Shahid Chamran University of Ahvaz

Irrigation Sciences and Engineering (JISE)
Vol. 45, No. 2, Summer 2022, p. 81-97

Journal homepage: https://jise.scu.ac.ir

Original Paper

Modeling Ground-Water Quality using Time Series Models (A Case Study: Dehloran Plain, Ilam)

F. Mohammadyari[1*], A. Zarandian[2] and F. Beigmohammadi[3]

1- Corresponding Author, PhD in Evaluation Environmental and Land Use Planning, Faculty of Natural Resources & Environment, Malayer University, Malayer, Iran.(m.fatima.1364@gmail.com).*
2- Assistant professor, research group of environmental assessment and risks, Research Center for Environment and Sustainable Development (RCESD), Department of Environment, Tehran, Islamic Republic of Iran.
3- Ph.D. Student of Environmental Pollution, Faculty of Natural Resources & Environment, Malayer University, Malayer, Iran.

ARTICLE INFO

Artlcle history:
Received: 19 July 2021
Revised: 19 November 2021
Accepted: 21 November 2021

Keywords:
Ground-water, Time series, Dehloran Plain, Auto-regressive.

TO CITE THIS ARTICLE:

Mohammadyari, F., Zarandian, A., Beigmohammadi, F. (2022). 'Modeling Ground-Water Quality using time Series Models (A Case Study: Dehloran Plain, Ilam)', Irrigation Sciences and Engineering, 45(2), pp. 81-97.

Abstract

The main purpose of the present study is to modeling the variation of ground-water quality parameters from 2001 to 2018 and predicting its quality for 2027. To achieve it, we accessed parameters which included total hardness (TH), total dissolved solids (TDS), sodium (Na), sulfates (SO_4), and chlorides (Cl) which acquired from thirty-four wells in Dehloran Plain, Ilam. Due to the large number of wells, the samples were classified through cluster analysis into six clusters. To determine the number of clusters, a hierarchical clustering method was used. Five time-series models of autoregressive (AR), moving-average (MA), auto-regressive moving-average (ARMA), autoregressive integrated moving-average (ARIMA), and seasonal auto-regressive integrated moving-average (SARIMA) were applied to predict the changing ground-water quality. The best model was selected based on the Autocorrelation function (ACF) and Partial autocorrelation function (PACF), Akaike Information Criterion (AIC), and Coefficient of determination (R^2). The results of the prediction indicated that the average concentration of Cl and Na will increase in all the clusters in 2027. Moreover, the average of the predicted SO_4 will increase in all clusters except for the sixth one. The average of TDS also will increase in the first to third clusters, while it will decline in the fourth, fifth, and sixth clusters. The average of the predicted TH in the first, second, third, and fifth clusters will rise, whereas it will be reduced in the fourth and sixth clusters. It can be concluded that the status of ground-water quality is worsening in Dehloran Plain and in 2027 its quality will become lower compared to previous years.

Introduction

The most necessary prerequisite that nature prepares to hold life for the human population is water. Groundwater is considered a critical natural resource for human-health, socioeconomic development, and ecosystem function (Dhayachandhran and Jothilakshmi, 2020). Groundwater constitutes about twenty percent of the world's fresh water supply. Alsalme et al (2021). The groundwater is susceptible to pollution due to excessive usage of fertilizers, pesticides, increased anthropogenic activities, and rapid-growing industries (Karthika et al., 2018). Groundwater contaminations threaten human health and intensely affect the environment (Kumar and Sangeetha, 2020). So, the study of water quality contributes to developing strategies to control surface and groundwater pollution (Kumar and Sangeetha, 2020).

In arid and semi-arid environments, groundwater has an important role in the ecosystem (Mirzavand and Ghazavi, 2015).

Dehloran Plain is a region with dry and mid-dry. in this plain rainfall is very variable and significantly lower than the evaporation rate; so, groundwater can be a main portion of the water. As regards, Ground-water management is more difficult than that for surface water resource management; so, there is a requirement to use for sensible and cost-effective procedures to evaluate the situation of these waters (Mirsanjari and Mohammadyari, 2017). In this regard, the present study was conducted to model the groundwater quality parameters of the Dehloran plain.

For this purpose, in the present study, time-series models were used to investigate the changes (from 2001 to 2018) and predict (for 2027) groundwater quality parameters. These models define a process of observations over time, examine time series analysis, and prediction of future values according to the historical observations of the variables (Taneja et al., 2017). The time-series model is developed in three phases: identification, assessment, tracking, and recognition (Shirmohammadi et al., 2013; He et al., 2014). In general, they are used for the prediction and production of the data (Adhikary et al., 2012). Time-series models are used extensively for different purposes of climate change (Cadenas and Rivera, 2010) air pollution (Anttila and Tuovinen, 2010; Chaudhur & Dutta; 2014; D'Urso et al, 2015; Taneja et al., 2017; Liu et al., 2019; Mirsangari et al., 2020), Ground-water quality and Ground-water level Panda and Kumar, (2011); Behnia and Rezaeian, (2015); Mirzavand and Ghazavi, (2015); Mirsanjari and Mohammadyari, (2017), and water quality (Faruk, 2010; Wang et al., 2014). Among these studies, The combination of time series models (AR, MA, ARMA, ARIMA and SARIMA) for groundwater prediction has been reported only in the study of Mirsanjari and Mohammadyari (2017). They used time series data Ground-water wells for agriculture in Mehran Plain, and finally Forecast the situation of ground-water quality parameters for the 2021 year by the best model obtained. the best model was selected according to information criterion or Akaike (AIC) and correlation coefficient. the results showed that the quality of groundwater for Agriculture Plain Mehran will decrease in future. so, the chief goal of this study is to evaluate the performances of time series models for groundwater predicting in Dehloran Plain.

Materials and Methods

Study Area

Dehloran Plain is located between 32◦ 2′- 33◦ 3′ N latitude and 53◦ 2′- 53◦ 40′ E longitude in an area of 4920 Km^2 in Ilam Province (Fig. 1). Climatically, according to Coupon classification, the area is categorized as a dry zone. There are six main geological structures including Sarvak, Ilam, Imam Hassan District, Pabdeh, Asmari, and Quaternary in the region. The sediments in the region include hillside deposits, alluvial fans, and plain deposits. Groundwater pollution sources in the region include natural and man-made resources. Geological formations are the most important source of natural groundwater pollution in this plain. Agricultural lands, agro-industrial complexes, slaughterhouses, urban and rural settlements, and sewage treatment plants are among the human activities that can cause groundwater pollution.

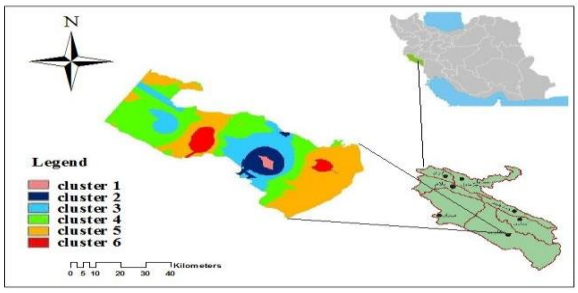

Fig. 1- Map of the wells clustering in the study area

Methodology

In this research, we used data acquired from ground-water quality parameters from the years 2001 to 2018 for 34 wells in Dehloran Plain and then modeling the variation process of these parameters and prediction for the next 9 years (2027) through the most fitted model. These data were taken from the Regional Water Department of Ilam Province. The quality parameters included TH (total hardness), TDS (total dissolved solids), SAR, EC, sodium, sulfate, and chlorine. The data was a continuously monthly-based measurement. Due to a large number of wells the samples were classified into 6 clusters using cluster analysis. In order to determine the number of clusters, a hierarchical clustering method was applied. The K-mean method was also used to specify the number of final clusters and the wells that are placed in one cluster. Drawing time series data is the first step in the analysis of time series design. The goal is to determine the presence or absence of the trend in the dataset. In the next step, the components of the process in time series were determined, and then they were removed in order to make the data static. After that, the appropriate model is fitted to the data to identify the best model and accordingly make the prediction. The third step was investigating the normality of the resulted prediction data; so that the Kolmogorov-Smirnov test was used to assess the normalization of data.

