



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Optimización de redes de monitoreo
de la carga hidráulica utilizando métodos
geoestadísticos espacio-temporales

Dr. Hugo Enrique Júnez Ferreira



**Dra. Graciela del Socorro
Herrera Zamarrón**
Asesor

JURADO ASIGNADO:

Presidente: DR. FRANCISCO JAVIER APARICIO MIJARES

Secretario: DR. JOSÉ JOEL CARRILLO RIVERA

1^{er.} Vocal: DRA GRACIELA DEL SOCORRO HERRERA ZAMARRÓN

1^{er.} Suplente: DR. MARTÍN ALFONSO GUTIÉRREZ LÓPEZ

2^{do.} Suplente: DR. ALDO IVÁN RAMÍREZ OROZCO

Lugar o lugares donde se realizó la tesis:

JIUTEPEC, MORELOS

TUTOR DE TESIS:

DRA. GRACIELA DEL SOCORRO HERRERA ZAMARRÓN


FIRMA

A Gina

AGRADECIMIENTOS

A mi familia, de quien recibo todo lo que me hace feliz, son mi motivación para seguir adelante. Gracias por su amor infinito.

A la Dra. Graciela del Socorro Herrera Zamarrón, por su confianza, sus enseñanzas, su disponibilidad, apoyo y guía durante todos mis estudios de posgrado.

A los Doctores Francisco Javier Aparicio Mijares y José Joel Carrillo Rivera, que a lo largo de mis estudios doctorales realizaron críticas constructivas y aportaciones que han enriquecido considerablemente el trabajo de investigación, les agradezco mucho su interés y exigencias en mi formación.

A los Doctores Martín Alfonso Gutiérrez López y Aldo Iván Ramírez Orozco, por sus aportaciones como miembros del comité doctoral.

Al CONACyT, por la beca que me fue otorgada, con la cual pude concluir mis estudios de doctorado.

A toda persona no mencionada que me ayudó durante los estudios.

Resumen

La información recabada por redes de monitoreo del agua subterránea es extremadamente valiosa en el entendimiento del funcionamiento de la misma y su presencia en los acuíferos. La carga hidráulica es una variable que depende no sólo del espacio sino también del tiempo y de las propiedades (densidad) del agua, por lo que es necesario un diseño de una red de monitoreo espacio-temporal para capturar la evolución de la carga hidráulica. Sin embargo, el monitoreo extensivo en el espacio y/o tiempo puede representar costos considerables. En esta tesis se presenta una nueva metodología para el diseño espacio-temporal óptimo de redes de monitoreo de la carga hidráulica. La elección espacio-temporal de posiciones de monitoreo se realiza mediante el uso del filtro de Kalman estático (FK) combinado con un método de optimización secuencial. El FK requiere como entrada una matriz de covarianza espacio-temporal, que en esta tesis es obtenida de un análisis geoestadístico. En cada paso, se empleó un método de optimización de inclusiones sucesivas que selecciona la posición espacio-temporal que minimiza una función de la varianza del error en la estimación. Se demostró la metodología para el diseño y evaluación del desempeño en el futuro de una red de monitoreo óptima, para un caso sintético basado en un modelo numérico de flujo del acuífero Valle de Querétaro, en el que se contaba con un alto nivel de información. En otro caso sintético se evaluó una red de monitoreo pobremente distribuida en el espacio, y en donde además se contaba con un bajo nivel de información espacio-temporal para el análisis, esto se realizó con el fin de emular un caso real. Finalmente se aplicó la metodología en el rediseño de la red de monitoreo del nivel del agua para el acuífero Valle de Querétaro con el objetivo de seleccionar de entre un conjunto de posiciones y tiempos de monitoreo, aquéllos que minimizan la redundancia espacio-temporal. Los resultados de la primera aplicación muestran que el 99% de la reducción máxima posible del valor de la raíz cuadrada de la varianza promedio del error en la estimación en el área de estudio, se obtiene cuando se monitorea en tan sólo 477 de las 2616 posiciones espacio-temporales disponibles. Para el segundo caso, esta reducción se alcanza con las primeras 37 de 912 disponibles. Para el caso real, el monitoreo en las 178 posiciones con la más alta prioridad ofrece prácticamente el mismo nivel de información que el emplear las 418 posiciones de monitoreo disponibles. Estos resultados muestran que el método propuesto es exitoso en propagar la información en el espacio y en el tiempo, lo que se traduce en una reducción significativa del costo asociado al monitoreo, si consideramos que éste es proporcional al número de posiciones monitoreadas. El error promedio en la estimación obtenida con la red de monitoreo óptima diseñada para el caso sintético con alto nivel de información de la carga hidráulica fue menor a 0.58 m, mientras que para el caso sintético con bajo nivel de información y para el caso real los errores fueron menores a 24.12 m y 27.53 m, respectivamente. Para el caso de un alto nivel de información de la variable de interés y una adecuada distribución en el espacio de la red de monitoreo fue posible obtener estimaciones de la carga hidráulica muy cercanas a las simulaciones de un modelo numérico de flujo, a pesar de no emplear las ecuaciones que describen la física del problema. Estos resultados confirman el potencial de las técnicas geoestadísticas combinadas con estimación espacio-temporal empleando el FK en el diseño de redes de monitoreo óptimas y en la evaluación de redes de monitoreo existentes.

Abstract

The information collected by groundwater monitoring-networks is extremely valuable in understanding functioning and presence of water in aquifers. As hydraulic-head (HH) is a variable that depends not only on space but also on time and on water properties (density), a space-time monitoring-network design is necessary to capture HH evolution in aquifers. However, extensive monitoring in space and/or time may involve considerable costs. This thesis presents a new methodology for the optimal design of space-time HH monitoring-networks. The choice of the space-time monitoring positions is done using a static Kalman filter (KF) combined with a sequential optimization method. The KF requires as input a space-time covariance matrix, which in this thesis is derived from a geostatistical analysis. A successive inclusions optimization method that selects the space-time position that minimizes a function of the estimate error variance, in each step, was used. The methodology was demonstrated by applying it to the design and evaluation of the future performance of an optimal monitoring-network for a synthetic case based on a numerical flow model of the Valle de Querétaro aquifer, which had a high level of information. A poorly-distributed in space monitoring-network was evaluated in another synthetic case, which had a low space-time level of information for the analysis, this was done in order to emulate a real case. Finally, the methodology was applied in the redesign of the water level monitoring-network for the Valle de Querétaro aquifer in order to select from a set of monitoring positions and times, those that minimize spatiotemporal redundancy. The first application results show that 99% of the maximum possible reduction of the square root of the average estimate error variance in the study area is obtained when monitoring in only 477 of the 2616 space-time available positions. For the second case, this reduction is achieved with the first 37 of 912 available. For the real case, monitoring in the 178 positions with the highest priority offers virtually the same level of information than when monitoring the 418 positions. These results show that the proposed method is successful in propagating the information in each datum in space and time, which implies a significant reduction in associated monitoring costs if we consider that it is proportional to the number of monitoring positions. The average estimate error obtained with the optimal monitoring-network designed for the synthetic case with high level of information was lower than 0.58 m, while errors for the synthetic case with low level of information and for the real case were lower than 24.12 m and 27.53 m, respectively. In the case of a high level of information for the variable of interest and a proper spatial distribution of the monitoring-network it was possible to obtain HH estimates very close to the simulations of a numerical flow model, in spite of not using the equations that describe the physics of the problem. These results confirm the potential of geostatistical techniques combined with space-time estimates using the KF in the design of optimal monitoring-networks and evaluation of existing monitoring-networks.

ÍNDICE

1. INTRODUCCIÓN	1
1.1 Antecedentes	1
1.2 Objetivos	11
1.3 Estructura de la tesis.....	11
2. LA CARGA HIDRÁULICA	13
2.1 Definición.....	13
2.2 Consideraciones para la correcta interpretación de mediciones de la carga hidráulica ...	15
2.3 Presencia de tendencia en la carga hidráulica	17
3. TEORÍA GEOESTADÍSTICA Y DEL FILTRO DE KALMAN.....	20
3.1 Conceptos básicos de estadística.....	20
3.2 Geoestadística.....	22
3.3 Filtro de Kalman.....	30
3.3.1 <i>Filtro de Kalman estático</i>	31
4. METODOLOGÍA GEOESTADÍSTICA PARA EL DISEÑO ESPACIO-TEMPORAL ÓPTIMO DE REDES DE MONITOREO DE LA CARGA HIDRÁULICA.....	33
4.1 Definición del objetivo del monitoreo	33
4.2 Análisis geoestadístico espacio-temporal.....	33
4.3 Optimización de la red de monitoreo	34
4.4 Criterio para determinar el número total de posiciones de monitoreo	35
5. DISEÑO DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON ALTO NIVEL DE INFORMACION ESPACIO-TEMPORAL.....	36
5.1 Modelo numérico de flujo	36
5.2 Objetivo del monitoreo.....	50
5.3 Análisis geoestadístico espacio-temporal de la carga hidráulica	51
5.4 Optimización de la red de monitoreo	55
5.5 Determinación del número total de posiciones de monitoreo	56
5.6 Red de monitoreo óptima	57
5.7 Estimaciones y varianzas del error en la estimación para el período de diseño.....	59
5.8 Evaluación del desempeño de la red de monitoreo en el futuro.....	66
5.9 Evaluación del desempeño de la red de monitoreo a través del tiempo.....	71
5.10 Evaluación del desempeño de la red de monitoreo en la captura de la distribución espacial	72
5.11 Conclusiones y recomendaciones.....	74
6. EVALUACIÓN DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON BAJO NIVEL DE INFORMACION ESPACIO-TEMPORAL	76
6.1 Base de datos de la carga hidráulica.....	76
6.2 Objetivo del monitoreo.....	76
6.3 Análisis geoestadístico espacio-temporal de la carga hidráulica	77
6.4 Optimización de la red de monitoreo	81

6.5 Determinación del número total de posiciones de monitoreo	82
6.6 Programa de monitoreo óptimo.....	83
6.7 Estimaciones y varianzas del error en la estimación para el período de diseño.....	85
6.8 Evaluación del desempeño de la red de monitoreo en el futuro.....	91
6.9 Evaluación del desempeño de la red de monitoreo a través del tiempo.....	98
6.10 Evaluación del desempeño de la red de monitoreo en la captura de la distribución espacial	106
6.11 Conclusiones y recomendaciones.....	108
7. DISEÑO DE UNA RED DE MONITOREO DEL NIVEL DEL AGUA PARA EL ACUÍFERO VALLE DE QUERÉTARO	109
7.1 Base de datos del nivel del agua subterránea	109
7.3 Análisis geoestadístico espacio-temporal del nivel del agua subterránea.....	111
7.4 Optimización de la red de monitoreo	118
7.5 Determinación del número total de posiciones de monitoreo	119
7.7 Estimaciones y varianzas del error en la estimación para el período de diseño.....	122
7.8 Conclusiones y recomendaciones.....	131
8. CONCLUSIONES Y RECOMENDACIONES.....	133
9. BIBLIOGRAFÍA.....	135
ANEXO A. Revisión del estado del arte del diseño de redes de monitoreo de la calidad del agua subterránea.....	142

ÍNDICE DE FIGURAS

Fig. 2.1 Carga hidráulica en piezómetro	14
Fig. 2.2 Superficie potenciométrica y tipo de pozos.....	15
Fig. 5.1 Malla de elemento finito y condiciones de frontera del modelo	37
Fig. 5.2 Coeficientes de infiltración usados en la estimación de la recarga vertical natural	39
Fig. 5.3 Recarga vertical inducida asignada al modelo (m/d).....	41
Fig. 5.4 Pozos localizados en diferentes zonas del acuífero	42
Fig. 5.5 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo L-246	43
Fig. 5.6 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo 644-A.....	44
Fig. 5.7 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo L-107	45
Fig. 5.8 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo 2082	46
Fig. 5.9 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo 1792	47
Fig. 5.10 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo 752	48
Fig. 5.11 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo 976	49
Fig. 5.12 Simulación y estimación de la carga hidráulica para el corte A-A' de la Fig. 5.4	50
Fig. 5.13 Área de estudio y localización de los puntos de monitoreo/estimación	51
Fig. 5.14 Semivariograma espacio-temporal muestral de la carga hidráulica	52
Fig. 5.15 Proyección del semivariograma espacio-temporal muestral de la carga hidráulica sobre el plano semivariograma-intervalos espaciales (círculos) y modelo de semivariograma ajustado (cuadros)	53
Fig. 5.16 Proyección del semivariograma espacio-temporal muestral de la carga hidráulica sobre el plano semivariograma-intervalos temporales (círculos) y modelo de semivariograma ajustado (cuadros)	54
Fig. 5.17 Varianza Total vs Número de posiciones de monitoreo espacio-temporal	56
Fig. 5.18 Raíz cuadrada (Varianza total/Posiciones de estimación espacio-temporal) vs Número de posiciones de monitoreo espacio-temporal	57
Fig. 5.19 Posiciones de monitoreo para cada mes	58
Fig. 5.20 Número de veces de monitoreo para cada pozo en el programa de monitoreo óptimo espacio-temporal (arriba) y nombre del pozo (abajo).....	59
Fig. 5.21 Estimación de la carga hidráulica para agosto del primer año	61
Fig. 5.22 Varianzas del error en la estimación y orden de prioridad de los pozos para agosto del primer año	62
Fig. 5.23 Estimación de la carga hidráulica para julio del segundo año.....	63
Fig. 5.24 Varianzas del error en la estimación y orden de prioridad de los pozos para julio del segundo año	64
Fig. 5.25 Estimación de la carga hidráulica para julio del tercer año.....	65
Fig. 5.26 Varianzas del error en la estimación y orden de prioridad de los pozos para julio del tercer año.....	66

Fig. 5.27 Estimación de la carga hidráulica para julio del cuarto año	68
Fig. 5.28 Varianzas del error en la estimación y orden de prioridad de los pozos para julio del cuarto año.....	69
Fig. 5.29 Estimación de la carga hidráulica para julio del quinto año.....	70
Fig. 5.30 Varianzas del error en la estimación y orden de prioridad de los pozos para julio del quinto año.....	71
Fig. 6.1 Área de estudio y localización de los puntos de monitoreo/estimación (círculos) y puntos únicamente de estimación (cuadros).....	77
Fig. 6.2 Semivariograma espacio-temporal muestral de la carga hidráulica	78
Fig. 6.3 Proyección del semivariograma espacio-temporal muestral de la carga hidráulica sobre el plano semivariograma-intervalos espaciales (círculos) y modelo de semivariograma ajustado (cuadros)	79
Fig. 6.4 Proyección del semivariograma espacio-temporal muestral de la carga hidráulica sobre el plano semivariograma-intervalos temporales (círculos) y modelo de semivariograma ajustado (cuadros)	80
Fig. 6.5 Varianza Total vs Número de posiciones de monitoreo espacio-temporal	82
Fig. 6.6 Raíz cuadrada (Varianza total/Posiciones de estimación espacio-temporal) vs Número de posiciones de monitoreo espacio-temporal.....	83
Fig. 6.7 Número de veces de monitoreo para cada pozo en el programa de monitoreo óptimo espacio-temporal (arriba) y nombre del pozo (abajo).....	84
Fig. 6.8 Estimación de la carga hidráulica para agosto del primer año	86
Fig. 6.9 Varianzas del error en la estimación para agosto del primer año.....	87
Fig. 6.10 Estimación de la carga hidráulica para julio del segundo año.....	88
Fig. 6.11 Varianzas del error en la estimación para julio del segundo año	89
Fig. 6.12 Estimación de la carga hidráulica para julio del tercer año.....	90
Fig. 6.13 Varianzas del error en la estimación para julio del tercer año.....	91
Fig. 6.14 Red de monitoreo (círculos) y nodo de estimación con los errores más altos (cuadro) 93	
Fig. 6.15 Datos en los pozos 6 y 12, y datos y estimaciones en el nodo 63	94
Fig. 6.16 Estimación de la carga hidráulica para julio del cuarto año	95
Fig. 6.17 Varianzas del error en la estimación para julio del cuarto año.....	96
Fig. 6.18 Estimación de la carga hidráulica para julio del quinto año.....	97
Fig. 6.19 Varianzas del error en la estimación para julio del quinto año	98
Fig. 6.20 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo L-246.....	99
Fig. 6.21 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo 644-A	100
Fig. 6.22 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo L-107.....	101
Fig. 6.23 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo 2082	102
Fig. 6.24 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo 1792	103
Fig. 6.25 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo 752	104
Fig. 6.26 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo 976	105

Fig. 6.27 Simulación y estimación de la carga hidráulica para el corte A-A' de la Fig. 5.4	106
Fig. 7.1 Área de estudio y ubicación de pozos con datos del nivel del agua subterránea	110
Fig. 7.2 Malla de estimación (cuadros) y posibles puntos de monitoreo espacial (círculos)	111
Fig. 7.3 Semivariograma espacio-temporal muestral de niveles del agua subterránea	112
Fig. 7.4 Proyección del semivariograma espacio-temporal muestral de niveles del agua subterránea sobre el plano semivariograma-intervalos espaciales.....	113
Fig. 7.5 Proyección del semivariograma espacio-temporal muestral de niveles del agua subterránea sobre el plano semivariograma-intervalos temporales	114
Fig. 7.6 Semivariograma espacio-temporal muestral de los residuos de un ajuste de un polinomio de primer grado	115
Fig. 7.7 Proyección del semivariograma espacio-temporal muestral de los residuos de un ajuste de un polinomio de primer grado sobre el plano semivariograma-intervalos espaciales (círculos) y modelo de semivariograma ajustado (cuadros).....	116
Fig. 7.8 Proyección del semivariograma espacio-temporal muestral de los residuos de un ajuste de un polinomio de primer grado sobre el plano semivariograma-intervalos temporales (círculos) y modelo de semivariograma ajustado (cuadros).....	117
Fig. 7.9 Varianza Total vs Número de posiciones de monitoreo espacio-temporal	119
Fig. 7.10 Raíz cuadrada (Varianza total/Posiciones de estimación espacio-temporal) vs número de posiciones de monitoreo espacio-temporal	120
Fig. 7.11 Posiciones de monitoreo para cada año	121
Fig. 7.12 Número de veces de monitoreo para cada pozo en el programa de monitoreo óptimo espacio-temporal (arriba) y nombre del pozo (abajo).....	122
Fig. 7.13 Estimación del nivel del agua subterránea para diciembre de 1997	124
Fig. 7.14 Varianzas del error en la estimación y orden de prioridad de los pozos para diciembre de 1997.....	125
Fig. 7.15 Estimación del nivel del agua subterránea para diciembre de 2001	126
Fig. 7.16 Varianzas del error en la estimación y orden de prioridad de los pozos para diciembre de 2001	127
Fig. 7.17 Estimación del nivel del agua subterránea para diciembre de 2004.....	128
Fig. 7.18 Varianzas del error en la estimación y orden de prioridad de los pozos para diciembre de 2004.....	129
Fig. 7.19 Estimación del nivel del agua subterránea para noviembre 2007	130
Fig. 7.20 Varianzas del error en la estimación y orden de prioridad de los pozos para noviembre de 2007.....	131

ÍNDICE DE TABLAS

Tabla 5.1 Tasas de bombeo promedio mensuales.....	38
Tabla 5.2 Factores de escalamiento del bombeo mensual	38
Tabla 5.3 Precipitación mensual promedio histórica (1921-1998).....	39
Tabla 5.4 Modelo de semivariograma esférico obtenido para una tormenta.....	40
Tabla 5.5 Parámetros del modelo de semivariograma espacio-temporal de la carga hidráulica ..	54
Tabla 5.6 Resultados de la validación cruzada para el periodo de agosto del primer año a julio del tercer año.....	55
Tabla 5.7 Desempeño de la red de monitoreo óptima	67
Tabla 5.8 Errores en la estimación a través del tiempo para algunos pozos utilizando la red de monitoreo propuesta.....	72
Tabla 5.9 Errores en la estimación de la carga hidráulica para el corte A-A' de la Fig. 5.4 utilizando la red de monitoreo propuesta.....	73
Tabla 6.1 Parámetros del modelo de semivariograma espacio-temporal de la carga hidráulica ..	80
Tabla 6.2 Resultados de la validación cruzada para el periodo de agosto del primer año a julio del tercer año.....	81
Tabla 6.3 Desempeño de la red de monitoreo.....	92
Tabla 6.4 Errores en la estimación a través del tiempo para algunos pozos utilizando la red de monitoreo y el programa propuesto	105
Tabla 6.5 Errores en la estimación de la carga hidráulica para el corte A-A' de la Fig. 5.4 utilizando la red de monitoreo propuesta.....	107
Tabla 7.1 Parámetros del modelo de semivariograma espacio-temporal de los residuos de un ajuste de un polinomio de primer grado	118
Tabla 7.2 Resultados de la validación cruzada para los últimos 10 años con información	118

1. INTRODUCCIÓN

1.1 Antecedentes

La extracción adecuada de los elementos naturales con los que se dispone es fundamental para lograr un desarrollo ordenado de la humanidad; un elemento primordial para la vida es el agua. Las dotaciones requeridas de agua para consumo humano, uso agrícola, industrial, generación de energía, etc., podrán ser satisfechas a largo plazo únicamente mediante un control eficiente en la captación y extracción de las fuentes de abastecimiento, alta eficiencia en la distribución, adecuado uso y posterior tratamiento. A su vez, acciones de este tipo propician una mayor producción en el campo, en la industria y en la generación de energía.

Del total de agua presente en la Tierra, aproximadamente el 96.5% es salada y se encuentra en los océanos, cerca del 1.74% se presenta en forma de hielo, alrededor del 1.7% es agua subterránea y menos del 0.06% se distribuye en ríos, lagos, atmósfera, suelo y biosfera (Gleick, 1996). Es claro entonces que del total de agua dulce asequible para consumo humano, la mayor parte se encuentra en el subsuelo (agua subterránea) y sólo un pequeño porcentaje es agua superficial (principalmente ríos y lagos).

El agua subterránea presenta algunas ventajas respecto al agua superficial como: una composición química casi constante, se encuentra menos expuesta a ser contaminada, los acuíferos pueden actuar como filtros naturales, se tiene menor pérdida por evaporación y su presencia se encuentra menos afectada por variaciones climáticas. La distribución de la precipitación pluvial que ocurre dentro del territorio mexicano es muy irregular, presentándose en abundancia al sureste y escasa en las regiones centro-norte y norte del país, por lo que en muchas de estas regiones, la única manera de contar con el vital líquido es mediante la extracción del agua subterránea.

Poco más de la tercera parte del agua utilizada en México es de origen subterráneo; con ella se atienden las necesidades de alrededor de dos terceras partes de la población y se abastece un tercio de la superficie total irrigada (Carabias *et al.*, 2005). Es un hecho que la mayor parte de las ciudades importantes del país se abastecen parcial o totalmente de agua proveniente del subsuelo. Para el año 2006, 104 de los 653 acuíferos considerados dentro de territorio nacional se han considerado como sobre-explotados.

En la gran mayoría de los centros urbanos se tiene fuerte concentración de captación de agua subterránea que ha producido efectos perjudiciales como: abatimiento significativo del nivel del agua subterránea, asentamiento del terreno, grietas, problemas de degradación de la calidad del agua, migración de agua fósil de mala calidad, etc. Puesto que, no obstante los problemas generados, la necesidad de agua continuará creciendo ininterrumpidamente, es evidente que debemos preservar el agua subterránea, mediante una extracción más racional. Esto ha implicado la necesidad de realizar estudios con los objetivos siguientes: cuantificar el volumen aprovechable de agua subterránea en cada zona, definir el manejo óptimo del agua subterránea en cada zona, encontrar medidas correctivas para alargar la vida útil de acuíferos sometidos a una sobreexplotación irreversible, etc. (Chávez, 1986). El conocimiento de la carga hidráulica y de su evolución en el tiempo es crucial para poder sugerir alternativas apropiadas en el manejo del agua subterránea.

La interacción del agua subterránea en el ambiente requiere de un programa de observación integral de variables que incorporen su funcionamiento en espacio y tiempo, así como de los

efectos esperados (Maderey-Rascón y Carrillo-Rivera, 2005). De esta manera resulta de suma importancia definir y establecer indicadores confiables a través de programas de detección (como es el caso de las redes de monitoreo), para poder tener elementos de juicio que permitan fundamentar la realización de cualquier tipo de medida preventiva y correctiva relacionada con el agua y su entorno ambiental. Las redes de monitoreo son herramientas de protección y control de la calidad que ofrecen niveles de información altos con inversiones mínimas (Szucs y Jordan, 1994).

