

# Modelling electrical conductivity of groundwater using an adaptive neuro-fuzzy inference system

B. Tutmez<sup>a,\*</sup>, Z. Hatipoglu<sup>b</sup>, U. Kaymak<sup>c</sup>

<sup>a</sup>*Hacettepe University, Department of Mining Engineering, Beytepe, 06532 Ankara, Turkey*

<sup>b</sup>*Mersin University, Department of Geological Engineering, Cifilikköy, 33342 Mersin, Turkey*

<sup>c</sup>*Erasmus University Rotterdam, School of Economics, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands*

Received 8 March 2005; received in revised form 11 July 2005; accepted 11 July 2005

## Abstract

Electrical conductivity is an important indicator for water quality assessment. Since the composition of mineral salts affects the electrical conductivity of groundwater, it is important to understand the relationships between mineral salt composition and electrical conductivity. In this present paper, we develop an adaptive neuro-fuzzy inference system (ANFIS) model for groundwater electrical conductivity based on the concentration of positively charged ions in water. It is shown that the ANFIS model outperforms more traditional methods of modelling electrical conductivity based on the total solids dissolved in the water, even though ANFIS uses less information. Additionally, the fuzzy rules in the ANFIS model provide a categorization of ground water samples in a manner that is consistent with the current understanding of geophysical processes.

© 2005 Elsevier Ltd. All rights reserved.

*Keywords:* Electrical conductivity; Charged ions; Groundwater quality; Fuzzy modelling; Adaptive networks

## 1. Introduction

Water quality assessment is an important activity in agricultural water management. The ionic composition of water has significant influence on plant growth. Irrigation with water of insufficient quality might retard plant growth and may contaminate soil, rendering it less suitable for agriculture because of its salinity. The concentration of positively charged ions, especially of sodium is more important than others. High concentrations of sodium, both in absolute and in relative terms, make water unsuitable for irrigation (Drever, 1997), since sodium adversely affects soil structure and soil

permeability by replacing calcium and magnesium (McNeely et al., 1979).

Electrical conductivity (EC) is an important parameter in groundwater quality assessments for drinking and irrigation, since it is related to the concentration of charged particles in the water. EC is measured by an electronic probe, which applies an electric voltage between two electrodes. The resistance of the water is measured by the drop in the voltage. EC, which is inversely proportional to the resistance, is then the conductance per unit distance. Pure liquid water has a very low electrical conductivity. The presence of charged particles in the water increases its conductivity. In general, as the concentration of total dissolved solids (TDS) in the water increases, its conductance also increases. Important water quality classification systems such as the Wilcox (1948) diagram and the USA Salinity

\*Corresponding author. Tel.: +90 312 2977600;  
fax: +90 312 2992155.

E-mail address: tutmez@hacettepe.edu.tr (B. Tutmez).

Laboratory (1954) classification system use EC as input. Wilcox diagram classifies water quality based on the EC and sodium concentration of water. For this diagram, an EC level between 0 and 750  $\mu\text{S}/\text{cm}$  is determined to be excellent-good and an EC level between 750 and 2000  $\mu\text{S}/\text{cm}$  is good-permissible for irrigation. Water with an EC level greater than 2000  $\mu\text{S}/\text{cm}$  is considered to be doubtful for irrigation. Alternatively, the USA Salinity Laboratory system uses EC together with the sodium, magnesium and calcium concentrations in the water. According to this system, water with EC level between 100 and 250  $\mu\text{S}/\text{cm}$  has low, between 250 and 750  $\mu\text{S}/\text{cm}$  has medium and above 750  $\mu\text{S}/\text{cm}$  has high salinity hazard.

Since EC is influenced by the composition of mineral salts in the groundwater, it is important to understand the relationships between mineral salt composition and electrical conductivity. Traditional methods of modelling EC are based on TDS, but then it is not possible to study how the specific composition of water influences EC. This paper shows that the relation between the composition of groundwater and EC can be modelled by using an adaptive neuro-fuzzy inference system (ANFIS). ANFIS models are attractive since they can learn the underlying relations from numerical data, while the fuzzy rules obtained can provide a transparent linguistic description for the working of the model. Fuzzy systems provide the possibility of integrating (logical) information processing with the attractive mathematical properties of general function approximators (Setnes et al., 1998). Various papers have already considered engineering applications of neural-fuzzy modelling in hydrogeological-based systems. In one of these studies, Chang and Chang (2001) have shown that using ANFIS for modeling of real-time reservoir operation is practicable and effective. Another paper presented by Lu and Lo (2002) has investigated reservoir water quality using self-organizing maps and fuzzy theory. Similarly, Hasebe and Nagayama (2002) studied the reservoir operation using neural networks and fuzzy systems. In an interesting study, minimizing the variance of reservoir systems operations using fuzzy-neural techniques has been discussed in Ponnambalam et al. (2003). In a recent paper, Nayak et al. (2004) have utilized ANFIS for modelling hydrological time series. In addition to these papers, a comprehensive book entitled Fuzzy Logic in Geology has been edited by Demicco and Klir (2004).

In this paper, we study the modelling of EC from the major positively charged ions in the water by using an ANFIS. The model is developed by using real-world data regarding the water sources in Mersin region of southern Turkey. Our results indicate that the ANFIS model outperforms more traditional methods of modelling electrical conductivity based on the total solids dissolved in the water, even though ANFIS uses less information.

The paper is organized as follows. Section 2 explains how EC is currently modelled based on TDS as input. Section 3 presents a brief description of ANFIS models. In this section, architecture of ANFIS and the corresponding learning algorithms are described. The first part of Section 4 describes the study area and sampling. In the second part, a short explanation of sampling is given. Section 5 explains the application of the ANFIS method to model the electrical conductivity of water from positively charged ions. Validation of the model is also given in this section. Finally, Section 6 concludes the paper.

## 2. Modelling EC from TDS

Electrical conductance mostly depends on major anions and cations in water. The major cations in natural groundwater are sodium ( $\text{Na}^+$ ), calcium ( $\text{Ca}^{+2}$ ), potassium ( $\text{K}^+$ ), magnesium ( $\text{Mg}^{+2}$ ), while the major anions are chloride ( $\text{Cl}^-$ ), sulfate ( $\text{SO}_4^{-2}$ ), carbonate ( $\text{CO}_3^{-2}$ ), and bicarbonate ( $\text{HCO}_3^-$ ). Electrical conductivity provides a good indication of the changes in the composition of water, especially in its total dissolved solids (McNeely et al., 1979). The presence of charged ionic species in a solution makes it conductive. The relationship between ionic concentration and conductivity is fairly simple and direct in dilute solutions of single salts (Fig. 1). As the concentration is increased, the slope of the relation decreases slightly. This general behaviour is typical of all salts, but the slope of the straight part of the curve and the degree it flattens with increasing concentration are different for different salts. To conduct a current, solute ions actually must move through the solution to transfer charges, and the effectiveness of a particular ion in this process depends

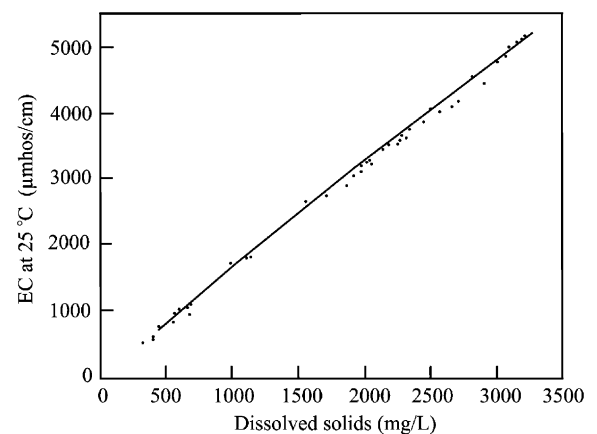


Fig. 1. Dissolved solids and EC, Gila River at Bylas, Ariz. (Hem, 1985).

on its charge, size, the way it interacts with the solvent and other factors. Increasing concentration, owing to interferences and interactions among the ions, decreases ionic mobility. Then, the relationships that affect conductivity may be complicated (Hem, 1985).

