Multi-objective optimization of groundwater monitoring network using a probability Pareto genetic algorithm and entropy method (case study: Silakhor plain)

Mehdi Komasi and Hesam Goudarzi

ABSTRACT

Optimal groundwater monitoring networks have an important role in water resources management. For this purpose, two scenarios were presented. The first scenario designs a monitoring network and the second scenario chooses optimal wells from the existing ones in the study area of the monitoring network. At the first step, a database including groundwater elevation in potential wells was produced using the Kriging method. The optimal monitoring network in the first scenario was determined by preset conventions and found by the non-dominated sorting genetic algorithm (NSGA-II). In the second scenario, the optimal monitoring network was determined by entropy theory through calculating entropy for each of the 29 observation wells. Finally, the first scenario obtained a network with 12 observation stations showing root mean square error (RMSE) value given as 0.61 m. Comparison between entropy of rainfall and groundwater level time series in the first scenario had the same variation. The optimal monitoring network in the first scenario has been able to reduce the number of monitoring stations by 60% in comparison with the existing observation network. The second scenario used entropy theory and calculated the energy of each of the 29 observation wells which obtained a monitoring network with 11 stations.

Key words | entropy, groundwater monitoring network, Kriging, NSGA-II, Silakhor

Mehdi Komasi (corresponding author)
Department of Civil Engineering,
Ayatollah Boroujerdi University,
Boroujerd,
Iran
E-mail: *komasi@abru.ac.ir*

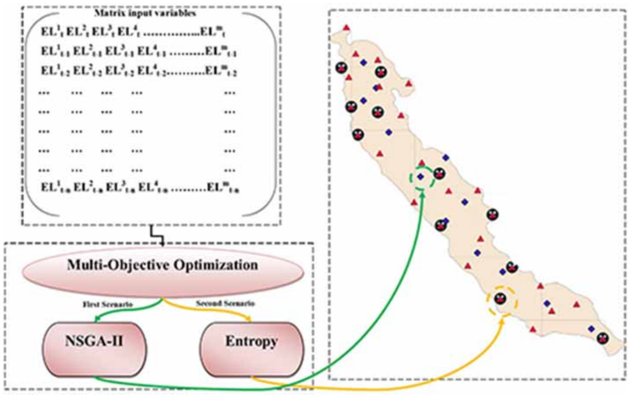
Hesam Goudarzi
Civil Engineering – Water Resource Management,
K. N. Toosy University of Technology,
Tehran,
Iran

HIGHLIGHTS

- Successful management of groundwater resource depends on groundwater monitoring network.
- A new multi-objective simulation-optimization model can design groundwater monitoring networks.
- NSGA-II was used for determining the optimal groundwater monitoring network.
- Entropy theory can determine optimal monitoring stations between existing stations.
- Optimal groundwater monitoring network can help to obtain missing data in non-point.

doi: 10.2166/hydro.2020.061

GRAPHICAL ABSTRACT



INTRODUCTION

Hydrometric monitoring networks play an important role in water resources planning and management. Groundwater monitoring network design can be divided into two streams: the stream designed for groundwater quality monitoring and that designed for groundwater quantity monitoring. This does not mean that one network could not perform both tasks, but it means that the approaches used in the design vary based on the intended use of the network. Hence, a suitable monitoring network needs to maintain and expand due to the importance of hydrometric data for water resources management practices as the availability and reliability of hydrometric data are important for planning and management purposes (Mishra & Coulibaly 2009).

Groundwater monitoring approaches have two main categories, hydrogeological and statistical simulation, variance or probability-based methods (Loaiciga *et al.* 1992). For optimal groundwater management, it is necessary to collect sufficient information from the quantitative and qualitative features of the aquifer. Assessing and anticipating the level of groundwater through specific models help to predict groundwater resources. The design of observation networks is a multi-objective optimization problem. In the case of multi-objective optimization, measurements are highly considered in terms of spatial and temporal frequencies. On the other hand, these calculations require a large amount of infrequent data intervals, resulting in costs going up significantly. To this end, the network must be optimized to a degree that adequately reflects the level of the station in the target area. In order to generalize the measurement results from available wells to other points without statistics, the well network should have a density appropriate to the area (Hosseini & Kerachian 2017; Keum *et al.* 2017; Ayvaz & Elçi 2018). To implement this, different models and algorithms such as land-based methods and optimization algorithms have been used. In recent years, with the demonstration of the capabilities of intelligent models, these models have been applied in groundwater modeling. Intelligent models such as artificial neural networks (ANN) and support vector machines (SVM) are more developed than other methods due to their processing speed, high accuracy, and relatively low cost. There have been a wide range of studies on the design of groundwater monitoring networks (e.g., Bashi-Azghadi & Kerachian 2010; Kollat *et al.* 2011; Barca *et al.* 2013; Reed & Kollat 2013; Khader & McKee 2014; Bhat *et al.* 2015; Júnez-Ferreira *et al.* 2016; Komasi *et al.* 2018; Pourshahabi *et al.* 2018). In this way, Asefa *et al.* (2004) used the SVM method based on Statistical Learning Theory (SLT) to design a long-term groundwater monitoring network to reduce spatial redundancy. In this study, the possibility of generalizing the developed method for the analysis of time redundancy has been mentioned. Another approach in terms of groundwater monitoring networks is entropy theory. In this way, Mogheir *et al.* (2003) used entropy theory in research to design a site for sampling groundwater quality monitoring networks. They defined the collection of groundwater quality variables using the information transfer model and the correlation model which determine the location transport model of the wells better than the correlation model. In another study, Leach *et al.* (2016) considered feeding groundwater as the basis for the design of a monitoring network using the dual entropy optimization model to design an optimal groundwater monitoring network. Utilizing entropy is a suitable method for finding common information among observation wells for optimization. Some studies have a scattered method for designing a groundwater monitoring network. Li & Hilton (2005) used the method of optimizing ants to solve the problem of optimizing the design of the groundwater monitoring network. In this study, while investigating a pollutant in the quality monitoring network, only a spatial optimization has been achieved. Also, Guo *et al.* (2011) compared several methods, sampling, hydrological analysis, and a method of non-homogeneous mean level approach to optimizing the observation network of groundwater in the field of designing a network for observing groundwater level, which results in the superiority of this method over other methods. In another study, Bhat *et al.* (2012) used the usual Kriging method to design a groundwater monitoring network in Florida. The results showed that 58 wells should be reduced to 44 wells in the region to reduce the estimation error. Zhou *et al.* (2013) utilized mapping to optimize the observation network for groundwater level. The results show that these maps can distinguish the regions with unique spatial and temporal characteristics, and can be used to design the grid. Also, Esquivel *et al.* (2015) used multi-criteria analysis to design an optimal groundwater table monitoring network. A multi-criteria analysis has been implemented with the help of GIS on an aquifer in Mexico. The analytic hierarchy process (AHP) method was used to weigh the criteria and the stability rate was 0.08.

Multi-objective optimization algorithms have been used in network design. For example, Kollat & Reed (2007) used the epsilon-dominance non-dominated sorted genetic algorithm II (NSGA-II) to design long-term groundwater monitoring networks based on cost, contaminant estimation error, uncertainty, and contaminant mass error. Later, Kollat *et al.* (2008) studied the next generation epsilon-dominance hierarchical Bayesian optimization algorithm (ϵ-hBOA) and its application to groundwater network design and found that it was superior for network design due to its model building capability. In a further study, Dhar & Patil (2012) presented a method for designing a groundwater quality monitoring network under cognitive uncertainty conditions in predicting the mass flow through the second version of the genetic algorithm (GA) and Kriging. The results of this study showed the effectiveness of the proposed method with two approaches. Leach *et al.* (2015) added objective functions for stream flow signatures and indicators of hydrologic alteration to explicitly consider the shape of the hydrograph at different gauging sites when building a hydrometric network. This modification increases the variety of flow regimes seen in optimal networks. Luo *et al.* (2016) aimed at optimizing the long-term groundwater monitoring networks using PPGA. In this research, four objective functions were considered for optimization of the observation network. Using the PPGA, efforts to obtain the Pareto-front of four targets have been considered.

