



# Application of Dempster–Shafer theory and fuzzy analytic hierarchy process for evaluating the effects of geological formation units on groundwater quality

Marzieh Mokarram<sup>1</sup> · Majid Hojati<sup>2</sup> · Ali Saber<sup>3</sup>

Received: 16 January 2019 / Accepted: 22 April 2019 / Published online: 9 May 2019  
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

## Abstract

This study investigates the impacts of different geological units on groundwater quality of an aquifer in southern Iran. The Kriging interpolation technique with a Gaussian semivariogram model was employed to prepare groundwater maps for different water quality constituents. In the next stage, two different models based on fuzzy analytic hierarchy process (AHP) and Dempster–Shafer theory (DST) were used to evaluate the overall water quality index based on the World Health Organization’s drinking water standard in different parts of the aquifer. The DST model was able to generate water quality maps with 99.5%, 99%, and 95% confidence levels. The water quality maps were subsequently compared with the geology map of the area to determine the effects of different soil types on the water quality of the aquifer. Both methods showed poor water quality indices in the areas with an Asmari formation containing elevated levels of chloride and sodium ions. Comparison of water quality maps generated by the fuzzy-AHP and DST model revealed that the DST could more reliably handle the uncertainty in the water quality data, and thus was able to generate more accurate water quality maps. Increasing the confidence level in the DST model yielded water quality maps with a decreased overall water quality index. Results of this study could assist water management practices to generate water quality maps for their groundwater resources with confidence levels commensurate socio-economic importance of the region.

**Keywords** Drinking water · Groundwater management · Water scarcity · Salt dome · Uncertainty analysis · Kriging · Decision-making

---

Responsible editor: Marcus Schulz

---

✉ Ali Saber  
sabersic@unlv.nevada.edu

Marzieh Mokarram  
m.mokarram@shirazu.ac.ir

Majid Hojati  
majid.hojati@ut.ac.ir

- <sup>1</sup> Department of Range and Watershed Management, College of Agriculture and Natural Resources of Darab, Shiraz University, Shiraz, Iran
- <sup>2</sup> Department of Remote Sensing and GIS, Tehran University, Tehran, Iran
- <sup>3</sup> Department of Civil and Environmental Engineering and Construction, University of Nevada Las Vegas, 4505 S. Maryland Pkwy., Las Vegas, NV 89154, USA

## Introduction

Groundwater is considered as an indispensable water source for potable, agricultural, and industrial purposes in arid and semi-arid areas. The increased population and industrialization over the recent years have intensified the water extraction from aquifers, deteriorating the quality of groundwater resources (Amer et al. 2012). Groundwater resources in arid and semi-arid areas are typically located in the vicinity of salt domes. Dissolution of minerals from the salt domes into the groundwater can significantly decrease the quality of these limited freshwater (Islam et al. 2017). Hence, monitoring the groundwater quality variations in both spatial and temporal scales is of great importance in water-scarce regions.

In recent years, various data-driven models including artificial neural networks (ANNs), adaptive-network-based fuzzy inference system (ANFIS), support vector regressions (SVRs), locally weighted projection regressions (LWPRs), and relevance vector machines (LVMs) have been employed

to simulate the groundwater quality (Bieroza et al. 2018; Diamantini et al. 2018; Rowles et al. 2018; Todorov et al. 2018; Villa-Achupallas et al. 2018; Wu et al. 2018). Among them, ANFIS and ANN systems have been proven to be promising in water resources quality modeling (Cordoba et al. 2014; Gazzaz et al. 2012; Lallahem and Hani 2017; Nadiri et al. 2017; Najah et al. 2013; Sarkar and Pandey 2015). Kuo et al. (2004) used three different ANN systems to model the groundwater quality variations in Taiwan. Their results indicated that including the data set's maximum and minimum values in the training process could increase the forecast accuracy of the model. Khalil et al. (2005) used ANN, SVM, LWPR, and RVM systems to predict groundwater quality in Sumas-Blaine aquifer, north-western Washington State. They reported that the learning algorithm in RVM, SVM, and LWPR could filter the noise from the inputs, whereas insufficient training of ANNs could lead to overfitting. Khashei-Siuki and Sarbazi (2015) used ANFIS, ANN, inverse distance weighted (IDW), kriging, and co-kriging geostatistical models to predict groundwater quality in Mashhad plain, Iran. They reported that compared with other methods, the ANN system could more accurately predict the spatial groundwater quality variations.

Rahimi and Mokarram (2012) used a linear fuzzy method coupled with GIS to generate fuzzy groundwater quality maps for an aquifer in southwestern Iran. Their results indicated that the coupled system could accurately map the aquifer's spatial water quality variations. Mokarram and Sathyamoorthy (2016) investigated the quality of groundwaters in western Iran, based on both concentrations of inorganic constituents and landform classes. Their results indicated that a trained fuzzy system based on the analytic hierarchy process (AHP) could successfully predict water quality of water resources based on the study site's digital elevation model (DEM).

Discrete water quality measurements might follow a specific statistical distribution system. Hence, an appropriate uncertainty model that best suits the system's statistical distribution should be used for uncertainty manage purposes. Non-definite model components, stemmed from either the nature of the data or model principles, can decrease the reliability of the model's outputs, increasing the uncertainty in the decision-making process. Different methods including the Bayesian theory (Gershman and Blei 2012), fuzzy theory (Zadeh 1965), and Dempster–Shafer (Shafer 1976) have been employed to incorporate the uncertainty in the decision-making process. Fuzzy systems are able to consider different principle and categorizations in the decision-making process (Venkatraman et al. 2015).

The Dempster–Shafer theory (DST) is particularly of interest as it can consider both observational and epistemological uncertainties. The mathematical theory of evidence was first introduced by Dempster in 1976 (Dempster 1968; Dempster 2008) and it was developed by Shafer in 1976 (Shafer 1976).

This theory is the generalized form of the Bayesian theory that simultaneously considers both uncertainty and inaccuracy in the system (Chaabane et al. 2008). This theory is of significant importance as it is capable of considering conditions or events with different probabilities. This theory uses a probability interval, ranging from the lowest probability, known as belief, to the highest probability (plausibility) for considering uncertainty in a system. The DST can be used as a tool to analyze uncertainty in the imprecise probability theory (Helton 1997). The theory of interval statistical models (Cumming and Fidler 2005) and the interval probability theory (Dempster 2008) are two methods providing upper and lower probability bounds for imprecise probabilities. Only a few studies have used the DST in the areas of water resources and groundwater management. Neshat and Pradhan (2015) used the DST and GIS for risk assessment of groundwater pollution. They were able to present groundwater pollution risk maps based on the DST. Rahmati and Melesse (2016) also used the DST for predicting groundwater quality in Khuzestan province, Iran. The results showed that the DST model is suitable for the determination of groundwater quality.

Extensive fieldwork and the high cost required for regular water quality monitoring of large aquifers can limit the availability of high-quality data from the entire extents of the aquifer. Hence, the use of flexible methods based on DST or fuzzy theory can substantially benefit the groundwater management practices to handle the uncertainties associated with non-inclusive, patchy but still precious water quality data, and to reliably estimate the groundwater quality of the entire area.

However, the use of differences in principles and methodologies used in Bayesian or Dempster–Shafer or fuzzy theories might result in different water quality predictions and thus different water quality maps, particularly when predictions with higher confidence levels are demanded.

Only a few studies have attempted to characterize the differences in water quality maps obtained from different theories. The current study compares the DST and fuzzy AHP methods for predicting water quality in Firoozabad aquifer, southern Iran. A prolonged drought in this region, combined with over-extraction from groundwater wells, and salinization from different geological units including salt dome have significantly influenced the groundwater quality of this aquifer. In addition, different geological formations including salt domes have significantly influenced the groundwater quality of this aquifer. This study compares the water quality maps with different confidence levels generated based on the DST and fuzzy AHP methods to investigate the effects of different geological units in the region on salinity and overall water quality index of the aquifer.

To date, no specific functional toolbox has been developed for incorporating DST in GIS and remote sensing packages. Hence, this study employed the imprecise probability

propagation (IPP) Toolbox in MATLAB R2017a for applying the DST to raster image data sets.

## Methodology

### Study site

This study focuses on Firoozabad aquifer with a surface area of 566.25 km<sup>2</sup>, located in a semi-arid climate (between 28° 36'–28° 57' N and 52° 16'–52° 46' E) in southern Fars Province, Iran (Fig. 1). This aquifer is an important freshwater source for drinking water and agriculture applications in the region.

