basin, Northern China. *J. Geochem. Explor.* **93** (1), 1–12.

Gupta, S. K., Deshpande, R. D., Agarwal, M. & Raval, B. R. 2005 Origin of high fluoride in groundwater in the North Gujarat-Cambay region, India. *Hydrogeol. J.* **13** (4), 596–605.

Gutiérrez-Estrada, J. C., de Pedro-Sanz, E., López-Luque, R. & Pulido-Calvo, I. 2004 Comparison between traditional methods and artificial neural networks for ammonia concentration forecasting in an eel (Anguilla anguilla L.) intensive rearing system. *Aquacult. Eng.* **31**, 183–203.

Handa, B. K. 1975 Geochemistry and genesis of fluoride-containing ground waters in India. *Ground Water* **13** (3), 275–281.

Haykin, S. S. 1998 *Neural Networks: A Comprehensive Foundation*. Prentice Hall, New Jersey, p. 842.

Hornik, K., Stinchcombe, M. & White, H. 1989 Multilayer feedforward networks are universal approximators. *Neural Netw.* **2** (5), 359–366.

Hudak, P. F. & Sanmanee, S. 2003 Spatial patterns of nitrate, chloride, sulfate, and fluoride concentrations in the Woodbine Aquifer of North-Central Texas. *Environ. Monit. Assess.* **82** (3), 311–320.

Hyndman, R. J. & Koehler, A. B. 2006 Another look at measures of forecast accuracy. *Int. J. Forecast.* **22** (4), 679–688.

Jacks, G., Bhattacharya, P., Chaudhary, V. & Singh, K. P. 2005 Controls on the genesis of some high-fluoride groundwaters in India. *Appl. Geochem.* **20** (2), 221–228.

Kadkhodaie-Ilkhchi, A. & Amini, A. 2009 A fuzzy logic approach to estimating hydraulic flow units from well log data: a case study from the Ahwaz oilfield, South Iran. *J. Petrol. Geol.* **32** (1), 67–78.

Kadkhodaie-Ilkhchi, A., Rahimpour-Bonab, H. & Rezaee, M. 2009 A committee machine with intelligent systems for estimation of total organic carbon content from petrophysical data: an example from Kangan and Dalan reservoirs in South Pars Gas Field, Iran. *Comput. Geosci.* **35** (3), 459–474.

Kaiser, H. F. 1958 The Varimax criterion for analytic rotation in factor analysis. *Psychometrika* **23** (3), 187–200.

Kaynak, O., Zadeh, L. A., Türksen, B. & Rudas, I. J. 1998 *Computational Intelligence: Soft Computing and Fuzzy-Neuro Integration with Applications (Nato ASI Series (closed))*. Springer, Berlin, p. 547.

Kharb, P. & Susheela, A. K. 1994 Fluoride ingestion in excess and its effects on organic and certain inorganic constituents of soft tissues. *Med. Res. Soc.* **22**, 43–44.

Labani, M. M., Kadkhodaie-Ilkhchi, A. & Salahshoor, K. 2010 Estimation of NMR log parameters from conventional well log data using a committee machine with intelligent systems: a case study from the Iranian part of the South Pars gas field, Persian Gulf Basin. *J. Petrol. Sci. Eng.* **72** (1–2), 175–185.

Lee, K. H. 2004 *First Course on Fuzzy Theory and Applications*. Springer, Berlin, p. 335.

Li, H., Chen, C. L. P. & Huang, H.-P. 2000 *Fuzzy Neural Intelligent Systems: Mathematical Foundation and the Application in Engineering*. CRC Press LLC, Boca Raton, FL, p. 388.

Lim, J.-S. 2005 Reservoir properties determination using fuzzy logic and neural networks from well data in offshore Korea. *J. Petrol. Sci. Eng.* **49** (3–4), 182–192.

Mamdani, E. H. 1976 Advances in the linguistic synthesis of fuzzy controllers. *Int. J. Man-Machine Stud.* **8** (6), 669–678.

Mamdani, E. H. 1977 Application of fuzzy logic to approximate reasoning using linguistic synthesis. *IEEE Trans. Comput.* **C-26** (12), 1182–1191.

Mamdani, E. H. & Assilian, S. 1975 An experiment in linguistic synthesis with a fuzzy logic controller. *Int. J. Man-Machine Stud.* **7** (1), 1–13.

Mark, O., Lacoursière, J. O., Vought, L. B.-M., Amena, Z. & Babel, M. S. 2010 Application of hydroinformatics tools for water quality modeling and management: case study of Vientiane, Lao P.D.R. *J. Hydroinform.* **12** (2), 161–171.

McGrail, B. P. 2001 Inverse reactive transport simulator (INVERTS): an inverse model for contaminant transport with nonlinear adsorption and source terms. *Environ. Model. Softw.* **16** (8), 711–723.

Merdun, H., Çınar, Ö., Meral, R. & Apan, M. 2006 Comparison of artificial neural network and regression pedotransfer functions for prediction of soil water retention and saturated hydraulic conductivity. *Soil Till. Res.* **90** (1–2), 108–116.

Nadiri, A. A. 2007 Water level evaluation in Tabriz underground area by artificial neural networks. MS Theses, University of Tabriz, p. 178.

