

Diseño óptimo de la red de monitoreo del acuífero “Saltillo-Ramos Arizpe” para el adecuado manejo del recurso hídrico, aplicando un enfoque geoestadístico

• Martín A. Díaz-Viera •
Instituto Mexicano del Petróleo

• Félix Canul-Pech* •
Instituto Nacional de Estadística y Geografía, México

*Autor de correspondencia

Resumen

Díaz-Viera, M. A., & Canul-Pech, F. (septiembre-octubre, 2014). Diseño óptimo de la red de monitoreo del acuífero “Saltillo-Ramos Arizpe” para el adecuado manejo del recurso hídrico, aplicando un enfoque geoestadístico. *Tecnología y Ciencias del Agua*, 5(5), 141-159.

El presente estudio tiene la finalidad de realizar el diseño óptimo de la red de monitoreo de los niveles piezométricos en el acuífero “Saltillo-Ramos Arizpe”, aplicando un enfoque geoestadístico. Para determinar los mejores pozos para el monitoreo se consideran dos alternativas: en la primera se fija un nivel de varianza del error de estimación, mientras que en la segunda se fija la cantidad de pozos, independientemente del error de estimación que resulte. En ambas opciones se aplica la técnica de optimización de las inclusiones sucesivas óptimas, en combinación con el método de kriging ordinario como estimador espacial. Los datos utilizados fueron tomados de 750 aprovechamientos hidráulicos (pozos y norias) de un estudio geohidrológico realizado en el acuífero en 2007. Los niveles piezométricos presentaron el fenómeno de tendencia espacial, es decir, falta de estacionaridad, la cual fue estimada por medio de un ajuste polinomial de primer grado y posteriormente restada a los datos originales, obteniéndose valores residuales. A dichos residuales se les calcularon semivariogramas y se les ajustó un modelo esférico e isotrópico. La red óptima consta de 144 pozos, determinada para una desviación estándar del error de estimación por kriging ordinario en bloques de 21 metros. Lo anterior representa un 19.2% de los 750 aprovechamientos hidráulicos existentes en el acuífero de estudio, lo cual implica un ahorro en un 80.8% de costo de monitoreo en pozos. Asimismo, se determinaron los mejores 50, 100, 200, 300, 400 y 500 pozos para el monitoreo, opciones que se pueden tomar en cuenta dependiendo de los recursos materiales, financieros y humanos disponibles para realizar dicha actividad.

Palabras clave: estimación, kriging, optimización, inclusiones sucesivas, geoestadística, acuífero, red de monitoreo, diseño óptimo de redes.

Abstract

Díaz-Viera, M. A., & Canul-Pech, F. (September-October, 2014). *A Geostatistical Approach to the Optimal Design of the “Saltillo-Ramos Arizpe” Aquifer Monitoring Network for Proper Water Resources Management*. *Water Technology and Sciences (in Spanish)*, 5(5), 141-159.

The purpose of this study is to determine the optimal design of the piezometric level monitoring network in the “Saltillo-Ramos Arizpe” aquifer using a geostatistical approach. To identify the best wells for monitoring, two options were considered: one fixed the level of variance in the estimation error and the other fixed the number of wells regardless of the resulting estimation error. Both options applied the optimal successive inclusions technique in combination with the ordinary kriging method as a spatial estimator. The data used were taken from 750 hydraulic installations (wells and treadmills) in a geohydrological study conducted in the aquifer in 2007. Groundwater levels showed spatial trends —i.e., lack of stationarity— which was estimated by a first degree polynomial fit and then subtracted from the original data, thereby obtaining residuals. With these residuals, semivariograms were calculated and an isotropic spherical model was fitted. The resulting optimal network consists of 144 wells, with a standard deviation of error of 21 meters determined by ordinary block kriging. This represented 19.2% of the 750 existing hydraulic installations in the study aquifer, and implies a 80.8 % savings in the cost of monitoring the wells. The best 50, 100, 200, 300, 400 and 500 monitoring wells were also determined, and represent options which can be considered depending on the material, financial and human resources available for this activity.

Keywords: Estimation, kriging, optimization, successive inclusions, geostatistics, aquifer, monitoring network, optimal network design.

Recibido: 05/12/12
Aceptado: 15/03/14

Introducción

El estudio del recurso hídrico dulce disponible se justifica por la demanda de consumo que se tiene por parte de los seres vivos. Por lo tanto, es importante monitorear el comportamiento del agua subterránea, sobre todo en los acuíferos de alta demanda.

El área metropolitana de Saltillo constituye un área de estudio interesante para la aplicación de la metodología del presente trabajo. Según los resultados de los censos de población publicados por el Instituto Nacional de Estadística y Geografía (INEGI, 2010), esta zona ha presentado un crecimiento urbano muy importante en los últimos años y la demanda de agua es abastecida principalmente por el acuífero "Saltillo-Ramos Arizpe", ya que no cuenta con otras fuentes importantes de agua dulce.

La metodología utilizada se basa en aprovechar que la varianza del error de estimación con kriging no depende del valor de la variable; lo anterior, según Samper y Carrera (1990). Por lo tanto, es posible calcular dicha varianza antes de realizar las medidas y obtener la localización de los puntos de medición, en este caso pozos, de forma que se minimice la incertidumbre de estimación. También se puede minimizar dicha incertidumbre dada una cantidad prefijada de pozos.

Algunos trabajos de referencia son: Prakash y Singh (2000), quienes realizaron un diseño de red óptima en la cuenca alta de Kongal, distrito de Nalgonda, A. P. (India), aplicando kriging; Banjevic y Switzer (2002) compararon estrategias de optimización de recocido simultáneo simulado y algoritmos de puntos de selección secuenciales; Dixon, Smyth y Chiswell (1999) aplicaron procedimientos de sistemas de información geográfica y algoritmos de recocido simulado; Lloyd y Atkinson (1999) realizaron diseños óptimos de configuración de muestreos utilizando kriging ordinario y kriging indicador; Cousens, Brown, Mcbratney, & Moerkerk (2002) aplicaron diferentes estrategias de muestreo para la

generación de mapas de malezas; Wang y Qi (1998) estudiaron el efecto de la distribución de muestras sobre el análisis de la estructura espacial de un suelo contaminado, entre otros.

El propósito de este estudio fue establecer y aplicar una metodología para determinar la red óptima de pozos para el monitoreo de los niveles piezométricos en el acuífero en cuestión con fines de manejo. Para conseguir lo anterior, se aplicó el enfoque geoestadístico: análisis preliminar y exploratorio de los datos, obtención de un modelo variográfico de la correlación espacial, estimación con kriging ordinario y optimización, aplicando la técnica de inclusiones sucesivas.

Objetivos

General

Establecer y aplicar a un caso de estudio una metodología que permita de manera sistemática obtener el diseño óptimo de redes de monitoreo con fines de manejo, a partir del análisis geoestadístico de los datos de niveles piezométricos.

Específicos

- Obtener un modelo de la variabilidad espacial de los datos de niveles piezométricos existentes mediante el análisis geoestadístico.
- Evaluar el comportamiento espacial de la varianza del error de estimación, aplicando kriging ordinario como método de estimación de los niveles piezométricos.
- Determinar la red óptima de pozos para el monitoreo de los niveles piezométricos en el acuífero "Saltillo-Ramos Arizpe" bajo diferentes escenarios.

Metodologías para el diseño óptimo de redes de monitoreo

Para el *diseño óptimo de redes de monitoreo de aguas subterráneas* se han propuesto, en diversas

partes del mundo, metodologías basadas en enfoques geoestadísticos o usando el "filtrado de Kalman". El primero permite evaluar el nivel de incertidumbre a través de la varianza del error de estimación en los puntos de una red, con base en la correlación espacial de una variable. El otro permite evaluar el grado de incertidumbre mediante el acoplamiento de un algoritmo de filtrado a un modelo de flujo hidrodinámico y ayuda a mejorar la exactitud de los resultados y la calibración del modelo. Se considera que ambos enfoques son equivalentes.

Para aplicar el enfoque geoestadístico se requiere de una función que describa la correlación espacial de la propiedad en estudio $Z(x)$. Usualmente puede utilizarse el semivariograma (Chiles & Delfiner, 1999; Olea, 1999; Isaaks & Srivastava, 1989; Moral-García, 2003).

El estimador más común del semivariograma o semivarianza según Omre (1984) se puede expresar como:

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

Donde N es el número de pares de $[Z(x+h) - Z(x)]$ separados a una distancia h . Cuando la función aleatoria es estacionaria, se cumple la relación:

$$\gamma(h) = \text{Var}(Z) - C(h) \quad (2)$$

Por lo tanto, en este caso se tiene definido tanto el semivariograma como la función de covarianza.

Normalmente, el semivariograma es una función monótona creciente, ya que al incrementarse la distancia h aumenta la diferencia $|Z(x+h) - Z(x)|$. Si Z es una función aleatoria estacionaria, γ alcanza un valor límite, denominado meseta o *sill*, equivalente a la varianza de Z ; la meseta se alcanza para un valor de h conocido como rango. El rango

determina la zona de influencia; fuera de ésta, la correlación espacial es nula. Por definición, $\gamma(0) = 0$, pero frecuentemente $\gamma(0)$ tiene valor positivo, denominado efecto pepita o nugget.