The lowest and highest elevations in the study site are 1134 m and 2885 m, respectively. The water quality of this aquifer is influenced by salt domes covering an area of 2.43 km<sup>2</sup> in the south and east of aquifer (Fig. 2a). These domes with an average height of 220 m (1320 m above sea level) are results of Hormuz series evapotranspiration during late-Precambrian to the middle-Cambrian period which are now exposed to the Earth surface. Saline springs can also be found in the vicinity of the salt dome. Figure 2a shows different geological formations in the study area. Table 1 summarizes the surface area of different landforms in Firoozabad aquifer. As shown in Fig. 2a and Table 1, the quaternary (alluvial terraces) and OMas (Asmari) geological units with areas of 281.62 km<sup>2</sup> and 119.82 km<sup>2</sup>, respectively cover a majority of surface area of the study site.

In order to evaluate the groundwater quality over the study period (23 September 2017–21 March 2017), water samples from 150 different monitoring wells (Fig. 2b) were obtained in a monthly manner and were subjected to water quality analysis to measure Cl<sup>-</sup>, Na<sup>+</sup>, total dissolved solids (TDS),

electrical conductivity (EC), Ca<sup>2+</sup>, Mg<sup>2+</sup>, SO<sub>4</sub><sup>2-</sup>, and Th concentrations in samples (Fars Regional Water Authority, 2016). Three samples were collected from each sampling point, and the average of the three concentrations for each water quality constituent was reported.

### Geostatistical analyses

Interpolations for different water quality data, i.e., Cl<sup>-</sup>, Na<sup>+</sup>, TDS, EC, Ca<sup>2+</sup>, Mg<sup>2+</sup>, SO<sub>4</sub><sup>2-</sup>, and Th concentrations, were performed based on the ordinary Kriging method using ArcGIS v.10.5 (ESRI 2018), which subsequently were used as the input data for DST and fuzzy-AHP methods. The ordinary kriging is a promising interpolation technique for data sets with special autocorrelation or directional bias. The ordinary kriging method interpolates the sample values in unmeasured locations based on the influence of surrounding measured values. In the kriging method, the influence of surrounding measured values can be determined based on the semivariance of data's spatial distribution. The semivariance ( $\gamma$ ) for water quality parameter  $Z$  measured at different locations can be determined as (Oliver and Webster 1990):

$$\gamma(h) = \frac{1}{2n} \sum_{i=1}^n [Z(x_i) - Z(x_i + h)]^2 \tag{1}$$

where  $h$  is the distance between the sampling locations of  $x_i$  and  $x_i + h$ ,  $Z(x_i)$  and  $Z(x_i + h)$  are the measured values of variable  $Z$  at the corresponding locations, and  $n$  is the number of pairs of sample measurements.

Basic spatial parameters including nugget and partial sill for different semivariogram models including circular, spherical, exponential, and Gaussian were calculated. Nugget defines the micro-scale variability and measurement error of

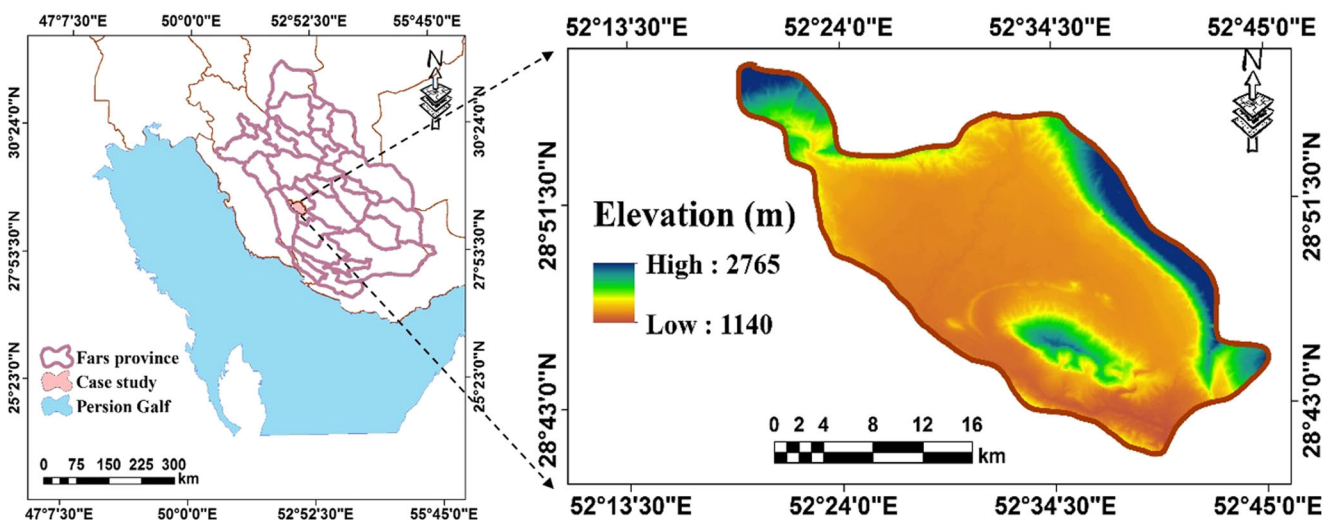
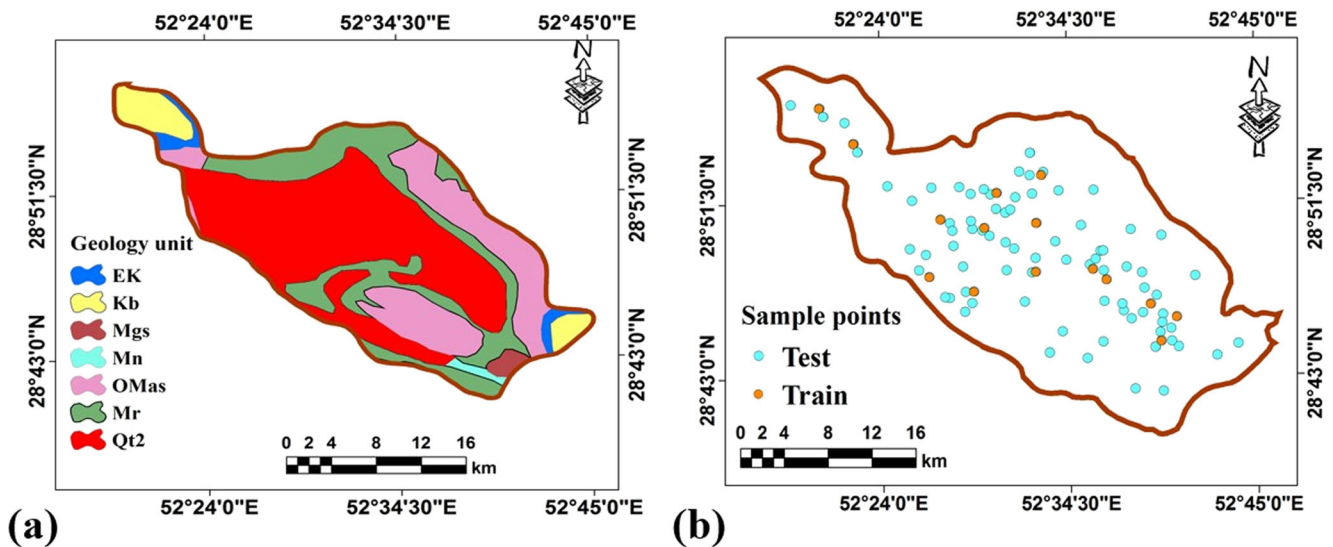


Fig. 1 Location and digital elevation model of Firoozabad aquifer (study area), and location of the salt dome in the south of the aquifer (source: <http://earthexplorer.usgs.gov>)



**Fig. 2** **a** Geological formations in the study area and **b** location of monitoring wells used for water quality analysis

collected data points, whereas partial sill indicates the amount of variation in the data (López-Granados et al. 2002).

Preliminary evaluations showed that among different models (circular, spherical, exponential, and Gaussian models), the Gaussian model could best describe the distribution of data points; hence, semivariance model based on the Gaussian model was used for interpolation purposes.