La información recabada por las redes de monitoreo del agua subterránea es sumamente valiosa en el entendimiento del funcionamiento del agua en sí y de los acuíferos que la contienen. Sin embargo, la toma de datos en campo se traduce en tiempo y dinero, por lo que en muchas ocasiones se ha dejado de realizar un monitoreo continuo en pozos, perdiéndose el seguimiento de variaciones de la carga hidráulica en muchos acuíferos. En México, es poco común el monitoreo espacio-temporal consistente de la carga hidráulica (ya sea por falta de planeación y/o por el alto costo que esto puede representar); como criterio práctico, las redes de monitoreo se conforman por pozos a los que es fácil acceder, tratando de cubrir lo más posible el acuífero y en específico, las zonas definidas como de recarga y descarga; usualmente las frecuencias de monitoreo son propuestas para temporada de estiaje y/o temporada de lluvias, muchos acuíferos del país no son siquiera monitoreados anualmente. En años recientes, con la finalidad de reducir el costo para la adquisición de datos y tratando a la vez de evitar perder información valiosa del funcionamiento del agua subterránea en acuíferos, se han desarrollado trabajos para el diseño óptimo de redes de monitoreo en los que se busca lograr un cierto nivel de información, evitando muestrear en sitios que aportan información redundante, mediante el uso de una función objetivo y algún método de optimización.

Uno de los problemas al momento de estimar la carga hidráulica es que su valor no solamente depende de su posición sino también del tiempo, por lo que es necesario el diseño de redes de monitoreo que permitan caracterizar de manera adecuada el comportamiento espacio-temporal de esta variable. En la literatura pueden encontrarse muchos trabajos dedicados al diseño óptimo de redes de monitoreo de la carga hidráulica, sin embargo, en la mayoría de éstos, no se considera al espacio y al tiempo de forma conjunta, es decir, se seleccionan posiciones de monitoreo pero no frecuencias y viceversa. En el presente trabajo se presenta una metodología para el diseño espacio-temporal óptimo de redes de monitoreo de la carga hidráulica, basada en un análisis geoestadístico de los datos históricos que se tengan de esta variable en un acuífero.

En las siguientes secciones se describirá el panorama actual de dos problemas íntimamente relacionados: la estimación de la carga hidráulica y su monitoreo. Estos dos problemas están relacionados debido a que cuando se determinan los sitios y frecuencias de monitoreo de una variable, por lo general se plantea la necesidad de estimar la variable en sitios y probablemente también en tiempos en los que no se monitoreó la variable. Por lo tanto, se requiere de un método para interpolar o estimar la variable. Por otro lado, una vez que se escoge un método de estimación, éste se puede utilizar para determinar los “mejores” sitios y tiempos de monitoreo, con algún criterio de optimización.

Estimación de variables hidrogeológicas

Una herramienta útil para obtener estimaciones confiables de la distribución de un parámetro ya sea en espacio y/o tiempo es la geoestadística. El uso de técnicas geoestadísticas para la determinación de variables hidrogeológicas se ha dado desde los años setenta del siglo

pasado (Delhomme, 1978; Gambolati y Volpi, 1979; Volpi *et al.* 1979), sin embargo, la aplicación e implementación de estos modelos no se ha agotado. La estimación espacial (en 3D) de la carga hidráulica es un problema al que se enfrenta todo hidrogeólogo. El especialista debe contar con metodologías capaces de representar los cambios en el almacenamiento y distribución del agua subterránea, de forma eficiente. Con frecuencia al llevar a cabo el análisis de la evolución de la carga hidráulica, se encuentra que los datos son escasos en el espacio y que no se cuenta para todos los años con datos en los mismos pozos. Esto dificulta dicho análisis debido a que, por un lado, la estimación para cada año puede contar con poca información y hacer que su incertidumbre sea grande, y por otro lado, en los diferentes años la estimación tendrá un nivel de incertidumbre diferente en zonas distintas, lo que hace que sea difícil compararlas.

La heterogeneidad que presentan los fenómenos hidrogeológicos hace impráctico describir la distribución de la carga hidráulica en un acuífero considerando la medición en un solo tiempo. Debido a que generalmente la información con que se cuenta es limitada, no es posible describir una variable regionalizada con precisión. Por lo tanto, es conveniente ver el dato asequible como una de muchas realizaciones, es decir, el parámetro se tratará como una variable aleatoria. La decisión de seleccionar una función aleatoria para modelar la variable regionalizada es sólo una manera de atacar el problema.

Usualmente, la información hidrogeológica disponible es multivariada, esto quiere decir que una variable se encuentra generalmente correlacionada a otra u otras; esta información adicional puede ser muy útil para robustecer la estimación de la variable. En geoestadística clásica se ha estudiado la correlación existente entre dos o más variables que varían en el espacio; la variación conjunta de dos atributos se mide a través de un semivariograma cruzado. De forma general, dentro de la geoestadística, a esta técnica se le conoce como cokrigado y también es referida como análisis geoestadístico multivariado.

Ahmed y De Marsily (1993) mostraron que es posible utilizar niveles piezométricos y la transmisividad para estimar un campo de transmisividades mediante el cokrigado. En este caso el nivel del agua subterránea se utiliza como variable secundaria para mejorar la estimación de la transmisividad. Trabajos de este tipo mantienen un punto de vista puramente espacial, sin embargo, los hidrogeólogos han tratado la variación de la carga hidráulica también desde el punto de vista temporal. Muchas propiedades varían en el tiempo, por ejemplo, la elevación del nivel estático que medimos en un pozo nunca es constante; dejando a un lado la incertidumbre de la medición, el nivel estático en un pozo varía debido a la extracción en el pozo o de una forma regional por una disminución de la recarga, que está íntimamente relacionada con el ciclo de lluvia que prevalece en la zona, o más allá de la zona de estudio. Las propiedades de los materiales como porosidad, conductividad hidráulica (LT^{-1}), permeabilidad (L^2) y coeficiente de almacenamiento, cambian debido a procesos de compactación, pero el tiempo en el que se desarrollan por lo general es muy largo; en este caso se consideran estables.

Los modelos que analizan la variación en el espacio son similares a los que varían en el tiempo. La diferencia radica en el hecho de que a las variables espaciales se les trata en una, dos o tres dimensiones, mientras que el tiempo se mueve siempre hacia adelante y los procesos tienen efecto en secuencia. Por ejemplo, si medimos una variable $Z(t)$, la carga hidráulica en una posición, utilizando intervalos discretos en el tiempo para su medición, obtenemos una serie de observaciones $z(t_i)$, $i=1,2,\dots,n$. En algunas situaciones, esta serie tendrá suficiente información para realizar predicciones de $Z(t)$ en un tiempo futuro o tal vez interpolar entre los tiempos de medición. Para realizarlo será necesario un modelo que caracterice el comportamiento dinámico de $Z(t)$.

Aunque se han empleado las técnicas univariada y multivariada, el desarrollo espacio-temporal es relativamente nuevo en la hidrogeología (y en la geoestadística); la modelación espacio-tiempo es resultado de la investigación de procesos dinámicos que involucran tanto la variabilidad espacial como la temporal (Stein, 1986; Christakos, 2000; Bierkens, 2001). En el campo de la hidrogeología los primeros en plantear el problema fueron Rouhani y Hall (1989). En su investigación suponen que el espacio y el tiempo pueden ser caracterizados por una función aleatoria que trata de forma conjunta el espacio y el tiempo, al integrarse este último como una dimensión adicional a la componente espacial. Este procedimiento fue aplicado por Bilonick (1985), obteniendo resultados interesantes en el cálculo de depósitos de sulfato en el noreste de los Estados Unidos.

Es necesario recalcar que existen diferencias fundamentales entre los ejes coordenados del espacio y el tiempo. Por ejemplo, para el tiempo es claro que existe un orden, podemos ubicar con facilidad el pasado, presente, y futuro, esto no puede hacerse en el espacio. En el tiempo, las mediciones son tomadas en un solo sentido, aquí la estimación es una predicción. En contraste en el espacio la estimación es enfocada a una interpolación de los datos. Las unidades de escala y distancia son distintas en el espacio y el tiempo, y no pueden ser comparadas en un sentido físico (Rouhani y Myers, 1990).

Revisión del estado del arte de métodos estadísticos para la estimación de niveles del agua subterránea

El desconocimiento de los gradientes geotérmicos que tienen influencia en el agua subterránea, la ausencia de mediciones piezométricas a diferentes profundidades en un mismo pozo, y el desconocimiento de las características fisicoquímicas del agua subterránea dificultan la estimación de la carga hidráulica en 3D (ver capítulo dos). Aunque muchos autores se refieren en forma indistinta a la carga hidráulica y a las elevaciones del nivel del agua, en esta tesis, para los trabajos basados en métodos estadísticos nos referiremos a la estimación de niveles del agua subterránea, ya que en estos trabajos no se toman en cuenta variaciones de la densidad y temperatura en el agua subterránea para ajustar los niveles medidos a su correspondiente valor comparable de carga hidráulica; en cambio, para los trabajos que involucran modelos de simulación numérica, en los que se incluyen hipótesis de flujo horizontal, temperatura uniforme y densidad constante en el agua, se considera que se estima la carga hidráulica. En esta sección se incluyen únicamente métodos de estimación estadísticos por lo que éstos sólo estiman niveles del agua subterránea y en un caso la profundidad de estos niveles.

Kemal y Guney (2007) realizaron un análisis espacial de los niveles del agua subterránea en una zona de irrigación en Turquía usando kriging universal. Aplicaron la técnica de kriging universal como método de interpolación, por ser ésta la que ha dado mejores resultados en el tratamiento de los niveles del agua subterránea. Los residuos producto del ajuste de un polinomio a los datos de los niveles del agua se consideran correlacionados espacialmente. Con los residuos se obtiene un semivariograma muestral al cual se le ajusta un modelo. El modelo seleccionado es aquél que permite obtener las estimaciones con los menores errores, al compararse con los datos medidos en campo.

Ahmadi y Sedghamiz (2007) analizaron la pertinencia del uso de técnicas geoestadísticas para la estimación ya sea espacial o temporal de los niveles del agua subterránea. Comprobaron la validez de los modelos seleccionados en cada caso mediante validación cruzada. Para el caso

espacial obtuvieron una subestimación de los niveles del agua de aproximadamente un 3% sobre toda el área de estudio; para el caso temporal la subestimación fue del 6%.

Mendoza y Herrera (2007) realizaron la estimación multivariada espacio-tiempo de niveles del agua subterránea en el valle de Querétaro-Obrajuelo. Se evaluó el trabajar con información de diferentes años para realizar una estimación multivariada de los niveles. Utilizaron los conceptos de cokriging y semivariograma cruzado para la estimación multivariada de los niveles del agua subterránea y compararon los resultados con las estimaciones univariadas para diferentes años. En general obtuvieron mejores resultados con la estimación multivariada que con la estimación univariada.

Ahmadi y Sedghamiz (2008) evaluaron el uso del kriging y cokriging en la estimación de la profundidad de niveles del agua subterránea en un acuífero de Irán. Compararon estimaciones en 39 pozos. Promediaron todos los valores máximos anuales, todos los mínimos y todos los valores medios, para un periodo de 12 años, a estas variables les llamaron H, L y A respectivamente. Con éstas obtuvieron estimaciones espaciales (mediante kriging ordinario) de profundidades del nivel del agua. Como variable secundaria para el empleo del cokriging utilizaron los valores máximos, mínimos y medios absolutos de todo el historial con el que contaban, a estas variables les llamaron α , β y γ , respectivamente. Obtuvieron estimaciones de la profundidad del nivel del agua subterránea para condiciones secas, húmedas y normales en el acuífero, y las compararon con los datos medidos en campo. Para el kriging, las subestimaciones fueron de 5.5, 2.2 y 5.3%, para el cokriging las subestimaciones fueron de 3.3, 2 y 2.2%, respectivamente.

Mendoza (2008) exploró métodos de estimación del nivel del agua subterránea utilizando variables espaciales, temporalmente correlacionadas. Para poder cumplir con ello, tuvo que remover la tendencia a los datos mediante el ajuste de un polinomio para así trabajar con los residuos resultantes. A los residuos se les ajustó un modelo de correlación (semivariograma teórico). La medición de los errores de estimación utilizando el modelo obtenido se realizó mediante un procedimiento llamado validación cruzada. La validación cruzada es un procedimiento que permite establecer si el modelo de semivariograma obtenido es representativo de la variabilidad espacial de los datos usados en la estimación del mismo. El método consiste en sacar un elemento de la muestra y estimar el valor en ese punto con kriging ordinario empleando el modelo de semivariograma obtenido. Los resultados que obtuvo fueron buenos en el sentido de que con la modelación espacio-temporal se logró una mayor reducción de la varianza del error de la estimación, comparada con la obtenida con la modelación univariada. Utilizó varios métodos para estimar los niveles del agua subterránea: el kriging ordinario (KO), el cokriging (Cok o multivariado) y la estimación espacio-temporal (ET). Los métodos que consideran tanto al espacio como al tiempo en conjunto son los que le dieron los mejores resultados.

Ta'any y Tahboub (2009) determinaron la variabilidad espacial y temporal de fluctuaciones de los niveles del agua subterránea para el periodo 2001-2005, en la cuenca Amman-Zarqa en Jordán. Obtuvieron modelos y estimaciones espaciales para temporada de lluvias y de secas dentro del periodo señalado. Para el análisis temporal obtuvieron autocorrelogramas para cada pozo y analizaron la dependencia temporal de los niveles del agua subterránea en cada uno de ellos. El grado de incertidumbre en las estimaciones espaciales resultó ser menor al 5% para todos los tiempos excepto para la temporada de lluvias del 2005. Con el análisis temporal se determinó mediante los autocorrelogramas que la dependencia temporal de la información en cada pozo es alta, salvo dos pozos donde no se encontró esta dependencia. Algunos pozos mostraron dependencia de 3 a 6 meses y otros de 10 a 21 meses. Esto quiere decir que durante ese intervalo de tiempo los niveles del agua subterránea pueden ser fácilmente predichos.

Diseño óptimo de redes de monitoreo del agua subterránea

El principal problema en la evaluación del agua subterránea es que se tiene una carencia casi absoluta de datos básicos como: características hidráulicas y límites físicos de los acuíferos - horizontal y vertical-, características constructivas y material litológico atravesado durante la perforación de los pozos registrados, etc. La evolución de los niveles y de la calidad del agua subterránea en las zonas de extracción es casi desconocida en la gran mayoría de los acuíferos del país, incluyendo el monto de la extracción; aunque es justo que técnicas con mucha visión, intuyendo la importancia de esta información, establecieron en sus zonas de control respectivas, la observación sistemática de los niveles del agua subterránea en pozos “piloto” (red de monitoreo). En la actualidad, se han realizado trabajos con la finalidad de optimizar el diseño de este tipo de redes. El diseño de redes se ha extendido al monitoreo de la calidad del agua subterránea ya que se han presentado problemas graves de contaminación del agua subterránea en numerosas áreas urbanas e industriales del país donde se enfrentan serias dificultades para desechar las aguas negras sin contaminar las fuentes de abastecimiento

El diseño de una red de monitoreo óptima del agua subterránea consiste en seleccionar de forma óptima las posiciones de los pozos y las frecuencias de monitoreo para satisfacer diferentes objetivos. Existen muchos trabajos dedicados al análisis del diseño de redes de monitoreo de la calidad del agua subterránea y que pueden servir de base en el diseño de redes que permitan una adecuada caracterización de la carga hidráulica en acuíferos (ver anexo A). Los primeros trabajos en el diseño del monitoreo se centraron en métodos para situar nuevos pozos de monitoreo. Recientemente, se han desarrollado métodos para identificar planes de monitoreo que minimizan la redundancia espacial y temporal en redes de monitoreo existentes. La mayoría de estos trabajos suponen que el diseño de una red de monitoreo involucra el espacio pero no el tiempo, esto es, se seleccionan únicamente las posiciones de los pozos de monitoreo pero no las frecuencias de monitoreo o viceversa.

De acuerdo con Herrera y Pinder (2005), tres puntos de vista han influenciado grandemente el diseño de redes de monitoreo del agua subterránea. En el primero de ellos, que llamamos marco hidrológico (Loaiciga *et al.*, 1992), la red y su programa de monitoreo se definen por consideraciones que toman en cuenta solamente las condiciones hidrológicas del sitio, sin recurrir a técnicas estadísticas o probabilísticas avanzadas. El segundo propone el análisis de datos dentro de un marco estadístico y define la red de monitoreo basado en inferencias obtenidas de los datos, éste se llama marco estadístico. El último punto de vista se llama marco de modelación, en éste se utilizan modelos numéricos de flujo del agua subterránea para determinar posiciones y frecuencias de monitoreo.

Loaiciga (1992) presenta una revisión del diseño de redes de monitoreo de calidad del agua subterránea que incluye un gran número de trabajos de este tipo. Existen trabajos de diversos autores que proponen métodos para el diseño de redes de monitoreo en los que las decisiones de monitoreo involucran tanto al tiempo como al espacio, Herrera (1998) llama a éstos diseños espacio-temporales.

Revisión del estado del arte del diseño de redes de monitoreo de niveles del agua subterránea

Dentro del marco estadístico, los métodos más relevantes para esta tesis se basan en geoestadística. Samper y Carrera (1990) realizaron la comparación de diferentes diseños de redes de observación del nivel del agua subterránea basados en la implementación del kriging. Dividen

el diseño de redes de monitoreo en dos grandes categorías: el método del punto ficticio y el caso de varios puntos. La primera comprende los casos cuando se realiza una estimación global (se estima en los pozos existentes y en el propuesto) y cuando se realiza una local (se estima en puntos representativos de una región como pueden ser los nodos de una malla que representa la zona a monitorear). La segunda categoría considera posiciones arbitrarias como puntos de medición: método de inclusiones sucesivas (se van probando puntos alternativos que se unen a un conjunto ya dado de pozos), intercambios sucesivos óptimos (se intercambian los elementos de un conjunto dado de puntos existentes con los elementos de un conjunto de puntos alternativos para encontrar la mínima varianza), enumeración total (se analizan subconjuntos del total de pozos a monitorear; se empieza con todos los subconjuntos de un elemento, luego los que se formen con dos elementos y así sucesivamente hasta llegar al subconjunto que proporcione la incertidumbre deseada) y "rama y límite" (inverso al anterior). Dentro de esta categoría también se considera el uso de mallas regulares para evaluar el efecto del cambio en el espaciamiento entre pozos de la red de observación sobre el nivel de incertidumbre. En la evaluación, las estimaciones son realizadas mediante kriging, y los diferentes métodos de optimización empleados son inclusiones sucesivas, intercambios sucesivos óptimos, enumeración total y "rama y límite". La dimensión del diseño de las redes es bidimensional. Del análisis realizado para el diseño de una red se concluye que el método de inclusiones sucesivas aunque es subóptimo resulta dar buenos resultados en la elección de puntos de monitoreo además de que resulta ser una metodología muy económica. Es por lo anterior que esta metodología es la que se ha utilizado frecuentemente, obteniendo buenos resultados.

Gangopadhyay *et al.* (2001) identificaron los pozos de monitoreo que son importantes para predecir la variación dinámica del nivel del agua subterránea. Llevaron a cabo un análisis de componentes principales para todos los pozos de una red de monitoreo existente de un acuífero, y asignaron un ranking a cada uno de éstos. El ranking se calculó dividiendo el número de veces que un pozo aparecía como pozo principal (cuando explica la mayor parte de la información para estimar en un sitio determinado) entre el número de veces que fue empleado en el análisis de componentes principales. Para cada sitio de estimación, se realizó un análisis de componentes principales que incluía únicamente a los pozos que se encontraban dentro de un radio de búsqueda, este radio se definió tratando de incorporar el comportamiento físico del acuífero, calculando un radio de 10 km como de influencia una vez que se bombea en un pozo. Se hicieron evaluaciones comparando la distribución del nivel del agua subterránea usando datos de todos los pozos disponibles contra la obtenida empleando los datos de los pozos seleccionados tomando como pauta el esquema de ranking encontrado. El valor del coeficiente de variación promedio para todo el periodo de diseño fue empleado para tomar la decisión del número de pozos a incluir en la red de monitoreo.

Nunes *et al.* (2004b) propusieron un método para elegir en forma óptima un subconjunto de pozos dentro de una red de monitoreo existente de niveles de agua subterránea. El método considera una función de optimización que suma una parte espacial (en la que se busca reducir la varianza del error en la estimación), una parte temporal (que trabaja con series de tiempo y en la que se determinan las redundancias temporales), y otros dos términos que incluyen los tiempos del monitoreo y sus costos (para recorrer todos los sitios elegidos y para realizar el monitoreo en cada estación). La estimación se realizó con kriging y el método de optimización seleccionado fue recodido simulado. Los autores concluyen que la reducción relativa en los costos de exploración compensa la relativa pérdida de información en datos representativos. En este artículo el espacio y el tiempo son independientes y luego se suman en la función objetivo, es

decir el método de estimación no incluye al espacio y al tiempo en forma conjunta. No se determina una frecuencia de monitoreo; las series de tiempo solo se utilizan para evaluar la redundancia en el tiempo de la información que se adquiere al monitorear en un sitio, tuvieron preferencia las posiciones menos redundantes en el tiempo.

Bravo (2005) calculó la matriz de covarianza a priori utilizada por el filtro de Kalman a partir de un modelo compuesto que se ajustó a dos semivariogramas muestrales obtenidos en el análisis geoestadístico de los abatimientos anuales de los periodos 1993-1995 y 1995-1999. Para evaluar la cantidad de información que aportaban los puntos adicionales incluidos en la red preliminar se hizo un estimación de los errores en el periodo 2010 al 2030 obtenidos si éstos no eran incluidos en la red de monitoreo. Para esto se utilizó como sustituto de los datos de abatimientos los resultados de un modelo numérico de flujo para ese periodo. Como resultado de este análisis se quitaron algunas puntos adicionales considerados en la red de monitoreo preliminar, obteniéndose así la red de monitoreo óptima.

Kumar *et al.* (2005) diseñaron una red de monitoreo del nivel del agua subterránea basándose en la teoría de las variables regionalizadas. Quitaron la tendencia a los datos al ajustarles un plano. Ajustaron un modelo lineal al semivariograma de los residuos obtenidos, estimaron valores mediante kriging universal y determinaron los errores en la estimación mediante un procedimiento de validación cruzada (jackknife). Se probaron diferentes conjuntos de pozos superimponiendo mallas cuadrículas de diferentes tamaños y seleccionando un pozo por celda. La estimación se realizó mediante kriging residual y el método de optimización fue de inspección. El diseño es bidimensional. Al dividir el área de estudio en cuadrículas de diferente tamaño de celda se llegó a determinar la densidad de pozos óptima para reducir los errores en las estimaciones.

Faisal *et al.* (2007) diseñaron una red de monitoreo óptima de la piezometría, necesaria para la estimación de un balance de agua subterránea. Se estimaron modelos teóricos para semivariogramas experimentales de datos de pozos en tres distintas campañas: dos con los niveles mínimos en el acuífero y una con los niveles máximos. Se realizó la validación cruzada del modelo seleccionado en cada campaña y se calcularon los residuos en cada pozo. Residuo se le llamó a la diferencia entre el valor medido y el valor estimado en cada pozo mediante validación cruzada. Para cada pozo se obtuvo el valor absoluto del promedio de los residuos de las tres validaciones (una para cada campaña). Se le dio un orden de prioridad a cada pozo. El pozo con el mayor valor absoluto del promedio de los residuos fue el de mayor prioridad y el pozo con el menor valor el menos prioritario. Se evaluaron la evolución del nivel piezométrico medio, el rendimiento específico, recarga por año y balance anual para distintos conjuntos de pozos siguiendo la prioridad obtenida en el paso anterior. El menor subconjunto con valores de estimación muy parecidos para todos los casos con respecto a los obtenidos utilizando todos los pozos de la red existente, es el que conforma la red óptima. Esto se traduce en una disminución en los costos de monitoreo para obtener el mismo nivel de información a la hora de realizar un balance del agua subterránea. La estimación se basó en la validación cruzada utilizando kriging y la optimización mediante inspección. La dimensión del diseño de la red es bidimensional. El monitoreo de los pozos seleccionados mediante la técnica propuesta ofrece prácticamente el mismo nivel de información para un balance del agua subterránea que el monitoreo utilizando todos los pozos existentes. El criterio de optimización a diferencia del filtro de Kalman no permite que el nivel de incertidumbre se actualice una vez que un pozo ha sido seleccionado para formar parte de la red de monitoreo. Con el uso del filtro de Kalman (método bayesiano) se

considera que, una vez seleccionado un pozo, el nivel de incertidumbre en el área de estudio disminuye.

Revisión del estado del arte del diseño de redes de monitoreo de la carga hidráulica

Aunque la metodología propuesta en el presente trabajo cae dentro del marco estadístico para el diseño de redes de monitoreo, ésta se deriva de una metodología propuesta en el marco de modelación (Herrera, 1998), por lo que a continuación se presenta la revisión que se realizó de trabajos basados en esta forma de diseño.

Andricevic (1990) presentó un método de tiempo real para el manejo y monitoreo de la carga hidráulica del agua subterránea. El modelo de manejo, estructurado como un problema de control de tiempo óptimo, identifica los gastos de extracción óptimos que satisfacen una función de costo de tipo sanción (penalti) de dos objetivos conflictivos: satisfacer las demandas de extracción y mantener los niveles de carga hidráulica deseados. Se incluye una acción de control obtenida en forma de retroalimentación que depende de la incertidumbre futura en la predicción de la carga hidráulica debido a la variabilidad en la conductividad hidráulica. Cuando se obtienen nuevas observaciones de campo, la conductividad hidráulica se actualiza en tiempo real.

