



**UNIVERSIDAD AUTÓNOMA DE ZACATECAS
"FRANCISCO GARCÍA SALINAS"
UNIDAD ACADÉMICA DE INGENIERÍA ELÉCTRICA
DOCTORADO EN CIENCIAS DE LA INGENIERÍA
(DOCII)**



DESARROLLO DE METODOLOGÍA Y SOFTWARE PARA EL DISEÑO DE REDES DE MONITOREO DE LA CALIDAD DEL AGUA SUBTERRÁNEA BASADA EN GEOESTADÍSTICA Y SIG

TESIS QUE PARA OPTAR POR EL GRADO DE: DOCTOR EN CIENCIAS DE LA INGENIERÍA

Presenta:
MC. Jorge Aceves De Alba

Director:
Dr. Hugo Enrique Júnez Ferreira

Co-director:
Dr. Julián González Trinidad

Zacatecas, Zac., agosto de 2019

AGRADECIMIENTOS

A Yoli, por el apoyo incondicional durante todo el tiempo, quien ha sido clave para culminar esta etapa en mi vida.

A Lía, mi hija y motor que me impulsa a seguir trabajando día a día.

A mi hermano Enrique, por la disponibilidad, rapidez y exactitud con la que apoyó las revisiones del idioma Inglés en el artículo.

A mi familia, por su soporte en todas las formas que tuvieron en sus manos y que no dudaron en hacerlo.

Al Doctor Hugo, a quien he tenido el gusto de conocerlo, ha sido un guía y compañero en este camino, ha sabido alentarme en los momentos de incertidumbre.

Al Doctor Toño quien, a su modo sencillo, franco y directo, me ha permitido conocer otra perspectiva de las cosas.

A los Doctores Julián, Zavala y Bautista por sus contribuciones para terminar este trabajo.

A todos aquellos que mostraron interés en saber como iba en este proceso, preguntaron una o más veces como iba, cuanto me faltaba y que sabiéndolo o no, me alentaron a terminar.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca otorgada durante el periodo de agosto de 2014 a febrero de 2016. También quisiera agradecer al Programa para el Desempeño Profesional Docente (PRODEP) por la beca otorgada de agosto de 2016 a julio de 2017. Este trabajo fue parcialmente financiado por CONACYT a través de la subvención 294537 y CONACYT-BMBF (Ministerio Federal de Educación e Investigación, Alemania) a través de la subvención 0207586

RESUMEN

La gestión del agua subterránea para su distribución y uso es un aspecto que debe ser atendido día con día. La extracción del agua mediante pozos cada vez más profundos representa un potencial problema, ya que puede existir un cambio en su calidad, lo cual requiere de un manejo dinámico para atender sus implicaciones. En esta situación se encuentran una importante cantidad de zonas que dependen en gran medida del agua subterránea para su abastecimiento por lo que se hace necesaria la búsqueda de estrategias para atenuar efectos negativos.

Como parte de las acciones necesarias para llevar a cabo la gestión del agua subterránea, se hace necesario establecer redes de calidad del agua subterránea, cuyo principal objetivo deberá ser la obtención de la mayor cantidad de información posible con la menor cantidad de puntos de muestreo, estableciendo cuáles serán los parámetros fisicoquímicos que se desean analizar.

Para el diseño de una red de calidad del agua subterránea, el considerar a la ubicación geográfica como parte de los atributos de los datos que se obtienen de los sitios donde se realiza el análisis de muestras de agua subterránea es primordial, ya que permite estudiar, relacionar y evaluar la información espacial por medio de la geoestadística y lograr representaciones que permitan un mejor entendimiento de lo analizado.

El análisis de las metodologías existentes para el diseño de redes de monitoreo de agua subterránea, permite la identificación de alguna que potencialmente podría aplicarse al caso de estudio correspondiente.

En el presente trabajo se propone e implementa una metodología mediante el desarrollo de un software amigable que además de considerar la varianza del error en la estimación total obtenida con la información de los sitios que la conforman, incorpora una función suplementaria que prioriza las posiciones perimetrales de puntos de muestreo. Este desarrollo permitirá a las personas encargadas de la toma de decisiones contar con una herramienta de fácil acceso y uso sencillo en la gestión del agua subterránea, recurso estratégico y complejo de manejar.

Otro punto a favor del desarrollo presentado es que las redes de monitoreo diseñadas o evaluadas permitirán considerar la dinámica de la calidad del agua subterránea, mediante un manejo ágil de la información facilitando la constante evaluación de alternativas de redes de calidad de agua subterránea que permitan tomar una mejor decisión teniendo referencias para un comparativo.

ABSTRACT

Groundwater management is an issue that must be addressed daily for its distribution and use. The sustained increasing depth of pumping wells is a potential problem, since significant changes in water quality can be induced, which requires a dynamic management over time to consider its implications. In this situation, there are significant areas that depend strongly on groundwater for their supply, so it is necessary to find strategies to mitigate negative effects.

As part of the actions to carry out groundwater management, it is necessary to establish groundwater quality networks, whose main objective should be to obtain as much information as possible with the least amount of sampling points, establishing the physicochemical parameters to be analyzed.

For the design of a groundwater quality network, it is crucial considering geographical location as part of the attributes obtained from the sites where the analysis of groundwater samples is done, since it allows studying, relate and evaluate spatial information through geostatistics and generate representations that permit a better understanding of the analyzed data.

The analysis of existing methodologies for the design of groundwater monitoring networks, allows the identification of one that could potentially be applied to the corresponding case study.

In this paper, a methodology is proposed and implemented through the development of friendly software that additionally considers the total estimate error variance obtained with the information of the monitoring sites that comprise it, furthermore, it incorporates a supplementary function that priorities peripheral sampling points. This development will allow decision makers to have a tool that is easily accessible and simple to use in groundwater management, a strategic and complex to manage resource.

Another aspect in favor of the presented development is that the designed or evaluated monitoring networks will allow considering the dynamics of groundwater quality, through quick information handling, simplifying the constant evaluation of

alternative groundwater quality networks in order to make a better decision by having comparison references.

Contenido

CAPÍTULO 1 INTRODUCCIÓN	1
Justificación	3
Hipótesis	4
Objetivos	4
Objetivo general	4
Objetivos específicos.....	4
1.1 Estado del arte	4
CAPÍTULO 2 MATERIALES Y MÉTODOS	12
2.1 Descripción de zona de estudio	12
2.2 Metodología	19
CAPÍTULO 3 SOFTWARE	37
3.1 Diseño del software	37
3.2 Desarrollo e implementación	41
CAPÍTULO 4 RESULTADOS Y DISCUSIONES	49
4.1 Uso de la metodología y herramienta desarrollada	49
4.2 Comparativo de metodologías.....	49
CAPÍTULO 5 CONCLUSIONES	57
Bibliografía	60

ÍNDICE DE FIGURAS

Figura 1.- Mapa de ubicación de la zona de estudio. Fuente: Mapa de La Región del Bajío modificado de Suárez-mota and Villase (2015); Acuífero Irapuato-Valle de National Water Information System (2019); División política de países de Tapiquén (2019).	13
Figura 2.- Ubicación de los puntos de muestreo	16
Figura 3.- Análisis de arsénico a) número de muestras por grupos y el límite de referencia de la OMS, b) correlación entre la concentración y la profundidad del pozo.	18
Figura 4.- Correlación entre Cloruro y la conductividad eléctrica.	19
Figura 5.- Diagrama de flujo y pseudocódigo de la generación del MEP.	24
Figura 6.- Caso 1 para la MEP	25
Figura 7.- Caso 2 para la MEP	25
Figura 8.- Caso 3 para la MEP	25
Figura 9.- Caso 4 para la MEP	26
Figura 10.- Malla de estimación personalizada (MEP).	27
Figura 11.- Alcance de los rangos de puntos de muestreo en los nodos de la MEP para conductividad eléctrica y arsénico.....	28
Figura 12.- Malla de estimación recortada (MER)	30
Figura 13.- Conteo del número de puntos de muestreo dentro de rango para cada nodo de la MER por parámetro	31
Figura 14.- Obtención de los valores necesarios para obtener la FS para la conductividad eléctrica	32
Figura 15.- Obtención de los valores necesarios para obtener la FS para el arsénico.....	32
Figura 16.- Ejemplo de malla de estimación recortada (MER) representando los valores en los nodos para la obtención de la FS.....	34
Figura 17.- Pantalla principal del software.....	39
Figura 18.- Pantalla de la malla de estimación y ponderación.....	40
Figura 19.- Pantalla del diseño de la red de monitoreo y comparativa de la seleccionada	41
Figura 20.- Apertura y despliegue de la tabla de datos	42
Figura 21.- Identificación de errores en la tabla de datos.....	42
Figura 22.- Actualización de la información en la tabla de datos.....	42
Figura 23.- Representación de la información mediante el histograma y mapa de ubicación para el análisis exploratorio.....	43
Figura 24.- Transformación de los datos de un parámetro	43
Figura 25.- Representación de los datos mediante el variograma y la validación cruzada para el análisis estructural	44
Figura 26.- Actualización de los datos después del cambio de una transformación y/o cambio de modelo matemático.....	45
Figura 27.- Representación de la MEP y la MER	46

Figura 28.- Representación de la Malla de estimación proveniente de un archivo y de la MER correspondiente	46
Figura 29.- Representación de la red de monitoreo de agua subterránea y la gráfica comparativa de la seleccionada contra la peor y la óptima	48
Figura 30.- Gráfica comparativa de los diferentes casos tomando los resultados obtenidos de PCT (%) vs número de puntos de muestreo	53
Figura 31.- Redes de monitoreo de agua subterránea y su orden de prioridad para los primeros 10, 30 y 50 puntos de muestreo para los casos: a) "0-100", b) "0.001-99.999", c) "0.005-99.995", d) "0.01-99.99", e) "50-50" y f) "100-0"	56

ÍNDICE DE TABLAS

Tabla 1.- Resumen de metodologías revisadas.....	9
Tabla 2.- Resultados obtenidos para el cálculo de FS para diferentes puntos de muestreo y parámetros	33
Tabla 3.- Estructura de la necesaria en la tabla para herramienta.....	41
Tabla 4.- Estructura del archivo de la malla de estimación	44

CAPÍTULO 1 INTRODUCCIÓN

El agua subterránea es un recurso vital para uso humano. De manera global un 38% de las zonas de riego utilizan el agua subterránea y casi el 50% de la población mundial la utiliza para su consumo (WWAP, 2016). Estos datos muestran la dependencia que existe del agua subterránea, la cual se incrementa en zonas áridas, donde las fuentes de agua superficial son muy escasas o nulas. En 2003 se publicó que el agua potable es una necesidad humana y por tanto un derecho (WHO, 2003) y en febrero de 2012 México lo incorporó a su Constitución (DOF, 2012).

El estudio del agua subterránea, por tanto, es un tema prioritario a nivel global, lo cual se puede observar en la gran cantidad de investigaciones desde diferentes enfoques tales como aquellos que tratan de identificar las fuentes de recarga y los porcentajes de contribución para cada caso (Batelaan & Smedt, 2007; Wang et al., 2016), estudios piezométricos (Pearson, Berg, & Macdonald, 2011; Yang, Cao, Liu, & Yang, 2008; Zhou, Dong, Liu, & Li, 2013), de identificación y/o seguimiento de fuentes de contaminación (Chen, Smith, & Beckie, 2012; Yakirevich et al., 2013) como algunos de los ejemplos más comunes, todas estas con la finalidad de conocer diferentes aspectos del agua subterránea.

Estos factores son importantes a considerar debido a la gran relevancia antes mencionada del agua subterránea y por tanto su manejo adecuado requiere entre otras cosas, de un entendimiento profundo de la distribución espacial y su calidad. Para mejorar el entendimiento del agua subterránea, es necesario obtener muestras representativas de la zona donde se desea estudiar, para ello se realiza en primera instancia un censo de aprovechamientos que permite identificar la distribución espacial, así como sus condiciones, para posteriormente de manera deseable, planear campañas de monitoreo de aquellos aprovechamientos que por sus condiciones (ubicación, equipamiento, disponibilidad), resultan los más convenientes. Una selección adecuada de los aprovechamientos del área de interés permite la generación de lo que es conocido como redes de monitoreo, que a partir de los aprovechamientos seleccionados permite estimar datos en el resto de las ubicaciones (Baalousha, 2010). Este tipo de redes son de gran relevancia

ya que el costo de monitoreo se puede reducir en más del 50% en dependencia de las condiciones de la zona que se está analizando.

La delimitación de la zona de estudio está en dependencia de los objetivos del mismo. Para propósitos administrativos, en México se tiene una delimitación por acuíferos en los que de manera deseable deberían de utilizarse las redes de monitoreo de agua subterránea, sin embargo, no siempre se cuenta con personal capacitado que pueda diseñar este tipo de configuraciones. Actualmente se cuenta con diferentes metodologías para la generación de redes de monitoreo desde diferentes enfoques dependiendo en gran medida de los objetivos de la red, es por ello que el establecer su propósito debe de ser el primer paso a seguir (Om Prakash, 2014).

El objetivo principal en el Valle de Toluca (Esquivel, Morales, & Esteller, 2015) fue determinar las áreas con las mayores fluctuaciones en los niveles de agua y para ello se tomaron en cuenta datos como la carga hidráulica, disminución de niveles, los cambios en periodos específicos en el tiempo y la densidad de aprovechamientos hídricos todo ello integrado dentro de un Sistema de Información Geográfica (SIG), sin embargo la dependencia de la conversión de los datos a capas en el SIG requiere de cierto nivel de conocimiento de estas herramientas. Otro ejemplo en la aplicación de las redes de monitoreo apoyado en los SIG mediante un proceso semiautomático se desarrolló para el diseño de mallas de diferentes densidades (Preziosi, Petrangeli, & Giuliano, 2013). También se han utilizado las redes de monitoreo para estudiar los procesos de decremento de los niveles de agua en Beijín, China, esto mediante el establecimiento de criterios para la delimitación de los diferentes tipos de zonas (Zhou et al., 2013). Uno de los problemas que se presenta en la generación de redes de monitoreo es determinar el espaciamiento ideal que debe de tener una malla de estimación y las consecuencias que se tienen en los casos extremos, lo cual se analizó en Texas (Uddameri & Andruss, 2013).

Existen algunas otras metodologías que pueden utilizar los datos de un censo de aprovechamientos para el diseño de la malla de estimación mediante el uso de la geoestadística, el filtro de Kalman y optimización heurística (Júnez-Ferreira,

Herrera, Gonzalez-Hita, Cardona, & Mora-Rodriguez, 2016a), sin embargo, requieren del conocimiento de la geoestadística, además de todo el proceso para poder aplicarlo.

Justificación

En México existe información con diferentes fuentes, una parte importante la generan la CONAGUA y los órganos operadores locales para cada entidad, sin embargo, en muchos casos no existe un análisis posterior que permita determinar los puntos que generen información redundante, así como la identificación de las zonas en donde se requiere una mayor cantidad de datos para reducir el error en la estimación cuando se utilizan modelos matemáticos en la interpolación. Además de ello en algunos casos en que ya existen redes de monitoreo de calidad de agua subterránea, no hay un análisis periódico de la información que permita determinar si prevalecen las condiciones de manera que la red siga teniendo validez pasado el tiempo, considerando que la calidad del agua a lo largo del tiempo puede presentar cambios, por lo cual es recomendable determinar la periodicidad de la revisión de la red, en función de que tan alta es la variabilidad de las condiciones de la zona que se está estudiando.

Los objetivos de la red de monitoreo determinarán los parámetros que se tomarán en cuenta, así como los pesos que se darán a cada uno de dichos parámetros, lo cual será decisivo para los resultados que se obtendrán, esto no limita el análisis de la información desde diferentes perspectivas en función de los objetivos del estudio y los diferentes resultados obtenidos, de manera que se pueden llegar a considerar diferentes redes de monitoreo con sus respectivos puntos de monitoreo y prioridades. Esta perspectiva implica que el análisis de los datos se pudiera hacer de manera dinámica, con usuarios con diferentes niveles de conocimiento mediante el uso de herramientas geoestadísticas y que requieren de un proceso automatizado para poder generar una red de monitoreo a partir de su información, así como el rediseño de redes en el caso de que exista una previa.

Hipótesis

Es posible combinar una metodología existente y una función suplementaria, para establecer diferentes diseños de redes de monitoreo de calidad de agua subterránea con sus respectivas prioridades y distribuciones espaciales, sin aumentar más del 0.5% el error en la estimación, aplicado en el desarrollo de un nuevo software

Objetivos

Objetivo general

Elaborar una herramienta de software con base en la adecuación de una metodología existente para la generación de redes de monitoreo de calidad de agua subterránea, que permita determinar el orden de importancia de los puntos de monitoreo, considere las zonas de baja densidad de información y automatice el proceso de generación de la red dentro de un acuífero.

Objetivos específicos

- a) Modificar una metodología para que permita orientar el diseño de la red hacia una distribución espacial específica de los pozos.
- b) Generar un software que permita la implementación de todo el proceso de la metodología.
- c) Aplicar la herramienta utilizando los datos de monitoreo de un acuífero y comparar los resultados.

1.1 Estado del arte

Existen algunas clasificaciones que se han desarrollado para agrupar el diseño de redes de monitoreo para agua subterránea. Júnez-Ferreira et al. (2016a) y Loaiciga, Member, Member, & Rouhani (1992) sugieren dos grupos principales: el hidrogeológico y el estadístico. Herrera & Pinder (2005) sugieren incluir un tercero que combina ambos al cual llaman de modelación. Zhang, Pinder, & Herrera (2005) llaman aproximación cuantitativa al grupo estadístico, consideran los subgrupos aproximación de simulaciones estocásticas, aproximación basada en la varianza y aproximación basada en métodos.

Andricevic (1990) clasifica como optimización y simulación, en esta última considera los análisis espacio-temporales. Singh & Katpatal (2017) clasifica en geoestadística, teoría de entropía, *simulated annealing* (recocido simulado) y análisis de componentes principales. Para el análisis de las diferentes propuestas se consideraron la aproximación hidrogeológica, estadística y la de modelación (Herrera & Pinder, 2005).

Derivado del análisis de diferentes metodologías es posible encontrar que algunas de ellas son una combinación de algunas ya desarrolladas. Considerando el grupo del análisis hidrogeológico, Zhou et al. (2013), emplean la información de zonas hidrogeológicas, saturación, recarga e influencia de algunos regímenes de agua subterránea combinándolos como diferentes capas de información utilizando Sistemas de Información Geográfica (SIG), analizando los resultados y con ellos definen las nuevas zonas de instalación de pozos de monitoreo. Singh & Katpatal (2017), evalúan el impacto de factores hidrogeológicos para la selección y localización de pozos de observación mediante niveles de agua, procesos de descarga, recarga y almacenamiento. Además de esto se tomó en cuenta el uso de suelo, vegetación y precipitación pluvial por medio del análisis de imágenes de satélite, integradas y analizadas por medio de SIG.

