

# Geostatistical analysis of static level evolution between 1995-2005 and 2007-2017 in San Luis Potosí Aquifer, México

Maria Ines Navarro-Hernández<sup>1</sup>, Abraham Cardenas Tristán<sup>1</sup>, Ratih Fitria Putri<sup>2\*</sup>

<sup>1</sup>Center for Research and Graduate Studies, Faculty of Engineering, San Luis Potosí Autonomous University, México

<sup>2</sup>Department of Environmental Geography, Faculty of Geography, Universitas Gadjah Mada, Indonesia

**Abstract.** Knowing the piezometric levels evolution of an aquifer is essential for planning the use of water resources in the cities of central México, such as San Luis Potosí, whose population depends almost entirely on groundwater. Applying geostatistical methods from georeferenced data can reliably provide information on the spatial variability of valley piezometry. The objective of this study is to evaluate the evolution of the static levels of the aquifer in San Luis Potosí Valley, through an interpolation analysis by the kriging method. For both periods, the experimental variogram was adjusted to the Gaussian model and the method with the best results was ordinary kriging. From this analysis it was obtained that the decay of the static levels is more critical in the central zone of the valley, while the zones of the northwest and towards the south, the aquifer levels have shown an important recovery; the important values for the decrease in static levels coincide with the areas where the population density is greater.

## 1 Introduction

Since the 1970s, geostatistics has emerged as a science that is responsible for the study of phenomena with spatial correlation, to be able to predict the values of the variables in places where sampling was not carried out or places for which there are no data. Statistical methods use stochastic theories for interpolation and the distribution of uncertainty. An important aspect of geostatistics is that it uses spatial autocovariance structures, represented by semivariograms, which indicates the degree of similarity of values that belong to a regionalized variable, taking into account the distance that this attribute is concerning others that influence about him. To be able to predict the interpolation of the values of points that were not sampled, geostatistics relies on several methods, such as the Kriging method [1].

Throughout the geostatistics history, geostatistical methods have been widely used to analyze spatial processes and more recently space-time processes in almost all applied sciences, such as hydrogeology, environmental engineering in studies of pollutants concentrations ( air, water, soil) and meteorology for climate predictions, etc. [2], [3].

## 2 Study area

The SLP valley is located in the physiological province between the Sierra Madre Oriental and the Sierra Madre Occidental. It has an approximate altitude of 1850 masl and occupies an extension of 1980 km<sup>2</sup>; It is surrounded to the west by the Sierra de San Miguelito and to the east

by the Sierra de Álvarez [4]. This basin is located in the southwestern part of the state of San Luis Potosí (Figura 1), and within the area covered by this valley are totally or partially the municipalities of San Luis Potosí, Soledad de Graciano Sánchez, Mexquitic de Carmona and Cerro de San Pedro [5], with a current population of 1,221,556.

The average annual temperature is 17.5 ° C, with precipitation of approximately 402 mm per year and average annual evaporation of 2038.7 mm, this confirms a type of dry or semi-arid climate because evaporation far exceeds the precipitation [5].

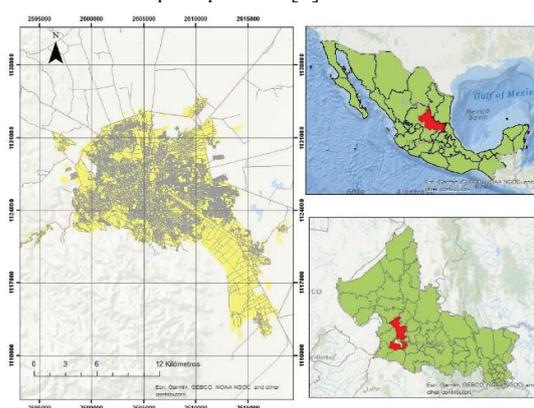


Fig. 1. Location of the study area.

Since 1995, there is a dropped cone in the city of San Luis Potosí, between 1971 and 1995 the drop in static levels was 95 m. In the period 1995-2001, this cone descended 25 m more towards the northern part mainly,

\* Corresponding author: [ratihfitria.putri@ugm.ac.id](mailto:ratihfitria.putri@ugm.ac.id)

even though there were areas in which the levels were kept constant [6].

by 1960 59% of the water for domestic use came from surface water and the rest was extracted from the aquifer, in 2005, this amount had increased to 92% and only 8% of the water destined for domestic use came from superficial water bodies [6]. The increase in the extraction of groundwater has been accelerated; and this went from being 0.97 m<sup>3</sup>/s in 1970 to 4.1 m<sup>3</sup>/s for 2007 [7].