Van Geer *et al.* (1991) utilizaron el filtro de Kalman para mejorar y cuantificar la incertidumbre de una simulación numérica del agua subterránea, para posteriormente diseñar una red de monitoreo. Emplearon un modelo numérico de flujo para obtener la matriz de covarianza de los parámetros hidrogeológicos y posteriormente su calibración. Usaron el código de modelación MODFLOW junto con el filtro de Kalman. El objetivo del diseño fue monitorear el cambio espacio-temporal de la carga hidráulica, causado por la extracción. El único término incierto considerado en la ecuación de flujo es el error del modelo. Los autores proponen usar un filtro de Kalman para estimar los parámetros de la ecuación de flujo (llamados parámetros deterministas) y algunos parámetros asociados con la matriz de covarianza del error del modelo (llamados parámetros estocásticos). La calibración es ejecutada para un periodo en el cual las condiciones estacionarias se satisfacen y en el cual todas las matrices requeridas por el algoritmo del filtro de Kalman se suponen fijas (no cambian con el tiempo). El método estima los parámetros secuencialmente: primero se propone una estimación de la covarianza del error del modelo y se realiza la primera calibración de parámetros deterministas, entonces estos parámetros se mantienen constantes y se obtienen los parámetros estocásticos. Estas dos etapas se repiten varias veces hasta alcanzar un criterio de error preestablecido para ambos sistemas de parámetros.

Yangxiao *et al.* (1991) analizaron dos ejemplos de campo en los que se aplicó el método de Van Geer *et al.* (1991). En ambos la frecuencia de monitoreo se mantuvo fija y las densidades de las redes se minimizaron bajo la restricción de un valor límite de la desviación estándar del error de estimación preestablecido. Se analizaron varias alternativas y se eligió la mejor por inspección. Los autores analizaron la importancia relativa de la densidad de la red espacial y las frecuencias de monitoreo relacionándolas con el tiempo de respuesta del sistema. Encontraron que si el sistema reacciona rápido, la optimización espacial de la red es importante. Si el sistema reacciona lento, tanto la optimización temporal como la espacial son importantes.

Wu (2003) diseñó una red de monitoreo óptima de la carga hidráulica para un acuífero profundo. La metodología consiste en lo siguiente: se calibra un modelo numérico de flujo del acuífero usando un algoritmo de estimación óptima que emplea el elemento finito acoplado con el filtro de Kalman. La densidad de la red de monitoreo es analizada simulando la superficie de

la desviación estándar sobre el acuífero para diferentes configuraciones de redes. Una ventaja del método es que el algoritmo de simulación es calculado sin la necesidad de datos reales, lo cual hace posible introducir posiciones y frecuencias de medición ficticias para calcular superficies de desviación estándar ficticias, esto puede lograrse cambiando la matriz de medición en el filtro de Kalman. Para seleccionar el número de posiciones a ser incluidas en la red se buscó que la desviación estándar promedio sobre todo el acuífero tuviera un valor previamente establecido de 0.5 m. La estimación se realiza con un modelo numérico de flujo acoplado con el filtro de Kalman, la optimización se realiza mediante inspección. La dimensión del diseño de la red es bidimensional e incluye el tiempo. El método no solo considera inferencias estadísticas de la variable medida, sino también la dinámica del acuífero mediante el uso del modelo numérico de flujo. La frecuencia de monitoreo es la misma para todos los pozos seleccionados.

La metodología propuesta en esta tesis para el diseño óptimo de redes de monitoreo de la carga hidráulica se basa en una desarrollada por Herrera (1998) para el diseño de redes de monitoreo de la calidad del agua subterránea, por lo que aunque este tipo de diseños se presentan en el anexo A, por su relevancia para esta tesis, se incluye a continuación.

Herrera (1998), propuso un método para el diseño espacio-temporal óptimo de redes de monitoreo de la calidad del agua subterránea combinando un filtro de Kalman, un modelo numérico de flujo y transporte estocástico, y un método de optimización de inclusiones sucesivas. El modelo estocástico utiliza un campo aleatorio correlacionado espacialmente para las conductividades. Además, la fuente del contaminante también se considera una variable aleatoria. Con la simulación estocástica se obtiene una primera estimación espacio-temporal de las concentraciones del contaminante y de la matriz de covarianza del error de esta estimación. Estas estimaciones se usan en el filtro de Kalman y éste se acopla con el método de optimización para escoger secuencialmente las posiciones y tiempos de muestreo que reducen más la varianza en los puntos de interés. Presenta un problema sintético para el cual diseña un programa de muestreo de un contaminante (es decir, determina dónde y cuándo muestrear) y evalúa las estimaciones obtenidas de las concentraciones de la pluma del contaminante para un periodo de dos años. El objetivo de la red de monitoreo fue minimizar una función de la varianza del error de la estimación (llamada varianza total) en todas las posiciones y tiempos para una malla que cubre el área de estudio. Los resultados muestran que la metodología propuesta produce redes de monitoreo que con un número pequeño de pozos permite obtener estimaciones con poca incertidumbre.

A diferencia de Herrera (1998), en la metodología presentada en esta tesis para el diseño óptimo de redes de monitoreo de la carga hidráulica, la matriz de covarianza espacio-temporal requerida para estimar con el filtro de Kalman, se obtiene de un análisis geoestadístico de los datos, por lo que cae dentro del marco estadístico de diseño. El método de optimización es también de inclusiones sucesivas, pero, a diferencia de Herrera, la evaluación espacio-temporal de las posibles posiciones de monitoreo es implementada en tiempo real (el significado será explicado en detalle más adelante).

Este tipo de metodología es muy útil en el diseño de nuevas redes de monitoreo o en el rediseño de una existente, incluyendo también su calendario óptimo de monitoreo de acuerdo a la función objetivo. La metodología propuesta únicamente requiere datos históricos de la variable, sin importar si el monitoreo fue o no constante en el espacio y/o tiempo. Se desarrolló la metodología para el diseño de redes que capturen la distribución de la carga hidráulica en 2D; este es un enfoque válido para el análisis del movimiento del agua subterránea debido a que normalmente las dimensiones horizontales de los acuíferos son dos o tres órdenes de magnitud

mayores que su dimensión vertical, por lo que para fines prácticos puede considerarse que el flujo del agua subterránea es bidimensional (CFE, 1983). Para los casos de flujo subterráneo en que se tienen identificados gradientes verticales pronunciados, se requeriría una red de monitoreo en 3D.

Hasta donde sabemos, la metodología propuesta es la primera de su tipo que selecciona las posiciones y los tiempos de monitoreo de la carga hidráulica a través del uso de un semivariograma espacio-temporal obtenido de un análisis geoestadístico. Se emplea una estimación espacio-temporal, a diferencia de la mayoría de los trabajos basados en geoestadística para el diseño de redes de monitoreo, que consideran únicamente al espacio o al tiempo.

1.2 Objetivos

El objetivo general de esta tesis es proponer un método para el diseño óptimo de redes de monitoreo de la carga hidráulica utilizando un método geoestadístico espacio-temporal acoplado con un filtro de Kalman.

El objetivo específico es:

- Diseñar y evaluar redes de monitoreo para algunos casos específicos utilizando el método propuesto.

1.3 Estructura de la tesis

El presente trabajo consta de ocho capítulos. El capítulo uno (introducción) trata los antecedentes que dan origen a este trabajo: la importancia del conocimiento de la evolución del agua subterránea, la necesidad del diseño espacio-temporal óptimo de redes de monitoreo, el problema que representa la estimación de la carga hidráulica y su monitoreo (donde además se incluye el estado del arte al respecto); después se presentan los objetivos planteados.

En el capítulo segundo se define la carga hidráulica y se mencionan consideraciones para su correcta interpretación en el análisis del flujo de agua subterránea.

En el capítulo tercero se introducen los conceptos básicos estadísticos utilizados en esta tesis, los fundamentos de la teoría geoestadística primero desde un punto de vista univariado y luego extendiéndola al análisis espacio-tiempo, por último se presenta la teoría del filtro de Kalman.

La metodología geoestadística propuesta para el diseño espacio-temporal óptimo de redes de monitoreo de la carga hidráulica se presenta en el capítulo cuatro.

En el capítulo cinco se presenta una aplicación de la metodología propuesta para un caso sintético basado en un modelo numérico de flujo del acuífero Valle de Querétaro. El modelo numérico de flujo permitió probar la metodología para un caso controlado en el que se cumplen la hipótesis de flujo horizontal, no existencia de recarga profunda, y temperatura y densidad del agua subterránea constantes. Con las simulaciones obtenidas se supuso un alto nivel de información espacio-temporal de la carga hidráulica en el acuífero; además el modelo permitió evaluar el desempeño en el futuro de la red diseñada.

La falta de un monitoreo espacio-temporal sistemático de la carga hidráulica propicia que se presenten muchos huecos de información en las bases de datos espacio-temporales de esta variable. Tratando de emular un caso real, en el capítulo seis se aplicó la metodología para

evaluar un subconjunto de pozos de la red de monitoreo diseñada en el capítulo cinco, considerando únicamente parte de la información espacio-temporal empleada en esa aplicación. Con ayuda de las simulaciones del modelo numérico de flujo fue posible evaluar el desempeño en el pasado y en el futuro de esta red, y realizar comparaciones con los resultados obtenidos en el capítulo cinco.

En el capítulo siete se presenta el diseño de una red de monitoreo de la elevación del nivel del agua para el acuífero Valle de Querétaro empleando la metodología propuesta en este trabajo.

Vale la pena señalar que para los casos sintéticos presentados en este trabajo no se trataron de reflejar condiciones específicas de ningún acuífero, aunque el modelo numérico de flujo utilizado está basado en el acuífero Valle de Querétaro.

Dentro del capítulo ocho se enuncian las conclusiones y recomendaciones generadas de este trabajo.

En el capítulo nueve se enlista la bibliografía consultada para la realización de la investigación.

Al final, esta tesis cuenta con un anexo en donde se presenta una revisión del estado del arte del diseño de redes de monitoreo de la calidad del agua subterránea.

2. LA CARGA HIDRÁULICA

2.1 Definición

La carga hidráulica es una medida específica de la presión del agua o la energía total por unidad de peso sobre un punto geodésico. Usualmente se mide como la elevación del nivel del agua (respecto al nivel del mar), expresada en unidades de longitud. En un acuífero donde las condiciones de salinidad y temperatura son constantes a lo largo del espesor total del acuífero, se puede calcular la carga hidráulica conociendo la profundidad del nivel del agua en un piezómetro y la elevación del brocal con respecto a un nivel de referencia (usualmente el nivel del mar). Aunque la medición realizada es del nivel del agua, la carga hidráulica representa una energía total. La carga hidráulica puede ser usada para determinar un gradiente hidráulico entre dos o más puntos por lo que conocer su distribución espacial en un acuífero puede ayudar a determinar las direcciones del flujo subterráneo, y posibles zonas de recarga y descarga.

En el caso del movimiento del agua a través de un medio poroso bajo condiciones saturadas o no saturadas se requiere que exista una diferencia de energía entre dos puntos en el medio para que se produzca un flujo neto de agua entre ellos. Si recordamos los conceptos básicos de Bernoulli podemos escribir que el nivel de energía en términos de altura o columna de agua está compuesto de tres términos: h_v altura de velocidad, h_p altura de presión y h_z altura o cota geométrica. De esta manera, la energía total por unidad de peso en un punto se puede escribir como:

$$h_T = h_v + h_p + h_z \quad (2.1)$$

Cada término puede ser escrito usando una notación más completa como:

$$h_T = \frac{v^2}{2g} + \frac{P}{\rho g} + z \quad (2.2)$$

donde z es la elevación del fondo del piezómetro con respecto al nivel de referencia, se le llama carga de posición y representa la energía potencial por unidad de peso del agua, $\frac{P}{\rho g}$ es la carga de presión en términos de la altura de la columna de agua, relativa al fondo del piezómetro, y $\frac{v^2}{2g}$ representa la energía de presión por unidad de peso del agua, $\frac{v^2}{2g}$ es la carga de velocidad y representa la energía cinética por unidad de peso del agua, ρ es la densidad del agua (ML^{-3}), v es la velocidad de movimiento del agua (LT^{-1}), y g es la aceleración local de la gravedad (LT^{-2}).

La presión ejercida por una columna de agua usualmente puede aproximarse como estática o de reposo (aún en el océano donde existen olas y corrientes), porque los movimientos producen cambios despreciables en la presión. La presión en cualquier punto de un fluido en reposo (estático) se llama presión hidrostática; en el fondo de un piezómetro este valor se encuentra definido por la siguiente expresión:

$$P = \rho g H \quad (2.3)$$

donde H es la altura de la columna de agua dentro del piezómetro.

Para el caso específico de movimiento del agua en un medio poroso permeable, la velocidad del movimiento es despreciable frente a las otras dos componentes, por lo cual la energía total h_T se escribe en forma resumida como:

$$h = \psi + z \quad (2.4)$$

donde h se conoce comúnmente como cota piezométrica o carga hidráulica y ψ se conoce como carga de presión (Bear, 1979; Freeze y Cherry, 1979; Fetter, 1994; Kresic, 2007).

La Fig. 2.1 muestra la descomposición de la carga hidráulica en un piezómetro, en el cual el punto de medición está ubicado en X.

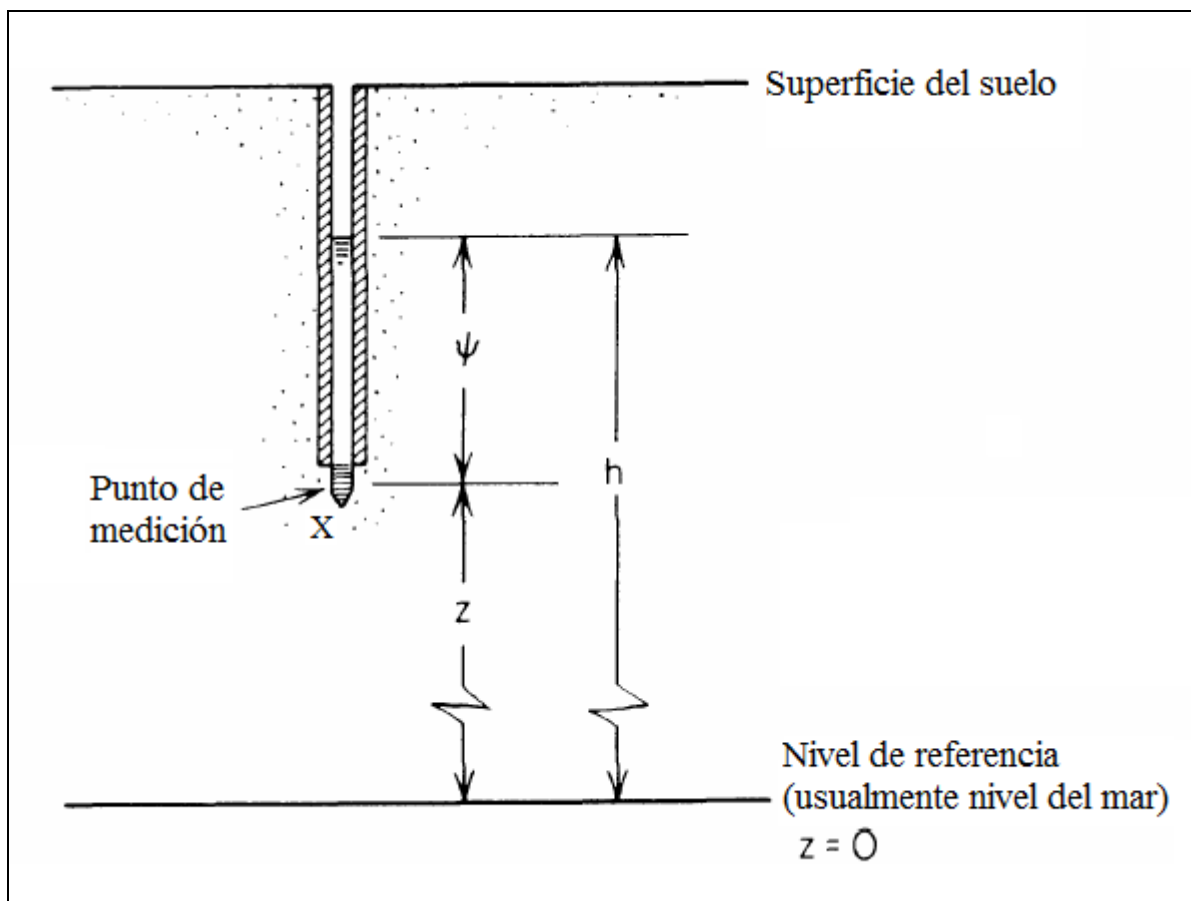


Fig. 2.1 Carga hidráulica en piezómetro

Se define como nivel freático o espejo del agua, en los acuíferos libres, al lugar geométrico de los puntos donde la presión hidrostática es igual a la presión atmosférica. Por encima del nivel freático los poros del suelo pueden contener aire o agua, por lo cual se la llama zona de aireación. En la zona freática, por debajo del espejo del agua, los intersticios están llenos de agua por lo cual se le llama también zona de saturación.

En los acuíferos confinados el agua se encuentra sometida a una presión generalmente mayor a la atmosférica por el peso de la sobrecarga y el propio nivel hidrostático. Si un pozo llega a penetrar la capa confinante, el agua subirá hasta alcanzar el nivel piezométrico, que es el equivalente artesiano del nivel freático, y si ese nivel se encuentra por encima de la superficie del terreno, el pozo descargará como pozo surgente.

El nivel piezométrico o potenciométrico se define como la elevación del nivel de agua en un pozo, o la elevación a la que el nivel de agua de un pozo artesiano ascenderá en un tubo lo suficientemente alto para que el flujo se detenga (Fig. 2.2).

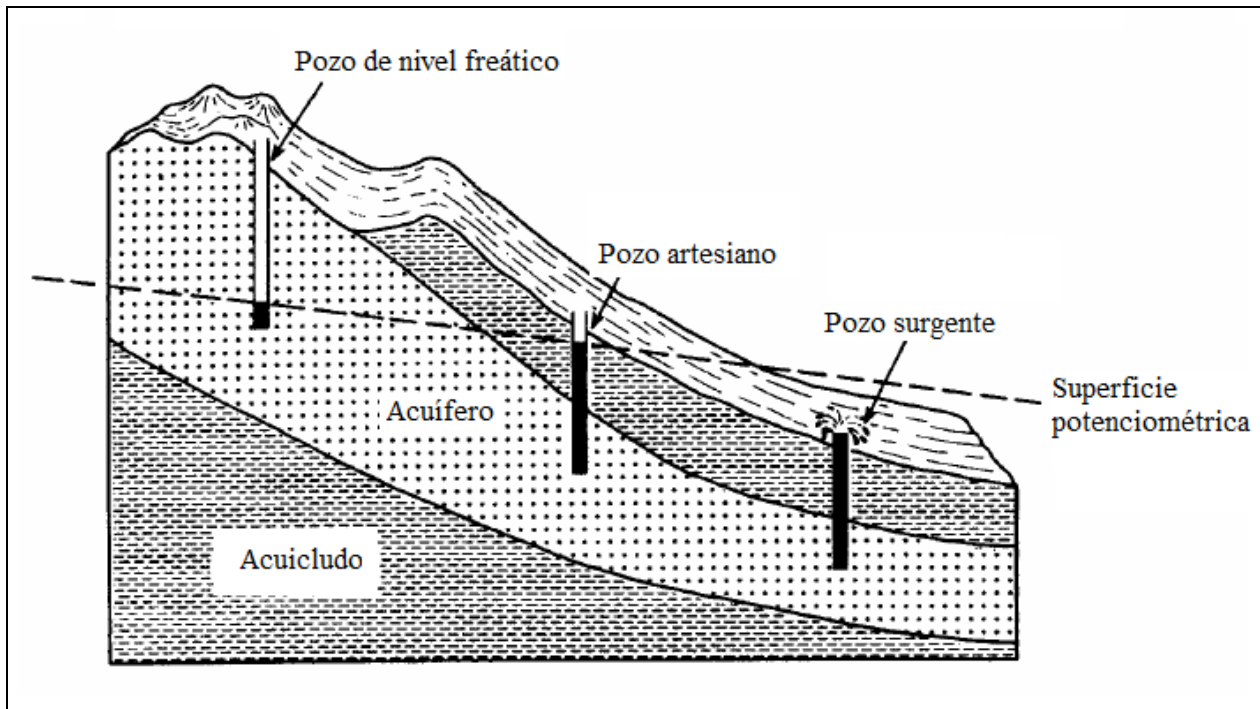


Fig. 2.2 Superficie potenciométrica y tipo de pozos

En un acuífero confinado la carga hidráulica puede disminuir pero el nivel piezométrico puede permanecer sobre la unidad confinante. En este caso una cantidad de agua es liberada desde el almacenamiento y el acuífero permanece saturado.

2.2 Consideraciones para la correcta interpretación de mediciones de la carga hidráulica

Los datos usados para construir mapas de nivel freático o piezométrico son elevaciones del nivel de agua medidas en algunos pozos habilitados en la zona de estudio. No todos los pozos son útiles para este efecto. Por ejemplo, si un pozo perfora más de un acuífero el nivel del agua dentro de él corresponderá a un promedio del nivel de energía en cada acuífero atravesado. Es importante que los datos que se vayan a comparar entre sí correspondan al mismo tipo de flujo, ya sea local, intermedio o regional.

Si se va a medir el nivel del agua en un pozo que normalmente es utilizado para abastecimiento, debe asegurarse que no haya existido bombeo durante un tiempo suficiente para que el nivel de agua se recupere a lo que se conoce como nivel estático (nivel del agua en un

pozo antes de iniciar el bombeo). El nivel estático del agua es equivalente al nivel freático para un acuífero libre o nivel piezométrico para un acuífero confinado en condiciones naturales.

El flujo subterráneo se compone por zonas de recarga, de tránsito y de descarga. Es obvio que en las zonas de recarga y descarga se presentan flujos descendentes y ascendentes, respectivamente. En cuanto a la zona de tránsito se refiere, generalmente se considera un flujo prácticamente horizontal con un cierto gradiente hidráulico que propicia el movimiento. Carrillo-Rivera (2000) demostró que la ecuación del balance de agua subterránea empleada para determinar algunos parámetros hidráulicos de dos acuíferos de México, ofrece resultados poco satisfactorios si se desprecia o no es adecuadamente definido el flujo vertical. Carrillo-Rivera identificó el flujo vertical en un acuífero por un incremento de la temperatura del agua subterránea, que llega a superar la temperatura ambiental en hasta 23 °C y estableció que la presencia de fallas y la extracción pueden favorecer el flujo vertical.

Los niveles del agua medidos en pozos de bombeo o artesianos que penetran acuíferos confinados son generalmente afectados por cambios en la temperatura de la columna de agua. Este efecto de la temperatura llega a ser significativo en pozos profundos. El problema de conocer la carga hidráulica puede evitarse monitoreando la presión en el fondo del pozo, pero esta opción es costosa y consecuentemente no siempre posible. Kawecki (1995), propuso un método para la corrección de la temperatura basado en un modelo que se supone describe el enfriamiento en el pozo durante su recuperación. El método está diseñado para determinar la transmisividad de un acuífero usando los datos de recuperación en el bombeo, y para calcular el abatimiento real durante el bombeo. Los únicos requerimientos adicionales a los datos de una prueba de bombeo son la medición de la temperatura inicial y estabilizada de la descarga. A pesar de las múltiples suposiciones en el modelo de enfriamiento, el método ha dado aparentemente resultados exitosos en algunos casos de estudio. Aunque se requiere de más trabajo para probar su efectividad, el método ha resultado útil como una segunda alternativa para los casos donde no sea posible la habilitación de pozos de observación o el monitoreo de la presión en el fondo del pozo de prueba.

Generalmente existe un gradiente natural geotérmico de incremento en la temperatura con la profundidad. Es un gradiente variable que depende de muchos factores. Los valores típicos de un gradiente promedio citados en la literatura se encuentran en el rango de 1.8 a 3.7 °C por cada 100 m (Kawecki, 1995). El efecto de la temperatura es aproximadamente proporcional a la profundidad del pozo. En un pozo de 100 m de profundidad el efecto de temperatura típicamente será de 1 o 2 cm, en un pozo de 500 m será de 0.5 a 1.0 m. En el caso de un acuífero que tenga un espesor de 1,000 m sería del orden de 5 a 10 m. Una forma para calcular el gradiente natural geotérmico es a través de geotermómetros. Los geotermómetros son herramientas que permiten estimar indirectamente la temperatura mínima de equilibrio a la que el agua llegó, con base en esta temperatura es factible determinar la profundidad a la que el agua circuló; eso se hace con base en la química del agua. Esto es, los geotermómetros calculan la temperatura del agua a la cual se alcanzó el último equilibrio de la interacción roca-agua en función de la temperatura o presión. Pueden estimarse a través de concentraciones de sílice (cuarzo o calcedonia) y proporciones relativas de cationes: Na-K-Ca, y Na/K (Carrillo-Rivera *et al.*, 1996).