Given the models used in this study, there is no need to determine the line fitting equation and to remove the process. In the ARIMA and SARIMA models, the seasonal data status is also removed through differentiation.

In this study we used R software to determine the best time series model and finally predict the data using the selected model. Out of 76 qualitative data of each parameter, 36 data were simulated and set aside for model calibration. After fitting the model, it is necessary to determine the accuracy of the selected trend. The best model was determined according to the coefficient of determination and the. Akaike's information criterion (AIC) was used to compare different models (p, q) ARMA and was calculated as Equation (1) (Mirsanjari and Mohammadyari, 2017).

$$AIC(K)= NLn(MSE)+2K \quad (1)$$

Where n is the number of data points (for calibration), and K is the number of free parameters, and MSE stands for mean square error. Usually, the preferred model gives a higher R2 or the smallest value of AIC. After validation of the best model fitted on time series, it can be used for future prediction. The prediction process is that the current period is shown by t and t+τ represents the prediction for the period of t+τ. The prediction is made by considering the mean at the origin of t from the model written at the time of t+τ. In general, prediction is provided for the time of t+τ-1, t+2+t+1. In this method, xt+j which occurs at the time of t is replaced with the predictions of ẍt+j (t), and εt+j that are not occurred at the time of t are substituted with zero. εt+j that are not occurred are replaced with a single-period prediction error of e1(t-j)=xt-j-ẍt-j(t-j-1). (Mohammadyari, 2021).

Time Series Models

AR Model

The auto-regressive model (AR) (p) can be expressed as Eq. (2):

$$z_t= \varphi_1 z_{t-1}+ \varphi_2 z_{t-2}+ \varphi_p z_{t-p}+ a_t \quad (2)$$

Where φ_1, φ_2 and φ_p are coefficient and model parameters and a_t is random term of the data that follows by normal distribution with a zero mean (Hannan, 1971; Mirsanjari and Mohammadyari, 2017).

MA Model

The moving average model (MA (q)) can be expressed as Eq. (3):

$$z_t= \theta_1 a_{t-1} + \theta_2 a_{t-2}+ \ldots\ldots \theta_q a_{t-q}+ a_t \quad (3)$$

Where θ_1, θ_2 and θ_q are coefficient and model parameters and at is a random term of the data that follows by normal distribution with a zero mean (Hannan, 1971).

ARMA Model

The auto-regressive moving average (ARMA) model ARMA (p, q) can be expressed as Eq. (4):

$$Y_t = \delta + \sum_{i=1}^{p} \Phi_i y_{t-1} + \sum_{j=1}^{q} \varphi_j e_{t-j} + e_t \quad (4)$$

Where δ is the constant term of the ARMAmodel, ϕ_i indicates the ith auto-regressive coefficient, φ_j is the jth moving average coefficient, e_t shows the error term at time period t, and Y_t refers the value of groundwater level observed or forecasted at time period t (Erdem & Shi, 2011).

ARIMA and SARIMA Models

Auto-regressive integrated moving average (ARIMA) models are one of the most important linear model types for time series forecasting. ARIMA models originated from the combination of autoregressive models (AR) and moving average models (MA). ARIMA fits a Box-Jenkins ARIMA model to a time series (Shirmohammadi et al., 2013). ARIMA was issued to model time series behavior and to generate forecasts. ARIMA modelling uses correlational techniques and can be used to model patterns that may not be visible in plotted data (Box et al., 1994). In ARIMA, the future value of a variable is assumed to be a linear function of several past observations and random errors. A SARIMA model can be explained as ARIMA (p, d, q) (P, D, Q) s, where (p, d, q) is the non-seasonal part of the model and (P, D, Q) s is the seasonal part of the model in which p is the order of non-seasonal auto regression, d is the number of regular differencing, q is the order of non-seasonal MA, P is the order of seasonal auto-regression, D is the number of seasonal differencing, Q is the order of seasonal MA, and s is the length of the season (Faruk, 2010).

Data Clustering

Owing to a large number of wells, the clustering analysis was used to convert the data into six clusters. To determine the cluster numbers, we used the hierarchical clustering method. Then, using the K-mean method we determined the number of final clusters and also which wells are located in a specific cluster. In cluster analysis, the number of attributes (P) on the number of measured elements (N) is measured and then is formed as a matrix of N*P from raw data. Then, the matrix of raw data is converted to the matrix of similarities or distances, and then using one of the classification methods, the N (number of elements) are categorized based on the similarities between them (Guler et al., 2002). After cluster analysis, it is determined which wells have the most similarities from the aspect of qualitative parameters. In the end, the annual average of all qualitative parameters is calculated for the wells which are lied in a cluster; this value is representative of all the wells which are in one specified cluster and used in the simulation of qualitative parameters.

Results and Discussion

To predict the parameters of ground-water quality for each cluster using monthly data, they were analyzed in four sections including random, seasonal, trend, and observed; the random parameter was selected as the model. Additionally, five models with twelve sub-models were evaluated. The results of clustering indicated that the wells were classified into 6 clusters based on their similarities. The number of wells in the clusters from first to the sixth was seven, two, five, seven, eight, and five wells, respectively. the results related to the AR, MA, ARMA, ARIMA and SARIMA models of all clusters are shown in table (1) After that 12 models were fitted on parameters, the best model was selected to predict groundwater parameters using Akaike and correlation coefficient indices.

Also Figure (2) to (7) shows real data and simulation of groundwater well parameters.

Table 1- Results of time series models in 6 clusters

Clusters models			1			2			3			4			5			6		
AR(1)	parameter		Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2
	Cl	φ1	-0.9	70.08	0.81	-0.3	89.9	0.46	-0.3	166.7	0.35	-0.3	180.3	0.79	-0.4	187	0.49	-0.35	170.2	0.56
	Tds		-0.6	459.2	0.53	-0.39	406.8	0.46	-0.3	496.9	0.51	-0.2	471	0.77	-0.4	421	0.45	-0.3	464.4	0.51
	Th		-0.8	422.4	0.61	-0.5	389.8	0.55	-0.2	453.5	0.49	-0.3	448	0.57	-0.4	378	0.71	-0.2	463.3	0.57
	Na		-0.4	203.3	0.35	-0.43	136.7	0.79	-0.3	190.6	0.76	-0.3	196	0.57	-0.4	103	0.55	-0.4	182.8	0.64
	So_4		-0.8	165.1	0.78	-0.6	112.1	0.3	-0.3	223.9	0.70	-0.3	212	0.48	-0.5	140	0.67	-0.3	333.3	0.52
AR(2)	Cl	φ1	-1.27	59/9	0.76	-0.54	80.72	0.45	-0.5	155.7	0.52	-0.5	178.6	0.77	-0.6	183	0.49	-0.5	161.9	0.57
		φ2	-0.8			-0.47			-0.6			-0.3			-0.4			-0.5	9	
			-0.4			-0.44			-0.5			-0.3			-0.5			-0.4		
	Tds		-0.8	442.2	0.29	-0.55	402.9	0.35	-0.5	485.9	0.61	-0.3	468	0.79	-0.5	413	0.47	-0.5	457	0.54
			-0.8			-0.41			-0.6			-0.4			-0.5			-0.4		
	Th		-0.9	412.8	0.52	-0.6	388.7	0.55	-0.4	443.07	0.52	-0.5	444	0.59	-0.6	369	0.68	-0.4	453.7	0.52
			-0.8			-0.3			-0.5			-0.4			-0.5			-0.5		
	Na		-0.7	193.6	0.38	-0.62	131.88	0.78	-0.6	180.7	0.76	-0.3	197.9	0.57	-0.6	94.5	0.59	-0.6	177.4	0.65
			-0.6			-0.43			-0.5			-0.05			-0.6			-0.4		
	So_4		-1.13	157.2	0.70	-0.74	112.6	0.39	-0.5	213.1	0.71	-0.4	210	0.51	-0.7	130	0.67	-0.4	331.7	0.53
			-0.8			-0.22			-0.5			-0.3			-0.6			-0.3		
MA(1)	Cl	θ1	-1	59/1	0.71	-1	89.96	0.45	-1	148.93	0.47	-1	167.8	0.76	-1	171.5	0.43	-1	151.8	0.56
																			1	
	Tds		-1	443.4	0.35	-1	387.6	0.38	-1	479.4	0.55	-1	458	0.76	-1	404	0.48	-1	446.6	0.51
	Th		-1	409	0.49	-0.9	374.1	0.55	-1	437.2	0.59	-1	432	0.59	-1	361	0.68	-1	444.3	0.58
	Na		-1	185.5	0.36	-1	120.5	0.79	-1	172.6	0.77	-0.3	195	0.58	-1	84	0.55	-1	164.9	0.66
	So_4		-1	153	0.70	-1	97.38	0.41	-1	205.9	0.71	-1	197.1	0.52	-1	123	0.65	-1	319	0.51
MA(2)	Cl	θ1	-1.62	52.4	0.69	-0.07	66.63	0.45	-1.5	147.3	0.48	-0.92	169.7	0.8	-1.2	172	0.49	-1.76	150.5	0.58
		θ2	0.62			-1			0.5			-0.07			0.2			0.76		
	Tds		-1.5	439.1	0.55	-1.97	372.1	0.38	-1.5	478.1	0.54	-0.7	459	0.8	-1.9	403	0.57	-1.5	446.8	0.5
			0.54			1			0.5			-0.2			0.99			0.5		
	Th		-1.6	403	0.48	-1.99	367.7	0.54	-1.42	438.6	0.65	-1.1	433	0.58	-1.9	355	0.67	-1.9	436.7	0.54
			0.63			0.99			-1			0.17			1			0.99		
	Na		-1.9	181.6	0.38	-1.31	121.46	0.79	-1.5	170.7	0.77	-0.3	197.8	0.58	-1.9	73	0.58	-1.9	157.8	0.64
			0.9			0.31			0.5			0.05			1			1		
	So4		-1.6	147.6	0.67	-1.97	77.48	0.30	-1.5	204.5	0.71	-0.9	199	0.57	-1.9	114	0.66	-1.95	315	0.53
			0.62			1			0.5			-0.09			1			0.95		