According to the literature review, the present study aims to assess the optimal groundwater monitoring network with the application of GA and Kriging as interpolation methods regarding the case study while focusing on the Silakhor plain located in Iran. This study designs an optimal groundwater network in two scenarios. The first scenario redesigns and evaluates the groundwater monitoring network by entropy theory, and the second scenario chooses optimal wells from the existing wells by considering entropy of each station in the study area concerning the groundwater monitoring network.

METHODOLOGY

Study area

The data used in this paper are from the Silakhor plain located in the west of Iran in Lorestan province. The Silakhor plain is an upper district of Karun River basin. Starting from Garrin Mountain, Gelerood River passes southern Borujerd and receives more creeks in lower Silakhor where it is called the Tireh River and is an artesian aquifer. The Silakhor plain is located between 33° 15′ 46″ and 34° 10′ 00″ northern latitude and 48° 28′ 53″ and 49°

30′ 25″ eastern longitude. The Silakhor aquifer area is 2,545.8 km^2 (Figure 1). For modeling and designing the optimal monitoring network, the statistical data of the Silakhor plain for a ten-year period were used. The statistical characteristics of its stations are presented in Table 1.

Figure 2 shows monthly rainfall time series and groundwater level changes in the Silakhor plain from 2003 to 2013. There are 29 different stations in this plain. The aquifer's parameter is measured by a piezometric well in each area.

Kriging

Kriging is a class of statistical techniques for optimal spatial prediction. It has been used in many other disciplines, including agriculture, mining, and environmental sciences. Kriging is a probabilistic predictor assumed as a statistical model for the data. Kriging predictors have standard errors that quantify the uncertainty associated with the predicted values. Kriging predictors are called optimal predictors because the prediction error is minimized and, on average, the predicted value and the true one coincide. Kriging uses a semivariogram (a function of the distance and direction separating two locations) to quantify the spatial dependence in data (Krivoruchko 2012).

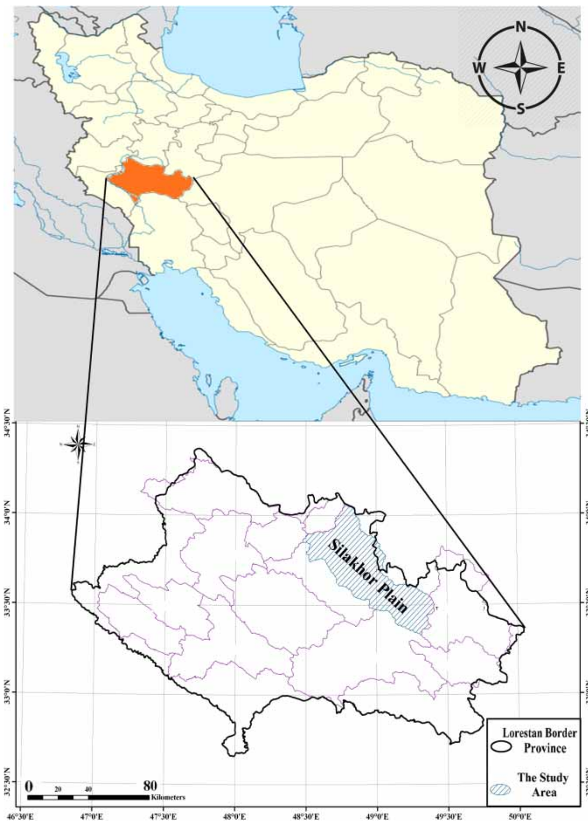


Figure 1 | Location of the Silakhor plain in Lorestan province.

Table 1 | Statistic characteristics of groundwater levels data for the Silakhor plain

Monthly time series	Average (m)	Min (m)	Max (m)
Groundwater level (m)	1,501.3	1,493.2	1,507.4
Rainfall (mm)	34.1	0.0	176.4

Inverse distance weighing (IDW)

The IDW is one of the most applied and deterministic interpolation techniques in the field of hydrology. IDW estimations were made based on nearby known locations. The weights assigned to the interpolating points are the inverse of their distance from the interpolation point. Consequently, the close points are made up to have more weights (more impact) than distant points and vice versa. The known sample points are implicit to be self-governing from each other (Preziosi *et al.* 2013).

Non-dominated sorted genetic algorithm (NSGA-II)

Genetic algorithms can find optimum results in a multi-objective optimization problem. In this way in the GA, unlike other classical methods, a random population solution is selected. Each solution for the problem is represented as a set of parameters which are known as genes. Joining genes creates a binary bit string of values, denoting each member of the population referred to as a chromosome. A chromosome evolves through iterations, called generation (Gen *et al.* 1997).

To solve a multi-objective problem, two different strategies can be applied either by aggregating the weighted objectives to form a single objective problem or by finding the multiple solutions on a Pareto front to generate the best alternatives. The first method provides the leverage to solve the problem as a single objective, but assigning weights can be challenging for most problems. The second method requires solving the problem for all the objectives to obtain the non-dominated or Pareto set of solutions. Non-dominated sets of solutions consist of feasible optimal solutions (Puri *et al.* 2017).

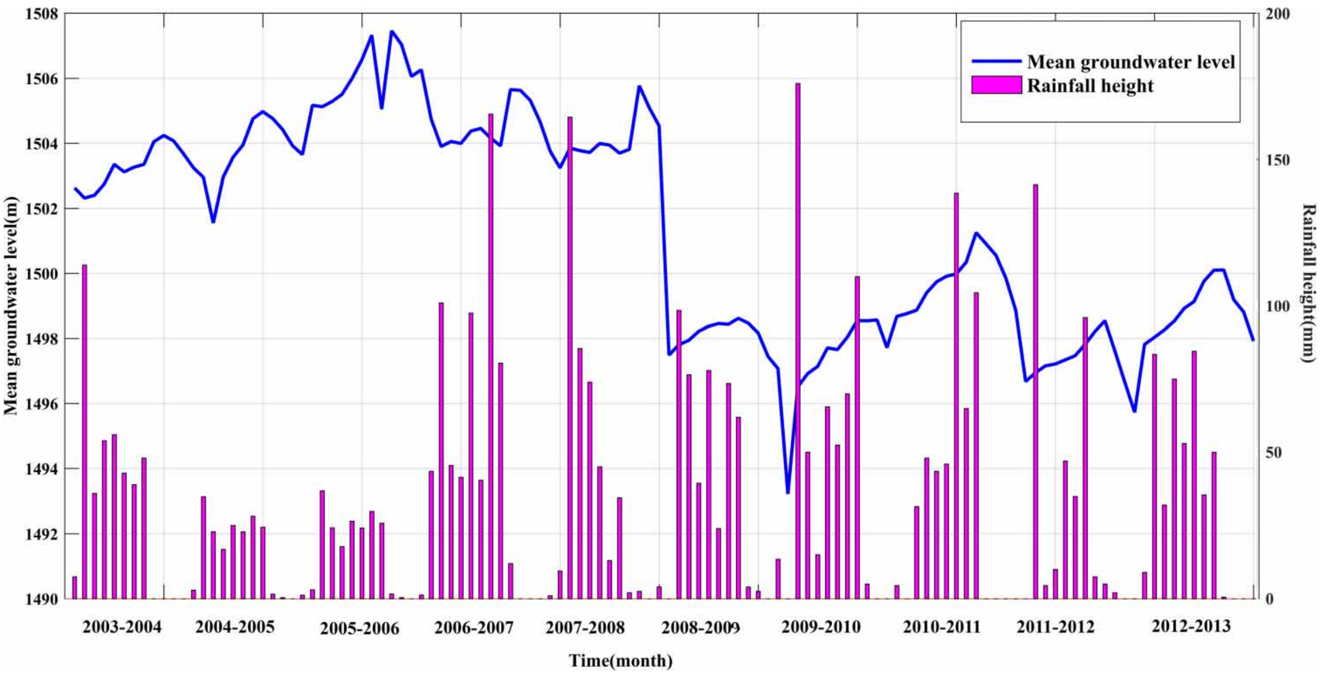


Figure 2 | Time series of mean groundwater and rainfall time series in the Silakhor plain.