**Fuzzy analytic hierarchy process (AHP)**

Membership functions (MFs) were used to prepare fuzzy maps for water quality parameters. In general, a fuzzy set of *Z* in *B* can be defined as:

$$Z = \{b, \mu A(b) \mid b \in B\} \tag{2}$$

where  $\mu A(b)$  is called the membership function of *b* in *B*. The MF linearly maps each element of *b* to a membership value between 0 and 1, indicating the lowest and the highest degrees for a membership, respectively (Mokarram and

Sathyamoorthy 2016; Shobha et al. 2013). Fuzzy maps for each input data in the study area were prepared based on the World Health Organization’s (WHO) drinking water quality standard (WHO, 2017). According to WHO, an elevated level for each water quality parameter indicates a decreased water quality (Shobha et al. 2013). The fuzzy rule equation for each parameter was defined as:

$$\begin{cases} 1, & x \leq v \\ Q-x/Q-i & v < x \leq Q \\ 0, & x > Q \end{cases} \tag{3}$$

In Eq. (3), *Q* was determined based on the permissible limit for each parameter in the WHO’s drinking water quality standard (WHO 2017). Table 2 summarizes the limits used based on WHO’s water quality standard for each water quality parameter.

An analytical hierarchy process (AHP) was used in order to overly the fuzzy maps and subsequently preparing the overall drinking water quality index map (Fig. 3). AHP is a multi-criteria decision-making method based on a pair-wise

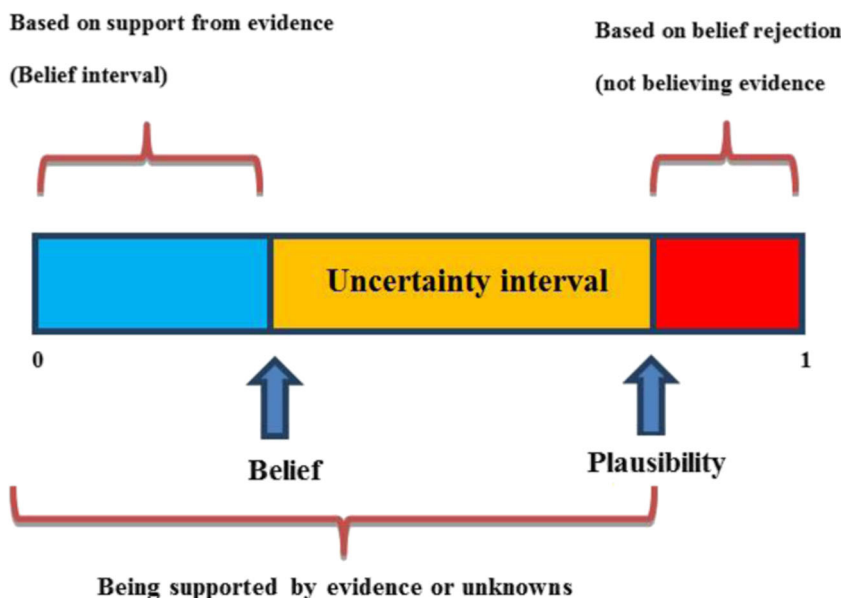
**Table 1** Surface area of different landforms in Firoozabad aquifer

Formation	Description	Area (km <sup>2</sup> )	Area (%)
OMas	Asmari	119.82	21.16
Mn	Mishan	4.17	0.74
Mr	Razak	115.93	20.47
Kb	Bangestan	32.90	5.81
EK	Pabdeh-Gurpi	5.39	0.95
Mgs	Gachsaran	6.43	1.14
Qt2	Quaternary	281.62	49.73
Total		566.25	100.00

**Table 2** Fuzzy limits considered for each water quality parameter based on WHO’s drinking water quality standard

Parameter	Limit
Ca <sup>2+</sup> (mg L <sup>-1</sup> )	200
Cl <sup>-</sup> (mg L <sup>-1</sup> )	200
Mg <sup>2+</sup> (mg L <sup>-1</sup> )	150
Th (mg L <sup>-1</sup> )	500
Na <sup>+</sup> (mg L <sup>-1</sup> )	200
SO <sub>4</sub> <sup>2-</sup> (mg L <sup>-1</sup> )	200
TDS (mg L <sup>-1</sup> )	500
EC (dS m <sup>-1</sup> )	0.3

**Fig. 3** Uncertainty in the belief space in Dempster–Shafer theory



comparison matrix (Saaty 1980). The matrix is called consistent if the transitivity (Eq. (4)) and reciprocity (Eq. (5)) rules are respected:

$$b_{ij} = b_{im} \cdot b_{mj} \tag{4}$$

$$b_{ij} = \frac{1}{b_{ji}} \tag{5}$$

where  $i, j$ , and  $m$  are all of the alternatives of the matrix.

In a consistent matrix (Eq. (5)), all the comparisons  $b_{ij}$  obey the equality  $a_{ij} = p_i/p_j$ , where  $p_i$  is the priority of the alternative  $i$ . A detailed description of the fuzzy AHP method can be found in Bellman and Zadeh (1970).

**Dempster–Shafer**

**The detection framework**

Dempster–Shafer theory (DST) was used to prepare water quality maps with different confidence levels. The DST is based on the beliefs that are concluded from evidence such that the belief structure of the control theory is related to the classical probability model (Shafer 1976). Suppose,  $\theta$  is a finite set of elements, where an element can be a hypothesis,

an objective, or a system status.  $\theta$  set is called the detection framework that can be determined by  $\Omega(\theta)$ :

$$\begin{aligned} \theta &= \{a, b, c\} \text{ and } \Omega(\theta) \\ &= \{\phi\{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\} \end{aligned} \tag{6}$$

$\Phi$  is an empty set that denotes the perfect status of the system.  $A = \{a, b\}$  is the subset of  $\theta$ , indicating  $A \subset \theta$ . Therefore,  $A$  presents a system malfunction in  $a$  or  $b$ , and  $\theta$  represents the system malfunction in  $a, b$ , or  $c$  (Shafer 1976).

**Mass function, focal elements, and core elements**

In order to show the reliability of any subset of the decision-making framework, the mass function is defined. The mass function is shown by  $m$  and is defined as Eqs. (7), (8), and (9) (Shafer 1976):

$$m : \Omega(\theta) \rightarrow [0, 1] \tag{7}$$

$$\Omega(\theta) = 0 \tag{8}$$

$$\sum_{A \subset \Omega(\theta)} m(A) = 1 \tag{9}$$

where  $m$  mass function is called a function of the basic probability assignment (BPA).

**Table 3** Meaning of different belief intervals in the DST model

Interval	Description
[0, 0]	Certainly no support is made of the hypothesis
[1, 1]	The theory should be fully supported
[0, 1]	The information is absolutely unknown, and no interpretation can be made based on the evidence
Interval	Supporting the theory and decision-making
[0.3, 1]	Intend to support the theory
[0, 0.6]	The support from the theory is decreased