Se tienen muchas limitaciones para definir en campo el valor de la carga hidráulica dado que se tienen diferentes tipos de flujo de agua subterránea en la vertical con diferente densidad de agua y que usualmente están conectados (en esa dimensión). A continuación se describe el ajuste que debe hacerse a la carga hidráulica cuando la densidad es variable.

Corrección de datos de carga hidráulica de densidad variable

El término de la carga de presión ψ que aparece en la ecuación 2.4, depende de la densidad del agua, la cual puede variar en función tanto de la temperatura como de la composición química (en particular, la salinidad). Esto significa que el cálculo de la carga hidráulica (si el pozo no penetra todo el acuífero) depende de la densidad del agua dentro del piezómetro. Si una o más mediciones de la carga hidráulica van a ser comparadas, necesitan ser estandarizadas, usualmente a su carga de agua dulce, la cual puede ser calculada de la siguiente forma (Freeze y Cherry, 1979):

$$h_{ad} = \psi \frac{\rho}{\rho_{ad}} + z \quad (2.5)$$

donde h_{ad} es la carga de agua dulce, ρ_{ad} es la densidad del agua dulce, ρ es la densidad del agua sin estandarizar, z es la carga de posición, y ψ es la carga de presión a estandarizar.

2.3 Presencia de tendencia en la carga hidráulica

El movimiento del agua subterránea se explica por los gradientes de energía que presenta la distribución de la carga hidráulica en el subsuelo, propiciados ya sea por la recarga/descarga vertical natural o artificial, o por el intercambio de agua entre acuíferos. La presencia de estos gradientes generalmente se manifiesta en los datos de la carga hidráulica en forma de tendencia. Una de las hipótesis de un análisis geoestadístico consiste en considerar al fenómeno físico como una función estacionaria de segundo orden. Esto implica que el valor esperado de la variable debe ser constante sobre toda el área de estudio, o en vecindades locales, lo cual no ocurre cuando se presenta una tendencia en los datos. Una manera para eliminar este problema consiste en ajustar un polinomio a los datos (cálculo de la tendencia) para trabajar con los residuos resultantes, éstos son tratados como realizaciones de variables aleatorias estacionarias correlacionadas, el semivariograma es calculado y modelado para el kriging. Finalmente la tendencia se agrega de nuevo a los estimados del kriging. El método es conocido como “kriging de regresión” o “kriging residual”. Usualmente el ajuste del polinomio se realiza mediante el método de los mínimos cuadrados. Mínimos cuadrados es una técnica de optimización matemática que, dada una serie de mediciones, intenta encontrar una función que se aproxime a los datos (un "mejor ajuste"). Intenta minimizar la suma de los cuadrados de las diferencias ordenadas (llamadas residuos) entre los valores generados por la función y los correspondientes a los datos.

Si el acuífero no presenta abatimientos locales significativos, los modelos de semivariograma de residuos resultado de ajustes de polinomios de primer o segundo orden son suficientes para obtener buenos resultados en la estimación. Sin embargo, cuando se tienen abatimientos locales significativos, estos ajustes no son suficientes. Lo anterior puede explicarse del hecho de que los polinomios de primer y segundo grado sólo ajustan el comportamiento global del nivel del agua subterránea en acuíferos y no los abatimientos locales. El ajustar únicamente el comportamiento global implica que las estimaciones se obtengan a la altura de la superficie del polinomio propuesto por lo que cuando se presentan comportamientos locales, las estimaciones en esas zonas quedarán muy por encima o muy por debajo de los datos de niveles del agua subterránea, como ocurre en el trabajo de Mendoza (2008).

Tonkin y Larson (2002) propusieron ajustar este efecto mediante la combinación del método de regresión lineal (krigeado simple) y de forma puntual utilizar una función logarítmica similar a la ecuación de Theis. El método propuesto calcula la tendencia sumando al modelo lineal la componente de descenso del nivel provocado en el pozo que está bombeando. Los resultados muestran que las estimaciones con este método son mejores en las zonas donde se encuentra un pozo de bombeo. Una de las limitantes para el empleo de este método es que considera al acuífero como homogéneo.

Otra alternativa consiste en trabajar con los residuos del ajuste de polinomios de alto grado. La idea de trabajar con polinomios de más alto grado es obtener una superficie que represente lo mejor posible la distribución de la carga hidráulica sobre todo el acuífero. Sin embargo, al intentar resolver los sistemas de ecuaciones que se plantean en el ajuste de este tipo de polinomios se presentan inestabilidades numéricas, asociadas al alto número de condicionamiento de las matrices que intervienen. Al momento de efectuar un análisis numérico, el número de condicionamiento asociado con un problema es una medida de la factibilidad de tratar con ese problema mediante cómputo digital, esto habla de qué tan bien condicionado numéricamente es un problema. Un problema con un número de condicionamiento bajo se dice que es bien condicionado, mientras que un problema con un número de condicionamiento alto se dice que es mal condicionado. Es posible reducir el número de condicionamiento debido a la magnitud de los datos de longitud y latitud de los pozos escalando las coordenadas, o trasladando el origen del sistema de coordenadas.

López y Samper (1997) propusieron técnicas para reducir el número de condicionamiento de la matriz que se resuelve para estimaciones del nivel del agua subterránea mediante kriging universal (desarrollado para un campo aleatorio con tendencia desconocida). Emplearon el escalamiento de las coordenadas de los datos, la factorización QR de la matriz que incluye los datos del polinomio que ajusta los datos, y funciones derivadas de los eigenvectores de la matriz del semivariograma para producir números de condicionamiento pequeños y muy similares entre ellos para la matriz que se resuelve para la estimación mediante kriging universal. Los resultados de la validación cruzada de los datos de carga hidráulica son iguales para las primeras dos técnicas y los peores resultados se obtienen para el tercer caso.

Existen dos desventajas del kriging residual o de regresión. Primero, la tendencia generalmente es estimada por mínimos cuadrados ordinarios (MCO), la cual es insesgada, pero no se obtienen estimados con varianza mínima a menos que los sitios de muestreo hayan sido seleccionados mediante un plan de muestreo en particular (Webster y Oliver, 2007). La segunda desventaja es que los estimados de las semivarianzas obtenidas de los residuos de la tendencia tienen sesgo. Esto se debe a que dependen de una manera no lineal de los parámetros de la tendencia, los cuales son estimados con error. Como resultado, el semivariograma es subestimado, y el sesgo se incrementa conforme la distancia del intervalo se incrementa. Una solución propuesta a estos problemas es usar mínimos cuadrados generalizados para estimar los parámetros de la tendencia. El método de los mínimos cuadrados generalizados requiere un semivariograma de los residuos, de tal manera que se sigue un procedimiento iterativo. Se obtienen estimados de los MCO, y se ajusta un semivariograma a los residuos. Este semivariograma se usa en mínimos cuadrados generalizados para re-estimar los parámetros de la tendencia, y el procedimiento se repite hasta que los estimados se estabilizan. Esta aproximación reduce la varianza del error de los parámetros de la tendencia, pero no remueve el sesgo de los estimados en el semivariograma porque aún dependen de los parámetros de la tendencia. Este sesgo podría no importar cuando los datos son densos porque es típicamente muy pequeño en

distancias de intervalo pequeñas, en esos casos solo los puntos a pequeñas distancias del punto de estimación tienen pesos apreciables en los sistemas de kriging. Finalmente, aún si se ignora el sesgo de la varianza de las predicciones tanto de la tendencia como del kriging de los residuos, el kriging de regresión no permite combinarlos en una predicción válida de la varianza para la estimación del kriging, aunque se pudiera computar la varianza del kriging universal (Webster y Oliver, 2007).

En resumen, para predecir valores de variables ambientales que tienen una tendencia espacial pronunciada y una variación aleatoria espacialmente dependiente, se requieren obtener estimados de la tendencia con varianza mínima para estimar el semivariograma de los residuos de la tendencia sin sesgo y para estimar, con varianza conocida, la suma de la tendencia y la variación aleatoria en sitios sin muestrear. Una manera práctica de hacerlo es computar el mejor estimador lineal insesgado empírico (MELIE) con un semivariograma estimado por residuos de máxima verosimilitud (REMV). El método fue recomendado por Stein (1999), y ahora con un mayor poder computacional y software complicado se está volviendo accesible en la práctica (Webster y Oliver, 2007).

Para evitarse problemas asociados a trabajar con datos con tendencia, Kashyap y Vakkalagadda (2009) propusieron un modelo de semivariograma de la carga hidráulica calculado a partir de datos de transmisividad y de la desviación estándar de flujo (σ_w). Sin embargo, el punto débil del modelo es el requerimiento de σ_w . Aunque este estadístico pudiera no ser conocido explícitamente, pudiera estimarse de una variación conocida de descensos/recarga en el área de estudio. Las suposiciones del modelo propuesto son: condición de estado permanente, acuífero homogéneo, y flujo con una distribución normal y media cero.

La extensión de este tipo de técnicas para estimación espacio-temporal resulta un buen campo para desarrollar investigaciones futuras.

3. TEORÍA GEOESTADÍSTICA Y DEL FILTRO DE KALMAN

En este trabajo, la geoestadística se emplea para obtener la matriz de covarianza espacio-temporal de la carga hidráulica que necesita el método propuesto para el diseño óptimo de redes de monitoreo. Sin embargo, una vez obtenida esta matriz, la optimización ocupa como método de estimación al filtro de Kalman. Ambos temas se abordan en este capítulo en el que primero se presentan algunos conceptos básicos de estadística, después se habla de geoestadística y finalmente se trata el filtro de Kalman.

3.1 Conceptos básicos de estadística

En esta sección se introducen los conceptos básicos estadísticos que se utilizarán en esta tesis, los cuales han sido tomados de Samper y Carrera (1990), a menos de que se especifique lo contrario en el texto.

Funciones de densidad y de distribución

La función de distribución $F(x)$ de una variable aleatoria X se define como $F(x) = \mathbf{Prob}(X \leq x)$.

Se dice que X es una variable aleatoria discreta si existe un número finito de valores x_1, x_2, \dots, x_n tales que $\mathbf{Prob}(X = x_i) = a_i \geq 0 \quad \forall i, \sum_{i=1}^n a_i = 1$.

Si $\mathbf{Prob}(X = x) = 0$ para toda x , la variable aleatoria se dice que es continua. Si existe una función $f(t)$ no negativa definida entre $-\infty < t < \infty$ tal que la función de distribución de X se puede expresar como $F(x) = \int_{-\infty}^x f(t)dt$, se dice que $f(t)$ es la función de densidad de X .

Dadas n variables X_1, X_2, \dots, X_n , la función de distribución n -dimensional se define como:

$$F(x_1, x_2, \dots, x_n) = \mathbf{Prob}(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n).$$

Esperanza matemática y momentos de orden n

La esperanza matemática $E(X)$ o valor esperado m de una variable aleatoria X se define como:

$$m = E(X) = \sum_i x_i \mathbf{Prob}(X = x_i) \quad \text{si } X \text{ es discreta,}$$

$$m = E(X) = \int_{-\infty}^{\infty} tf(t)dt \quad \text{si } X \text{ es continua.}$$

La esperanza matemática o media es una medida de la tendencia central de una variable aleatoria. Otras medidas de tendencia central son la mediana y la moda. Para una variable aleatoria continua la moda x_0 es aquel valor de x para el que $f(x)$ es máxima (valor más probable) mientras que la mediana x_e es aquel valor para el que $F(x_e) = 0.5$.

El momento de orden n de una variable aleatoria X continua con respecto al origen se define como $E(X^n) = \int_{-\infty}^{\infty} t^n f(t) dt$.

Para variables discretas la definición es idéntica salvo que $f(t)$ se sustituye por $\mathbf{Prob}(X = t)$ y la integral por la sumatoria. De forma similar, se definen los momentos de orden n con respecto a la media m como $\alpha_n = E[(X - m)^n] = \int_{-\infty}^{\infty} (t - m)^n f(t) dt$.

El momento de orden uno con respecto a la media es 0, mientras que el momento de orden dos es la varianza σ^2 , es decir, $\sigma^2 = E[(X - m)^2] = \int_{-\infty}^{\infty} (t - m)^2 f(t) dt$.

La desviación estándar σ se define como la raíz cuadrada de la varianza. La varianza de una variable aleatoria es una medida de dispersión de su función de distribución. El coeficiente de variación λ se define como $\frac{\sigma}{m}$.

Otras características de la distribución de una variable aleatoria son el coeficiente de asimetría γ_1 y el coeficiente de curtosis γ_2 . El primero de ellos mide la simetría de la distribución y se define como $\gamma_1 = \frac{\alpha_3}{\sigma^3}$, donde α_3 es el momento de orden tres con respecto a la media. El coeficiente de curtosis mide el grado de apuntamiento y se define como $\gamma_2 = \frac{\alpha_4}{\sigma^4} - 3$.

Funciones de densidad y de distribución condicional

La función de distribución condicional de X dado $Y = y$ se define como $F_{x|y}(x|y) = \frac{\mathbf{Prob}(X \leq x, Y = y)}{\mathbf{Prob}(Y = y)}$, si $\mathbf{Prob}(Y = y) > 0$.

La función de densidad condicionada de X dado $Y = y$ se define como $f_{x|y}(x|y) = \frac{f_{xy}(x, y)}{f_y(y)}$.

Si $g(X)$ es una función de X , entonces el valor esperado condicionado de $g(X)$ dado que $Y = y$ se denota como $E[g(X)|Y = y]$ y se calcula por medio de $E[g(X)|Y = y] = \int_{-\infty}^{\infty} g(x) f_{x|y}(x|y) dx$.

Proceso gaussiano

Según Gelb *et al.* (1996), un proceso gaussiano es un proceso aleatorio continuo caracterizado por la propiedad de que sus funciones de distribución de probabilidad conjunta de todos los órdenes son distribuciones normales multidimensionales. Para un proceso gaussiano, la distribución de la variable aleatoria X es la distribución normal, para la cual la función de densidad se expresa por:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] \quad (3.1)$$

Se dice que n variables aleatorias X_1, X_2, \dots, X_n son normales en su conjunto o que tienen una distribución conjunta normal n -variable si su función de densidad n -dimensional tiene la forma:

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} |\mathbf{P}|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\mathbf{m})^T \mathbf{P}^{-1}(\mathbf{x}-\mathbf{m})\right] \quad (3.2)$$

donde:

\mathbf{x} = vector n -dimensional (x_1, x_2, \dots, x_n) .

\mathbf{m} = vector n -dimensional de las medias, $m_i = E(X_i)$.

\mathbf{P} = Matriz $n \times n$ de covarianza de las n variables aleatorias, definida como $\mathbf{P}_{ij} = Cov(X_i, Y_j)$.

\mathbf{P}^{-1} = Matriz inversa de \mathbf{P} , y $|\mathbf{P}^{-1}|$ su determinante.

El superíndice T quiere decir transpuesta.

Todas las propiedades de los procesos aleatorios gaussianos quedan definidas por los momentos de distribución de primer y segundo orden.

Por su conveniencia analítica, los procesos gaussianos son los más frecuentemente encontrados en el análisis de sistemas. Por ello, es apropiado introducir la notación que debe tener la información de la ecuación 3.2. La notación usada en la secuela es $\mathbf{x} \sim N(\mathbf{m}, \mathbf{P})$, lo que indica que \mathbf{x} es un vector aleatorio gaussiano (normal) con media \mathbf{m} y covarianza \mathbf{P} .

Ruido blanco gaussiano

Un proceso de ruido blanco gaussiano es un proceso gaussiano de variables aleatorias no correlacionadas en el tiempo, con media cero y covarianza finita (la cual se denota como \mathbf{P}).

Formalmente, $\{\mathbf{e}_n\}$ es un proceso de ruido blanco si $E(\mathbf{e}_n) = 0$ para toda n , $E(\mathbf{e}_i^T \mathbf{e}_i) = \mathbf{P}$ y $E(\mathbf{e}_i \mathbf{e}_j^T) = 0$ para i diferente de j .

3.2 Geoestadística

La geoestadística se define como una rama de la estadística aplicada que se especializa en el análisis y la modelación de la variabilidad espacial en ciencias de la Tierra. Su objeto de estudio es el análisis y la predicción de fenómenos en espacio y/o tiempo (Díaz, 2002). En ella se aplica la teoría de las variables regionalizadas; se busca la correlación que existe entre la información con que se cuenta.

Variable regionalizada

Se nombra como variable regionalizada $z(x)$ a la variable distribuida en el espacio de manera tal que presenta una estructura espacial de correlación.

Función aleatoria

Si a cada punto x que pertenece a un dominio en el espacio le hacemos corresponder una variable aleatoria $z(x)$, entonces el conjunto de variables aleatorias espacialmente distribuidas $\{z(x), x \in R^n\}$ será una función aleatoria $Z(x)$.

Función de distribución y momentos de una función aleatoria

Sea $Z(x)$ una función aleatoria definida en R^n , entonces el vector aleatorio $\{Z(x_1), Z(x_2), \dots, Z(x_k)\}$ se caracteriza por su función de distribución de probabilidad k -variada: $F_{Z(x_1), Z(x_2), \dots, Z(x_k)}(z_1, z_2, \dots, z_k) = \mathbf{Prob}[Z(x_1) \leq z_1, Z(x_2) \leq z_2, \dots, Z(x_k) \leq z_k]$.

El conjunto de todas las distribuciones para todo valor de k y para cualquier selección de puntos en R^n constituye la ley espacial de probabilidad de la función aleatoria $Z(x)$. En geoestadística lineal son suficientes los dos primeros momentos de la distribución de $Z(x)$. De hecho, en la mayoría de las aplicaciones prácticas la información disponible no permite inferir momentos de mayor orden.

El momento de primer orden es la esperanza matemática definida como $E[Z(x)] = m(x)$, en donde la función $m(x)$ se conoce con el nombre de tendencia. Para ilustrar la dependencia espacial de la esperanza matemática se da la expresión de ésta para $Z(x)$ cuando es una función de variables continuas, $E[Z(x)] = \int_{-\infty}^{+\infty} z(x) f(z(x)) dz$.

Los tres momentos de segundo orden de $Z(x)$ considerados en geoestadística son:

a) La varianza o momento de segundo orden de $Z(x)$ respecto a $m(x)$:

$$\sigma^2(x) = \text{Var}[Z(x)] = E\{[Z(x) - m(x)]^2\}.$$

b) La covarianza de dos variables aleatorias $Z(x_i)$ y $Z(x_j)$, $C(x_i, x_j)$, definida como:

$$C(x_i, x_j) = E\{[Z(x_i) - m(x_i)][Z(x_j) - m(x_j)]\},$$

es en general una función de x_i y x_j . Esta función se llama a veces función de autocovarianza y es una medida estadística de la correlación entre dos variables.

c) El semivariograma $\gamma(x_i, x_j)$ que se define como:

$$\gamma(x_i, x_j) = \frac{1}{2} E\{[Z(x_i) - Z(x_j)]^2\},$$

es también conocido como función de semivarianza, semivariograma o variograma.

El semivariograma es la herramienta central de la geoestadística. Al igual que la covarianza, el semivariograma es una función que describe el grado de correlación espacial de la propiedad que se estudia.

Funciones aleatorias estacionarias

Se dice que una función aleatoria es estrictamente estacionaria si su función de distribución es invariante respecto a cualquier traslación de vector h , o lo que es lo mismo, la función de distribución del vector aleatorio $[Z(x_1), Z(x_2), \dots, Z(x_k)]$ es idéntica a la del vector $[Z(x_1 + h), Z(x_2 + h), \dots, Z(x_k + h)]$ para cualquier h . Sin embargo, puesto que la geoestadística lineal se basa en los dos primeros momentos de la función aleatoria, es suficiente imponer que estos dos momentos existen y limitar la hipótesis de estacionariedad a los dos primeros momentos.

Funciones aleatorias estacionarias de segundo orden

Se dice que una función aleatoria $Z(x)$ es estacionaria de segundo orden si:

a) $E[Z(x)]$ existe y no depende de x , es decir,
 $E[Z(x)] = m$ para todo x .

b) Para toda pareja de variables aleatorias $\{Z(x+h), Z(x)\}$ su covarianza existe y sólo depende del vector separación h , es decir,
 $C(h) = C(x+h, x) = E[Z(x+h)Z(x)] - m^2$.

La estacionariedad de la varianza implica que la varianza existe, es finita y no depende de x , es decir $\sigma^2 = C(0) = \text{Var}[Z(x)]$. Asimismo, bajo esta hipótesis el semivariograma también es estacionario y se cumple que:

$$\gamma(h) = \gamma(x+h, x) = \frac{1}{2} E\{[Z(x+h) - Z(x)]^2\} = \frac{1}{2} \text{Var}[Z(x+h) - Z(x)].$$

Además existe una relación directa entre el semivariograma y la función de covarianza,
 $\gamma(h) = C(0) - C(h)$.

Funciones aleatorias intrínsecas

Para las funciones aleatorias $Z(x)$ que representan fenómenos físicos que tienen una capacidad casi ilimitada de variación, no están definidas la varianza ni la covarianza. Sin embargo, existen casos en que sus incrementos o diferencias $Z(x+h) - Z(x)$ tienen una varianza finita. Esta es la motivación para definir el concepto de funciones aleatorias intrínsecas como aquéllas cuyos incrementos $Z(x+h) - Z(x)$ tienen esperanza matemática y varianza definidas e independientes de x para todo vector h , es decir:

$$E[Z(x+h) - Z(x)] = m(h),$$

$$\text{Var}[Z(x+h) - Z(x)] = g(h).$$

La función $m(h)$ es la función de tendencia que es necesariamente lineal en h ya que $Z(x+h_1+h_2) - Z(x) = [Z(x+h_1+h_2) - Z(x+h_1)] + [Z(x+h_1) - Z(x)]$ y tomando la esperanza matemática se tiene:

$$m(h_1+h_2) = m(h_2) + m(h_1).$$

Aunque no es indispensable, es habitual suponer $m(h) = 0$. Si no fuese así, la función aleatoria se definiría como $[Z(x) - m(x)]$.

Cuando $m(h) = 0$, el valor esperado de las diferencias es

$$E[Z(x+h) - Z(x)] = 0$$

y la varianza de las diferencias es

$$\text{Var}[Z(x+h) - Z(x)] = 2\gamma(h).$$

Este conjunto de condiciones se conoce como hipótesis intrínseca.

Una función aleatoria estacionaria de segundo orden es siempre intrínseca. Sin embargo, el recíproco no siempre es cierto.

Estimación del semivariograma para variables intrínsecas

La estimación más común del semivariograma, a la que llamaremos semivariograma muestral, es la siguiente:

$$\gamma^*(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i+h) - Z(x_i)]^2 \quad (3.3)$$

donde $N(h)$ es el número de pares $Z(x_i)$ y $Z(x_i+h)$ que están separados por una distancia $h = |h|$, conocida como “intervalo”. Este estimador del método de momentos de Matheron se caracteriza por su robustez.

Si el semivariograma es acotado, el valor del intervalo a partir del cual el semivariograma no se incrementa es conocido como alcance o rango (radio de correlación) y marca el límite de la dependencia espacial de la propiedad. La varianza máxima es conocida como “sill” o meseta del semivariograma y teóricamente debe coincidir con la varianza a priori σ^2 de la muestra de la función aleatoria $Z(x)$.

Un aspecto del semivariograma que es importante señalar es que por definición $\gamma(0) = 0$ pero en la práctica el semivariograma muestral $\gamma^*(h)$ cuando $|h|$ tiende a cero no necesariamente se anula. Esto es conocido como efecto “pepita” (nugget en inglés) o microvarianza. En principio esto puede ocurrir solamente si existen discontinuidades en la función aleatoria. En la práctica su existencia se debe a la variación espacial que no puede explicar el semivariograma debido a la escala del muestreo.

Deducción del estimador del método de momentos de Matheron

Una forma de medir la dispersión de información consiste en calcular el momento de inercia de los datos con respecto a una línea a 45°, $y = x$ (dispersión nula), este momento de inercia puede calcularse con la siguiente fórmula:

$$\text{Momento de inercia} = \frac{1}{2n} \sum_{i=1}^n (x_i - y_i)^2$$

que es la mitad del promedio de las diferencias cuadradas entre las coordenadas x y y de cada uno de los n pares de puntos en una gráfica de dispersión, el factor $\frac{1}{2}$ es consecuencia del hecho de que el momento de inercia se calcula con la distancia perpendicular de los puntos a una línea a 45°, ver Fig. 3.1.