Entropy method

Entropy has been discussed previously in the literature for water resources and hydrology (Alfonso *et al.* 2014; Komasi & Sharghi 2019). It is a nonparametric method with no prior assumptions about the data being used (Mishra & Coulibaly 2010; Samuel *et al.* 2013). In information theory, the Shannon entropy (or marginal entropy) uses $H(X)$ of a discrete random variable (monitoring well); X provides a measure of the information content from a finite sample in bits for a single station shown as follows:

$$H(x) = -\sum_{i=1}^{n} P(x_i) \log P(x_i) \tag{1}$$

where, $P(x_i)$ is the probability of the value x_i in the bin of values n. The number of events (binned values) n was determined by discretizing the groundwater table level time series over a number of bins (Shannon 1948).

Objective functions

Groundwater monitoring network design optimization needs objective functions that limit optimization. For this purpose, two objective functions are defined in groundwater monitoring network design: (1) minimizing the number of observation wells as the representative of maintenance costs and water level readings and (2) minimizing root mean square error (RMSE) between observation value and simulation value on potential wells for accuracy of monitoring network design. These functions are suggested in Equations (2) and (3).

$$Minimization\ \{f_1(number\ of\ wells) = f_1(X_1, X_2, X_3, \ldots, X_N)\} \tag{2}$$

$$Minimization\ \{f_2(RMSE) = f_2(RMSE_1, RMSE_2, RMSE_3, \ldots, RMSE_N)\} \tag{3}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T}\sum_{i=1}^{N}(Hobs_i^t - Hest_i^t)^2}{TN}} \tag{4}$$

where, f_1 is the first objective function, X is the representative of potential wells, N is the number of potential wells, f_2 is the second objective function, $Hobs_i^t$ is the observation groundwater level on i point in t period of time, $Hest_i^t$ is the estimation of groundwater level on i point and in t period,

T is the total period of time, and $RMSE$ is the root mean square error between $Hest_i^t$ and $Hobs_i^t$.

Optimization processes

In this study, monthly observation groundwater levels of the Silakhor plain from 2003 to 2013 were used in the optimal model. Using the Kriging method and time series data obtained from observation wells, potential wells were interpolated in ten years by Arc GIS. This information is considered as observation data with simulation data. Simulation data have been produced by the IDW method for ten years in potential wells by MATLAB software. According to Equation (4), the second step for determining the optimal groundwater monitoring network is to calculate RMSE between observation and simulation values for each well as the objective function. After these steps, the optimization algorithm starts to solve the optimization problem and find the optimal monitoring network. The optimization process is illustrated in Figure 3. Results of interpolation values in potential wells during ten years are shown in Figure 4.

RESULTS AND DISCUSSION

In this research, two scenarios were examined: (1) redesigning the groundwater monitoring network in the Silakhor plain and (2) choosing the optimal wells from the existing ones in the study area for the groundwater monitoring network. Multi-objective optimization design problems of the groundwater monitoring network were solved by NSGA-II. The NSGA-II parameters include number of iterations, total population, crossover percent, and mutation percent given as 200, 1,000, 0.7, and 0.2, respectively.

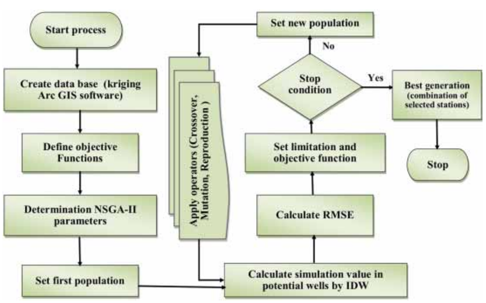


Figure 3 | Optimization processes for groundwater network design.

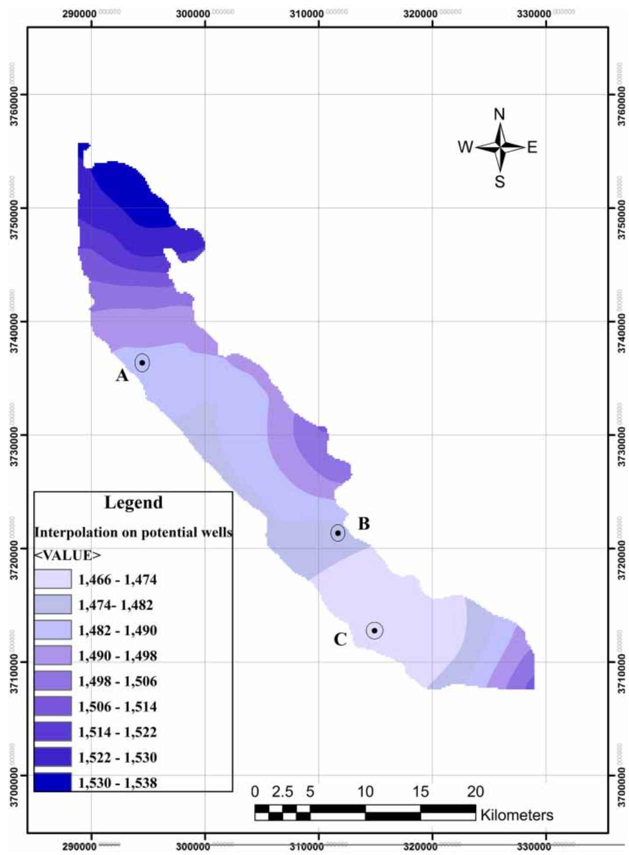


Figure 4 | Interpolation of groundwater level in the Silakhor plain for 2003–2013 by using existing monitoring network. A, B, and C are used for cross validation.

First scenario

In the first scenario, 29 wells were used and interpolated as observation wells in the Silakhor plain to determine the groundwater level on potential wells. Figure 5 shows the existing and proposed potential wells in the Silakhor plain. To specify potential wells' locations, ASTM standard D 5092-90 (1996) was used. In this regulation, a grid of 492 points with 1 km distance between each point was used. A groundwater monitoring network was designed with potential wells. After determining groundwater level in potential wells as observation data, the simulated groundwater level was calculated by the IDW method in Matlab software. A matrix was introduced into the optimization model of 492 rows and 120 columns presenting the number of potential stations and monthly groundwater level in the potential stations, respectively. Then, NSGA-II looks for the optimal stations among the 492 potential stations and specifies the number of optimal stations.