Currently, several factors exacerbate this decrease in piezometric levels, one of them being the accelerated population growth and the expansion of agricultural and industrial activities. However, the low permeability of the lithological materials that border the valley, the low precipitation and high evaporation of the area that characterizes semi-arid climates, erosion, and deforestation, as well as the paving of recharge areas, affect the impediment of retention and infiltration of runoff water into the aquifer. This has caused the deficit in the hydrological balance between the period from 1995 to 2002 to have increased from 36 million m<sup>3</sup>/year to 42.5 million m<sup>3</sup>/year [8].

The volume of water withdrawn for 2006 was 92.2 million m<sup>3</sup>/year. This was obtained from the exploitation of 125 supply sources so that the depths of these wells vary between 270 and 700 meters. [9].

### 3 Methodology

Kriging is a geostatistical interpolation technique that returns the best linear unbiased prediction of intermediate data and avoids clustering effects. This method assumes that the data is generated from the Gaussian distribution with the correlation structure defined by the estimated semivariogram [10]. The semivariogram is a function that allows analyzing the spatial behavior of a variable within an area. The experimental semivariogram is defined by the following mathematical expression [11]:

$$(h) = \frac{1}{2N(h)} \sum_{\gamma=1}^{N(h)} [Z(X_i + h) - Z(X_i)]^2 \quad (1)$$

Where,  $Z(X_i)$  is the value of the variable at site  $x$ ;  $Z(X_{i+h})$  is another sample value separated from the previous one by a distance  $h$ , and  $N$  is the number of pairs that are separated by the distance  $h$ . The semivariance function is calculated for various distances  $h$ . Thus, the experimental semivariogram corresponds to an average distance between pairs of sites within each interval and not to a specific distance  $h$ . To interpret the semivariogram, the criterion is that the smaller the distance between the sites, the greater the spatial correlation between the sample values. [12].

Ordinary kriging does not require information beyond the measurement and its geographical coordinates, it is based on the assumption that the variation is random and spatially dependent, and that the random process is intrinsically stationary, with a constant measurement and a variance that depends only on the separation of distance and direction between places, but not in the absolute position. Kriging

prediction is a linear sum of data that can have one, two, or three dimensions [13].

Assuming that the following sampling points,  $X_1, X_2, \dots, X_N$  have been recorded for the values of a random variable  $Z$  to give  $N$  data,  $Z(X_i), i=1,2,\dots, N$ . For spot kriging,  $Z$  is predicted at any new point  $X_0$ , using the equation:

$$Z(X_0) = \sum_{i=1}^N \lambda_i Z(X_i), \quad (2)$$

Where  $\lambda_i$  are the weights, to guarantee that the estimate is impartial, the weights must add 1.

$$\sum_{i=1}^N \lambda_{y0} = 1 \quad (3)$$

The expected difference is  $E[Z(x_0) - z(x_0)] = 0$  and the variance of the prediction are calculated using the equation:

$$\begin{aligned} Var [Z(x_0)] &= E\{[Z(x_0) - z(x_0)]^2\} \\ &= 2 \sum_{i=1}^N \lambda_i \gamma(x_i - x_0) - \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j \gamma(x_i - x_j) \end{aligned} \quad (4)$$

Where  $\gamma(x_i - x_0)$  is the semivariance of  $Z$  between the sampling point  $x_i$  and the target point  $x_0$ , and the quantity  $\gamma(x_i - x_j)$  is the semivariance between the  $i$ -th and  $j$ -th sampling points. These semi-variances are derived from the variogram model, to ensure that the variances are not negative. If a target point is also a sampling point, kriging returns the observed value and the estimated variance will be zero [11].

As is known, Kriging is a geostatistical method that uses a set of geostatistical tools to predict the values in places where there was no sampling. It is important to highlight that to model with these methods, it is first necessary to verify that the data complies with the conditions, analyzing elements such as normal distribution without trend, histogram, and normal QQ graph, to determine the best way to model the spatial structure of the data, that is, the type of semivariogram (spherical, exponential, etc.) and from this, to be able to select the best kriging method [14].

Two geostatistical analyzes were performed in this study, the first is an analysis of the piezometric evolution from 1995 to 2005; the second corresponds to the analysis of the piezometric evolution for the period from 2007 to 2017. The field information that supports these statistical analyzes was provided by CONAGUA, however, the piezometric measurements for the year 2017 are found on the official website of the INEGI, in a hydrogeological study of the San Luis Potosí aquifer.