Nadiri, A. A., Chitsazan, N., Tsai, F. T.-C. & Asghari Moghaddam, A. 2013a Bayesian artificial intelligence model averaging for hydraulic conductivity estimation. *J. Hydro. Eng.* (in press).

Nadiri, A. A., Asghari Moghaddam, A., Tsai, F. T.-C. & Fijani, E. 2013b Hydrogeochemical analysis for tasuj plain aquifer, Iran. *Journal of Earth System Science* (in press).

Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models part I: a discussion of principles. *J. Hydrol.* **10** (3), 282–290.

Nayak, P. C., Sudheer, K. P., Rangan, D. M. & Ramasastri, K. S. 2004 A neuro-fuzzy computing technique for modeling hydrological time series. *J. Hydrol.* **291** (1–2), 52–66.

Newton, S. C., Pemmaraju, S. & Mitra, S. 1992 Adaptive fuzzy leader clustering of complex data sets in pattern recognition. *IEEE Trans. Neural Netw.* **3** (5), 794–800.

Nourani, V., Komasi, M. & Alami, M. T. 2012 Hybrid wavelet-genetic programming approach to optimize ANN modeling of rainfall–runoff process. *J. Hydrol. Eng.* **17** (6), 724–741.

Nourani, V., Moghaddam, A. A. & Nadiri, A. O. 2008 An ANN-based model for spatiotemporal groundwater level forecasting. *Hydrol. Proc.* **22** (26), 5054–5066.

Nourani, V. & Sayyah Fard, M. 2012 Sensitivity analysis of the artificial neural network outputs in simulation of the evaporation process at different climatologic regimes. *Adv. Eng. Softw.* **47** (1), 127–146.

Ozsvath, D. L. 2006 Fluoride concentrations in a crystalline bedrock aquifer Marathon County, Wisconsin. *Environ. Geol.* **50** (1), 132–138.

Palani, S., Liong, S. Y. & Tkalich, P. 2008 An ANN application for water quality forecasting. *Mar. Pollut. Bull.* **56** (9), 1586–1597.

Pulido-Calvo, I. & Gutiérrez-Estrada, J. C. 2009 Improved irrigation water demand forecasting using a soft-computing hybrid model. *Biosyst. Eng.* **102** (2), 202–218.

Ranković, V., Radulović, J., Radojević, I., Ostojić, A. & Čomić, L. 2012 Prediction of dissolved oxygen in reservoirs using adaptive network-based fuzzy inference system. *J. Hydroinform.* **14** (1), 167–179.

Rao Nagendra, C. R.. 2003 Fluoride and environment – A review. In: *Proceedings of the Third International Conference on Environment and Health* (M. J. Bunch, V. Madha Suresh & T. Vasantha Kumaran, eds). Chennai, India 15–17 December 2003. Department of Geography, University of Madras and Faculty of Environmental Studies, York University, pp. 386–399.

Saxena, V. & Ahmed, S. A. 2001 Dissolution of fluoride in groundwater: a water–rock interaction study. *Environ. Geol.* **40** (9), 1084–1087.

Saxena, V. & Ahmed, S. 2003 Inferring the chemical parameters for the dissolution of fluoride in groundwater. *Environ. Geol.* **43** (6), 731–736.

Sharma, V., Negi, S. C., Rudra, R. P. & Yang, S. 2003 Neural networks for predicting nitrate-nitrogen in drainage water. *Agric. Water Manage.* **63** (3), 169–183.

Shekofteh, H., Afyuni, M., Hajabbasi, M. A., Iversen, B. V., Nezamabadi-pour, H., Abassi, F. & Sheikholeslam, F. 2013 Nitrate leaching from a potato field using adaptive network-based fuzzy inference system. *J. Hydroinform.* **15** (2), 503–515.

Sugeno, M. 1985 *Industrial Application of Fuzzy Control*. North-Holland, New York, p. 269.

Takagi, T. & Sugeno, M. 1985 Fuzzy identification of systems and its application to modeling and control. *IEEE Trans. Syst. Man Cybernet.* **SMC-15** (1), 116–132.

Tayfur, G. & Singh, V. P. 2006 ANN and fuzzy logic models for simulating event-based rainfall-runoff. *J. Hydraul. Eng.* **132** (12), 1321–1330.

Tutmez, B. & Hatipoglu, Z. 2010 Comparing two data driven interpolation methods for modeling nitrate distribution in aquifer. *Ecol. Inform.* **5** (4), 311–315.

USEPA 2009 *National Primary Drinking Water Regulations*. US Environmental Protection Agency, EPA 816-F-09-004. Available from: http://water.epa.gov/drink/contaminants/. May 2009.

WHO 2006 *Guidelines For Drinking-Water Quality*. Vol 1, World Health Organization, First addendum to third edition, Recommendations. WHO, Geneva.

Zadeh, L. A. 1965 Fuzzy sets. *Inform. Control* **8** (3), 338–353.

Zietsman, S. 1991 Spatial variation of fluorosis and fluoride content of water in an endemic area in Bophuthatswana. *J. Dental Assoc. SA* **46** (1), 11–15.

Zounemat-Kermani, M. & Teshnehlab, M. 2008 Using adaptive neuro-fuzzy inference system for hydrological time series prediction. *Appl. Soft Comput.* **8** (2), 928–936.

First received 13 January 2013; accepted in revised form 15 April 2013. Available online 27 May 2013