Mohammadyari *et al.,45 (2) 2022* *DOI:* 10.22055/JISE.2021.37997.1984

Table 1- (continued)

Clusters models		1			2			3			4			5			6		
ARMA (1,1)	parameter	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2
	Cl	φ1 θ1	-0.6 -1	54.3	0.77	-0.07 -1	73.21	0.42	-0.1 -1	150.3	0.56	-0.03 -1	169.7	0.79	-0.1 -1	173.1	0.95	-0.1 -1	153.4	0.56
	Tds		-0.4 -0.99	441.7	0.56	-0.21 -1	388.2	0.45	-0.1 -1	480.9	0.59	0.08 -0.9	460	0.79	-0.09 -1	406	0.51	-0.08 -1	448.3	0.79
	Th		-0.5 -1	406.2	0.62	-0.2 -1	373.8	0.55	-0.01 -1	439.2	0.54	-0.04 -1	434	0.61	0.2 -1	361	0.71	-0.1 -1	445.9	0.54
	Na		-0.1 -1	186.3	0.35	-0.11 -1	122.09	0.79	-0.1 -1	174	0.77	-0.1 -0.2	197.8	0.76	-0.2 -1	84	0.56	-0.1 -1	165.6	0.89
	So_4		-0.5 -1	149	0.72	-0.42 -1	93.62	0.45	-0.1 -1	207.3	0.71	0.03 -1	199	0.45	-0.2 -1	122	0.69	-0.05 -1	320.9	0.51
ARMA (1,2)	Cl	φ1 θ1 θ2	-0.02 -1.6 0.6	54.4	0.76	-0.08 -1 -0.4	87.79	79.2	0.2 -1.7 0.7	148.3	0.54	0.5 -1.9 1	166.7	0.79	0.4 -1.9 0.99	170	0.39	-0.7 -0.09 -0.9	154.7	0.63
	Tds		0.39 -1.9 0.9	440.7	0.53	-0.99 0 -0.99	385.7	0.37	-0.4 -0.5 -0.4	483.3	0.61	-0.5 -0.2 -0.7	460	0.81	-0.5 -0.3 -0.6	408	0.47	-0.6 -0.2 -0.7	450.2	0.61
	Th		0.2 -1.9 0.99	405.2	0.62	0.25 -1.98 1	367.8	0.56	-0.3 -0.4 -0.5	440.1	0.59	-0.6 0 -1	432	0.54	0.2 -1.9 1	355	0.71	-0.8 0 -1	446.8	0.62
	Na		0.3 -1.9 1	180.6	0.25	0.47 -1.98 0.99	118.5	0.79	0.2 -1.7 0.7	171.7	0.77	-0.4 0.02 -0.09	199.8	0.59	-0.9 -0.05 -0.9	87	0.48	0.2 -1.98 1	158.1	0.73
	So_4		0.3 -1.9 0.99	149	0.73	-0.2 -1.97 1	78.08	0.79	-0.3 -1.7 0.7	205.4	0.72	-0.7 -0.05 -0.9	198	0.49	0.1 -1.9 1	115	0.85	-0.3 -0.5 -0.4	322.9	0.59

Table 1- (continued)

Clusters models			1			2			3			4			5			6		
	parameter		Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2
ARMA (2,1)	Cl	φ1 φ2 θ1	-0.4 -0.6 -1	52.8	0.91	-0.08 -0.43 1	75.12	0.68	-0.2 -0.5 -1	141.9	0.53	-0.02 -0.3 -1	168.4	0.88	-0.1 -0.2 -1	173	0.38	-0.1 -0.4 -1	148.3	0.84
	Tds		-0.4 -0.6 -1	436.9	0.76	-0.3 -0.5 -1	378.9	0.75	-0.2 -0.5 -1	472.6	0.61	0.5 -0.3 -1	467	0.79	-0.1 -0.5 -1	400	0.5	-0.1 -0.4 -1	444	0.59
	Th		-0.4 -0.7 -1	402.6	0.77	-0.3 -0.2 -1	373.09	0.55	-0.07 -0.5 -1	430	0.56	-0.4 -0.7 0.67	498	0.51	-0.2 -0.4 -1	358	0.7	-0.1 -0.5 -1	435.7	0.85
	Na		0.2 -0.3 -1	183.6	0.35	-0.16 -0.28 -1	121.39	0.79	-0.2 -0.5 -1	166.5	0.78	-1.2 -0.3 0.8	199	0.57	-0.3 -0.6 -1	77	0.55	-0.2 -0.3 -1	162.66	0.73
	So_4		-0.4 -0.6 -1	148.4	0.85	-0.58 -0.36 -1	91.61	0.39	-0.2 -0.5 -1	199.9	0.73	0.03 -0.33 -1	197.3	0.89	0.2 -0.2 -1.9	115	0.67	-0.1 -0.4 -0.9	316.2	0.59
ARMA (2,2)	Cl	φ1 φ2 θ1 θ2	-0.4 -0.6 -1 0.01	54.8	0.78	-0.39 -0.43 -1 0.96	483.6	0.35	-0.1 -0.5 -1.1 0.1	143.6	0.52	0.6 -0.3 -1.9 1	165.1	0.86	0.5 -0.2 -1.9 0.99	171	0.45	0.4 -0.2 -1.99 0.99	147.5	0.64
	Tds		-0.6 -0.7 -0.6 -0.3	437.8	0.48	-0.1 -0.6 -1.9 1	359.02	0.53	-0.1 -0.5 -1 0.2	474.2	0.61	0.6 -0.4 -1.9 1	451	0.82	0.4 -0.3 -1.9 0.9	399	0.42	0.4 -0.2 -1.9 1	442.9	0.56
	Th		-0.3 -0.6 -1.1 0.09	404.6	0.67	0.27 -0.1 -1.98 0.99	354.9	0.87	-0.01 -0.5 -1 0.1	432.8	0.61	-0.8 -0.2 0 -1	432	0.65	0.3 -0.2 -1.9 1	355	0.7	-0.2 -0.4 -1.9 0.99	431.7	0.64
	Na		0.2 -0.3 -1 0	183.6	0.35	0.52 -0.19 -1.9 0.99	119.4	0.79	-0.1 -0.5 -1.1 0.1	168	0.79	0.2 0.3 -0.7 -0.2	199.6	0.59	0.1 -0.4 -1.9 1	70	0.54	0.2 -0.2 -1.98 1	157.6	0.62
	So_4		-0.3 -0.5 -1.1 0.1	150.3	0.79	-0.36 -0.54 -1.87 1	72.04	0.51	-0.1 -0.5 -1.1 0.1	201.6	0.73	-0.7 -0.1 -0.08 0.9	199	0.45	0.2 -0.2 -1.9 1	115	0.67	-0.3 -0.5 -1.8 1	308.4	0.63