In the second instance, two databases were prepared, which include the measurements of the piezometric changes for both periods; this database contains information on the coordinates in UTM, and the static levels for each well in both years, so that in the 1995-2005 analysis 22 wells were used, while for the 2007-2017 period 24 wells were used. The third phase of this analysis involves making an exploratory analysis of the variable evolution of the piezometric level, for example, observing the distribution of the data, since, in case of

not having a normal trend, some transformations will be necessary to achieve this behavior.

The fourth phase is the application of geostatistical methods, for this, it is necessary to make a structural analysis, that is, study the semivariogram of the primary variable, choose the semivariogram that best fits, and subsequently, the type of Kriging that best fits to the behavior of the data. In the final stage, once the best method to simulate the data has been determined, the spatial prediction map is constructed through the ArcGIS geostatistics tool, and finally, interpretations about the phenomenon are made.

## 4 Results and Discussion

### 4.1 Analysis of the evolution of the piezometric level for the period 1995-2005

In this first analysis that includes ten years, 22 wells distributed in the San Luis Potosí Valley were used. Thus, it was pertinent to carry out an exploratory study of the data, in this case, of the variable that contains the changes in the piezometric level between 1995 and 2005. Therefore, a histogram was generated, as shown in Fig. 2 where, in the upper right corner appear the values for the measures of central tendency, asymmetry, and kurtosis measures.

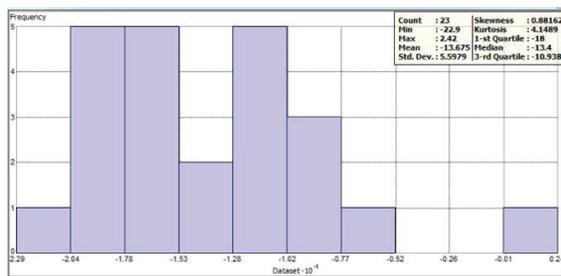


Fig. 2. Histogram of change in static level 1995-2005.

As observed in the histogram, if a curve could be constructed, this would be of the leptokurtic type, that is, a little pointed, which translates into more homogeneous data, this can be validated with the value of kurtosis that is 4.14, what which means that it is very close to 3, which is the homogeneity value. On the other hand, the skewness is 0.881, and when this value is close to zero, it can be said that it is a low skewness; however, being positive indicates that the skew is to the right of the curve or histogram.

Next, a trend analysis is performed, which consists of a three-dimensional graph of the data of the variable studied. In Fig. 3, it is observed how these values are distributed, which are also projected on each of the axes, with this, the projections of the trends (blue line and green line) can be observed.

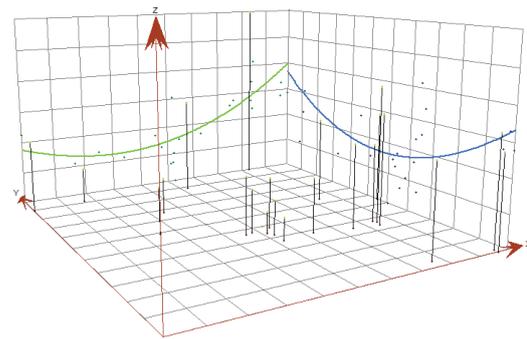
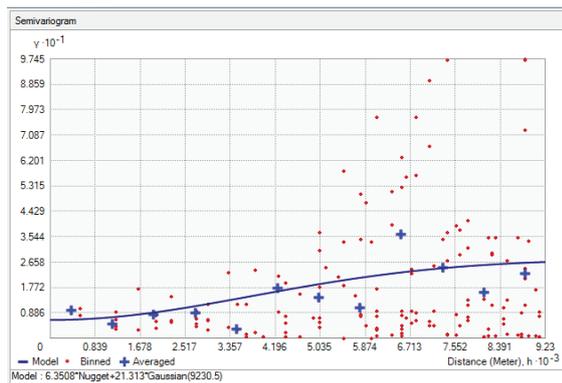


Fig. 3. Trend analysis for static levels 1995-2005.

For the projection of the XZ plane (green line) it is noted that the trend increases from west to east, that is, from this information we can begin to deduce that the piezometric decays are towards the east and central part of the area of study. On the other hand, concerning the projection of the YZ plane (blue line), a significant drift and greater dispersion are observed, indicating that the main decays are in the central part, while to the north and south the level remains stable or increases.

