

METHODOLOGY FOR THE OPTIMIZATION OF GROUNDWATER QUALITY MONITORING NETWORKS ORIENTED TO SATISFY A SPECIFIC SPATIAL COVERAGE

ACEVES-DE-ALBA, J.^{1,2} – JÚNEZ-FERREIRA, H. E.^{1*} – GONZÁLEZ-TRINIDAD, J.¹ – CARDONA-BENAVIDES, A.² – BAUTISTA-CAPETILLO, C. F.¹

¹*Doctorado en Ciencias de la Ingeniería, Universidad Autónoma de Zacatecas, Av. Ramón López Velarde No. 801, Col. Centro, Zacatecas, Zacatecas, México, C.P. 98000
(phone: +52-492-924-2432; fax: +52-492-925-6690 ext. 1613)*

²*Facultad de Ingeniería, Universidad Autónoma de San Luis Potosí, Av. Dr. Manuel Nava No. 8, Zona Universitaria Poniente, San Luis Potosí, San Luis Potosí, México, C.P. 78290
(phone/fax: +52-444-826-2300)*

**Corresponding author
e-mail: hejunez@uaz.edu.mx*

(Received 19th Mar 2019; accepted 3rd May 2019)

Abstract. Through the optimal design of groundwater monitoring networks, it is possible to obtain the maximum level of information for certain variables (water level, some quality indicator parameters, and/or the presence of some contaminant) at minimum cost; the methodologies for this kind of designs follow a procedure to discard wells that provide redundant information. The information collected by monitoring networks is fundamental for supporting decision making that aims for a sustainable management of groundwater. In this paper, an existing methodology for the design of optimal groundwater quality monitoring networks is modified to include a new algorithm that allows to obtain alternative spatial distributions of the wells that provide the highest level of information, expressed in terms of the estimated uncertainty over an estimation grid. Additionally, different weights can be assigned to each monitoring site and to the analyzed variable(s) within the objective function in order to accomplish specific objectives without a significant loss of information. The final result represents a significant improvement to the original methodology by adjusting the spatial coverage of the optimal monitoring network to satisfy specific objectives.

Keywords: *geostatistics, Kalman filter, estimation grid, water management, sustainable*

Introduction

Groundwater is a vital resource for human beings; it provides 50% of the total water used worldwide (WWAP, 2016). This data shows the dependence of people on groundwater, mainly in regions with extensive cropping areas for food production and industrial activities (Li et al., 2015).

Considering the importance of groundwater, it is necessary to know, in detail, the spatial distribution of the groundwater flow systems and its physical-chemical characteristics to improve its management (Esquivel et al., 2015; Kim et al., 2007). It is relevant to assure sustainable levels of groundwater quality when it is used for water supply (Yang et al., 2018). The implementation of groundwater quality monitoring programs is of paramount importance; for instance, it was detected that in almost 15% of the monitoring stations in Europe, the standard for nitrates in drinking water established by the World Health Organization (WHO) was exceeded (WWAP, 2018). It is possible to characterize the groundwater quality in some areas of interest (AOI) by measuring some physical-chemical parameters of water extracted by wells and

estimating the spatial distribution of their concentrations; however, the costs and time involved in visiting and sampling all the existing wells are often not economically viable; specially in low income countries.

The optimal design of groundwater monitoring networks is an alternative to reduce time and costs associated to the acquisition of data to accomplish specific objectives (Júnez-Ferreira et al., 2016a); in addition the development of a sustainable strategy for groundwater extraction management is required.

The methodologies for the optimal design of groundwater monitoring networks can be subdivided in three main groups: 1) those based on elements of hydrogeological conceptual models, 2) those that execute data analysis with statistical techniques, and 3) those that use groundwater flow numerical models (Júnez-Ferreira et al., 2016a; Loaiciga et al., 1992; Herrera and Pinder, 2005; Zhang et al., 2005). The latter can be considered as the most robust framework, since it does not only consider hydrogeological aspects (as group one) but also takes into consideration groundwater levels evolution (in space and time). It is important to note that some methodologies combine aspects of two or the three groups.

The hydrogeological framework usually uses a Geographic Information System (GIS) in order to incorporate as much areal information as possible of the hydrogeological factors that control the groundwater functioning such as geology, recharge and discharge zones, rainfall data, groundwater levels, land use/land cover images derived from satellites, among others (Singh and Katpatal, 2017; Zhou et al., 2013). The statistical-based methodologies usually employ a variance-based objective function evaluated by an optimization method, they are commonly based on spatial and/or spatiotemporal geostatistical techniques, entropy theory, harmonic series models and principal component analysis (Mirzaie-Nodoushan et al., 2017; Júnez-Ferreira et al., 2016b; Andricevic, 1990; Kim et al., 2007; Júnez-Ferreira and Herrera, 2013), some of them are multicriteria (Hosseini and Kerachian, 2017a, b). The combination of hydrogeological criteria through the use of GIS and statistical techniques has been used in (Baalousha, 2010; Uddameri and Andruss, 2014; Sizirici and Tansel, 2015; Júnez-Ferreira et al., 2016a; Alizadeh and Mahjouri, 2017). The methodologies based on a numerical flow model include optimization methods and are always combined with a GIS (Kim, 2015) or with statistical estimation methods such as the Kalman Filter (KF), Monte Carlo simulations (Herrera and Pinder, 2005; Zhang et al., 2005; Kollat et al., 2011; Luo et al., 2016; Jiang et al., 2018), or can have characteristics of the three groups of methodologies for the optimal design of groundwater monitoring networks (Nobre et al., 2007).

Most of these methodologies require a large amount of information about the study site that includes a comprehensive database of sampling campaigns, geological data, and an advanced understanding of flow systems.

In those areas with scarce data, it is necessary to implement methodologies to start with the data generation. In the literature, the KF or different types of kriging have been widely used as estimation methods for the design of optimal monitoring networks oriented to the reduction of the estimate error variance in areas with few hydrogeological data (Júnez-Ferreira et al., 2016a; Júnez-Ferreira and Herrera, 2013); however, the spatial distribution of the wells with higher priority selected by these methodologies preferably covers the central area of a pre-defined estimation grid, frequently assigning a lower priority to those wells located at the periphery of the study area. This is considered as a deficiency because a poor spatial coverage of the

monitoring network may lead to the undesirable extrapolation of the monitored variable(s) at those areas. In addition, a resilient monitoring network in regard to sampling station distribution, in practice could be easier to implement and to explain to stakeholders. The design of an optimal monitoring network results in an academic exercise not feasible to real-world limitations (data availability, budget), hence the implementation of alternative methodologies is advisable (Bode et al., 2018).

In Mexico there are 653 administrative aquifers, 105 of them have been classified with intensive abstraction resulting in negative environmental effects such as the ongoing deterioration of water quality (Esteller et al., 2011), sea water intrusion problems, presence of brackish water (CONAGUA, 2017), and high fluoride and arsenic concentrations in groundwater (Banning et al., 2012; Cardona et al., 2018). However, the natural evolution of the groundwater quality has not received the required attention as a tool for management planning.

Mexico and a number of countries around the world are involved in the Sustainable Development Goals (SDG) to address global challenges such as environmental (water) degradation. Specifically, goal 6: Clean Water and Sanitation, indicator 6.3.2: Proportion of water bodies with good water quality requires the implementation of groundwater monitoring networks to compile data for some core parameters at existing sampling stations (UN WATER).

The lack of properly designed wells for groundwater quality monitoring purposes makes necessary employing the available long screened production wells. The spatial distribution of these production wells depends on factors such as the location of irrigation lands and populated areas. Therefore, it is relevant for water management institutions to consider methodologies that contribute to the design of optimal groundwater monitoring networks, with water quality data obtained from production wells. This is an important issue for making decisions regarding groundwater sustainable development, not only in Mexico, but in several other countries around the world.

This paper proposes a methodology for the design of multivariable groundwater monitoring networks. The approach includes 1) irregular shape estimation grid according with the spatial sampling station distribution 2) a ponderation grid map based on sampling station density per unit area 3) the definition of an objective function that considers simultaneously, the ponderation grid map, priority determination for any water quality parameters included in the monitoring network and/or for sampling stations and a KF for the reduction of the estimate uncertainty for each variable provided by the selected sampling stations. The methodology allows obtaining alternative spatial distributions of the wells that provide the highest level of information, expressed in terms of the estimated uncertainty over an estimation grid.

Materials and methods

The methodology presented in this paper is based on that proposed in Júnez-Ferreira et al. (2016a); the original methodology selects the optimal monitoring positions by using spatial correlations expressed as covariance matrices derived from semivariogram models obtained through geostatistical analyses to data of groundwater quality parameters. The optimization of the monitoring network is carried out by successive inclusions employing the KF as the estimation method, minimizing a function called the joint total (JT) normalized variance of the estimation error (Júnez-Ferreira et al., 2016a).