En la parte del grupo del análisis estadístico, se reconoce la importancia de considerar la complejidad hidrogeológica mediante el cálculo del error en la estimación de la varianza y la variabilidad de los sistemas de flujo de agua subterránea en espacio-tiempo utilizando el filtro de Kalman para optimizar el proceso (Andricevic, 1990). Se utiliza una simulación estocástica de flujo de agua subterránea y un algoritmo de simulación para obtener las varianzas de las estimaciones de los niveles. Utiliza el concepto de umbral de información como una restricción de confiabilidad para tener en cuenta diferentes áreas de sensibilidad. Kim et al. (2007), evalúan redes de monitoreo existentes en Corea para determinar la frecuencia óptima de monitoreo de agua subterránea mediante un análisis de modelos de series armónicas para encontrar las fluctuaciones periódicas. La estimación de la varianza fue utilizada para determinar la frecuencia de muestreo. Baalousha (2010), combina la metodología DRASTIC con el análisis

de la interpolación de los valores de nitrato mediante el uso de SIG en Nueva Zelanda resultando un mapa que combina varianza y vulnerabilidad. Para la interpolación de los valores de nitrato fue utilizado un modelo de semivariograma esférico tomando como criterios la raíz del error cuadrático medio y el criterio de información de Akaike (RMSE y AIC respectivamente por sus siglas en inglés). H. E. Júnez-Ferreira & Herrera (2013), utiliza la covarianza espacio-temporal y el filtro de Kalman estático para el diseño óptimo espacio-tiempo de redes de monitoreo de la carga hidráulica en Querétaro. Después de un análisis estadístico de manera conjunta de varios eventos de monitoreo fue posible calcular que el 57% de los datos obtenidos en los diferentes muestreos resultaban redundantes en cuanto a la información que proporcionaban. El resultado óptimo obtenido está en función del error en la estimación de la varianza. Uddameri & Andruss (2014), presentan una metodología multicriterio para la creación de una red de monitoreo en Texas. Se tomaron en cuenta seis variables con la posibilidad de la asignación de pesos a cada una de ellas. Utilizando SIG se llevó a cabo la interpolación de cada una de estas variables y fue posible analizar el error en la predicción. La intersección de cada una de estas variables también permitió definir las áreas susceptibles a ser consideradas para la ubicación de los puntos de monitoreo. Sizirici & Tansel (2015), realizan un análisis histórico en Florida para determinar la redundancia de datos basados en la dirección de flujo del agua subterránea, la profundidad de los pozos y la concentración de contaminantes. Se evaluó el factor de retardo de los compuestos orgánicos y las estadísticas de la información histórica del monitoreo del agua subterránea para determinar la mejor ubicación de los puntos de muestreo, así como su frecuencia. H. Júnez-Ferreira et al. (2016b) combinaron la metodología DRASTIC para determinar el índice de vulnerabilidad del acuífero y el filtro de Kalman para su optimización en Zacatecas. Se definieron diferentes pesos a las variables consideradas en el cálculo del índice como primer estrategia y un aumento de nodos en la malla de estimación en áreas de importancia como una segunda estrategia que permitió evaluar el desempeño del diseño de la red de monitoreo de calidad de agua subterránea. Alizadeh & Mahjouri (2017) exploran el uso del modelo Bayesiano de Máxima Entropía para revisar una red de monitoreo

de calidad de agua subterránea en Irán. Se utilizó información espacio-temporal para el análisis, además de considerar la profundidad del agua subterránea, precipitación, la conductividad hidráulica y la pendiente del terreno natural. El análisis es realizado mediante un modelo Bayesiano y se emplea la estimación del error en la varianza para la simulación y optimización.

En la parte del grupo de modelación, Zhang et al. (2005), utiliza un algoritmo genético combinado con el filtro de Kalman y un modelo estocástico para la simulación en espacio-tiempo de un contaminante en el agua subterránea mediante el análisis de una red de monitoreo de agua subterránea de largo plazo. La simulación del flujo y transporte del agua subterránea se realizó utilizando datos aleatorios de conductividad hidráulica obtenidos del Latin Hypercube Sampling (LHS) y el coeficiente de variación fue reducido mediante el uso de algoritmos genéticos que permitió maximizar el bombeo reduciendo los costos de la remediación. El filtro de Kalman fue utilizado para la actualización de la matriz de covarianza espacio-temporal. Nobre et al. (2007), obtienen una solución óptima mediante el análisis dinámico de los datos utilizando una versión modificada de la metodología DRASCTIC combinada con una metodología de fuzzy hierarchy para evaluar el índice potencial de una fuente de contaminación en Brasil. Utilizan SIG, MODFLOW y MODPATH como herramientas para la modelación, cálculo de la vulnerabilidad del agua subterránea y un mapa de riesgo. Kollat et al. (2011) analizan una red de monitoreo de agua subterránea multi-objetivo combinando modelación de flujo y transporte de contaminantes, y el filtro de Kalman en conjunto con bias-aware, una búsqueda multi-objetivo (optimización Bayesiana jerárquica) y un soporte de decisiones basado en el análisis visual. Esta metodología es potente y permite el uso eficiente en redes de monitoreo y predicción de largo plazo. Kim (2015), analiza la distribución óptima de la red de monitoreo de pozos mediante MODFLOW para la estimación de las direcciones de flujo de agua subterránea que se ven afectadas por la construcción de barreras. Esquivel et al. (2015), utiliza seis factores y dos restricciones mediante el uso de SIG y un análisis multicriterio para determinar las áreas óptimas para el diseño de una red de monitoreo en el Estado de México. Mediante el SIG se pudo intersectar

la información y hacer una combinación lineal ponderada como un análisis multicriterio para la evaluación y la generación del mapa final. Luo et al. (2016), utiliza un método probabilístico consistente en un algoritmo genético de Pareto bajo incertidumbre para la optimización en una red de monitoreo de agua subterránea de largo plazo multi-objetivo en Indiana. Los objetivos para la optimización consistieron en la minimización de costos, el error en la estimación de la masa, el primer momento del error en la estimación y el segundo momento del error en la estimación. El método probabilístico se utilizó para la modelación, MODFLOW y MT3DMS se utilizaron para el flujo y transporte. La evaluación de la efectividad de la metodología fue hecha mediante Monte Carlo. Hosseini & Kerachian (2017b), proponen una metodología combinada para el rediseño de una red de monitoreo de agua subterránea en Irán utilizando redes neuronales artificiales, un modelo kriging espacio-temporal para la simulación de las fluctuaciones de los niveles de agua, además de la determinación del tamaño y la distribución de las áreas de alta prioridad mediante la generación de mallas hexagonales de diferentes tamaños y la creación de polígonos de Thiessen, para áreas de menor prioridad. Hosseini & Kerachian (2017a), utiliza también una malla hexagonal y añade el método de máxima entropía bayesiana para el análisis espacio-temporal en el diseño de redes de monitoreo de agua subterránea. Jiang et al. (2018), utiliza el método del filtro de Kalman y el alpha-cut para la identificación de las fuentes de contaminación en agua subterránea y la caracterización de la morfología de la pluma. El LHS fue utilizado para la generación de información de conductividad hidráulica y un modelado Monte Carlo para la simulación del flujo y transporte considerando las fuentes potenciales de contaminación, la asignación de diferentes pesos y el cálculo de la matriz de covarianza para la actualización de la pluma contaminante mediante el filtro de Kalman. En la **Tabla 1** se puede observar un resumen de las metodologías revisadas.

Tabla 1.- Resumen de metodologías revisadas

Tipo	Autor	Parámetros	Metodología
Análisis hidrogeológico	(Zhou et al., 2013)	Geología Geomorfología Profundidad del agua Propiedades del suelo Uso de suelo Precipitación Cuerpos de agua Zonas de pozos	Zonificación mediante SIG
Análisis hidrogeológico	(Singh & Katpatal, 2017)	Niveles de agua Descarga Recarga Almacenamiento Uso de suelo Vegetación	Analytical hierarchical process (AHP)- geoestadístico Fussy Logic Análisis mediante SIG
Análisis estadístico	(Andricevic, 1990)	Carga hidráulica	Filtro de Kalman Statistical monitoring power
Análisis estadístico	(Kim, Lee, Lee, & Yi, 2007)	Nivel del agua	Modelo autoregresivo Series armónicas
Análisis estadístico	(Baalousha, 2010)	* Nivel del agua (D) * Recarga (R) * Medio del acuífero (A) * Medio del suelo (S) * Topografía (T) * Influencia de la zona vadosa (I) * Conductividad hidráulica (C) Nitratos	DRASTIC combinado con geoestadística
Análisis estadístico	(H. E. Júnez-Ferreira & Herrera, 2013)	Carga hidráulica	Estimación espacio temporal de la carga hidráulica con geoestadística y

			Filtro de Kalman
Análisis estadístico	(Uddameri & Andruss, 2014)	Producción Estimación de recarga Distancia de pozos a los límites del distrito Distancia a ríos Distancia a corrientes Variabilidad	Monitoring priority index (MPI) Geoestadística SIG
Análisis estadístico	(Sizirici & Tansel, 2015)	53 reduciéndolo a 13	Análisis de dirección, carga hidráulica y ubicación
Análisis estadístico	(H. Júnéz-Ferreira, González, Reyes, & Herrera, 2016b)	Arsénico Cloruros Conductividad eléctrica Fluoruros Magnesio Nitratos Sodio Temperatura	Filtro de Kalman Método de optimización heurística Zonas prioritarias
Análisis estadístico	(Alizadeh & Mahjouri, 2017)	Nivel estático Precipitación Conductividad hidráulica Pendiente Nitratos Sólidos disueltos totales	Teoría bayesiana de la máxima entropía
Modelación	(Zhang et al., 2005)	Uso de suelo Cobertura de suelo Conductividad hidráulica	Algoritmo genético Filtro de Kalman Simulación estocástica
Modelación	(Nobre, Rotunno Filho, Mansur, Nobre, & Cosenza, 2007)	Nivel estático Recarga Medio del acuífero Suelo/uso de suelo y cobertura de suelo Topografía Conductividad hidráulica	Mapa de vulnerabilidad Fuzzy Hierarchy Model Modelo numérico

Modelación	(Kollat, Reed, & Maxwell, 2011)	Gradiente hidráulico Medio Trazador	Modelación de flujo y transporte Filtro de Kalman
Modelación	(Kim, 2015)	Topografía Hidrogeología Dirección de flujo	Modelo numérico
Modelación	(Esquivel et al., 2015)	Taza de decremento del nivel de agua Gradiente hidráulico Densidad de pozos Grietas	Analytical hierarchical process (AHP)
Modelación	(Luo, Wu, Yang, Qian, & Wu, 2016)	Porosidad Grosor del acuífero Dispersividad longitudinal Relación de dispersividad Flujo constante Correlación de conductividad hidráulica	Pareto genetic algorithm (PGA) MODFLOW
Modelación	(Hosseini & Kerachian, 2017b)	Nivel estático	Red neuronal artificial Modelo S-T Kriging
Modelación	(Hosseini & Kerachian, 2017a)	Nivel estático	Teoría bayesiana de la máxima entropía MODFLOW
Modelación	(Jiang, Fan, Xia, Li, & Zhang, 2018)	Conductividad hidráulica	Filtro de Kalman MODFLOW

CAPÍTULO 2 MATERIALES Y MÉTODOS

2.1 Descripción de zona de estudio

Para la evaluación del desempeño de la metodología propuesta, se aplicó en el diseño de una red de monitoreo en la Región del Bajío en la parte central de México (**Figura 1**). La Región del Bajío, específicamente donde se localiza el acuífero Irapuato-Valle ha recibido recientemente interés debido a la extracción intensiva de agua para el uso en la agricultura, población e industria. La extracción del agua subterránea (alrededor de 600×10^6 m³/año) mediante más de 2900 pozos de producción en el área proporcionan agua para más de 900,000 habitantes en las zonas urbanas y rurales, 600,000 hectáreas de riego, el consumo de agua de una refinera, una planta de producción de electricidad y varios parques industriales (Comisión Nacional del Agua, 2015; Esteller et al., 2011; J. A. M. Gómez & Sandoval, 2004). Las unidades rocosas en la zona de estudio incluyen una secuencia de rocas plutónicas mesozoicas y vulcano-sedimentarias afectadas con diferentes grados de metamorfismo y afloramiento de rocas intrusivas terciarias en la Sierra de Guanajuato ubicada al norte de la zona de estudio que constituye el basamento hidrogeológico. La secuencia terciaria incluye rocas volcánicas (ignimbritas, flujo de lava y tobas) y sedimentarias continentales (conglomerados y relleno sedimentario de cuenca). La secuencia volcánica del Oligoceno está compuesta por andesita y riolita, las rocas volcánicas del Mioceno y Plioceno que se encuentran por encima son principalmente andesita y basalto asociadas con el Cinturón Neo-volcánico Mexicano (Nieto-samaniego, Ojedagarcía, Alaniz-álvarez, & Xu, 2012). Durante este periodo de deformación, varias fallas fueron reactivadas, la N-S (Falla Taxco-San Miguel de Allende) y direcciones preferentes NE-SW, junto con la Falla El Bajío con dirección principal NW-SE formando un semigraben donde se ubica la cuenca del Bajío (Botero-santa et al., 2015). Al mismo tiempo, la erosión de las montañas ocasionó el transporte del material clástico a la estructura del graben. La deposición del relleno de los sedimentos de relleno de cuenca en la estructura del graben produjo un acuífero libre regional permeable con un grosor máximo entre los 350 a 400 m. Los

sedimentos de relleno de la cuenca están pobremente consolidados con intercalaciones de arenas, limos, arcillas y suelo residual.

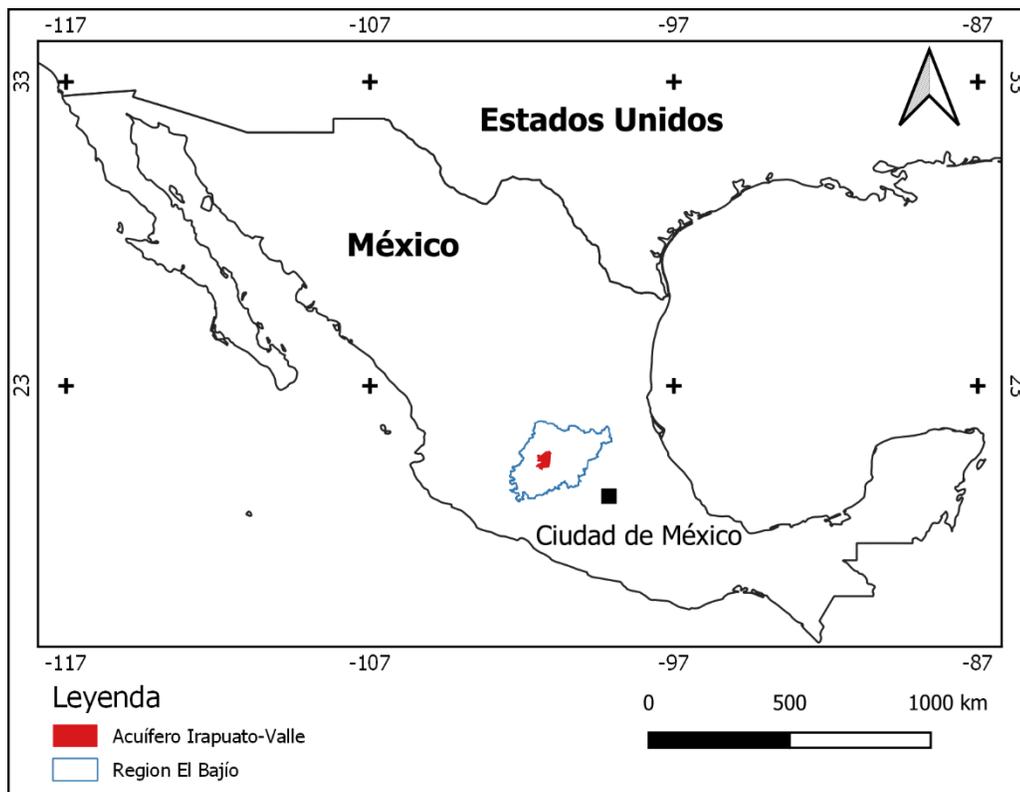


Figura 1.- Mapa de ubicación de la zona de estudio. Fuente: Mapa de La Región del Bajío modificado de Suárez-mota and Villase (2015); Acuífero Irapuato-Valle de National Water Information System (2019); División política de países de Tapiquén (2019).

De la información de pruebas de bombeo, los valores típicos de conductividad hidráulica del acuífero van de 1 a 10 m/día (Júnez-Ferreira et al., 2016a). Las rocas volcánicas del Terciario son cortadas por una serie de fallas normales NW-SE y N-S, generando la posibilidad de un acuífero heterogéneo fracturado con un espesor de 200 a 300 m. Las rocas basálticas del Pleistoceno que afloran en la porción sur de la zona de estudio están en contacto con los sedimentos de relleno de la cuenca, conformando un acuífero libre con una conductividad hidráulica con valores de entre 6 a 18 m/día (Júnez-Ferreira et al., 2016a). Las rocas volcánicas del Oligoceno y Mioceno debajo de los sedimentos también constituyen el acuífero profundo, se han identificado en pozos profundos (600 – 700 m) que producen

agua subterránea termal (35 – 45 °C en la salida), condiciones de fracturamiento a lo largo de estructuras de tendencias NW-SE.

En diciembre de 2003 se realizó una campaña de muestreo para determinar la calidad del agua subterránea a partir de la medición in situ de seis parámetros y el análisis de 20 parámetros fisicoquímicos (Júnez-Ferreira et al., 2016a), en total se obtuvieron 140 muestras a partir de pozos de producción dentro de la zona de estudio. La distribución de los pozos muestreados se puede observar en la **Figura 2**. Los pozos se encuentran en un rango de profundidad de entre 46 a 700 m con una media de 180 m y una desviación de estándar de 115 m. La mayoría de los pozos tienen la parte ranurada a lo largo de la zona saturada del pozo, lo que facilita mezclas artificiales de agua al interior del pozo e implica que las muestras de agua están compuestas por una mezcla de varias capas que contienen agua. Se ha identificado que el agua en zonas de recarga tiene un tipo predominante de $\text{HCO}_3\text{-Ca}$ (650 mg/L de sólidos disueltos en promedio). A lo largo del flujo del agua subterránea, la química del agua evoluciona de un predominante $\text{HCO}_3\text{-Mix}$ y $\text{HCO}_3\text{-Na}$. Los análisis de la calidad del agua subterránea indican la contribución de fuentes de contaminación geogénicas y antropogénicas. Las rocas volcánicas, particularmente las de composición riolítica y los rellenos de sedimentos de cuenca derivados de estas son fuentes potenciales de elementos traza, en especial para el arsénico. Las muestras de agua fueron analizadas para la concentración de arsénico (0.011 mg/L promedio y un máximo de 0.083 mg/L), 35% de las muestras revelan concentraciones por encima del valor de la línea de referencia de la OMS (0.010 mg/L) (**Figura 3a**).