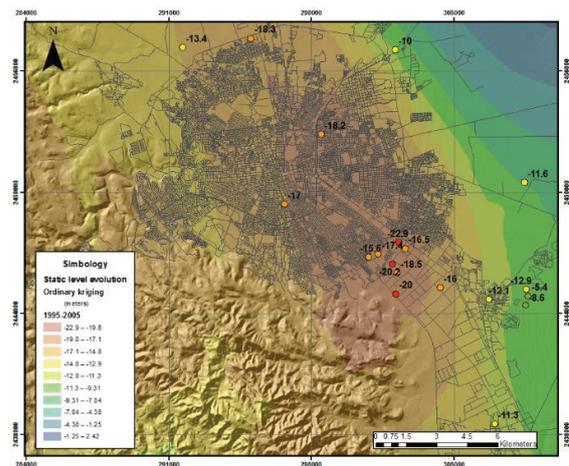
Then, simple kriging is carried out, for which no transformation was done to normalize the data, nor was it necessary to eliminate any trend. Simple kriging is a method in which the value of the mean must be known, in this case, that value is -13.7; It also requires knowing the covariance, however, the disadvantage of this type of kriging is that it produces very generalized results. For this simple kriging analysis, the semivariogram model that best fit the data was the Gaussian (this model is commonly used in data with normal distribution), as shown in Fig. 4, anisotropy was also incorporated. After performing simple kriging, ordinary kriging was constructed, for which the same Gaussian-type semivariogram model was maintained; ordinary kriging, unlike simple kriging, does not require knowing the mean, but it must be constant; Another difference is that it uses more complex equations for its construction and generates more detailed information, despite exaggerating the outliers and consequently presenting “portholes” effects in the interpolation results. Finally, universal kriging was also performed taking into account the first and second-order drift, as observed in the trend diagram, this type of kriging is similar to ordinary kriging, with the difference that it takes into account the trend of the data, so this is subtracted from the results. The disadvantage of this type of kriging is that it is stricter than the previous ones, therefore, it does not interpolate areas where there is no data, and this poses a problem for the study area since there is very little information to the west of the VSLP, therefore this surface lacks interpolation.



**Fig. 4.** Semivariogram, static level data 1995-2005.

When completing the construction of the three types of kriging, it is necessary to make a comparison between the values of the mean, the root means, the standardized mean, the root mean square error, and the standard mean error, to discriminate which method is which provides better results, or in other words, which method offers better interpolation. When comparing simple kriging with ordinary kriging, it was found that ordinary kriging shows better results in its predictions, even though the differences are minimal, because when comparing the mean, the one that is closest to zero is that of ordinary kriging, that is, it is less scattered, they both have a positive quadratic error, which means that they are undervaluing the prediction data, however, ordinary kriging is closer to unity. Fig. 5 shows the prediction map of the piezometric evolution level in San Luis Potosí valley, with ordinary kriging.

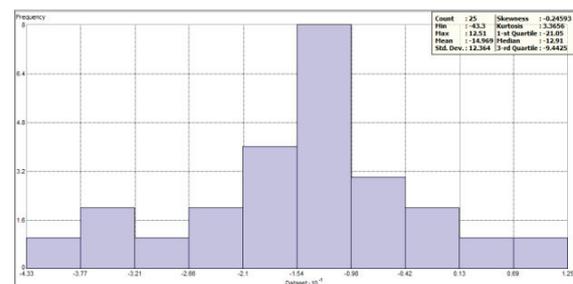
As can be seen in the prediction map built from ordinary kriging, there is a concentration of wells points with piezometric measurements for both years (1995 and 2005) in the central area, as a consequence, the prediction error spreads to the west, where there is less information available, therefore, the kriging prediction is sharper towards that extreme. On the other hand, from these results, we can confirm what had been inferred from the trend analysis in the data exploration stage, and that is that the area where there is more piezometric decay is in the central area and the recovery of the level Static occurs more to the northeast of the study area. This can be attributed to the fact that during this decade many agricultural wells that were located towards the northeast were closed due to the decrease in the need for irrigation water as a consequence of the growth of the municipal seat. Thus, by decreasing agricultural activity in that area, it could be a favorable scenario in this area for recovery of the aquifer's static levels.