This methodology requires of a spatial estimation grid where the objective function is evaluated. During the optimization process, the addition of weights to the groundwater quality parameters, or different densities in the estimation grid have been useful to satisfy specific objectives, like assigning a higher priority to contaminated sites or to highly dangerous parameters (Júnez-Ferreira et al., 2016a; Esquivel et al., 2015).

For the proposed methodology, the objective function incorporates the spatial coverage of the available sampling stations (by means of a supplementary function, SF) besides a modified version of the joint total (JT) normalized variance of the estimation error. The modified JT, besides having a weight for each groundwater quality parameter includes a weight for each sampling station, this new algorithm is now called the weighted joint total normal variance of the estimation error (WJT). Both, the WJT and the SF are weighted in order to define different alternatives of the spatial coverage for the selected monitoring networks without a significant loss of information.

The objective function (*OF*) (Eq. 1) for the proposed methodology is:

$$OF = V \times WJT + W \times SF \quad (\text{Eq.1})$$

where *V* and *W* are the weighing factors for the WJT and the SF, respectively; $V + W = 1$. Based in Júnez-Ferreira et al. (2016a) to obtain WJT, the KF is implemented through the use of \hat{q}^0 and P^0 that represent an initial state vector and the initial error covariance matrix, respectively. The KF process employs the following equations (Eqs. 2, 3 and 4):

$$\hat{q}^{n+1} = \hat{q}^n + K_{n+1}(z_{n+1} - H_{n+1}\hat{q}^n) \quad (\text{Eq.2})$$

$$P^{n+1} = P^n - K_{n+1}H_{n+1}P^n \quad (\text{Eq.3})$$

$$K_{n+1} = P^n H_{n+1}^T (H_{n+1} P^n H_{n+1}^T + r_{n+1})^{-1} \quad (\text{Eq.4})$$

where K_{n+1} is the Kalman gain, H_{n+1} is the sampling matrix, z_{n+1} is the measurement vector, and r_{n+1} is the measurement error covariance. Superscripts refer to the number of measurements employed in the estimation; subscripts are used to indicate the transition when a new measurement is going to be used in the process.

The spatial covariance matrix for each parameter is calculated as follows (Eq. 5):

$$C(h) = C(0) - \gamma(h) \quad (\text{Eq.5})$$

where $C(h)$ represent the covariance matrix, $C(0)$ is the variance for each analyzed parameter, and $\gamma(h)$ is the variogram model function.

The WJT value is calculated with (Eq. 6):

$$WJT = \sum_{p=1}^{np} \sigma_{w,p}^2 \times \alpha_w \times \alpha_p \quad (\text{Eq.6})$$

where $\sigma_{w,p}^2$ represents the normalized estimate error variance (located at the principal diagonal of the corresponding updated covariance matrix) at each estimation position

for the p parameter, when the w well has been sampled, α_w represents the w well weight, α_p represents the p parameter weight, and np is the number of parameters.

The main advantages of this proposal are: a) it considers the possibility of obtaining alternative monitoring networks designs oriented to a desired spatial coverage b) it allows the assignation of different weights to a particular sampling point, which gives the opportunity to take into account specific hydrogeological conditions such as contaminated areas without affecting, within the optimization procedure, those wells located around this specific sampling point.

Custom estimation grid

As part of the proposed methodology, it has been developed an algorithm for the generation of a custom estimation grid (CEG). The KF estimates will be obtained and the OF evaluated at the CEG; it offers advantages when compared to a regular grid as implemented in commercial software such as XTools Pro 18 (XTools Pro), create fishnet (ArcGIS) and sampling design tool (Buja, 2012) because it avoids extrapolation areas since it is an adjusted envelope of the sampling points.

The definition of the CEG starts with the calculation of the distance between the extreme coordinates of the wells (maximum distance), followed by the calculation of the grid density with half the maximum distance divided by ten; this density (separation between the grid nodes) is selected to draw a rectangular grid that covers all the area where the wells with available data are located, an initial value of 0 is set to all these nodes. The CEG will be composed by values different than 0, the following procedure is carried out from the top (North) to the bottom (South) row of the original rectangular grid to the assignation of these values:

1. Values of 1 are assigned to the nodes that are next to the wells.
2. For each row it is given a value of 2 to the nodes that are between those with a value of 1.
3. A value of 3 is assigned to the nodes in empty rows to connect the nodes with values different to 0.
4. Values of 4 will be assigned to the nodes with an original value of 0 in such a manner that each row of the CEG cannot be longer one node to the West and one node to the East than the next row at the south.

The flowchart and pseudocode for this process are shown in *Figure 1*.

The colours in *Figure 2* are directly related to the value that indicates in which part of the algorithm the node for the CEG was selected.

Supplementary function

For the case study, the SF is proposed trying to establish a higher priority to those wells that have influence over the highest number of nodes of the estimation grid, and at the same time, with small influence of the other wells on their estimates (i.e., there is a poor spatial coverage of wells at that area). For that purpose, it is necessary to define a Custom Smooth Grid (CSG) based on the CEG and the maximum range used in the semivariogram models. For each node of the CEG, it is calculated the sum of wells located at a distance less than the model range, obtaining the values of the CSG (*Fig. 3*).

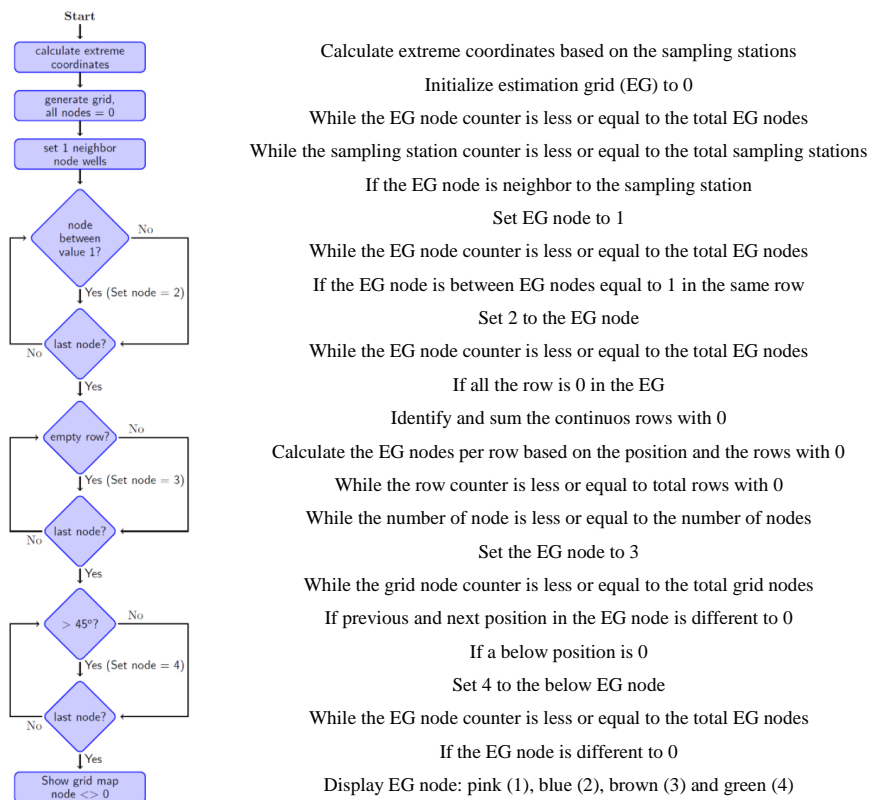


Figure 1. CEG generation flowchart and pseudocode

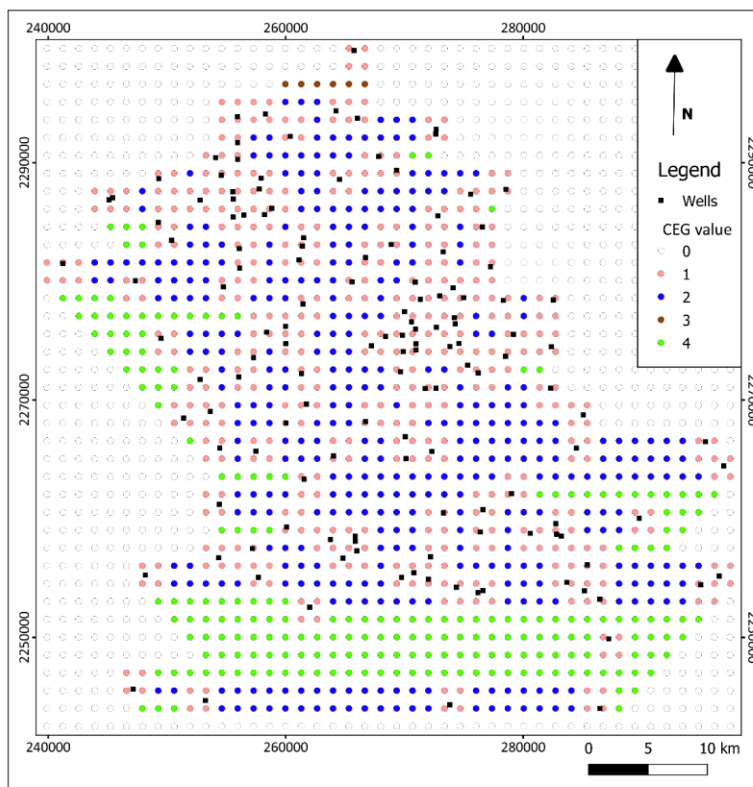


Figure 2. Custom estimation Grid (CEG)