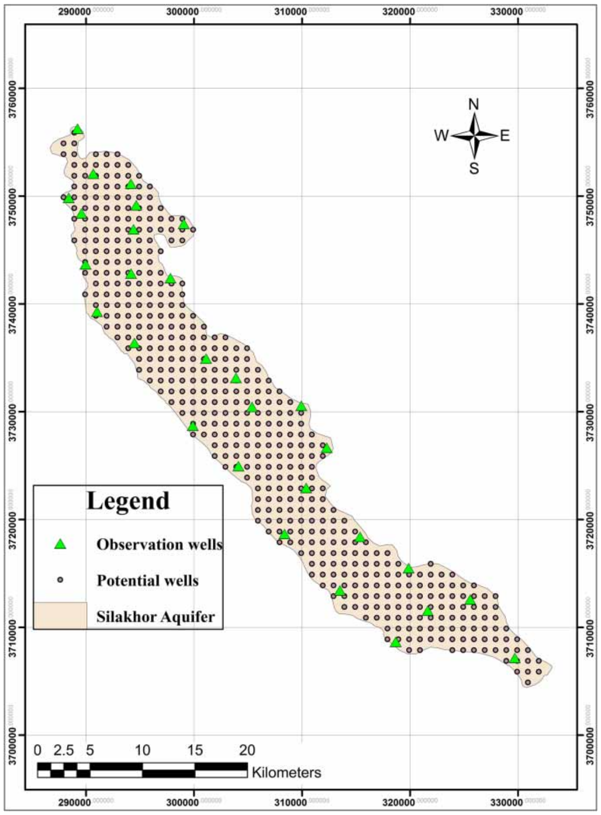


Figure 5 | Location of observation wells and potential wells in the Silakhor plain.

According to the results of the calculations conducted by NAGA-II in Matlab software, the optimal monitoring networks were obtained with 12 observation stations. Since the response obtained from the optimizer algorithm consists of a combination of the number of stations and their position, it maps the response optimizer algorithm in the form of a Pareto-front which represents the number of optimal stations versus RMSE. The Pareto-front for the first scenario is shown in Figure 6.

As shown in Figure 6, 12 optimal wells have RMSE value of 0.61 m. Figure 7 shows the spatial distribution of 12 optimal monitoring stations. As shown in Figure 7, the program identifies a monitoring network with 12 stations, which has reduced the number of stations by 60% as compared to the existing monitoring network (i.e., 29 observation wells). The RMSE corresponding to the number of 29 stations is 0.54 m. At first glance, it can be concluded that increasing the number of monitoring stations will increase network accuracy and reduce the value of RMSE. But the other objective function for minimizing network monitoring costs avoids network recommendation with more stations because increasing the number of stations will increase the costs of building and monitoring

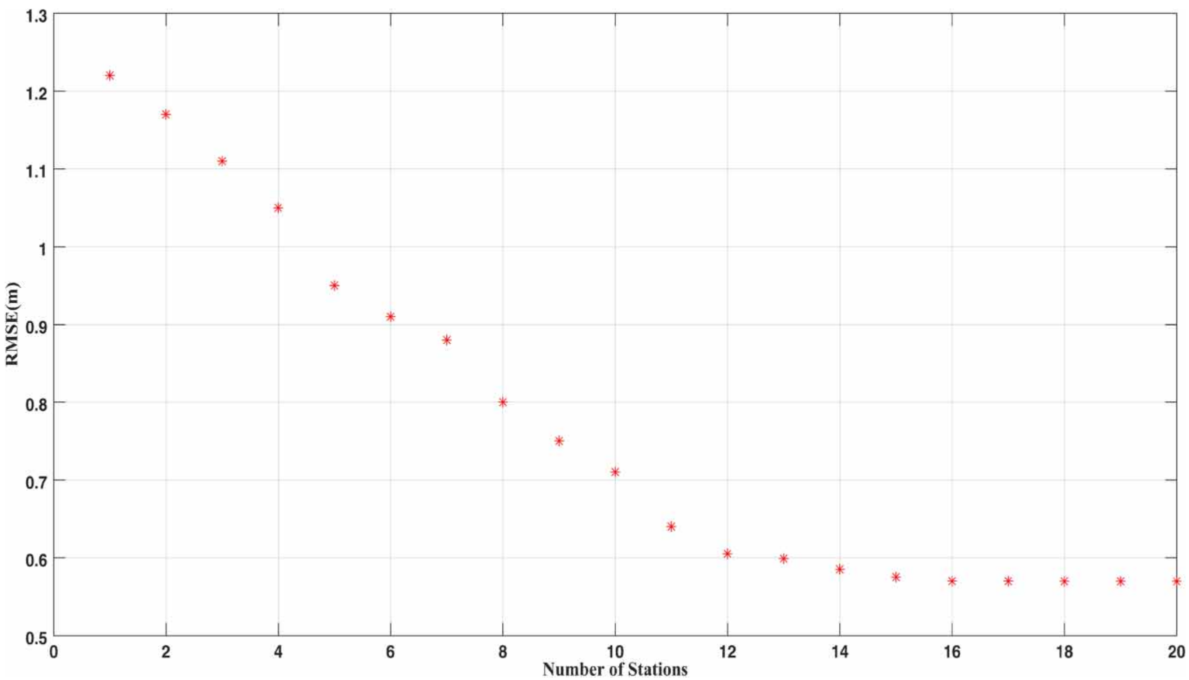


Figure 6 | Convergence graph in first scenario.

the network. Comparing the network with 12 stations with the existing network (29 stations) shows more stations (17 stations) could reduce only 12% of the network error. Another important point is that the new monitoring network, by changing the location of monitoring stations, improves spatial distribution and increases forecasting accuracy of groundwater level in an area which has no data. The optimal network in the northwest of the aquifer recommends three monitoring stations. Also, a more precise survey of this area shows a wetland that has a significant impact on groundwater level in the Silakhor aquifer; therefore, recommending three monitoring stations for this area is justifiable (see Figure 7).

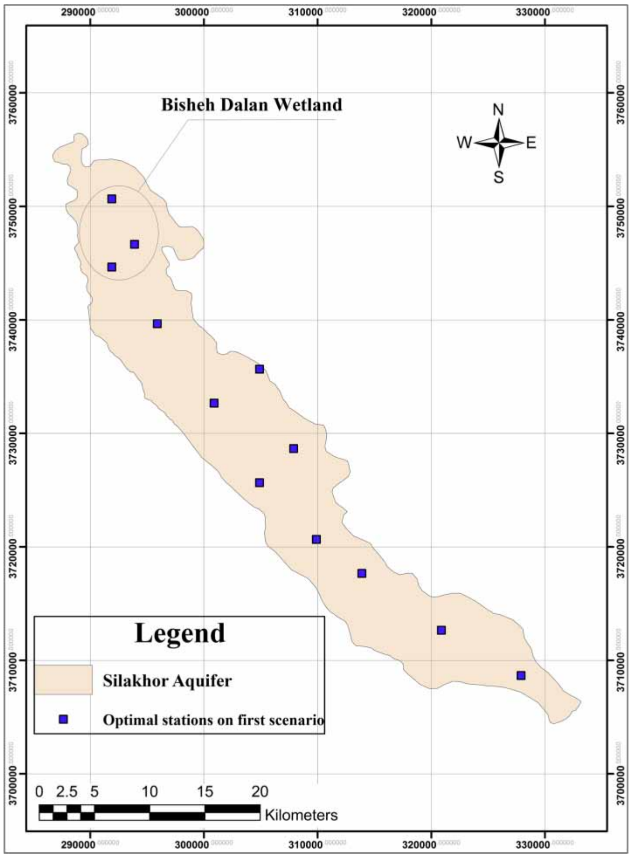


Figure 7 | Spatial distribution of 12 optimal monitoring stations in the first scenario.

After designing the groundwater monitoring network using the IDW method and NSGA-II optimization algorithm, 12 optimal stations have been evaluated by the entropy method. For this purpose, a groundwater level time series in optimal stations was chosen and then it calculated the signal energy of stations in a period of time. After calculating the signals' energy, the entropy is calculated according to Equation (1). The amount of rainfall entropy and its variations at different intervals affect the amount and diversity of the groundwater level entropy over the same time period. Comparison of the groundwater level entropy obtained from these 12 stations with rainfall entropy in the Silakhor plain indicates whether the data obtained from these stations can accurately identify these changes or these stations' information is inadequate. To investigate this, the monthly rainfall and groundwater levels were arranged over a period of ten years. Table 2 shows the entropy value for the rainfall time series, and Table 3 shows the entropy of groundwater level value for each selected station.