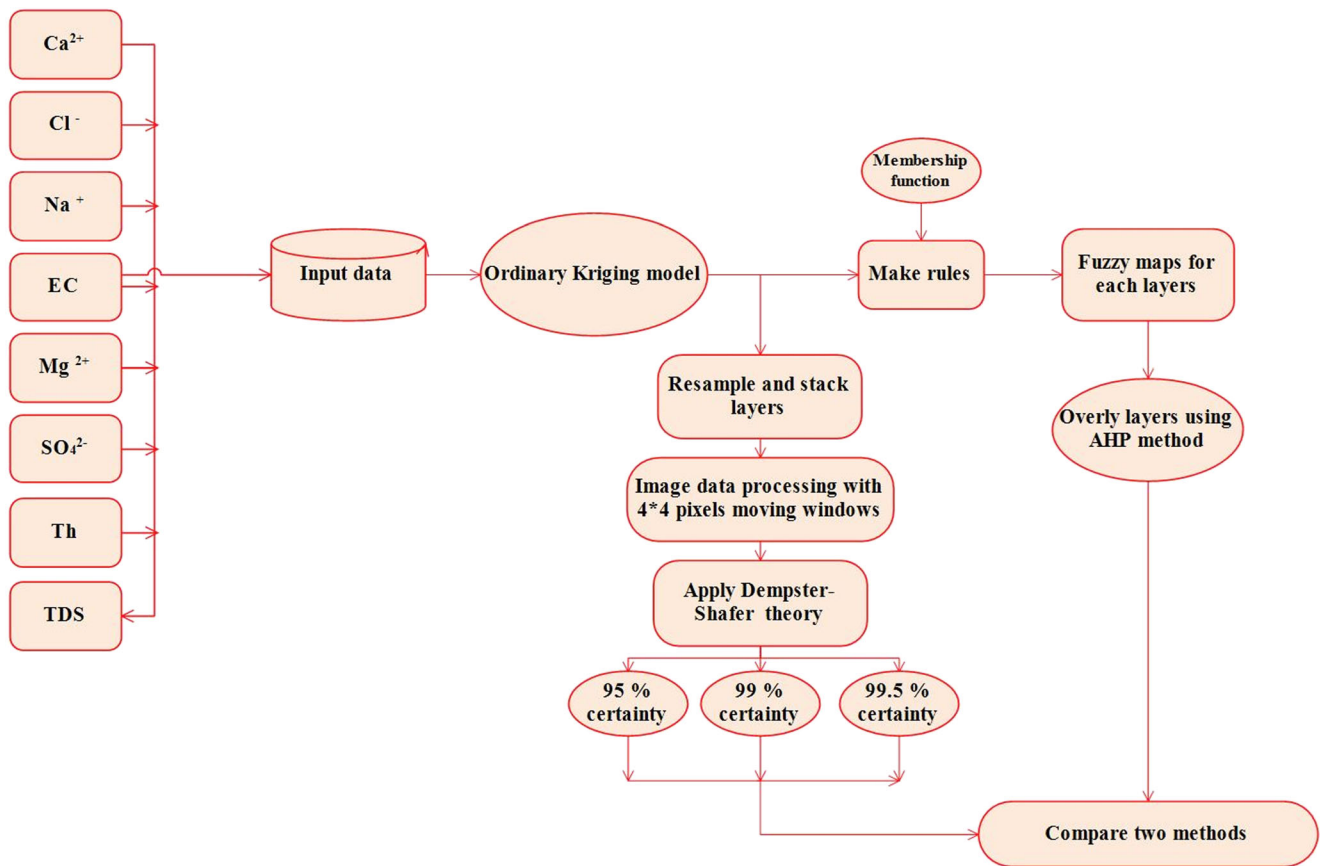


Fig. 4 Flowchart used for generating the groundwater quality maps based on fuzzy-AHP and DST methods

Equation (8) means that the scope of this function is on the entire set of the detection framework that is  $(\Omega^2)$  and its range is  $[0, 1]$ . The function  $m(A)$  represents the proportion of the share owned by of the set  $A$  from all the relevant and available evidence and support for the claim about a particular element of  $\theta$ .

In evaluating a condition in which the system suffers from a defect,  $m(A)$  can be represented as a degree of belief that has been obtained as a result of the observations related to a particular defect. Different evidence and information create various degrees of belief with respect to the manifested defect. Each subset  $A$  from  $\theta$  is called a focal element such that  $m(A) > 0$ . Moreover,  $c$  is defined as a core element of the mass function in  $\theta$  as represented in Eq. (10) (Shafer 1976).

$$c = U_{m(A) \neq 0} A \tag{10}$$

The belief function is defined as Eq. (11):

$$Bel(A) = \sum_{B \subset A} m(B) \text{ and } Bel : \Omega(\theta) \rightarrow [0, 1] \tag{11}$$

The plausibility function is determined as Eq. (12):

$$Pl(A) = 1 - Bel(\bar{A}) = \sum_{B \cap A \neq \emptyset} m(B) \text{ where } Pl : \Omega(\theta) \rightarrow [0, 1] \tag{12}$$

where  $Bel(A)$  function measures the total amount of probability that should be among the set  $A$  elements. As shown in Fig.

3,  $Bel(A)$  denotes the significant certainty of the belief  $A$ , or in other words, it shows the lower limit for probability  $A$ . The function  $Pl(A)$  measures the maximum amount of probability that can be distributed among the elements of  $A$ .  $Pl(A)$  describes the degree of the general belief related to  $A$  and is regarded as the upper limit function for probability  $A$  (Shafer 1976). Belief Interval  $[Bel(A), Pl(A)]$  reflects the uncertainty belief space and the space size  $Pl(A) - Bel(A)$  describes the unknown related to set  $A$  (Fig. 3). Table 3 summarizes the meaning of different belief intervals in DST (Shafer 1976).

Table 4 Concentrations of different quality parameters obtained from 150 monitoring wells during for the study period

Parameter	Range	Mean	Standard deviation
Ca <sup>2+</sup> (mg L <sup>-1</sup> )	1.5–596	195	89
Cl <sup>-</sup> (mg L <sup>-1</sup> )	25.12–437	84	45
Na <sup>+</sup> (mg L <sup>-1</sup> )	1.82–458	51	45
Mg <sup>2+</sup> (mg L <sup>-1</sup> )	3.2–569	137	80
SO <sub>4</sub> <sup>2-</sup> (mg L <sup>-1</sup> )	0.12–584	137	77
TH (mg L <sup>-1</sup> )	72.5–437	180	77
TDS (mg L <sup>-1</sup> )	55–945	295	117
EC (dS m <sup>-1</sup> )	0.39–1.75	0.71	0.15

**Table 5** Parameters in different semivariogram models employed for generating interpolation maps using the kriging method for different water quality constituents

Parameter	Circular				Gaussian				Spherical				Exponential			
	Nugget	Sill	Nugget/Sill	RMSE	Nugget	Sill	Nugget/Sill	RMSE	Nugget	Sill	Nugget/Sill	RMSE	Nugget	Sill	Nugget/Sill	RMSE
TDS	0.73	0.82	0.89	0.80	0.24	1.01	0.24	0.79	0.41	0.79	0.52	0.80	0.55	0.66	0.83	0.81
Th	0.62	0.68	0.91	0.80	0.53	0.58	0.91	0.80	0.61	0.89	0.69	0.80	0.69	0.74	0.93	0.82
Ca <sup>2+</sup>	0.10	0.89	0.11	0.92	0.30	0.98	0.31	0.91	0.05	0.94	0.05	0.90	0.21	0.77	0.27	0.93
Mg <sup>2+</sup>	0.43	0.91	0.47	0.61	0.33	0.94	0.35	0.60	0.43	0.93	0.46	0.60	0.54	0.82	0.66	0.62
Na <sup>+</sup>	0.42	0.53	0.79	0.90	0.80	0.85	0.94	0.90	0.87	0.98	0.89	0.90	0.89	0.97	0.92	0.91
Cl <sup>-</sup>	0.49	0.77	0.64	0.77	0.45	0.88	0.51	0.76	0.48	0.78	0.62	0.78	0.58	0.70	0.83	0.78
SO <sub>4</sub> <sup>2-</sup>	0.54	0.63	0.86	0.91	0.33	0.65	0.51	0.89	0.51	0.56	0.91	0.91	0.62	0.75	0.83	0.92
EC	0.43	0.60	0.72	0.56	0.08	0.87	0.09	0.53	0.41	0.60	0.68	0.56	0.51	0.58	0.88	0.62

The plausibility function is related to the belief function by a mediator function called doubt function as shown in Eq. (13) (Shafer 1976):

$$\begin{cases} Pl(A) = 1 - Bel(\bar{A}) \\ Pl(A) = 1 - Doubt(A) \end{cases} \quad (13)$$

where  $\bar{A}$  is the complement of  $A$  and  $Doubt(A)$  represent the doubt function for set  $A$ .

The methodology used in this research is summarized in Fig. 3.

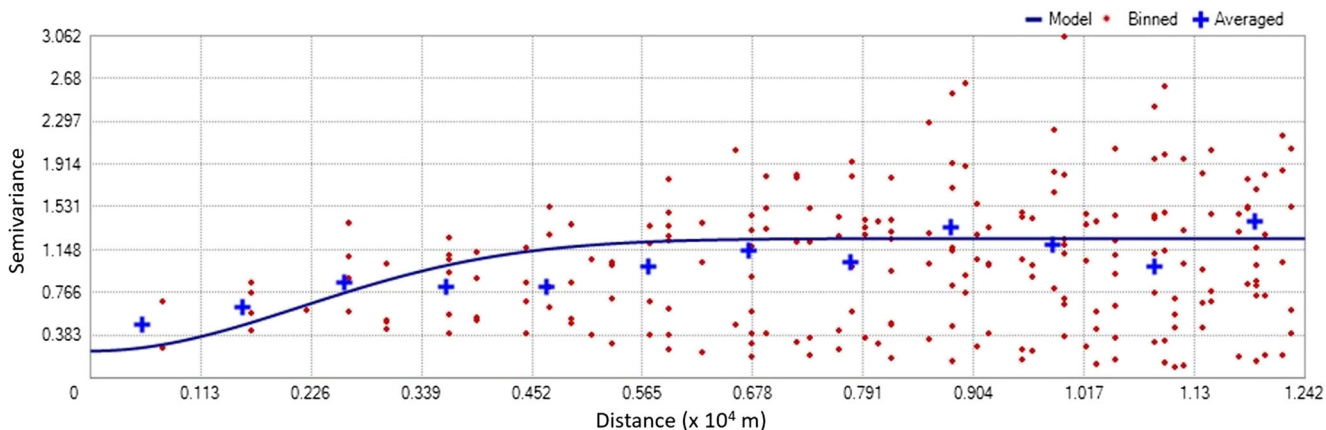
To date, there is no specific functional toolbox for implementing the DST in GIS and remote sensing packages. In this study, a model based on DST was implemented using imprecise probability propagation (IPP) Toolbox in MATLAB R2017a. As the design of this toolbox is based on linear matrices, some modifications were performed to be able to process raster image data sets. As all input image data needed to be of the same size, hence images with a pixel size of 1 km × 1 km were prepared. In the next step, in order to increase a

high computation speed, windows of 4 × 4 pixels were selected and subjected to the DST model analysis. This loop was continued until the DST model was implemented on entire pixels.