Sin embargo, no todos los semivariogramas alcanzan una meseta. Es posible que alguno no tienda asintóticamente a la varianza, sino que tienda a infinito cuando h tienda a infinito. Además, el rango no tiene que ser igual en todas las direcciones, reflejando en esa situación la existencia de anisotropía. También puede ser que para una dirección determinada haya más de un rango, lo cual indica la presencia de distintas estructuras de correlación actuando a diferentes escalas.

El sistema de kriging puntual ordinario (Isaaks & Srivastava, 1989) se puede escribir como:

$$\hat{Z}_0 = \sum_{i=1}^n W_i Z(x_i) \quad (3)$$

$$\sum_{j=1}^n W_j \hat{C}_{ij} + \mu = \hat{C}_{i0} \quad \forall i = 1, \dots, n$$

$$\sum_{i=1}^n W_i = 1$$

Donde \hat{Z}_0 es el estimador; W_i , los pesos para kriging puntual ordinario; μ , un multiplicador de Lagrange; \hat{C}_{ij} y \hat{C}_{i0} , las covarianzas que para el caso de la hipótesis intrínseca, es decir, estacionaridad en segundo orden de la variable a estimar, se pueden sustituir por las semivarianzas.

La varianza del error minimizado se puede expresar como:

$$\hat{\sigma}_R^2 = \hat{\sigma}^2 - \left(\sum_{i=1}^n w_i \hat{C}_{i0} + \mu \right) \quad (4)$$

Esto minimiza la varianza del error de estimación que normalmente se conoce como la varianza kriging ordinario.

El estimador de kriging en bloques (Moral-García, 2003) tendrá la forma:

$$Z_{KB}^*(x) = \sum_{i=1}^n w_i Z(x_i) \quad (5)$$

Siendo w_i los pesos para kriging ordinario en bloques que se asigna a los datos $Z(x_i)$. Como en el caso del kriging puntual ordinario, el estimador para el bloque debe ser insesgado, tal que minimice el error de la varianza:

$$\sum_{j=1}^n w_j C(x_i - x_j) + \mu = \bar{C}(x_i - V(x)) \quad (6)$$

$$i = 1, \dots, n$$

$$\sum_{i=1}^n w_i = 1$$

El sistema es idéntico al puntual, salvo que el término de la derecha, $\bar{C}(x_i - V(x))$ se refiere a la covarianza entre el punto x_i y el bloque $V(x)$, o sea, la covarianza media entre la variable aleatoria $Z(x_i)$, y la variable aleatoria $Z(x)$, siendo ésta la correspondiente a todos los puntos del bloque V .

En la práctica, la covarianza $\bar{C}(x_i - V(x))$ se aproxima con la media aritmética de la covarianza entre los puntos muestra x_i y los m puntos obtenidos al discretizar el bloque V :

$$\bar{C}(x_i - V(x)) = \frac{1}{m} \sum_{j=1}^m C(x_i - x'_j) \quad (7)$$

La varianza de kriging en bloques es:

$$\sigma_{ku}^2 = \bar{C}(V(x), V(x)) - \sum_{i=1}^n w_i \bar{C}(x_i, V(x)) - \mu \quad (8)$$

Donde la covarianza en el bloque $\bar{C}(V(x), V(x))$ se aproxima con la media aritmética de las covarianzas entre los puntos en los cuales se discretiza el bloque:

$$\bar{C}(V(x), V(x)) = \frac{1}{m^2} \sum_{j=1}^m \sum_{k=1}^m \bar{C}(x'_j - x'_k) \quad (9)$$

Para el diseño óptimo de redes de monitoreo, utilizando el enfoque geoestadístico se aplican los siguientes métodos (Samper & Carrera, 1990):

Método del punto ficticio

Si se tiene un número n de mediciones correspondientes a n puntos $\{x_i, i = 1, \dots, n\}$ espacialmente distribuidos y se desea conocer el valor de la medición en un punto adicional x_{n+1} .

1. Se considera primero un punto adicional ficticio $n + 1$.
2. Se estima el valor en x_{n+1} mediante kriging.
3. Se obtiene la varianza de la estimación de kriging.

Finalmente, se repiten los puntos 1, 2 y 3, variando la posición del punto hasta obtener el óptimo, que minimice la varianza.

Los métodos se pueden clasificar como *globales* y *locales*, según se quiera estimar el valor medio de la propiedad de manera global o se pretenda mejorar localmente la estimación de la misma.

Estimación global

Pretende minimizar la varianza de la estimación del valor medio de la propiedad de manera general, es decir, se considera a la región donde se diseña la red como un todo. Se define el índice $R(x)$ de reducción relativa de la incertidumbre:

$$R(x) = \frac{\sigma_0^2 - \sigma^2(x)}{\sigma_0^2} \quad (10)$$

Donde:

- σ_0^2 = varianza de la estimación con los n puntos existentes.
- $\sigma^2(x)$ = varianza de la estimación con un punto adicional $n + 1$ en la posición x .

Algoritmo 1

Si existen posiciones predefinidas para x_{n+1} , entonces se calcula $\sigma^2(x)$ para todas las posiciones y se selecciona la que tiene mayor $R(x)$.

Algoritmo 2

Si no hay restricciones para la posición de x_{n+1} , entonces se puede calcular $\sigma^2(x)$ en una malla y dibujar las curvas de nivel para $R(x)$.

Estimación local

Persigue la minimización de la varianza en cierta zona de la red de manera local.

Este enfoque provoca que existan tantas varianzas como puntos a estimar, por lo que sólo se puede aspirar a eliminar los máximos de la incertidumbre.

1. Se añade un punto en las zonas de mayor incertidumbre.
2. Se calculan las curvas de nivel de las desviaciones típicas para los casos con y sin el punto, y se comparan los resultados.

Esto continúa hasta encontrar el punto que más reduzca la incertidumbre.

En ambos enfoques se podría repetir el algoritmo agregando otros puntos. Existen estrategias eficientes para la inclusión de más de un punto, como son:

- El método de inclusiones sucesivas.
- El método de intercambios sucesivos óptimos.
- El método de enumeración total.
- El método de rama y límite, etcétera.

Estrategias de optimización usando más de un punto

Para el caso de redes irregulares, éste representa un problema discreto, donde se trata de

seleccionar el mejor subconjunto de puntos de un conjunto dado.

a) Método de inclusiones sucesivas

Se añade sucesivamente el punto que minimiza la varianza en cada paso hasta completar un número de puntos prefijados o llegar a un nivel de varianza dado.

b) Método de intercambios sucesivos óptimos

Se elige un conjunto X_0 de puntos tales que al intercambiar uno de sus elementos por uno de los restantes no se disminuye la varianza, lo cual es una condición necesaria, pero no suficiente para obtener el conjunto óptimo.

Algoritmo

1. Sea $X_0 = \{t_1, t_2, \dots, t_{n-k}\}$ y $\{X_1 = t_{n-k+1}, \dots, t_N\}$ y j_1 .
2. Se intercambia el elemento j -ésimo de X_0 sistemáticamente con todos los de X_1 y se calculan las varianzas correspondientes.
3. Cuando la varianza disminuye, entonces se hace permanente el intercambio.
4. Se modifica el valor de j según:

$$j = \begin{cases} j+1, & \text{si } j < N \\ 1, & \text{si } j = N \end{cases}$$

y se retorna al paso 2. El proceso se detiene cuando ninguno de los elementos de X_0 puede ser intercambiado de forma tal que disminuya la varianza de la estimación.

c) Método de enumeración total

Se examinan todos los subconjuntos del conjunto $T = \{t_1, t_2, \dots, t_N\}$ que tengan $n - k$ elementos. Aquel que resulte de varianza mínima es el óptimo.

El número total de subconjuntos a evaluar sería:

$$\binom{N}{n-k} = \frac{N!}{(N-n+k)!(n-k)!} \quad (11)$$

Una manera sistemática sería empleando una búsqueda de árbol. Se comienza con el conjunto nulo o vacío. Tiene la dificultad de que el número de operaciones a realizar es muy grande.

d) Método de rama y límite

Similar al anterior, con las diferencias:

- El árbol se construye partiendo del conjunto total de puntos.
- Cuando se avanza por el árbol se elimina un punto y cuando se retrocede se añade.
- Se compara una condición adicional, que la varianza de la estimación sea menor o igual que la más pequeña de las calculadas hasta este momento con $n - k$ puntos; en caso contrario, no se continúa avanzando, ya que de hacerlo se incrementaría la misma. Tiene la ventaja con respecto al anterior que reduce el tiempo de cálculo.

La técnica de optimización utilizada en este estudio es el de inclusiones sucesivas y el algoritmo (Carrera, Usunoff, & Szidarovszky, 1984; Carrera & Szidarovszky, 1985; Díaz-Viera, 1997, 1998) consiste en: al conjunto m de puntos existentes lo designamos por $X = \{x_1, x_2, \dots, x_m\}$ y al subconjunto óptimo de n puntos que minimizan la varianza de estimación lo designaremos como $X_o = \{x_{o1}, x_{o2}, \dots, x_{on}\}$.