Comparison of Tables 2 and 3 shows that entropy variations for each selected station are similar to the entropy variations of the rainfall time series. This indicates that the selected stations react to the entropy changes of the precipitation itself. In other words, the readings of groundwater level at these stations in different time periods are in accordance with the changes in rainfall and the results are far from real and not expected. Figure 8 shows the zoning of the study area based on the 12 optimal stations. As shown in Figure 8, the spatial distribution has been able to provide a good estimate of the groundwater level in the Silakhor plain, which results in precision and accuracy in modeling at unmatched locations as well as management plans in the context of groundwater resources management. As an example, three monitoring stations between 29 existing stations were used to estimate groundwater level by the first scenario. Table 4 shows the results of groundwater level in these points according to observation data and the first scenario. Figure 9 shows the difference between the interpolation map in the first scenario and the existing network.

Table 2 | Entropy for the rainfall time series in the Silakhor plain

Period of time	2003–2004	2004–2005	2005–2006	2006–2007	2007–2008	2008–2009	2009–2010	2010–2011	2011–2012	2012–2013
Rainfall time series	0.28	0.37	0.48	0.52	0.67	0.71	0.63	0.85	0.42	0.37

Table 3 | Entropy for the groundwater time series for 12 stations in the first scenario

Station number	Period of time										CC
	2003–2004	2004–2005	2005–2006	2006–2007	2007–2008	2008–2009	2009–2010	2010–2011	2011–2012	2012–2013	
18	0.18	0.27	0.38	0.42	0.56	0.59	0.51	0.73	0.33	0.29	0.94
69	0.25	0.33	0.44	0.48	0.63	0.67	0.59	0.82	0.41	0.36	0.97
136	0.29	0.35	0.48	0.53	0.66	0.71	0.65	0.87	0.44	0.37	0.95
163	0.27	0.36	0.47	0.51	0.66	0.73	0.62	0.83	0.41	0.36	0.96
194	0.14	0.23	0.32	0.38	0.53	0.57	0.52	0.74	0.31	0.25	0.92
241	0.76	0.83	0.94	0.98	1.14	1.20	1.13	1.34	0.9	0.86	0.94
279	0.06	0.15	0.26	0.30	0.45	0.49	0.44	0.63	0.17	0.14	0.93
321	0.32	0.40	0.52	0.55	0.72	0.74	0.66	0.88	0.47	0.41	0.92
365	0.56	0.62	0.73	0.77	0.94	0.96	0.91	1.12	0.69	0.64	0.89
409	0.26	0.35	0.46	0.48	0.63	0.67	0.59	0.82	0.35	0.29	0.87
427	0.09	0.16	0.26	0.3	0.43	0.48	0.41	0.60	0.19	0.14	0.95
468	0.02	0.09	0.21	0.24	0.42	0.46	0.39	0.60	0.16	0.12	0.97

CC, correlation coefficient.

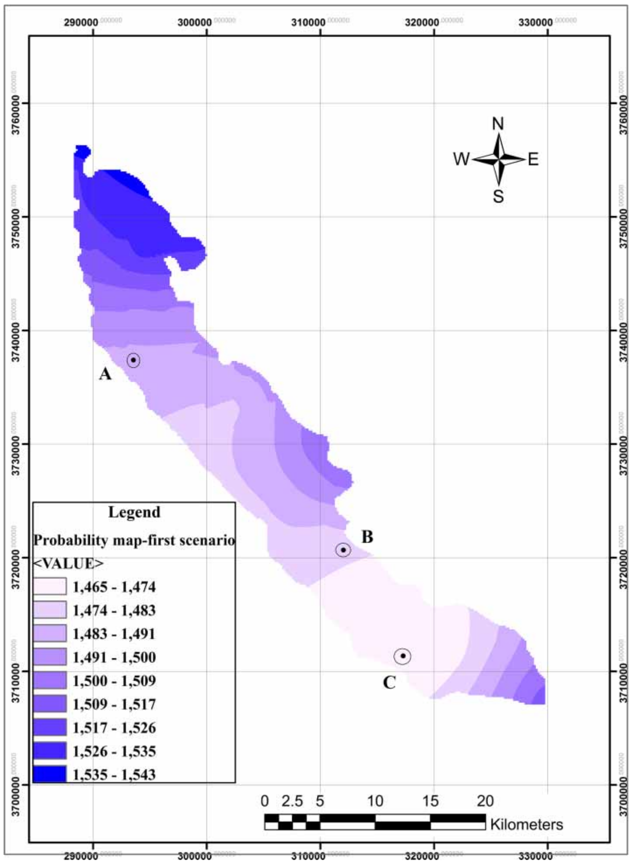


Figure 8 | Interpolation map of optimal network for augmenting the existing ones in the first scenario.

Table 4 | Groundwater level for three stations for cross validation in the first scenario

Station name	Observation level	First scenario level	Error %
A	1,490.1	1,485.3	4.8
B	1,498.5	1,494.6	3.9
C	1,470.7	1,466.2	4.5

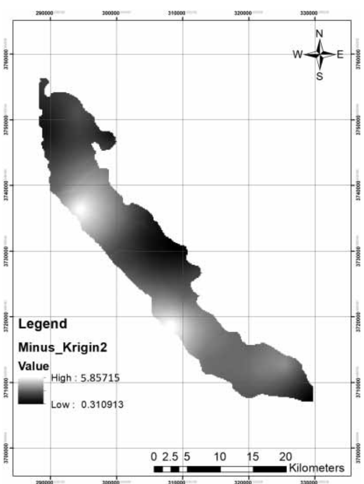


Figure 9 | Difference between interpolation map in the first scenario and the existing network.

Second scenario

The second scenario chooses optimal wells from the 29 observation wells shown in Figure 5. In the second scenario, the monitoring network was obtained from the existing observation wells in the Silakhor plain. For this purpose, 29 wells' data are introduced into the model to determine the optimal network in accordance with the objective functions. In this scenario, the entropy of these 29 stations was calculated; then, these stations were classified as the maximum amount of entropy. After that, according to the variations of each station, the optimal network was chosen. In other words, comparison between the entropy of rainfall time series and groundwater level in 29 stations can decide on the optimal monitoring network. According to Venetsanou *et al.* (2016) and Abdullahi & Garba (2015) who investigated the effect of rainfall on quantitative and qualitative changes in aquifer water, the time dependency of rainfall is effective in aquifer

Table 5 | Entropy for the groundwater time series for 29 stations in the second scenario