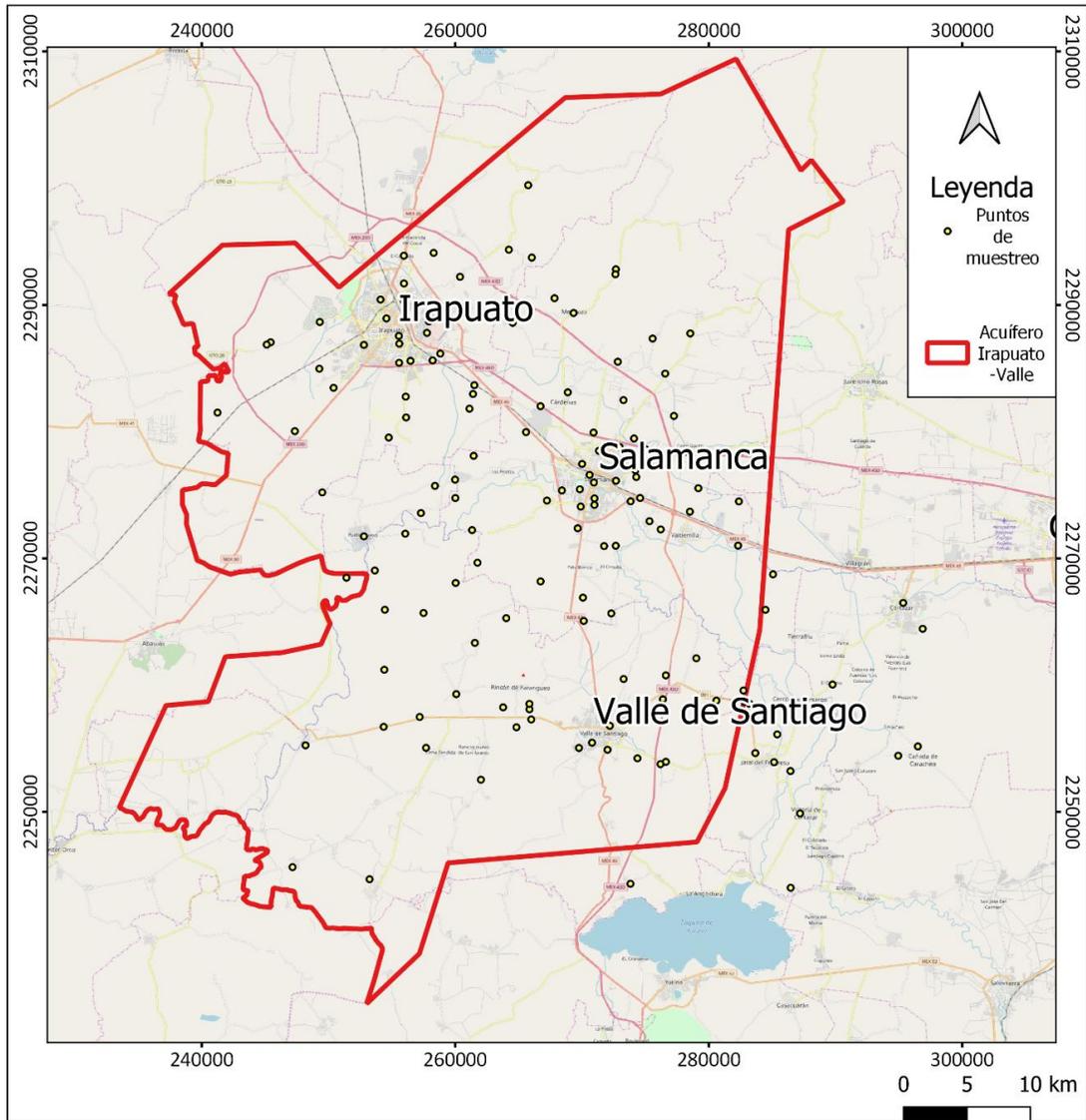


Figura 2.- Ubicación de los puntos de muestreo

El arsénico disuelto es encontrado normalmente como arsénico (III) en ambientes reductores, que generalmente se desarrollan a cierta profundidad en los acuíferos, mientras que cerca de la superficie del agua la condición que prevalece es la óxica (Appelo & Postma, 2005). Existe una falta de información en la calidad del agua de los pozos en el perfil vertical, por esta razón no es posible demostrar la existencia de la transición vertical entre las condiciones óxica y anóxica en el acuífero explotado. Sin embargo, las condiciones aeróbica y óxicas (oxígeno disuelto por encima de 1 mg/L, valores de Eh por encima de 100 mv) fueron detectadas en

todas las muestras de agua subterránea, lo que sugiere que las condiciones óxicas prevalecen en la mayoría de las formaciones geológicas captadoras de agua explotada por los pozos de producción. Teniendo en cuenta que la profundidad de la muestra se puede definir a la mitad de la sección ranurada del pozo, no existe correlación entre la profundidad de la muestra con los valores de arsénico ($R^2=0.07$), lo cual indica que el arsénico disuelto no se puede explicar por la desorción de los minerales de los óxidos e hidróxidos de hierro en condiciones de reducción (**Figura 3b**). La mayoría de las muestras de agua con valores altos de arsénico están asociadas con el tipo dominante $\text{HCO}_3\text{-Na}$, lo que indica que la lixiviación de rocas volcánicas con composición riolítica es un control importante. Existe evidencia en el norte-centro de México que sugiere la disolución de vidrio volcánico de composición félsica aparenta ser el proceso dominante para la liberación del arsénico en el agua subterránea (Banning et al., 2012); la concentración en las regiones centro y noroeste de la zona de estudio parecen derivadas de esta fuente. Los sedimentos de relleno de cuenca constituyen una fuente adicional, los valores de arsénico por encima del estándar de agua potable de la OMS en la región suroeste de la zona de estudio probablemente deriven de dichos sedimentos. La descarbonización de los sedimentos de relleno de cuenca podrían constituir un proceso adicional en el proceso de movilización del arsénico (Banning et al., 2012). La identificación del agua subterránea impactada de fuentes antropogénicas, consideradas tipos de aguas como Cl-Mg, Mezclas-Ca, Mezclas-Na, y SO_4 -mezclado; ejemplifican agua subterránea impactada por contaminación difusa. El impacto de la contaminación difusa producido por el riego puede ser evaluado con la combinación de indicadores como el cloruro, sulfato, nitrato, fosfato entre otros (Cardona, Carrillo-Rivera, Castro Larragoitia, & Graniel-Castro, 2008). El cloruro es un soluto conservador, es útil para detectar la infiltración de la contaminación difusa de actividades agrícolas. Adicionalmente, la conductividad eléctrica es una función de la concentración total y la carga de iones; la relación entre ellos es similar a lineal en la mayoría de los casos. La relación entre el cloruro y la conductividad eléctrica de las muestras de agua en la zona de estudio se muestra en la **Figura 4**. Las características de las aguas

superficiales y residuales locales se muestran como una referencia adicional. A pesar de la dispersión, muestra claramente una tendencia para la mayoría de las muestras, lo que sugiere que la infiltración del agua residual modifica en cierta medida la composición del agua subterránea. Unas cuantas muestras no siguen la tendencia general, particularmente los valores altos en la conductividad eléctrica con valores bajos de cloruro; para estas muestras se muestra una alta alcalinidad (> 500 mg/L CaCO_3) ligado a valores bajos de concentración de nitrato (aproximadamente 4-5 mg/L N-NO_3), lo que sugiere una influencia limitada de los flujos de retorno de irrigación o infiltración de aguas residuales.

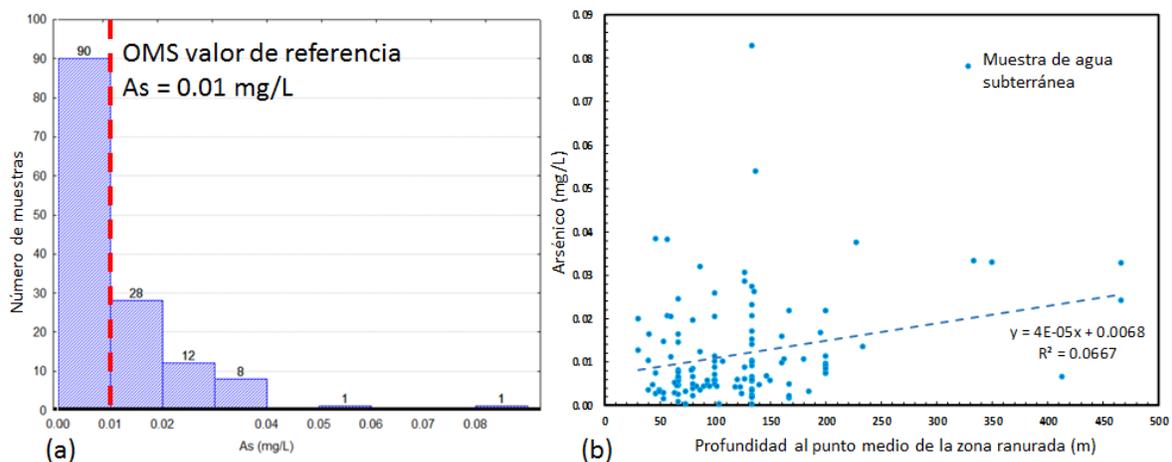


Figura 3.- Análisis de arsénico a) número de muestras por grupos y el límite de referencia de la OMS, b) correlación entre la concentración y la profundidad del pozo.

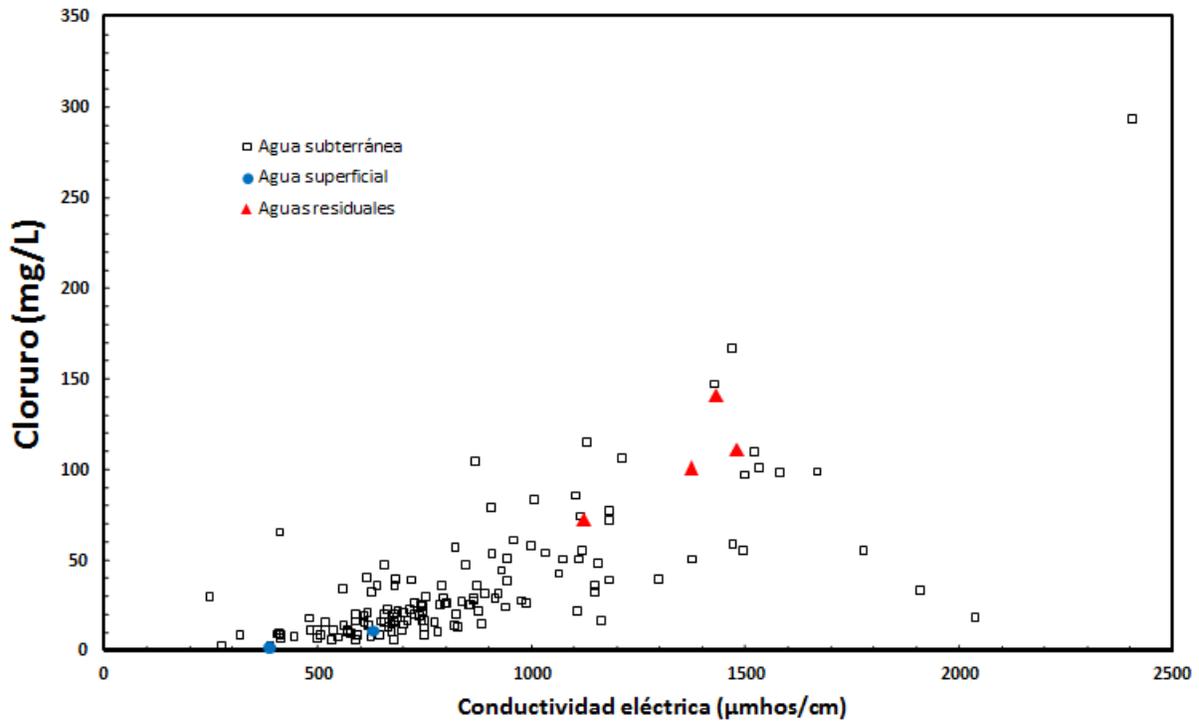


Figura 4.- Correlación entre Cloruro y la conductividad eléctrica.

2.2 Metodología

La metodología desarrollada está basada en la propuesta de Júnez-Ferreira *et al.* (2016a), donde se hace una selección óptima de posiciones de monitoreo utilizando una correlación espacial expresada mediante matrices de covarianza derivadas de modelos de semivariogramas obtenidos a través de análisis geoestadísticos de los datos de los parámetros de calidad del agua subterránea. La optimización de la red de monitoreo se genera mediante el uso cíclico del filtro de Kalman como método de estimación, minimizando la función llamada *joint total* (JT) varianza normalizada del error en la estimación (Júnez-Ferreira *et al.*, 2016a). Esta metodología necesita una malla de estimación espacial donde su objetivo sea la evaluación. Durante el proceso de optimización, la inclusión de pesos a los parámetros de calidad de agua subterránea o de diferentes densidades en la malla de estimación han resultado útiles para la asignación de prioridades más altas a sitios contaminados o a parámetros muy peligrosos. (Júnez-Ferreira *et al.*, 2016a), (Esquivel *et al.*, 2015).

Para la metodología propuesta, la función objetivo incorpora una cobertura espacial de las estaciones de muestreo disponible (por medio de una función suplementaria, FS) además de una versión modificada de la JT. La JT modificada, además de tener un peso para cada uno de los parámetros de calidad del agua subterránea incluye un peso propio para cada uno de los puntos de muestreo, este nuevo algoritmo fue llamado varianza normal total ponderada conjunta del error en la estimación (PCT). Tanto a la PCT y la FS les son definidos diferentes pesos para generar diferentes alternativas de cobertura espacial de la red de monitoreo sin tener pérdidas significativas de información.

La función objetivo (FO) de la metodología propuesta es (Ec. 1):

$$FO = V \times PCT + W \times FS \quad (\text{Ec.1})$$

donde V y W son los factores de peso para PCT y FS respectivamente; $V + W = 1$.

Basado en Júnez-Ferreira *et al.* (2016a) para obtener PCT, el filtro de Kalman es implementado a través del uso de \hat{q}^0 y P^0 que representan el vector de estado inicial y la matriz de varianza del error inicial respectivamente. El proceso del FK emplea las siguientes ecuaciones (Ec.2, Ec.3 y Ec.4)

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

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

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

donde K_{n+1} es la ganancia de Kalman, H_{n+1} es la matriz de muestreo, z_{n+1} es el vector de medida y r_{n+1} es la covarianza del error medido. Los superíndices hacen referencia al número de mediciones empleadas en la estimación; los subíndices son utilizados para indicar la transición cuando una nueva medición va a ser utilizada en el proceso.

La matriz de covarianza espacial para cada uno de los parámetros es calculada como se observa (Ec.5).

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

donde $C(h)$ representa la matriz de covarianza, $C(0)$ es la varianza para cada parámetro analizado, y $\gamma(h)$ es la función del modelo del variograma.

El valor de PCT es calculado con (Ec.6):

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

donde $\sigma_{w,p}^2$ representa la varianza en la estimación del error normalizada (ubicada en la diagonal principal de la matriz de covarianza actualizada correspondiente) en cada posición de estimación para el parámetro p , cuando el punto de muestreo w ha sido muestreado, α_w representa el peso del punto de muestreo w , α_p representa el peso del parámetro p y np es el número de parámetros

Las principales ventajas que se pueden observar en la metodología son: a) considera la posibilidad de obtener diferentes diseños alternos de redes de monitoreo considerando una perspectiva de cobertura espacial de la zona de muestreo b) permite evaluar los pesos asignados a las funciones que conforman la FO mediante una gráfica que sirve como referencia de los casos extremos (“óptimo” y “peor”) que permiten identificar el aumento del error del diseño que se está evaluando c) permite la asignación de diferentes pesos a uno o más puntos de muestreo en particular, lo que concede la posibilidad de tomar en cuenta las condiciones hidrogeológicas específicas como podrían ser áreas contaminadas sin afectar el proceso de optimización.

Malla de estimación personalizada

Como parte de la metodología propuesta, se ha desarrollado un algoritmo que permite la generación de una malla de estimación personalizada (MEP). Se obtienen las estimaciones del FK y se evalúa la FO en la MEP; esto ofrece ventajas comparativas contra mallas regulares como son las implementadas en algunos softwares comerciales como el XTools Pro 18 (XTools Pro), la herramienta de create fishnet (ArcGIS) y la sampling design tool (Buja), ya que al contar con una malla de estimación ajustada a la zona de los puntos de muestreo evita la extrapolación.

Para la definición de la MEP se inicia con el cálculo de la distancia entre los puntos de muestreo extremos (distancia máxima), posteriormente se hace el cálculo de la densidad de la malla de muestreo, la cual se obtiene tomando la distancia máxima y se divide entre diez; esta densidad (la separación tanto en horizontal como en vertical que existe entre los nodos de la malla de estimación) es seleccionada para trazar una malla rectangular que cubra la totalidad de la zona donde existen puntos de muestreo disponibles y a cada uno de ellos se les asigna un valor inicial de 0. Los valores de los nodos que conforman la MEP, podrán ser actualizados de manera que aquellos que cambien su valor conformarán la MEP final. La manera como esta actualización se lleva a cabo es mediante cuatro recorridos de arriba (Norte) a la parte inferior (Sur) en donde si cumple la condicionante analizada se verá afectado el valor y dicha actualización solo se llevará a cabo para nodos con valor de 0, si es diferente de 0, ya no se llevará a cabo un segundo cambio de valor.

El resumen de este proceso se presenta a continuación:

1.- Se asignan valores de 1 a aquellos nodos que son vecinos de los puntos de muestreo, entendiendo por vecinos aquellos nodos a izquierda, derecha, arriba y abajo que no tienen otro nodo intermedio entre ellos y el punto de muestreo.

2.- Se asigna el valor de 2 a aquellos nodos que haciendo una revisión por fila, se encuentran entre nodos que ya contienen el valor de 1.

3.- Se asigna el valor de 3 para aquellos nodos que se encuentran en uno o más renglones totalmente en 0 y que conectan nodos diferentes de 0 de diferentes renglones.

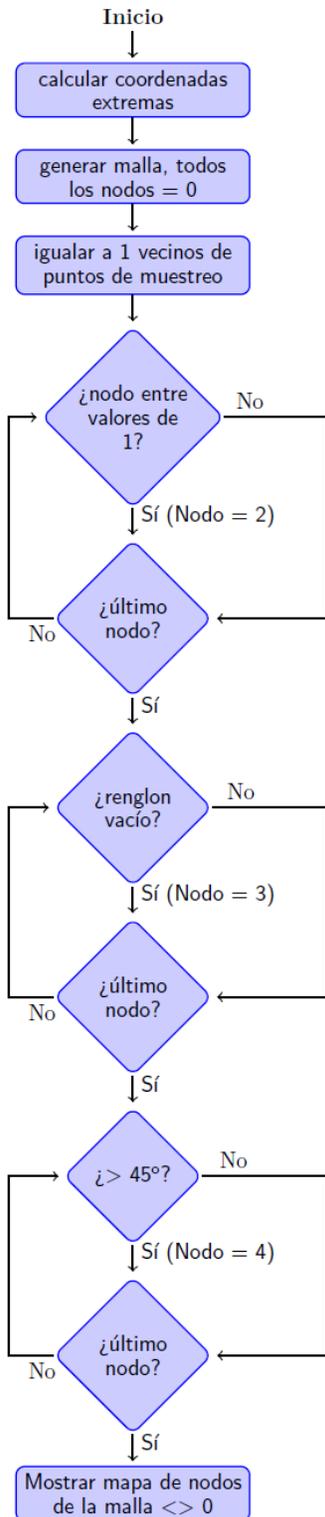
4.- Se asigna el valor de 4 en aquellos nodos que permiten delimitar la forma de la malla de tal forma que no existan cambios de más de un nodo hacia el Este y Oeste cuando se realiza un recorrido de Norte a Sur y viceversa.

El diagrama de flujo y pseudocódigo para este propósito se muestra en la **Figura 5**.

De acuerdo al pseudocódigo anteriormente descrito, se puede observar el detalle de asignación de valores para cada uno de los casos que se presentan en las diferentes etapas de desarrollo del pseudocódigo. En la **Figura 6** podemos observar la representación mediante un cuadro en color negro la representación de los puntos de muestreo, en círculos en color rosa los nodos que representan los cuatro vecinos (los más cercanos arriba, abajo, izquierda y derecha) y en círculos sin relleno los nodos de la malla de estimación que aún se encuentran sin seleccionar para la MEP.

Una vez que se han seleccionado los vecinos, se procede a la selección de los puntos intermedios entre los puntos del caso 1, de manera que si existen nodos vacíos entre nodos del mismo reglón del caso 1, será establecido un valor de 2, que son los que se pueden observar como ejemplo en color azul en la **Figura 7**.

Habiendo establecido los valores de los nodos para los casos 1 y 2, se hace un recorrido de toda la malla para detectar renglones totalmente vacíos entre renglones que si tienen información, si se llegan a detectar renglones en esa condición se le asignan a los nodos dentro de esos renglones vacíos, el valor de 3 para conectar entre renglones, estos nodos se pueden observar como ejemplo en color café en la **Figura 8**.



Obtener las coordenadas extremas con base en los puntos de muestreo

Inicializar en 0 los nodos de la malla de estimación

Mientras el contador de nodos de la malla de estimación sea menor al total de nodos

Mientras el contador de los puntos de muestreo sea menor o igual al total de puntos de muestreo

Si el nodo de la malla de estimación es vecino de un punto de muestreo

Igualar el valor del nodo a 1

Mientras el contador de la malla de muestreo es menor al total de nodos de la malla de muestreo

Si el nodo de la malla de estimación se encuentra entre dos nodos de la malla de muestreo del mismo renglón

Igualar a 2 el nodo de la malla de muestreo

Mientras el contador de nodos de la malla de estimación sea menor al total de nodos de la malla de estimación

Si en el renglón todos los nodos de la malla de estimación son cero

Identificar y sumar el número de renglones continuos totalmente en 0

Calcular el número de nodos de la malla de estimación por renglón basados en la posición y los renglones con 0

Mientras el contador de renglones es menor al total de renglones con 0

Mientras el número de nodos es menor al total de nodos del renglón

Igualar a 3 el nodo de la malla de muestreo

Mientras el contador de la malla de estimación sea menor al total de nodos de la malla de estimación

Si las posiciones anterior y siguiente en el nodo de la malla de estimación es 0

Si la posición en el renglón inmediato inferior es 0

Iguala a 4 el nodo de la malla de muestreo

Mientras el contador de la malla de estimación sea menor al total de nodos de la malla de estimación

Si el nodo de la malla de estimación es diferente de 0

Mostrar el nodo de la malla de estimación rosa (1), azul (2), café (3) y verde (4)

Figura 5.- Diagrama de flujo y pseudocódigo de la generación del MEP.

