Model. Earth Syst. Environ. (2016) 2:3
DOI 10.1007/s40808-015-0059-5

CrossMark

ORIGINAL ARTICLE

Modeling of multiple regression and multiple linear regressions for prediction of groundwater quality (case study: north of Shiraz)

M. Mokarram[1]

Received: 29 October 2015 / Accepted: 17 November 2015 / Published online: 12 December 2015

Abstract The aim of the study area was investigation groundwater quality and determination of relationship between effective parameters in groundwater quality in north of Fars province, southeast Iran. For determination of groundwater quality, parameters of calcium (Ca), pH, potassium (k), chlorine (Cl), magnesium (Mg), sodium (Na), electrical conductivity, sulfate (So_4), total dissolved solids (TDS) were used. Using inverse distance weighting spatial distribution of each parameters was determined. Also using multiple linear regressions (MLR) relationship between each of parameters was determined. It was found that in the study area, all of the parameters expect Hco_3 was low value in the north of the study area (high groundwater quality). While the maximum value of parameters was located in south of the study area (low groundwater quality). So the north of the study area was better quality than the south of the study area. The relationship between parameters by MLR showed that Cl and TDS had the strong positive correlation ($r = 0.97$). Also calcium and magnesium showed strong positive correlation ($r = 0.68$) and there was strong positive significant correlation between Ca, Na and Mg with So_4 (about $r = 0.6$).

✉ M. Mokarram
m.mokarram.313@gmail.com; m.mokarram@shirazu.ac.ir

[1] Department of Range and Watershed Management, College of Agriculture and Natural Resources of Darab, Shiraz University, Darab 71946-84471, Iran

Keywords Groundwater · Inverse distance weighting (IDW) · Multiple linear regressions (MLR)

Introduction

Groundwater is the first source of water for human consumption, as well as for agriculture, drinking and industrial uses (Jalali 2009). Increasing knowledge of geochemical processes that control groundwater chemical composition could lead to improved understanding of hydrochemical systems in such areas. Understanding relations can improve management and utilization of the groundwater resource by clarifying relations among groundwater quality, aquifer lithology, and recharge type (Ostovari et al. 2013). Recently by using geography information system (GIS) and sample data of well, prepare interpolation map. The GIS is an effective tool the estimation of the spatial distribution of environmental variables (Ordu and Demir 2009; Rabah et al. 2011). Spatial prediction and surface modeling of water properties has become a common topic in water science research. Interpolation can be undertaken utilizing simple mathematical models (e.g., inverse distance weighting, trend surface analysis and splines), or more complex models (e.g., geostatistical methods, such as kriging) (Negreiros et al. 2011). Ordianary kriging (OK), Universal kriging (UK), inverse distance weighting (IDW) and radial basis function (splines) are four ways to interpolate water properties. Past applications of these methods have given a range of results which have not always been consistent (Falivene et al. 2010).

Some authors found that kriging methods out performed IDW or Radial Basis Function (Splines) (Kravchenko and Bullock 1999; Panagopoulos et al. 2006). However, others showed that kriging was not better than the other two methods (Wollenhaupt et al. 1994; Gotway et al. 1996).