Mohammadyari *et al.,45 (2) 2022* *DOI:* 10.22055/JISE.2021.37997.1984

Table 1- (continued)

Clusters models			1			2			3			4			5			6		
	parameter		Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2
ARIMA (1,1,2)	Cl	φ1 d θ1 θ2	-0.5 1 -1.9 0.93	62.91	0.61	-0.35 1 -1.56 -1	74.5	0.32	-0.8 1 -0.2 -0.7	170.9	0.61	0.08 1 -1.98 0.99	172.8	0.8	-0.06 1 -1.9 1	176	0.4	-0.06 1 -1.98 1	157.13	0.63
	Tds		-0.3 1 -1.9 0.99	436.6	0.56	-0.1 1 -1.9 1	384.4	0.35	-0.8 1 -0.2 -0.7	490.7	0.63	-0.4 1 0.9 -1	470	0.77	0.4 1 -0.3 -1.9	399	0.42	-0.04 1 -1.9 -1	442.8	0.57
	Th		-1 1 -1.7 0.8	403.9	0.43	-0.2 1 -1.98 1	370.9	0.55	-0.8 1 -0.1 -0.8	447.7	0.58	-0.5 1 0 -1	450	0.58	-0.1 1 -1.9 1	359	0.7	-0.07 1 -1.97 1	440.3	0.61
	Na		-0.1 1 -1.9 1	189.2	0.29	-0.06 1 -1.98 1	126.08	0.78	-0.8 1 -0.2 -0.7	194.2	0.76	-0.2 1 -1 0.09	197.1	0.59	-0.2 1 -1.9 1	90	0.56	-0.1 1 -1.97 1	70.21	0.63
	So_4		-0.5 1 1.9 1	154	0.74	-0.40 1 -1.97 1	99.34	0.37	-0.8 1 -0.2 -0.7	226.8	0.71	0.08 1 -1.9 1	201	0.42	-0.2 1 -1.9 1	127	0.69	-0.01 1 -1.95 0.99	318.7	0.61
ARIMA (1,2,1)	Cl	φ1 d θ1	-0.7 2 -1	95.2	0.51	-0.43 2 -1	0.87	0.41	-0.4 2 -1	186.9	0.59	-0.6 2 -1	195.1	0.81	-0.5 2 -1	207	0.32	-0.4 2 -1	191.36	0.61
	Tds		-0.4 2 -0.99	460.3	0.22	-0.5 2 -1	411.2	0.33	-0.4 2 -1	496.3	0.59	0.07 2 -0.4	456	0.85	-0.05 2 -1.9	426	0.51	-0.4 2 -1	466.7	0.57
	Th		-0.5 2 -1	424.7	0.38	-0.6 2 -1	394.9	0.53	-0.4 2 -1	454.1	0.6	-0.07 2 -0.3	431	0.81	-0.5 2 -1	388	0.73	-0.4 2 -1	464.4	0.58
	Na		-0.5 2 -1	221.7	0.38	-0.5 2 -1	160.17	0.79	-0.5 2 -1	210.1	0.74	-0.6 2 -1	207.2	0.59	-0.5 2 -1	127	0.59	-0.5 2 -1	203.54	0.64
	So_4		-0.6 2 -1	182.6	0.40	-0.7 2 -1	133.82	0.32	-0.4 2 -1	241.6	0.68	-0.5 2 -1	226.8	0.42	-0.5 2 -1	164	0.68	-0.5 2 -1	338.7	0.57

Table 1- (continued)

Clusters models		1			2			3			4			5			6		
	parameter	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2
SARIMA (1,1,0) (1,1,1)(4)	Cl	φ1 d Φ1 D Θ1	-0.4 1 -0.5 1 0.7	88.3	0.78	-0.61 1 -0.37 1 -0.38	88.84	0.35	-0.5 1 -0.6 1 1	159.8	0.46	-0.6 1 -0.1 1 -0.5	183.1	0.78	-0.6 1 -0.3 1 -0.4	199	0.21	-0.5 1 0.19 1 -1	178.3	0.53
	Tds		-0.4 1 -0.5 1 -0.6	416.7	0.29	-0.6 1 -0.6 1 0	291.1	0.27	-0.5 1 -0.6 1 0.99	438.2	0.6	-0.5 1 -0.5 1 -0.1	427	0.73	-0.4 1 0.1 1 -0.6	389	0.57	-0.5 1 0.3 1 -0.8	424.9	0.43
	Th		-0.7 1 -0.07 1 0.1	382.1	0.39	-0.6 1 -0.6 1 0	341.6	0.35	-0.2 1 -0.99 1 0.7	406.7	0.48	-0.4 1 -0.5 1 -0.2	409	0.45	-0.5 1 -0.1 1 -0.4	356	0.63	-0.5 1 0.5 1 -1	418.7	0.08
	Na		-0.6 1 0.1 1 -0.05	203.5	0.28	-0.64 1 -0.65 1 -0.52	149	0.70	-0.5 1 -0.4 1 0.3	184	0.79	-0.6 1 -0.02 1 -0.6	193	0.48	-0.4 1 0.3 1 -0.7	119	0.57	-0.6 1 -0.1 1 0.6	189.4	0.73
	So_4		-0.7 1 -0.9 1 0.7	164.9	0.75	-0.5 1 0 1 -0.99	-1.17	0.28	-0.5 1 -0.8 1 1	212.8	0.82	-0.5 1 -0.3 1 -0.4	210	0.31	-0.5 1 0.1 1 -0.6	155	0.65	-0.5 1 0.66 1 0.7	310.4	0.55

Mohammadyari *et al.*,*45 (2) 2022* *DOI:* 10.22055/JISE.2021.37997.1984

Table 1- (continued)

Clusters models		1			2			3			4			5			6		
	parameter	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2	Model coefficient	AIC	R^2
SARIMA (1,1,1) (1,1,1)(4)	Cl	φ1 d Θ1 Φ1 D Θ2	-0.81 −1.000 0.051 -0.3 -0.61 −1.000 0.30 -0.81 −1.000 -0.21 0.10 0.41 −1.000 0.51 -0.2 -0.81 −1.000 -0.61 0.41	76	0.75	-0.421 −1.000 -0.451 -0.29 -0.41 −1.000 -0.41 -0.21 -0.41 −1.000 0.41 -0.21 -0.44 1 −1.000 -0.37 1 -0.36 -0.33 1 −1.000 0.20 1 1.25	73.77	0.32	-0.21 −1.000 -0.991 0.99 -0.21 −1.000 -0.71 0.99 -0.21 −1.000 0.71 -0.91 -0.31 −1.000 -0.21 0.1 -0.3 1 −1.000 0.01 1 0.01	145.1	0.49	-0.41 −1.000 -0.051 -0.4 -0.2 -0.21 −1.000 -0.31 -0.2 -0.21 −1.000 -0.41 -0.31 -0.31 −1.000 -0.8 0.1 -0.2 1 −1.000 -1.2 1 -0.2	171.8	0.7	-0.41 −1.000 -0.41 -0.2 -0.31 −1.000 0.21 -0.5 -0.41 −1.000 -0.11 -0.3 -0.31 −1.000 0.4 -1 -0.3 1 −1.000 0.19 1 -0.5	179	0.25	-0.41 −1.000 0.291 -1 -0.41 −1.000 0.21 -0.7 -0.31 −1.000 -1 1 0.66 -0.5 1 −1.000 -0.07 1 -0.6 -0.4 1 −1.000 0.39 1 -0.09	162.6	0.49
	Tds		405	0.33		277.6	0.23		423.3	0.79		415	0.7		373	0.86		409	0.53
	Th		396.9	0.34		327.5	0.25		394.2	0.54		398	0.45		339	0.66		403.6	0.48
	Na		186.9	0.83		134.4	0.75		169.1	0.79		186	0.52		103	0.65		173.7	0.69
	So_4		152.9	0.74		-16.13	0.18		197.8	0.87		199	0.38		138	0.75		292.9	0.78