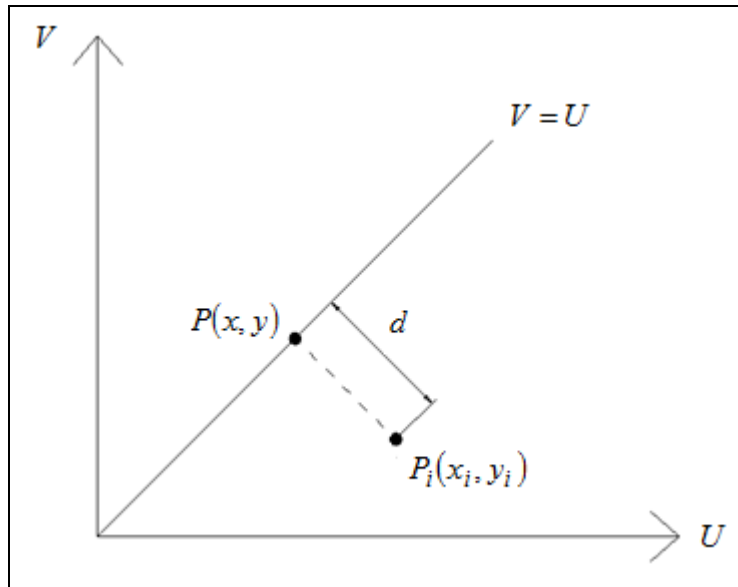


Fig. 3.1 Gráfica de dispersión de las variables U y V

El momento de inercia de una masa alrededor de un eje se calcula con $I = md^2$, siendo m la masa y d la distancia perpendicular del eje a la masa. Para el caso de una gráfica de dispersión se supone que todos los puntos tienen la misma masa y para simplificar se dice que $m = 1$ por lo que el momento de inercia se calcula con la expresión $I = d^2$.

La distancia perpendicular a una línea a 45° ($y = x$) de un punto P_i con coordenadas (x_i, y_i) se puede calcular con la expresión $d = \sqrt{\frac{1}{2}(x_i - y_i)^2}$, por lo que el momento de inercia de un punto con respecto a un eje a 45° (línea que representa una dispersión nula) es $I = \frac{1}{2}(x_i - y_i)^2$. Ahora bien, la suma de los momentos de inercia de varios puntos en una gráfica de dispersión de dos atributos da lugar a la ecuación 3.3 para el cálculo del semivariograma.

Función aleatoria espacio-temporal

Considere un dominio espacial D , y un dominio temporal T . Una variable aleatoria espacio-temporal (VA) $z(\mathbf{x}, t)$ es una variable que depende de la posición espacial $\mathbf{x} = (x, y) \in D$ y de un instante de tiempo $t \in T$. Una función aleatoria espacio-temporal (FA) $\{Z(\mathbf{x}, t), (\mathbf{x}, t) \in D \times T\}$ es definida como un grupo de VA dependientes $z(\mathbf{x}, t)$. La VA $z(\mathbf{x}, t)$ puede caracterizarse completamente si se conoce su función de distribución, que provee la probabilidad de la variable z en la posición espacial \mathbf{x} y el instante de tiempo t :

$$F(\mathbf{x}, t; z) = \mathbf{Prob}\{Z(\mathbf{x}, t) \leq z\}, \forall z(\mathbf{x}, t) \in D \times T \quad (3.4)$$

El concepto de función aleatoria espacial es análogo al de función aleatoria espacio-temporal, de esta manera, las hipótesis de estacionariedad de segundo orden y la intrínseca también aplican (De Iaco *et al.*, 2001).

Semivariograma espacio-temporal muestral

En geoestadística clásica, la covarianza y el semivariograma son funciones que describen la correlación espacial de la propiedad estudiada, en la región de interés. En la conceptualización espacio-temporal, esta definición se extiende para considerar también la correlación temporal.

El semivariograma muestral para los incrementos $\Delta\mathbf{x} = (\Delta x, \Delta y)$ y Δt se basa en la siguiente expresión:

$$\gamma_{st}(\Delta\mathbf{x}, \Delta t) = \frac{1}{2N(\Delta\mathbf{x}, \Delta t)} \sum_{i=1}^{N(\Delta t)} \sum_{k=1}^{N(\Delta\mathbf{x})} [Z(\mathbf{x}_k + \Delta\mathbf{x}, t_i + \Delta t) - Z(\mathbf{x}_k, t_i)]^2 \quad (3.5)$$

donde $N(\Delta\mathbf{x}, \Delta t)$ es el número de pares $z(\mathbf{x}_k, t_i)$ y $z(\mathbf{x}_k + \Delta\mathbf{x}, t_i + \Delta t)$ separados por los incrementos $\Delta\mathbf{x}$ y Δt .

Para capturar la estructura del semivariograma espacio-temporal muestral es necesario emplear modelos autorizados.

Modelos espacio-temporales

La geoestadística ofrece una gran variedad de métodos para modelar datos espacialmente, pero en procesos espacio-temporales, este tipo de aproximación que considera únicamente al espacio, puede conducir a la pérdida de información muy valiosa en la dimensión temporal.

Modelo lineal

Rouhani y Hall (1989) fueron los primeros investigadores que plantearon estimar niveles del agua subterránea utilizando información conjunta en el dominio del espacio y el tiempo. Los autores muestran utilizaron un modelo del tipo lineal para realizar esta estimación:

$$C_{st}(h_s, h_t) = C_s(h_s) + C_t(h_t) \quad (3.6)$$

donde $C_{st}(h_s, h_t)$ es la función de covarianza espacio-temporal, $C_s(h_s)$ y $C_t(h_t)$ son las funciones de covarianza espacial, y covarianza temporal, respectivamente. El modelo lineal presenta problemas ya que la matriz de covarianza, para ciertas configuraciones, es singular (Myers y Journel, 1990; Rouhani y Myers, 1990).

Modelo producto

Una forma sencilla de modelar la covarianza en espacio-tiempo es expresarla como el producto de una componente espacial y una temporal (De Cesare *et al.*, 1997).

$$C_{st}(h_s, h_t) = C_s(h_s)C_t(h_t) \quad (3.7)$$

donde la dependencia espacial es separada de la temporal, en este modelo se puede asumir un alcance independiente del tiempo.

El modelo producto en la parte espacial C_s , es positivo definido en el espacio R^d y C_t es positivo definido en R ; se pueden combinar modelos permitidos de covarianza espacial y modelos de covarianza temporal en forma de productos para obtener modelos de covarianza espacio-temporal (Cressie, 1991).

Modelo producto-suma

Un modelo geoestadístico muy general de covarianza espacio-temporal es el producto-suma (De Cesare *et al.*, 2001; De Iaco *et al.*, 2001). Este modelo tiene la siguiente estructura:

$$C_{st}(h_s, h_t) = k_1 C_s(h_s)C_t(h_t) + k_2 C_s(h_s) + k_3 C_t(h_t) \quad (3.8)$$

o de manera equivalente, para semivariogramas espacio-temporales:

$$\gamma_{st}(h_s, h_t) = [k_2 + k_1 C_t(0)]\gamma_s(h_s) + [k_3 + k_1 C_s(0)]\gamma_t(h_t) - k_1 \gamma_s(h_s)\gamma_t(h_t) \quad (3.9)$$

donde C_s y C_t son las funciones de covarianza espacial y temporal, respectivamente; γ_s y γ_t son sus correspondientes funciones de semivariograma. $C_{st}(0)$ es la varianza de γ_{st} , $C_s(0)$ es la varianza de γ_s y $C_t(0)$ es la varianza de γ_t . Por definición $\gamma_{st}(0,0) = \gamma_s(0) = \gamma_t(0) = 0$.

La siguiente condición está implícita en la transformación covarianza-semivariograma:

$$k_1 C_s(0)C_t(0) + k_2 C_s(0) + k_3 C_t(0) = C_{st}(0,0) \quad (3.10)$$

de la ecuación 3.9, se obtiene que

$$\gamma_{st}(h_s, 0) = [k_2 + k_1 C_t(0)]\gamma_s(h_s) \quad (3.11)$$

$$\gamma_{st}(0, h_t) = [k_3 + k_1 C_s(0)]\gamma_t(h_t) \quad (3.12)$$

se supone que

$$k_2 + k_1 C_t(0) = 1 ; \quad k_3 + k_1 C_s(0) = 1 \quad (3.13)$$

para estimar y modelar $\gamma_s(h_s)$ y $\gamma_t(h_t)$ mediante $\gamma_{st}(h_s, 0)$ y $\gamma_{st}(0, h_t)$, respectivamente.

De la ecuación 3.8, es claro que $k_1 > 0$, $k_2 \geq 0$ y $k_3 \geq 0$ son condiciones suficientes para obtener una covarianza positiva definida. De las ecuaciones 3.9 y 3.13, se obtiene:

$$\begin{aligned} k_1 &= [C_s(0) + C_t(0) - C_{st}(0,0)] / C_s(0)C_t(0) \\ k_2 &= [C_{st}(0,0) - C_t(0)] / C_s(0) \\ k_3 &= [C_{st}(0,0) - C_s(0)] / C_t(0) \end{aligned} \quad (3.14)$$

Para modelar por separado los semivariogramas espacial y temporal, es necesario asegurarse que las varianzas sean elegidas de tal manera que las ecuaciones 3.14 sean positivas (De Cesare *et al.* 2001).

Para un campo aleatorio espacio-temporal estacionario de segundo orden (De Iaco *et al.* 2001):

$$C_{st}(h_s, h_t) = C_{st}(0,0) - \gamma_{st}(h_s, h_t) \quad (3.15)$$

Estos modelos no solo permiten estimar valores en ubicaciones en las que no se tienen datos sino que también permiten la predicción en tiempos futuros, por lo que son útiles para analizar la contaminación del aire, datos meteorológicos, o datos de agua subterránea (De Iaco *et al.*, 2003).

No se presentan las ecuaciones para el krigeado espacial ni para el espacio-temporal porque estos métodos de estimación no se usan en la optimización.

Tendencia

El valor esperado de una función aleatoria puede ser constante o depender de las coordenadas de la posición. Para que una función aleatoria sea estacionaria de segundo orden, o satisfaga la hipótesis intrínseca, no debe existir tendencia. Cuando el valor esperado de los datos depende de la posición espacial, del tiempo o de ambos, la función aleatoria no satisface las hipótesis de estacionariedad de segundo orden ni la intrínseca. En ese caso, es posible trabajar con una variable modificada llamada el residuo. A continuación se describe cómo proceder en ese caso.

Una función aleatoria espacio-temporal $Z(\mathbf{x}, t)$ puede escribirse como:

$$Z(\mathbf{x}, t) = m(\mathbf{x}, t) + R(\mathbf{x}, t) \quad (3.16)$$

donde $m(\mathbf{x}, t)$ es una función determinista (conocida como tendencia) de coordenadas espaciales \mathbf{x} y temporal t , y $R(\mathbf{x}, t)$ (conocida como el residuo) es una función aleatoria espacio-temporal

estacionaria de media cero que modela la fluctuación espacio-temporal alrededor de $m(\mathbf{x}, t)$ (Kyriakidis y Journel , 1999).

Para representar la función $m(\mathbf{x}, t)$, una superficie de tendencia (típicamente funciones polinomiales para las coordenadas espaciales y funciones periódicas para el tiempo, aunque otras funciones pueden usarse) puede obtenerse mediante un ajuste por mínimos cuadrados a los datos. De esta manera, se obtienen datos estacionarios de media cero (residuos) y éstos son usados para calcular el semivariograma, esta técnica es llamada kriging residual.

3.3 Filtro de Kalman

El filtro de Kalman es un conjunto de ecuaciones matemáticas que proveen una estimación lineal insesgada de varianza mínima para el estado de un sistema utilizando datos con ruido (Jazwinski 1970). El filtrado se refiere a la estimación de un vector estado en un tiempo actual, basándose en mediciones pasadas. Los filtros recursivos son aquéllos que no necesitan almacenar mediciones pasadas con el propósito de calcular estimaciones presentes. El filtro de Kalman provee una solución recursiva eficiente del método de mínimos cuadrados.

En el filtrado de Kalman se definen tres modelos, al conjunto de los cuales se le suele llamar modelo del proceso, y dos fases que constituyen el filtrado de Kalman propiamente dicho:

Modelo del sistema. Describe la evolución en el tiempo de la cantidad que se desea estimar, expresada mediante un vector de estado \mathbf{C}_n . La transición entre estados $\mathbf{C}_n \rightarrow \mathbf{C}_{n+1}$ se caracteriza por la matriz de transición $\boldsymbol{\varphi}_n$ y la adición de un ruido blanco gaussiano \mathbf{w}_n con matriz de covarianza \mathbf{Q}_n .

$$\mathbf{C}_{n+1} = \boldsymbol{\varphi}_n \mathbf{C}_n + \mathbf{w}_n, \quad \mathbf{w}_n \sim N(0, \mathbf{Q}_n) \quad (3.35)$$

Modelo del sensor o de medición. Relaciona el vector de medida \mathbf{z}_n con el estado del sistema \mathbf{C}_n a través de la matriz de medición \mathbf{H}_n y la adición de un ruido blanco gaussiano \mathbf{v}_n con matriz de covarianza \mathbf{R}_n .

$$\mathbf{z}_n = \mathbf{H}_n \mathbf{C}_n + \mathbf{v}_n, \quad \mathbf{v}_n \sim N(0, \mathbf{R}_n) \quad (3.36)$$

Modelo a priori. Describe el conocimiento previo sobre el vector de estado en el instante inicial \mathbf{C}^0 en cuanto a valor esperado y matriz de covarianza \mathbf{P}^0 . Se suponen no correlacionados los ruidos correspondientes al proceso y al sensor.

$$E[\mathbf{C}^0] = \hat{\mathbf{C}}^0, \quad \mathbf{P}^0 = E\left\{(\mathbf{C}^0 - \hat{\mathbf{C}}^0)(\mathbf{C}^0 - \hat{\mathbf{C}}^0)^T\right\},$$

$$E[\mathbf{w}_k \mathbf{v}_j^T] = 0 \text{ para todo } j, k \quad (3.37)$$

Fase de propagación. En esta fase se predice cuál va a ser el nuevo valor de la cantidad que se desea estimar utilizando el modelo del sistema. Para ello, la estimación del estado anterior $\hat{\mathbf{C}}_n^n$ y su matriz de covarianza \mathbf{P}_n^n son extrapoladas para formar el vector de estado predicho $\hat{\mathbf{C}}_{n+1}^n$ y su matriz de covarianza \mathbf{P}_{n+1}^n . Donde

$$\hat{\mathbf{C}}_n^k = E\{\mathbf{C}_n | z_1, \dots, z_k\} \quad \text{y} \quad \mathbf{P}_n^k = E\left\{(\mathbf{C}_n - \hat{\mathbf{C}}_n^k)(\mathbf{C}_n - \hat{\mathbf{C}}_n^k)^T\right\},$$

$$\hat{\mathbf{C}}_{n+1}^n = \boldsymbol{\Phi}_n \hat{\mathbf{C}}_n^n \quad (3.38)$$

$$\mathbf{P}_{n+1}^n = \boldsymbol{\Phi}_n \mathbf{P}_n^n \boldsymbol{\Phi}_n^T + \mathbf{Q}_n \quad (3.39)$$

Fase de actualización. En esta fase se calcula el nuevo vector de estado $\hat{\mathbf{C}}_{n+1}^{n+1}$ y su matriz de covarianza \mathbf{P}_{n+1}^{n+1} . Para ello se utiliza la covarianza predicha para calcular la ganancia de Kalman \mathbf{K}_{n+1} , escalando ésta por el valor del residuo de medición $\mathbf{z}_{n+1} - \mathbf{H}_{n+1} \hat{\mathbf{C}}_{n+1}^n$, o sea por la estimación del error cometido en la predicción, y sumándose al vector de estado predicho $\hat{\mathbf{C}}_{n+1}^n$ se calcula el nuevo vector de estado $\hat{\mathbf{C}}_{n+1}^{n+1}$.

$$\hat{\mathbf{C}}_{n+1}^{n+1} = \hat{\mathbf{C}}_{n+1}^n + \mathbf{K}_{n+1} (\mathbf{z}_{n+1} - \mathbf{H}_{n+1} \hat{\mathbf{C}}_{n+1}^n) \quad (3.40)$$

$$\mathbf{P}_{n+1}^{n+1} = (\mathbf{I} - \mathbf{K}_{n+1} \mathbf{H}_{n+1}) \mathbf{P}_{n+1}^n \quad (3.41)$$

$$\mathbf{K}_{n+1} = \mathbf{P}_{n+1}^n \mathbf{H}_{n+1}^T (\mathbf{H}_{n+1} \mathbf{P}_{n+1}^n \mathbf{H}_{n+1}^T + \mathbf{R}_{n+1})^{-1} \quad (3.42)$$

3.3.1 Filtro de Kalman estático

En su forma general, el filtro se basa en dos ecuaciones: una ecuación dinámica y una de medición. En este trabajo utilizamos lo que llamamos el filtro de Kalman estático que usa únicamente la ecuación de medición, incorporando el tiempo mediante el uso de un vector de estado espacio-temporal (Herrera, 1998; Herrera y Pinder, 2005); esto se debe a que no se utiliza un modelo para describir la dinámica del parámetro que se estima. A continuación se describe el filtro de Kalman estático de la forma en que se emplea en esta tesis.

La ecuación lineal de medición del filtro de Kalman discreto que relaciona al vector estado de la variable \mathbf{h} en las posiciones y tiempos de interés, con datos medidos en \mathbf{z} es:

$$\mathbf{z}_j = \mathbf{H}_j \mathbf{h} + \mathbf{v}_j \quad (3.43)$$

donde $\{\mathbf{z}_j, j=1,2,\dots\}$ es una sucesión de mediciones de la carga hidráulica. La j -ésima matriz de medición \mathbf{H}_j , es una matriz $1 \times N$ que no es cero únicamente en la posición correspondiente a la entrada de \mathbf{h} en donde la j -ésima medición es tomada, y N es la dimensión del vector \mathbf{h} . $\mathbf{h} = \{\mathbf{h}_{ip}\}$ es el vector espacio-temporal con los valores estimados de la carga hidráulica en las posiciones y tiempos de interés (\mathbf{h}_{ip} es el valor de la variable en la posición x_i y tiempo t_p). El vector $\{\mathbf{v}_j, j=1,2,\dots\}$ representa los errores en la medición, se trata de una sucesión blanca gaussiana, con media cero y covarianza \mathbf{r}_j . La sucesión del error de medición $\{\mathbf{v}_j\}$ y el vector \mathbf{h} son independientes.

La matriz de covarianza del error en la estimación es:

$$\mathbf{P}^n = E\left\{\left(\mathbf{h} - \hat{\mathbf{h}}^n\right)\left(\mathbf{h} - \hat{\mathbf{h}}^n\right)^T\right\} \quad (3.44)$$

donde $\hat{\mathbf{h}}^n = E\left\{\left(\mathbf{h} / \mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n\right)^T\right\}$ es el valor esperado de \mathbf{h} , dadas las mediciones $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n$. En esta notación, el superíndice identifica el número n de vectores de medición usados para obtener una estimación.

Para implementar el filtro, se requiere proponer una estimación a priori de \mathbf{h} (llamada $\hat{\mathbf{h}}^0$) y de la matriz de covarianza del error en la estimación (\mathbf{P}^0). Nótese que (\mathbf{P}^0) es una matriz de covarianza espacio-temporal. Dadas estas estimaciones a priori, la estimación lineal de varianza mínima para \mathbf{h} puede obtenerse secuencialmente mediante las siguientes fórmulas:

$$\hat{\mathbf{h}}^{n+1} = \hat{\mathbf{h}}^n + K_{n+1}\left(\mathbf{z}_{n+1} - \mathbf{H}_{n+1}\hat{\mathbf{h}}^n\right) \quad (3.45)$$

$$\mathbf{P}^{n+1} = \mathbf{P}^n - K_{n+1}\mathbf{H}_{n+1}\mathbf{P}^n \quad (3.46)$$

$$K_{n+1} = \mathbf{P}^n\mathbf{H}_{n+1}^T\left(\mathbf{H}_{n+1}\mathbf{P}^n\mathbf{H}_{n+1}^T + \mathbf{r}_{n+1}\right)^{-1} \quad (3.47)$$

Existen muchas formas para estimar $\hat{\mathbf{h}}^0$, se presentan algunas en los diversos diseños presentados en esta tesis. La matriz de covarianza del error en la estimación a priori (\mathbf{P}^0) se obtiene de un análisis geoestadístico mediante el ajuste de un modelo de semivariograma espacio-temporal al semivariograma muestral. Una vez seleccionado el modelo de semivariograma, los elementos de la matriz de covarianza espacio-temporal son calculados con la ecuación 3.15.

4. METODOLOGÍA GEOESTADÍSTICA PARA EL DISEÑO ESPACIO-TEMPORAL ÓPTIMO DE REDES DE MONITOREO DE LA CARGA HIDRÁULICA

En este capítulo se presenta una nueva metodología para el diseño espacio-temporal óptimo de redes de monitoreo de la carga hidráulica, con el objetivo de seleccionar de entre un conjunto de posiciones y tiempos de monitoreo propuestos, aquéllos que minimizan la redundancia espacio-temporal.

La elección de las posiciones de monitoreo y tiempos que definirán la red de monitoreo y su calendario se lleva a cabo usando un filtro de Kalman estático (Herrera, 1998) y un método de optimización. El filtro de Kalman estático requiere una matriz de covarianza espacio-temporal a priori calculada a través de un modelo producto-suma (De Cesare *et al.*, 2001). El método de optimización es de inclusiones sucesivas y selecciona, una a la vez, las posiciones de monitoreo espacio-temporal que minimizan una función objetivo. La función objetivo depende de la varianza del error en la estimación y pudiera cambiar de acuerdo al problema particular de implementación. El procedimiento de optimización busca reducir el valor de esta función sobre el conjunto espacio-temporal de posiciones en las que se requiere estimar.

En el presente trabajo se considera, a diferencia de Herrera, el efecto de los datos en la varianza del error en la estimación únicamente en tiempos presentes y futuros, pero además, la matriz de covarianza es actualizada únicamente en el presente y el futuro. Esto es, cuando una posición de monitoreo es seleccionada para formar parte de la red de monitoreo, la matriz de covarianza es actualizada por el filtro de Kalman estático únicamente en tiempos presentes y futuros. La aplicación de este procedimiento en el proceso de optimización es a lo que llamamos la optimización en tiempo real de la red de monitoreo. En el presente trabajo se propone también un nuevo criterio para determinar el número total de posiciones para el programa de una red de monitoreo espacio-temporal.

La metodología propuesta se divide en cuatro pasos que son: definición del objetivo del monitoreo, análisis geoestadístico espacio-temporal, aplicación del método de optimización, y determinación del número total de posiciones de monitoreo; éstos se describen a continuación.

4.1 Definición del objetivo del monitoreo

Un diseño espacio-temporal de una red de monitoreo de la carga hidráulica considera estimar la variable sobre áreas de interés (como pueden ser áreas de extracción intensiva, áreas de recarga o descarga, o sobre toda el área que define un acuífero) para un cierto periodo de tiempo. Cada área se representa por un conjunto discreto de posiciones espaciales, y el periodo de tiempo por un conjunto discreto de tiempos. La unión de ambos conjuntos define un conjunto espacio-temporal de posiciones en las que se necesita estimar. Este conjunto puede ser más denso en áreas y tiempos con la más alta prioridad, para asignarles un mayor peso en el proceso de optimización.

4.2 Análisis geoestadístico espacio-temporal

Es necesario llevar a cabo un análisis geoestadístico para determinar la estructura de correlación espacio-temporal de los datos (ver capítulo tres), el resultado de este análisis es un modelo de

semivariograma espacio-temporal producto-suma (ecuación 3.9) con el que se calcula la matriz de covarianza a priori (con la ecuación 3.15) que necesita el filtro de Kalman estático.

4.3 Optimización de la red de monitoreo

Considere el conjunto de todas las posibles posiciones de monitoreo $\{x_1^M, x_2^M, \dots, x_{Npm}^M\}$ y tiempos de monitoreo $\{t_1^M, t_2^M, \dots, t_{Ntm}^M\}$. Todos los posibles pares del conjunto son $M = \{(x_i^M, t_p^M), i = 1, \dots, Npm; p = 1, \dots, Ntm\}$. Para las diferentes aplicaciones presentadas en esta tesis, M es el conjunto de posibles posiciones de monitoreo espacio-temporal, aunque la metodología puede ser aplicada aún si los tiempos de monitoreo no son los mismos para cada pozo. De este conjunto, se seleccionarán las posiciones que minimizan una función de la varianza del error en la estimación sobre las posiciones de estimación espacio-temporal $E = \{(x_j^E, t_q^E), j = 1, \dots, Npe; q = 1, \dots, Nte\}$.