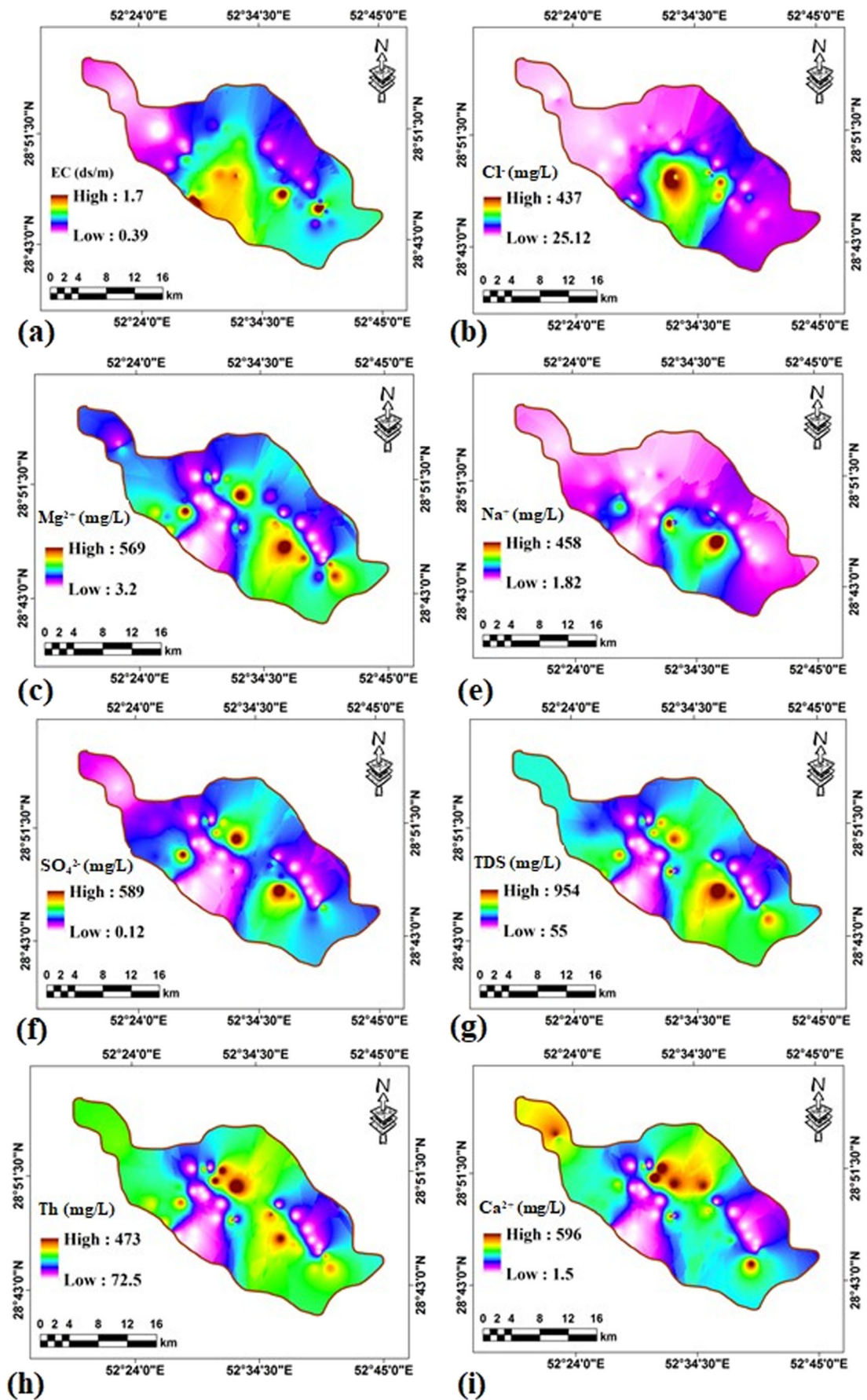
### Model comparison

In order to evaluate the effects of salt domes on aquifer’s water, groundwater quality maps based on both fuzzy AHP and DST models were prepared. In both models, the overall water quality index for models was defined based on WHO’s drinking water quality standard (Table 2). Figure 4 represents the flowchart used for comparing the water quality maps generated by the fuzzy AHP and DST models. The DST model could generate water quality maps with three confidence levels of 99.5%, 99%, and 95%.

Water quality data from 150 monitoring wells were used in this study. The data from 110 monitoring wells were used for generating the water quality maps based on both the DST and fuzzy AHP methods (train point in Fig. 2b). Water quality data



**Fig. 5** Example semivariogram with a Gaussian model fitted to the TDS observations in the study area





**Fig. 6** Maps of different water quality constituents including EC (a), Mg<sup>2+</sup> (b), Na<sup>+</sup> (c), Cl<sup>-</sup> (d), SO<sub>4</sub><sup>2-</sup> (e), TDS (f), Th (g), and Ca<sup>2+</sup> (h). The maps were generated by interpolating the measured values across the aquifer based on the ordinary kriging interpolation method with a Gaussian semivariance model

from 40 different monitoring wells (test point in Fig. 2b) were used to evaluate the water quality maps by the two models.

Root mean squared error (RMSE) was used to evaluate the performance of different interpolation models. RMSE can be defined as:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N [S(x_i) - \hat{S}(x_i)]^2} \tag{14}$$

where  $\hat{S}(x_i)$  is the estimated value,  $S(x_i)$  is the observed value, and  $N$  is the number of values in the dataset.

## Results and discussion

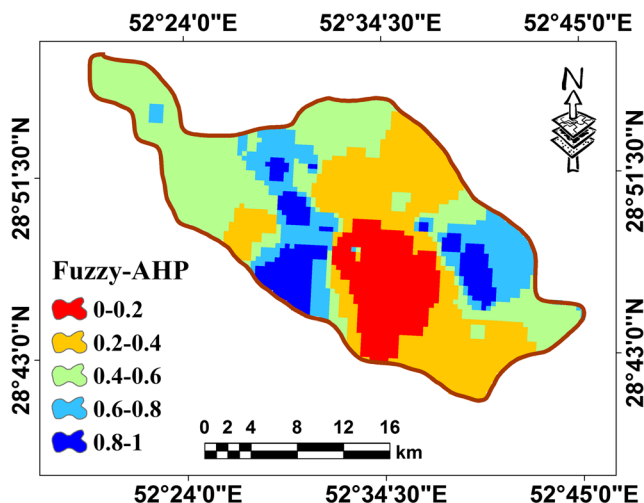
### Geostatistical analysis

Table 4 summarizes the range of concentrations observed for different water quality parameters in 150 monitoring wells during September 23, 2017–March 21, 2017.