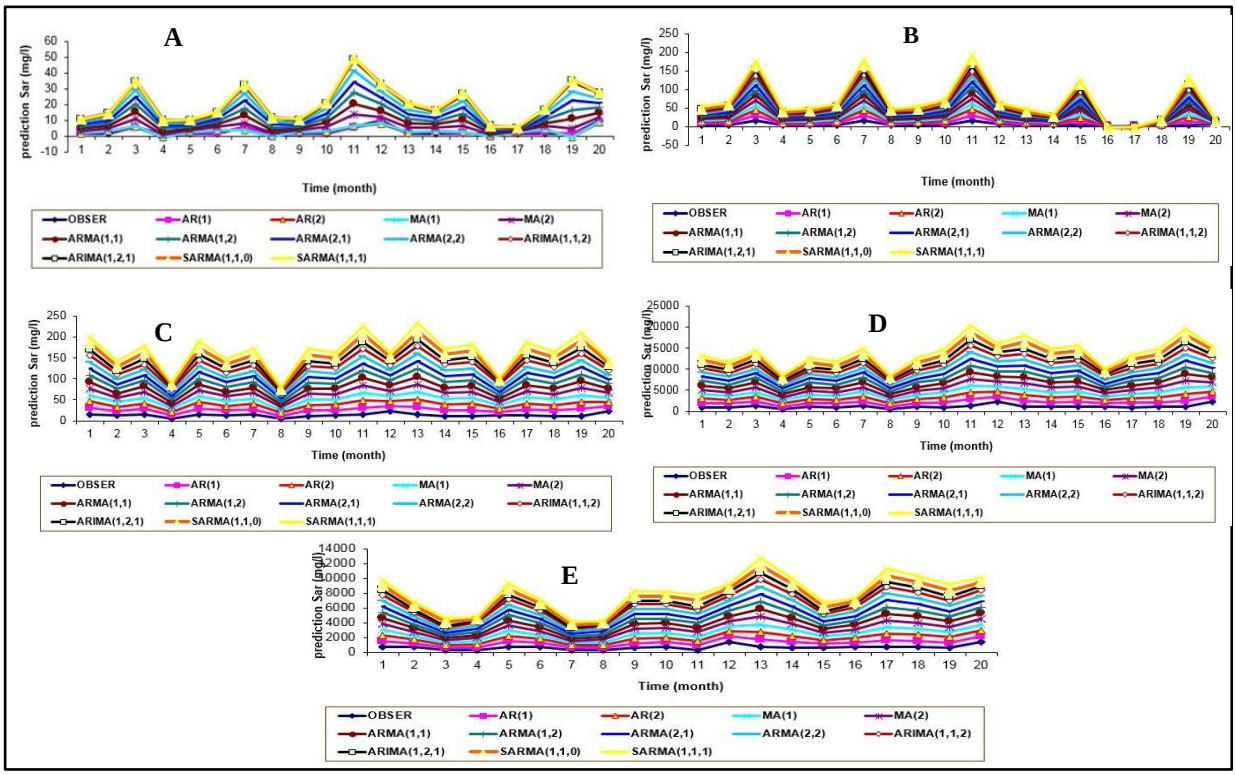

Fig. 2- Models prediction versus observed data in cluster 1
A: Cl; B: Na; C: SO_4; D: TDS; E: TH

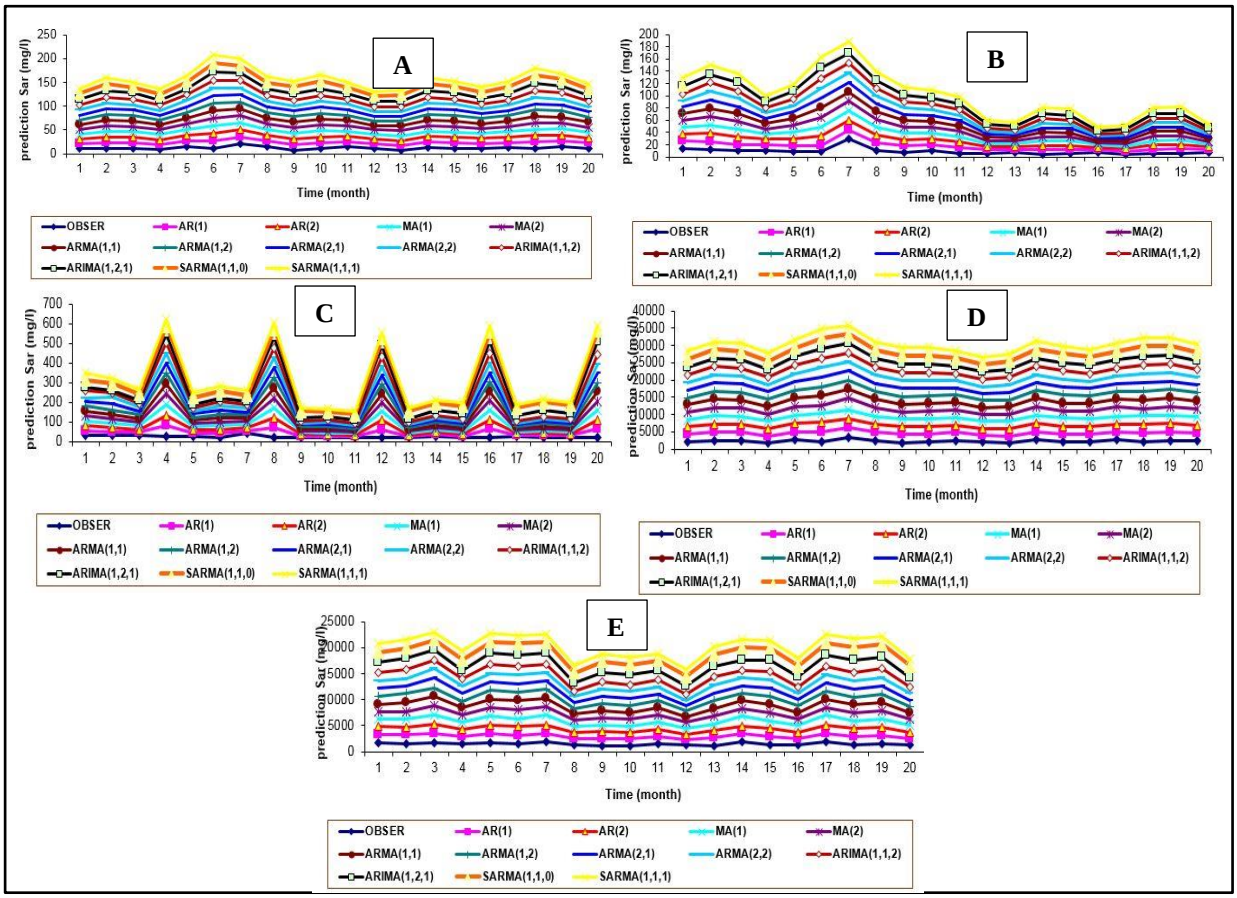

fig. 3- Models prediction versus observed data in cluster 2
A: Cl; B: Na; C: SO_4; D: TDS; E: TH