**Fig. 5.** Prediction map of the change in static level 1995-2005.

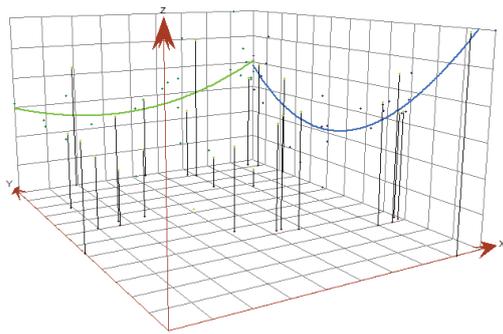
#### 4.2 Analysis of the piezometric level evolution for the period 2007-2017.

For the second analysis for the 2007-2017 period, 24 wells distributed in the San Luis Potosí Valley were used. Fig. 6 in the upper right corner shows the values for the measures of central tendency, asymmetry, and kurtosis measurements. When observing the histogram, it is noted that if a curve could be constructed, it would be of the leptokurtic type, that is, a little pointed, which translates into more homogeneous data, so this can be validated with the value of kurtosis that is 3.36, which means that it is very close to 3, which is the homogeneity value. On the other hand, the skewness is -0.246, and this value, being close to zero, can be said to be low skewness; however, and being negative, indicates that the skew is to the left of the curve or histogram.



**Fig. 6.** Histogram of change in static level 2007-2017.

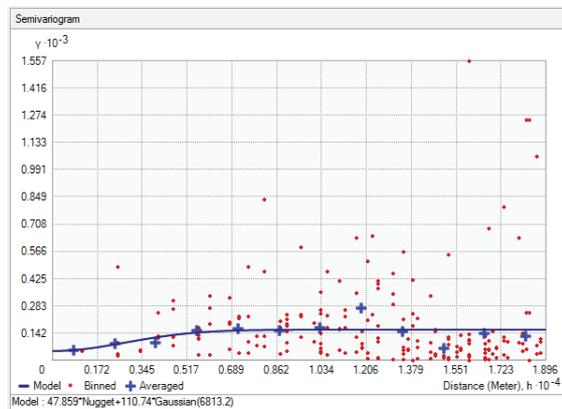
When performing a trend analysis, which consists of a three-dimensional graph of the data from studied variable (Fig. 7), it can be seen how these values are distributed, which are also projected on each of the axes.



**Fig. 7.** Trend analysis for static levels 2007-2017.

For the projection of the XZ plane (green line), we note that the trend increases from west to east, that is, that from this information we can begin to deduce that the piezometric decays are towards the east and central part of the study area. On the other hand, for the projection of the YZ plane (blue line), a significant drift and greater dispersion are observed, which indicates that the main decays are in the central part, while towards the south the static levels are considerably recovered.

Next, we proceed to build simple kriging, without the need to do any kind of transformation to normalize the data, nor was it necessary to eliminate any trend. For this simple kriging analysis, the semivariogram model that best fit the data was the Gaussian, as shown in Fig. 8. Anisotropy was also taken into account.

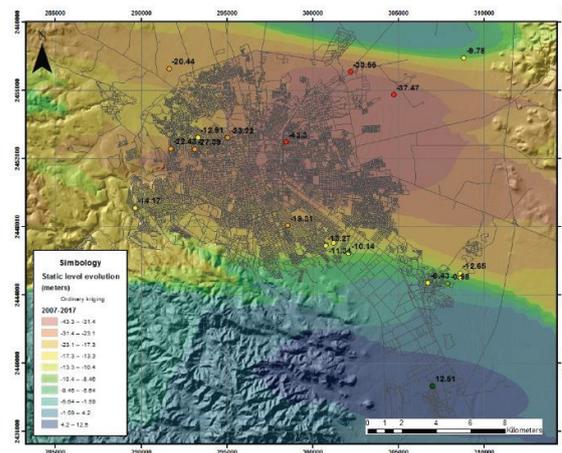


**Fig. 8.** Semivariogram, static level data 2007-2017.

After simple kriging, ordinary kriging was performed, this was adjusted to the same Gaussian-type semivariogram model; ordinary kriging, unlike simple kriging, does not require knowing the mean, but it must be constant. Finally, universal kriging was also carried out, taking into account the first and second-order drift as observed in the trend diagram. When completing the construction of the three types of kriging, a comparison was made between the mean values, the root means, the standardized mean, the mean root error, and the standard mean error, to be able to discriminate the method with the one with the best results, or in other words, which method offers the best interpolation. When comparing simple kriging with ordinary kriging, it was obtained that the ordinary one shows better results in its

predictions because when comparing the mean, the mean that is closest to zero is that of ordinary kriging, that is, it is less dispersed, for, on the other hand, they both have a positive squared error, which means that they are undervaluing the prediction data, however, the ordinary is closer to unity.