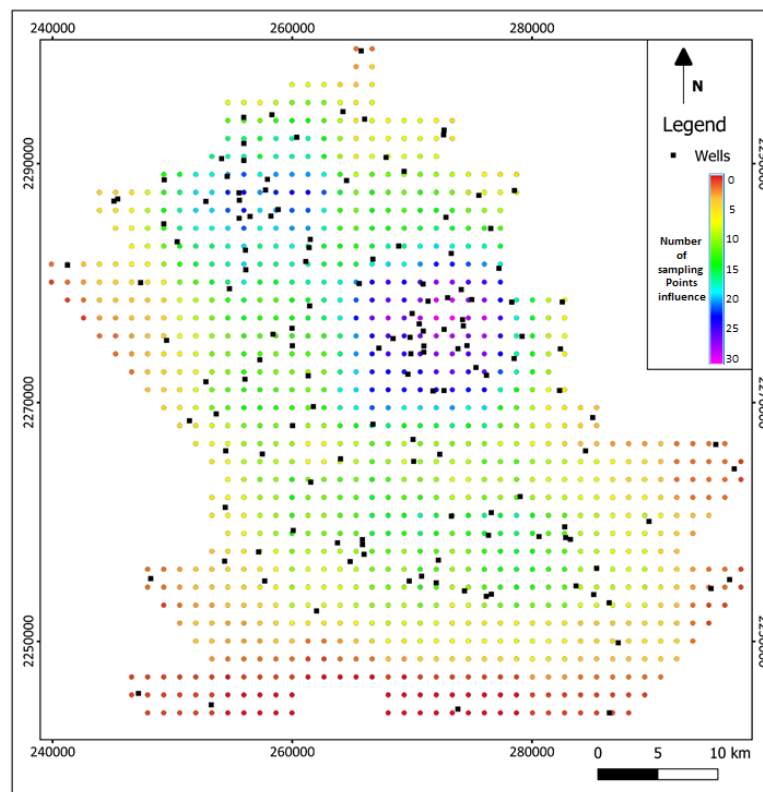


Figure 3. Custom Smooth Grid (CSG)

The SF is defined as follows:

The lower values in the CSG represent zones with a poor spatial coverage of wells.

The formulas for the values of the CSG are (Eq. 7):

$$ie_{w,p} = \sum_{n=1}^{nn} node_n \quad (\text{Eq.7})$$

where $ie_{w,p}$ represents the influence of the well w over the nodes of the CEG, corresponding to the p parameter, $node_n$ is 1 when the node is inside a circumference with center on the w well and radius equal to the range of the semivariogram model in the p parameter; $node_n$ is 0 when the node is outside the circumference, nn is the total number of nodes in the CEG (Eq. 8);

$$iw_{n,p} = \sum_{w=1}^{nw} well_w \quad (\text{Eq.8})$$

where $iw_{n,p}$ represents the influence of the sampling wells over the node n , corresponding to the p parameter, $well_w$ is 1 when the well is inside a circumference with center on the n node and radius equal to the range of the semivariogram model in

the p parameter, $well_w$ is 0 when the well is outside that circumference, nw is the total number of sampling wells (Eq. 9);

$$inw_{w,p} = \sum_{n=1}^{nw} (iw_{n,p} \times r_{n,p}) \quad (\text{Eq.9})$$

where $inw_{w,p}$ represents the influence of the sampling wells over circular zones of the CEG, $r_{n,p}$ is 1 when the node is inside a circumference with center on the w well and radius equal to the range of the semivariogram model in the p parameter, $r_{n,p}$ is 0 when the node is outside the circumference (Eq. 10);

$$f_{w,p} = \frac{\alpha_w}{ie_{w,p} \times inw_{w,p}} \quad (\text{Eq.10})$$

where $f_{w,p}$ represents the w well factor for the p parameter (Eq. 11);

$$ff_w = \sum_{p=1}^{np} f_{w,p} \alpha_p \quad (\text{Eq.11})$$

where ff_w represents the w well final factor, np represents the total number of parameters considered in the network design.

In the case study, it is proposed that $SF = ff_w$.

Results

Case study: hydrogeological setting

In order to evaluate the performance of the proposed methodology, it was applied for the design of a monitoring network in the “El Bajío” Region of Central Mexico (Fig. 4). The Bajío Region, specifically the Irapuato-Valle aquifer has become recently a focus of attention due to the intensive water extraction regime for agricultural, population, and industrial use. Groundwater extraction (about 600×10^6 m³/year) through more than 2900 production wells in the area supports the water supply for more than 900,000 inhabitants in urban and rural areas, the irrigation of approximately 60,000 hectares and the water consumption for a refinery, a power production plant and several industrial parks (Comisión Nacional del Agua, 2015; Esteller et al., 2011; Gómez and Sandoval, 2004). The rock units in the study area include a sequence of Mesozoic plutonic and vulcano-sedimentary rocks affected by several metamorphism stages and Tertiary intrusive rocks outcropping in the Sierra de Guanajuato located at the North of the study area that constitutes the hydrogeological basement. The Tertiary sequence include volcanic (ignimbrites, lava flow and tuffs) and continental sedimentary rocks (conglomerates and basin fill sediments). The Oligocene volcanic sequence is comprised of andesite and rhyolite; the overlying Miocene and Pliocene volcanic rocks are mainly andesite and basalt associated with the Mexican Neo-volcanic Belt (Nieto-samaniego et al., 2012). From the Oligocene, the Sierra de Guanajuato underwent a rapid uplift, the main geologic structures are NW-SE trending normal faults, displacing

more than 500 m the Oligocene volcanic sequence (Gómez et al., 1989; Nieto-samaniego, 1990). During this stage of deformation, several faults were reactivated, the N-S (Taxco-San Miguel de Allende fault) and NE-SW trending directions, together with the “El Bajío” Fault with NW-SE trend, formed half grabens, leading the formation of the “El Bajío” region (Botero-santa et al., 2015). At the same time, the erosion from the mountains led to transfer clastic material to the graben structure. The deposition of basin fill sediments in the graben structure produced a granular permeable regionally unconfined aquifer with a maximum thickness of 350 -400 m. Basin fill sediments are inter-bedded poorly consolidated fluvio-lacustrine sediments, including sands, inter-bedded sandy-clay and silts. From pumping test data, typical hydraulic conductivity values for that aquifer are 1 to 10 m/day (Júnez-Ferreira et al., 2016a). The Tertiary volcanic rocks are cut by a set of NW-SE and N-S normal faults, creating the possibility of a fractured heterogeneous aquifer (200-300 m thickness). The basaltic Pliocene rocks outcropping in the southern portion of the study area are in contact with the basin fill sediments, they confirm an unconfined aquifer with hydraulic conductivity values from 6 to 18 m/day (Júnez-Ferreira et al., 2016a). The Oligocene and Miocene volcanic rocks below the basin-fill sediments can also constitute a deep aquifer, fractured conditions along the NW-SE trending structures have been identified in deep wells (600-700 m deep) producing thermal groundwater (well-head temperature 35-45 °C).