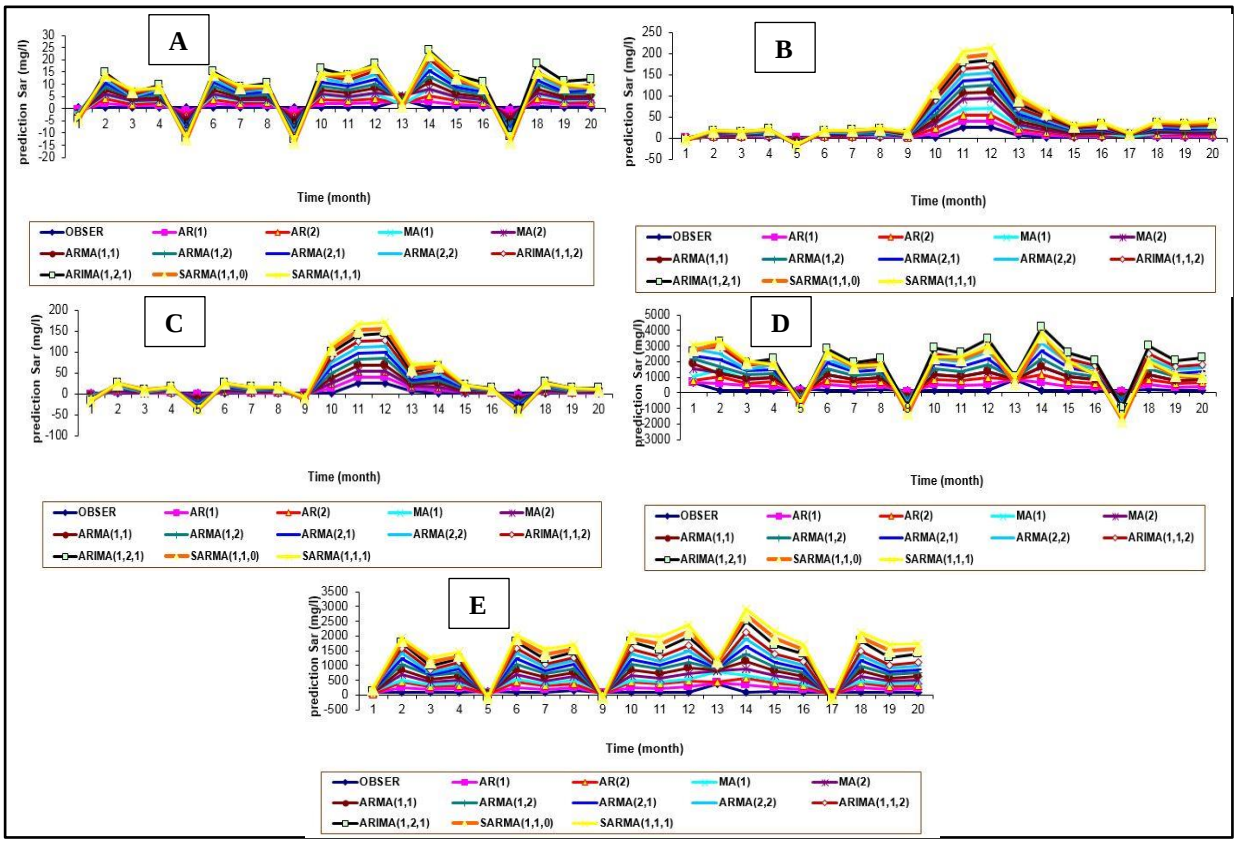

fig. 4- Models prediction versus observed data in cluster 3
A: Cl; B: Na; C: SO_4; D: TDS; E: TH

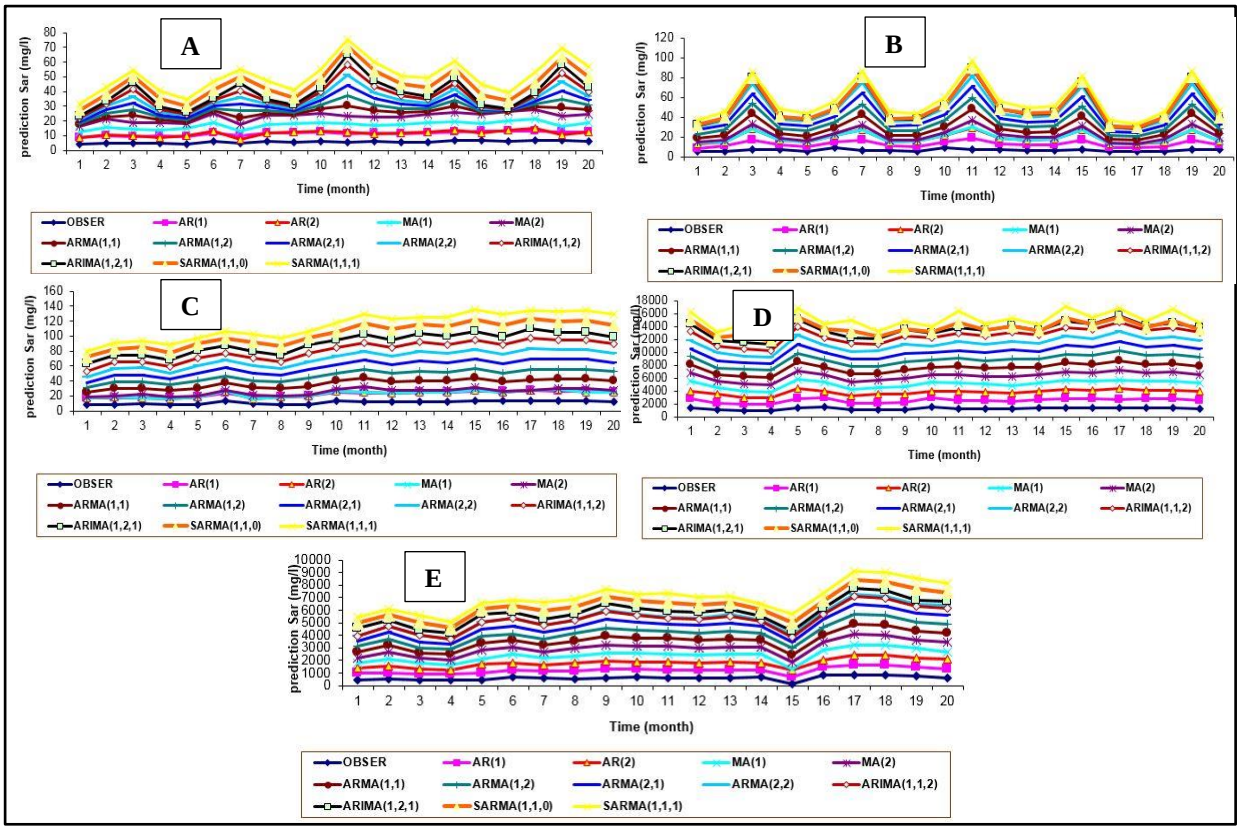

fig. 5- Models prediction versus observed data in cluster 4
A: Cl; B: Na; C: SO_4; D: TDS; E: TH