Fig. 9 shows the prediction map of the evolution of the piezometric level in San Luis Potosí valley, with ordinary kriging. From this prediction map made with ordinary kriging, we note that there is a concentration of wells points with piezometric measurements for the years 2007 and 2017 in the central-eastern area, as a consequence, the prediction error propagates to the West and Southwest, where less information is available, therefore, the kriging prediction is sharper towards that extreme. On the other hand, from these results, we can confirm what had been inferred from the trend analysis in the data exploration stage, and that is that the area where there is more piezometric decay is in the central-eastern area (mainly the urban spot) and the recovery of the static level in these years occurs more towards the northeast and southeast of the study area.



**Fig. 9.** Prediction map of the change in the static level 2007-2017.

## 5 Conclusion

When comparing these results of the changes in the piezometric levels for the period of 1995-2005 and 2007-2017, we observe that the zone of the decay of the static levels is maintained towards the center of the metropolitan zone of the San Luis Potosí Valley. And this phenomenon also affects the municipality of Soledad de Graciano Sánchez; With the difference that this area of influence is currently migrating to the east, the reason for this may be the accelerated increase of the population during the last decade towards the newly affected areas.

Also, it is important to note that the area with the greatest decays in the piezometric levels coincides with the area where there is the greatest thickness of deformable materials (quaternary deposits of sand, silt, and clay, poorly or not consolidated) and the greatest population accumulation. Thus, monitoring changes in the static levels of an aquifer is essential for planning

and managing water resources. Furthermore, such pronounced changes in the static levels of the aquifer combined with geological factors can generate physical phenomena of deformation such as the phenomenon of subsidence, which has affected the civilian population in the San Luis potosí valley and this has caused significant economic losses related to damage to buildings.

This research was supported by the National Council of Science and Technology (CONACYT) and the Center for Research and Postgraduate Studies of the Autonomous University of San Luis Potosí. Piezometric level data was provided by the National Water Commission (CONAGUA) and the National Institute of Statistics and Geography (INEGI).

## References

1. Burrough P. A., Environ. Ecol. Stat. **8**, 361 (2001)
2. S. De Iaco, M. Palma, and D. Posa, Spat. Stat. **14**, 546 (2015)
3. E. A. Varouchakis, *Background of Spatiotemporal Geostatistical Analysis: Application to Aquifer Level Mapping* (Elsevier Inc., 2018)
4. E. L. Flores-márquez, I. K. Ledesma, and C. Arango-galván, 425 (2011)
5. J. Martínez-Hernández, Identificación De La Recarga Natural E Inducida En La Cuenca De San Luis Potosí, Universidad Autónoma de San Luis Potosí, 2008
6. V. Trueba, A. Jiménez, and Q. Nguyen, *Estudio Para Determinar El Modelo de Explotación de Agua Disponible En El Valle de San Luis Potosí, Considerando La Demanda Actual y Futura, Bajo Un Marco de Sustentabilidad* (San Luis Potosí, 2008)
7. T. Hergt, J. Castro-Larragoitia, A. Cardona-Benavides, and J. J. Carrillo-Rivera, Ing. Hidraul. En Mex. **24**, 37 (2009)
8. M. C. Noyola-Medrano, J. A. Ramos-Leal, E. Domínguez-Mariani, L. F. Pineda-Martínez, H. López-Loera, and N. Carbajal, Rev. Mex. Ciencias Geol. **26**, 395 (2009)
9. CONAGUA, *Actualización de La Disponibilidad Media Anual de Agua En El Acuífero San Luis Potosí (2411) Estado de San Luis Potosí* (San Luis Potosí, 2015)
10. L. Zou, J. Kent, N. S. Lam, H. Cai, Y. Qiang, and K. Li, 1 (2015)
11. A. M. Oliver and R. Webster, *Basic Steps in Geostatistics: The Variogram and Kriging* (2015)
12. R. Giraldo, Univ. Nac. Colomb. 94 (2011)
13. F. E. Ikuemonisan, V. C. Ozebo, and O. B. Olatinsu, Geod. Geodyn. 1 (2020)
14. M. J. Abedini, M. Nasser, and D. H. Burn, Comput. Geosci. **41**, 136 (2012)