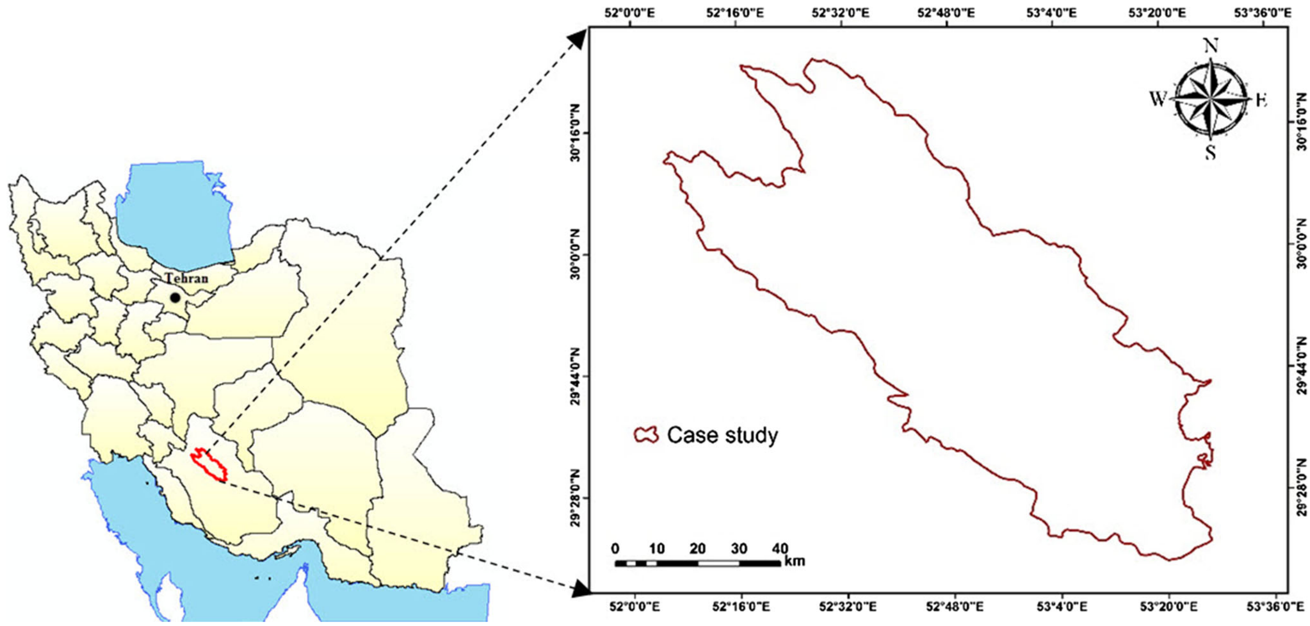


Fig. 1 Position of the study area

IDW (Panagopoulos et al. 2006) and Splines (Kravchenko 2003; Keshavarzi and Sarmadian 2010) are also commonly used classical interpolation methods to analyze the spatial variability of water properties.

The aim of this research was to employ the geostatistic analysis (IDW) for prediction of water quality. Also in order to determination of relationship between each of the groundwater parameters used the multiple linear regressions in North of Fars province, southeast Iran.

Material and method

Case study

This study was located in north of Fars province, southeast of Iran. It has an area of about 6657.25 km^2, and is located at longitude of N 29°18′–30°25′ and latitude of E 52°04′ to 53°26′ (Fig. 1). The altitude of the study area ranges from the lowest of 1530 m to the highest of 3099 m.

In order to predict the variability of groundwater quality, calcium (Ca), pH, potassium (k), chlorine (Cl), magnesium (Mg), sodium (Na), electrical conductivity (EC), sulfate (So_4), total dissolved solids (TDS) were prepared (Table 1) (Fars Regional Water Authority). For preparing the spatial distribution map of each parameters was used 122 wells that local position of the wells show in Fig. 2.

Table 1 Summary statistics of chemical compositions of major ions (mg/l) in the groundwater's of the study area

	Minimum	Maximum	Average	STDV
K	0.03	1.6	0.210492	0.322472
Na	0.12	122.77	14.06016	21.96956
Mg	0.5	85	10.50943	14.94863
Ca	1.5	80	8.906148	11.798
So_4	0.15	56.35	6.923279	8.715891
Cl	0.25	173	21.76066	37.98701
pH	6.85	8.13	7.296148	0.242526
TDS	315	11,540	2040.91	2581.938
EC	467	19,371	3167.746	4166.778

Method

Inverse distance weighted (IDW)

Inverse distance weighting interpolation explicitly implements the assumption that things that are close to one another are more alike than those that are farther apart. To predict a value for any unmeasured location, IDW will use the measured values surrounding the prediction location. Assumes value of an attribute z at any unsampled point is a distance-weighted average of sampled points lying within a defined neighborhood around that unsampled point. Essentially it is a weighted moving average (Burrough et al. 1998):