For most natural waters with EC in the range of 100–5000  $\mu\text{S}/\text{cm}$ , empirical relationships between electrical conductivity and total dissolved solids are used. It is observed that multiplying the TDS by 1.56 closely approximates EC when EC is in  $\mu\text{S}/\text{cm}$  and TDS is in  $\text{mg}/\text{l}$  (Todd, 1980; Hubbert, 1940; Luszczyński, 1961) proposes to use

$$\text{TDS} = AC$$

for conversion between the conductivity values and the TDS, where  $C$  is the conductance expressed in  $\mu\text{S}/\text{cm}$ , TDS is the total dissolved solids in  $\text{mg}/\text{l}$  and  $A$  is a conversion factor. Although this kind of linear relationship is used often the conversion factor is not constant for all waters, and it varies mostly between 0.55 and 0.75 (Hem, 1985). Therefore, a nonlinear modelling approach might capture this variation. The following section describes the basics of the ANFIS method that is used in this paper.

### 3. Adaptive neuro-fuzzy inference system (ANFIS)

In this paper, ANFIS proposed by Jang (1993) is used for modelling the electrical conductivity. ANFIS can construct an input–output mapping based on a given initial fuzzy system and available input–output data pairs by using a learning procedure. The adaptive network simulates a fuzzy inference system represented by the fuzzy if–then rules. The hybrid network of ANFIS system is functionally equivalent to Sugeno’s inference mechanism (Fuller, 1999).

#### 3.1. Outline of ANFIS

Among various fuzzy inference systems, Takagi–Sugeno (TS) systems (Takagi and Sugeno, 1985) have been applied successfully for data-driven-based fuzzy modelling. The TS model consists of a set of local input–output relations that describe the overall system. The rules in a first-order TS model have the following structure:

$R_i$ : (if  $x_1$  is  $A_{i1}$  and ... and  $x_n$  is  $A_{in}$ ) then

$$y_i = \mathbf{a}_i^T \mathbf{x} + b_i, \quad I = 1, 2, \dots, K,$$

where  $R_i$  is the  $i$ th rule in the rule base,  $\mathbf{x} = [x_1, \dots, x_n]^T$  is the input (antecedent) vector, and  $A_{i1}, \dots, A_{in}$  are the fuzzy sets defined for the respective antecedent variables. The rule consequent  $y_i$  is an affine combination of the inputs with parameters  $\mathbf{a}_i$  and  $b_i$  (Setnes et al., 1998).

An ANFIS system can be considered as an implementation of a TS system in a neural network architecture. In the following, we sketch briefly the outline of an ANFIS system by using a model with two inputs as an example (Jang et al., 1997). Let the inputs of the fuzzy system be  $x$  and  $y$ , and let the output be  $z$ . We consider a TS system with first-order consequents (Takagi and Sugeno, 1983) and two rules as follows:

**Rule 1.** If  $x$  is  $A_1$  and  $y$  is  $B_1$ , then  $f_1 = p_1x + q_1y + r_1$ .

**Rule 2.** If  $x$  is  $A_2$  and  $y$  is  $B_2$ , then  $f_2 = p_2x + q_2y + r_2$ .

To construct the adaptive system, five layers are used as shown in Fig. 2. Each layer involves several nodes described by a node function. The circles in the network represent nodes that possess no variable parameters, while the squares represent nodes that possess adaptive parameters to be determined by the network during training. The node function in each layer is described below.

Layer 1: The nodes in this layer represent the fuzzy sets in the antecedents of the fuzzy rules. It has parameters that control the shape and the location of the centre of each fuzzy set. In this study, we choose  $\mu_{A_i}(x)$  to be Gaussian with height equal to 1. The membership function is given by

$$\mu_{A_i}(x) = e^{-(x-c_i)^2/2\sigma_i^2}, \quad (1)$$

where  $c_i$  represent the centre of the Gaussian function, and  $\sigma_i$  represent the spread of the membership function. The outputs of this layer are the values of the antecedent membership functions corresponding to the fuzzified inputs of the system.

Layer 2: Every node in this layer computes the product of its inputs. The output of the layer is given by

$$w_i = \mu_{A_i}(x) * \mu_{B_i}(y), \quad i = 1, 2, \quad (2)$$

where  $\mu_{A_i}$  and  $\mu_{B_i}$  are the fuzzy sets defined for the variables  $x$  and  $y$ , respectively. The aim of this layer is to compute the degrees of activation (firing strength) of particular fuzzy rules (Gorzalczany, 2001).

Layer 3: The nodes in this layer normalize the firing strength of the rules by calculating the ratio of the  $i$ th rule’s firing strength to the sum of all rules firing strengths.

$$w_i^* = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2. \quad (3)$$

Layer 4: Nodes in this layer are adaptive where each node function represents a first order model with the consequent parameters. Thus, the output from this layer is expressed by

$$O_i^4 = w_i^* f_i = w_i^* (p_i x + q_i y + r_i), \quad (4)$$

where  $w_i^*$  is the output of Layer 3, and  $\{p_i, q_i, r_i\}$  is the parameter set.

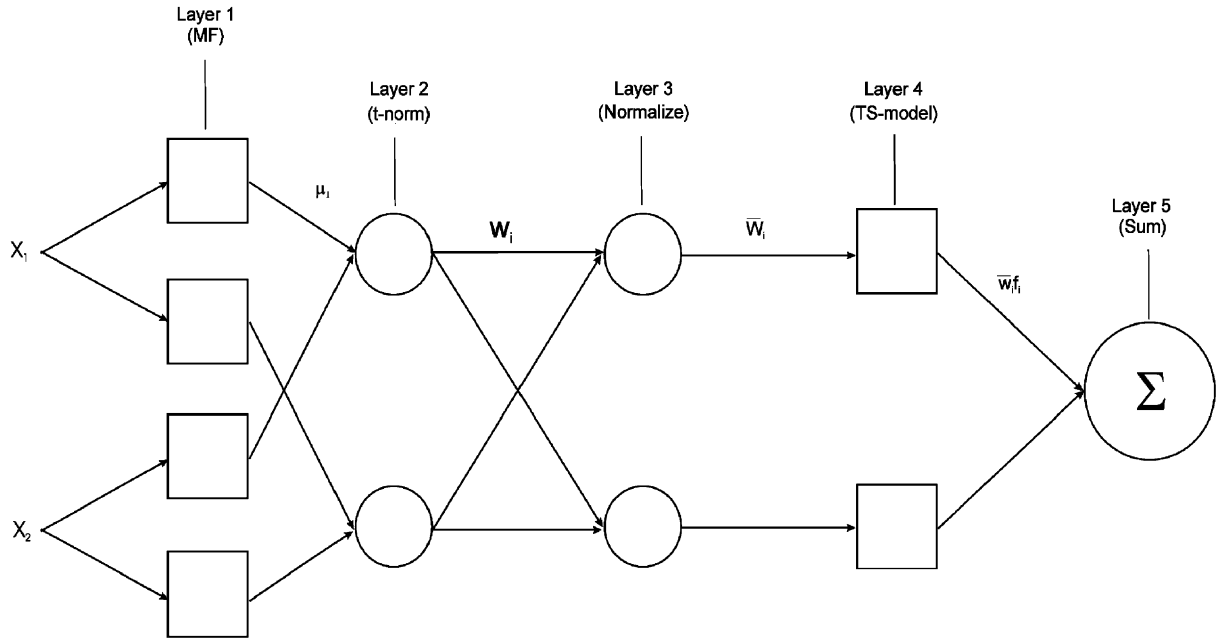


Fig. 2. ANFIS architecture.