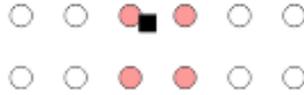


Figura 6.- Caso 1 para la MEP

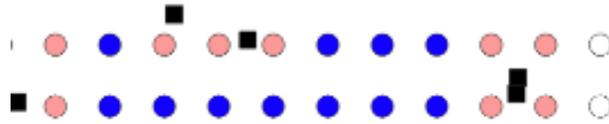


Figura 7.- Caso 2 para la MEP

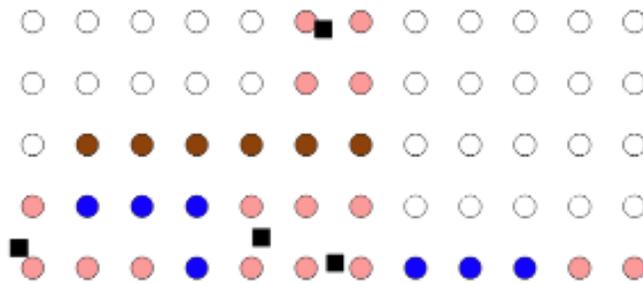


Figura 8.- Caso 3 para la MEP

Finalmente, una vez conectados todos los renglones, se procede a hacer un último recorrido por renglones de norte a sur con la intención de agregar nodos en la malla de estimación con valor de 4 en aquellos casos en los que exista una diferencia de más de una columna al conectar entre renglones, es decir, si al hacer el recorrido en el renglón superior en la columna 5 existe un elemento diferente de 0 y en el renglón siguiente la primer columna con valor diferente de 0 es la 10, se hará una asignación del valor 4 hasta llegar a la columna 6, un ejemplo de la asignación del valor de 4 a ambos lados de la malla de estimación se puede observar en la **Figura 9**.

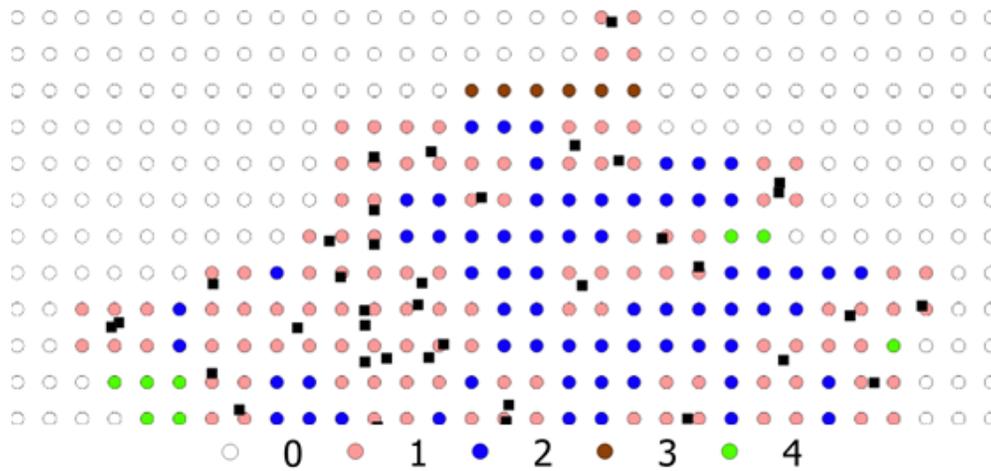


Figura 9.- Caso 4 para la MEP

Los colores en la **Figura 10** corresponden a los valores que fueron asignados y por tanto en que parte del algoritmo fueron asignados a los nodos de la MEP, aquellos nodos en la malla de estimación que no muestran relleno alguno significa que han sido excluidos para los próximos cálculos por no estar dentro de los criterios que se establecieron para su selección mencionados anteriormente.

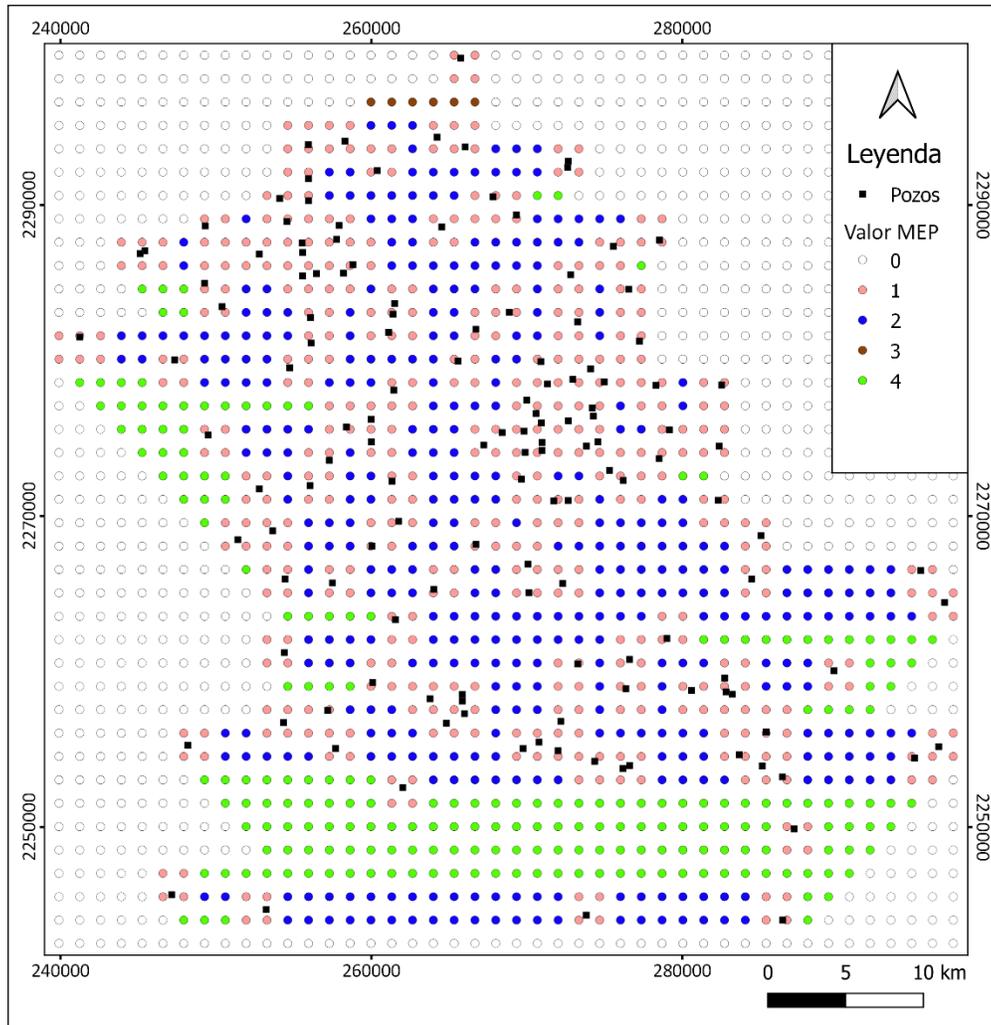


Figura 10.- Malla de estimación personalizada (MEP).

Una vez establecida la MEP, se hace el cálculo para cada uno de los nodos de la malla de estimación por parámetro comenzando por la suma de los puntos de muestreo que están dentro del rango del modelo del parámetro en evaluación, en cada uno de los nodos también se considerará cuantos puntos de muestreo son los que quedan dentro del rango. A partir de la información obtenida para cada punto de muestreo y nodo de la malla de estimación, se hará una suma para cada punto de muestreo en donde se agrupen los nodos de la malla de estimación por grupos de acuerdo al número de puntos de muestreo que están dentro del rango, de tal manera que todos aquellos nodos de la malla de estimación que tienen un solo punto de muestreo en rango formarán un grupo que se sumará y se dividirá

entre los puntos de muestreo en rango, esto se hará para todos los nodos de la malla de estimación dentro del rango de los puntos de muestreo y las sumas y divisiones correspondientes se harán para todos los grupos que sea necesario generar en función de los nodos que estén dentro del rango de los puntos de muestreo. Este proceso se repetirá para cada uno de los parámetros que se están considerando en la red de monitoreo.

Esto se puede ejemplificar con un par de parámetros (Conductividad eléctrica y Arsénico) y un par de puntos de muestreo (PM1 y PM2) como se observa en la **Figura 11**. En primera instancia se representa mediante círculos rojos el rango del PM1 para cada parámetro, en el caso de la Conductividad eléctrica es mayor el rango, por lo que el número de nodos de la MEP (representados mediante puntos con relleno en color gris) es mayor, así como también es mayor el número de puntos de muestreo que están dentro del rango de cada uno de los nodos la MEP. Se puede observar que en ambos parámetros existe una zona de intersección de los círculos que representan los rangos de los puntos de muestreo, en esa zona los nodos de la MEP tienen una suma cuando menos de dos en el número de puntos de muestreo que estén en rango.

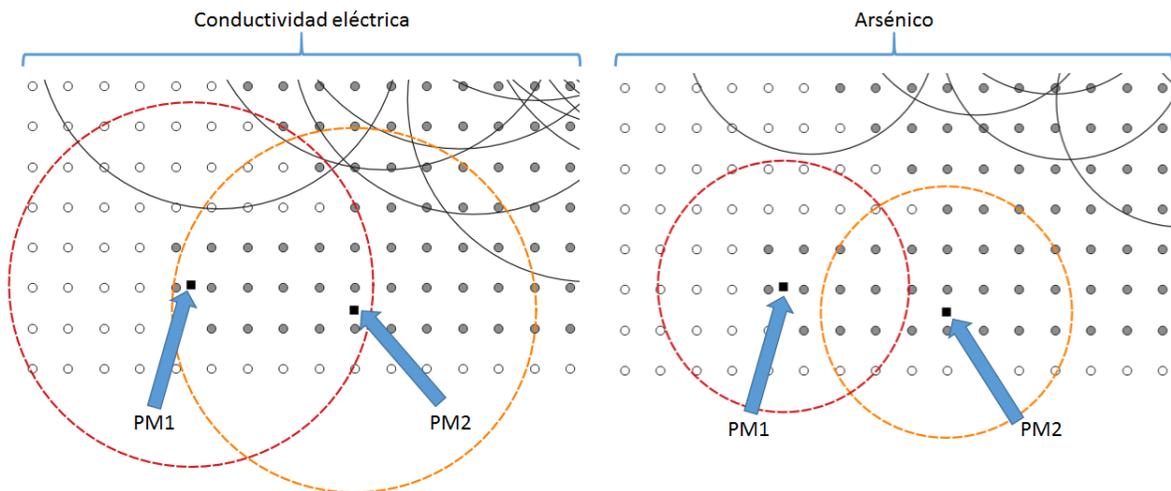


Figura 11.- Alcance de los rangos de puntos de muestreo en los nodos de la MEP para conductividad eléctrica y arsénico

Para la zona considerada en el rango del PM2 (representado mediante círculo anaranjado), se pueden identificar una gran cantidad de nodos de la MEP que solo

están contenidas en ese círculo, lo cual resulta relevante ya que esto implica que solo el PM2 tiene el suficiente rango para poder estimar el valor en esos nodos, además de ello, en el caso del parámetro del Arsénico, existen algunos elementos que no están contenidos en ninguno de los círculos que están representando los rangos de los puntos de muestreo, lo que equivale a que no podrán ser estimados con los rangos que se están estableciendo.

El establecimiento de los valores de los rangos de los modelos para cada parámetro, permite acotar la MEP mediante la identificación de aquellos nodos que no existe la posibilidad de estimar por ninguno de los parámetros al estar fuera de los rangos. Teniendo eso en cuenta, se tomaron los modelos de los semivariogramas de todos los parámetros, para posteriormente elegir el máximo rango considerado para definir cuales nodos se dejan fuera de la malla de estimación partiendo de la MEP, esto da como resultado la Malla de Estimación Recortada (**Figura 12**) que puede considerar todos los nodos de la MEP o excluir uno o más de esta, pero no puede considerar los que originalmente ya fueron descartados por la MEP.

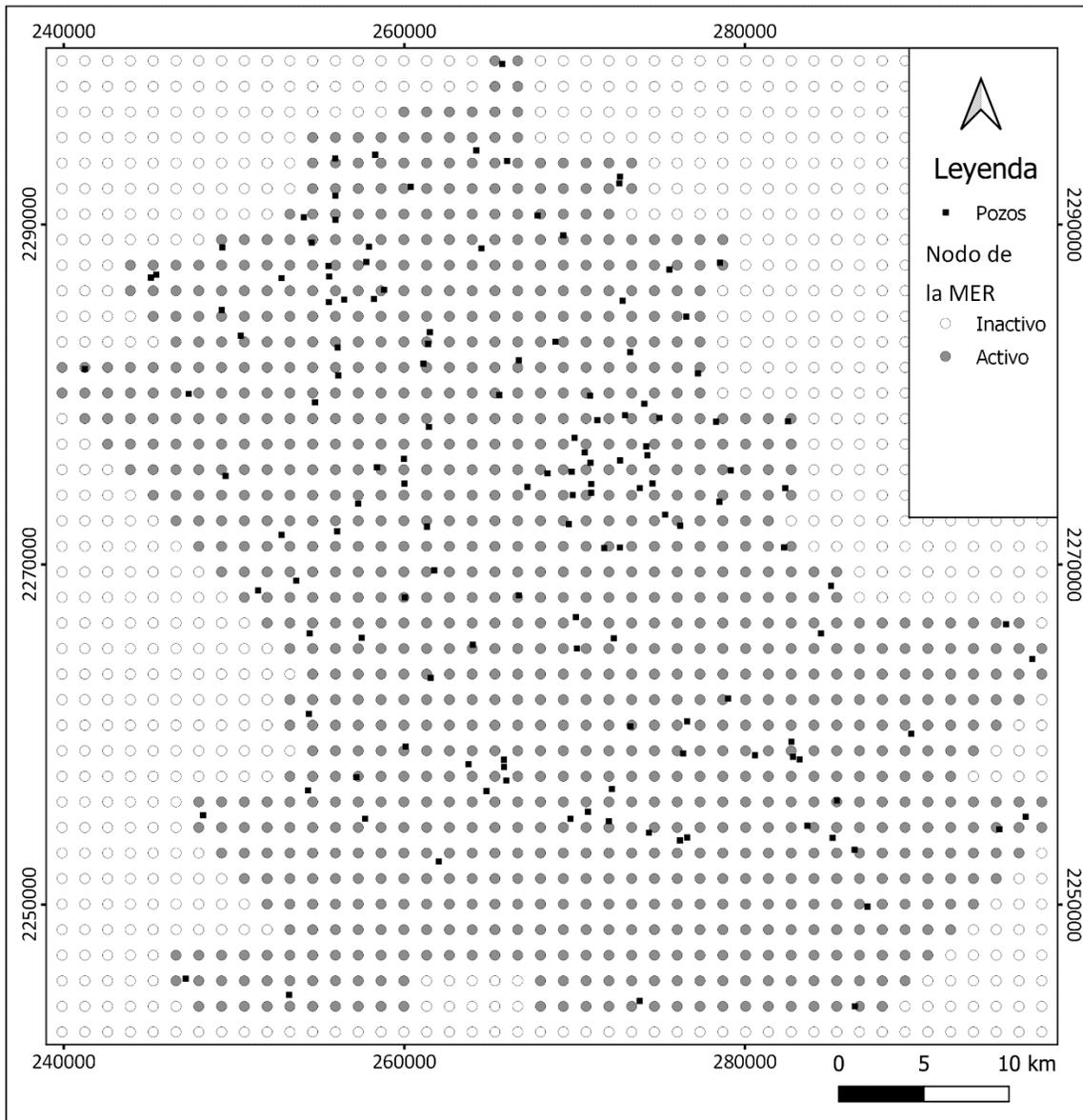


Figura 12.- Malla de estimación recortada (MER)

Función suplementaria

Para el caso de estudio, la Función Suplementaria (FS) se propone para establecer una prioridad mayor a aquellos puntos de muestreo que tienen influencia sobre una mayor cantidad de nodos de la malla de estimación y al mismo tiempo, es el único o existen muy pocos puntos de muestreo que cuentan con influencia sobre dichos nodos, lo que puede aumentar el interés de mantener

los puntos de muestreo para no perder cobertura espacial (e. g. en zonas con poca cobertura espacial en algunas zonas).

Para la obtención de la FS, se parte de aquellos nodos de la MER que están dentro del rango de los modelos para cada parámetro, se hace un conteo de los puntos de muestreo que están dentro de ese rango, dando como resultado que para cada nodo de la MER por parámetro se pueda saber ese número (**Figura 13**)

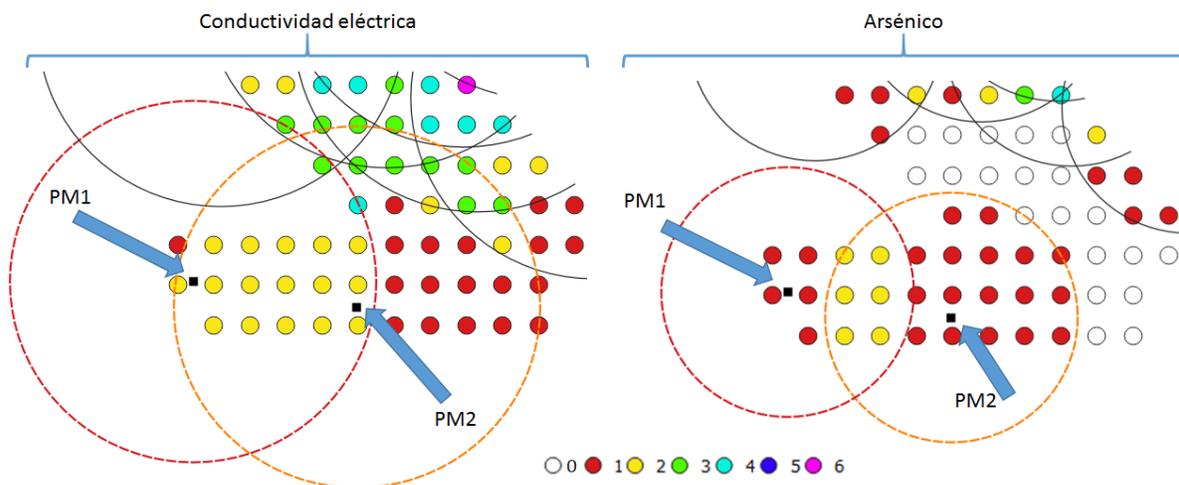


Figura 13.- Conteo del número de puntos de muestreo dentro de rango para cada nodo de la MER por parámetro

Una vez que cada nodo de la MER cuenta con la información del número de puntos de muestreo, se hace una sumatoria por grupo de nodos clasificados en función del número de puntos de muestreo dentro del rango, esto permite saber en el caso del ejemplo que para el PM1 se cuenta con un 1 nodo en el grupo de los que solo están en el rango de 1 punto de muestreo, cuenta con 16 nodos en el grupo de los que su rango tiene 2 puntos de muestreo y así sucesivamente hasta complementar todos los nodos que están en cada uno de los puntos de muestreo. Posteriormente se hace una sumatoria de todos los nodos y para cada grupo se divide el número de elementos entre el número de puntos de muestreo que representa ese grupo. Este proceso se hace para cada uno de los parámetros que se estén analizando para considerar el rango definido en el modelo en la

determinación del tipo de grupo que habrá de tener cada nodo de la MER (**Figura 14 y Figura 15**)