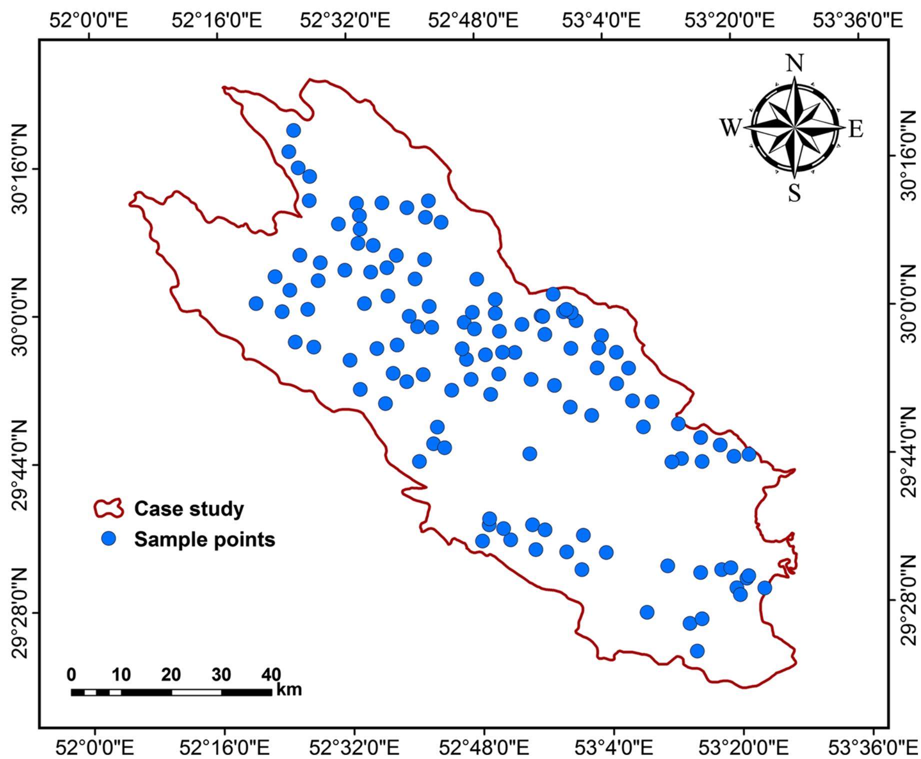


Fig. 2 Local position of the wells in the study area

$$\hat{z}(x_0) = \frac{\sum_{i=1}^{n} z(x_i) d_{ij}^{-r}}{\sum_{i=1}^{n} d_{ij}^{-r}} \quad (1)$$

where x_0 is the estimation point and x_i are the data points within a chosen neighborhood. The weights (r) are related to distance by d_{ij}.

Multiple linear regressions (MLR)

The general purpose of multiple regressions is to learn more about the relationship between several independent or predictor variables and a dependent or criterion variable. The general form of the regression equations is according to Eq. 2:

$$Y = A_0 + A_1 X_1 + A_2 X_2 + \cdots + b_n X_n \quad (2)$$

where Y is the dependent variable, A_0 is the intercept, $A_1 \ldots b_n$ are regression coefficients, and X_1– X_n are independent variables referring to basic soil properties.

The relation between variability of groundwater quality, calcium (Ca), pH, potassium (k), chlorine (Cl), magnesium (Mg), sodium (Na), electrical conductivity (EC), sulfate (So_4), total dissolved solids (TDS) were made related to groundwater quality by constructing regression equations in software of SPSS (2002).

Results

In the study area used from the IDW model for interpolation produces that show in Fig. 3. The lowest and the maximum output in IDW was 1.5 and 79.93 mg/l for Ca. 0.28 and 172.25 mg/l was lowest and maximum value for Cl. The minimum value for Na and So_4 was 0.12 and 0.21 mg/l respectively. While the maximum value for Na and So_4 was 120.94 and 56.31 mg/l respectively. Also the high and low value of Hco_3 was 2.01 and 10.79 mg/l. 317.76 and 11,533.4 mg/l was maximum and minimum value of TDS. The all of the parameters expect Hco_3 was low value in the north of the study area. While the maximum value of parameters was located in south of the study area. So the north of the study area was better quality than the south of the study area.