Layer 5: This is the output layer where each node is fixed. The single node labelled  $\Sigma$  computes the overall output as the summation of all the inputs from the previous layer, i.e.,

$$O_i^5 = \sum_i w_i^* f_i = \frac{\sum_i w_i f_i}{w_i} \quad (5)$$

### 3.2. Training and learning in ANFIS

Optimizing the values of the adaptive parameters is of vital importance for the performance of the adaptive system. In particular, the premise parameters in Layer 1 and the consequent parameters in Layer 4 need to be determined. Layer 1 parameters define the centre and the spread of the antecedent membership function. Layer 4 parameters defined by the set  $\{p_i, q_i, r_i\}$  corresponds to the consequent coefficients in Eq. (4).

Jang (1993) proposed a hybrid-learning algorithm for determining the parameters of an ANFIS. A hybrid learning algorithm combines gradient descent and the least square techniques for optimizing the network parameters. In this proposed approach, the output of the system,  $f$  is written as

$$\begin{aligned} f &= \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2 = w_1^* f_1 + w_2^* f_2 \\ &= (w_1^* x) p_1 + (w_2^* y) q_1 + (w_1^*) r_1 \\ &\quad + (w_2^* x) p_2 + (w_2^* y) q_2 + (w_2^*) r_2, \end{aligned} \quad (6)$$

where  $(p_1, q_1, r_1, p_2, q_2, r_2)$  are the consequent parameters of the linear sub-systems. If the objective function to minimize is chosen to be the minimization of the squared prediction errors, the objective function is linear in the consequent parameters. Therefore, least squares estimation can be used to determine these parameters, assuming that the Layer 1 parameters are fixed. Afterwards, the Layer 4 parameters can be fixed, and a back propagation approach is used to adapt the premise parameters in Layer 1. Iterating between optimizing the Layer 1 parameters and Layer 4 parameters, the optimal values for all free parameters are computed. In Section 5, an ANFIS model for electrical conductivity is developed, where the hybrid learning approach is used to optimize the parameters of the adaptive network.

### 4. Study area and sampling

Mersin is located in the Southern part of Turkey (Fig. 3). The study area is situated between latitude  $34^\circ 20'$  and  $34^\circ 57'$  and longitude  $36^\circ 38'$  and  $37^\circ 00'$  and it covers an area of approximately  $800 \text{ km}^2$ . This study area can be divided in to two parts as its hydrogeological properties. The north part is low-productive Hillside aquifer which is mainly consist of sedimentary rocks and the south part is most-productive Coastal aquifer which is characterized by a deltaic environment. Maximum specific discharge rates are  $0.65 \text{ l/s/m}$  for Hillside aquifer

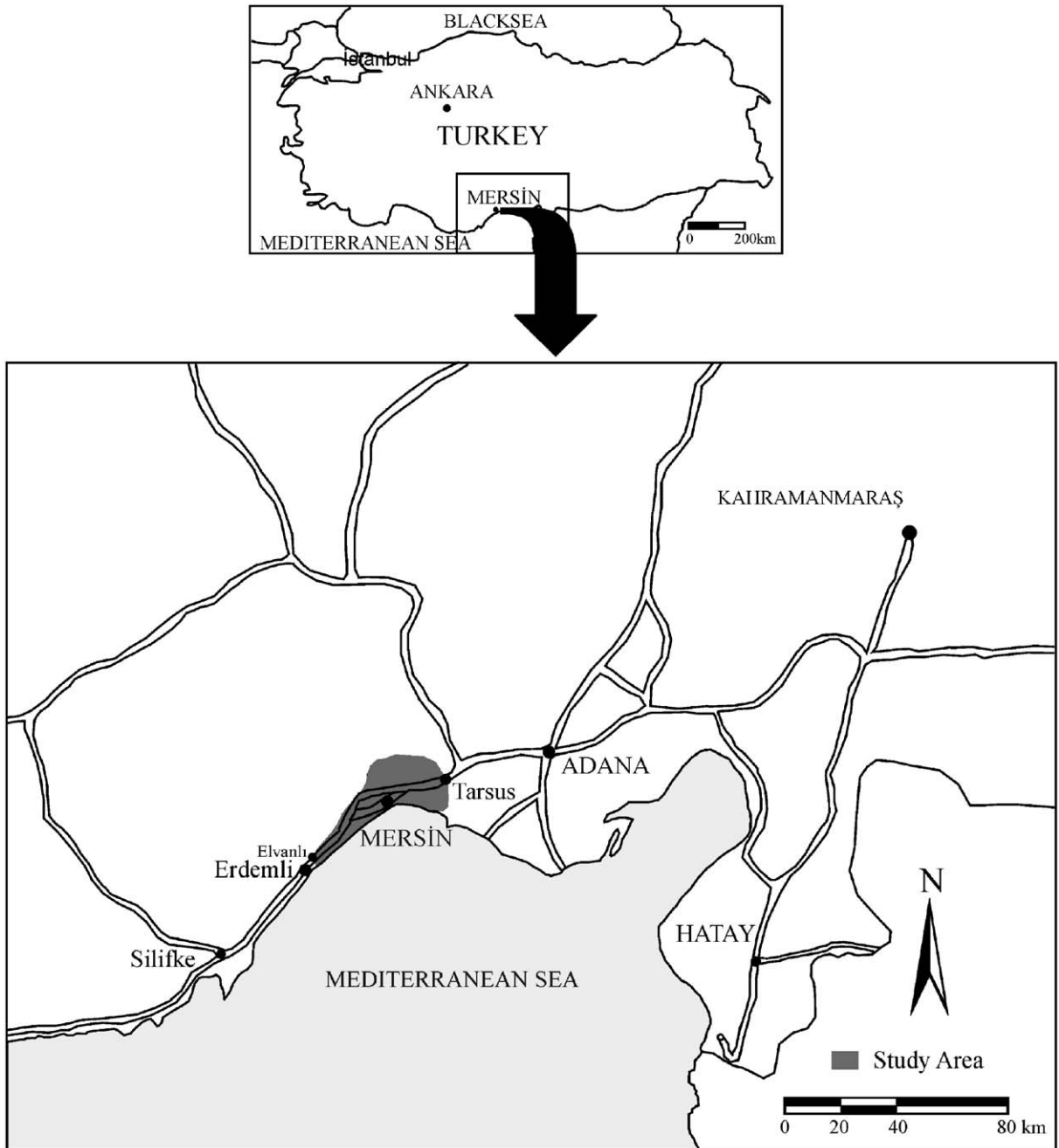


Fig. 3. Locality map of study area.