High spatial variations were observed for most of water quality parameters. This suggests the effects of specific geological units on water quality of the aquifer. The ordinary kriging interpolation method was used to evaluate the spatial variations of water quality parameters across the aquifer. Table 5 represents the RMSE between the measured and fitted semivariogram models in 40 test points for each water parameter. Among different theoretical models, the Gaussian model was found as the best fit. Table 5 also represents the calculated semivariogram parameters for each water parameter. The nugget values indicate the magnitude of spatial variability (including measurement errors) in data at scales smaller than sampling distances, and the sill values reflect to the total variance observed in data. The difference between the sill and the nugget (i.e., partial sill) for each parameter indicates the amount of variance that could be explained by the observations obtained



**Fig. 7** Weights assigned to each water quality constituent for creating the overall water quality index in the fuzzy AHP model

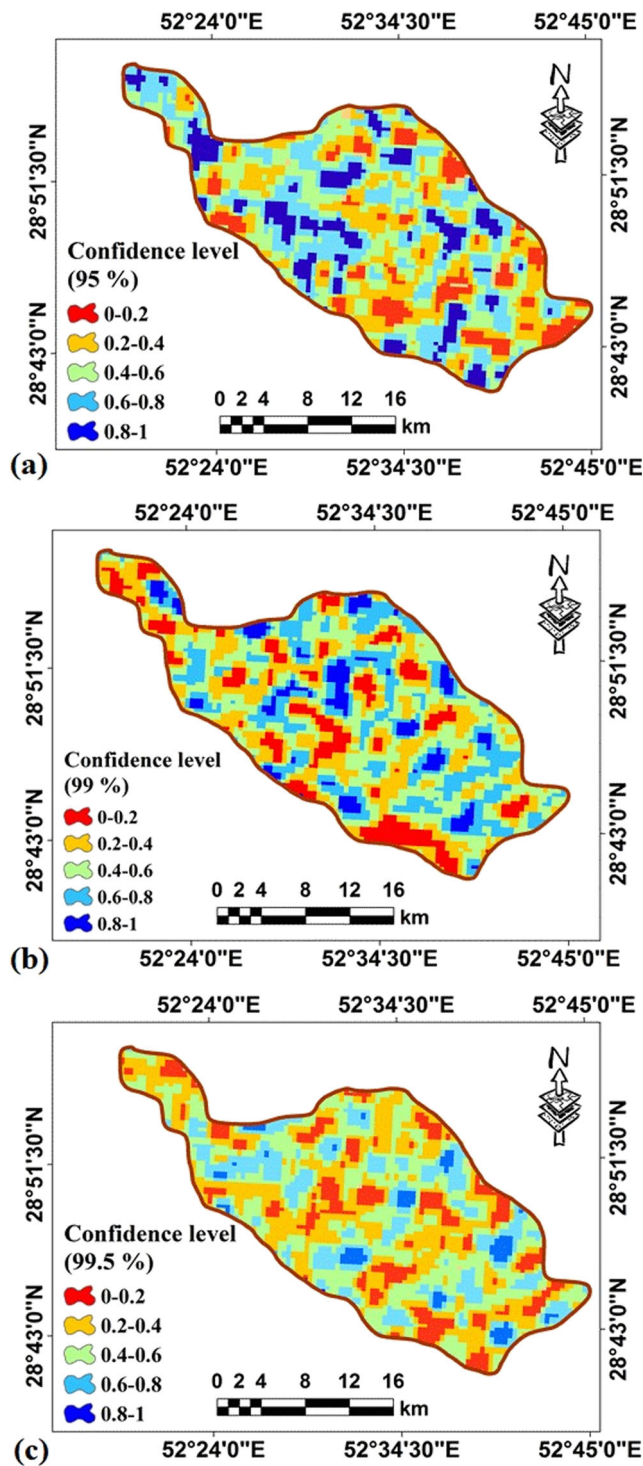


**Fig. 8** The average water quality index map generated based on the fuzzy AHP method for the period of September 23, 2017–March 21, 2017. A low water quality index (red) indicates a poor groundwater quality and a high water quality index (blue) indicates a high groundwater quality

from the sampling points with specific distances from each other (Karl and Maurer 2010; Oliver and Webster 1990). The nugget to sill ratio indicates the fraction of the total variation that cannot be described by observed spatial dependence of the variable (Karl and Maurer 2010). Figure 5 shows an example Gaussian semivariogram model fitted to the observed TDS data.

As shown in Table 5, the Gaussian model typically exhibited smaller nugget to sill ratios and yielded relatively lower RMSE values for all water quality constituents compared with other models. This means that the Gaussian exhibited higher spatial correlations at longer distances and less interpolation errors for all water quality constituents. Therefore, the ordinary kriging method based on a Gaussian semivariogram model was used for analyzing the spatial variations of different water quality constituents in the study area (Fig. 6).

Figures 6a and 4b show the distribution of EC (ranging between 0.39 and 1.7 ds m<sup>-1</sup>) and Cl<sup>-</sup> (ranging between 25.12 and 437 mg L<sup>-1</sup>) in the study area, respectively. The lowest concentrations of these water quality constituents occurred in northern and northwestern areas. Concentration Mg<sup>2+</sup> ranged between 3.2 and 569 mg L<sup>-1</sup> where the highest



**Fig. 9** The overall water quality index based on the Dempster–Shafer model with 99.5% (a) confidence level, 99% confidence level (b), and 95% confidence level (c) for the period of September 23, 2017–March 21, 2017

and the lowest concentrations occurred in southern and western parts of the aquifer, respectively (Fig. 6c).  $\text{Na}^+$  in the entire study area except for small parts of western parts was almost less than  $10 \text{ mg L}^{-1}$  (Fig. 6d). The concentration of  $\text{SO}_4^{2-}$  was

**Table 6** The percentage of the area of each class of water quality for each method

Overall water quality	DST			Fuzzy AHP
	Confidence level			
	95%	99%	99.5%	
Poor (0–0.2)	12.64	13.45	13.84	13.26
Low (0.2–0.4)	26.21	26.54	31.58	29.77
Moderate (0.4–0.6)	25.37	30.23	32.35	34.00
Good (0.6–0.8)	24.93	22.35	18.51	14.52
High (0.8–1)	10.84	7.43	3.71	8.45

between  $0.12$  and  $589 \text{ mg L}^{-1}$  with the lowest concentrations in western and northern areas (Fig. 6e). Nearly similar spatial variations were for TDS and Th concentrations for which the lowest concentrations occurred in western areas (Fig. 6f and g). The spatial variation for  $\text{Ca}^{2+}$  was slightly different as its highest concentrations (more than  $500 \text{ mg L}^{-1}$ ) occurred in northern and northeastern parts (Fig. 6h).

### Comparing the overall water quality index based on fuzzy-AHP and DST methods

Equation (3) was used to define MF of between  $[0, 1]$  for each water quality parameter. In each MF, the very low and very high concentrations for each constituent were close to 1 and 0, respectively. Hence, for  $\text{Ca}^{2+}$ ,  $\text{Cl}^-$ ,  $\text{Mg}^{2+}$ , Th,  $\text{Na}^+$ , EC,  $\text{SO}_4^{2-}$ , and TDS concentrations of greater than 200, 200, 150, 500, 200, 3000, 200, and  $500 \text{ mg L}^{-1}$ , respectively, MFs = 0 were considered. Figure 7 shows the weight assigned to each water quality parameter for generating the overall water quality index based on the fuzzy AHP method (Mokarram and Sathyamoorthy 2016).

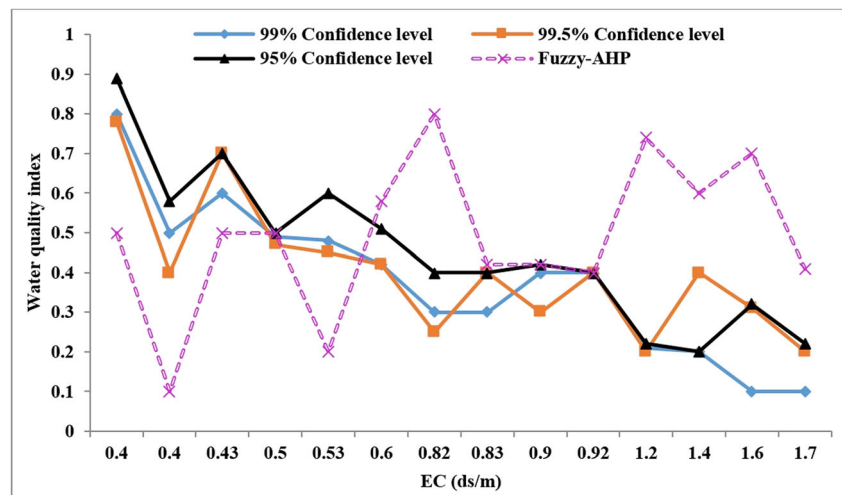
Figure 8 represents the overall water quality map generated based on the fuzzy AHP system. The map shows the average index during for the period of September 23, 2017 to March 21, 2017.

As seen, the fuzzy-AHP water quality index ranged between 0 and 0.94. In Fig. 8 the blue color (high water quality index) indicates the areas with a high groundwater quality; the red color (low water quality index) indicates areas with a low groundwater quality. As seen, central areas towards the southern parts of the aquifer showed the lowest drinking water quality index, whereas small areas in the eastern, western, and northern parts showed a high water quality index.