Complex relations between dissolved species can reveal the origin of solutes and the process that generated the observed water compositions. Calcium and magnesium

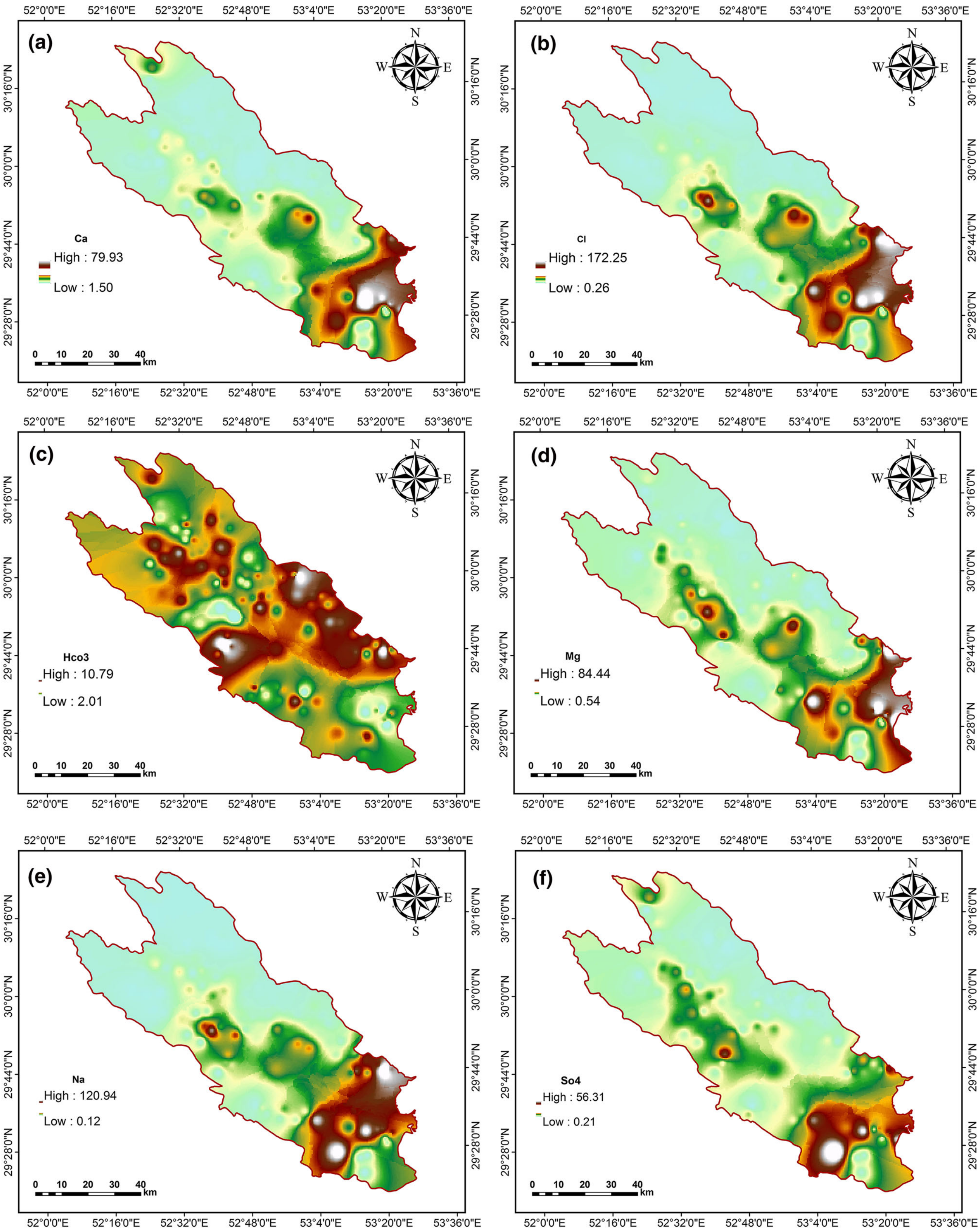


Fig. 3 Interpolated maps of the groundwater quality parameters generated using by IDW. **a** Ca; **b** Cl; **c** Hco_3; **d** Mg; **e** Na; **f** So_4; **g** TDS