	Period of time											
Station number	2003–2004	2004–2005	2005–2006	2006–2007	2007–2008	2008–2009	2009–2010	2010–2011	2011–2012	2012–2013	CC	
1	0.16	0.25	0.36	0.40	0.54	0.57	0.49	0.71	0.31	0.27	0.94	*
2	0.14	0.22	0.30	0.37	0.49	0.55	0.48	0.69	0.31	0.25	0.96	*
3	0.02	0.01	0.06	0.00	0.52	0.38	0.11	0.14	0.68	0.07	0.26	
4	0.08	0.14	0.27	0.31	0.38	0.49	0.41	0.66	0.23	0.16	0.84	*
5	0.28	0.37	0.48	0.52	0.67	0.71	0.63	0.84	0.42	0.37	0.92	*
6	0.01	0.01	0.44	0.00	0.01	0.01	0.43	0.25	0.05	0.00	0.32	
7	0.14	0.49	0.22	0.14	0.63	0.65	0.31	0.75	0.34	0.42	0.67	
8	0.12	0.21	0.28	0.34	0.39	0.47	0.51	0.72	0.32	0.23	0.95	*
9	0.01	0.09	0.48	0.02	0.55	0.26	0.01	0.54	0.39	0.01	0.55	
10	0.01	0.05	0.00	0.10	0.10	0.02	0.03	0.27	0.03	0.00	0.69	
11	0.06	0.15	0.26	0.3	0.45	0.49	0.4	0.6	0.17	0.14	0.96	*
12	0.03	0.68	0.86	0.46	0.13	0.18	1.14	0.75	0.62	0.40	0.19	
13	0.02	0.19	0.58	0.02	0.64	0.45	0.72	0.30	0.10	0.68	0.34	
14	0.31	0.37	0.48	0.52	0.69	0.71	0.65	0.87	0.44	0.39	0.93	*
15	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.15	0.00	1.25	0.24	
16	0.02	0.01	0.00	0.00	0.03	0.05	0.00	0.03	0.00	1.27	0.30	
17	0.09	0.15	0.28	0.33	0.46	0.51	0.45	0.67	0.24	0.17	0.86	*
18	0.06	0.04	0.54	0.03	0.02	0.01	0.02	0.36	0.01	0.14	0.12	
19	0.47	0.55	0.67	0.7	0.85	0.89	0.81	1.03	0.62	0.55	0.91	*
20	0.48	0.06	0.01	1.33	0.11	0.46	0.17	0.03	0.03	0.03	0.10	
21	0.07	0.00	0.03	0.07	0.00	0.05	0.07	0.04	0.04	0.06	0.11	
22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.01	0.54	
23	0.01	0.01	0.05	0.00	0.03	0.02	0.00	0.00	0.00	0.04	0.17	
24	0.10	0.00	0.04	0.03	0.00	0.09	0.06	0.08	0.02	0.02	0.25	
25	0.00	0.02	0.13	0.02	0.09	0.27	0.00	0.30	0.00	0.02	0.71	
26	0.25	0.32	0.43	0.47	0.63	0.69	0.62	0.83	0.39	0.35	0.89	*
27	0.00	0.01	0.00	0.00	0.00	0.01	0.11	0.05	0.33	0.02	0.28	
28	0.20	0.29	0.4	0.44	0.59	0.63	0.55	0.76	0.34	0.29	0.91	*
29	0.01	0.02	0.14	0.02	0.20	0.06	0.04	0.06	0.03	0.38	0.21	

feeding. In fact, the entropy value of the precipitation and its changes in different time intervals affect the value and variations of groundwater level entropy during the same time intervals. In this sense, the value of entropy obtained in each station in the monitoring network must be in good agreement with the time series of entropy changes. If the value of entropy obtained from the station is not proportional to the entropy variation of the time series, it indicates that the station cannot show the water balance based on the time dependence of the time series and the feeding of the aquifer.

Comparison of Tables 2 and 5 shows that entropy variations for 11 stations from the existing stations are similar to the entropy variations of the time series and, also, the entropy of these stations is higher than other stations. This indicates that the selected stations react to the entropy changes of the precipitation itself. In other words, the readings of groundwater level at these stations in different periods of time are in accordance with the changes in rainfall and the results are far from real and unexpected ones. These 11 stations are marked with stars in Table 5. The number of stations at some intervals indicates the value of entropy as zero. This value indicates that the time interval of the station in accordance with the precipitation fluctuations cannot be a precise measure of the station level in the Silakhor plain. The results of the second scenario represent a network with 11 stations among the 29 available ones. As shown in this table, 18 stations do not have the same variations according to the entropy of rainfall time series and have a few entropy variations between other stations. For this reason, these stations cannot affect the monitoring network. An investigation into the entropy of the network with 11 stations compared to the network with 29 observation stations has shown that the network with 11 stations could accommodate variations in the value of entropy according to the trend of rainfall time series. The optimal monitoring stations in the second scenario are shown in Figure 10.

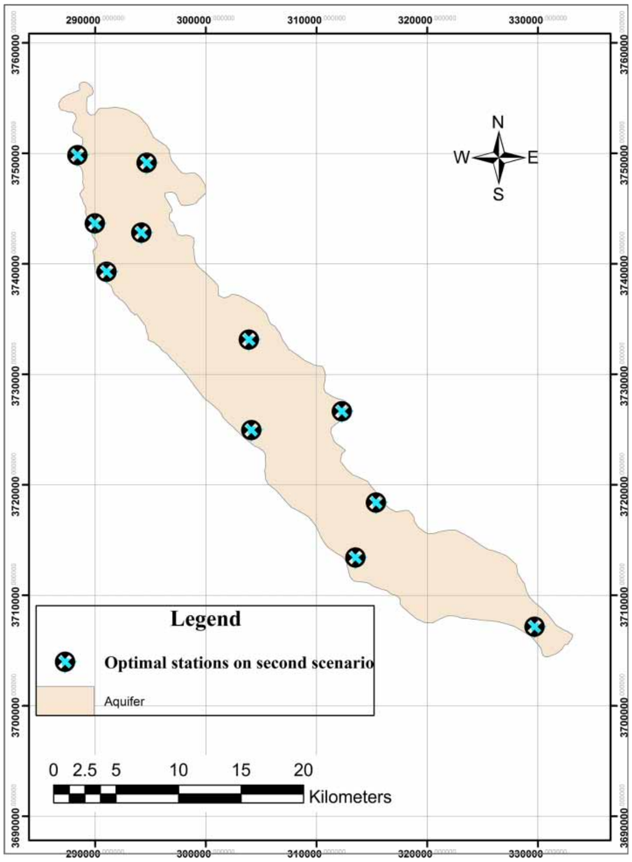


Figure 10 | Optimal monitoring stations selected in the second scenario.

According to the optimal monitoring station and data, the aquifer was zoning. Figure 11 shows the spatial interpolation map in the second scenario. As shown in Figure 11, this scenario could not predict the groundwater level as well as the first scenario on non-point in the Silakhor plain because the first scenario increased spatial monitoring accuracy by changing the location of stations. In comparison between the optimal network and existing network, it has been shown that the optimal network could reduce the cost of monitoring stations and have a similar zoning in the Silahkor plain rather than the existing monitoring network. As an example, three monitoring stations among 29 existing stations were used to estimate groundwater level by the second scenario. Table 6 shows the results of groundwater level in these points according to observation data and the second scenario. Figure 12 shows the difference between the interpolation map in the second scenario and the existing network.

The flexible nature of the first scenario as compared to the second scenario is clear in Figure 13. It is evident that optimizing the locations of the monitoring stations regardless of existing observation wells in the study area leads to the efficient collection of accurate groundwater levels.

Also, results show that taking into account the existing observation wells provides useful information for optimizing the entire monitoring network. In the second scenario, the

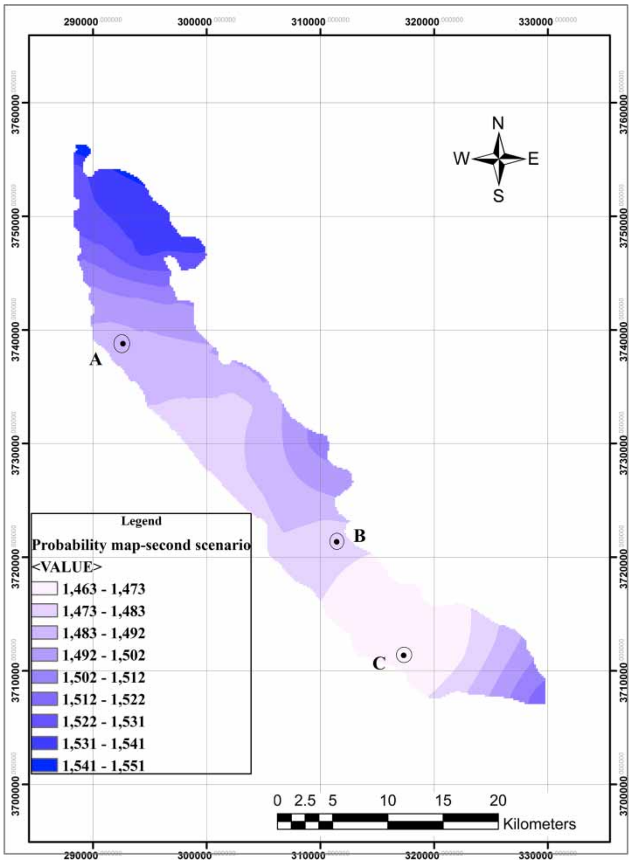


Figure 11 | Interpolation map of optimal network for augmenting the existing ones in the second scenario.