Figure 9 shows the water quality maps with different confidence levels generated based on the DST model.

Generally five water classes of poor, low, moderate, good, and high with the overall water quality index ranges of (0, 0.2), (0.2, 0.4), (0.4, 0.6), (0.6, 0.8), and (0.8, 1), respectively, were defined in both DST and fuzzy AHP

**Fig. 10** EC values and their corresponding water quality indices determined based on fuzzy-AHP and Dempster–Shafer theory with confidence levels of 95%, 99%, and 99.5% in 14 randomly select groundwater quality stations



models (Table 6). Table 6 represents the surface area of the study site covered by each water quality class based on the fuzzy AHP and DST models.

Based on fuzzy AHP model, the majority of the study area (34%) was ranked with as moderate (0.4–0.6) water quality. In Dempster–Shafer models with different confidence levels, the majority of the study area was ranked as moderate or low (Table 6). Both fuzzy AHP and DST models recognized the western areas of the aquifer with a quaternary soil type as the regions with the highest water quality index.

In general, the results of the DST model showed that (Fig. 9) by decreasing the confidence level, the areas with higher

water quality indices tended to increase. In comparison, the fuzzy model had a faster computation speed. However, the DST model was capable of generating water quality maps with different confidence levels. Therefore, the DST model can be used for generating management and planning water quality zoning maps with preferred confidence levels based on the economic conditions and importance of the study area.

Formation type could influence the water quality of an aquifer. Figure 2a was used to evaluate the effects of geological formation type in the region on water quality of the aquifer. As seen from Fig. 2a, quaternary sediments cover a majority of the study area. The coarse-grained sediments have resulted in a high porosity in and thus a high permeability in these areas (Panno and Hackley 2010). Hence, relatively higher water quality indices could be observed in areas with quaternary sediments. In contrast, the areas with an Asmari (OMas) formation composed of limestone, salt domes, dolomitic limestone, and marl (Fig. 2a) contain elevated levels of calcium, sodium, magnesium, and chloride ions (Panno and Hackley 2010). Razak formation composed of mainly conglomerates, marls, and shale with interactions of limestone (Ghazavi and Emami 2017) can also contribute to the elevated concentrations of minerals in central and southern areas. Hence, a poor water quality index could be observed in these areas. In areas with a Bangestan formation, water quality is better than the Asmari formation, which could be due to less salt content of the Bangestan (Panno and Hackley 2010).

Comparison of the geological maps and the water quality maps prepared based on the DST and fuzzy-AHP models indicates that in areas with Kb (Bangestan) and Mm (Mishan) formations, both methods yielded moderate water quality indices. In southeastern, and central towards the southern areas, the OMas formation could affect the quality of groundwater, and therefore both fuzzy-AHP and DST models showed low water quality indices in these regions. The elevation of OMas formation (Asmari) in the eastern and northeastern areas is

**Table 7** Comparison of water quality index in DST and fuzzy AHP models based on EC values in randomly selected 14 observation points

Number	X	Y	EC (dS m <sup>-1</sup> )	DST			Fuzzy-AHP
				Confidence level			
				99%	99.5%	95%	
1	52.37°	28.93°	0.4	0.8	0.78	0.89	0.5
2	52.68°	28.73°	0.4	0.5	0.4	0.58	0.1
3	52.43°	28.86°	0.43	0.6	0.7	0.7	0.5
4	52.53°	28.90°	0.5	0.49	0.47	0.5	0.5
5	52.63°	28.86°	0.53	0.48	0.45	0.6	0.2
6	52.64°	28.83°	0.6	0.42	0.42	0.51	0.58
7	52.49°	28.80°	0.82	0.3	0.25	0.4	0.8
8	52.55°	28.72°	0.83	0.3	0.4	0.4	0.42
9	52.54°	28.84°	0.9	0.4	0.3	0.42	0.42
10	52.66°	28.69°	0.92	0.4	0.4	0.4	0.4
11	52.66°	28.76°	1.2	0.21	0.2	0.22	0.74
12	52.51°	28.78°	1.4	0.2	0.4	0.2	0.6
13	52.50°	28.76°	1.6	0.1	0.31	0.32	0.7
14	52.52°	28.80°	1.7	0.1	0.2	0.22	0.41

about 1625 m higher than the central areas, resulting in a much higher depth of groundwater table in these areas. This would reduce the influence of OMAs formation on groundwater quality in eastern and northeastern areas compared with central and southern areas as the steep slope of the ground results in a decreased percolation of surface runoff and soil layers beneath the OMAs and could also decrease the percolation rate of minerals to the groundwater.

In order to compare the Fuzzy-AHP and DST models, 14 monitoring wells among the 40 test wells (test points in Fig. 2b) were randomly selected and the predicted water quality indices by fuzzy-AHP and DST models at this points were compared with measured EC values. The EC values and water quality indices were determined based on the two models which are shown in Fig. 10 and Table 7.

According to Fig. 10 and Table 5, by increasing EC water quality parameters, both methods showed decreased water quality indices. Moreover, a comparison of the two methods showed that the DST model was better for water quality and had better integration. Also, a comparison of EC values in Fig. 10 indicates that the results of the DST model were more accurate than the fuzzy-AHP method in determining the water quality index.

## Conclusions

In this study, fuzzy AHP and Dempster Shafer theory were used to determine the drinking water quality maps for an aquifer. The results of fuzzy-AHP method showed that most parts of the study areas had a moderate water quality index, and eastern and western parts had the best drinking water quality index. Compared with the fuzzy AHP model, the Dempster Shafer model had the capability of generating maps with various confidence levels. However, in comparison with the Dempster–Shafer model, the Fuzzy AHP model had faster computation speed. The results of the DST model showed that by decreasing the confidence level, the areas with higher water quality indices increased. Comparison of the two models showed that the DST model could more effectively handle the uncertainty in the data and thus generated more accurate water quality maps. One of the advantages of the DST model is the capability of computing the outputs with different confidence levels. Therefore, this model can be used for generating water quality zoning maps with confidence levels that best suits the economic conditions and importance of the study area.

**Acknowledgments** The authors would like to thank the personnel of Agricultural Jihad of Fars province for their kind assistance.

**Funding information** This study was financially supported by Shiraz University (238726-116).

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Amer, Reda, Robert Ripperdan, Tao Wang, and John Encarnación. 2012. “Groundwater quality and management in arid and semi-arid regions: case study, Central Eastern Desert of Egypt.” *J Afr Earth Sci* 69: 13–25. <https://www.sciencedirect.com/science/article/pii/S1464343X12000696> (August 24, 2018).
- Bellman RE, Zadeh LA (1970) Decision-making in a fuzzy environment. *Manag Sci* 17(4):B-141–B-164. <https://doi.org/10.1287/mnsc.17.4.B141>
- Bieroza MZ, Heathwaite AL, Bechmann M, Kyllmar K, Jordan P (2018) The concentration-discharge slope as a tool for water quality management. *Sci Total Environ* 630:738–749 <https://www.sciencedirect.com/science/article/pii/S0048969718306569> (August 16, 2018)
- Chaabane S, Ben M, Sayadi FF, Brassart E (2008) Color image segmentation based on dempster-shafer evidence theory. In: MELECON 2008—14th IEEE Mediter. Electrotech. Conf., IEEE, pp 862–866 <http://ieeexplore.ieee.org/document/4618544/>
- Cordoba GAC, Tuhovčák L, Tauš M (2014) Using artificial neural network models to assess water quality in water distribution networks. *Procedia Engineering* 70:399–408 <https://www.sciencedirect.com/science/article/pii/S1877705814000472> (September 15, 2018)
- Cumming Geoff, Fiona Fidler. 2005 “Interval estimates for statistical communication: problems and possible solutions.” *IASE/ISI Satellite*: 1–7
- Dempster AP (1968) A generalization of Bayesian inference. *J R Stat Soc Ser B* 30:205–247 <https://www.jstor.org/stable/2984504>. Accessed 4 Jan 2019
- Dempster AP (2008) Upper and lower probabilities induced by a multivalued mapping. *Stud Fuzziness Soft Comput* 219:57–72
- Diamantini E, Lutz SR, Mallucci S, Majone B, Merz R, Bellin A (2018) Driver detection of water quality trends in three large European river basins. *Sci Total Environ* 612:49–62 <https://www.sciencedirect.com/science/article/pii/S004896971732171X> (August 16, 2018)
- ESRI (2016) ArcMap 10.5, Redlands, California, USA, : Esri Inc. <https://www.esri.com/en-us/arcgis/about-arcgis/overview>. Accessed 4 Jan 2019
- Fars Regional Water Authority (FRWA) (2016) <https://www.frw.ir/>. Accessed 4 Jan 2019
- Gazzaz NM, Yusoff MK, Aris AZ, Juahir H, Ramli MF (2012) Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors. *Mar Pollut Bull* 64(11):2409–2420 <https://www.sciencedirect.com/science/article/pii/S0025326X12004043>
- Gershman SJ, Blei DM (2012) A tutorial on Bayesian nonparametric models. *J Math Psychol* 56(1):1–12
- Ghazavi M, and Emami N (2017) “Landslides and slope failures due to saturated soft soil: a case study.” In *Soft soil engineering*, pp. 103–109. Routledge, . <https://doi.org/10.1201/9780203739501>
- Helton JC (1997) Uncertainty and sensitivity analysis in the presence of stochastic and subjective uncertainty. *J Stat Comput Simul* 57:3–76
- Islam, Abu Reza Md. Towfiqul, Nasir Ahmed, Md. Bodrud-Doza, and Ronghao Chu. 2017. “Characterizing groundwater quality ranks for drinking purposes in Sylhet District, Bangladesh, using entropy method, spatial autocorrelation index, and geostatistics.” *Environ Sci Pollut Res* 24(34): 26350–26374. <https://doi.org/10.1007/s11356-017-0254-1> (January 15, 2019).