and 21.41/s/m for Coastal aquifer (Hatipoglu, 2004). Fig. 4 illustrates the distribution of Hillside and Coastal aquifer according to the outcrop of the rocks and alluvium material. It is determined from the geological map of the the study area which is prepared by Senol et al. (1998). Fifty one groundwater samples collected from springs and wells for the current study are shown

in Fig. 4. These samples were taken during September through November 2002 for chemical analyses, each sample being analysed for the major, minor ions and nutrients. Samples were preserved in the field and analysed in International Research and Application Center for Karst Water Resources Laboratory in Hacettepe University. EC, pH, and temperature were measured at the

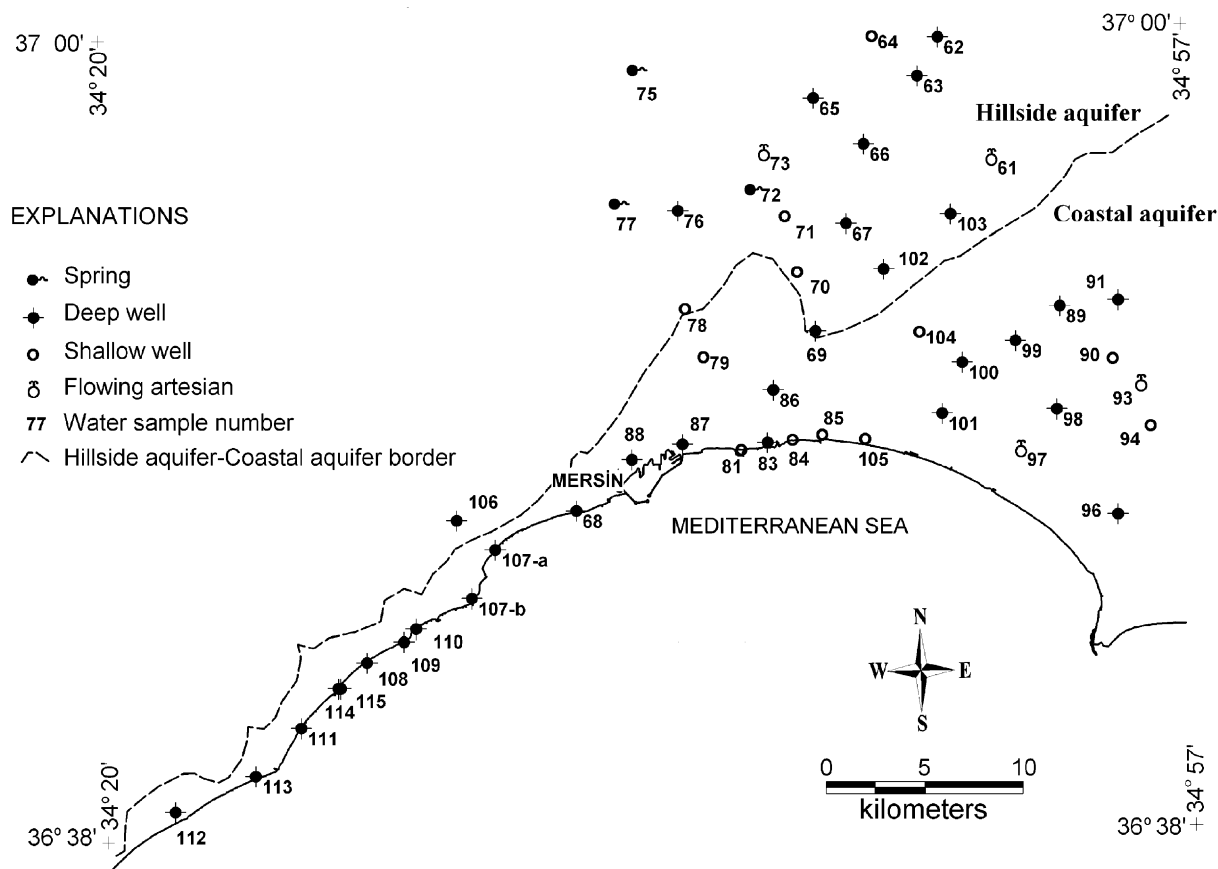


Fig. 4. Map of study area showing location of sampled wells and spring.

time of sampling. EC was measured using portable YSI EC meter, pH and temperature were measured using Orion pH meter. Samples were collected in two new 250 ml polyethylene bottles. All sampling bottles were washed with de-ionized water. Chemical analyses were performed Ion Chromatography Dionex 650. All samples had ionic balance errors below 5%.

**5. ANFIS model for electrical conductivity**

*5.1. Rule base design and optimization*

We have chosen a data driven to design the ANFIS model. First, the available data set consisting of 51 records was divided into two subsets randomly: the training set and the validation set, respectively. The training set consisted of 35 samples, and the validation set of 16 samples. To model the electrical conductivity, a network with four inputs was selected, with input variables corresponding to the concentration of Sodium (Na<sup>+</sup>), potassium (K<sup>+</sup>), calcium (Ca<sup>+2</sup>), and magnesium (Mg<sup>+2</sup>) ions. Fuzzy c-means clustering algorithm

(Bezdek, 1981) was used to design an initial rule base. After the data were clustered, the antecedent membership functions have been chosen such that the centres of the Gaussians correspond to the centres of the fuzzy clusters. The spreads of the membership functions were selected manually to provide sufficient coverage over the domain of discourse on which the variables were defined. A crucial point in the rule base design is selecting the number of rules. When fuzzy systems are designed by using fuzzy clustering, each cluster corresponds to a fuzzy rule. Hence, the number of clusters determines the number of rules. We have determined the number of clusters experimentally, by developing various models and studying the rules and their consequent parameters. The appropriate number of clusters resulted to be two. The cluster centres determined by the clustering algorithm lied at the following coordinates:

|           | Na <sup>+</sup> | K <sup>+</sup> | Ca <sup>+2</sup> | Mg <sup>+2</sup> |
|-----------|-----------------|----------------|------------------|------------------|
| Cluster 1 | 31.0            | 1.92           | 84.9             | 30.0             |
| Cluster 2 | 110             | 5.10           | 62.0             | 42.0             |



A first-order Takagi–Sugeno system was designed. Thus, the rule base consisted of two rules, each with five consequent parameters. The rules obtained from fuzzy clustering looked like as follows:

IF (Na is low) AND (*K* is low) AND  
 (Ca is high) AND (Mg is low) THEN  $y = Y$   
 IF (Na is high) AND (*K* is high) AND  
 (Ca is low) AND (Mg is high) THEN  $y = Y$ .