Hasta ahora hemos probado dos opciones para la función objetivo de la optimización. La primera fue propuesta originalmente por Herrera (1998) y es la varianza del error en la estimación, sumada sobre todos las posiciones de estimación espacio-temporal. Esta función es llamada la varianza total del error en la estimación y se calcula usando la siguiente fórmula:

$$\sigma_T^2 = \sum_{j=1}^{Npe} \sum_{q=1}^{Nte} \sigma_{j,q}^2 \quad (4.1)$$

donde $\sigma_{j,q}^2$ es la varianza del error en la estimación $e(x_j^E, t_q^E) = h(x_j^E, t_q^E) - \hat{h}(x_j^E, t_q^E)$, en la j -ésima posición de estimación y el q -ésimo tiempo de estimación. Nótese que al usar esta función objetivo, al agregar una posición de monitoreo, ésta contribuye en la reducción de la varianza del error en la estimación en todos los tiempos considerados. Esto es, cada posición de monitoreo contribuye a reducir la varianza del error en la estimación en el pasado, el presente y el futuro. Las fórmulas usadas para minimizar la varianza total son presentadas en Herrera (1998) y Herrera y Pinder (2005).

La segunda opción es considerar el efecto de las posiciones de monitoreo en la varianza del error en la estimación únicamente en tiempos presentes y futuros. Esta función objetivo es usada para optimizar una de las redes de monitoreo de la carga hidráulica presentadas en Herrera y Pinder (2005).

En el presente trabajo se considera también el efecto de los datos en la varianza del error en la estimación únicamente en tiempos presentes y futuros, pero además, la matriz de covarianza es actualizada únicamente en el presente y el futuro. Esto es, cuando se selecciona una posición para formar parte de la red de monitoreo, la matriz de covarianza se actualiza por el filtro de Kalman estático únicamente en tiempos presentes y futuros. La aplicación de este procedimiento en el proceso de optimización es a lo que llamamos la optimización en tiempo real de la red de monitoreo.

El problema es optimizado secuencialmente. El método de optimización puede ser considerado como un método de inclusiones sucesivas (Samper y Carrera 1990). El problema de optimización puede escribirse como:

Para $n = 1, 2, \dots, Npm \times Ntm$, elige una posición $(x_i^M, t_p^M) \in M$ que minimice la función:

$$\sigma_{TR}^2(n-1) = \sum_{j=1}^{Npe} \sum_{p \leq q}^{Nte} \sigma_{j,q}^2(n-1) \quad (4.2)$$

donde $\sigma_{TR}^2(n-1)$ es la varianza total en tiempo real y $\sigma_{j,q}^2(n-1)$ es la varianza del error en la estimación, cuando datos de $n-1$ posiciones de monitoreo elegidas previamente de manera óptima son usadas en el proceso de estimación. Cuando $n > 1$, las varianzas $\sigma_{j,q}^2(n-1)$ son obtenidas del filtro de Kalman estático y son los elementos que corresponden a las posiciones y tiempos de estimación en la diagonal principal de la matriz de covarianza \mathbf{P}^{n-1} . Por otro lado, $\sigma_{j,q}^2(0)$ es la varianza total obtenida de la matriz de covarianza derivada del análisis geoestadístico espacio-temporal.

Al aplicar una optimización secuencial es necesario determinar dónde parar el proceso, esto es, determinar el número total de posiciones a ser incluidas en la red de monitoreo espacio-temporal óptima.

4.4 Criterio para determinar el número total de posiciones de monitoreo

Existen muchos criterios para determinar el número total de posiciones en el programa de la red de monitoreo. En esta tesis hemos implementado un criterio basado en los resultados de la varianza total del error en la estimación, y que considera el orden de prioridad asignado por el método.

Para definir el número de posiciones de la red de monitoreo espacio-temporal, se propone usar el valor $\sqrt{\frac{\text{Varianza total}}{N}}$, donde $N = Npe * Nte =$ Número de posiciones de estimación espacio-temporal.

Puede demostrarse que $\sqrt{\frac{\sum_{j=1}^{Npe} \sum_{q=1}^{Nte} \sigma_{j,q}^2}{N}} \geq \frac{1}{N} \sum_{j=1}^{Npe} \sum_{q=1}^{Nte} \sigma_{j,q}$ para $N \geq 1$ y $\sum_{j=1}^{Npe} \sum_{q=1}^{Nte} \sigma_{j,q} \geq 0$. De esta

forma, el valor estadístico propuesto es una frontera superior para el error estándar promedio sobre la malla de estimación espacio-temporal, el cual puede ayudarnos para considerar un escenario conservador de diseño.

El criterio consiste en seleccionar las posiciones de monitoreo espacio-temporal donde se alcanza un 99% del valor de la máxima reducción posible (MRP):

$$MRP = \text{Max} \sqrt{\frac{\sum_{j=1}^{Npe} \sum_{q=1}^{Nte} \sigma_{j,q}^2}{N}} - \text{Mín} \sqrt{\frac{\sum_{j=1}^{Npe} \sum_{q=1}^{Nte} \sigma_{j,q}^2}{N}} \quad (4.3)$$

5. DISEÑO DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON ALTO NIVEL DE INFORMACION ESPACIO-TEMPORAL

En este capítulo se presenta una aplicación de la metodología propuesta en el capítulo cuatro para el diseño óptimo de una red de monitoreo espacio-temporal (seleccionando posiciones y tiempos de monitoreo) que permita estimar adecuadamente las variaciones mensuales de la carga hidráulica en un acuífero. Para lograr este fin, se trabajó con un caso sintético: un modelo numérico de flujo desarrollado para el acuífero Valle de Querétaro (Herrera *et al.*, 2007), lo que da la posibilidad de generar una “rica” base de datos de la carga hidráulica.

5.1 Modelo numérico de flujo

A continuación se presentan las principales características del modelo numérico de flujo con el que se generaron los datos de la carga hidráulica para el diseño de la red de monitoreo; se pone especial énfasis en los elementos que se modificaron con respecto al modelo original desarrollado por Herrera *et al.* (2007) y que contribuyen significativamente en las variaciones espaciales y temporales de la carga hidráulica. Los interesados en una revisión detallada del modelo numérico de flujo en su forma original pueden consultar el trabajo de Herrera *et al.* (2007).

El modelo seleccionado representa un acuífero libre que cubre un área de 379.25 km², las zonas más profundas son cercanas a 600 m. Las hipótesis del modelo consideran que el flujo en el acuífero es horizontal, no se presenta recarga profunda, y la temperatura y densidad del agua son constantes.

Se usó el simulador Princeton Transport Code (PTC), Babu *et al.*, (1997), que resuelve las ecuaciones diferenciales parciales que describen el flujo subterráneo y el transporte de contaminantes (en este caso sólo se empleó la solución para flujo). Para la discretización de las ecuaciones, PTC usa el método del elemento finito (horizontalmente) y diferencias finitas (verticalmente). La Fig. 5.1 muestra el dominio del modelo y la malla numérica triangular, ésta es más refinada en las áreas con una mayor densidad de pozos. La malla consiste en 6,676 nodos y 13,066 elementos. El modelo es de una sola capa (modelo bidimensional). Los valores de la conductividad hidráulica empleados en el modelo oscilan entre 0.001 y 15 m/d, mientras que los valores del rendimiento específico varían desde 0.10 en la parte más baja de la cuenca superficial hasta 0.07 en las partes más elevadas. Se obtuvieron simulaciones mensuales con el modelo numérico de flujo para un periodo de cuatro años.

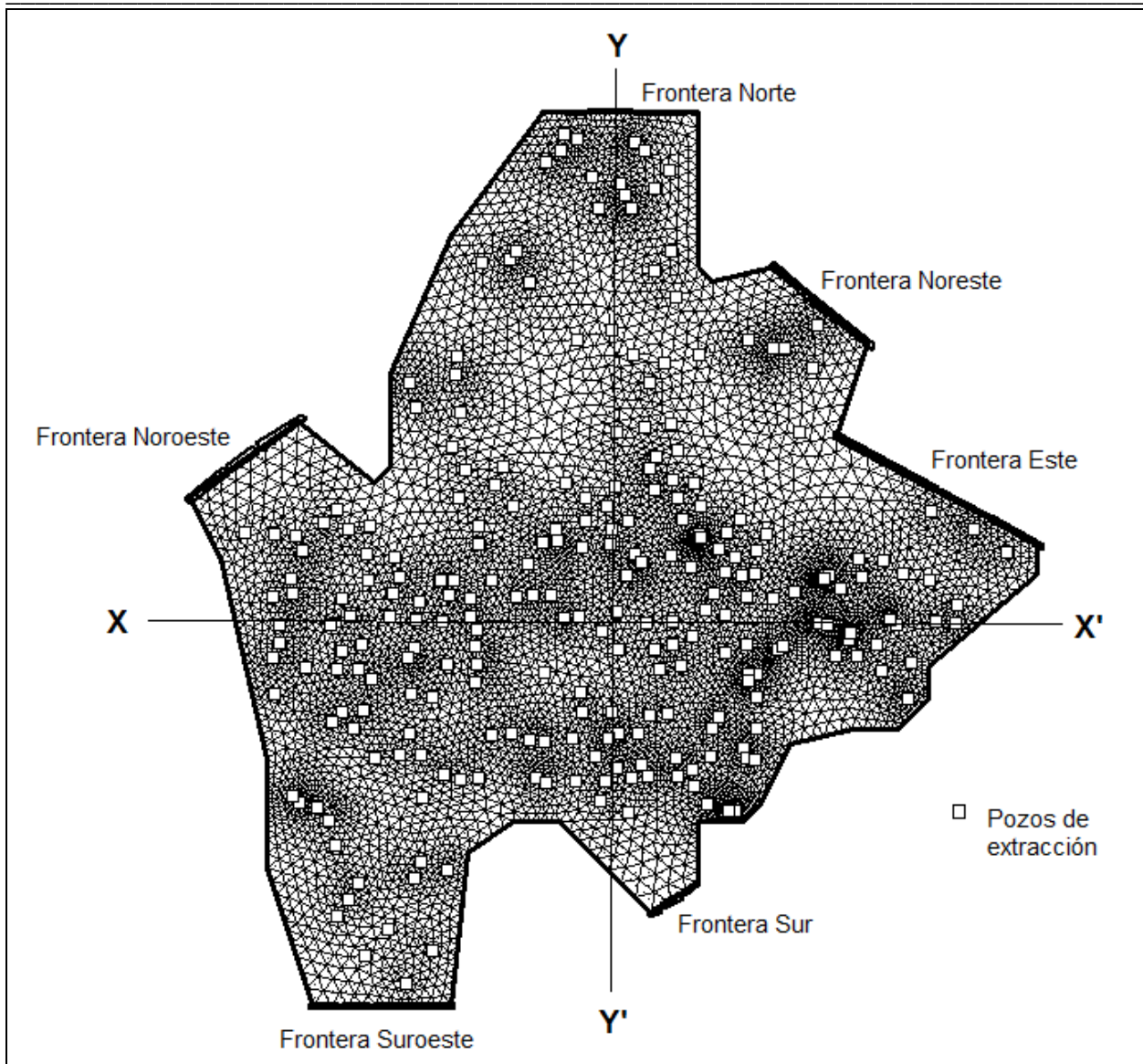


Fig. 5.1 Malla de elemento finito y condiciones de frontera del modelo

Generación de los datos de la carga hidráulica

A continuación se describe el procedimiento para la generación de la base de datos usada en el diseño y evaluación del desempeño de una red de monitoreo espacio-temporal.

Se corrió el modelo en estado transitorio, con una condición inicial basada en datos de la carga hidráulica medidos en agosto del primer año del periodo de simulación. Se usaron condiciones de frontera de carga asignada y sin flujo (Fig. 5.1). Se incluyeron los pozos de extracción con tasas de bombeo mensuales, modificando la entrada original de bombeo anual en cada uno de los pozos, de la manera que se explica a continuación.

Se calculó una tasa de bombeo promedio mensual con la información de un año específico, para el cual contábamos con la información de la Tabla 5.1.

Tabla 5.1 Tasas de bombeo promedio mensuales

	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
Tasa de bombeo promedio (l/s)	31.14	31.08	31.11	30.83	30.69	31.05	30.47	30.47	30.44	30.37	29.82	29.82

Con la Tabla 5.1 se calculó la tasa de bombeo promedio anual (30.61 l/s) y dividiendo este valor por la tasa de bombeo promedio de cada mes, se obtuvieron factores de escalamiento del bombeo mensual (Tabla 5.2). Para cada pozo se obtuvo una tasa de bombeo mensual multiplicando la tasa de bombeo original por su correspondiente factor de escalamiento.

Tabla 5.2 Factores de escalamiento del bombeo mensual

	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
Factor de escalamiento del bombeo	1.017	1.015	1.016	1.007	1.003	1.014	0.996	0.996	0.995	0.992	0.974	0.974

El cálculo de la recarga vertical natural se realizó usando diferentes coeficientes de infiltración multiplicados por la lámina de lluvia correspondiente para cada mes del año. El coeficiente de infiltración va de 2% en la zona urbana a 12% en las zonas topográficamente más elevadas de la cuenca (Fig. 5.2).

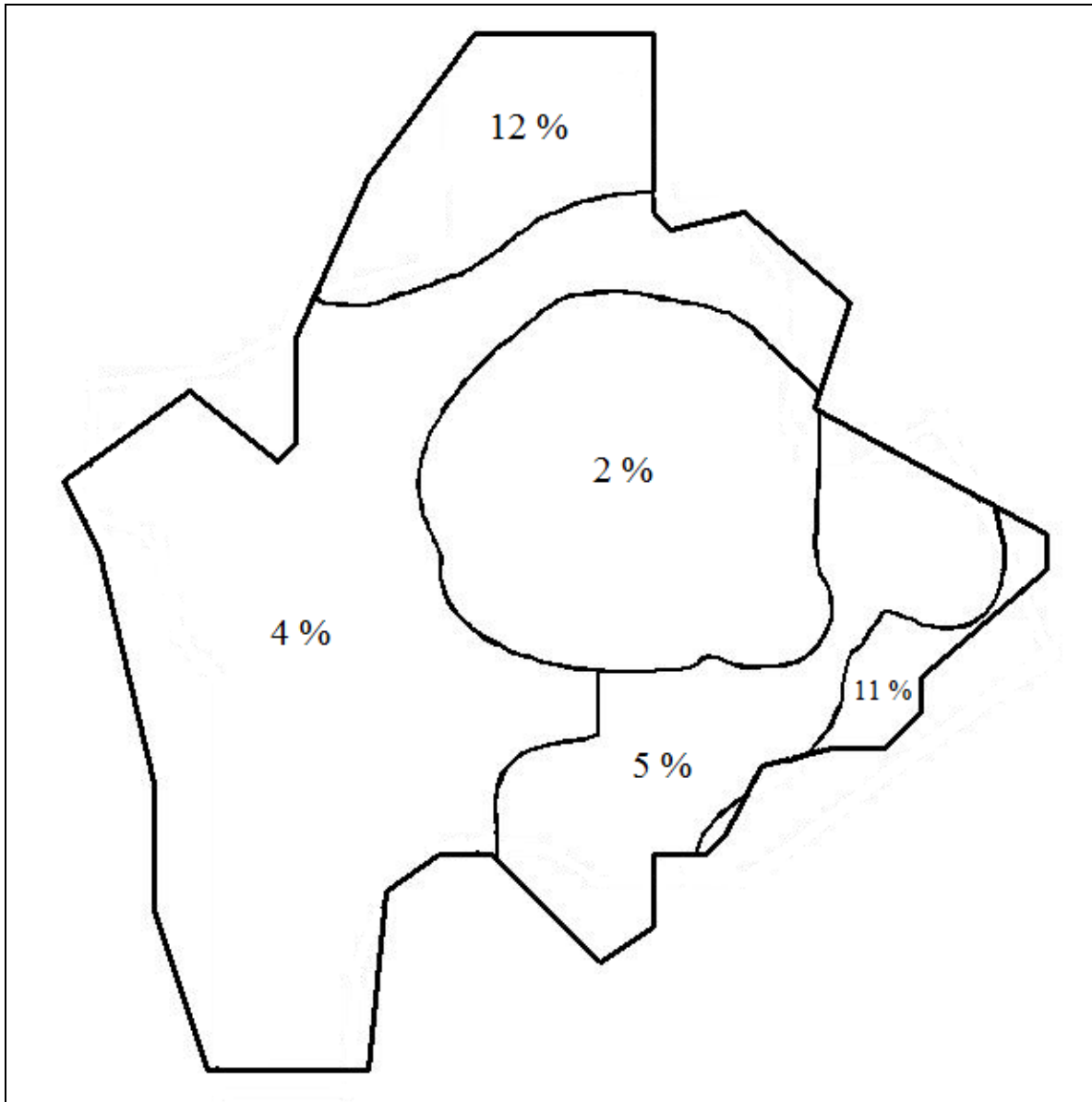


Fig. 5.2 Coeficientes de infiltración usados en la estimación de la recarga vertical natural

La información de la lámina de lluvia mensual histórica (Tabla 5.3) se perturbó antes de introducirse al modelo con el objetivo de no tener un patrón cíclico idéntico en la recarga durante el periodo de simulación, procedimiento que se explica a continuación.

Tabla 5.3 Precipitación mensual promedio histórica (1921-1998)

	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
Precip (mm)	11.3	4.6	7.1	17.9	36.6	102.6	127.1	92.3	89.4	39.0	11.1	9.0

Para cada mes se obtuvieron cuatro realizaciones espaciales de la precipitación (a cada realización se le asignó un año), usando un modelo de semivariograma espacial ajustado a datos

del logaritmo natural de lámina de lluvia de un caso real (Becerra-Soriano, 2009) (Tabla 5.4). El programa utilizado para obtener las realizaciones fue el SGSIM de la librería del GSLIB (Deutsch and Journel, 1998). Este programa está escrito en el lenguaje de programación Fortran y aplica el método de simulación secuencial gaussiana (SGS, por sus siglas en inglés) para obtener las realizaciones. SGS es un algoritmo de simulación estocástica muy conocido que se ha usado para obtener campos aleatorios gaussianos (para una revisión de la teoría, ver Deutsch y Journel, 1998).

Tabla 5.4 Modelo de semivariograma esférico obtenido para una tormenta

Semivariograma	Pepita (ln mm)	Meseta-pepita (ln mm)	Alcance (m)
ln precipitación pluvial	0.20	0.95	20,000

Existe recarga vertical inducida (retornos de riego y fugas en la red de agua potable) que se supuso constante para todos los periodos de modelación (Fig. 5.3).

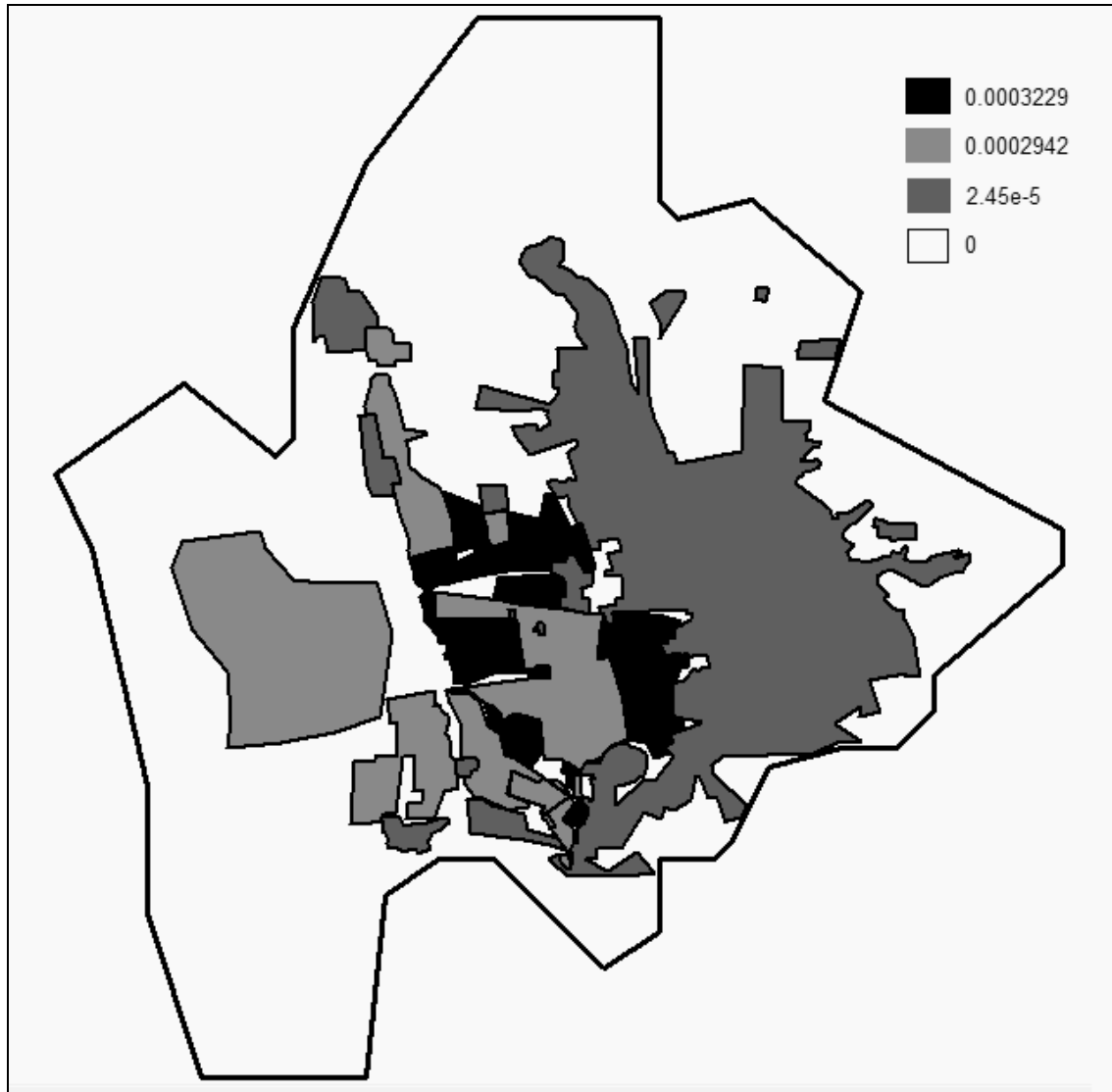


Fig. 5.3 Recarga vertical inducida asignada al modelo (m/d)

Con el modelo numérico de flujo se generaron datos mensuales de la carga hidráulica para un periodo de cuatro años en los 6,676 nodos de la discretización por elemento finito. Algunos pozos fueron seleccionados en diferentes zonas del acuífero (Fig. 5.4) para comparar el comportamiento temporal de la carga hidráulica (Figuras 5.5 - 5.11). En los pozos L-246, 644-A, L-107 y 2082 localizados respectivamente en la parte suroeste, sur, oeste y central del acuífero, donde se tienen los valores más bajos del coeficiente de infiltración y donde además se cuenta con un número importante de aprovechamientos, se presenta una tendencia temporal de descenso de la carga hidráulica; mientras que los pozos 752, 976 y 1792, ubicados respectivamente al este, noreste y norte del acuífero, muestran una recuperación de los niveles de la carga hidráulica a través del tiempo, siendo éste un efecto combinado de los coeficientes de infiltración de esas zonas y el relativamente bajo número de extracciones que allí se localizan. El pozo 752 parece alcanzar una estabilización de su nivel a partir de noviembre del segundo año, mientras que los pozos 976 y 1792 aumentan continuamente su nivel a través del tiempo, probablemente debido a que en esas zonas se presentan recargas importantes hacia el acuífero. En la Fig. 5.12 se

presentan simulaciones de la carga hidráulica para las fechas de inicio y final del periodo de diseño correspondientes al corte A-A' de la Fig. 5.4; se eligió esta zona porque es donde se tienen detectados los abatimientos más pronunciados ocasionados por el bombeo a través del tiempo.