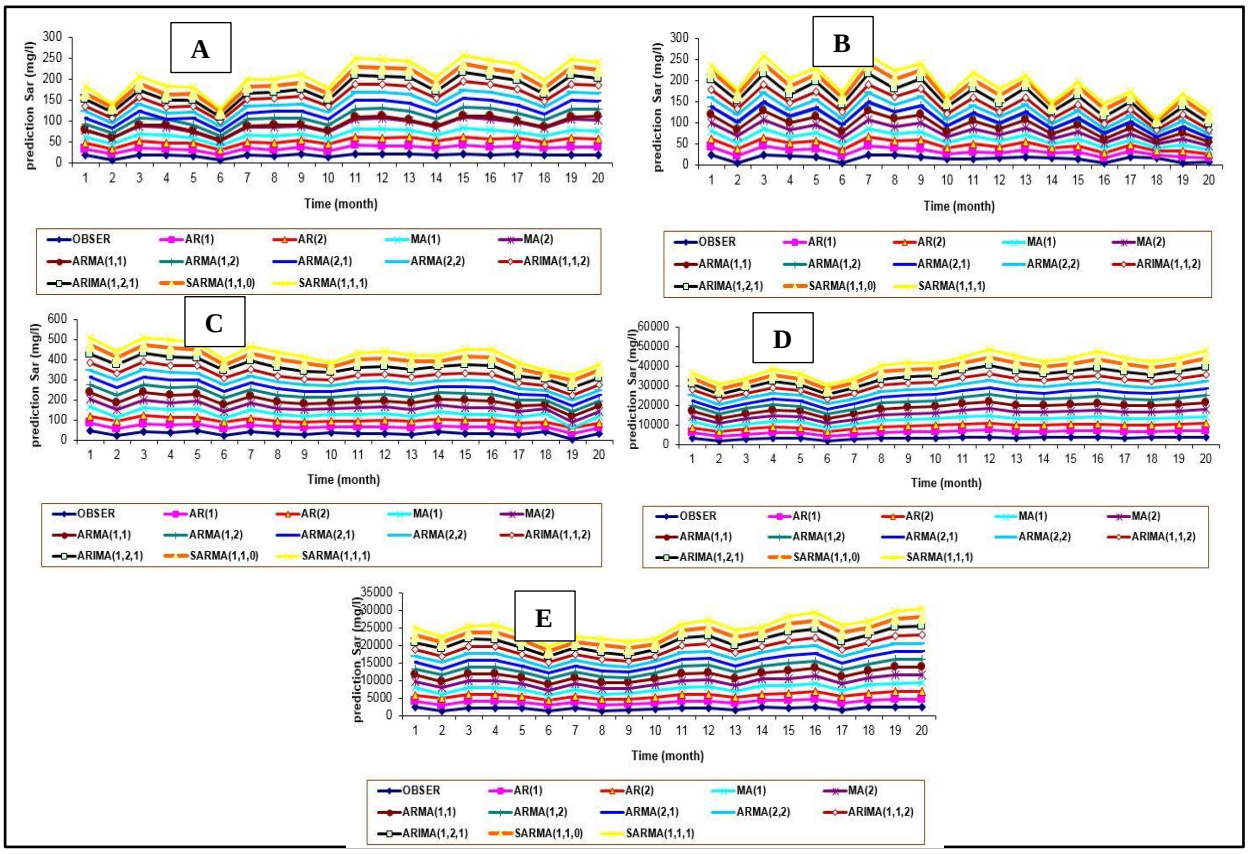

fig. 6- Models prediction versus observed data in cluster 5
A: Cl; B: Na; C: SO_4; D: TDS; E: TH

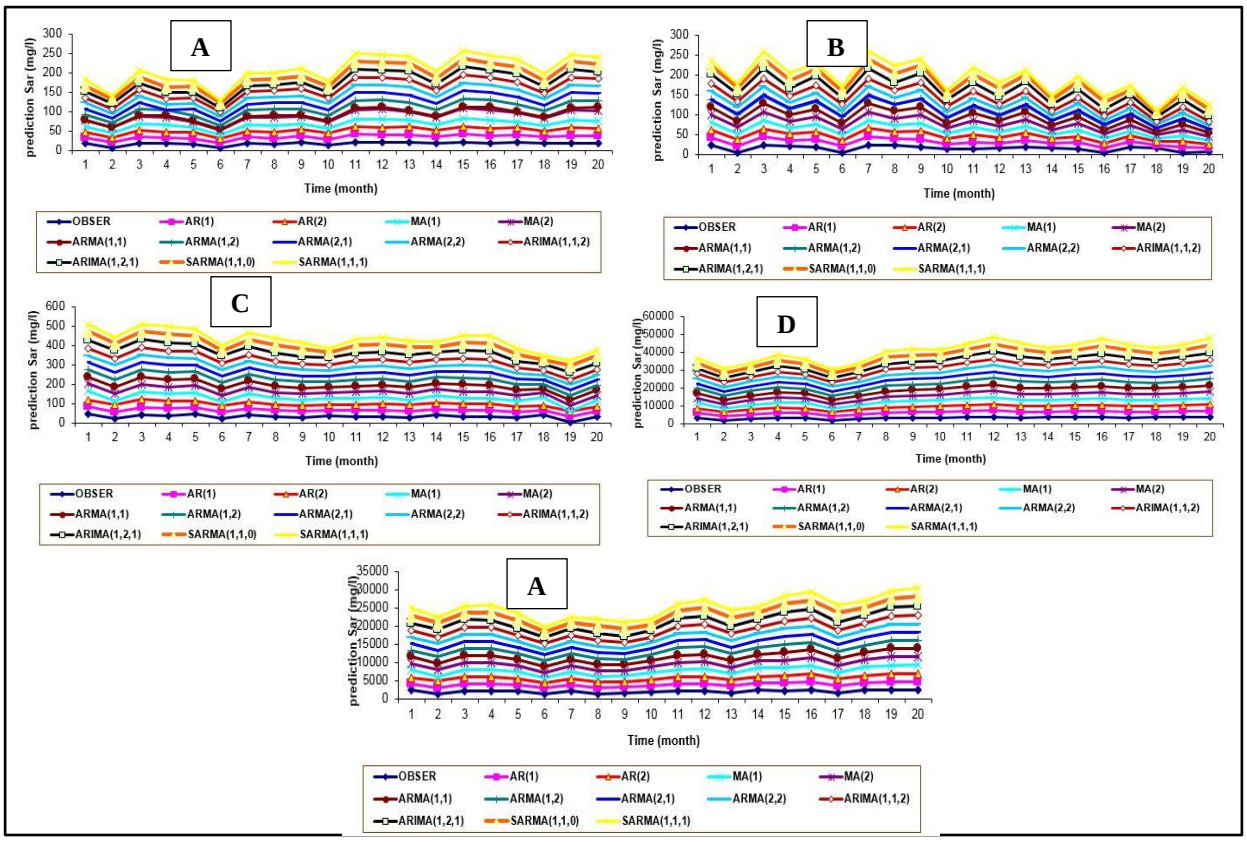

fig. 7- Models prediction versus observed data in cluster 6
A: Cl; B: Na; C: SO_4; D: TDS; E: TH

AIC and R^2 criteria for model selection were used to make accurate predictions. Using the results illustrated in Table 1 and AIC and R^2 coefficient values for the different models, the final models of quality parameters were determined for each cluster (Table 2).

In these models, R^2 and AIC had respectively the maximum and the minimum values, and the absolute value for the parameters of the selected models did not exceed 1. To evaluate the normalization of the predicted data, the Kolmogorov-Smirnov test was used. The results of normalizing the data indicated the normalization of the predicted data. The results showed that the predicted average for Cl and Na, in 2027 in all clusters will increase compared to their value in 2018. Also, the average of predicted SO_4 in the first to fifth clusters in 2027 compared to that in 2018 will rise and it will decline in the fifth cluster. Increasing Cl, Na and SO_4 are due Existence of gypsum, calcareous and evaporitic formations in the region. The most important reasons for the increase in Na are the decrease in groundwater reserves and the decrease in precipitation. The TDs average will rise in the first to third clusters and it will reduce in the fourth, fifth, and sixth clusters. Increased TDs reduces the quality of groundwater. In addition, the average of the predicted TH in the first, second, third, and fifth clusters will increase compared to that in 2018, while it will drop in the fourth and sixth clusters. This increase indicates that the water hardness of the plain is increasing. The reason for this difficulty is gypsum, calcareous and evaporitic geological formations (Table 3).

Tabel 2- Cluster final model Time Series

cluster	Parametr (mg/l)	Models
1	CL	ARMA (2,2)
	TDS	AR (1)
	TH	AR (1)
	NA	AR (1)
	SO_4	AR (1)
2	CL	AR (1)
	TDS	AR (1)
	TH	AR (1)
	NA	AR (1)
	SO_4	AR (1)
3	CL	AR (1)
	TDS	AR (1)
	TH	AR (1)
	NA	AR (1)
	SO_4	AR (1)
4	CL	AR (1)
	TDS	AR (1)
	TH	AR (1)
	NA	AR (1)
	SO_4	AR (1)
5	CL	AR (1)
	TDS	AR (1)
	TH	AR (1)
	NA	AR (1)
	SO_4	AR (1)
6	CL	ARMA (2,2)
	TDS	AR (1)
	TH	AR (1)
	NA	AR (1)
	SO_4	AR (1)