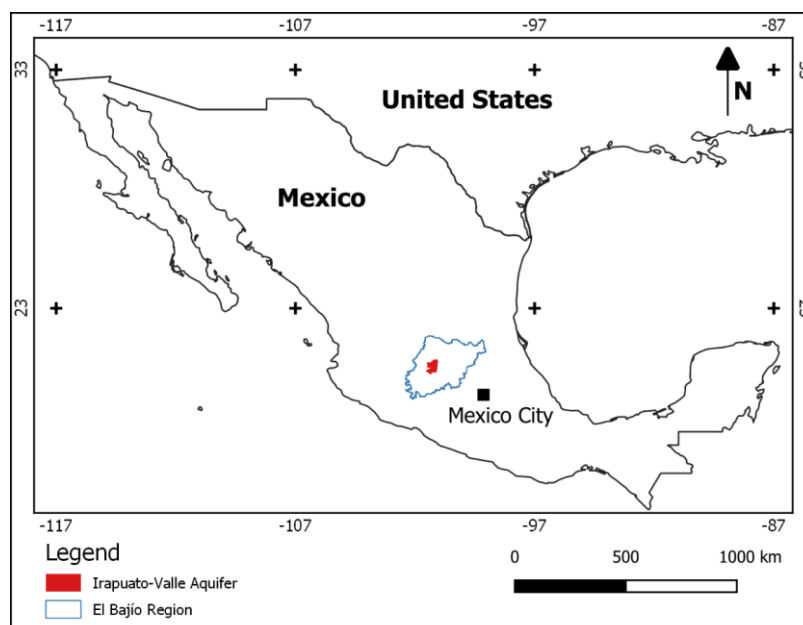


Figure 4. Location map of the study area. (Sources: Simplified El Bajío Region map modified Suárez-mota and Villase, 2015; Irapuato-Valle Aquifer from National Water Information System, 2019; countries boundaries from Tapiquén, 2019)

A sampling campaign for groundwater quality determination was carried out in December 2003 for laboratory measurement of 20 physicochemical parameters and in situ measurements of six field parameters (Júnez-Ferreira et al., 2016a), a total of 140 water samples were taken in production wells from the study area. A map displaying the sampled wells is presented in Figure 2 (UTM coordinates datum WGS84, Zone 14N). Wells ranging from 46 to 700 m deep (mean 180 m, standard deviation 115 m), most of

them screened along the saturated length of the well, promoting intra-well artificial mixing. This implies that sampled water represents a mixture of several water-bearing layers. It was identified that water in recharge areas is HCO₃-Ca dominant type (650 mg/L of total dissolved solids in average). Along the groundwater flow, water chemistry evolves to dominant HCO₃-Mix and HCO₃-Na. Groundwater quality analysis indicates the contribution of geogenic and anthropogenic contamination sources. The volcanic rocks, especially those of rhyolitic composition and the basin fill sediments derived from them, are potential trace element sources specifically for Arsenic. The water samples were analyzed for arsenic composition (mean 0.011 mg/L, max. 0.083 mg/L), 35% of the samples exhibit concentrations in excess of the WHO guideline value (0.010 mg/L) (*Fig. 5a*). Dissolved arsenic is usually found as arsenic (III) in reducing environments, which usually develops at some depth within aquifers, while near the water table oxic condition prevails (Appelo and Postma, 2005). There is a lack of vertical water quality profiles information in wells, for this reason it is not possible to demonstrate the existence of any vertical transition between oxic and anoxic conditions in the tapped aquifer. However, aerobic and oxic conditions (dissolved oxygen above 1 mg/L, Eh values above 100 mV) were detected in all the groundwater samples, suggesting that oxic conditions should prevail in most of the water bearing geologic formations tapped by the production wells. Considering that the sample depth could be assigned to mid-well screen depth, no correlation between sample depth and arsenic values was identified ($R^2 = 0.07$), indicating that dissolved arsenic is not explained by desorption from Fe-(hydr)oxide minerals in reducing conditions (*Fig. 5b*). Most of the high arsenic water samples are associated with HCO₃-Na dominant type, indicating that leaching from volcanic rocks with rhyolitic composition is a major control. Evidence from north-central Mexico suggests that dissolution of volcanic glass of felsic composition seems to be the dominant process for arsenic release into groundwater (Banning et al., 2012); the concentration in the central and northwest region of the study area seem to be derived from this source. The basin fill sediment constitutes an additional source; arsenic values above the WHO drinking water standard in the southwest region of the study area are probably derived from the basin fill sediments. Decarbonatization of basin-fill sediments could constitute an additional process for arsenic mobilization (Banning et al., 2012). The identification of groundwater impacted from anthropogenic sources, considered water types such as Cl-Mg, Mixed-Ca, Mixed-Na, and SO₄-Mixed; they exemplify impacted groundwater by diffuse contamination. Impact of diffuse contamination produced by irrigation can be evaluated with a combination of indicators, such as chloride, sulphate, nitrate, phosphate among others (Cardona et al., 2008). Chloride is a conservative solute; it is useful to detect the infiltration of diffuse contamination from agricultural activities. In addition, electrical conductivity is a function of the total concentration and charge of the ions; the relationship between them is approximately linear in most of the cases. The relation between chloride and electrical conductivity for water samples from the study area is shown in *Figure 6*. The characteristics of the local surface water and wastewater are presented as an additional reference. Despite the scatter, there is an evident general trend for most of the samples, suggesting that infiltration of wastewater modify to some extent groundwater composition. A few samples are not following the general trend, especially high electrical conductivity values with low chloride values; for these samples high alkalinity values (> 500 mg/L CaCO₃) linked with low nitrate

concentration (about 4-5 mg/L N-NO₃), suggest a limited influence of irrigation return flow or wastewater infiltration.

From the above considerations, the methodology proposed in this paper was applied for the design of an optimal groundwater quality monitoring network oriented to characterize the spatial distribution within the study area of only three parameters of the database presented in Júnez-Ferreira et al. (2016a): arsenic, chloride and electrical conductivity. Six cases were analyzed: 1) $V = 1$ and $W = 0$ within the OF (“100 – 0”), 2) $V = 99.999$ and $W = 0.001$ (“99.999 – 0.001”), 3) $V = 99.995$ and $W = 0.005$ (“99.995 – 0.005”), 4) $V = 99.99$ and $W = 0.01$ (“99.99 – 0.01”), 5) $V = 0.5$ and $W = 0.5$ (“50 – 50”) and 6) $V = 0$ and $W = 1$, but in this case the optimization procedure is applied in such a manner that the highest priority in wells is assigned to the highest estimate error variance (“Worst” case). For the six cases $\alpha_w = \alpha_p = 1$.

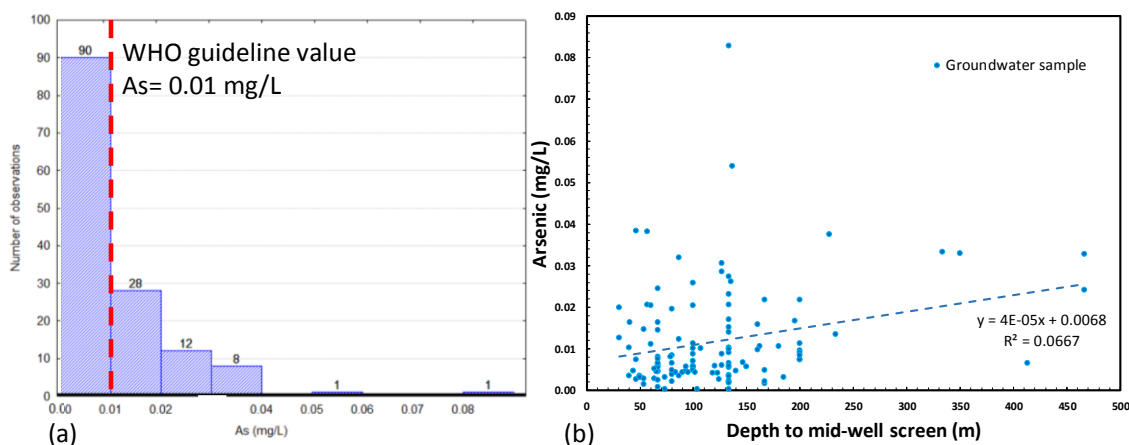


Figure 5. Arsenic analysis in a) the number of samples for groups and the reference of the WHO, in b) the correlation between concentration and well depth

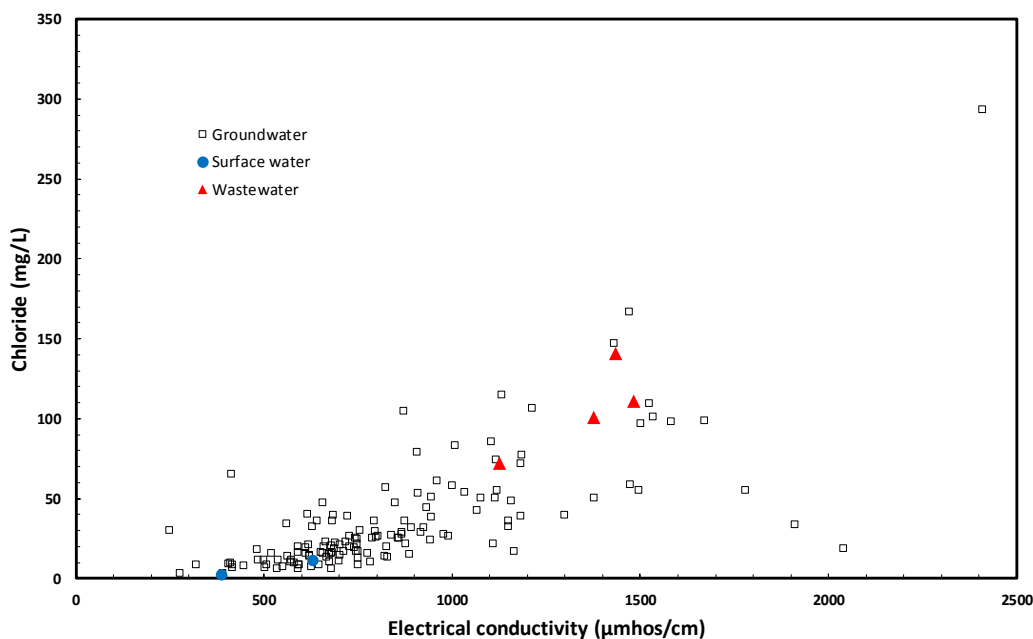


Figure 6. Correlation between chloride and electrical conductivity

Geostatistical analyses

As a result of the geostatistical analyses carried out to the selected groundwater quality parameters, semivariogram models for each parameter were obtained (*Table 1*). With these models, covariance matrices were derived and normalized for each parameter and used in the optimization process, using the CEG and the CSG values.