Algoritmo

1. Para $j = 0$, $X = \{x_1, x_2, \dots, x_m\}$ y $X_o = \{\emptyset\}$, donde $\emptyset =$ vacío.
2. Se calcula la varianza del error de estimación por kriging $\sigma^2(x_i)$ para i desde 1 hasta m y se selecciona el que tiene menor valor, de manera que el punto x_i se extrae del conjunto X y se incluye en el conjunto X_o resultando:

$$X = \{x_1, x_2, \dots, x_{m-j}\} \text{ y } X_o = \{x_{o1}\}$$

3. Se incrementa $j = j + 1$.

4. Para todos los valores de i , desde 1 hasta $(m - j)$, se evalúa la expresión (10), o sea, $R(x_i) = \frac{\sigma_j^2 - \sigma_{j+1}^2(x_i)}{\sigma_j^2}$, que es el índice de reducción relativa.

donde:

σ_j^2 = varianza del error de la estimación con los j puntos que pertenecen a X_o .

σ_{j+1}^2 = varianza del error de la estimación con un punto adicional $j + 1$ en la posición x_i .

5. El punto x_i que más reduzca la varianza del error de estimación, es decir, que tenga mayor valor de $R(x_i)$, se selecciona de manera que se extrae del conjunto X y se agrega al conjunto X_o resultando:

$$X = \{x_1, x_2, \dots, x_m\} \text{ y } X_o = \{x_{o1}, x_{o2}, \dots, x_{oj}\}$$

6. Mientras se cumpla que $j < n$ o $\sigma_j^2 <$ tolerancia, se retorna al paso 3; en caso contrario, se termina el procedimiento.

Las ecuaciones para el cálculo de la inversa de la matriz de kriging y la varianza de estimación con $j + 1$ y $j - 1$ puntos, a partir de los valores obtenidos con j puntos, según Samper y Carrera (1990), se deduce: sea, $G \cdot \lambda = \gamma$ y sea λ_0 la solución de este sistema. Si se añade un nuevo punto x_{j+1} las nuevas ecuaciones pueden escribirse como:

$$\begin{pmatrix} G & \vdots & g \\ \dots & \dots & \dots \\ g^t & \vdots & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ \vdots \\ \lambda_{j+1} \end{pmatrix} = \begin{pmatrix} \gamma \\ \vdots \\ \gamma_{vj+1} \end{pmatrix} \quad (12)$$

Donde $g^t = (\gamma_{j+1,1}, \dots, \gamma_{j+1,j+1})$. La inversa de la nueva matriz de coeficientes viene dada por:

$$\begin{pmatrix} X & \vdots & y \\ \dots & \dots & \dots \\ y^t & \vdots & s \end{pmatrix} \quad (13)$$

donde:

$$s = \frac{1}{g^t G^{-1} g} \quad (14)$$

$$y = -G^{-1} g s$$

$$X = G^{-1} + s^{-1} y y^t$$

Y la solución del sistema puede expresarse como:

$$\lambda_{j+1} = s (\gamma_{vj+1} - g^t \lambda_o) \quad (15)$$

$$\lambda = \lambda_o - G^{-1} g \lambda_{j+1} \quad (16)$$

La nueva varianza, tras añadir el punto x_{j+1} es:

$$\sigma_{j+1}^2 = \sigma_j^2 - (\gamma_{vj+1} - g^t G^{-1} g)^2 / (g^t G^{-1} g) \quad (17)$$

Estas expresiones son de gran utilidad en los cálculos de las varianzas del error de estimación al adicionar un punto.

La metodología geoestadística utilizada consta de las siguientes etapas: análisis preliminar de los datos, análisis exploratorio, análisis variográfico y estimación espacial.

En la etapa de *análisis preliminar* se calculan los niveles piezométricos, restando las alturas de brocal de los pozos y las profundidades del nivel estático; se genera un archivo vectorial en formato shape de puntos, con los atributos de los aprovechamientos hidráulicos asociados (clave, profundidad del nivel estático, altura de brocal y nivel piezométrico), cambiando de datum de NAD27 a ITRF92 y de proyección cartográfica de universal transversa de mercator (UTM) a cónica conforme de Lambert (CCL), a través de las ecuaciones citadas en Leick (1990).

Se realiza la sobreposición de la capa de curvas de nivel topográficas a escala 1:250 000, límites de las localidades urbanas (datos del Instituto Nacional de Estadística y Geografía) y el polígono del acuífero en cuestión, para verificar la posición de los pozos y sus atributos.

En la etapa de *análisis exploratorio* se calculan los parámetros estadísticos básicos, como media, varianza, desviación estándar, cuartiles, valor máximo, valor mínimo, coeficiente de curtosis y asimetría. Asimismo, generar el histograma de frecuencias, las gráficas de Q-Q Plot y de caja. Se plantea detectar los valores atípicos, separarlos de los demás datos y volver a calcular los parámetros estadísticos básicos. La idea es realizar un análisis gráfico-numérico de la naturaleza de los datos, evaluar algunos supuestos, como la normalidad, linealidad y homocedasticidad (Armstrong & Delfiner, 1980); evaluar la presencia de los valores atípicos, los cuales son los valores que exceden los límites establecidos por la siguiente expresión (Díaz-Viera & Barandela, 1994; Díaz-Viera, Hernández-Maldonado, & Méndez-Venegas, 2010):

$$\text{Límite superior} = q(0.75) + 1.5 iq$$

$$\text{Límite inferior} = q(0.25) - 1.5 iq$$

Donde $q(p)$ son los cuantiles, mientras que iq es el rango intercuartil (diferencia entre el tercer y el primer cuartil); los valores atípicos pueden influir de manera significativa en etapas posteriores del análisis geoestadístico (Cressie & Hawkins, 1980; Díaz-Viera, 2002), y en el impacto de la ausencia de datos.

En la etapa de *análisis variográfico* se estiman, en primer término, las semivarianzas adireccionales; del mismo modo, las semivarianzas para las direcciones de 45, 90 y 270° a diferentes distancias, analizándose para determinar la posible presencia de anisotropía. En caso de existir anisotropía, se estima la dirección de ésta, y los alcances de menor y mayor valor para su consideración. Asimismo, se obtiene el modelo de semivariograma que mejor se ajuste a los datos siguiendo el criterio de información de Akaike (AIC) (Akaike, 1974, 1977), el cual relaciona la complejidad del modelo (número de parámetros) y la bondad de ajuste (error medio cuadrático); se le aplica la validación cruzada, en específico el método "leave one out" mencionado por Journel y

Huïjbregts (1978), que consiste en dejar un punto conocido fuera de la estimación, y proceder a calcular su valor con kriging ordinario, utilizando los parámetros del modelo del semivariograma obtenido; de forma análoga se procede con los demás puntos, consiguiendo las diferencias entre el valor real y el estimado; se analizan las estadísticas de estas diferencias y se espera que el valor medio sea cercano a cero, una varianza pequeña y una varianza normalizada cercano a uno, lo anterior según Omre (1984).

Si se detecta presencia de tendencia espacial $M(x, y)$, se procede a estimar dicha tendencia a partir de polinomios de primer o segundo grado; posteriormente se realiza la diferencia al valor conocido $Z(x, y)$, quedando un residual $R(x, y)$, es decir, $R(x, y) = Z(x, y) - M(x, y)$; entonces todo el procedimiento geoestadístico se aplicará sobre los residuales $R(x, y)$.

La finalidad principal de esta etapa es estimar y modelar una función que refleje la correlación espacial de los niveles piezométricos a partir de la adopción razonada de la hipótesis más adecuada acerca de su variabilidad.

En la *estimación espacial* se generará una malla de puntos en el área de estudio, considerando las coordenadas extremas del polígono del acuífero; la resolución de esta malla será de 500 por 500 metros, lo anterior, considerando la resolución espacial de los datos conocidos; posteriormente, basado en los parámetros del modelo de semivariograma que se obtenga en el análisis variográfico, se procederá a estimar en cada nodo de la malla la variable en estudio (nivel piezométrico o el residual de éste) y la variancia del error de estimación; para lo anterior, se buscarán los 16 pozos más cercanos a cada nodo. Moral-García (2003) recomienda poco más de una decena para esta operación. Para estas estimaciones se utilizará kriging puntual ordinario.

Características del área y datos de estudio

El área de estudio (figura 1) es el acuífero "Saltillo-Ramos Arizpe". Se localiza en el

sureste de Coahuila de Zaragoza, México, y cubre parte de los municipios de Saltillo, Ramos Arizpe y Arteaga; se enmarca entre las coordenadas geográficas $101^{\circ} 06'$ y $100^{\circ} 47'$ longitud oeste, y $25^{\circ} 35'$ y $25^{\circ} 20'$ de latitud norte, datum ITRF92. Al sur y sureste colinda con la sierra de Zapalinamé; al oriente, con la sierra de Arteaga, sierra San Lucas y sierra San José de los Nuncios; y al poniente, con la sierra El Asta y sierra Palma Gorda. Dentro del área que ocupa el acuífero se encuentra la ciudad de Saltillo, capital del estado de Coahuila de Zaragoza, y las poblaciones de Ramos Arizpe y Arteaga.

Las coordenadas de los vértices que conforman la poligonal de este acuífero y la ubicación de los pozos se transformaron al datum ITRF92 y en la proyección cartográfica cónica conforme de Lambert (CCL), utilizando los siguientes parámetros:

- Falso este: 2 500 000 m.
- Falso norte: 0 m.
- Primer paralelo: 17.5° .
- Segundo paralelo: 29.5° .
- Paralelo de origen: 12° norte.
- Meridiano de origen: 102° oeste.