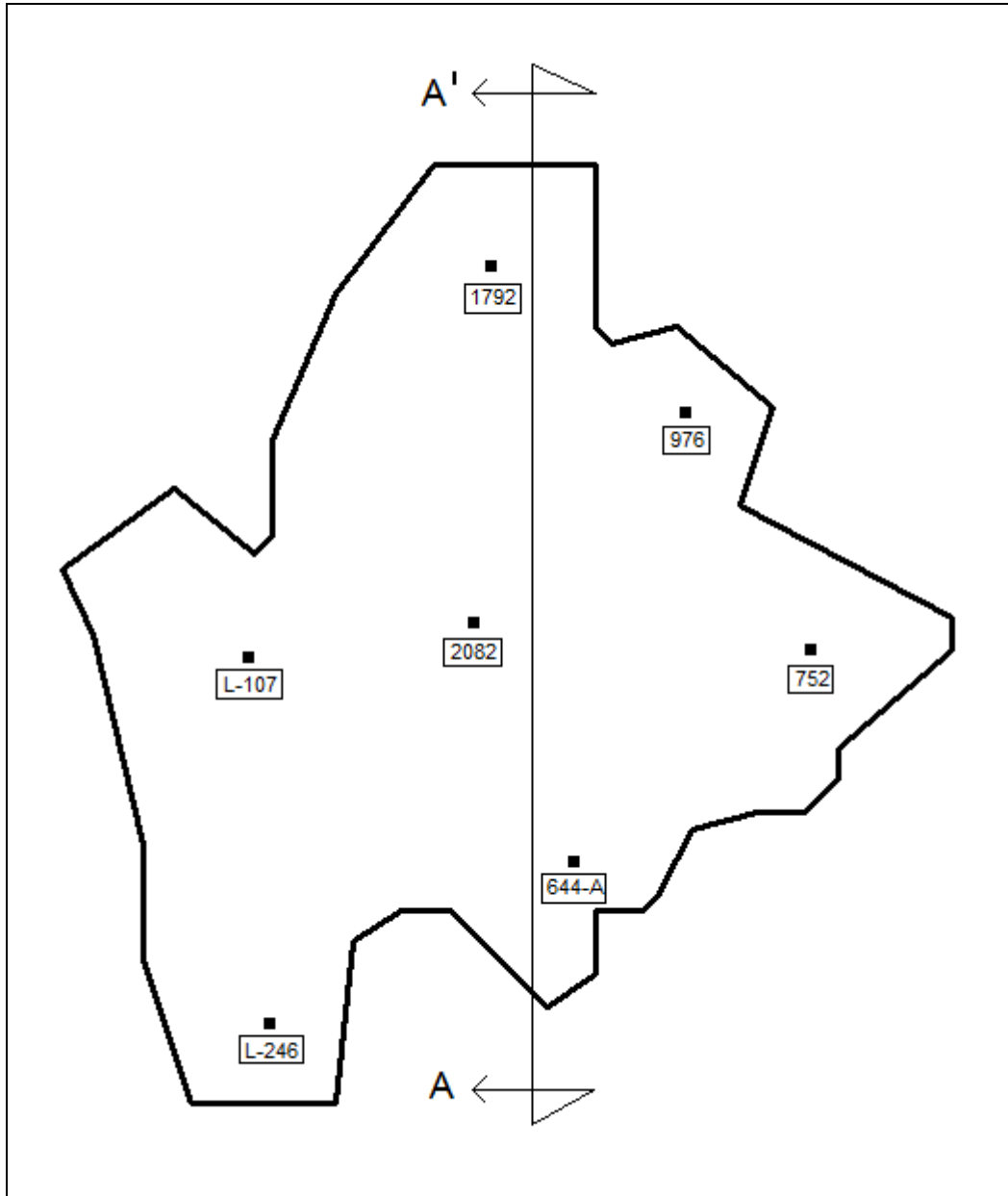


Fig. 5.4 Pozos localizados en diferentes zonas del acuífero

DISEÑO DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON ALTO NIVEL DE INFORMACIÓN ESPACIO-TEMPORAL

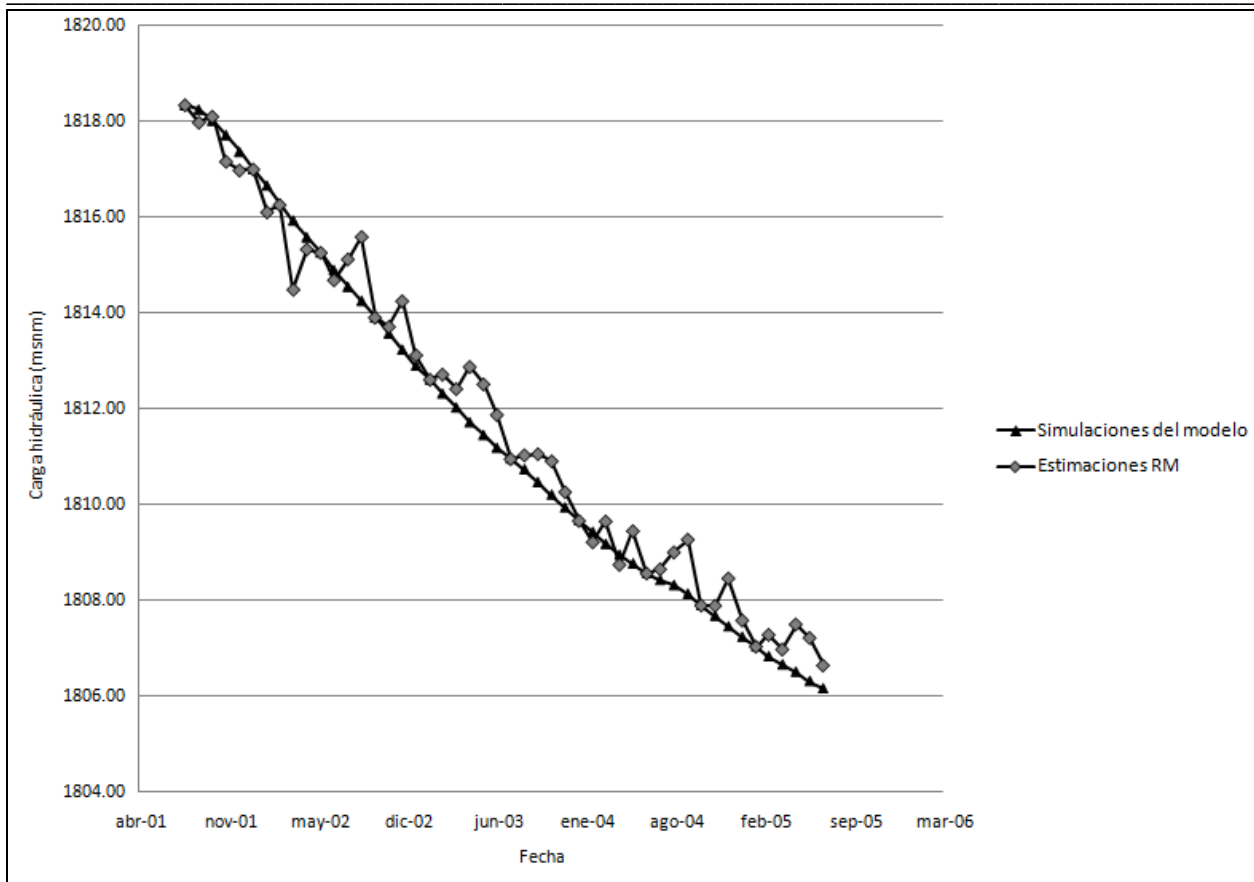


Fig. 5.5 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo L-246

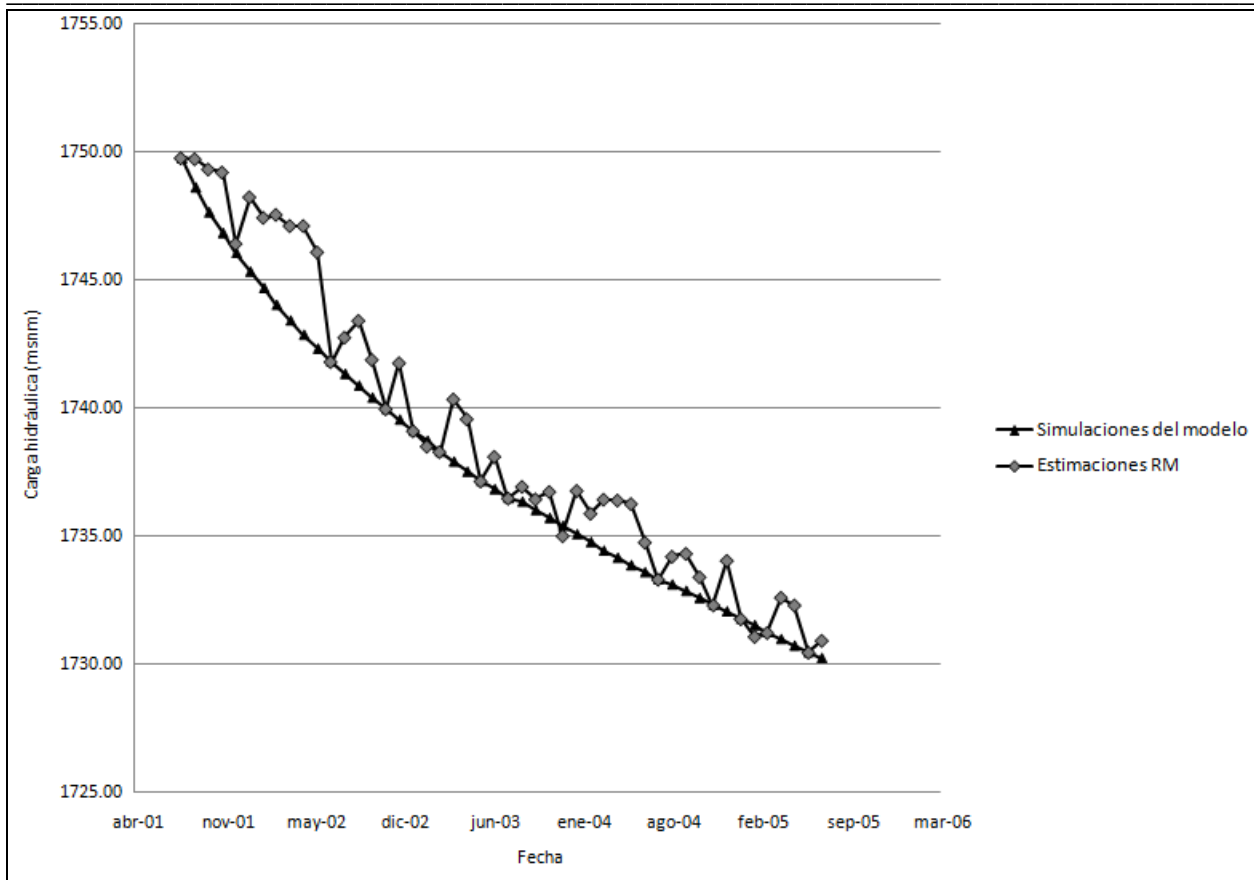


Fig. 5.6 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo 644-A

DISEÑO DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON ALTO NIVEL DE INFORMACIÓN ESPACIO-TEMPORAL

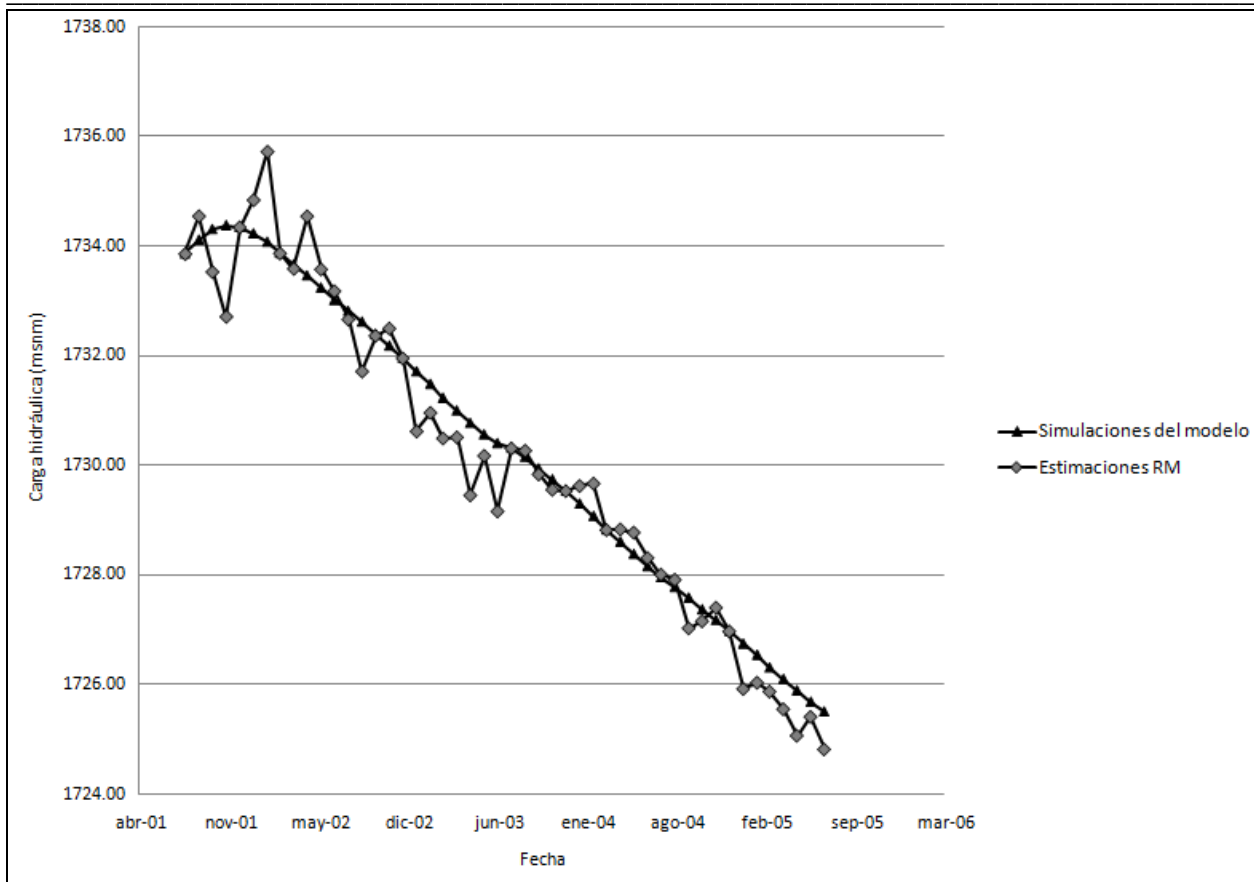


Fig. 5.7 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo L-107

DISEÑO DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON ALTO NIVEL DE INFORMACIÓN ESPACIO-TEMPORAL

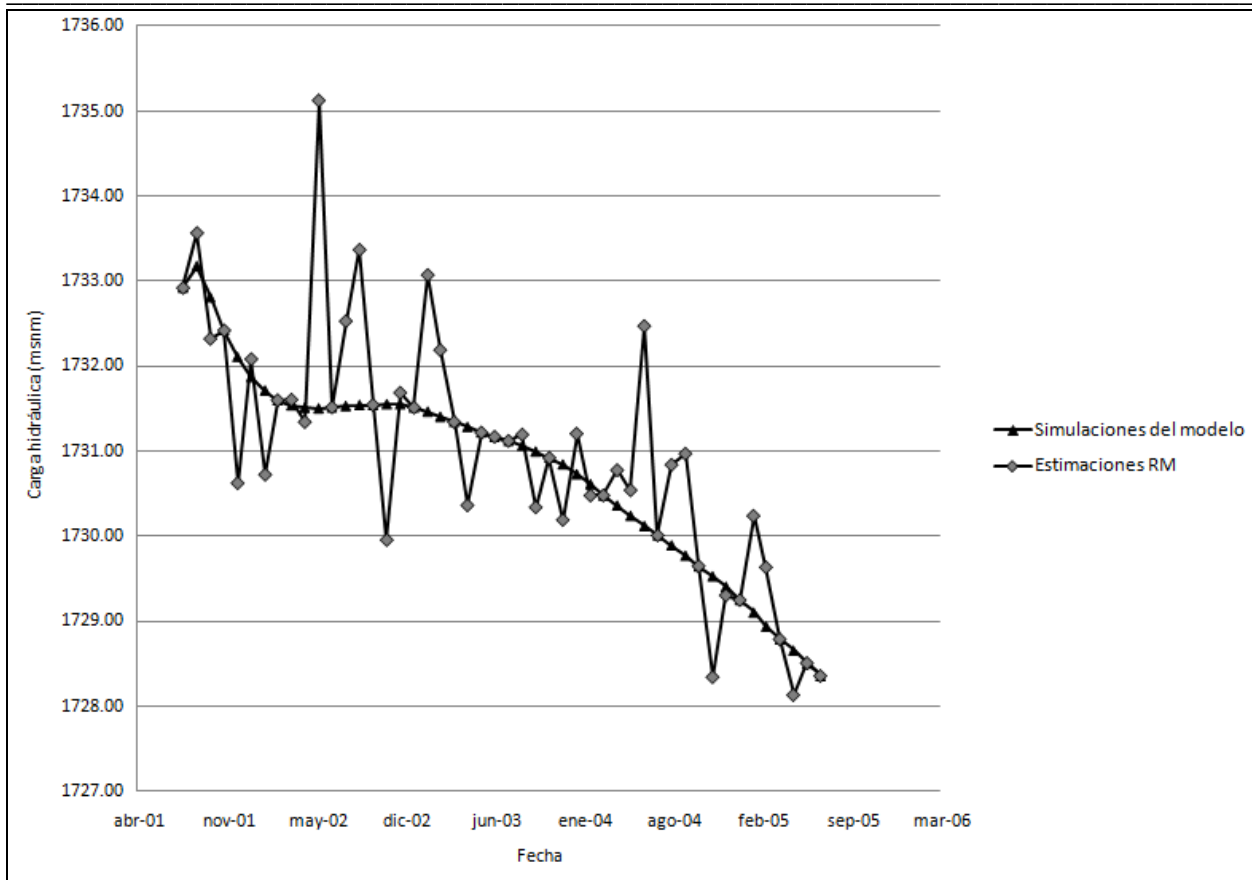


Fig. 5.8 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo 2082

DISEÑO DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON ALTO NIVEL DE INFORMACIÓN ESPACIO-TEMPORAL

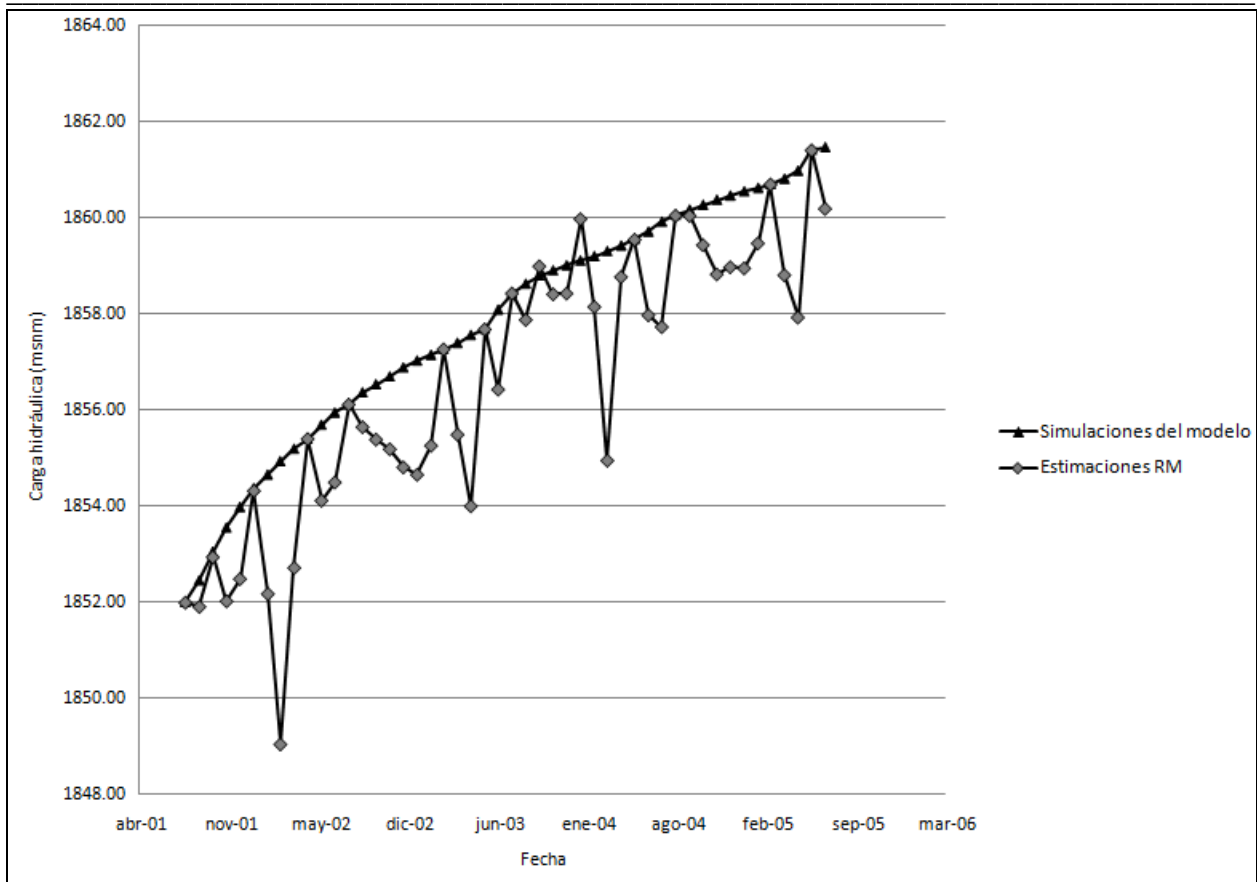


Fig. 5.9 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo 1792

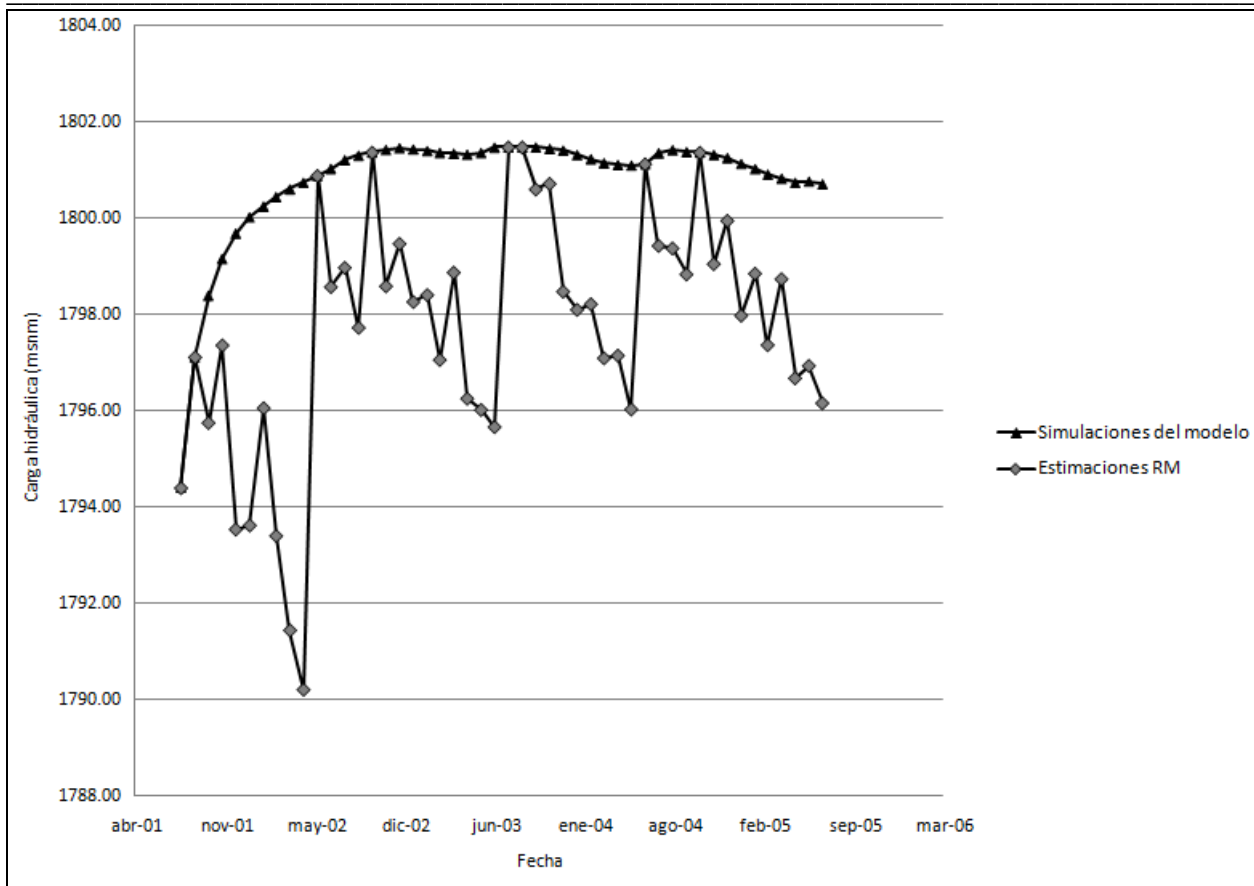


Fig. 5.10 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo 752

DISEÑO DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON ALTO NIVEL DE INFORMACIÓN ESPACIO-TEMPORAL

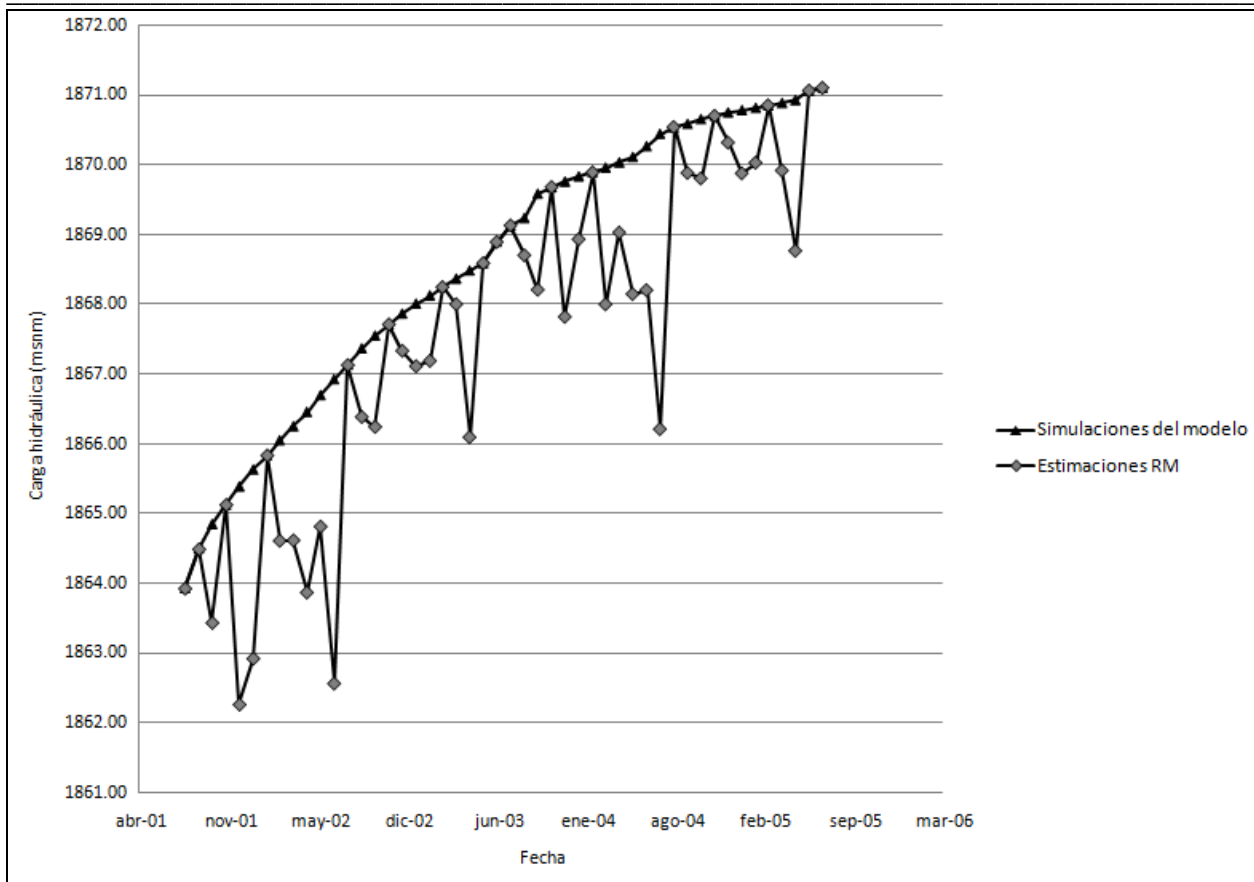


Fig. 5.11 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo 976

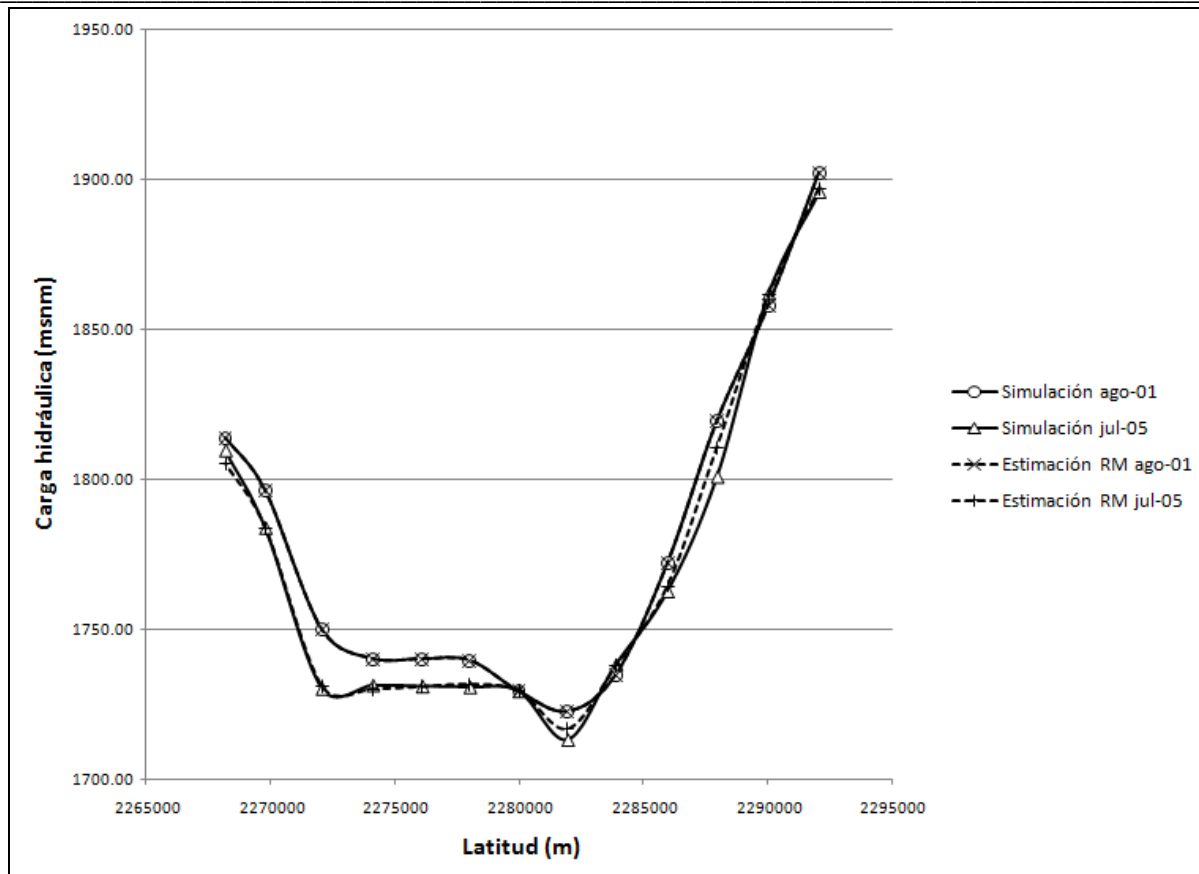


Fig. 5.12 Simulación y estimación de la carga hidráulica para el corte A-A' de la Fig. 5.4

En el diseño de la red de monitoreo se emplearon simulaciones mensuales de la carga hidráulica en las 109 posiciones de la Fig. 5.13 para un periodo de cuatro años (agosto del primer año a julio del quinto año); estas posiciones corresponden a nodos de la discretización mediante elemento finito del modelo numérico de flujo descrito anteriormente. La información de los dos primeros años se empleó para el diseño de la red de monitoreo (agosto del primer año a julio del tercer año) y la información de los dos años siguientes (agosto del tercer año a julio del quinto año) se usó para la evaluación de su desempeño en el futuro.

5.2 Objetivo del monitoreo

El objetivo del diseño es seleccionar las posiciones y tiempos de monitoreo necesarios para estimar adecuadamente las variaciones mensuales de la carga hidráulica en el acuífero en diferentes tiempos. Para este ejemplo, se asignó el mismo peso a la incertidumbre en todo el acuífero y sobre todo el periodo analizado, por lo que 109 nodos de una malla semi-regular con elementos cuadriláteros de aproximadamente 2 km de lado fueron elegidos como posiciones de estimación, estos cubren de manera homogénea el área del acuífero (Fig. 5.13). Un posible tiempo de estimación fue seleccionado para cada uno de los meses dentro del periodo de agosto del primer año a julio del tercer año de la simulación. En este ejemplo, estas posiciones y tiempos se tomaron también como posibles pozos y tiempos de monitoreo, respectivamente. De esta manera, 109 puntos y 24 tiempos fueron seleccionados para la malla de estimación espacio-temporal y para las posibles posiciones de monitoreo espacio-temporal.