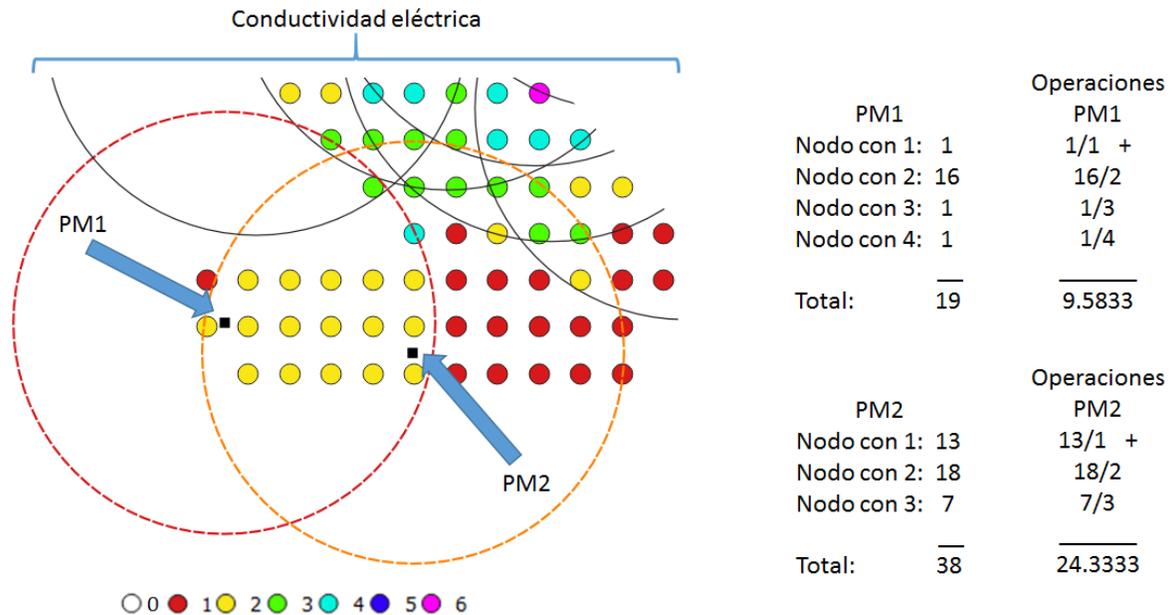


Figura 14.- Obtención de los valores necesarios para obtener la FS para la conductividad eléctrica

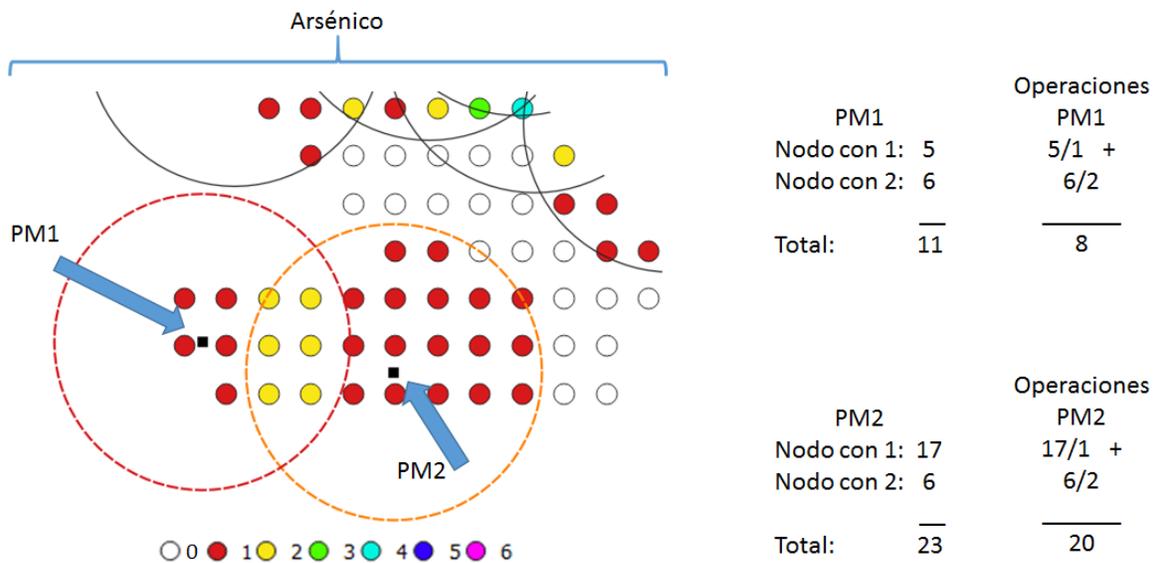


Figura 15.- Obtención de los valores necesarios para obtener la FS para el arsénico

Tabla 2.- Resultados obtenidos para el cálculo de FS para diferentes puntos de muestreo y parámetros

FS	Punto de muestreo	Valor
	PM1-Conductividad eléctrica	$9.5833/19 = 0.5043$
	PM2-Conductividad eléctrica	$24.3333 / 38 = 0.7373$
	PM1-Arsénico	$8 / 11 = 0.7272$
	PM2-Arsénico	$20 / 23 = 0.8695$

Una vez que se cuenta con los valores en los nodos que permitieron el cálculo de la FS, se pueden representar sus valores de manera tal que su representación nos denote en colores del rojo al anaranjado, las zonas donde existe una menor densidad de puntos de muestreo y por consiguiente su importancia bajo el enfoque de la FS, en el caso extremo los puntos de muestreo en colores del azul al morado.(Figura 16)

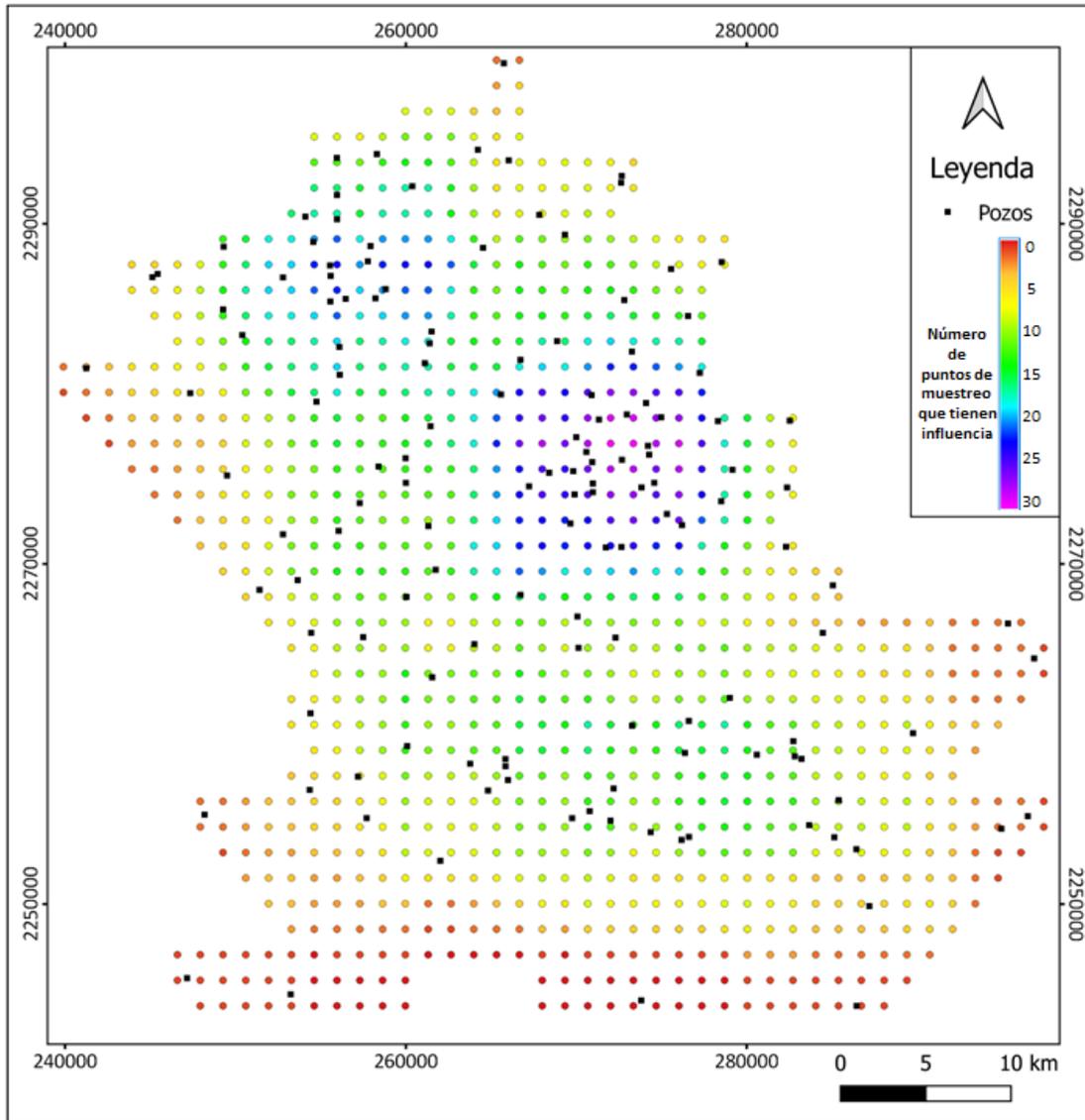


Figura 16.- Ejemplo de malla de estimación recortada (MER) representando los valores en los nodos para la obtención de la FS.

Ya que fue analizado el proceso para la obtención de los valores que permiten el cálculo de la FS se puede definir de acuerdo a lo siguiente:

Los valores más bajos en la MER representan las zonas con la menor cobertura espacial de puntos de muestreo.

Las fórmulas para obtener los valores de la MER son (Ec.7):

$$ie_{w,p} = \sum_{n=1}^{nn} nodo_n \quad (Ec.7)$$

donde $ie_{w,p}$ representa la influencia del punto de muestreo w sobre los nodos de la MER, correspondiente al parámetro p , $nodo_n$ es 1 cuando el nodo está dentro de la circunferencia con el centro en el punto de muestreo w y el radio igual al rango del semivariograma del modelo en el parámetro p ; $nodo_n$ es 0 cuando el nodo está fuera de la circunferencia, nn es el total de nodos en la MER (Ec.8);

$$iw_{n,p} = \sum_{w=1}^{nw} \text{punto de muestreo}_w \quad (\text{Ec.8})$$

donde $iw_{n,p}$ representa la influencia de los puntos de muestreo sobre el nodo n , correspondiente al parámetro p , $\text{punto de muestreo}_w$ es 1 cuando el punto de muestreo está dentro de la circunferencia con centro en el nodo n y el radio es igual al rango del semivariograma del modelo en el parámetro p , $\text{punto de muestreo}_w$ es 0 cuando el punto esta fuera de la circunferencia, nw es el total de puntos de muestreo (Ec.9);

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

donde $inw_{w,p}$ representa la influencia en los puntos de muestreo en zonas circulares de la MER, $r_{n,p}$ es 1 cuando el nodo dentro de la circunferencia con centro en el punto de muestreo w y un radio igual al rango del semivariograma del modelo en el parámetro p , $r_{n,p}$ is 0 cuando el nodo esta fuera de la circunferencia (Ec.10);

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

donde $f_{w,p}$ representa el factor del punto de muestreo w para el parámetro p (Ec.11);

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

donde ff_w representa el factor final del punto de muestreo w , np representa el número total de parámetros considerados en el diseño de la red.

La metodología planteada toma $FS = ff_w$.

Una parte fundamental en la propuesta de la modificación en la metodología de JÚnez-Ferreira *et al.* (2016a) consistió en la integración de lo que se le denominó como factor de influencia, en el cual se realizan cálculos en los nodos de la malla de estimación, así como en los puntos de muestreo para determinar el número de nodos que se encuentran dentro del rango de influencia de cada punto de muestreo de acuerdo a la geoestadística y de esos nodos, en qué porcentaje influyen, considerando si existe más de un punto de muestreo dentro del rango y posteriormente hacer una sumatoria de cuantos nodos están dentro de su rango y que importancia relativa implican.

CAPÍTULO 3 SOFTWARE

Tomando como base la propuesta metodológica, se plantearon las opciones para la optimización del proceso por medio de la integración de todo el proceso en la generación del diseño de la red de monitoreo de calidad de agua subterránea

3.1 Diseño del software

Para la aplicación de la metodología, se analizaron diferentes lenguajes de programación en los que se tuviera la posibilidad de integrar todos los elementos de la metodología, que permitiera la representación gráfica de los resultados, que contemplara herramientas tales que las operaciones matriciales propias del desarrollo de la parte del FK facilitaran su desarrollo y que permitiera la integración en un SIG.

Se tomaron en el JavaScript, Java, Python, PHP y C++ por ser los lenguajes de desarrollo más utilizados actualmente y el punto clave para la determinación de la selección del lenguaje fue la integración en un SIG, por lo que se optó por el Python.

La idea de la integración al SIG fue debido a que inicialmente se había planteado la opción de que funcionara como un complemento a un software ya existente, inicialmente a un software comercial, posteriormente a un software libre y finalmente el desarrollo se hizo de manera independiente, aunque cuando se optó por esta última opción ya se había iniciado el desarrollo en Python y se pudo observar que contaba con las librerías de Numpy, la cual considera instrucciones para operaciones matriciales utilizando el procesador matemático de la computadora, lo que mejora el tiempo de respuesta de las operaciones, el Matplotlib, que permite diferentes formas de representación de la información necesaria para los mapas, variogramas y gráficas, el PyQt que permite la generación de las pestañas, ventanas, botones, barras de desplazamiento, listas desplegables y tablas, por lo que se constató que cumplía con lo necesario para el desarrollo de la herramienta.

Además de la selección del Python como lenguaje, se analizaron 3 IDE's para el desarrollo los cuales fueron Eclipse, el PyCharm e Spyder, optando por el PyCharm en su versión educativa por considerarse el más amigable en su entorno.

Debido al tiempo que se llevó el desarrollo, se fueron actualizando las versiones utilizadas en el lenguaje y por consiguiente de las librerías anteriormente mencionadas. El desarrollo comenzó en el Python 2.8 y se terminó en la versión 3.7.

La forma de distribuir el proceso de la metodología consistió en una ventana con un solo menú y con 3 pestañas como se observa en la **Figura 17** en donde se puede observar que el número 1) que representa el menú de Archivo, el 2) la pestaña del análisis exploratorio y estructural, la 3) la malla de estimación y de ponderación, la 4) el diseño de red de monitoreo de agua subterránea, la 5) la tabla de datos, la 6) la estadísticas básicas, la 7) el histograma, la 8) el mapa de distribución, la 9) los modelos y sus parámetros, la 10) el variograma, la 11) la validación cruzada, las barras de desplazamiento de 12) la pepita, 13) de la Meseta, 14) del Rango, la 15) la lista desplegable del parámetro y la 16) el botón de graficar.

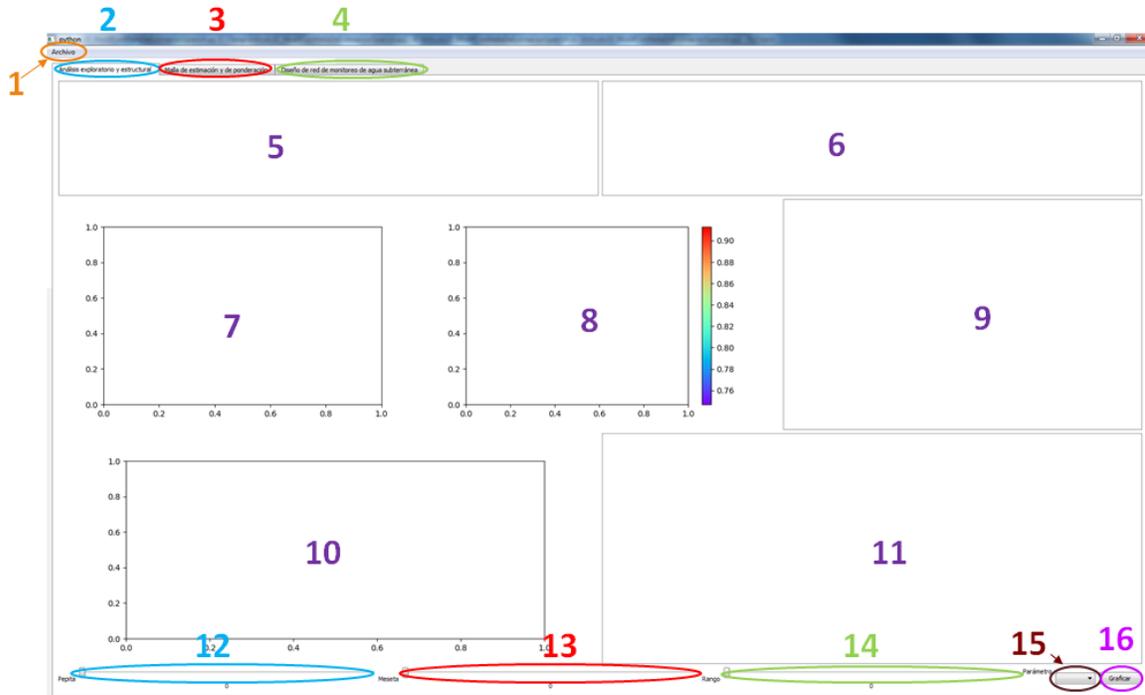


Figura 17.- Pantalla principal del software

En la pestaña de la malla de estimación y ponderación (**Figura 18**), se tiene en la 17) el mapa de distribución de la malla de estimación, en la 18) el mapa de distribución de la malla de estimación personalizada, en el 19) el botón para graficar la malla de estimación y ponderación y en el 20) se tiene el botón para cargar una malla de estimación a partir de un archivo de texto.

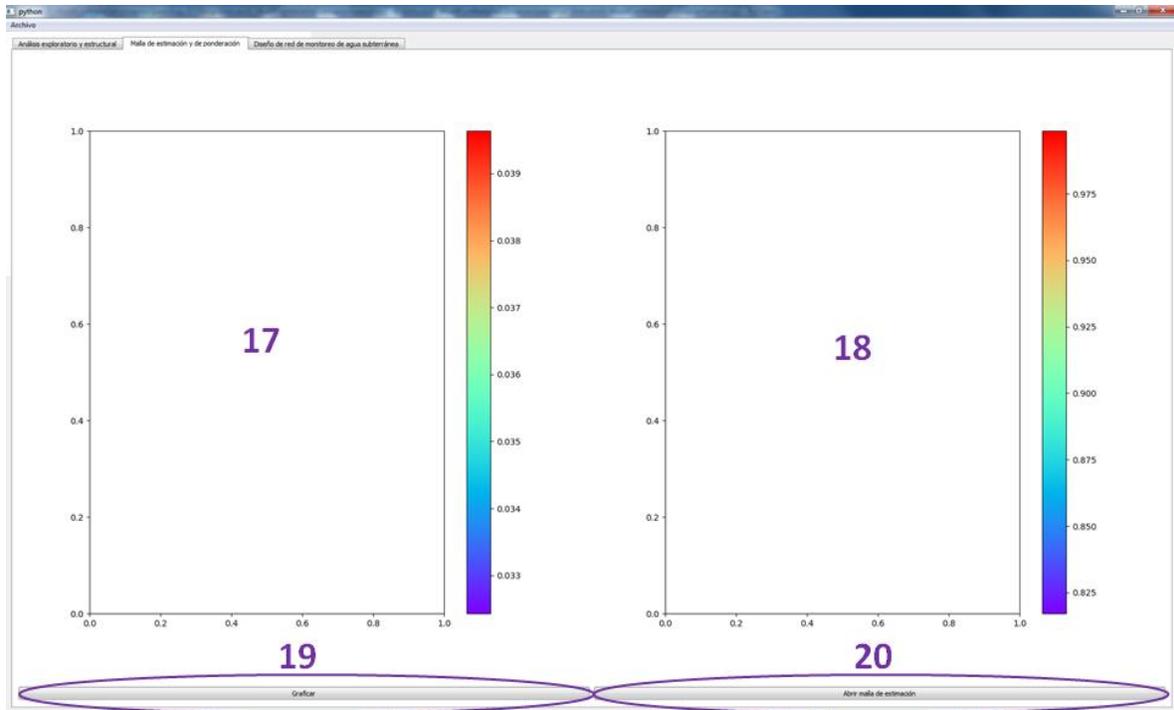


Figura 18.- Pantalla de la malla de estimación y ponderación

En la pestaña del diseño de red de monitoreo de agua subterránea (**Figura 19**), se tiene en el 21) la barra de desplazamiento para definir los pesos del FK vs FS, en la 22) el mapa de distribución de los puntos de muestreo con su correspondiente orden de prioridad, el 23) la gráfica comparativa de la óptima, peor y seleccionada, la 24) para calcular y mostrar el mapa de distribución y la gráfica, el 25) para mostrar el mapa de distribución, la 26) la barra de desplazamiento para indicar a cuantos puntos de muestreo se les quiere mostrar la etiqueta con su orden de prioridad, 27) si se desea que el cambio en la barra de desplazamiento muestre automáticamente los cambios.