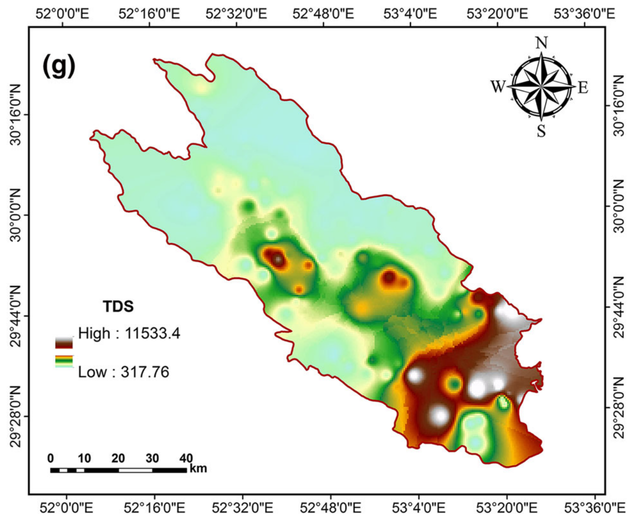


Fig. 3 continued

showed strong positive correlation (r = 0.68) which indicating a same geological formation (Fig. 3a). Also the results show that there was strong positive significant correlation between Ca, Na and Mg with So_4 (about r = 0.6) (Figs. 3b, c, 4b, c, d).

Figure 5 illustrate relation between Mg/Ca and Na/Ca ratios for groundwater (r = 0.60) (Fig. 5).

Also Fig. 6 shows the value of Cl as a function of Na in the groundwater samples and there was a strong correlation (r = 0.84) between them.

Figure 7 illustrate relations between various anions and cations with TDS. There were correlation between TDS and Mg (r = 0.0.74), Na (r = 0.86), Ca (r = 0.85), SO4 (r = 0.62) and Cl (r = 0.97) (Fig. 6). So Cl and TDS has the strong positive correlation.