As this is the initial rule base, the rule consequents were determined from global least squares regression, leaving it up to ANFIS to optimize the consequent parameters. Hence, all the rules in the rule base initially had the same consequent determined by the least square estimation as

$$Y = 3.4Na + 13.4K + 4.0Ca + 8.4Mg - 10.1.$$

After the initialization, the ANFIS system was trained by using a hybrid training algorithm. Figs. 5–8 show the initial and the final membership functions for the four inputs that have been determined by training. For each input, the membership functions are labelled as ‘low’ and ‘high’. The training algorithm has produced a considerable change in the membership functions for

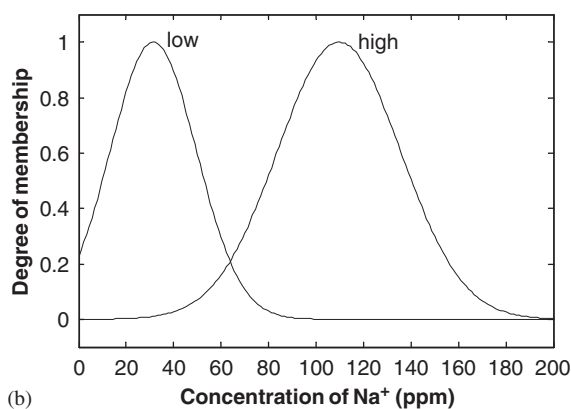
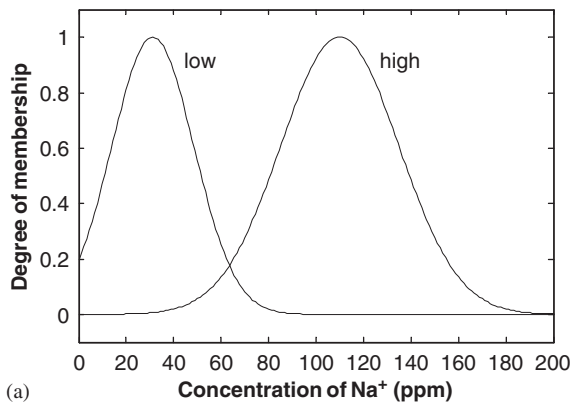
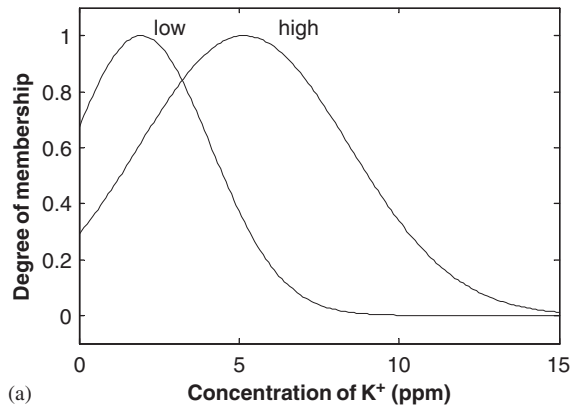
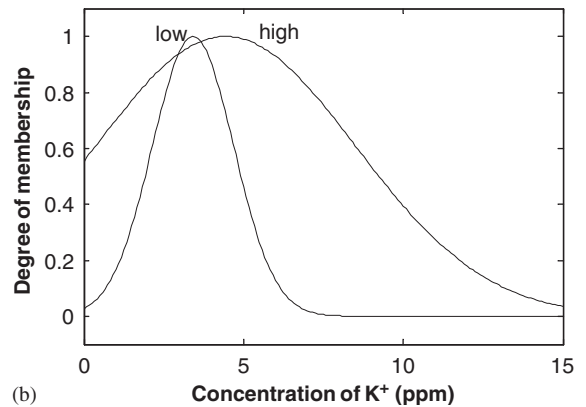


Fig. 5. Initial and final membership functions for  $Na^+$ .



(a)



(b)

Fig. 6. Initial and final membership functions for  $K^+$ .

$K^+$ . The membership functions for other inputs have been modified only slightly. After optimization, the following rule consequents have been determined.

**Rule 1.**  $y_1 = 1.89Na + 15.11K + 4.01Ca + 5.08Mg + 110.9$ .

**Rule 2.**  $y_2 = 3.44Na + 9.92K + 4.64Ca + 8.00Mg + 14.2$ .

The fuzzy system has identified two relations: one is for water characterized by high calcium content and low other cations, the other for water characterized by low calcium and high other cations. Compared with global linear regression, the coefficient for sodium is lower in the former case, while the coefficient for sodium is higher in the latter case. Furthermore, one can also observe a significant change in the coefficients for potassium, compared to the global regression. Therefore, we conclude that the learning algorithm has significantly modified the local models.

### 5.2. Evaluation of the prediction performance

To assess the performance of the identified ANFIS model, we have plotted the measured EC values against the predicted EC values. Fig. 9 illustrates the results

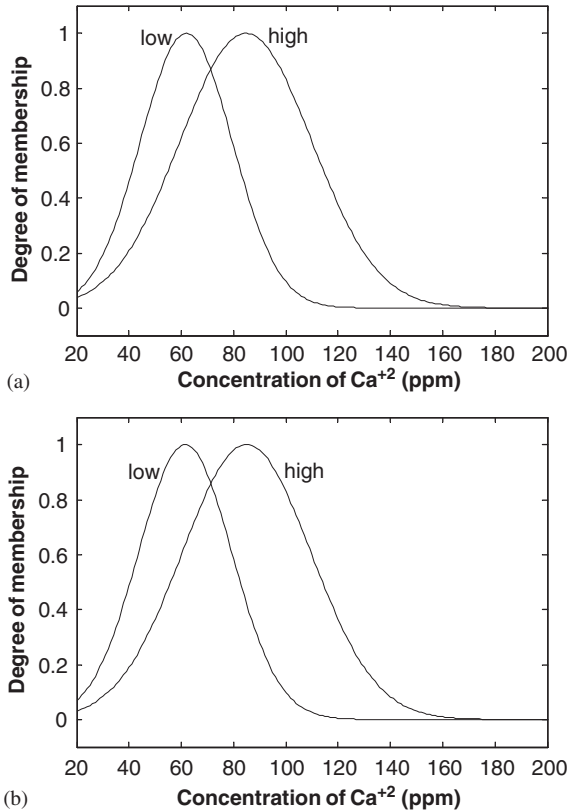


Fig. 7. Initial and final membership functions for  $\text{Ca}^{+2}$ .

together with the cross-correlations between predicted and measured data both for the training set and the validation set, respectively. The large determination coefficient indicates that the model has good prediction capability. We have also compared the performance of the ANFIS model against the performance of a traditional method for modelling EC based on TDS, as explained in Section 2. Fig. 10 shows the prediction from the TDS method against the measured EC together with the cross-correlations between predicted and measured data both for the training set and the validation set, respectively. As observed from the determination coefficient, the ANFIS method outperforms the method based on TDS.

We have also investigated to what degree our model can explain the variance in the data by using the “variance account for” VAF in (7) as a performance index (Grima, 2000):

$$\text{VAF} = \left( 1 - \frac{\text{var}(y - y^*)}{\text{var}(y)} \right) 100\%, \quad (7)$$

where ‘var’ denotes the variance,  $y$  is the measured value,  $y^*$  is the predicted value. In general, the closer the VAF to 100%, the better the model performs. Table 1 shows the VAF values obtained from ANFIS and from

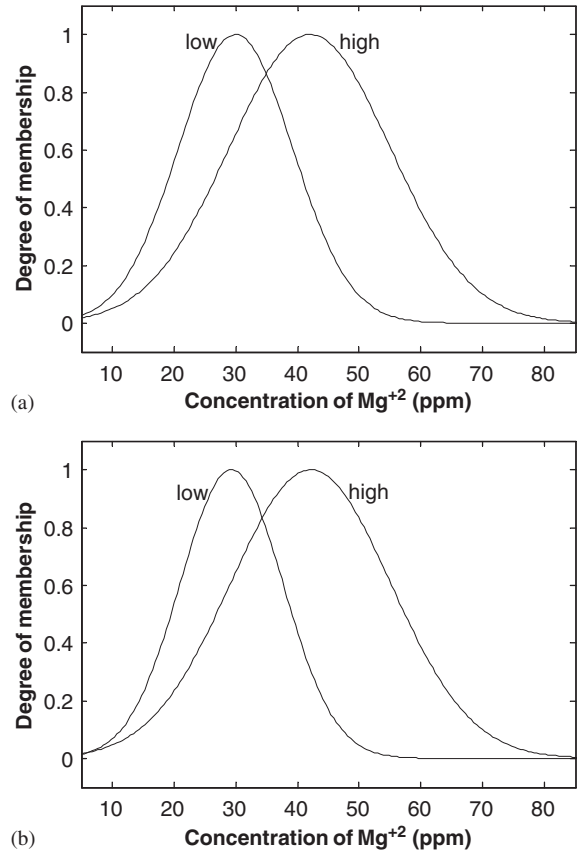


Fig. 8. Initial and final membership functions for  $\text{Mg}^{+2}$ .