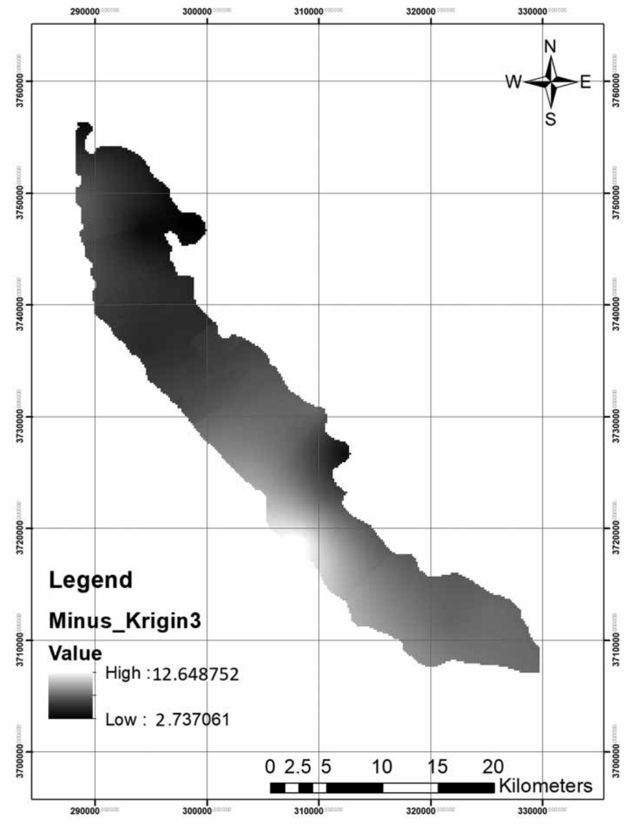


Figure 12 | Difference between interpolation map in second scenario and the existing network.

Table 6 | Groundwater level for three stations for cross validation in the second scenario

Station name	Observation level	Second scenario level	Error%
A	1,490.1	1,484.9	5.2
B	1,498.5	1,492.1	6.4
C	1,470.7	1,464.1	6.6

main constraint is the fact that existing monitoring wells are used in the optimized network. This caused a rise in the RMSE in the network associated with the second scenario as compared with the RMSE associated with the first scenario (Table 7).

CONCLUSIONS

Considering that the implementation of water resources monitoring programs is costly and time-consuming, a method for optimizing the existing network is necessary. In the process of optimizing the groundwater monitoring network, it is important to identify the number and location of wells when faced with high volumes of data in the studied time span. A small amount of groundwater monitoring plays a special role in underground water studies and the prediction of the future behavior of groundwater flows in aquifers. Therefore, the aim of this research was to design a groundwater level monitoring network by providing an optimization method. The groundwater level monitoring network is designed by taking into account all potential points within the aquifer area. This means that the design of the monitoring network is carried out from well to well, and the search algorithm space optimizes all the potential points of the aquifer range. Proper characterization of groundwater conditions relies on well-designed groundwater monitoring networks. This study introduced the multi-objective

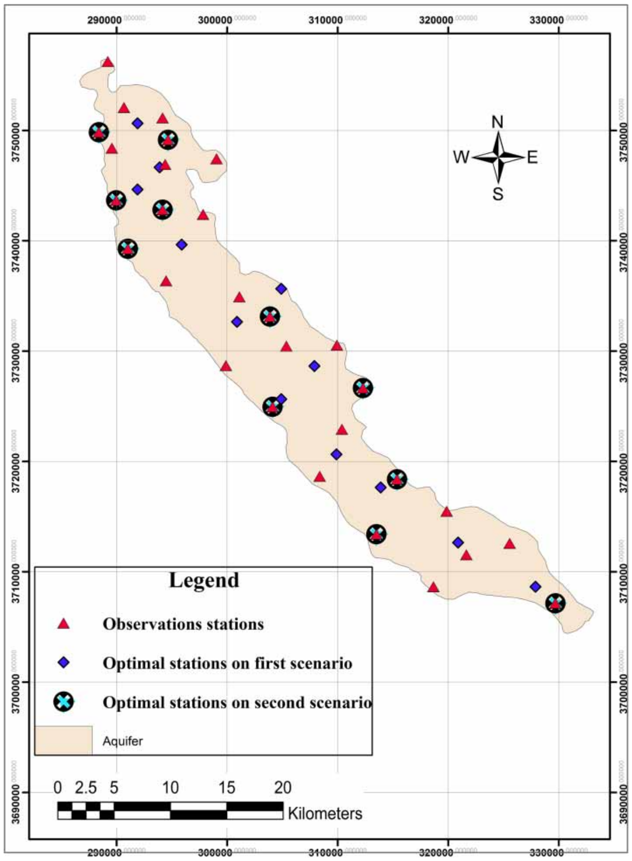


Figure 13 | Location of selected observation stations in two scenarios versus existing observation stations.

Table 7 | Comparison of existing stations and two scenarios

Scenarios	Number of stations	RMSE(m)	Decreased number of stations	Network entropy
Existing stations	29	0.54	–	0.75
First scenario	12	0.61	60%	0.84
Second scenario	11	0.78	62%	0.79

optimization of groundwater monitoring network design with the evolutionary algorithm NSGA-II and entropy theory in two scenarios. The novel methodology was applied to the Silakhor plain aquifer, Iran. The first part of this paper created a time series of groundwater level values over the entire area of the aquifer by Kriging and stored the time series into a comprehensive database. The second part of this paper optimized the network of observation wells employing the NSGA-II and optimized groundwater levels obtained by the IDW. The objectives of the optimization problem were the minimization of the number of monitoring wells and the minimization of the RMSE.

In the first scenario, groundwater monitoring wells obtained by 12 monitoring stations have RMSE $= 0.61$ m. The rehabilitation of the groundwater monitoring network in the Silakhor plain has been able to improve the spatial distribution of observation wells, which was examined in the first scenario. Also, changing the location of new stations in the first scenario has increased the accuracy of the aquifer zoning at points without statistics of groundwater level. On the other hand, the new network can provide a good method for managing groundwater resources, spatial modeling of the aquifer, and subsequently, the prospecting of groundwater flow and the risk of subsidence. This scenario has no similar results to the first scenario concerning RMSE but can reduce drilling and maintenance costs.

Comparison of the alignment lines obtained from the monitoring network in different scenarios with the initial state indicates that the optimization algorithm solutions have been effective and optimal and have been able to increase the accuracy of the network based on the objective functions.

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

REFERENCES

Abdullahi, M. G. & Garba, I. 2015 Effect of rainfall on groundwater level fluctuation in Terengganu, Malaysia. *Journal of Geophysics & Remote Sensing* **4** (2), 142–146.

Alfonso, L., Ridolfi, E., Gaytan-Aguilar, S., Napolitano, F. & Russo, F. 2014 Ensemble entropy for monitoring network design. *Entropy* **16** (3), 1365–1375.

Asefa, T., Kemblowski, M. W., Urroz, G., McKee, M. & Khalil, A. 2004 Support vectors–based groundwater head observation networks design. *Water Resources Research* **40**, 11.

Ayvaz, M. T. & Elçi, A. 2018 Identification of the optimum groundwater quality monitoring network using a genetic algorithm based optimization approach. *Journal of Hydrology* **563**, 1078–1091.