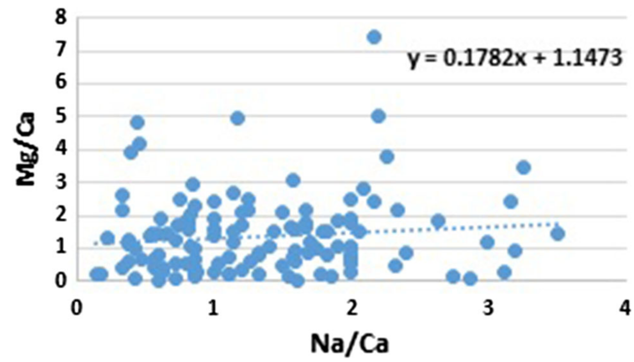


Fig. 5 Plot of Mg/Ca versus Na/Ca for the ground waters in the study area

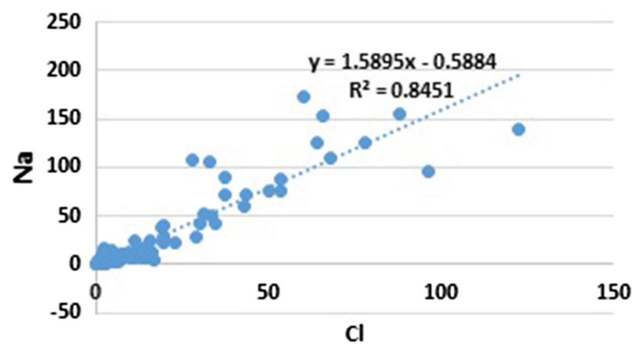
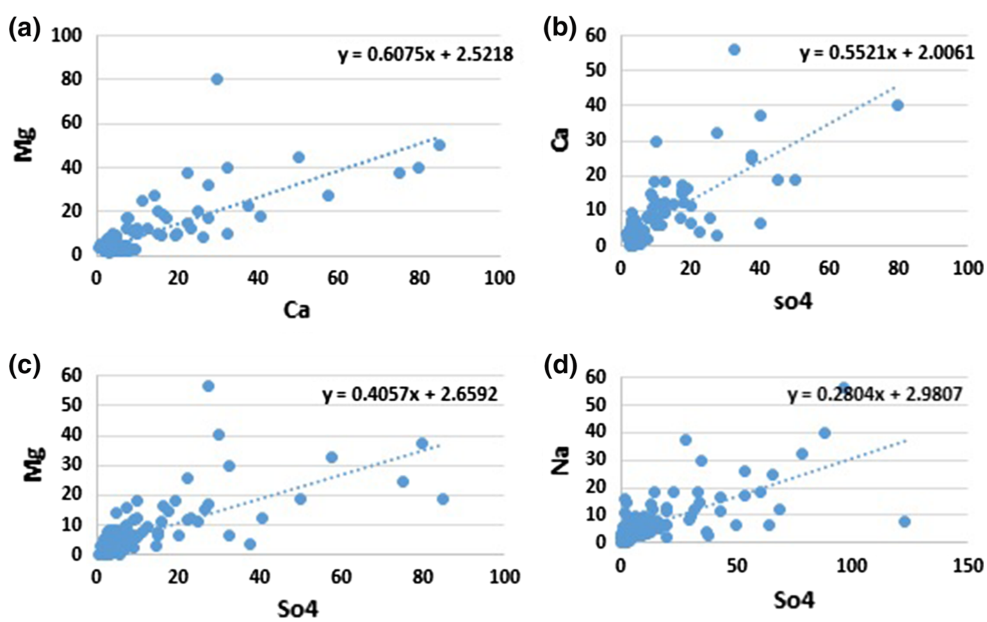


Fig. 6 Plots of Na+ versus Cl in the study area

Fig. 4 Relationship between groundwater quality parameters. **a** Mg versus Ca; **b** Ca versus So_4; **c** Mg versus So_4; **d** Na versus So_4