Table 1. Semivariogram models used for the generation of the groundwater monitoring network

Parameter	Model	Nugget (u^2)*	Sill (u^2)*	Range (m)
Electrical conductivity	Spherical	0	125097	6792
Arsenic	Spherical	6.5×10^{-5}	0.000142	4685
Chloride	Spherical	60.98	1490	6515

*u in (mg/L) for arsenic and chloride, in (μ mhos/cm) for electrical conductivity

Discussion

The methodology proposed by JÚnez-Ferreira et al. (2016a) considered the selection of sampling points by choosing the well positions that minimize the estimate error variance of the selected indicator parameters with the option to include priority zones according to the groundwater quality data. The optimization process is supported by a Kalman filter to calculate WJT. The comparison between JÚnez-Ferreira et al. (2016a) and the methodology proposed in this paper was carried out for various weighted values of the Kalman filter and the ponderation map proposed also in this new methodology. For the first case (100 – 0), the 100 indicates the percentage used for the calculations of the WJT with the Kalman filter and 0 the percentage of the SF; meaning that this is the equivalent of the JÚnez-Ferreira et al. (2016a) methodology.

Figure 7 shows the percentage of reduction of the initial WJT that is obtained on the CEG each time a new well is added to the monitoring network for the six analyzed cases. The fastest reduction in the WJT is, as expected, produced for case “100 – 0”, on the other hand, the lowest reduction corresponds to what we called the “Worst” case. The introduction of non-zero values to the SF produces a slower reduction in the WJT compared to the “100 – 0”, it can be shown that assigning the same weight to both functions of the OF (“50 – 50”) results in a reduction of the WJT that is found approximately in the middle of the “100 – 0” and the “Worst” cases. Assigning a very small weight to the SF produces almost the same estimates than in the optimal case in terms of the estimate uncertainty value (“100 – 0”); however, a different set of wells is now selected.

Figure 8a shows that for the “100 – 0” case, the first 10 positions are selected mainly at the central part of the cloud of the sampling stations, when the optimal monitoring network includes 30 positions; it extends to cover a larger area of the cloud of available sampling stations. However, it does not include positions in the periphery. An optimal monitoring network with 50 positions reaches some of the peripheral sampling positions with a high density at the central zone of the cloud.

Even though a priority order was assigned to all the monitoring wells, only those with the highest priority will be monitored in an optimal scheme since they achieve the highest reduction in the estimation grid uncertainty. In order to compare the set of wells selected for cases “100 – 0”, “99.999 – 0.001”, “99.995 – 0.005”, “99.99 – 0.01”, and

“50 – 50”, an analysis of *Figure 8* was done, where 10, 30 and 50 wells comprise the selected monitoring networks for each case. The resulting monitoring network for the “worst” case is not analyzed further; their WJT values were only used as a reference in *Figure 7*, to visualize how close the different cases are to the best and worst possible estimation scenarios.