- Karl JW, Maurer BA (2010) Spatial dependence of predictions from image segmentation: a variogram-based method to determine appropriate scales for producing land-management information. *Eco Inform* 5(3):194–202. <https://doi.org/10.1016/j.ecoinf.2010.02.004>
- Khalil, Abedalrazq, Mohammad N. Almasri, Mac McKee, and Jagath J. Kaluarachchi. 2005. “Applicability of statistical learning algorithms in groundwater quality modeling.” *Water Resour Res* 41(5). <https://doi.org/10.1029/2004WR003608> (August 16, 2018).
- Khashei-Siuki A, Sarbazi M (2015) Evaluation of ANFIS, ANN, and geostatistical models to spatial distribution of groundwater quality (case study: Mashhad Plain in Iran). *Arab J Geosci* 8(2):903–912. <https://doi.org/10.1007/s12517-013-1179-8> August 16, 2018
- Kuo Y-M, Liu C-W, Lin K-H (2004) Evaluation of the ability of an artificial neural network model to assess the variation of groundwater quality in an area of blackfoot disease in Taiwan. *Water Res* 38(1):148–158 <https://www.sciencedirect.com/science/article/pii/S0043135403005013> (August 16, 2018)
- Lallahem S, and Hani A (2017) “Artificial neural networks for defining the water quality determinants of groundwater abstraction in coastal aquifer.” In *AIP Conf. Proc.*, AIP Publishing LLC, 20013. <https://doi.org/10.1063/1.4976232>.
- López-Granados F, Jurado-Expósito M, Atenciano S, García-Ferrer A, Sánchez de la Orden M, García-Torres L (2002) Spatial variability of agricultural soil parameters in Southern Spain. *Plant Soil* 246(1): 97–105. <https://doi.org/10.1023/A:1021568415380> January 5, 2019
- Mokarram M, Sathyamoorthy D (2016) Investigation of the relationship between drinking water quality based on content of inorganic components and landform classes using fuzzy AHP (case study: South of Firozabad, West of Fars Province, Iran). *Drink Water Eng Sci* 9(2): 57–67 <https://www.drink-water-eng-sci.net/9/57/2016/> (August 16, 2018)
- Nadiri AA, Gharekhani M, Khatibi R, Moghaddam AA (2017) Assessment of groundwater vulnerability using supervised committee to combine fuzzy logic models. *Environ Sci Pollut Res* 24(9): 8562–8577 <http://www.ncbi.nlm.nih.gov/pubmed/28194673> (January 15, 2019)
- Najah A, El-Shafie A, Karim OA, El-Shafie AH (2013) Application of artificial neural networks for water quality prediction. *Neural Comput & Applic* 22(S1):187–201. <https://doi.org/10.1007/s00521-012-0940-3>
- Neshat A, Pradhan B (2015) Risk assessment of groundwater pollution with a new methodological framework: application of Dempster–Shafer theory and GIS. *Nat Hazards* 78(3):1565–1585. <https://doi.org/10.1007/s11069-015-1788-5> August 16, 2018
- Oliver MA, Webster R (1990) Kriging: a method of interpolation for geographical information systems. *Int J Geogr Inf Syst* 4(3):313–332. <https://doi.org/10.1080/02693799008941549> August 9, 2018
- Panno S, Hackley K. 2010. “Geologic influences on water quality.” *Geology of Illinois*: 337–50.
- Rahimi D, Mokarram M (2012) *2 International Journal of Environmental Sciences Assessing the groundwater quality by applying fuzzy logic in gis environment—a case study in Southwest Iran*. Integrated Publishing Association. <http://www.indianjournals.com/ijor.aspx?target=ijor:ijes&volume=2&issue=3&article=061> (August 16, 2018).
- Rahmati O, Melesse AM (2016) Application of Dempster–Shafer Theory, spatial analysis and remote sensing for groundwater potentiality and nitrate pollution analysis in the semi-arid region of Khuzestan, Iran. *Sci Total Environ* 568:1110–1123 <https://linkinghub.elsevier.com/retrieve/pii/S004896971631350X> (August 16, 2018)
- Rowles LS et al (2018) Perceived versus actual water quality: community studies in Rural Oaxaca, Mexico. *Sci Total Environ* 622(623):626–634 <https://www.sciencedirect.com/science/article/pii/S0048969717333673> (August 16, 2018)
- Saaty TL (1980) *The analytic hierarchy process: planning, priority setting, resource allocation*. McGraw-Hill International Book Co. [https://books.google.com/books/about/The\\_Analytic\\_Hierarchy\\_Process.html?id=Xxi7AAAAIAAJ](https://books.google.com/books/about/The_Analytic_Hierarchy_Process.html?id=Xxi7AAAAIAAJ).
- Sarkar A, Pandey P (2015) River water quality modelling using artificial neural network technique. *Aquatic Procedia* 4:1070–1077 <https://www.sciencedirect.com/science/article/pii/S2214241X15001364> (September 15, 2018)
- Shafer G (1976) Dempster-Shafer theory. *Int J Approx Reason* 21(2):1–2
- Shobha G, Gubbi J, Raghavan KS, Kaushik LK, Palaniswami M (2013) A novel fuzzy rule based system for assessment of ground water potability: A case study in South India. *Magnesium (Mg)* 30(35-41): 10. <https://www.semanticscholar.org/paper/A-novel-fuzzy-rule-based-system-for-assessment-of-%3A-Shobha-Gubbi/d7aece145da74107fb10986a452101e5e15b9532>
- Todorov D, Driscoll CT, Todorova S, Montesdeoca M (2018) Water quality function of an extensive vegetated roof. *Sci Total Environ* 625:928–939 <https://www.sciencedirect.com/science/article/pii/S0048969717335118> (August 16, 2018)
- Venkatramanan S, Chung SY, Rajesh R, Lee SY, Ramkumar T, Prasanna MV (2015) Comprehensive studies of hydrogeochemical processes and quality status of groundwater with tools of cluster, grouping analysis, and fuzzy set method using GIS platform: a case study of Dalcheon in Ulsan City, Korea. *Environ Sci Pollut Res* 22(15): 11209–11223 <http://www.ncbi.nlm.nih.gov/pubmed/25779109> (January 15, 2019)
- Villa-Achupallas M, Rosado D, Aguilar S, Galindo-Riaño MD (2018) Water quality in the tropical Andes hotspot: the Yacuambi River (Southeastern Ecuador). *Sci Total Environ* 633:50–58 <https://www.sciencedirect.com/science/article/pii/S0048969718309240> (August 16, 2018)
- World Health Organization (WHO) (2017) *Guidelines for drinking-water quality: fourth edition incorporating the first addendum*. <http://www.ncbi.nlm.nih.gov/pubmed/28759192> (October 24, 2018).
- Wu Z, Wang X, Chen Y, Cai Y, Deng J (2018) Assessing river water quality using water quality index in Lake Taihu Basin, China. *Sci Total Environ* 612:914–922 <https://www.sciencedirect.com/science/article/pii/S0048969717323148> (August 16, 2018)
- Zadeh LA (1965) Fuzzy sets. *Inf Control* 8(3):338–353

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.