La zona se localiza en la provincia fisiográfica de las sierras y cuencas, en el límite de la provincia de la Sierra Madre Oriental (Conagua, 2010). Las características fisiográficas son de dos tipos: a) la zona de topografía relativamente suave corresponde al Valle de Saltillo-Ramos Arizpe, en donde la elevación del terreno disminuye hacia el norte, con valores de 1 800 a 1 200 msnm; y b) la zona de sierras del poniente, orientadas este-oeste, que se caracterizan por una topografía abrupta, con elevaciones máximas de 2 400 msnm.

Está formado en la parte del valle por los depósitos aluviales, conglomerado Reynosa y lutitas fracturadas de la formación Parras. Las rocas areniscas fracturadas del grupo Difunta forman parte de este acuífero en las zonas de las sierras localizadas al oeste y noroeste del valle.

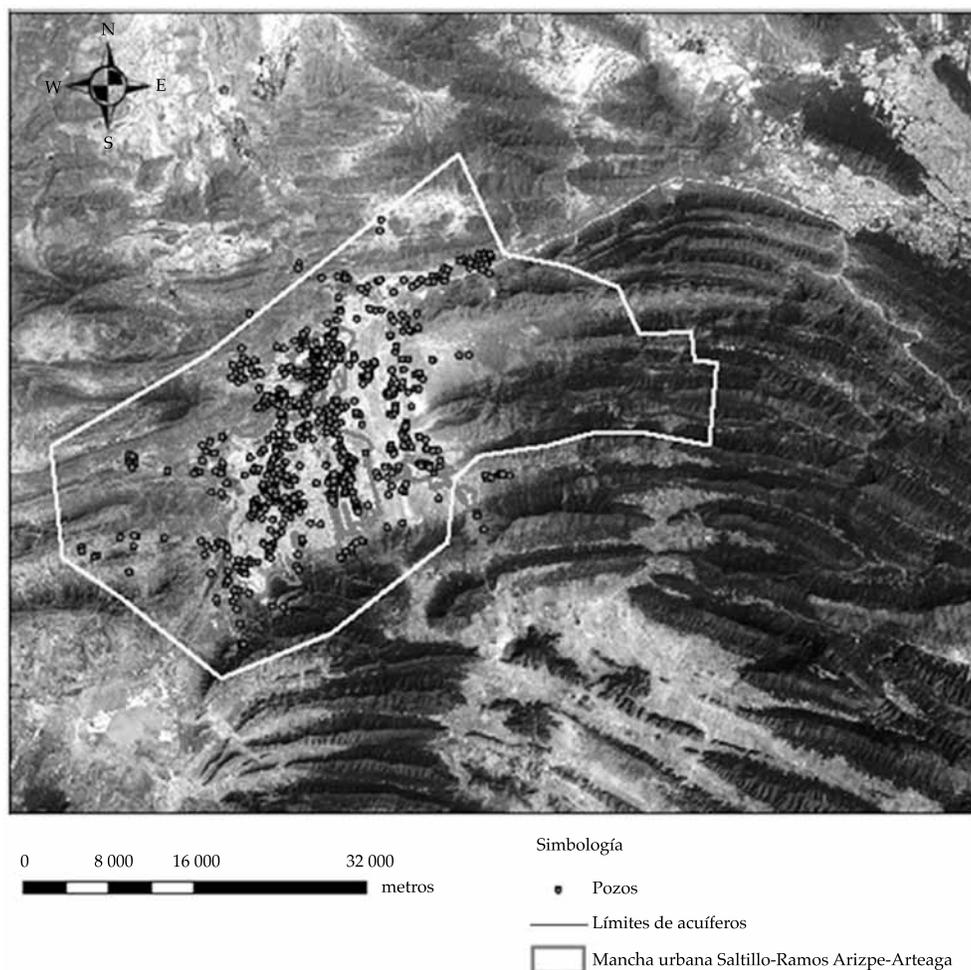


Figura 1. Ubicación del acuífero "Saltillo_Ramos Arizpe" y los 750 aprovechamientos hidráulicos.

Es un acuífero de tipo semiconfinado y el basamento hidrogeológico —definiendo basamento a la profundidad a la cual las fracturas en la Formación Parras desaparecen o se cierran a profundidad— se ubica a 250 y 450 m.

Los límites del acuífero de carácter impermeable son el contacto entre la Formación Parras fracturada y la Formación Indidura de baja permeabilidad, al oeste y sureste del valle.

El acuífero consta de 750 aprovechamientos hidráulicos. Los registros de éstos fueron proporcionados por la Comisión Nacional del Agua (Conagua) de un estudio geohidrológico realizado en 2007.

Resultados

Análisis preliminar

De 750 aprovechamientos hidráulicos (figura 1), 413 cuentan con registros de profundidad de nivel estático. Se revisaron las alturas de los brocales de estos últimos pozos con respecto a las curvas de nivel de la carta topográfica escala 1:250 000, serie III de INEGI; detectándose inconsistente una altura de brocal (pozo SRA-071), con un valor de 5 261 m, ubicándose entre las curvas de nivel topográficas de 1 500 y 1 600 m; asimismo, cuatro no contaban con dicho valor (claves SRA-292, SRA-423, SRA-424, SRA-425),

resultando 408 pozos con datos consistentes. Se calcularon los niveles piezométricos para los análisis subsecuentes.

Análisis exploratorio

Los 408 niveles piezométricos presentaron tendencia espacial, ya que las semivarianzas que se calcularon no se acotan en la línea de la varianza (figura 2). Para eliminar la tendencia espacial, se realiza un ajuste polinomial de primer grado a los datos. Se obtienen los siguientes coeficientes:

$$\begin{aligned} a &= 29\,674.64084 \text{ m} \\ b &= -0.002119162 \text{ m} \\ c &= -0.0151855557 \text{ m} \end{aligned}$$

Para expresar esta variable en la forma:

$$Z(x,y) = M(x,y) + R(x,y) \quad (18)$$

Donde $M(x,y) = a + b * X + c * Y$ es la tendencia y $R(x) = Z(x) - M(x,y)$ son los residuales de los niveles piezométricos.

Se realiza el análisis exploratorio de los residuales $R(x)$; se detectan y eliminan 26 valores atípicos.

En el cuadro 1 se visualiza que la asimetría es positiva, con un valor de 0.2883 m; la media (-2.9250) es mayor que la mediana (-5.2860), obteniéndose un rango intercuartil de 71.24 m.

En el histograma de frecuencias (figura 2) se alcanza a distinguir la asimetría positiva. En la gráfica de cajas se observa que la media es mayor que la mediana, notando aún algunos valores atípicos.

Análisis variográfico

Se estimaron las semivarianzas, las cuales ya se acotan a la línea de la varianza (figura 3), obteniéndose un modelo de semivariograma adireccional, resultando con los siguientes parámetros:

- Modelo: esférico.
- Pepita: 130.4765 m.

- Meseta-pepita: 739.8475 m.
- Alcance: 6 708.8846 m.
- AIC: 4527.1123.

Asimismo, se visualiza gráficamente la validación cruzada del modelo de semivariograma, obteniéndose un coeficiente de regresión de 0.984. Las estadísticas obtenidas de la validación son:

Media de las diferencias entre el valor real y el estimado	-0.21699 m
Varianza de las diferencias entre el valor real y el estimado	59 365.26
Media de las diferencias normalizadas	-0.111604
Varianza de las diferencias normalizadas	9.755

La diferencia normalizada es la diferencia entre el valor real de la variable y el estimado, inversamente proporcional con la desviación estándar del error de estimación.

Análisis de anisotropía

En la figura 4 se visualiza que los valores de las semivarianzas son similares en todas direcciones; se observa un comportamiento isotrópico, determinándose que no hay presencia de anisotropía o al menos no existen los suficientes elementos para determinar su presencia.

Estimación espacial

Con los parámetros del modelo de semivariograma obtenidos en el análisis variográfico se realiza la estimación espacial, aplicando kriging puntual ordinario y utilizando una malla de 500 por 500 metros, considerando el espacio que cubre el polígono del acuífero en cuestión.

En la figura 5 se muestra el mapa en 3D de los residuales estimados de los niveles piezométricos; los valores representados van de -60.8 a 73.9 metros.