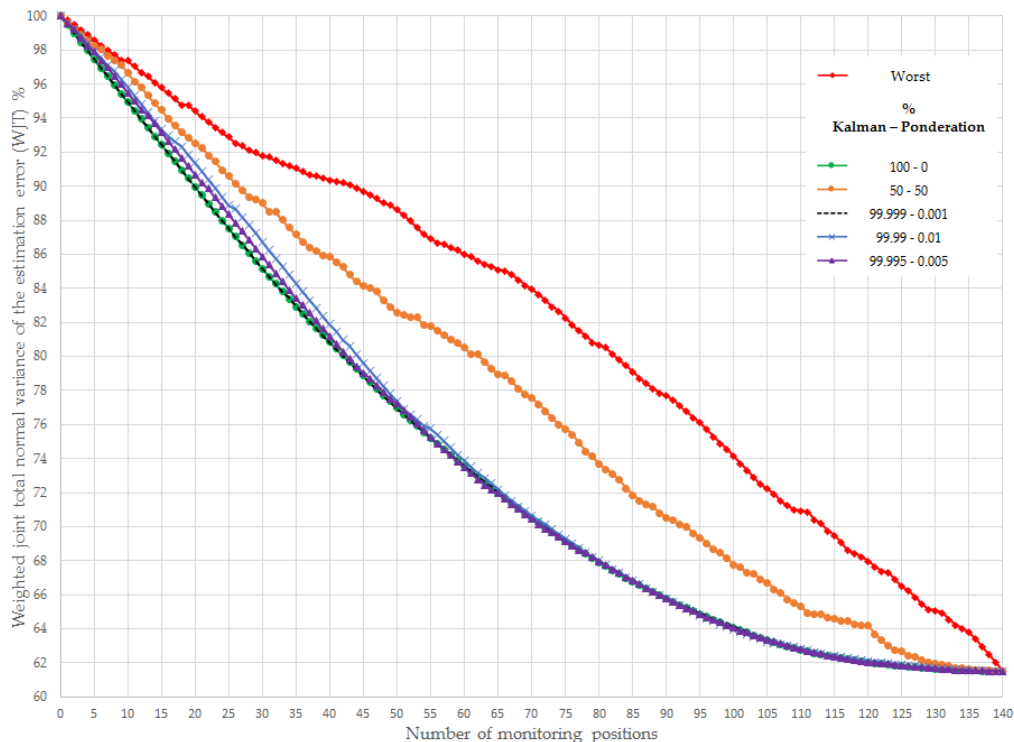


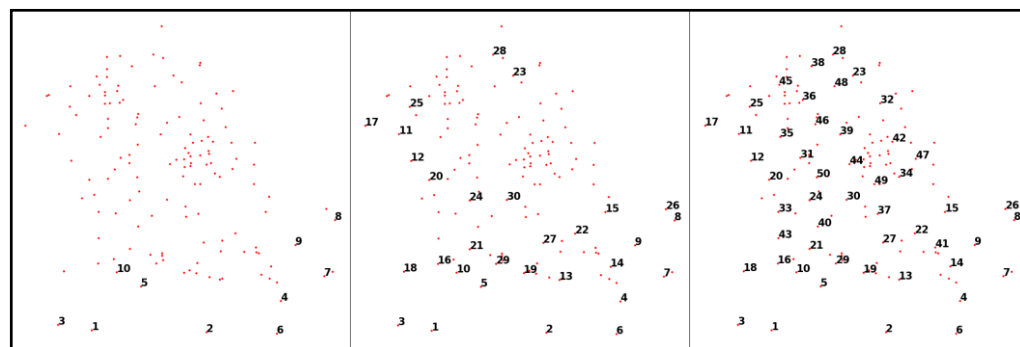
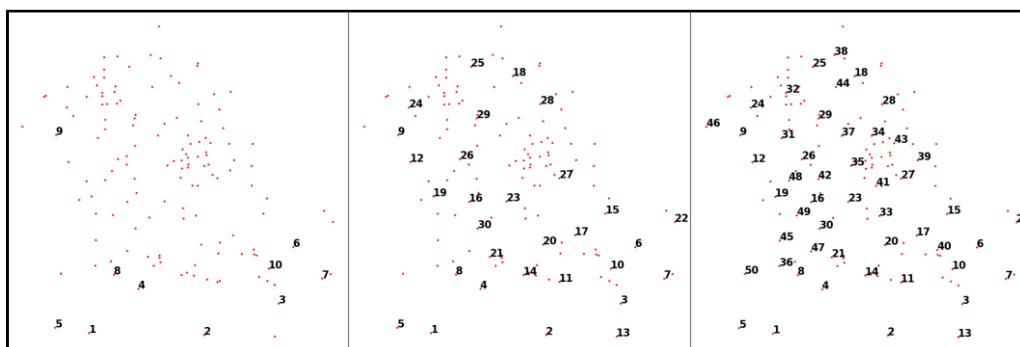
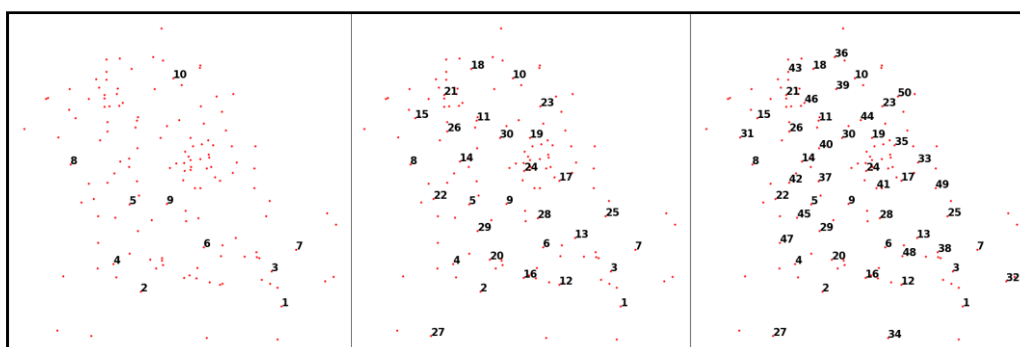
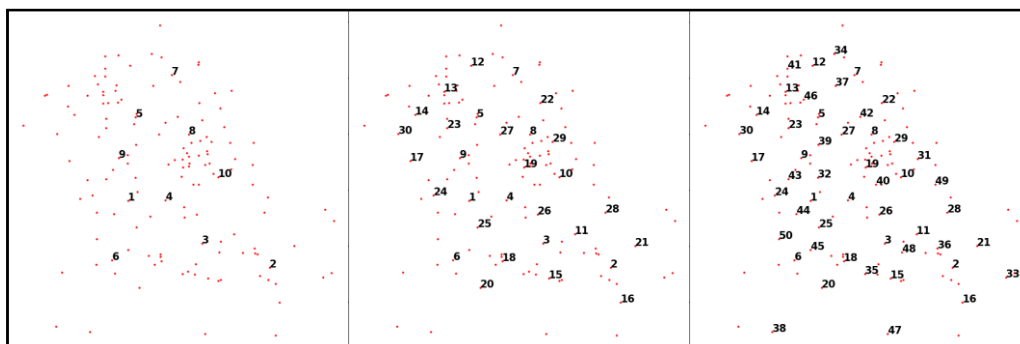
Figure 7. Comparison of the WJT (%) vs number of monitoring positions

When analyzing the “99.999 – 0.001” case (*Fig. 8b*), the first 10 positions are slightly biased to the south, including some of the peripheral wells. When 30 positions are selected, the wells are more homogeneously distributed within the cloud of points, but unlike the “100 – 0” case, the peripheral well with a priority value of 27 is included. With 50 positions, the same spatial coverage than the “100 – 0” case is virtually obtained.

For the “99.995 – 0.005” case shown in *Figure 8c*, the first 10 positions are notoriously biased to the extreme Southern part; only the position with a priority value of 9 is located at the Northwest. For 30 positions, the distribution is considerably better, with small areas uncovered at the central part. Finally, with 50 positions, the spatial coverage is better than in cases “100 – 0”, “99.999 – 0.001”, considering various wells at the Western periphery. The increase in the WJT value when selecting 50 wells for the “99.995 – 0.005” and the “100 – 0” cases is of 0.25%, this means that the loss of information is minimum for this alternative.

In *Figure 8d* it can be shown that for the “99.99 – 0.01” case, the first 10 and 30 selected positions are more biased to the south compared to the previously analyzed cases. When 50 positions are included in the monitoring network, the coverage is very similar to *Figure 8c*; however, the WJT value has increased in 0.40% with respect to the “100 – 0” case.

The “50 – 50” case (*Fig. 8e*), is clearly biased to the selection of wells along the Southern periphery when 10 positions are included in the monitoring network. The selection of 30 wells now includes the Western periphery and one well is selected at the North. When 50 positions are selected, the monitoring networks is denser at those peripheral zones, leaving the central part of the cloud uncovered.



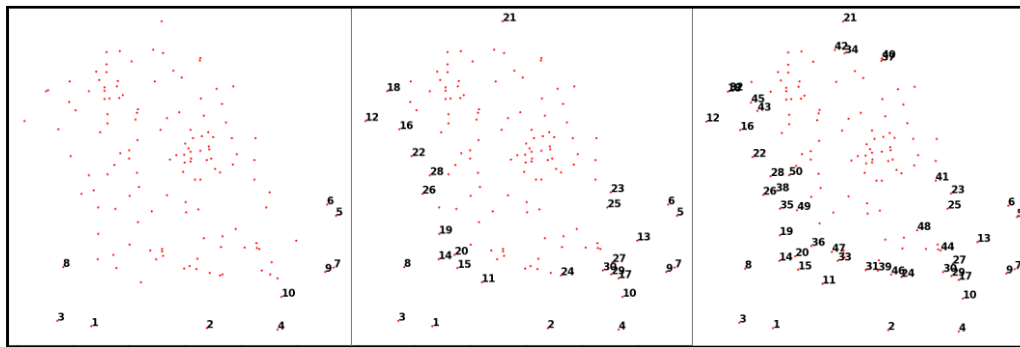


Figure 8. Groundwater monitoring networks and priority order for cases: a) “0-100”, b) “0.001-99.999”, c) “0.005-99.995”, d) “0.01-99.99” and e) “50-50”

Conclusions

Various methodologies for the design of optimal groundwater monitoring networks are based on the reduction of the estimated error variance of the variable(s) involved in the design; these estimates are obtained over an estimation grid. In most cases, the shape of the estimation grid is square or rectangular; therefore, at poorly sampled peripheral areas, the estimates are extrapolations. On the other hand, these methodologies are biased to assign a higher priority to those wells that have more influence over a larger number of nodes of the estimation grid, thus assigning a low priority to those wells closer to the periphery of the estimation grid (Júnez-Ferreira et al., 2016a).

When the design is oriented to reduce the estimated error variance only, without assigning specific weights to the nodes of the estimation grid, the monitoring network could fail in covering some pre-defined priority zones or wells with valuable information. Therefore, it is necessary to evaluate alternative designs to consider those areas or sites of interest, minimizing the loss of information when estimating with the selected optimal monitoring network.

The proposed methodology includes a new algorithm for the definition of the estimation grid that takes into account the shape of the cloud of available sampling positions, trying to avoid extrapolation. The inclusion of a supplementary function in the optimization procedure allows the monitoring design to achieve a better spatial coverage, especially at the periphery of the estimation grid, without having a significant loss of information. The objective function considers also the assignment of weights to the wells; this aspect is important when some of them have been previously identified as relevant.

Through this methodology, it is possible to analyze alternative designs in which the spatial coverage of the monitoring networks changes, but its estimates are close to the optimal in terms of the estimate error variances.

Geochemistry evolution, along with groundwater flow in thick heterogeneous aquifers can be interpreted by specific methodologies; hydrogeochemistry is useful to solve questions of origin and attenuation of geogenic and anthropogenic contaminants. The proposed methodology considers statistical techniques to analyze the spatial distribution of specific water quality parameters as a result of sampling campaigns. In this regard, the understanding of groundwater chemical evolution and/or impacts of

diffuse contamination with the monitoring network is beyond of the scope of the proposed methodology. This methodology is useful for the selection of the most convenient set of sampling stations in selected areas from a number of alternatives, it was specifically oriented for the implementation of drinking water monitoring networks. Once the monitoring network is designed, the collected data will include information from the whole selected area, data redundancy will be avoided and WTJ customized accordingly for example with financial resources and/or available time. The implementation of this methodology should facilitate to decision makers the classification of water quality in aquifers.

Acknowledgements. Jorge Aceves De Alba greatly appreciates the support of the Consejo Nacional de Ciencia y Tecnología (CONACyT) for a scholarship grant from August 2014 to February 2016. Jorge Aceves De Alba would also like to acknowledge the support of the Programa para el Desempeño Profesional Docente (PRODEP) for a scholarship grant from August 2016 to July 2017, as well as the reviewers for their comments and suggestions which helped us to improve this paper. This work was partially financially supported by the National Council of Science and Technology (CONACyT) through grant 294537 (Consolidación del Laboratorio Nacional de Espectrometría de Masas con Aceleradores) and CONACyT-BMBF (Federal Ministry of Education and Research, Germany) through grant 0207586.