TDS models. The ANFIS model performs better than the TDS model, both on the training data and on the validation data.

In order to test the sensitivity of the ANFIS modelling approach to sampling variation, we have randomly drawn the training and the validation data sets 13 more times. In each case, the training set had 35 records and the validation set had 16 records. With each training data, ANFIS models and TDS models were developed and the performance on both the training data and the validation data was quantified by using the VAF index. The values of the VAF for 13 experiments are given in Table 2.

Table 2 shows that on average, ANFIS models outperform the TDS model. This indicates a good generalization capability for the ANFIS model. The VAF values predicted by the ANFIS and the TDS models for both the training and the evaluation data are shown in Figs. 11 and 12, respectively.

### 5.3. Discussion

Our results show that the ANFIS modelling that we propose for electrical conductivity has high predictive



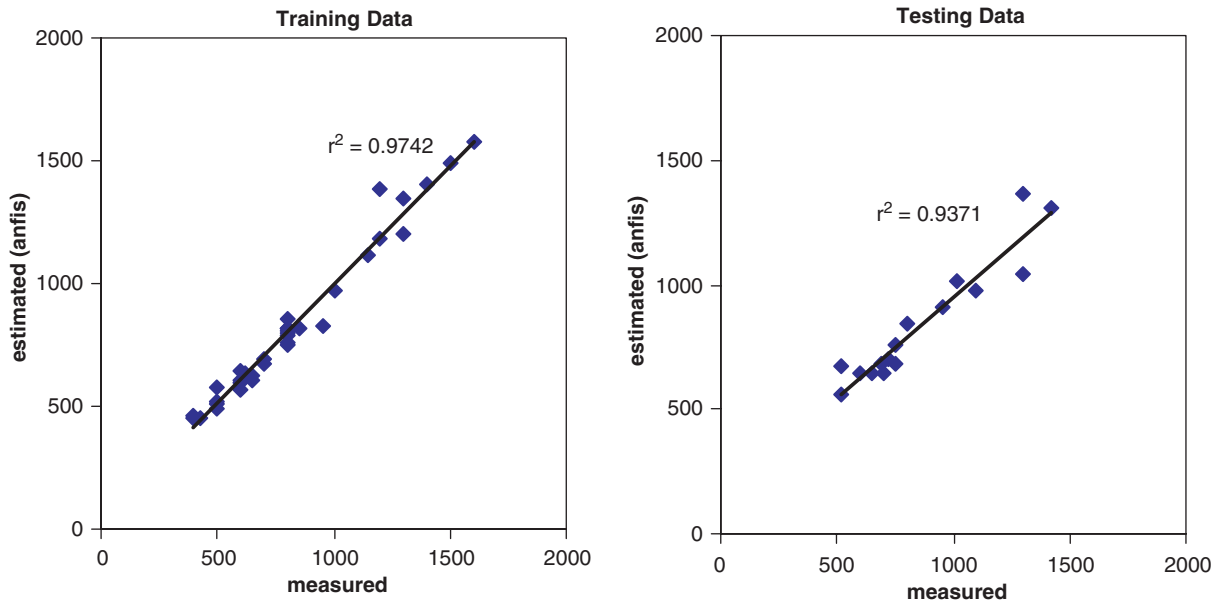


Fig. 9. Scatter plot of measured and predicted values for ANFIS model.

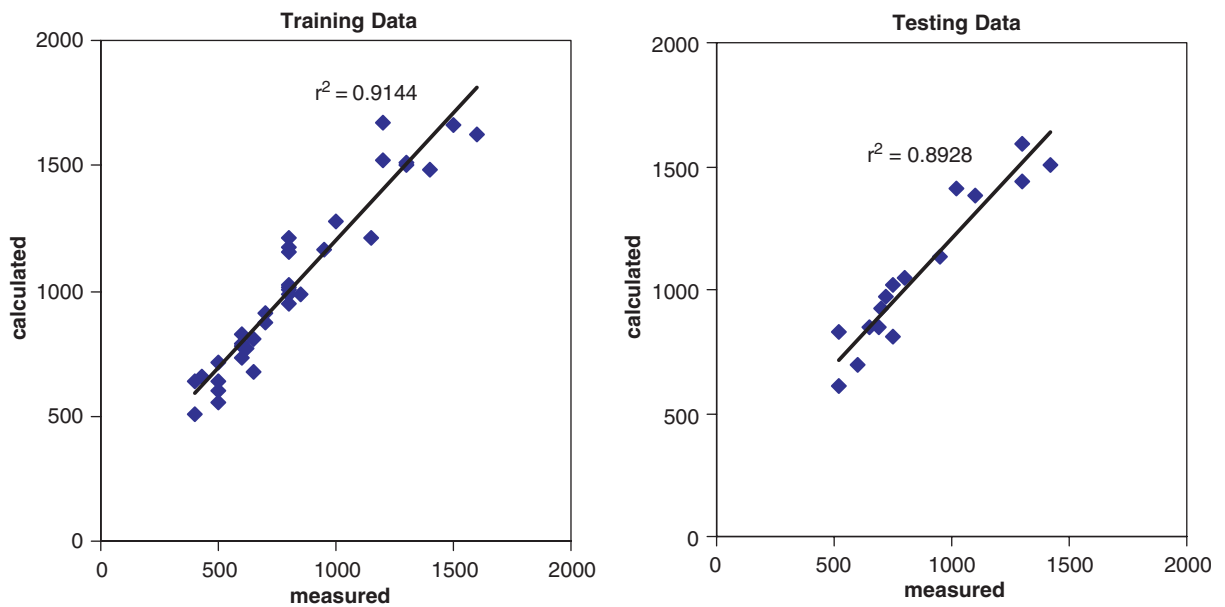


Fig. 10. Scatter plot of measured and predicted values for TDS method.

power. It outperforms methods based on total dissolved solids, as indicated by the VAF performance index. However, model validation is more important with ANFIS models, since they have more parameters than the models based on TDS. Hence, there is a danger of over fitting, which must be managed by the model validation procedure. Note that the method-based total-

dissolved solids needs the identification of all ionic species in the solution. In this sense, it needs more information for the modelling. The ANFIS model, however, has as inputs the concentrations of the major cations in the solution. Hence, it can achieve better performance, while it uses less information. This is possible because of the nonlinear function

Table 1  
Performances comparison of ANFIS and TDS model

| VAF–ANFIS/Training | VAF–TDS/Training | VAF–ANFIS/Testing | VAF–TDS/Testing |
|--------------------|------------------|-------------------|-----------------|
| 0.9742             | 0.9027           | 0.8957            | 0.8897          |