Barca, E., Passarella, G., Vurro, M. & Morea, A. 2013 Optimal redesign of groundwater monitoring networks: a data-driven multi-approach software. In *8th International Conference of the EWRA: Water Resources Management in an Interdisciplinary and Changing Context, Porto (PT)*.

Bashi-Azghadi, S. N. & Kerachian, R. 2010 Locating monitoring wells in groundwater systems using embedded optimization and simulation models. *Science of the Total Environment* **408** (10), 2189–2198.

Bhat, S., Motz, L. H., Pathak, C. & Kuebler, L. 2012 Designing a groundwater-level monitoring network using geostatistics: a case study for South and Central Florida, USA. In *World Environmental and Water Resources Congress 2012: Crossing Boundaries*. pp. 48–58.

Bhat, S., Motz, L. H., Pathak, C. & Kuebler, L. 2015 Geostatistics-based groundwater-level monitoring network design and its application to the Upper Floridan aquifer, USA. *Environmental Monitoring and Assessment* **187** (1), 4183.

Dhar, A. & Patil, R. S. 2012 Multiobjective design of groundwater monitoring network under epistemic uncertainty. *Water Resources Management* **26** (7), 1809–1825.

Esquivel, J. M., Morales, G. P. & Esteller, M. V. 2015 Groundwater monitoring network design using GIS and multicriteria analysis. *Water Resources Management* **29** (9), 3175–3194.

Gen, M., Cheng, R. & Wang, D. 1997 Genetic algorithms for solving shortest path problems. In *Evolutionary Computation, 1997, IEEE International Conference on*. pp. 401–406.

Guo, Y., Wang, J. F. & Yin, X. L. 2011 Optimizing the groundwater monitoring network using MSN theory. *Procedia-Social and Behavioral Sciences* **21**, 240–242.

Hosseini, M. & Kerachian, R. 2017 A Bayesian maximum entropy-based methodology for optimal spatiotemporal design of groundwater monitoring networks. *Environmental Monitoring and Assessment* **189** (9), 433.

Júnez-Ferreira, H. E., Herrera, G. S., González-Hita, L., Cardona, A. & Mora-Rodríguez, J. 2016 Optimal design of monitoring networks for multiple groundwater quality parameters using a Kalman filter: application to the Irapuato-Valle aquifer. *Environmental Monitoring and Assessment* **188** (1), 39.

Keum, J., Kornelsen, K. C., Leach, J. M. & Coulibaly, P. 2017 Entropy applications to water monitoring network design: a review. *Entropy* **19** (11), 613.

Khader, A. I. & McKee, M. 2014 Use of a relevance vector machine for groundwater quality monitoring network design under uncertainty. *Environmental Modelling & Software* **57**, 115–126.

Kollat, J. B. & Reed, P. M. 2007 A computational scaling analysis of multi-objective evolutionary algorithms in long-term groundwater monitoring applications. *Advances in Water Resources* **30** (3), 408–419.

Kollat, J. B., Reed, P. M. & Kasprzyk, J. R. 2008 A new epsilon-dominance hierarchical Bayesian optimization algorithm for large multiobjective monitoring network design problems. *Advances in Water Resources* **31** (5), 828–845.

Kollat, J. B., Reed, P. M. & Maxwell, R. M. 2011 Many-objective groundwater monitoring network design using bias-aware ensemble Kalman filtering, evolutionary optimization, and visual analytics. *Water Resources Research* **47** (2).

Komasi, M. & Sharghi, S. 2019 Recognizing factors affecting decline in groundwater level using wavelet-entropy measure (case study: Silakhor plain aquifer). *Journal of Hydroinformatics* **21** (3), 510–522.

Komasi, M., Sharghi, S. & Safavi, H. R. 2018 Wavelet and cuckoo search-support vector machine conjugation for drought forecasting using standardized precipitation index (case study: Urmia Lake, Iran). *Journal of Hydroinformatics* **20**, 975–988.

Krivoruchko, K. 2012 Empirical Bayesian kriging. *ArcUser Fall* 6–10.

Leach, J. M., Kornelsen, K. C., Samuel, J. & Coulibaly, P. 2015 Hydrometric network design using streamflow signatures and indicators of hydrologic alteration. *Journal of Hydrology* **529**, 1350–1359.

Leach, J. M., Coulibaly, P. & Guo, Y. 2016 Entropy based groundwater monitoring network design considering spatial distribution of annual recharge. *Advances in Water Resources* **96**, 108–119.

Li, Y. & Hilton, A. B. C. 2005 Reducing spatial sampling in long-term groundwater monitoring networks using ant colony optimization. *International Journal of Computational Intelligence Research* **1** (1), 9–28.

Loaiciga, H. A., Charbeneau, R. J., Everett, L. G., Fogg, G. E., Hobbs, B. F. & Rouhani, S. 1992 Review of ground-water quality monitoring network design. *Journal of Hydraulic Engineering* **118** (1), 11–37.

Luo, Q., Wu, J., Yang, Y., Qian, J. & Wu, J. 2016 Multi-objective optimization of long-term groundwater monitoring network design using a probabilistic Pareto genetic algorithm under uncertainty. *Journal of Hydrology* **534**, 352–363.

Mishra, A. K. & Coulibaly, P. 2009 Developments in hydrometric network design: a review. *Reviews of Geophysics* **47**, 2.

Mishra, A. K. & Coulibaly, P. 2010 Hydrometric network evaluation for Canadian watersheds. *Journal of Hydrology* **380** (3–4), 420–437.

Mogheir, Y., De Lima, J. L. M. P. & Singh, V. P. 2003 Assessment of spatial structure of groundwater quality variables based on the entropy theory. *Hydrology and Earth System Sciences* **7** (5), 707–721.

Pourshahabi, S., Talebbeydokhti, N., Rakhshandehroo, G. & Nikoo, M. R. 2018 Spatio-temporal multi-criteria optimization of reservoir water quality monitoring network using value of information and transinformation entropy. *Water Resources Management* **32**, 3489–3504.

Preziosi, E., Petrangeli, A. B. & Giuliano, G. 2013 Tailoring groundwater quality monitoring to vulnerability: a GIS procedure for network design. *Environmental Monitoring and Assessment* **185** (5), 3759–3781.

Puri, D., Borel, K., Vance, C. & Karthikeyan, R. 2017 Optimization of a water quality monitoring network using a spatially referenced water quality model and a genetic algorithm. *Water* **9** (9), 704.

Reed, P. M. & Kollat, J. B. 2013 Visual analytics clarify the scalability and effectiveness of massively parallel many-objective optimization: a groundwater monitoring design example. *Advances in Water Resources* **56**, 1–13.

Samuel, J., Coulibaly, P. & Kollat, J. 2013 CRDEMO: Combined regionalization and dual entropy-multiobjective optimization for hydrometric network design. *Water Resources Research* **49** (12), 8070–8089.

Shannon, C. E. 1948 A mathematical theory of communication. *Bell System Technical Journal* **27** (3), 379–423.

Standard practice for Design and Installation of Ground water monitoring wells in Aquifer, Designation – (1996): D 5092-90, Annual Book of ASTM standard.

Venetsanou, P., Kazakis, N., Anagnostopoulou, C. & Voudouris, K. S. 2016 Impacts of rainfall changes on groundwater balance of coastal aquifers: a case study of the Thermaikos Gulf, North Greece. *Global NEST Journal* **18**, 185–195.

Zhou, Y., Dong, D., Liu, J. & Li, W. 2013 Upgrading a regional groundwater level monitoring network for Beijing Plain, China. *Geoscience Frontiers* **4** (1), 127–138.

First received 10 April 2020; accepted in revised form 24 August 2020. Available online 27 October 2020