REFERENCES

- [1] Alizadeh, Z., Mahjouri, N. (2017): A spatiotemporal Bayesian maximum entropy-based methodology for dealing with sparse data in revising groundwater quality monitoring networks: the Tehran region experience. – *Environ. Earth Sci.* 76(12): 1-15.
- [2] Andricevic, R. (1990): Cost-effective network design for groundwater flow monitoring. – *Stoch. Hydrol. Hydraul.* 4(1): 27-41.
- [3] Appelo, C. a. J., Postma, D. (2005): *Geochemistry, Groundwater and Pollution*. 2nd Ed. – A. A. Balkema Publishers, Leiden.
- [4] ArcGIS, “Create Fishnet” – <http://desktop.arcgis.com/en/arcmap/10.3/tools/data-management-toolbox/create-fishnet.htm> (accessed: 21-Oct-2018).
- [5] Baalousha, H. (2010): Assessment of a groundwater quality monitoring network using vulnerability mapping and geostatistics: A case study from Heretaunga Plains, New Zealand. – *Agric. Water Manag.* 97(2): 240-246.
- [6] Banning, A., Cardona, A., Rude, T. R. (2012): Applied geochemistry uranium and arsenic dynamics in volcano-sedimentary basins - an exemplary study in North-Central Mexico. – *Appl. Geochemistry* 27(11): 2160-2172.
- [7] Bode, F., Ferré, T., Zigelli, N., Emmert, M., Nowak, W. (2018): Reconnecting stochastic methods with hydrogeological applications: a utilitarian uncertainty analysis and risk assessment approach for the design of optimal monitoring networks. – *Water Resour. Res.* 54(3): 2270-2287.
- [8] Botero-Santa, P. A., Alaniz-Álvarez, S. A., Nieto-Samaniego, Á. F., López-Martínez, M., Levresse, G., Xu, S. (2015): Origin and development of the El Bajío basin in the central sector of the Transmexican Volcanic Belt [Origen y desarrollo de la cuenca El Bajío en el sector central de la Faja Volcánica Transmexicana]. – *Rev. Mex. Ciencias Geológicas* 23(1): 84-98.
- [9] Buja, K. (2012): Sampling Design Tool (ArcGIS 10.0). – <https://www.arcgis.com/home/item.html?id=ecbe1fc44f35465f9dea42ef9b63e785> (accessed: 10-Oct-2018).
- [10] Cardona, A., Carrillo-Rivera, J. J., Castro Larragoitia, G. J., Graniel-Castro, E. H. (2008): Combined use of Indicators to Evaluate Waste-Water Contamination to Local Flow Systems in Semi-Arid Regions: San Luis Potosi. – In: Rivera, J. J. C., Guerrer, M. A. O.

- (eds.) Mexico in Groundwater Flow Understanding from Local to Regional Scale. 1st Ed. CRC, London, pp. 85-104.
- [11] Cardona, A., Banning, A., Carrillo-Rivera, J. J., Aguillón-Robles, A., Rűde, T. R., Aceves de Alba, J. (2018): Natural controls validation for handling elevated fluoride concentrations in extraction activated Tóthian groundwater flow systems: San Luis Potosí, Mexico. – *Environ. Earth Sci.* 77(4): 121.
- [12] Comisión Nacional del Agua (2015): Update of the Average Annual Availability of Water in the Irapuato-Valle Aquifer (1119), State of Guanajuato [Actualización de la disponibilidad media anual de agua en el acuífero Irapuato-Valle (1119), Estado de Guanajuato]. – Comisión Nacional del Agua, Mexico.
- [13] CONAGUA (2017): Water Statistics in Mexico [Estadísticas del agua en México]. – CONAGUA, Ciudad de México.
- [14] Esteller, M. V., Rodríguez, R., Cardona, A. (2011): Evaluation of hydrochemical changes due to intensive aquifer exploitation : case studies from Mexico. – *Environ Monit Assess* 184(9): 5725-41.
- [15] Esquivel, J. M., Morales, G. P., Esteller, M. V. (2015): Groundwater monitoring network design using GIS and multicriteria analysis. – *Water Resour. Manag.* 29(9): 3175-3194.
- [16] Gómez, J. A. M., Sandoval, R. (2004): Use of groundwater in the Irapuato-Valle de Santiago aquifer region (Mexico) and its impact on the hydrogeological system [Uso del agua subterránea en la región acuífera Irapuato-Valle de Santiago (México) y su impacto sobre el sistema hidrogeológico]. – *Boletín Geológico y Min.* 115): 311-318.
- [17] Gómez, J. J. A., Gómez, J. M. A., Samaniego, Á. F. N. (1989): Considerations about the tectonic evolution during the Cenozoic of the Sierra de Guanajuato and the southern part of the Mesa Central [Consideraciones acerca de la evolución tectónica durante el Cenozoico de la Sierra de Guanajuato y la parte meridional de la Mesa Central]. – *Rev. del Inst. Geol. UNAM* 8(1): 33-46.
- [18] Herrera, G. S., Pinder, G. F. (2005): Space-time optimization of groundwater quality sampling networks. – *Water Resour. Res.* 41(12): 25-49.
- [19] Hosseini, M., Kerachian, R. (2017a): A Bayesian maximum entropy-based methodology for optimal spatiotemporal design of groundwater monitoring networks. – *Env. Monit Assess* 189: 1-24.
- [20] Hosseini, M., Kerachian, R. (2017b): A data fusion-based methodology for optimal redesign of groundwater monitoring networks. – *J. Hydrol.* 552: 267-282.
- [21] Jiang, S., Fan, J., Xia, X., Li, X., Zhang, R. (2018): An effective Kalman filter-based method for groundwater pollution source identification and plume morphology characterization. – *Water (Switzerland)* 10(8). DOI: 10.3390/w10081063.
- [22] Jűnez-Ferreira, H. E., Herrera, G. S. (2013): A geostatistical methodology for the optimal design of space-time hydraulic head monitoring networks and its application to the Valle de Querétaro aquifer. – *Environ. Monit. Assess.* 185(4): 3527-3549.
- [23] Jűnez-Ferreira, H. E., Herrera, G. S., Gonzalez-Hita, L., Cardona, A., Mora-Rodríguez, J. (2016a): Optimal design of monitoring networks for multiple groundwater quality parameters using a Kalman filter: application to the Irapuato-Valle aquifer. – *Environ. Monit. Assess.* 188(1): 39.
- [24] Jűnez-Ferreira, H., González, J., Reyes, E., Herrera, G. S. (2016b): A geostatistical methodology to evaluate the performance of groundwater quality monitoring networks using a vulnerability index. – *Math. Geosci.* 48(1): 255-44.
- [25] Kim, G.-B. (2015): Optimal distribution of groundwater monitoring wells near the river barrages of the 4MRRP using a numerical model and topographic analysis. – *Environ. Earth Sci.* 73(9): 5497-5511.
- [26] Kim, G.-B., Lee, K.-K., Lee, J.-Y., Yi, M.-J. (2007): Case study for determination of a water level monitoring frequency for nationwide groundwater monitoring networks in Korea. – *J. Hydrol.* 342(3-4): 223-237.