En la figura 6 se observan las desviaciones estándares del error de estimación más grandes

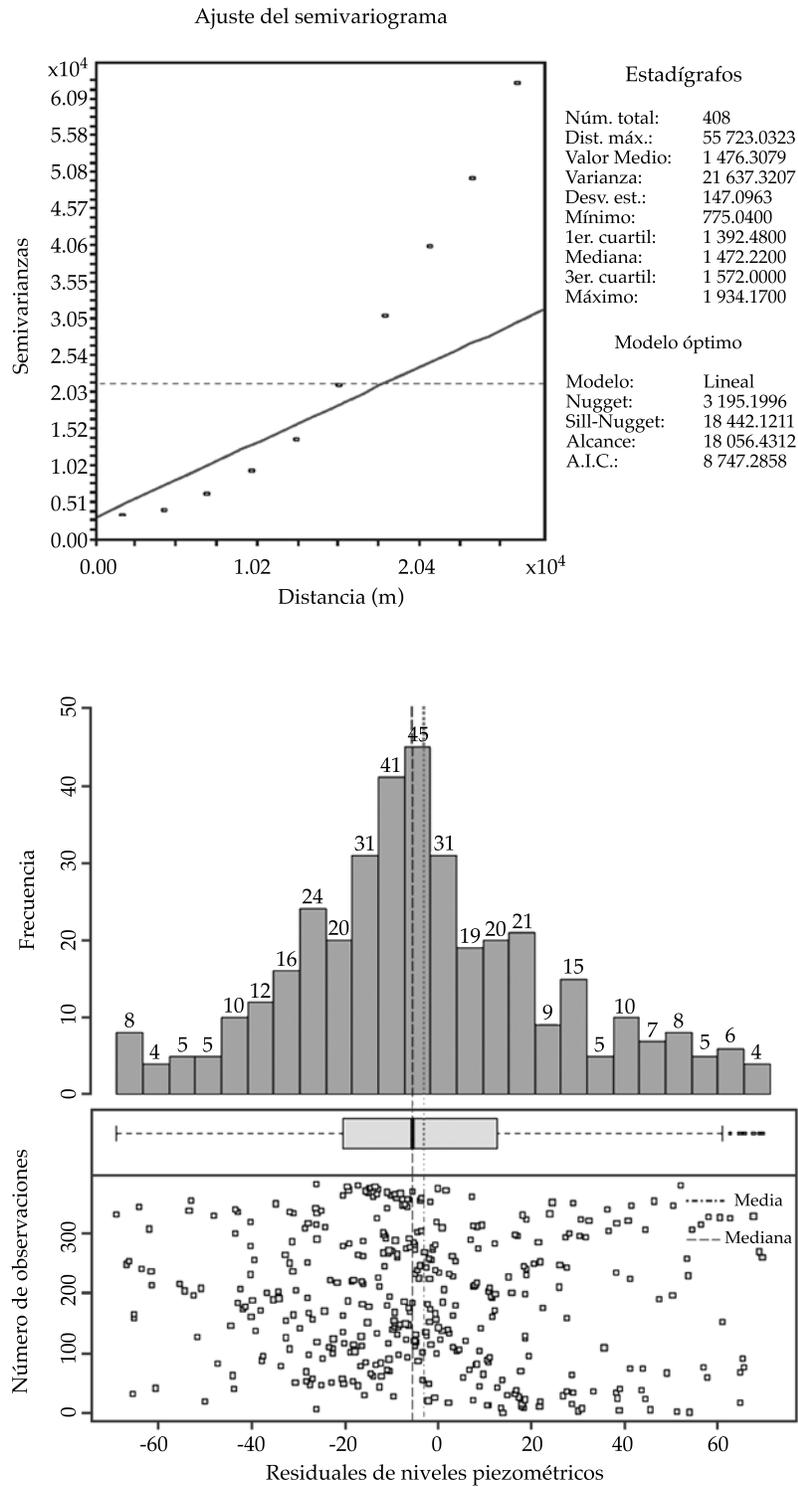


Figura 2. Presencia de tendencia espacial, histograma de frecuencias y gráfica de cajas (*box plot*) de los residuales de los niveles piezométricos sin los 26 *outliers* (382 datos).

Cuadro 1. Parámetros estadísticos básicos de los residuales sin los 26 valores atípicos.

Parámetros	Valor
Total de datos	382
Distancia máxima	50 078.43
Media	-2.9250
Varianza	826.2219
Desviación estándar	28.7441
Coefficiente de variación	9.8285
Mediana	-5.2860
Primer cuartil (Q1)	-20.32
Segundo cuartil (Q3)	13.24
Rango intercuartil	71.24
Asimetría	0.2883
Curtosis	3.0251
Valor máximo	71.2412
Valor mínimo	-68.9487

Cuadro 2. Variación de la desviación estándar de la estimación con respecto al número de pozos.

Núm. de pozos	Desviación estándar del error de estimación (m)
1	862.96
50	24.73
60	23.54
70	22.78
80	22.22
90	21.84
100	21.58
110	21.38
120	21.22
130	21.11
140	21.03
144	21.00
150	20.96
160	20.9
200	20.76
300	20.65
400	20.63
450	20.62

(29.5 m) ubicadas en la periferia de la zona mapeada; y las desviaciones más pequeñas (14.2 m) se distribuyen a lo largo del mapa

de suroeste a noreste, cercanos a los pozos utilizados para determinar el modelo del semivariograma.

Una vez estimados los residuales $\hat{R}(x,y)$, se les adicionó la tendencia $M(x,y)$ utilizando la ecuación (18), para obtener los niveles piezométricos estimados $\hat{Z}(x,y)$.

En la figura 7 se observa el mapa en 2D de los niveles piezométricos estimados; van en la dirección suroeste a noreste, con valores de 1 786.96 a 1 181.94 m.

Optimización utilizando inclusiones sucesivas

A continuación se muestran los resultados de la optimización. Se presentan dos variantes:

Optimización considerando una tolerancia en la desviación estándar del error de estimación

En la figura 8 se visualiza el comportamiento de la desviación estándar del error de estimación con kriging ordinario en bloques, a medida que se adicionan pozos por inclusiones sucesivas.

A partir de 144 pozos, el valor de la desviación estándar del error de estimación presenta variación a nivel de centímetros, resultando un valor de 21 m, como se puede observar en el cuadro 2.

En la figura 9 se visualiza la distribución espacial de los 144 pozos. Se observa que cubren la mayor parte del acuífero, excepto en la zona noreste, en donde no se cuenta con pozos.

Optimización prefijando cantidad de pozos

En el cuadro 3 se presenta la comparación de los parámetros estadísticos básicos de la desviación estándar del error de estimación con kriging puntual ordinario de los mejores pozos para el monitoreo obtenidos en la optimización, prefijando una cantidad de éstos.

Se observa que la media y mediana disminuyen conforme aumenta la cantidad de pozos; caso contrario sucede con la varianza y, por lo tanto, con la desviación estándar.

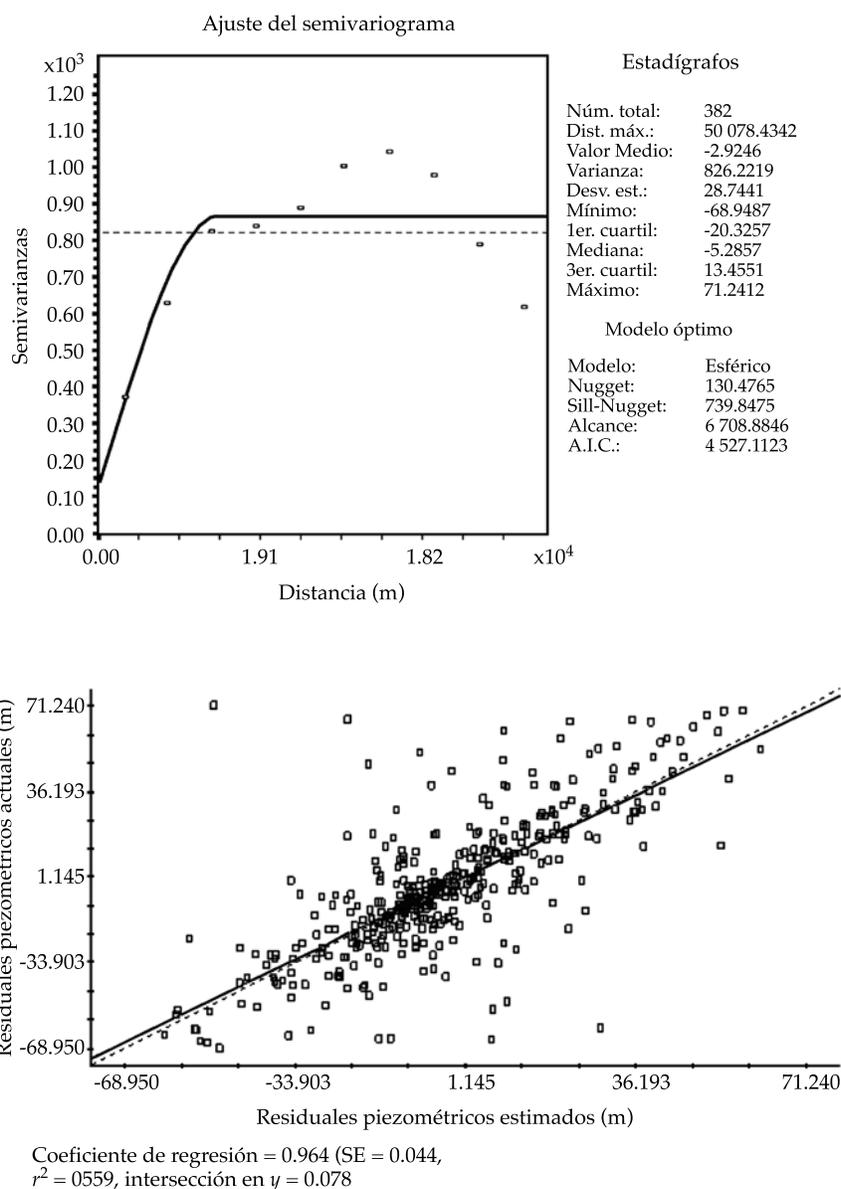


Figura 3. Ajuste y validación cruzada del modelo de semivariograma.