Table 2  
VAF statistics for ANFIS and TDS model

| VAF–ANFIS/Training | VAF–TDS/Training   | VAF–ANFIS/Testing  | VAF–TDS/Testing    |
|--------------------|--------------------|--------------------|--------------------|
| 0.9640             | 0.9031             | 0.9310             | 0.7696             |
| 0.9623             | 0.9148             | 0.9230             | 0.8356             |
| 0.9855             | 0.9096             | 0.7016             | 0.9376             |
| 0.9715             | 0.8613             | 0.9091             | 0.9338             |
| 0.9688             | 0.9030             | 0.9066             | 0.8158             |
| 0.9766             | 0.9070             | 0.9081             | 0.8920             |
| 0.9603             | 0.8894             | 0.9591             | 0.8739             |
| 0.9828             | 0.9170             | 0.8555             | 0.8962             |
| 0.9743             | 0.9030             | 0.9120             | 0.9066             |
| 0.9764             | 0.9222             | 0.8529             | 0.8768             |
| 0.9674             | 0.8959             | 0.9784             | 0.8769             |
| 0.9513             | 0.8899             | 0.9763             | 0.9263             |
| 0.9770             | 0.8856             | 0.9147             | 0.9125             |
| Average<br>0.9707  | Average<br>0.9002  | Average<br>0.9022  | Average<br>0.8811  |
| St. Dev.<br>0.0095 | St. Dev.<br>0.0160 | St. Dev.<br>0.0713 | St. Dev.<br>0.0490 |

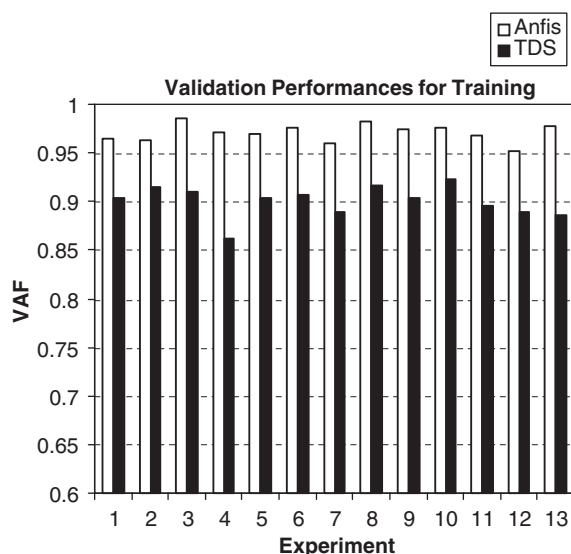


Fig. 11. VAF values from ANFIS and TDS models for training sets.

approximation characteristic of ANFIS. Since determining and measuring all ionic species in the water is more time consuming and more expensive, our modelling approach could reduce costs of sampling and monitoring water resources in water management.

In addition to the numerical prediction, ANFIS models provide more information regarding the process compared to the TDS based models. In particular, the fuzzy rules identified can be studied and the degree of firing of the rules can provide information about the type of composition of cations in the studied sample. For all the samples in our data set, we have studied to what degree the two local models represented by the rule consequents are active by looking at the normalized degree of fulfilment (relative degree of activation) of the fuzzy rules (i.e.  $w_i^*$  in Layer 3 of the ANFIS). One can use the normalized degree of fulfilment as a simple classifier by assigning a sample to the rule that fires most. Fig. 13 shows the dominant rule in each of the available data points in our total data set.

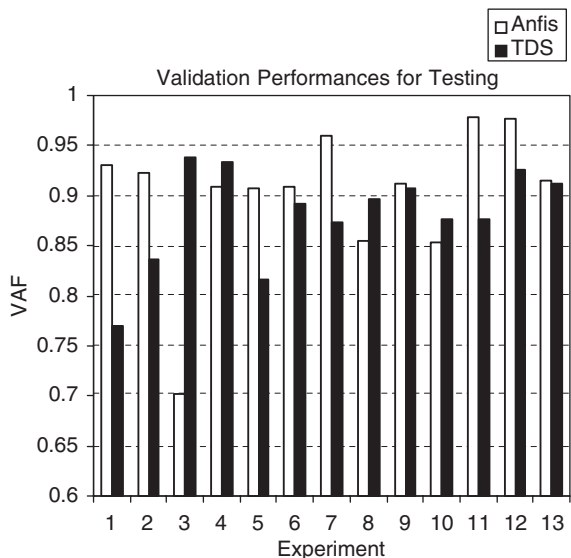


Fig. 12. VAF values from ANFIS and TDS models for evaluation sets.

Recall that Rule 1 indicates sources with low sodium content but high calcium content. Conversely, Rule 2 indicates sources with high sodium content and low calcium content. This implies that the fuzzy model has captured an essential characteristic of geophysical process in coastal areas. As a result of dissolution of calcite,  $Ca^{+2}$  and  $HCO_3^-$  ions dominate in fresh water in coastal areas. When seawater intrudes in a coastal fresh water aquifer, an exchange of cations takes place, whereby sodium replaces calcium (Appelo and Postma, 1994). The fuzzy rules have partitioned the water sources according to this geophysical phenomenon. As expected, Rule 2 dominates in samples that are close to the coastline, while Rule 1 dominates in the hillside samples. However, Rule 2 is also active on some hillside samples, indicating a possible mixing with seawater or dissolution of salts.

Finally, we have also considered to what degree our fuzzy system is able to distinguish samples with high sodium content from samples with low sodium content. For that reason, we have shown in Fig. 13 also the relative sodium content (in percentages) of the data samples. We find that in sources where Rule 2 is active,