- [27] 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 Resour. Res.* 47(2): 1-18.
- [28] Li, Z., Deng, X., Wu, F., Hasan, S. S. (2015): Scenario analysis for water resources in response to land use change in the middle and upper reaches of the Heihe River Basin. – *Sustain.* 7(3): 3086-3108.
- [29] Loaiciga, B. H. A., Member, A., Member, A., Rouhani, S. (1992): Review of groundwater quality monitoring network design. – *J. Hydraul. Eng.* 118(1): 11-37.
- [30] 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. – *J. Hydrol.* 534): 352-363.
- [31] Mirzaie-Nodoushan, F., Bozorg-Haddad, O. Loaiciga, H. A. (2017): Optimal design of groundwater-level monitoring networks. – *J. Hydroinformatics* 19(6): 920-929.
- [32] National Water Information System, “National Water Information System [Sistema Nacional de Información del Agua]” – <http://sina.conagua.gob.mx/sina/> (accessed: 14-Apr-2019).
- [33] Nieto-Samaniego, Á. F. (1990): Cenozoic faulting and stratigraphy in the southeastern part of the Sierra de Guanajuato [Fallamiento y estratigrafía cenozoicos en la parte sudoriental de la Sierra de Guanajuato]. – *Rev. del Inst. Geol. UNAM* 9(2): 146-155.
- [34] Nieto-Samaniego, , Á. F., Ojeda-García, Á. C., Alaniz-Álvarez, S. A., Xu, S. (2012): Geology of the Salamanca region, Guanajuato, Mexico [Geología de la región de Salamanca, Guanajuato, México]. – *Boletín la Soc. Geológica Mex.* 64(3): 411-425.
- [35] Nobre, R. C. M., Rotunno, O. C. Filho, Mansur, W. J., Nobre, M. M. M., Cosenza, C. a, N. (2007): Groundwater vulnerability and risk mapping using GIS, modeling and a fuzzy logic tool. – *J. Contam. Hydrol.* 94(3-4): 277-292.
- [36] Singh, C. K., Katpatal, Y. B. (2017): Evaluating control of various hydrological factors on selection of groundwater-level monitoring networks in irrigated areas using a geospatial approach. – *J. Irrig. Drain. Eng.* 143(8): 05017003.
- [37] Sizerici, B., Tansel, B. (2015): Parametric fate and transport profiling for selective groundwater monitoring at closed landfills : A case study. – *Waste Manag.* 38): 263-270.
- [38] Suárez-Mota, M. E., Villase, J. L. (2015): The Bajío region, Mexico and the conservation of its floristic diversity. – *Rev. Mex. Biodivers.* 86: 799-808.
- [39] Tapiquén, E. P. (2019): Efraín Porto Tapiquén Geography, GIS and Environmental Studies [Geografía, SIG y Estudios Ambientales]. – <https://tapiquen-sig.jimdo.com/> (accessed: 14-Apr-2019).
- [40] Uddameri, V., Andruss, T. (2014): A GIS-based multi-criteria decision-making approach for establishing a regional-scale groundwater monitoring. – *Environ. Earth Sci.* 71(6): 2617-2628.
- [41] UN WATER, “Indicator 6.3.2 - Water quality” – <http://www.sdg6monitoring.org/indicators/target-63/indicators632/> (accessed: 25-Oct-2018).
- [42] WWAP (2016): Water and Jobs. – WWAP, Paris.
- [43] WWAP (2018): Nature-Based Solutions for Water. – WWAP, Paris.
- [44] XTools Pro, “XTools Pro 18.” – https://help.xtools.pro/pro/18.0/en/XTools_Pro_Components/Analysis_Tools/Create_Fish_net.htm (accessed: 11-Oct-2018).
- [45] Yang, Y., Xu, W., Chen, J., Chen, Q., Pan, Z. (2018): Hydrochemical characteristics and groundwater quality assessment in the diluvial fan of Gaoqiao, Emei Mountain, China. – *Sustainability* 10(12): 4507.
- [46] Zhang, Y., Pinder, G. F., Herrera, G. S. (2005): Least cost design of groundwater quality monitoring networks. – *Water Resour. Res.* 41(8): 1-12.
- [47] Zhou, Y., Dong, D., Liu, J., Li, W. (2013): Upgrading a regional groundwater level monitoring network for Beijing Plain, China. – *Geosci. Front.* 4(1): 127-138.

APPENDIX

Samples

UTM X	UTM Y	Electrical conductivity ($\mu\text{mhos/cm}$)	Arsenic (mg/l)	Chloride (mg/l)
255942	2293883	721	0.02835	38.85
258294	2294106	865	0.032	29.1
264225	2294358	563	0.0232	13.8
266037	2293737	698	0.0153	10.8
249273	2284968	248	0.0034	29.75
250397	2283463	739	0.0056	19.2
247341	2280044	581	0.0097	10
241239	2281506	674	0.0101	14.9
260375	2292217	1113	0.0107	50.22
264528	2288593	502	0.0069	6.84
269331	2289364	787	0.0066	25.43
254750	2279539	839	0.0057	26.9
256120	2281124	802	0.0059	26.7
256090	2282764	821	0.0051	13.7
258212	2285624	857	0.0053	25.43
266723	2282010	594	0.0061	8.56
268873	2283108	876	0.0068	21.77
274098	2279465	666	0.0085	13.45
270904	2279933	749	0.0101	12.96
261117	2281813	500	0.0038	11.25
265579	2279966	678	0.0047	20.3
261324	2272237	1471	0.0046	166.98
259981	2276218	748	0.0036	24.7
261502	2283669	482	0.0074	17.86
271738	2270976	1474	0.0102	58.96
272305	2265668	916	0.0146	29.1
272657	2271001	977	0.0165	27.64
264019	2265292	2410	0.0383	293.52
261750	2269664	1105	0.01285	85.48
278536	2287752	626	0.0042	7.09
275567	2287348	590	0.0068	6.06
276558	2284583	750	0.0047	8.31
272818	2285517	647	0.0171	8.56
273267	2282492	828	0.0232	13.45
261399	2282976	413	0.0041	65.43
257767	2287796	655	0.0263	15.9
256467	2285587	572	0.0136	11.49
255569	2285435	1378	0.00345	50.17
255547	2287556	1000	0.002	58.21
254566	2288936	663	0.0141	22.74
267229	2274569	1034	0.0085	53.81

270070	2266911	867	0.0206	27.39
265991	2257299	703	0.0056	14.75
264823	2256684	681	0.0082	16.38
277246	2281242	570	0.0102	10.27
278314	2278407	727	0.0205	26.58
282529	2278425	690	0.0176	22.01
258815	2286160	551	0.006	7.09
255955	2290276	485	0.008	11.69
254101	2290424	610	0.0114	19.56
252790	2286848	628	0.0042	32.28
249309	2288653	590	0.004	16.38
255582	2286945	933	0.0044	44.51
265851	2258103	673	0.0059	10.51
263772	2258257	746	0.0088	17.12
260077	2259299	703	0.0045	21.03
261545	2263341	1067	0.0297	42.395
257189	2257498	446	0.00365	8.055
285064	2268748	1299	0.0103	39.62
282307	2271012	1182	0.006	39.13
282371	2274499	686	0.01695	18.91
278506	2273692	925	0.0131	31.55
279172	2275540	943	0.0081	24.21
270591	2276591	748	0.0219	21.77
269822	2275460	1119	0.0046	55.28
269893	2274092	945	0.083	38.4
270976	2274737	823	0.054	56.75
270973	2274224	1076	0.0107	50.39
269655	2272381	1149	0.0096	32.29
262022	2252546	406	0.002	8.875
257694	2255060	782	0.004	10.27
254349	2256718	507	0.0034	8.8
254400	2261222	887	0.0035	14.84
257494	2265699	1779	0.01635	54.84
284469	2265951	1166	0.0047	16.63
283214	2258544	1117	0.0033	74.36
276376	2258892	520	0.0066	15.65
272009	2254911	1533	0.0029	100.77
276613	2253958	615	0.0057	40.11
271331	2278492	620	0.0218	14.43
272962	2278788	679	0.0206	15.41
274982	2278628	726	0.0114	19.81
268410	2275364	874	0.0057	35.96
270001	2277455	827	0.0205	19.57
272675	2276129	621	0.0168	13.94
270933	2275990	1430	0.0067	146.76
261448	2278096	775	0.0049	15.9

260003	2274766	792	0.0027	36.2
258394	2275726	680	0.0077	6.35
257289	2273581	1158	0.0147	48.76
279012	2262127	611	0.0103	15.41
280594	2258783	1523	0.0123	109.58
285398	2256122	1582	0.0044	97.84
274367	2254234	590	0.0032	19.57
273276	2260493	682	0.0071	36.2
269754	2255053	2040	0.0089	18.83
272189	2256803	945	0.0091	50.88
270786	2255466	1910	0.0027	33.37
251411	2268482	960	0.0178	61.15
254443	2265954	1495	0.0315	54.845
253651	2269050	859	0.0384	25.44
256054	2271954	1213	0.0092	106.65
266727	2268190	1670	0.0026	98.82
273817	2244339	756	0.0023	30.08
286457	2244021	991	0.0074	26.67
287199	2249886	796	0.0052	29.35
286441	2253231	892	0.0095	31.55
285147	2253929	848	0.0038	47.29
283668	2254647	559	0.0016	34
295339	2266495	1501	0.0024	96.86
296876	2264452	537	0.0007	11.5
252791	2271746	797	0.0112	25.93
282816	2258695	1009	0.0017	82.9
270132	2265073	1182	0.0112	71.9
276607	2260779	1131	0.0029	114.6
282733	2259587	907	0.0007	79.2
276186	2253776	642	0.0021	36.2
255951	2291694	716	0.0048	22.8
257925	2288696	658	0.0032	19.9
260031	2268069	1184	0.0197	77.2
265749	2299445	276	0.001	3.2
272667	2292812	632	0	11.1
272632	2292422	390	0	3.2
267827	2290531	535	0.0019	6.3
245432	2287054	593	0.0046	8.7
245118	2286870	712	0.0045	16.8
249504	2275210	751	0.005	17
265846	2258531	1109	0.0099	21.4
253236	2244695	410	0.00035	10
247158	2245650	417	0.0035	6.6
248184	2255261	619	0.0023	20.9
294946	2254438	320	0.00035	8.7
296494	2255175	417	0.0023	8.3

289764	2260054	1149	0.0023	35.9
273830	2274505	872	0.033	104.9
275319	2272939	684	0.0066	39.3
276204	2272290	909	0.0242	53.4
274578	2274772	657	0.0333	47.6
274275	2276426	743	0.0376	25
274206	2276950	650	0.0328	16.3