En la figura 10 se muestra la ubicación de los mejores 50 pozos; se nota que se distribuyen por el acuífero, faltando en la zona este, donde no existen pozos

Discusión

Este estudio fue realizado considerando los niveles piezométricos de los pozos del acuífero en cuestión. Dichos datos presentaron tendencia

espacial, la cual fue estimada por medio de un ajuste polinomial de primer grado y restada a los datos originales obteniendo residuales.

El análisis geoestadístico subsecuente y la optimización se realizó utilizando los residuales obtenidos para no verse afectados por la falta de estacionaridad.

El modelo de semivariograma obtenido es esférico, no se detectó la presencia de anisotropía, es decir, resultó isotrópico.

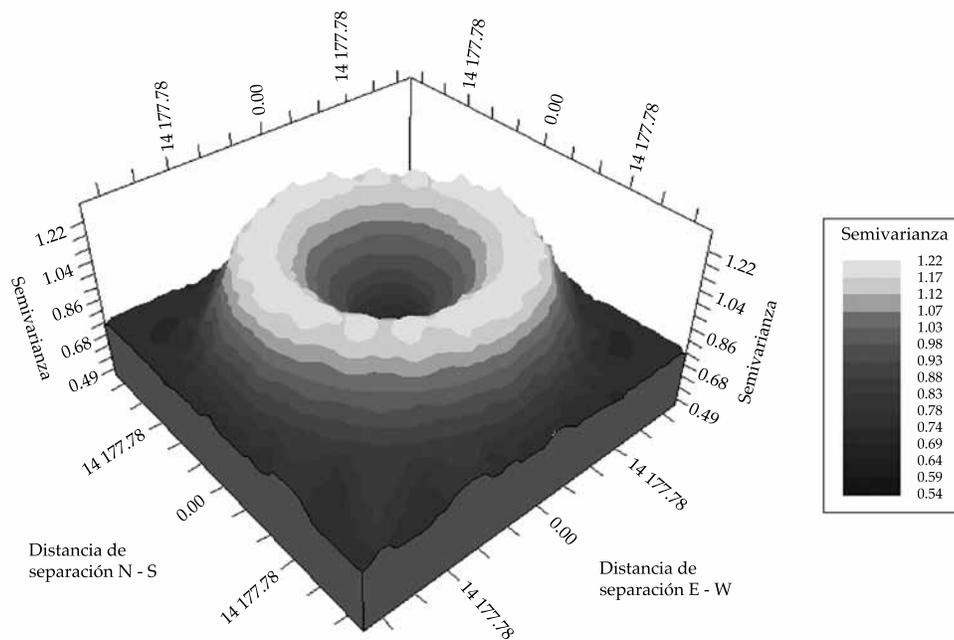


Figura 4. Semivarianzas a direcciones 0, 45, 90 y 135°, y a diferentes distancias.

Estimación de residuales de los niveles piezométricos. Kriging puntual ordinario

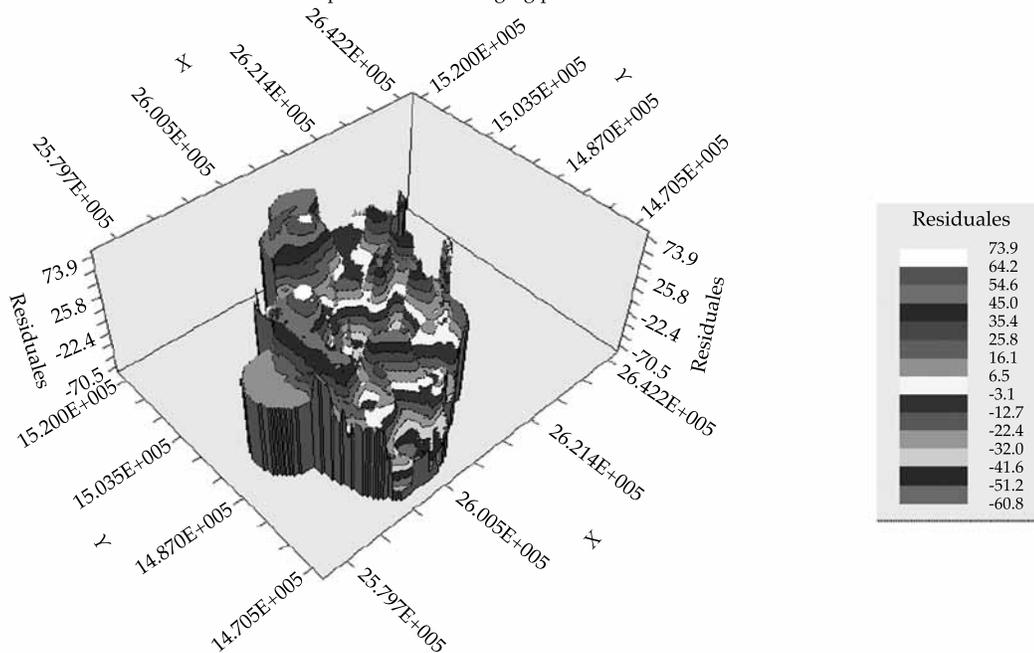


Figura 5. Mapa en 3D de los residuales estimados de los niveles piezométricos.

Desviación estándar de la estimación sin optimizar. Kriging puntual ordinario

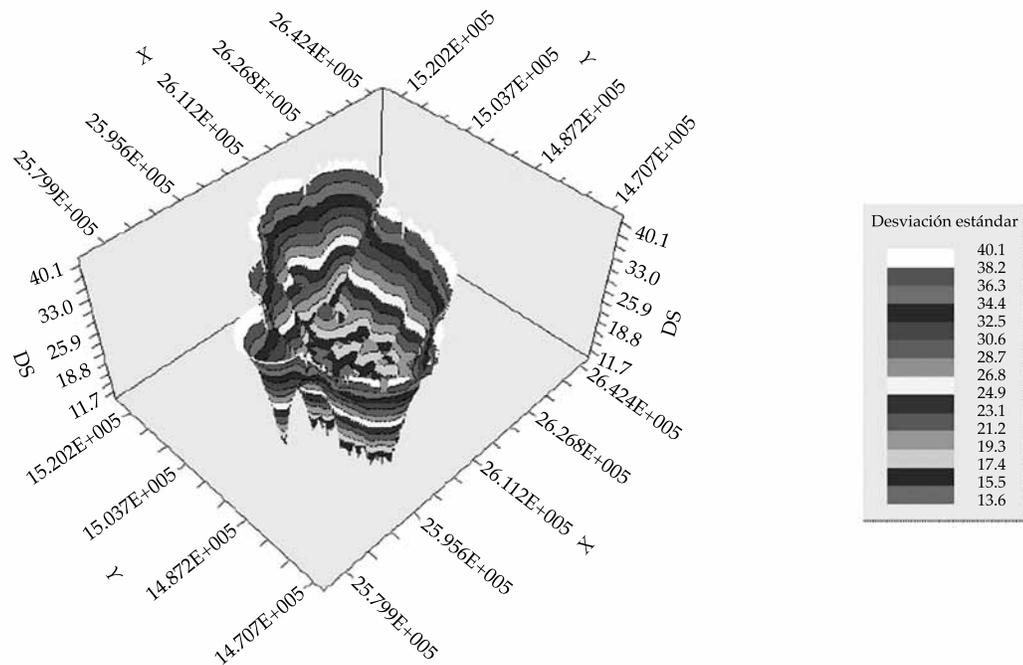


Figura 6. Vista en 3D de las desviaciones estándar del error de estimación de los residuales.

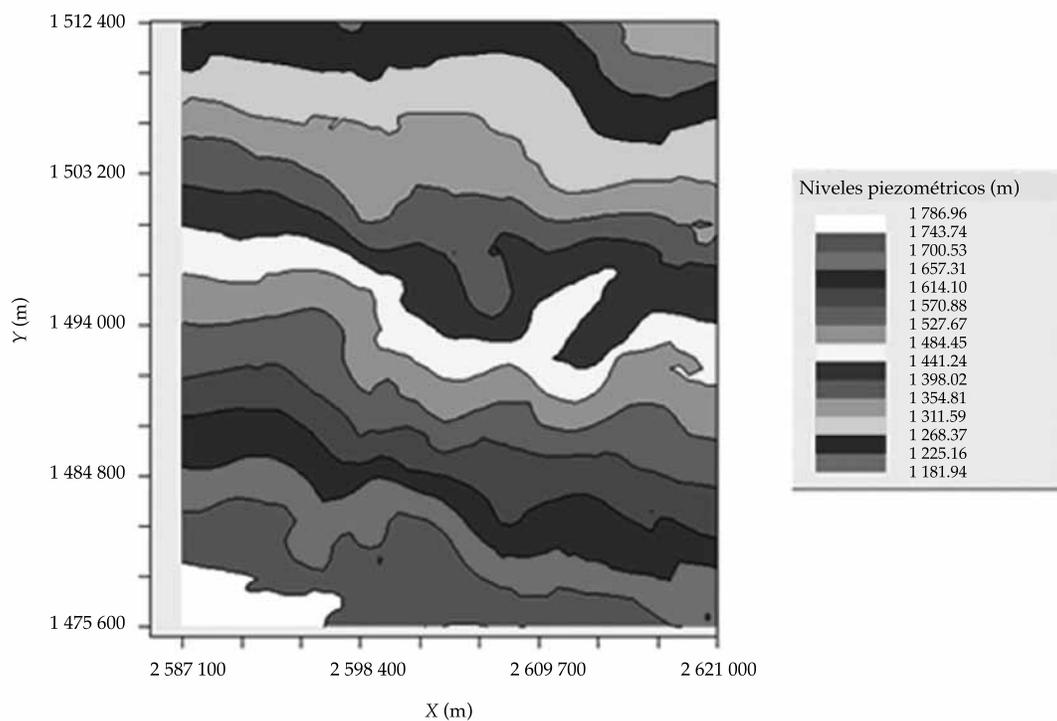


Figura 7. Mapa en 2D de los niveles piezométricos estimados utilizando kriging puntual ordinario.