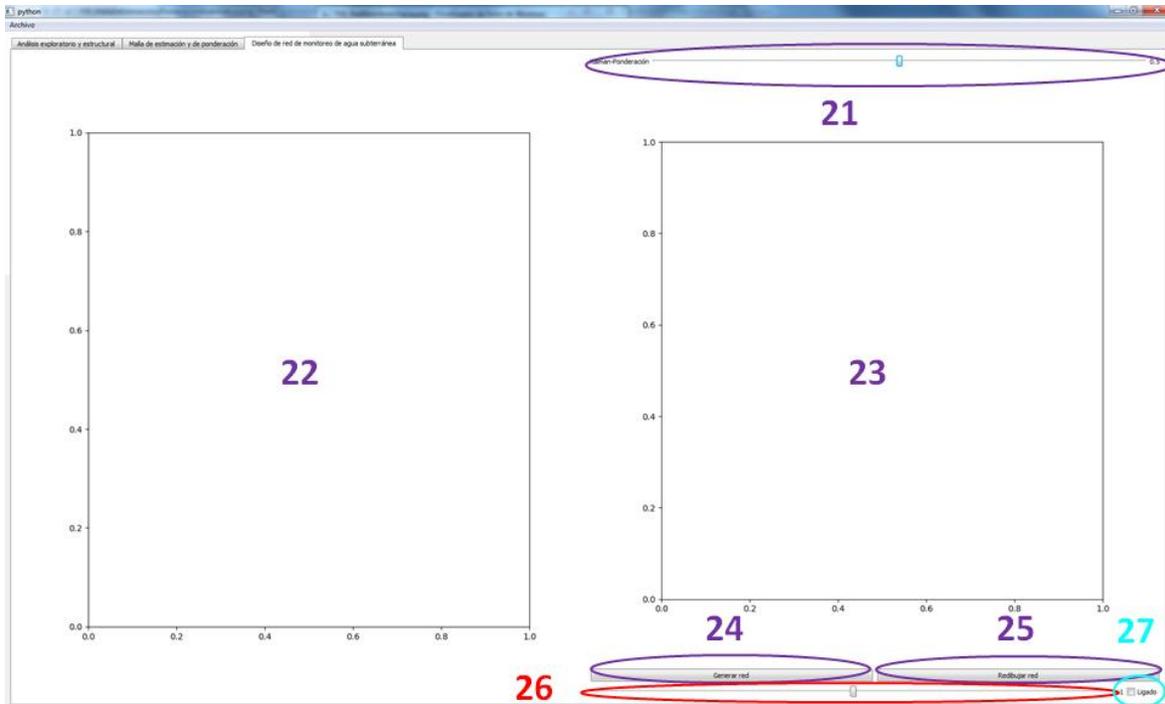


Figura 19.- Pantalla del diseño de la red de monitoreo y comparativa de la seleccionada

3.2 Desarrollo e implementación

Para la carga de los datos en la herramienta, es necesario contar con un archivo de texto que tenga la estructura que se puede visualizar en la **Tabla 3**, donde se observan cuáles son los elementos que debe de tener la tabla de datos y el acomodo de dichos elementos, donde a partir de la 6 columna irán todos aquellos parámetros que se desean considerar en la generación de la red de monitoreo, es importante considerar que de no tener la estructura establecida con sus correspondientes encabezados (primer renglón), no se hará la carga de los datos.

Tabla 3.- Estructura de la necesaria en la tabla para herramienta

Columna	Nombre	Descripción	Comentarios
1	Clave	Identificador único	≤ 8 caracteres
2,3	X, Y	Coordenadas en UTM	≤ 2 decimales
4	Z	Elevación en metros	En metros
5	W	Peso del punto de muestreo	1, >0 <1, >1
6, ..., n	P1, ..., Pn	Parámetros 1 hasta n	En las unidades del parámetro

Teniendo el archivo de texto con la estructura señalada, el formato con el que deberá de guardarse es el CSV, con esto, se asegura la correcta lectura del

archivo a través del menú de archivo (**Figura 20**), donde se toma la opción de Abrir y se busca la ruta del archivo CSV para posteriormente seleccionarlo para su apertura y abrirlo. Después de su apertura, se mostrará la totalidad de los datos, así como sus estadísticas básicas, donde se podrán detectar algunos errores en la información como valores extremos o falta de datos. Si se llegara a presentar alguno de estos casos, no es necesario hacer la corrección desde el archivo de texto y volver a cargar la información, ya que se puede hacer la corrección directamente desde la tabla de datos (**Figura 21** y **Figura 22**) y volver a presionar el botón de graficar para que considere los cambios hechos directamente sobre la tabla de datos.



Figura 20.- Apertura y despliegue de la tabla de datos

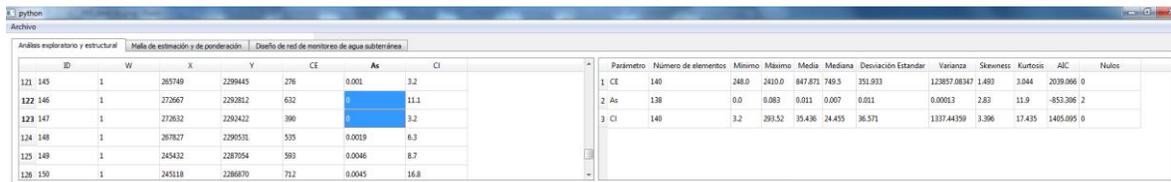


Figura 21.- Identificación de errores en la tabla de datos

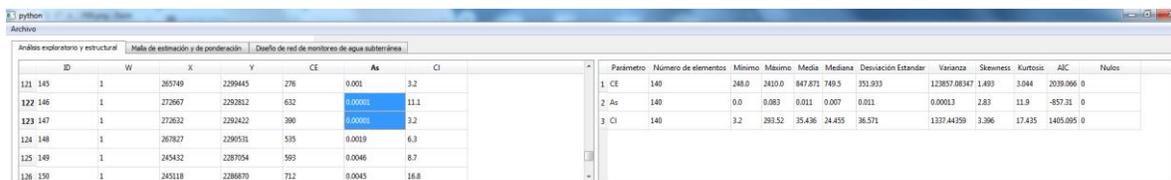


Figura 22.- Actualización de la información en la tabla de datos

Habiendo revisado la información desde su tabla de datos y sus estadísticas básicas, se puede seguir trabajando con el análisis exploratorio de los datos mediante el histograma que nos permitirá ver la distribución de nuestros datos por parámetro, esta opción tiene una variante si se escoge la transformación logarítmica (**Figura 23** y **Figura 24**). Al lado derecho del histograma podemos observar el mapa de distribución de los puntos que nos permite identificar puntos fuera de la zona de la nube de puntos y al igual que en la identificación de datos nulos, nos permite la corrección de la información directamente en la tabla de datos cargada para nuevamente graficar la información. Además de mostrar la distribución de puntos, asigna colores de acuerdo con una rampa de colores que representa los valores del parámetro analizado, esto permitirá identificar alguna tendencia en los datos y/o valores extremos como parte del análisis exploratorio.

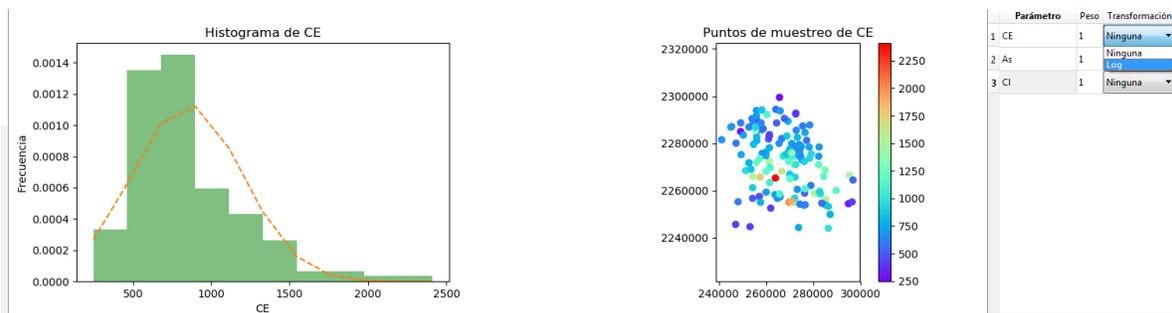


Figura 23.- Representación de la información mediante el histograma y mapa de ubicación para el análisis exploratorio

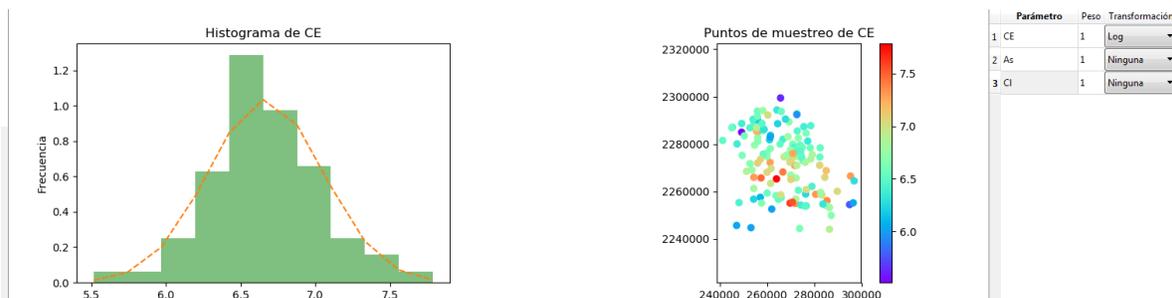


Figura 24.- Transformación de los datos de un parámetro

Después de hacer el análisis exploratorio, se prosigue con el análisis estructural de los datos, donde se definen los valores del modelo (pepita, meseta, rango) que se utilizarán para generar el variograma para cada uno de los parámetros (**Figura**

25). La herramienta cuenta con la posibilidad de escoger entre el modelo esférico y el de gauss y por default se establece el esférico.

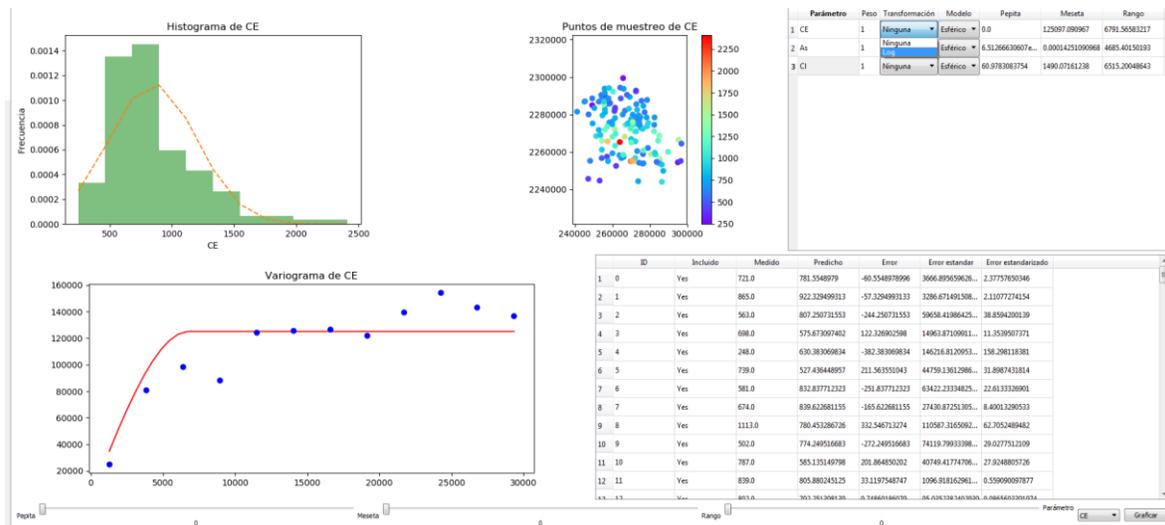


Figura 25.- Representación de los datos mediante el variograma y la validación cruzada para el análisis estructural

La información que se muestra al presionar el botón de graficar corresponde a la del parámetro que se tiene seleccionado en la lista desplegable de la esquina inferior derecha, si se quiere analizar otro de los parámetros de la tabla de datos, se deberá de hacer la selección en la lista desplegable y nuevamente presionar el botón de graficar. Finalmente, en esta misma pestaña en la parte baja derecha se cuenta con la ventana de la validación cruzada que nos permite conocer el error en la estimación para el parámetro que se esté analizando (**Figura 26**).

En la segunda pestaña correspondiente a la malla de estimación y al mapa de ponderación se puede generar automáticamente la MEP si se presiona directamente el botón de graficar o en caso de contar con una malla generada previamente con un archivo que contenga la estructura de la **Tabla 4**, se puede optar por el botón de abrir malla de estimación.

Tabla 4.- Estructura del archivo de la malla de estimación

Columna	Nombre	Descripción	Comentarios
1	Clave	Identificador único	≤ 8 caracteres
2,3	X, Y	Coordenadas en UTM	≤ 2 decimales

Si se tomó esta segunda opción, después de abrir el archivo con la malla de la estimación, se deberá de presionar el botón de graficar.

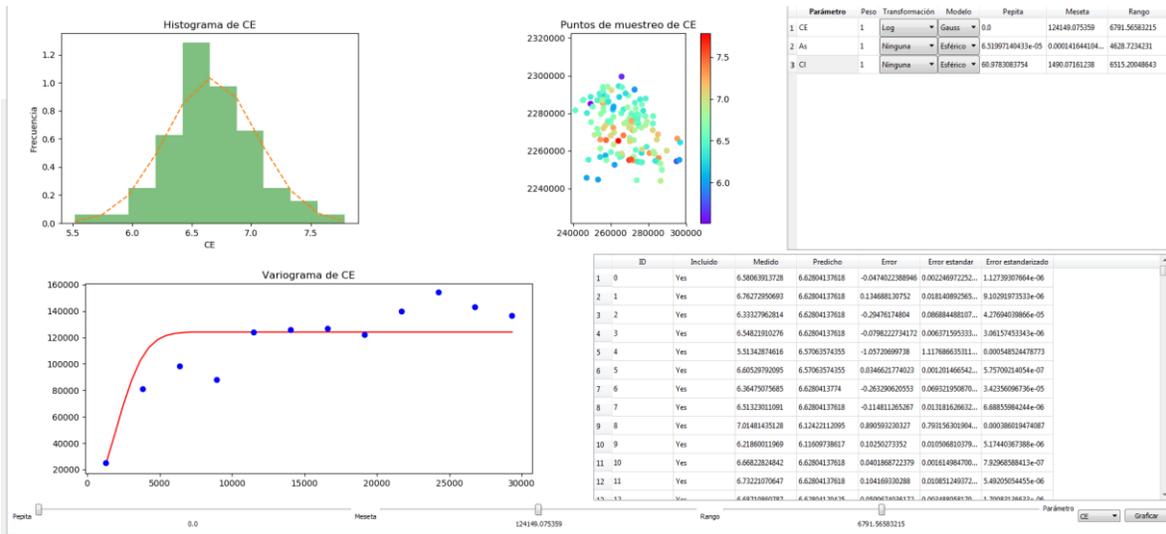


Figura 26.- Actualización de los datos después del cambio de una transformación y/o cambio de modelo matemático

Se puede observar un ejemplo de MEP (Figura 27) y con apertura de archivo (Figura 28), en ambos casos además de la malla de estimación, se genera del lado derecho la MER con una representación nombrada mapa de ponderación donde se muestran con colores los valores calculados en los nodos de la malla de estimación de la FS que permite observar las zonas de mayor y menor densidad de puntos de muestreo.

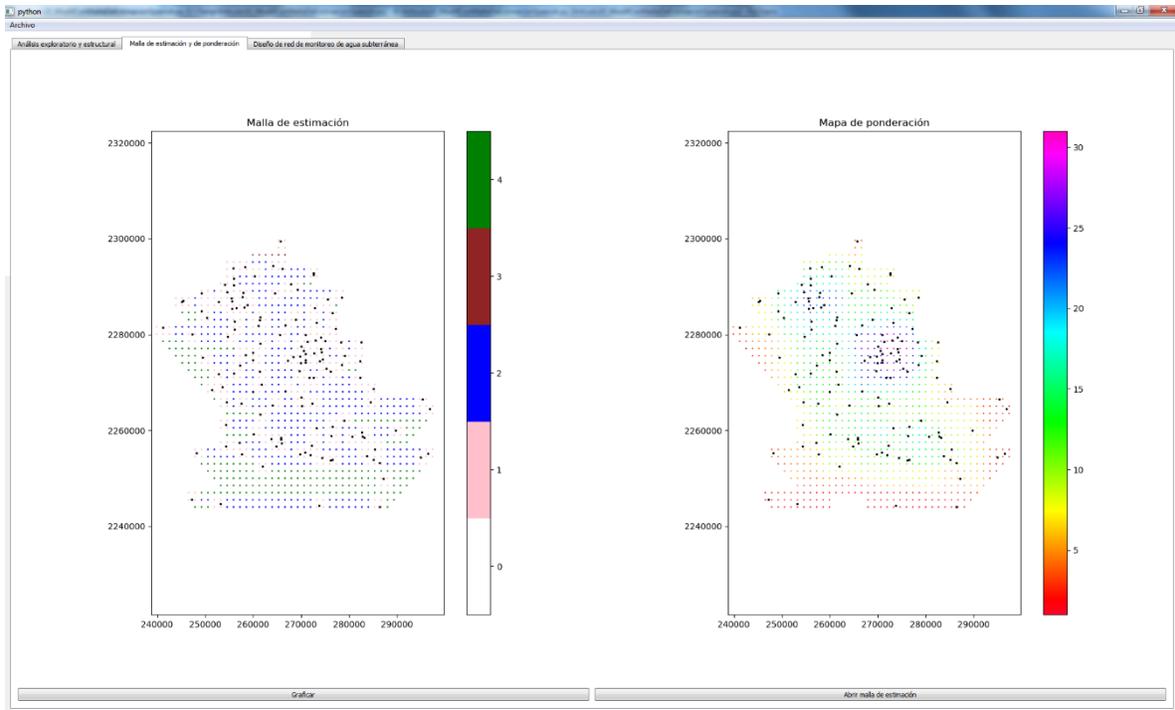


Figura 27.- Representación de la MEP y la MER

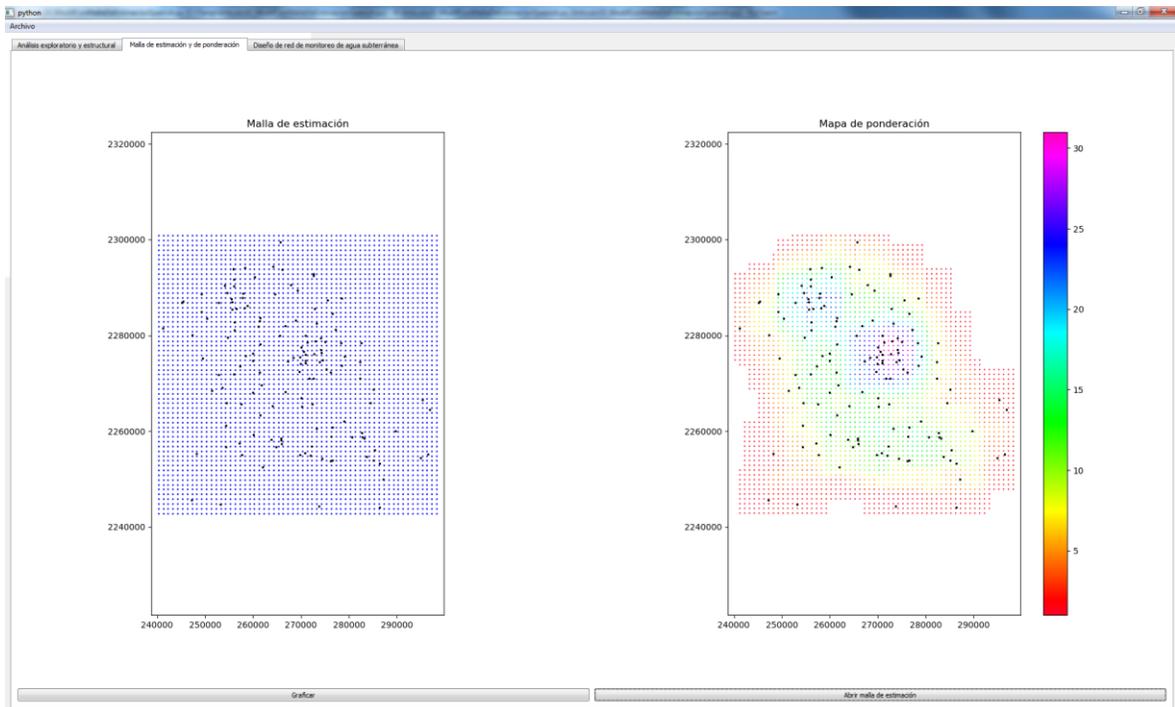


Figura 28.- Representación de la Malla de estimación proveniente de un archivo y de la MER correspondiente

Finalmente, en la última pestaña se observa la gráfica de distribución de puntos de muestreo con su correspondiente orden de prioridad, así como la gráfica comparativa de la selección efectuada, contra el escenario peor y óptimo de acuerdo con los valores establecidos para los modelos en cada parámetro y los pesos para el FK y la FS que se deberán de establecer en la parte superior derecha de esta pestaña. Después de esto se presiona el botón de generar red y después de hacer los cálculos correspondientes, se mostrarán las etiquetas de los primeros 10 elementos de prioridad de la red de monitoreo. En caso de que se deseen ver más o menos elementos, se puede activar la opción de ligado y utilizar la barra de desplazamiento para indicar la cantidad de elementos que se quieren etiquetar, la actualización se hará de manera automática y en eso consiste la función de ligado. En caso de que no se desee utilizar esta función se deja inactiva y se elige el número de elementos a etiquetar y posteriormente se presiona el botón de redibujar la red, lo que permitirá el etiquetado del número de elementos indicad.