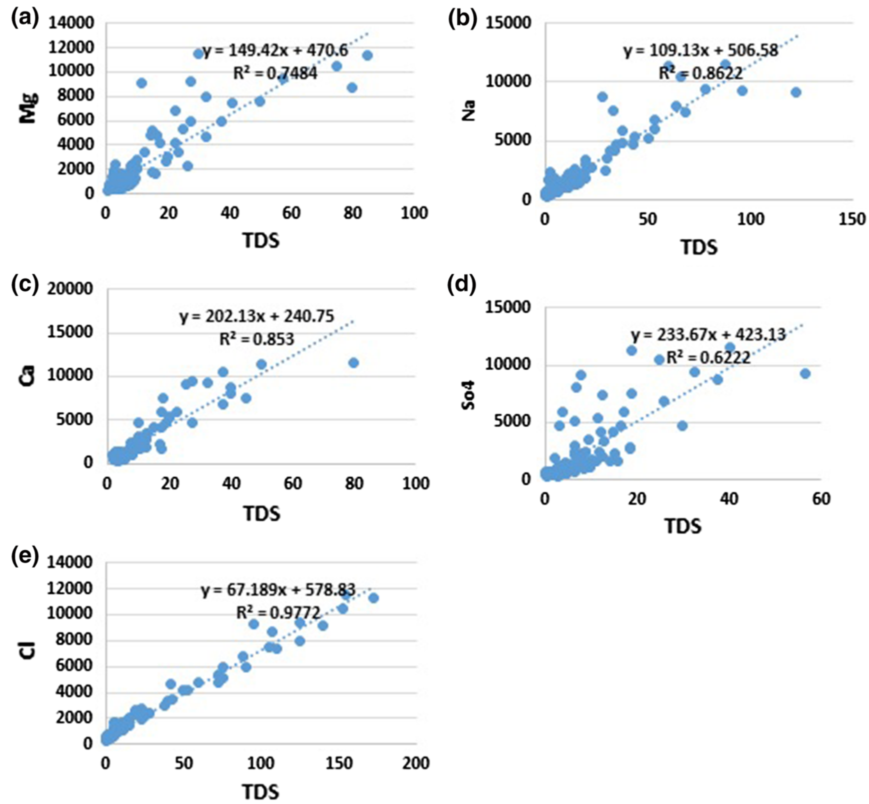


Fig. 7 Relationship between TDS and anions/cations in groundwater for the study area

Conclusion

This aim of the study was determination of the groundwater quality and relationship between variability of groundwater quality, calcium (Ca), pH, potassium (k), chlorine (Cl), magnesium (Mg), sodium (Na), electrical conductivity (EC), sulfate (So_4), total dissolved solids (TDS) in north of the Fars province, southeast of Iran. The results of spatial distribution each of parameters by IDW show that the all of the parameters expect Hco_3 was low value in the north of the study area. While the maximum value of parameters was located in south of the study area. So the north of the study area was better quality than the south of the study area. The relationship between parameters by MLR show that Calcium and Magnesium showed strong positive correlation ($r = 0.68$) and there was strong positive significant correlation between Ca, Na and Mg with So_4 (about $r = 0.6$). Also Cl and TDS has the strong positive correlation ($r = 0.97$).

References

Burrough PA, Goodchild MF, McDonnell RA, Switzer P, Worboys M (1998) Principles of geographic information systems. Oxford University Press, Oxford

Falivene O, Cabrera L, Tolosana-Delgado R, Saez A (2010) Interpolation algorithm ranking using cross-validation and the role of smoothing effect: a coal zone example. Comput Geosci 36(4):512–519

Gotway CA, Ferguson RB, Hergert GW, Peterson TA (1996) Comparison of kriging and inverse-distance methods for mapping soil parameters. Soil Sci Soc Am J 60(4):1237–1247

Jalali M (2009) Geochemistry characterization of groundwater in an agricultural area of Razan, Hamadan, Iran. Environ Geol 56:1479–1488

Keshavarzi A, Sarmadian F (2010) Comparison of artificial neural network and multivariate regression methods in prediction of soil cation exchange capacity (case study: Ziaran region). Desert 15(2010):167–174

Kravchenko A, Bullock DG (1999) A comparative study of interpolation methods for mapping soil proper- ties. Agron J 91(3):393–400

Kravchenko AN (2003) Influence of spatial structure on accuracy of interpolation methods. Soil Sci Soc Am J 67(5):1564–1571

Negreiros J, Costa AC, Painho M (2011) Evaluation of stochastic geographical matters: morphologic geostatistics, conditional sequential simulation and geographical weighted regression. Trends Appl Sci Res 6(2011):237–255

Ordu S, Demir A (2009) Determination of Land data of Ergene basin (Turkey) by planning geographic information systems. J Environ Sci Technol 2(2):80–87

Ostovari Y, Zare S, Harchegani HB, Asgari K (2013) Effects of geological formation on groundwater quality in Lordegan Region, Chahar-mahal-va-Bakhtiyari, Iran. Int J Agric Crop Sci 5(17):1983–1992

Panagopoulos T, Jesus J, Antunes MDC, Beltrao J (2006) Analysis of spatial interpolation for optimising management of a salinized field cultivated with lettuce. Eur J Agron 24(1):1–10

Rabah FKJ, Ghabayen SM, Salha AA (2011) Effect of GIS interpolation techniques on the accuracy of the spatial representation of groundwater monitoring data in gaza strip. J Environ Sci Technol 4(6):579–589

SPSS Inc (2002) SPSS for Windows, Release 9.5 Spss. Inc

Wollenhaupt NC, Wolkowski RP, Clayton MK (1994) Mapping soil test phosphorus and potassium for variable rate fertilizer application. J Prod Agric 7(4):441–447