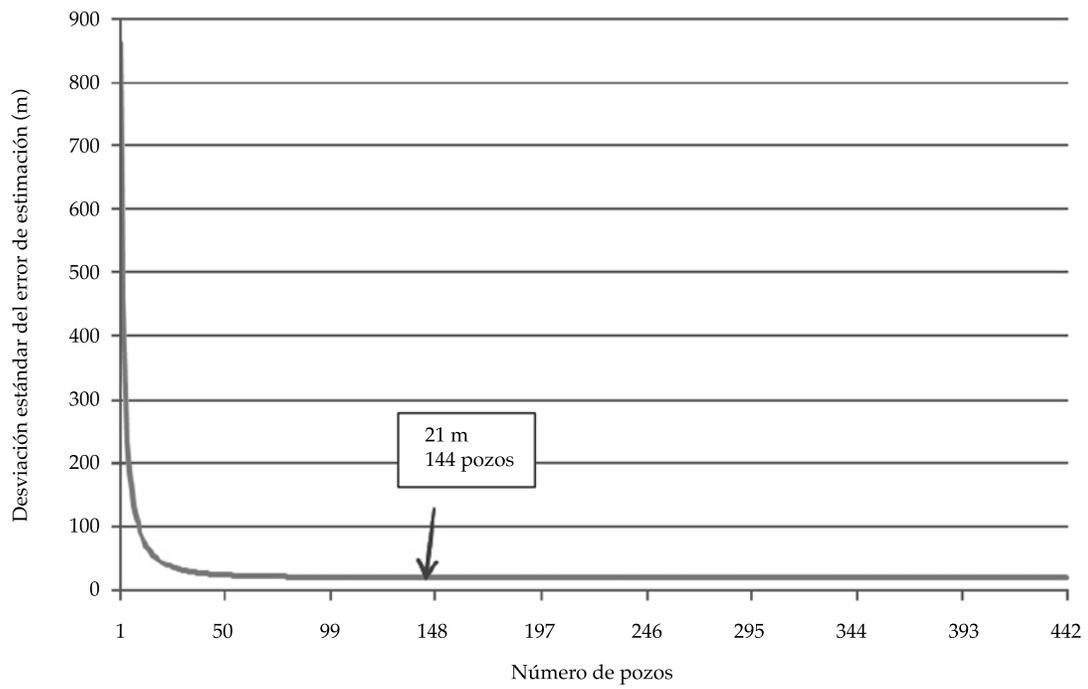


Figura 8. Comportamiento de la desviación estándar del error de estimación por kriging ordinario en bloques.

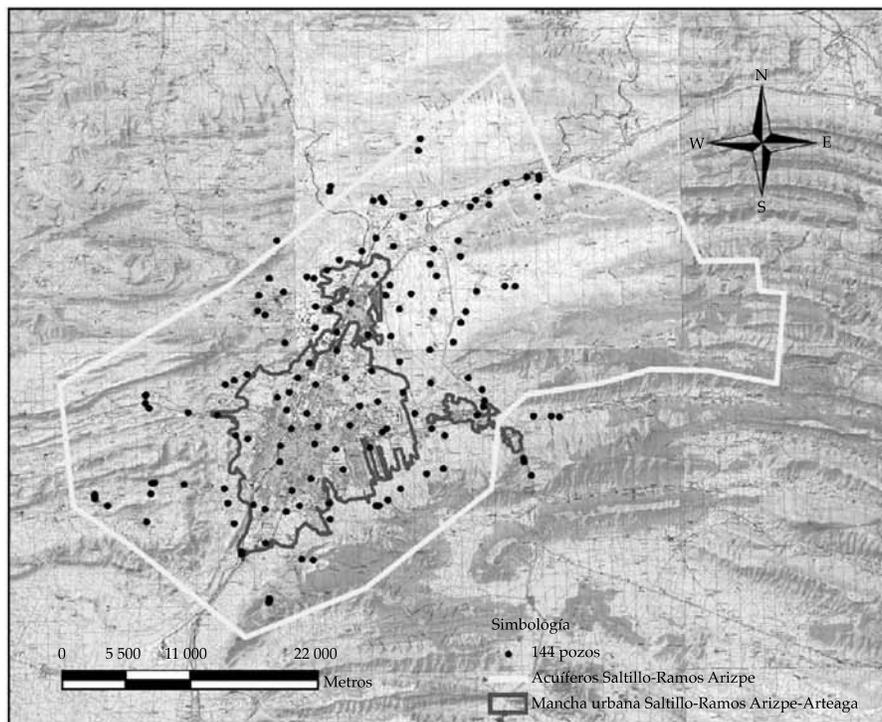


Figura 9. Distribución espacial de los 144 pozos que conforman la red óptima determinada.

Cuadro 3. Comparación de parámetros estadísticos básicos de la desviación estándar del error de estimación de los mejores pozos.

Parámetro estadístico	Toda la red	Red óptima (21 m)	Pozos que ofrecen mayor certidumbre en la estimación					
			50	100	200	300	400	500
Número de pozos	750	3 144	50	100	200	300	400	500
Media	26.9734	27.6769	29.8205	28.1804	27.3404	27.1512	27.0601	26.9889
Varianza	87.9695	73.9683	54.0006	68.5659	78.8585	82.3256	84.8264	86.5767
Desviación estándar	9.3792	8.6004	7.3485	8.2805	8.8802	9.0733	9.2101	9.3047
Coficiente de variación	0.3477	0.3107	0.2464	0.2938	0.3248	0.3342	0.3404	0.3448
Valor mínimo	11.6435	11.7095	11.7134	11.7095	11.7089	11.7089	11.6436	11.6436
Primer cuartil (0.25)	18.4486	20.0717	24.4888	21.1941	19.2468	18.8975	18.6577	18.4234
Mediana	26.1632	29.1670	27.8962	26.4394	26.0529	26.0700	26.0860	26.1025
Tercer cuartil (0.75)	35.1362	35.0940	35.6206	35.2252	35.0515	35.0223	35.0644	35.0713
Valor máximo	43.2565	43.2565	43.2565	43.2565	43.2565	43.2565	43.2565	43.2565
Asimetría	0.1899	0.3257	0.3394	0.3487	0.2971	0.2594	0.2314	0.2144
Kurtosis	1.7406	1.8237	2.1486	1.8832	1.7776	1.7635	1.7464	1.7401

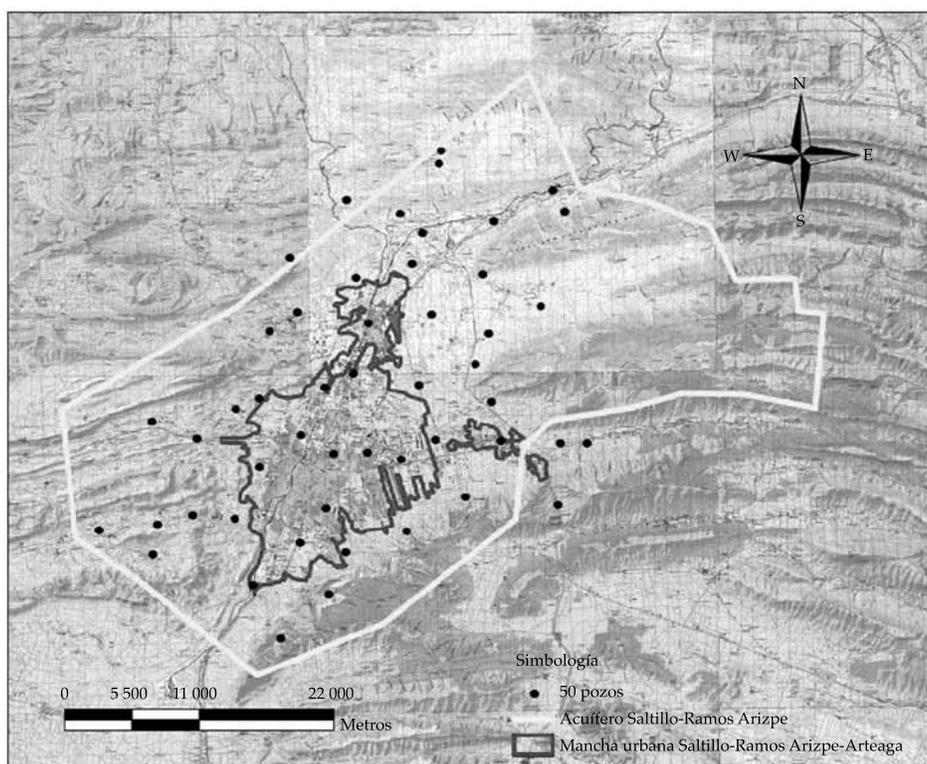


Figura 10. Ubicación de los 50 pozos que ofrecen mayor certidumbre de estimación.