Si se desee reasignar el peso de FK y la FS, se puede utilizar la barra de desplazamiento de la parte superior derecha y una vez especificados los nuevos valores, se presiona el botón de generar red (**Figura 29**).

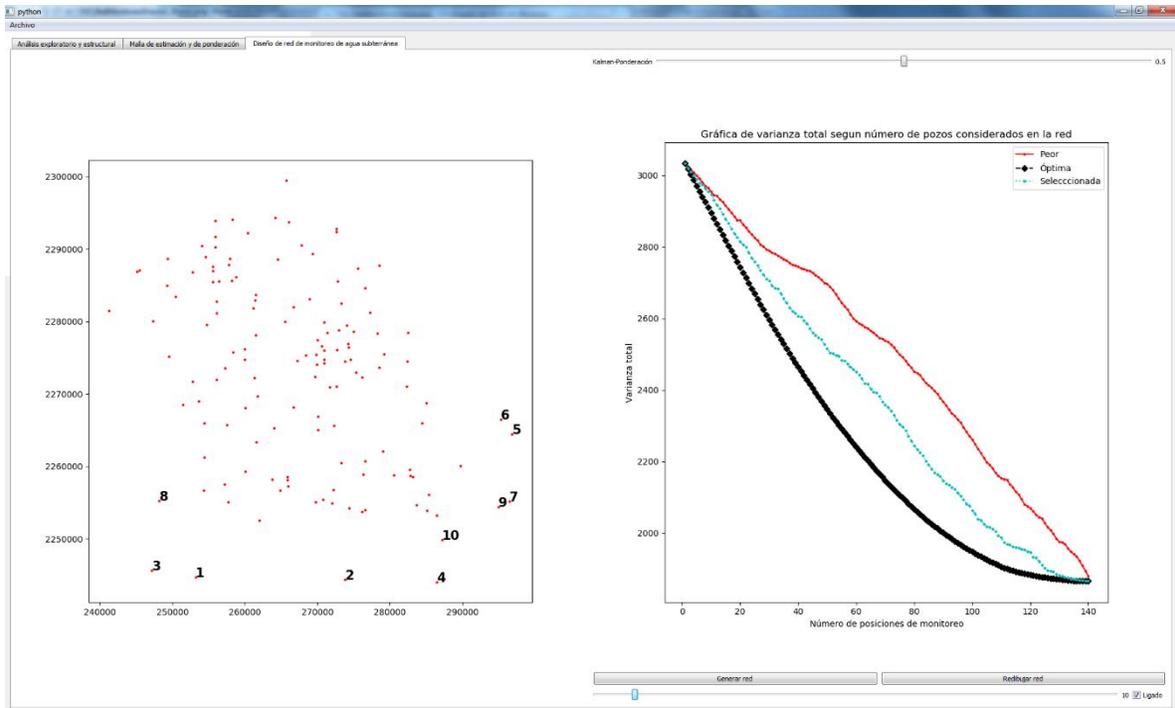


Figura 29.- Representación de la red de monitoreo de agua subterránea y la gráfica comparativa de la seleccionada contra la peor y la óptima

CAPÍTULO 4 RESULTADOS Y DISCUSIONES

4.1 Uso de la metodología y herramienta desarrollada

La metodología propuesta se aplicó al diseño óptimo de redes de monitoreo de calidad de agua subterránea orientada a caracterizar la distribución espacial de la zona de estudio tomando como referencia solo tres parámetros de la base de datos presente en Júnez-Ferreira *et al.* (2016a): arsénico, cloruro y conductividad eléctrica. Los resultados del análisis geoestadístico a los datos del caso de estudio generados por la herramienta de manera automática para cada uno de los parámetros utilizando modelos de variograma esférico se presentan en la Tabla 1. Con esta información, se generaron dentro de la herramienta las matrices de covarianza tomando como referencia la malla de estimación suavizada y acotada mediante los procesos de MEP y MER.

Tabla 1. Modelos y parámetros del variograma utilizados para la generación de la red de monitoreo de agua subterránea

Parámetro	Modelo	Pepita (u^2)*	Meseta (u^2)*	Rango (m)
Conductividad eléctrica	Esférico	0	125097	6792
Arsénico	Esférico	6.5×10^{-5}	0.000142	4685
Cloruro	Esférico	60.98	1490	6515

* u en (mg/L) para Arsénico and Cloruro, en (μ mhos/cm) para la Conductividad eléctrica

4.2 Comparativo de metodologías

En la metodología propuesta por Júnez-Ferreira *et al.* (2016a) se hace una selección de puntos de muestreo que permiten minimizar el error en la estimación de la varianza de los parámetros que fueron considerados en la generación de red de la red, además de la opción de generar zonas prioritarias de acuerdo con la calidad del agua subterránea. La optimización de este proceso es generada mediante el filtro de Kalman para calcular la PCT. La comparación entre Júnez-Ferreira *et al.* (2016a) y la metodología propuesta se realizó dando diferentes porcentajes de peso al filtro de Kalman y al mapa de ponderación. En el primer

caso (100 – 0), el 100 indica el porcentaje que fue asignado a para el cálculo de la PCT con el filtro de Kalman y el porcentaje de 0 corresponde a la FS, con lo cual se puede observar el resultado obtenido utilizando la metodología de Júnez-Ferreira *et al.* (2016a).

En la **Figura 30** se muestra el porcentaje de reducción que se va obteniendo en la varianza en el error en la estimación según se van incrementando el número de puntos de muestreo a la red de monitoreo para cada uno de los casos que se están analizando. En esta se puede ver que la reducción más rápida se consigue según lo esperado para el caso de “100 – 0”, por otro lado, con fines de conocer los casos extremos, se cuenta con el denominado “peor”, donde estratégicamente durante el proceso de selección de los puntos de muestreo, se fueron agregando en el orden en el que contribuía con el menor decremento del error en la varianza. Entre estos casos extremos podemos observar el comportamiento de la función mediante la asignación de valores diferentes a la FS. Para conocer la respuesta en diferentes casos, se asignaron 6 diferentes porcentajes de peso a PCT y a la FS para el análisis de sus respuestas tanto en la reducción del error, así como en la distribución espacial de los elementos obtenidos. Los porcentajes asignados de PCT y FS respectivamente en cada caso fueron de “100 – 0”, “99.999 – 0.001”, “99.995 – 0.005”, “99.99 – 0.01”, “50 – 50” y “0 – 100”, donde se analizó la reducción del error para los 140 puntos de muestreo. De manera espacial, se analizó la distribución señalando el orden de prioridad asignado para los primeros 10, 30 y 50 puntos de muestreo, dejando de lado caso denominado “peor”, ya que este solo se estableció como referencia para contrastar los errores que se obtienen en la asignación de diferentes pesos en los diferentes casos.

En síntesis se analizaron siete casos: 1) $V = 1$ y $W = 0$ (“100 – 0”), 2) $V = 99.999$ y $W = 0.001$ (“99.999 – 0.001”), 3) $V = 99.995$ y $W = 0.005$ (“99.995 – 0.005”), 4) $V = 99.99$ y $W = 0.01$ (“99.99 – 0.01”), 5) $V = 0.5$ y $W = 0.5$ (“50 – 50”), 6) $V = 0$ and $W = 1$ (“0 – 100”) y en el último caso el proceso de optimización se aplicó a la inversa, de tal manera que la prioridad más alta fue asignada a los

puntos de muestreo con mayor error en la estimación del error en la varianza (caso “peor”). Para los siete casos $\alpha_w = \alpha_p = 1$.

En la **Figura 31a** se puede observar el caso con el porcentaje de “100 – 0”, el cual corresponde a la metodología de JÚnez-Ferreira *et al.* (2016a), donde en las primeras 10 posiciones no tocan los puntos extremos, la selección está focalizada principalmente en la zona central de la nube de puntos de estaciones de muestreo, cuando se incrementa a 30 posiciones, se puede observar cómo se incrementa la cobertura de la zona, sin embargo, las zonas más extremas aún no son cubiertas. Cuando se consideran las 50 se logra una mayor densidad en la zona central y se logra una cobertura parcial de las zonas extremas, aunque se sigue observando que existe una mayor prioridad en las zonas con mayor densidad de puntos de estaciones de muestreo. El análisis de la asignación del orden de prioridad de manera espacial resulta primordial, ya que, si bien a cada uno de los puntos de estaciones de muestreo se les asigna una prioridad, dentro de los objetivos de la generación de redes de monitoreo se encuentra la obtención de la mayor cantidad posible de información con la menor cantidad de puntos reduciendo el error en la estimación, además de estas consideraciones, se podría añadir a esto, que la ubicación espacial de las estaciones de muestreo permitan tener la mayor cobertura posible y este último punto es el que se busca mejorar con la propuesta de modificación de la metodología de JÚnez-Ferreira *et al.* (2016a), por lo que el comparativo de los resultados obtenidos dando diferentes pesos a la PCT y a la FS, permite observar las diferencias en la distribución espacial de las estaciones de muestreo.

El resultado de la asignación de los pesos de “99.999 – 0.001” se puede observar en la **Figura 31b**, donde en las diez primeras posiciones, la distribución en principio se carga más a la parte perimetral comparativamente con lo observado en la **Figura 31a**, así como un mayor porcentaje de puntos de muestreo a la parte inferior de la zona de distribución. El contar con una selección de puntos de muestreo en la zona más perimetral, permite abarcar una zona más amplia con la misma cantidad de puntos y como se observó en la **Figura 30**, sin aumentar significativamente el error (menor al 0.0013% en las primeras 50 posiciones). Con

el análisis de las primeras treinta posiciones, ya se observa una distribución de los puntos de muestreo bastante similar, sin embargo, la parte baja con el punto de muestreo con una prioridad de 30 se cuenta con mejor cobertura con esta asignación de pesos. Al llegar a las cincuenta posiciones nos encontramos con una cobertura prácticamente igual a la que se logra en la **Figura 31a**, con la única diferencia como resulta lógico, que el orden de prioridad para cada uno de los puntos de muestreo en algunos casos difiere de manera importante.

Los resultados de la asignación de los pesos de “99.995 – 0.005”, se pueden observar en la **Figura 31c**, donde el peso de la FS resulta aún más evidente en los resultados. Observando las diez primeras posiciones, ocho de ellas se encuentran en la parte inferior de los puntos de muestreo y solo dos de los puntos (las posiciones 7 y 10) se encuentra en la mitad superior de la zona de puntos, lo que no resulta en una distribución espacial muy apropiada para lograr una mayor cobertura, sin embargo, si se analizan las treinta primeras posiciones, la distribución espacial logra una cobertura incluso más amplia que la **Figura 31a** y **Figura 31b** al considerar puntos de muestreo aún más extremos que los anteriormente citados y finalmente en las cincuenta primeras posiciones logra una distribución bastante similar que las Figuras anteriores, mejorando el alcance espacial al considerar una mayor cantidad puntos de muestreo en zonas más extremas. Al igual que en el caso anterior, el aumento en el error es mínimo (menos del 0.07% en las primeras cincuenta posiciones), por lo que representa una mejora al ampliar la cobertura de muestreo.

En la **Figura 31d** se puede observar el resultado obtenido considerando los valores de “99.99 – 0.01”, donde nueve de las primeras diez posiciones de muestreo se encuentran en la parte inferior con una distribución perimetral, dejando casi la mitad del área de muestreo sin seleccionar algún punto de muestreo. Con treinta posiciones seguimos viendo un patrón de selección perimetral de posiciones mejorando ligeramente la distribución de cuando se consideraban solo diez elementos, ya que ahora se puede ver un cubrimiento en la parte superior, aunque se ve claramente una zona en el centro en la que no se

ha seleccionado ningún punto de muestreo. Finalmente, con la asignación de las cincuenta primeras posiciones, se observa una distribución uniforme, muy similar a la obtenida en la **Figura 31c**, con la diferencia más notable de la posición 44 que no había sido asignada en ninguna de las Figuras anteriores, con un error cercano al 0.18% en las primeras cincuenta posiciones, lo que representa un error mayor que en la **Figura 31b** y **Figura 31c** sin que represente una mejoría notable en la distribución de la prioridad de los puntos de muestreo si se considera el incremento del error (menor a 0.07% en la **Figura 31c** y menor a 0.18 en la **Figura 31d**).

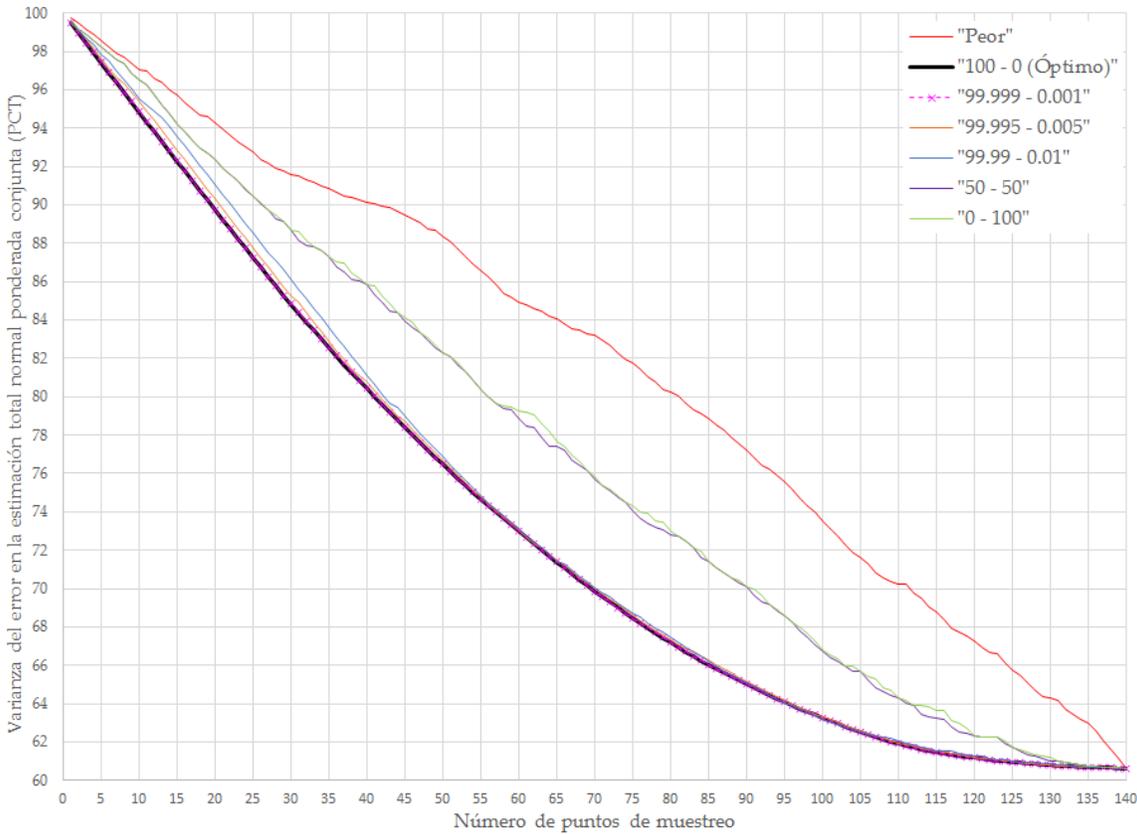
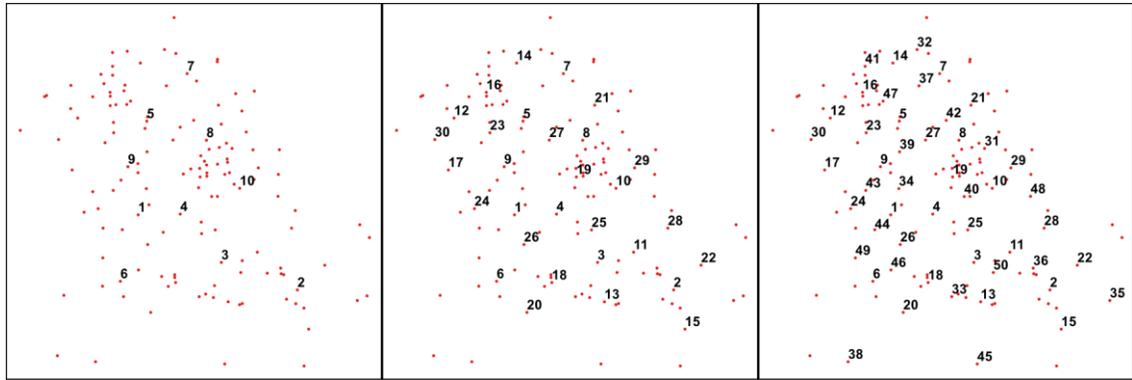


Figura 30.- Gráfica comparativa de los diferentes casos tomando los resultados obtenidos de PCT (%) vs número de puntos de muestreo

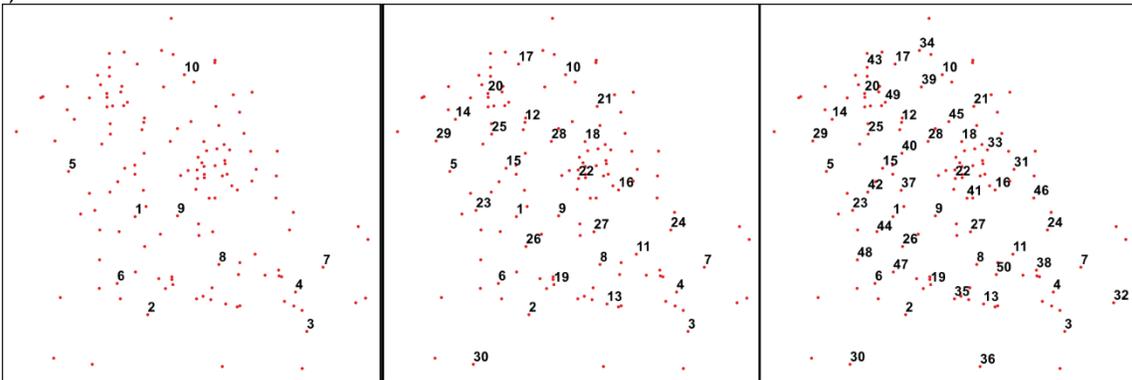
En la **Figura 31e**, donde la asignación de pesos es de "50 – 50", se puede observar claramente que la FS tiene una mayor influencia en el resultado. En las diez primeras posiciones se ve un efecto similar a la **Figura 31d**, donde solo se tiene un punto de muestreo en la mitad superior y la selección se da exclusivamente en la parte perimetral, sin embargo, si bien es cierto que la

selección en las treinta y cincuenta posiciones ya consideran la parte superior de la zona de muestreo, se sigue manteniendo un patrón de selección totalmente perimetral, lo cual deja una zona en el centro claramente sin puntos de muestreo lo que también repercute de manera importante al incrementar el error hasta más del 33.8%.