Tabel 3- A comparison of ground-water changing trends in the 2018 and 2027 years

cluster	Parametr (mg/l)	2018	2027
	CL	23.4	24.4
	TDS	1107.6	1118.4
1	TH	620.4	622.2
	NA	23.24	23.84
	SO_4	28.74	29.81
	CL	21.2	23.7
	TDS	1090.6	1777.39
2	TH	711.4	1096.1
	NA	21.2	22.01
	SO_4	30.12	39.9
	CL	18.88	20.01
	TDS	153.6	186.47
3	TH	90.2	109.8
	NA	20.31	21.58
	SO_4	18.64	21.55
	CL	32.98	34.02
	TDS	2715	2705.9
4	TH	1726.4	1717.4
	NA	26.74	28.85
	SO_4	38.82	40.5
	CL	21.72	22.72
	TDS	817.4	811.3
5	TH	380.4	382.16
	NA	23.84	24.5
	SO_4	25.43	26.10
	CL	28.1	29.02
	TDS	1949.2	1921.09
6	TH	1219.4	1213.9
	NA	23.02	23.09
	SO_4	34. 7	33.06

A time series model was used in this study to predict the quality parameters of groundwater. To this end, two models of AR (1) and ARMA (2,2) for Ground-water quality parameters in Dehloran Plain were identified through the autocorrelation and partial correlation functions. The fitness of the model was confirmed by the analysis of the remaining fitted model. The comparison of the data from the prediction provided by different methods and the original data in Table 1 and Figures 2 to 7 show that the selected time series models had an acceptable performance in predicting the time series of ground-water quality parameters. This is in line with the findings by Mirzaee et al (2010), Samadihabashi Samadi (2013), Mirzavand and Ghazavi, (2015) and Mirsanjari and Mohammadyari, (2017).

Unauthorized wells, the type of formations, the chemical geological status and irregular operation of groundwater aquifers are the most important reasons of groundwater quality decrease in Dehloran Plain. Protecting approach should contain pressurized irrigation plans, recharge artificial wells, unauthorized filling of wells and eventually increasing knowledge. as regards the studied area which is located in arid and semi-arid area and dealing with water scarcity, the results shows a clear view of the groundwater quality parameters in future and can be used by experts and planners.

Conclusion

According to the random nature and non-deterministic of the Ground-Water topics, time series are one of the suitable methods with which to anticipation Ground-Water phenomena. In this study, We integrated various time series models for a superior efficiency of anticipation of groundwater Quality. According to the results, it can be said that the integration of time series models has an benefit in terms of groundwater quality anticipating. Given that the region under study

was located in an arid and semi-arid climate with the problems of water shortage, the results of this study can provide a clear vision of ground-water quality parameters in the future which can be used by researchers and planners authorities in the water sector.

Acknowledgement

The authors of this article would like to thank Regional Water Department of Ilam Province for provision of the data.

References

1- Adhikary, S.K., Rahman, M. and Gupta, A.D., 2012. A stochastic modelling technique for predicting groundwater table fluctuations with time series analysis. *International journal of applied science and engineering research*, *1*(2), pp.238-249.

2- Alsalme, A., Al-Zaqri, N., Ullah, R. and Yaqub, S., 2021. Approximation of ground water quality for microbial and chemical contamination. S*audi Journal of Biological Sciences*, 28(3), pp.1757-1762.

3- Anttila, P. and Tuovinen, J. P., 2010. Trends of primary and secondary pollutant concentrations in Finland in 1994–2007. *Atmos. Environ*, 44, pp. 30–41.

4- Behnia, N. and Rezaeian, F., 2015. Coupling wavelet transform with time series models to estimate groundwater level. *Springer Berlin Heidelberg*, 8, pp. 1866-7538.

5- Box, G.E.P., Jenkins, G.M. and Reinsel, G.C., 1994. Time series analysis: forecasting and control. Prentice Hall, Englewood Cliffs.

6- Cadenas, E. and Rivera, W., 2010. Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA-ANN model. *Renew. Energy*, 35, pp. 2732-2738.

7- Chaudhuri, C.h. and Dutta, D., 2014. Mann–Kendall trend of pollutants, temperature and humidity over an urban station of India with forecast verification using different ARIMA models. *Environ Monit Assess*, 186, pp.4719–4742.

8- Dhayachandhran, K.S. and Jothilakshmi, M., 2020. Quality assessment of ground water along the banks of Adyar river using GIS. *Materials Today: Proceedings*, 45 (4), pp.6234-6241.

9- D'Urso, P., De Giovanni, L. and Massari, L., 2015. Time series clustering by a robust auto regressive metric with application to air pollution. *Chemometrics and Intelligent Laboratory Systems*, 141, pp. 107–124.

10- Erdem, E. and Shi, J., 2011. ARMA based approaches for forecasting the tuple of wind speed and direction. *Appl Energy*, 84(2), pp. 1405–1414.

11- Faruk, D., 2010. A hybrid neural network and ARIMA model for water quality time series prediction. *Eng Appl Artif Intell*, 23(4), pp. 586–594.

12- Guler, C., Thyne, G. D., McCray, J. E. and Turner, A. K. 2002. Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. *Hydrogeology journal*, 10, 455-474.

13- Hannan, E.J. 1971. Multiple time series. Wiley, New York.

14- He, Z., Zhang, Y., Guo, Q. and Zhao, X., 2014. Comparative study of artificial neural networks and wavelet artificial neural networks for groundwater depth data forecasting with various curve fractal dimensions. *Water resources management*, 28(15), pp.5297-5317.

15- Karthika, I.N., Thara, K. and Dheenadayalan, M.S., 2018. Physico-Chemical Study of the Ground Water Quality at Selected Locations in Periyakulam, Theni district, Tamilnadu, India. *Materials Today: Proceedings*, 5(1), pp.422-428.

16- Kumar, S. and Sangeetha, B., 2020. Assessment of ground water quality in Madurai city by using geospatial techniques. *Groundwater for Sustainable Development*, 10, p.100297.

17- Liu, H., Wu, H., Lv, X., Ren, Z., Liu, M., Li, Y. and Shi, H., 2019. An Intelligent Hybrid Model for Air Pollutant Concentrations Forecasting: Case of Beijing in China. *Sustainable Cities and Society*. 47, p.101471.

18- Mirsangari, M, M., Zarandian, A., Mohammadyari, F. and suziedelyte-visockiene, j., 2020. Investigation of the impacts of urban vegetation loss on the ecosystem service of air pollution mitigation in Karaj metropolis, Iran. *Journal Environmental Monitoring and Assessment*, 192, pp. 1-23.

19- Mirsanjari, M. M. and Mohammadyari, F., 2017, Application of Time-series Model to Predict Groundwater Quality Parameters for Agriculture: (Plain Mehran Case Study), International Conference on Renewable Energy and Environment November 1-3, 2017 Toronto, Canada.

20- Mirzaee, S., Chitsazan, M., CHinipardaz, R. and Samady, H., 2010. Forecast plain SHAHREKORD using time-series models and examine ways to improve, The first regional conference on optimal utilization of water resources and river basins Karon, 46, pp. 1-8.

21- Mirzavand, M. and Ghazavi, R., 2015. A stochastic modelling technique for groundwater level forecasting in an arid environment using time series methods. *Water resources management*, *29*(4), pp.1315-1328.

23- Mohammadyari, F. 2021. Evaluating and Modeling Selected Ecosystem Services with the Approach of Urban Expansion Impacts on landscape patterns in Karaj metropolis. PhD dissertation. Malayer University Faculty of Natural Resources and Environment

24- Panda, D.K. and Kumar, A., 2011. Evaluation of an over-used costal aquifer (Orissa, India) using statistical approaches. *Hydrol. Sci. Jour*, 56, pp. 486-497.

25- Samadihabashi, R., 2014. Groundwater level prediction using time series model (Case study: Plain Urmia), Master's thesis, University of Urmia, 160pp.

26- Shirmohammadi, B., Vafakhah, M., Moosav,i V. and Moghaddamnia, A. 2013. Application of several data-driven techniques for predicting groundwater level. *Water Resour Manag*, 27, pp. 419–432.

27- Taneja, K., Ahmad, S.h., Ahmad, K. and Attri, S.D., 2017. Time series analysis of aerosol optical depth over New Delhi using BoxeJenkins ARIMA modeling approach. *Atmospheric Pollution Research*, 7, pp. 585-596.

28- Wang, H.R., Wang, C., Lin, X. and Kang, J., 2014. An improved ARIMA model for precipitation simulations. Nonlinear Process. *Geophys*. 21, pp. 1159-1168.