En la estimación espacial de los niveles piezométricos, a los residuales estimados se les adicionó la tendencia espacial calculada. Se

observó que en gran parte del acuífero se realiza la estimación con kriging puntual ordinario, dejando sólo algunas zonas sin cubrir debido

a que no existen pozos vecinos para realizar la estimación; se observan conos de abatimiento en el acuífero.

En la optimización, considerando una tolerancia de la desviación estándar del error de estimación, se observó que existe una convergencia del error cercano a los 50 pozos adicionados por la técnica de las inclusiones sucesivas; sin embargo, a partir de 144 pozos, se observó una variación a nivel de centímetros, y ese fue el criterio elegido para detener la inclusión de más pozos.

Es importante mencionar que los pozos obtenidos de la optimización, prefijando cantidad de pozos, son los que ofrecen los valores más altos del índice de reducción relativa; es decir, son los mejores 50, 100, 200, 300 400 y 500 pozos para el monitoreo del acuífero.

Conclusiones

En el presente trabajo se cumplieron los objetivos planteados, aplicando la metodología definida para obtener el diseño óptimo de la red de monitoreo con fines de manejo del acuífero "Saltillo-Ramos Arizpe", a partir del análisis geoestadístico de los datos de niveles piezométricos, utilizando kriging ordinario para la estimación y la técnica de optimización de inclusiones sucesivas.

El diseño óptimo de la red se realizó considerando los límites del acuífero definidos por la Conagua, sin considerar otros parámetros geohidrológicos de la zona de estudio. Cabe señalar que la metodología utilizada es aplicable para cualquier área y tipo de variable regionalizada en estudio.

Se utilizaron 382 datos de niveles piezométricos de pozos para la determinación del modelo de semivariograma de los 750 que existen, representando un 50.9% del total, por lo que se concluye que es una cantidad suficiente para la modelación de la correlación espacial de dicha variable.

La red óptima determinada consta de 144 pozos, alcanzando una desviación estándar

del error de estimación por kriging ordinario en bloques de 21 metros, lo que representa un 19.2% de los 750 pozos existentes en el acuífero y significaría ahorrar un 80.8% en el costo de monitoreo en pozos.

La optimización, dada una cantidad fija de pozos, permite determinar los mejores pozos para el monitoreo, escogiendo el número de éstos que se consideren convenientes de acuerdo con el recurso que se tenga disponible.

Para futuros trabajos de tal naturaleza en este acuífero, se recomienda subdividir el acuífero considerando su naturaleza hidrogeológica e incorporar los modelos de flujo de agua subterránea que existan.

Agradecimientos

Al Instituto de Geofísica de la Universidad Nacional Autónoma de México, al Instituto Nacional de Estadística y Geografía, Dirección Regional Noreste, y a la Comisión Nacional del Agua, Delegación Saltillo, Coahuila de Zaragoza, por las facilidades brindadas para realizar este estudio.

Referencias

- Akaike, H. (1974). A New Look at Statistical Model Identification. *IEEE Transactions on Automatic Control*, AC-19, 716-722.
- Akaike, H. (1977). An Entropy Maximization Principle. In P. R. Krishnaiah (Ed.). *Applications of Statistic*. Amsterdam: North Holland.
- Armstrong, M., & Delfiner, P. (1980). Towards a More Robust Variogram: Case Study on Coal. Note n_671. Fontainebleau, France: CGMM.
- Banjevic, M., & Switzer, P. (2002). *Optimal Network Designs in Spatial Statistics*. Stanford: Stanford University.
- Carrera, J., Usunoff, E., & Szidarovszky, F. (1984). A Method for Optimal Network Design for Groundwater Management. *Journal of Hydrology*, 73, 147-163.
- Carrera, J., & Szidarovszky, F. (1985). Numerical Comparison of Network Design Algorithms for Regionalized Variables. *Appl. Math. Comput.*, 16, 189-202.
- Chiles, J. P., & Delfiner, P. (1999). *Geostatistics Modeling Spatial Uncertainty* (157-158 pp.). New York: John Wiley & Sons.
- Cressie, N., & Hawkins, D. M. (1980). Robust Estimation of the Variogram. *Math. Geol.*, 12, 115-126.

- Conagua (junio de 2010). *Sección Aguas Nacionales. Sistema Nacional de Información del Agua (Sina)*. Comisión Nacional del Agua. Recuperado de <http://www.cna.gob.mx/>.
- Cousens, R. D., Brown, R. W., Mcbratney, A. B., & Moerkerk, M. (2002). Sampling Strategy is Important for Producing Weed Population Maps: A Case Study Using Kriging. *Weed Science*, 50, 542-546.
- Díaz-Viera, M. (2002). *Notas del Curso: Geoestadística Aplicada* (135 pp.). México, DF: Instituto de Geofísica, UNAM, Instituto de Geofísica y Astronomía, Ministerio de Ciencia, Tecnología y Medio Ambiente de Cuba.
- Díaz-Viera, M., Hernández-Maldonado, V., & Méndez-Venegas, J. (2010). *RGEOESTAD: un programa de código abierto para aplicaciones geoestadísticas basado en R-Project*, México. Recuperado de <http://mmc2.geofisica.unam.mx/gmee/paquetes.html>.
- Díaz-Viera, M., & Barandela, R. (agosto de 1994). *GEOESTAD: un sistema de computación para el desarrollo de aplicaciones geoestadísticas*. Memorias del II Taller Internacional Informática y Geociencias, GEOINFO'94, Ciudad de la Habana, Cuba.
- Díaz-Viera, M. (11-15 noviembre de 1997). *DISRED: una herramienta para el diseño de redes de monitoreo usando un enfoque geoestadístico*. Primer Congreso Nacional de Aguas Subterráneas, Mérida, México.
- Díaz-Viera, M. (24-27 de marzo de 1998). El diseño de redes óptimas de monitoreo en aguas subterráneas usando un enfoque geoestadístico. *Memorias del IV Taller Internacional Informática y Geociencias* (pp. 66-69). GEOINFO'98, Ciudad de la Habana, Cuba.
- Dixon, W., Smyth, G. K., & Chiswell, B. (March, 1999). Optimized Selection of River Sampling Dites. *Water Research*, 33(4), 971-978.
- Journel, A. G., & Huijbregts, Ch. J. (1978). *Mining Geostatistics* (590 pp.). London: Academic Press.
- INEGI (junio de 2010). *Apartado Geografía*. Instituto Nacional de Estadística y Geografía. Recuperado de <http://www.inegi.org.mx/>.
- Isaaks, E. H., & Srivastava, R. M. (1989). *An Introduction to Applied Geostatistics* (278-279 pp.). New York/ Oxford: Oxford University Press.
- Leick, A. (1990). *GPS Satellite Surveying* (pp. 284-301). Orono, USA: University of Maine.
- Lloyd, C. D., & Atkinson, P. M. *Designing optimal sampling configurations with ordinary and indicator kriging*. In Proceedings of the 4th International Conference on GeoComputation. 1999.
- Moral-García, F. A. (2003). *La representación gráfica de las variables regionalizadas. geoestadística lineal* (pp. 71-74). Cáceres, España: Universidad de Extremadura.
- Olea, R. A. (1999). *Geostatistics for Engineers and Earth Scientists* (pp. 7-16) Lawrence, USA: Kansas Geological Survey, The University of Kansas.
- Omre, H. (1984). *The Variogram and its Estimation. Geostatistics for Natural Resources Characterization* (107-125 pp.). Vol. 1. Verly et al. (Eds.). Hingham: NATO ASI Series. Reidel.
- Prakash, M. R., & Singh, V. S. (April, 2000). Network Design for Groundwater Monitoring – A Case Study. *Environmental Geology*, 39(6), 628-632.
- Samper, C. F. J., & Carrera, R. J. (1990). *Geoestadística. Aplicaciones a la Hidrogeología Subterránea* (pp. 409-427). Barcelona: Centro Internacional de Métodos Numéricos en Ingeniería.
- Wang, X. J., & Qi, F. (December 11, 1998). The Effects of Sampling Design on Spatial Structure Analysis of Contaminated Soil. *The Science of the Total Environment*, 224(1-3), 29-41.

Dirección institucional de los autores

Dr. Martín A. Díaz-Viera

Investigador
 Instituto Mexicano del Petróleo
 Programa de Recuperación de Hidrocarburos
 Eje Central Norte Lázaro Cárdenas 152
 Col. San Bartolo Atepehuacan
 Delegación Gustavo A. Madero
 Apdo. Postal 14-805
 07730 México, D.F., MÉXICO
 Teléfonos: +52 (55) 9175 6473 y 9175 6993
 Fax: +52 (55) 9175 6993
 mdiazv@imp.mx
 mdiazv64@yahoo.com.mx

M. en C. Félix Canul-Pech

Instituto Nacional de Estadística y Geografía
 Dirección Regional Noreste
 Pino Suárez No. 602 Sur, Col. Centro
 64000 Monterrey, Nuevo León, MÉXICO
 Teléfonos: +52 (81) 8152 8250 y 8152 8249
 felix.canul@inegi.org.mx
 fcanul2003@yahoo.com.mx