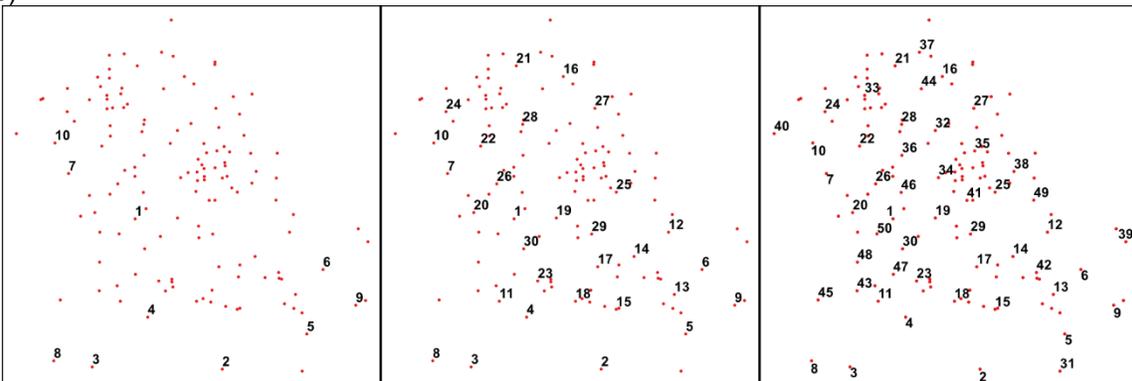
Finalmente, en la **Figura 31f**, donde se la asignación fue de “0 – 100”, se puede observar prácticamente el mismo patrón que lo observado en la **Figura 31e**, un patrón totalmente perimetral en las 10, 30 y 50 primeras posiciones y se hace evidente la necesidad de considerar la FS como tal, es decir, suplementaria a la ya existente, ya que de no ser así, los resultados resultan pobres para distribuciones de este tipo. El error se mantiene al igual que en la anterior, hasta más del 33.8 %, prácticamente su comportamiento es a la mitad de lo observado entre el óptimo y el peor.



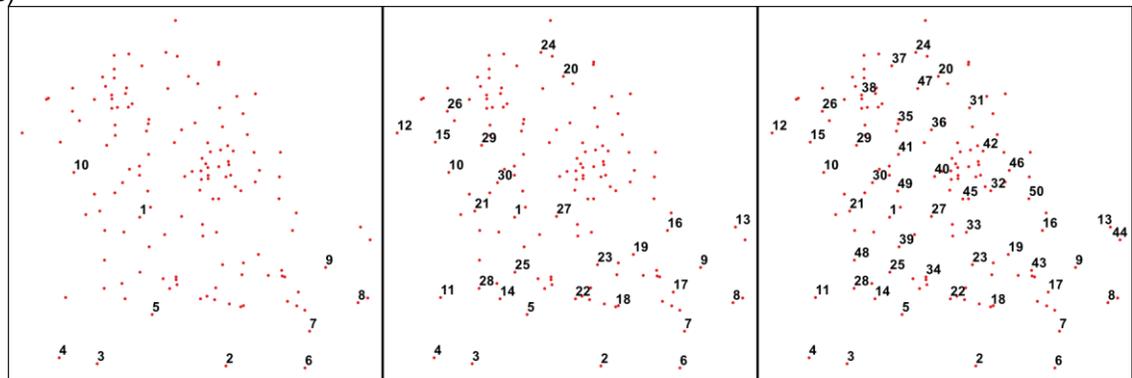
(a)



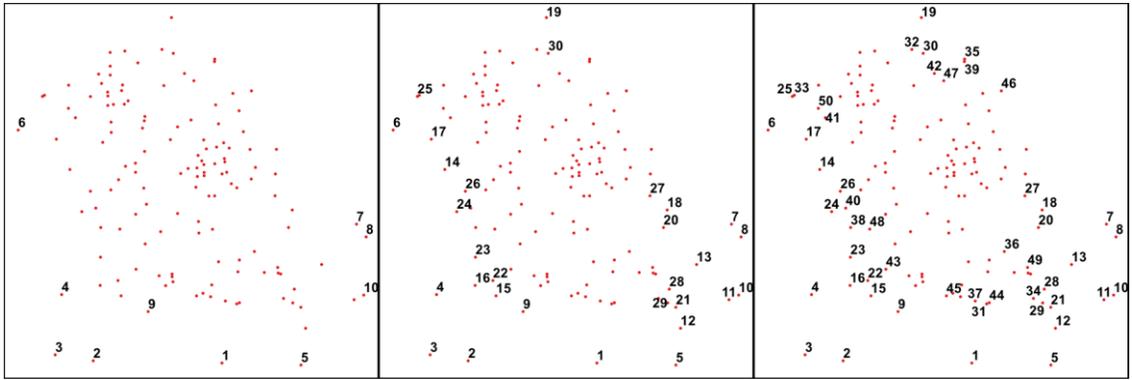
(b)



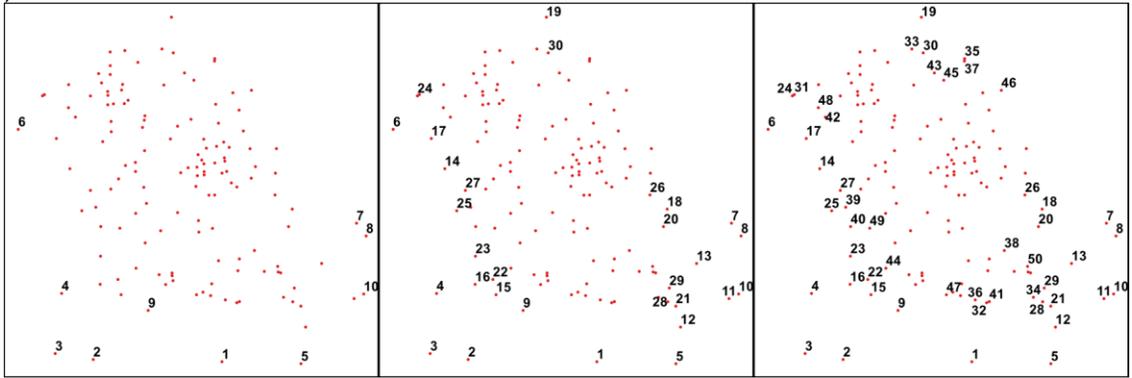
(c)



(d)



(e)



(f)

Figura 31.- Redes de monitoreo de agua subterránea y su orden de prioridad para los primeros 10, 30 y 50 puntos de muestreo para los casos: a) "0-100", b) "0.001-99.999", c) "0.005-99.995", d) "0.01-99.99", e) "50-50" y f) "100-0".

CAPÍTULO 5 CONCLUSIONES

Varias metodologías para el diseño óptimo de redes de monitoreo del agua subterránea están basadas en la reducción de la estimación del error en la varianza de una o más variables consideradas en el diseño; dichas estimaciones se obtienen a partir de una red de puntos de estimación. En la mayoría de los casos, se genera una red con límites extremos a partir de las mínimos y máximos en las coordenadas que se tienen de los puntos de muestreo, sin embargo esto genera redes de puntos de estimación con forma rectangular o cuadrada que no se ajustan necesariamente a las particularidades de las formas que se generan a partir de la distribución espacial de los puntos de muestreo, dando como resultado que en muchos de los casos la densidad sea muy variable y en algunos casos con escasez de puntos de muestreo, sobre todo en las zonas perimetrales. El resultado de la heterogeneidad de la distribución espacial y la forma rectangular o cuadrada de la red de puntos de estimación originan la asignación de prioridades más altas a aquellos puntos de muestreo que tienen un mayor número de nodos de estimación cercanos a su posición contrariamente a lo que ocurre con los puntos de muestreo que se encuentran ubicados en la periferia de la malla de estimación (Júnez-Ferreira *et al.* 2016a).

Cuando el diseño está solamente orientado a la reducción de la estimación del error en la varianza, sin la asignación de pesos específicos a los nodos de la malla de estimación, la red de monitoreo puede tener algunas deficiencias en la cobertura de algunas zonas prioritarias o puntos de muestreo con información valiosa. Considerando estos puntos, resulta importante la evaluación de alternativas en el diseño de la malla de estimación que consideren estas áreas de interés y se minimice la pérdida de información en la estimación con la red de monitoreo óptima seleccionada.

La metodología propuesta incluye la generación y programación de un algoritmo para la generación de la malla de estimación que toma en consideración la forma de la distribución espacial de los puntos de muestreo con el propósito de evitar la extrapolación. La inserción de la FS, así como el despliegue de los

resultados obtenidos del proceso de optimización permite el análisis de la distribución espacial de los puntos de muestreo seleccionados de manera que se pueda identificar la mejor cobertura, sobre todo en la zona perimetral de la zona de muestreo sin que incremente significativamente el error en la estimación. La FO considera la asignación de pesos a los puntos de muestreo, lo cual es de gran importancia cuando se ha identificado uno o más puntos que resultan de importancia y se desea conservar como parte de la red de monitoreo.

A través de esta metodología y su implementación a través de una herramienta independiente, es posible el análisis de diferentes alternativas en el diseño de la red de monitoreo mediante la asignación de diferentes pesos a la PCT y a la FS, con su representación puntual y la actualización de sus prioridades cuando se cambian los pesos. Además de eso, se puede observar la gráfica que representa el error en el caso óptimo, el peor y los datos generados a partir de la FO de acuerdo con los pesos seleccionados. La parte de la FO se actualiza al igual que las prioridades, con la asignación de nuevos pesos a la PCT y a la FS, permitiendo de esta manera hacer el comparativo con los escenarios extremos y evaluar cuales valores permiten tener una buena distribución de puntos sin incrementar drásticamente el error.

Los flujos de agua subterránea en acuíferos heterogéneos pueden ser interpretados mediante metodologías específicas que se basan en la evolución geoquímica; la hidrogeoquímica es útil en la resolución de cuestiones relacionadas con el origen y la atenuación de contaminantes geogénicos y antropogénicos. La metodología propuesta considera técnicas estadísticas para analizar la distribución espacial de la calidad de agua en ciertos parámetros como resultado de una campaña de muestreo.

Esta metodología permite la selección de los puntos de muestreo más adecuados a partir de la información obtenida de una campaña de muestreo y los parámetros seleccionados con base en los objetivos de la red de monitoreo mediante la asignación directa de pesos a los puntos de muestreo y de ser necesario a los parámetros de mayor interés. Una vez que la red de monitoreo ha sido diseñada, los datos obtenidos pueden analizarse y ajustarse a uno o más

criterios tales como el presupuesto con el que se cuenta para hacer el monitoreo, el tiempo disponible, las distancias y las dificultades encontradas en campo para muestrear todos los puntos de la red de monitoreo. La herramienta de software generada, permite el análisis y ajuste de los criterios y diseños de la red de manera fácil teniendo en todo momento los dos escenarios (óptimo y peor) de referencia para tomar en cuenta que tan alejado y cercano está de los dos extremos. Esta propuesta metodológica incluye una función que permite el diseño de una red de monitoreo de agua subterránea con una mejor cobertura y distribución de los puntos de muestreo en la periferia de la zona de estudio.

La propuesta metodológica propuesta e implementada en una herramienta de software independiente, facilita a los tomadores de decisiones, la selección de los puntos de muestreo para la generación de las redes de monitoreo en función de los objetivos de la red.

Bibliografía

- 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. *Environmental Earth Sciences*, 76(12), 1–15. <https://doi.org/10.1007/s12665-017-6767-6>
- Andricevic, R. (1990). Cost-effective network design for groundwater flow monitoring. *Stochastic Hydrology and Hydraulics*, 4(1), 27–41. <https://doi.org/10.1007/BF01547730>
- Baalousha, H. (2010). Assessment of a groundwater quality monitoring network using vulnerability mapping and geostatistics: A case study from Heretaunga Plains, New Zealand. *Agricultural Water Management*, 97(2), 240–246. <https://doi.org/10.1016/j.agwat.2009.09.013>
- Batelaan, O., & Smedt, F. De. (2007). GIS-based recharge estimation by coupling surface – subsurface water balances, 337–355. <https://doi.org/10.1016/j.jhydrol.2007.02.001>
- Chen, Y., Smith, L., & Beckie, R. (2012). Modeling of strategies for performance monitoring of groundwater contamination at sites underlain by fractured bedrock. *Journal of Contaminant Hydrology*, 134–135, 37–53. <https://doi.org/10.1016/j.jconhyd.2012.03.009>
- Díaz Viera, M. A. (2002). Geoestadística aplicada. *Instituto de Geofísica*, 144. Retrieved from <http://mmc2.geofisica.unam.mx/cursos/geoest/GeoEstadistica.pdf>
- DOF, D. O. de la F. (2012). 26/3/2015 DOF - Diario Oficial de la Federación. Retrieved from http://dof.gob.mx/nota_detalle.php?codigo=5232952&fecha=08/02/2012
- Esquivel, J. M., Morales, G. P., & Esteller, M. V. (2015). Groundwater Monitoring Network Design Using GIS and Multicriteria Analysis. *Water Resources Management*, 29(9), 3175–3194. <https://doi.org/10.1007/s11269-015-0989-8>
- Giraldo Henao, R. (2011). Introducción a la geoestadística. Teoría y aplicación, 94.
- Herrera, G. S., & Pinder, G. F. (2005). Space-time optimization of groundwater quality sampling networks. *Water Resources Research*, 41(12), 25–49. <https://doi.org/10.1029/2004WR003626>
- Hosseini, M., & Kerachian, R. (2017a). A Bayesian maximum entropy-based methodology for optimal spatiotemporal design of groundwater monitoring networks. *Environ Monit Assess*, 189, 1–24. <https://doi.org/https://doi.org/10.1007/s10661-017-6129-6>
- Hosseini, M., & Kerachian, R. (2017b). A data fusion-based methodology for optimal redesign of groundwater monitoring networks. *Journal of Hydrology*, 552, 267–282. <https://doi.org/10.1016/j.jhydrol.2017.06.046>
- 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). <https://doi.org/10.3390/w10081063>
- Júnez-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. *Environmental Monitoring and Assessment*, 185(4), 3527–3549. <https://doi.org/10.1007/s10661-012-2808-5>

- Júnez-Ferreira, H. E., Herrera, G. S., Gonzalez-Hita, L., Cardona, A., & Mora-Rodriguez, J. (2016a). Optimal design of monitoring networks for multiple groundwater quality parameters using a Kalman filter: application to the Irapuato-Valle aquifer. *Environmental Monitoring and Assessment*, *188*(1), 39. <https://doi.org/10.1007/s10661-015-5036-y>
- Júnez-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. *Mathematical Geosciences*, *48*(1), 255–44.
- Kim, G.-B. (2015). Optimal distribution of groundwater monitoring wells near the river barrages of the 4MRRP using a numerical model and topographic analysis. *Environmental Earth Sciences*, *73*(9), 5497–5511. <https://doi.org/10.1007/s12665-014-3802-8>
- 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. *Journal of Hydrology*, *342*(3–4), 223–237. <https://doi.org/10.1016/j.jhydrol.2007.05.034>
- Kollat, J. B., Reed, P. M., & Maxwell, R. M. (2011). Many-objective groundwater monitoring network design using bias-aware ensemble Kalman filtering, evolutionary optimization, and visual analytics. *Water Resources Research*, *47*(2), 1–18. <https://doi.org/10.1029/2010WR009194>
- Loaiciga, B. H. A., Member, A., Member, A., & Rouhani, S. (1992). Review of Ground-Water Quality Monitoring Network Design. *Journal of Hydraulic Engineering (J HYDRAUL ENG-ASCE)*, *118*(1), 11–37.
- Luo, Q., Wu, J., Yang, Y., Qian, J., & Wu, J. (2016). Multi-objective optimization of long-term groundwater monitoring network design using a probabilistic Pareto genetic algorithm under uncertainty. *Journal of Hydrology*, *534*, 352–363. <https://doi.org/10.1016/j.jhydrol.2016.01.009>
- Nobre, R. C. M., Rotunno Filho, O. C., 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. *Journal of Contaminant Hydrology*, *94*(3–4), 277–292. <https://doi.org/10.1016/j.jconhyd.2007.07.008>
- Om Prakash, B. D. (2014). Optimal monitoring network design for efficient identification of unknown groundwater pollution sources. *International Journal of GEOMATE*, 785–790.
- Pearson, S., Berg, J., & Macdonald, M. (2011). Minnesota Groundwater Level Monitoring Network — Guidance Document for Network Development, (May).
- Preziosi, E., Petrangeli, a. B., & Giuliano, G. (2013). Tailoring groundwater quality monitoring to vulnerability: A GIS procedure for network design. *Environmental Monitoring and Assessment*, *185*(5), 3759–3781. <https://doi.org/10.1007/s10661-012-2826-3>
- 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. *Journal of Irrigation and Drainage Engineering*, *143*(8), 05017003. [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0001213](https://doi.org/10.1061/(ASCE)IR.1943-4774.0001213)
- Sizirici, B., & Tansel, B. (2015). Parametric fate and transport profiling for selective groundwater monitoring at closed landfills : A case study. *Waste Management*, *38*, 263–270.

- Uddameri, V., & Andruss, T. (2013). A statistical power analysis approach to estimate groundwater-monitoring network size in Victoria County Groundwater Conservation District, Texas. *Environmental Earth Sciences*, 71, 2605–2615. <https://doi.org/10.1007/s12665-013-2901-2>
- Uddameri, V., & Andruss, T. (2014). A GIS-based multi-criteria decision-making approach for establishing a regional-scale groundwater monitoring. *Environmental Earth Sciences*, 71(6), 2617–2628. <https://doi.org/10.1007/s12665-013-2899-5>
- Wang, T., Franz, T. E., Yue, W., Szilagyi, J., Zlotnik, V. A., You, J., ... Young, A. (2016). Feasibility analysis of using inverse modeling for estimating natural groundwater recharge from a large-scale soil moisture monitoring network. *Journal of Hydrology*, 533, 250–265. <https://doi.org/10.1016/j.jhydrol.2015.12.019>
- WHO, W. H. O. (2003). The Right to Water. *Health and Human Rights Publication Series*. Retrieved from http://www.who.int/water_sanitation_health/en/righttowater.pdf
- WWAP. (2016). *The United Nations World Water Development Report 2016: Water and Jobs*. Paris.
- Yakirevich, a., Pachepsky, Y. a., Gish, T. J., Guber, a. K., Kuznetsov, M. Y., Cady, R. E., & Nicholson, T. J. (2013). Augmentation of groundwater monitoring networks using information theory and ensemble modeling with pedotransfer functions. *Journal of Hydrology*, 501, 13–24. <https://doi.org/10.1016/j.jhydrol.2013.07.032>
- Yang, F., Cao, S., Liu, X., & Yang, K. (2008). Design of groundwater level monitoring network with ordinary kriging. *Journal of Hydrodynamics*, 20(3), 339–346. [https://doi.org/10.1016/S1001-6058\(08\)60066-9](https://doi.org/10.1016/S1001-6058(08)60066-9)
- Zhang, Y., Pinder, G. F., & Herrera, G. S. (2005). Least cost design of groundwater quality monitoring networks. *Water Resources Research*, 41(8), 1–12. <https://doi.org/10.1029/2005WR003936>
- Zhou, Y., Dong, D., Liu, J., & Li, W. (2013). Upgrading a regional groundwater level monitoring network for Beijing Plain, China. *Geoscience Frontiers*, 4(1), 127–138. <https://doi.org/10.1016/j.gsf.2012.03.008>