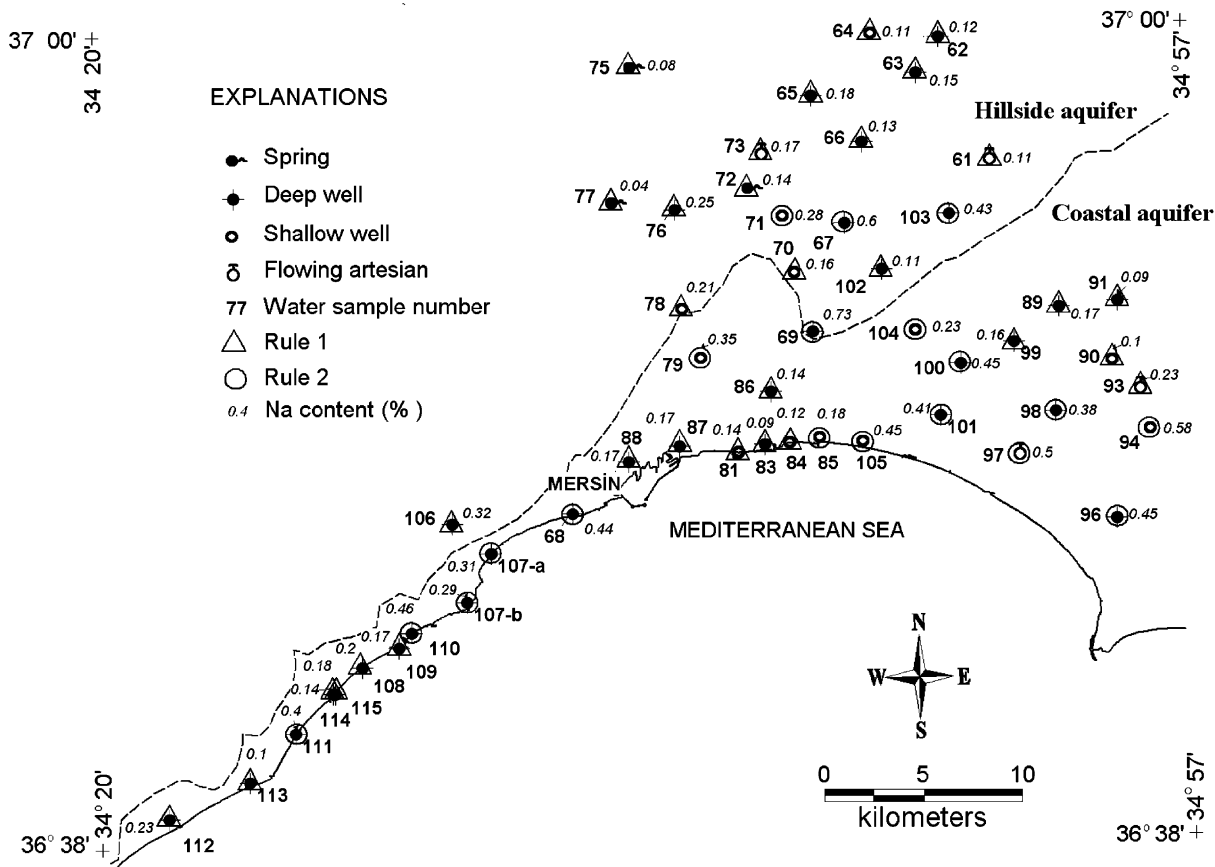


Fig. 13. Rule activation and sodium content for data points.

the relative sodium content is high (above 0.25%), while in sources where Rule 1 is active, the relative sodium content is low. There were two exceptions to this rule: our fuzzy model did not classify data sources 85 and 106 correctly.

## 6. Conclusions

We have presented an adaptive neuro-fuzzy inference system (ANFIS) for modelling the relation between the major cations in groundwater and its electrical conductivity. It has been observed that the ANFIS model outperforms more traditional methods of modelling electrical conductivity based on the total solids dissolved in the water. In addition to the numerical prediction power, another attractive property of the fuzzy model provided by ANFIS is its transparency. We have observed that the fuzzy rules in the ANFIS model provide a categorization of ground water samples in a manner that is consistent with the current understanding of geophysical processes in coastal areas. In our two-rule model, it has been found that one of the rules describes the relation between cationic composition and electrical conductivity for low sodium and high calcium samples, while the other rule describes the relation between high sodium and low calcium samples.

The approach can easily be extended to other hydrogeology-based domains such as evaluating the special reservoir topics and pollution processes. It is clear that the soft computing approaches such as fuzzy sets and neural networks can have a large impact in geological and mining systems.

## Acknowledgements

The authors would like to extend their deep appreciation to Professor Serdar Bayari of Hydrogeology Division of Geological Engineering Department at Hacettepe University.

This work was partly conducted when the first author was a visiting research fellow at Department of Computer Science, Erasmus University Rotterdam, supported by the Scientific and Technical Research Council of Turkey (Tubitak) within the scope of Münir Birsal Foundation Fellowships Programme.

## References

Appelo, C.A.J., Postma, D., 1994. *Geochemistry, Groundwater and Pollution*. A.A. Balkema, Rotterdam 500pp.  
 Bezdek, J.C., 1981. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York 256pp.

Chang, L-C., Chang, F.-J., 2001. Intelligent control for modeling of real-time reservoir operation. *Hydrogeological Processes* 15 (9), 1621–1634.  
 Demicco, R.V., Klir, G.J., 2004. *Fuzzy Logic in Geology*. Elsevier Academic Press, Amsterdam 346pp.  
 Drever, J.I., 1997. *The Geochemistry of Natural Waters. Surface and Groundwater Environments*. Prentice-Hall, Englewood Cliffs, NJ 436pp.  
 Fuller, R., 1999. *Introduction to Neuro-Fuzzy Systems*. Physica-Verlag, Heidelberg 289pp.  
 Gorzalczy, M.B., 2001. *Computational Intelligence Systems and Applications*. Physica-Verlag, Heidelberg 362pp.  
 Grima, M.A., 2000. *Neuro-Fuzzy Modeling in Engineering Geology*. A.A. Balkema, Rotterdam 252pp.  
 Hasebe, M., Nagayama, Y., 2002. Reservoir operation using the neural network and fuzzy systems for dam control and operation support. *Advances in Engineering Software* 33, 245–260.  
 Hatipoglu, Z., 2004. *Hydrogeochemistry of Mersin-Tarsus coastal aquifer*. Ph.D. dissertation, Hacettepe University, Pure and Applied Science, Ankara, 142pp. (in Turkish).  
 Hem, J.D., 1985. Study and interpretation of the chemical characteristics of natural water. US Geological Survey Water Supply Paper 2254, 263pp.  
 Hubbert, M.K., 1940. The theory of groundwater motion. *Journal of Geology* 48, 785–944.  
 Jang, J.-S.R., 1993. ANFIS: adaptive network-based fuzzy inference system. *IEEE Transactions on Systems, Man and Cybernetics* 23 (3), 665–685.  
 Jang, J.-S.R., Sun, C.T., Mizutani, E., 1997. *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. Prentice-Hall International (UK) Limited, London 614pp.  
 Lu, R.-S., Lo, S.-L., 2002. Diagnosing reservoir water quality using self-organizing maps and fuzzy theory. *Water Research* 36, 2265–2274.  
 Lusczynski, N.J., 1961. Head and flow of groundwater of variable density. *Journal of Geophysical Research* 66, 4247–4256.  
 McNeely, R.N., Neimanis, V.P., Dwyer, L., 1979. *Water Quality Sourcebook, A Guide to Water Quality Parameters*. Inland Waters Directorate, Water Quality Branch, Ottawa, Canada 88pp.  
 Nayak, P.C., Sudheer, K.P., Rangan, D.M., Ramasastri, K.S., 2004. A neuro-fuzzy computing technique for modeling hydrological time series. *Journal of Hydrology* 291, 52–66.  
 Ponnambalam, K., Karray, F., Mousavi, S.J., 2003. Minimizing variance of reservoir systems operations benefits using soft computing tools. *Fuzzy Sets and Systems* 139, 451–461.  
 Senol, M., Sahin, S., Duman, T.Y., 1998. *The geological investigation of Mersin Region*. General Directorate of Mineral Research and Exploration of Turkey (unpublished report), Ankara, 46pp. (in Turkish).  
 Setnes, M., Babuska, R., Verbruggen, H.B., 1998. Transparent fuzzy modeling. *International Journal of Human-Computer Studies* 49, 159–179.  
 Takagi, H., Sugeno, M., 1983. Derivation of fuzzy control rules from human operator's control actions. In: *Proceedings of IFAC Symposium on Fuzzy Information, Knowledge Representation and Decision Analysis*, pp. 55–60.

- Takagi, H., Sugeno, M., 1985. Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics* 15, 116–132.
- Todd, D.K., 1980. *Groundwater Hydrology*. Wiley, New York 535pp.
- US Salinity Laboratory Staff, 1954. *Diagnosis and Improvement of Saline and Alkali Soils*. US Department of Agriculture Handbook No: 60, Washington, DC.
- Wilcox, L.V., 1948. *The quality of water for irrigation use*. US Department of Agriculture, Technical Bulletin 962. Washington DC.