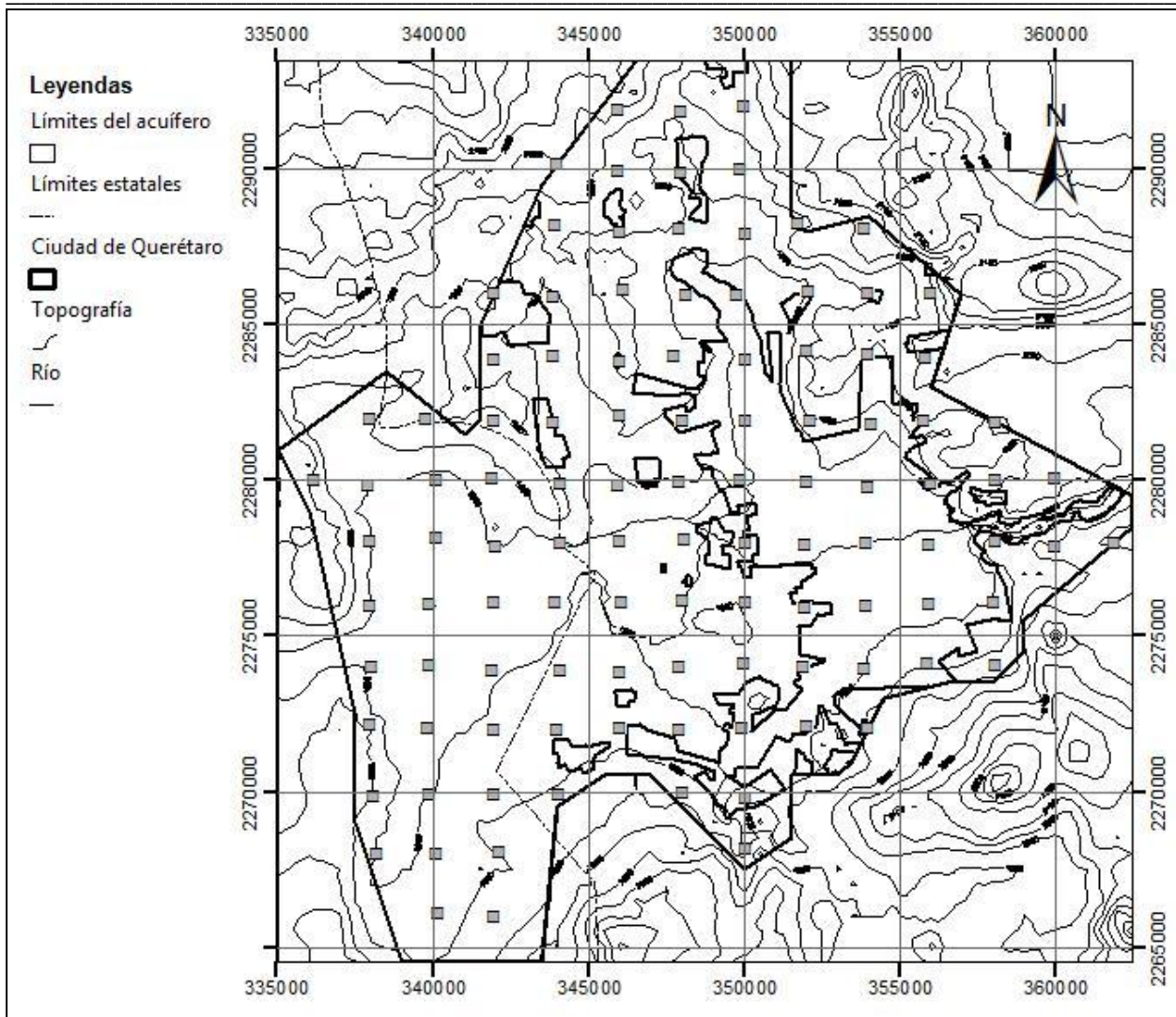


Fig. 5.13 Área de estudio y localización de los puntos de monitoreo/estimación

5.3 Análisis geoestadístico espacio-temporal de la carga hidráulica

El análisis se basó en la construcción del semivariograma espacio-temporal muestral con 2616 datos piezométricos (lo que equivale a información mensual en un periodo de dos años para las 109 posiciones señaladas anteriormente, obtenida del modelo numérico de flujo) (Fig. 5.14). Para llevar a cabo este análisis utilizamos una rutina modificada del GAMV (Deutsch *et al.*, 1997) por De Cesare *et al.* (2002) para la modelación espacio-temporal de semivariogramas. Empleamos 20 intervalos espaciales, con 1388.60 m de separación de intervalo espacial, y 20 intervalos temporales con una separación de intervalo de 0.095 de año.

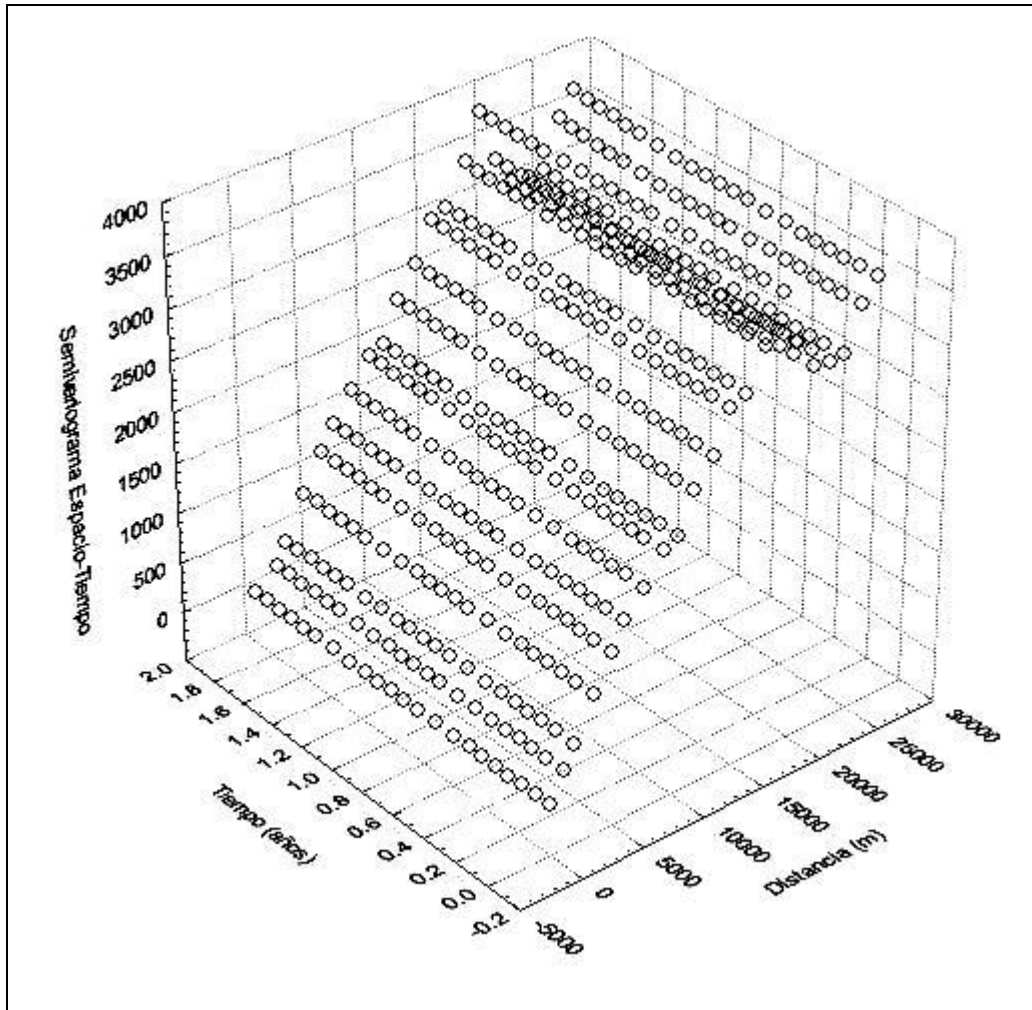


Fig. 5.14 Semivariograma espacio-temporal muestral de la carga hidráulica

Para analizar mejor el comportamiento del semivariograma espacio-temporal muestral, presentamos su proyección sobre el plano semivariograma-intervalos espaciales (Fig. 5.15) y sobre el plano semivariograma-intervalos temporales (Fig. 5.16). La proyección espacial muestra un comportamiento bien definido con ligeras fluctuaciones a partir de los 18,000 metros. La proyección temporal muestra que la componente de varianza temporal tiene poca influencia en la forma del semivariograma espacio-temporal muestral. No se advirtió la presencia de tendencia en los datos.

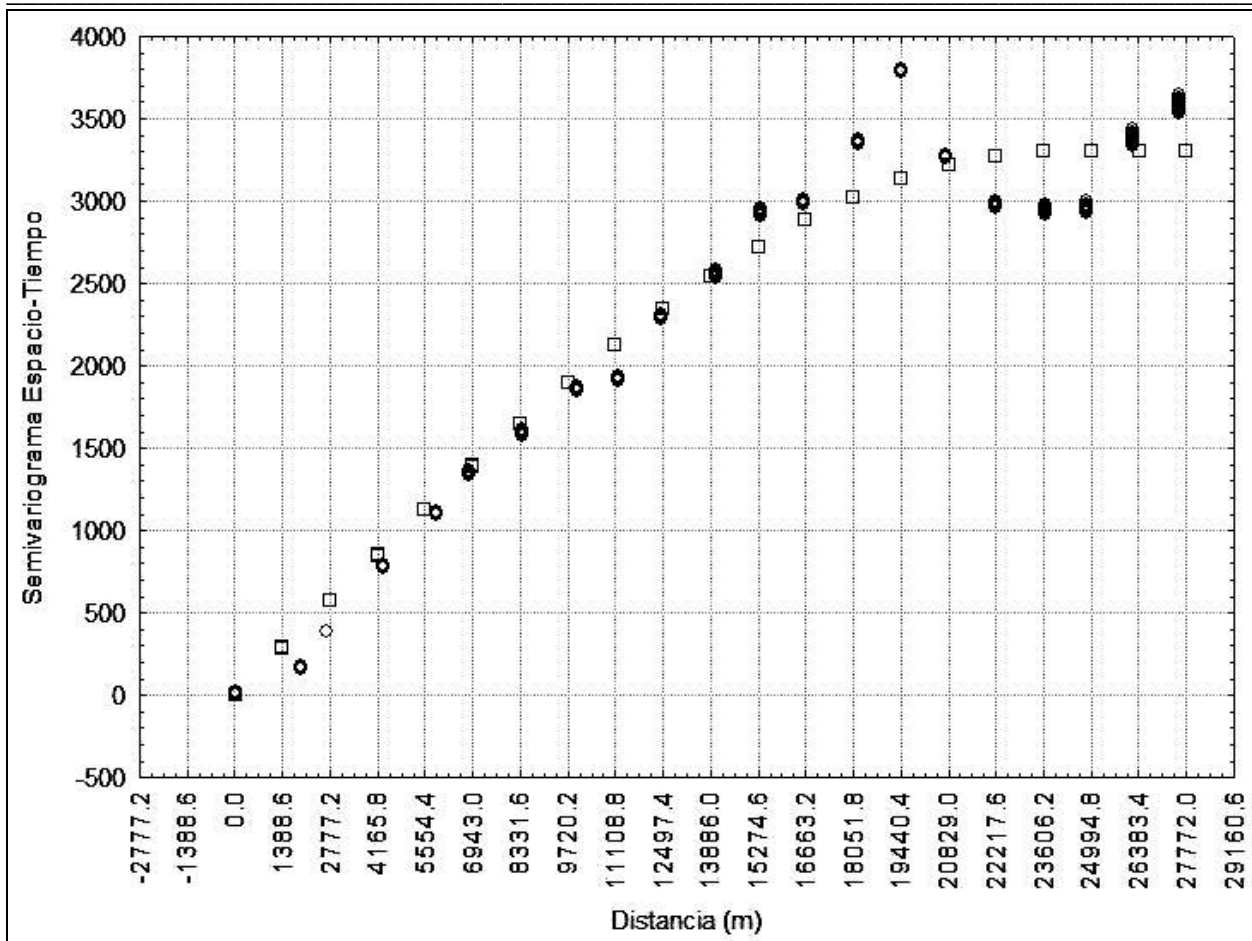


Fig. 5.15 Proyección del semivariograma espacio-temporal muestral de la carga hidráulica sobre el plano semivariograma-intervalos espaciales (círculos) y modelo de semivariograma ajustado (cuadros)

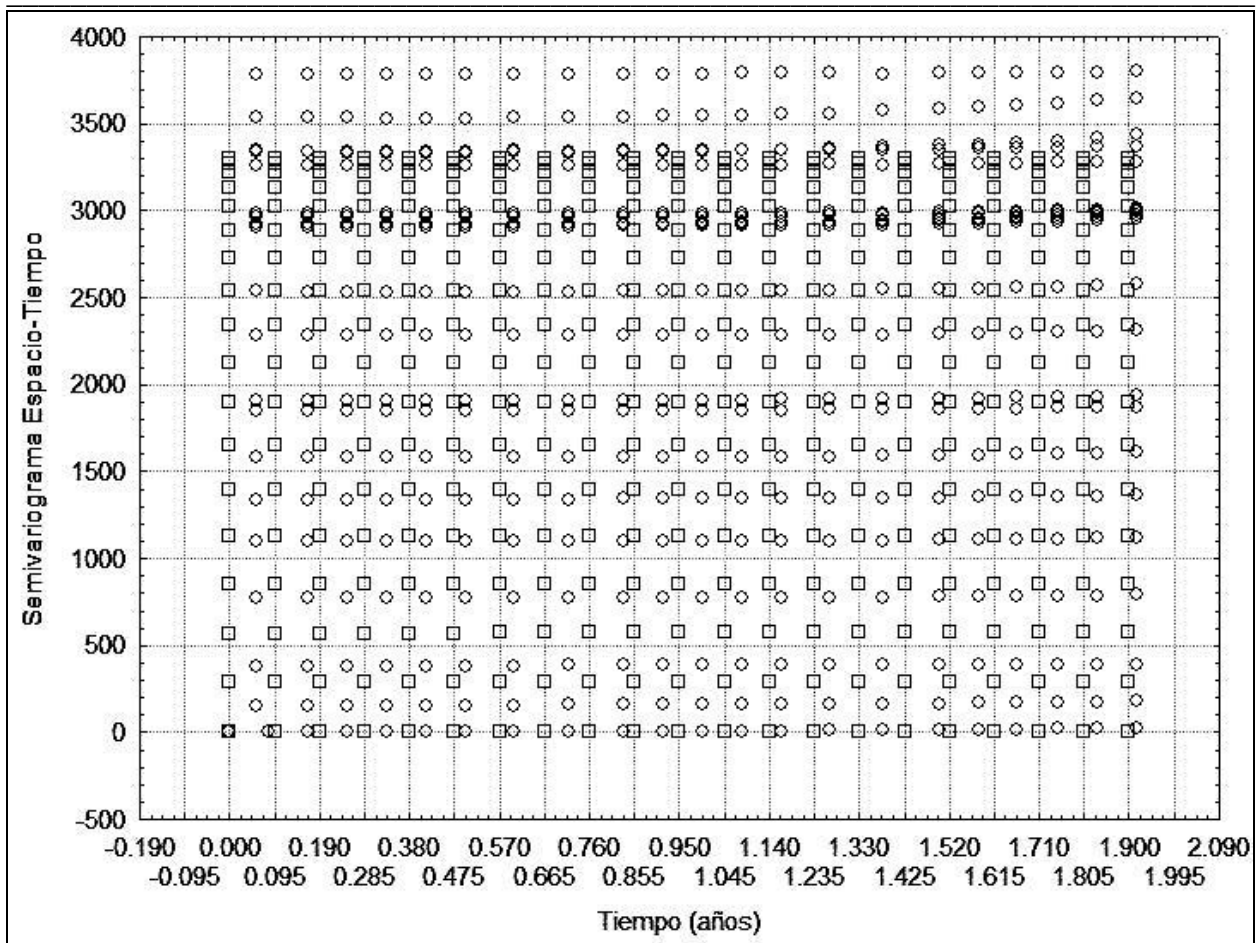


Fig. 5.16 Proyección del semivariograma espacio-temporal muestral de la carga hidráulica sobre el plano semivariograma-intervalos temporales (círculos) y modelo de semivariograma ajustado (cuadros)

Ajuste del modelo de semivariograma

Siguiendo a De Iaco *et al.* (2002), un modelo producto-suma se ajustó manualmente al semivariograma espacio-temporal muestral (esto quiere decir que diferentes modelos fueron propuestos “a ojo” y evaluados para seleccionar aquél con los mejores resultados en la validación cruzada para el periodo de agosto del primer año a julio del tercer año). Los parámetros del modelo seleccionado se presentan en la Tabla 5.5.

Tabla 5.5 Parámetros del modelo de semivariograma espacio-temporal de la carga hidráulica

Modelo	Componente	Pepita	Meseta	Alcance
Esférico	Espacial	0	3300	24000
Esférico	Temporal	0	2.98	1.264
-	Global	-	3300	-

Los resultados de la validación cruzada para el periodo de agosto del primer año a julio del tercer año se presentan en la Tabla 5.6; como puede verse, los errores máximos y mínimos son pequeños, considerando la alta variabilidad de los datos.

Tabla 5.6 Resultados de la validación cruzada para el periodo de agosto del primer año a julio del tercer año

Error mín (m)	Error máx (m)	Error medio (m)	ECM (m²)	RECM (m)	ECME
-4.86	3.79	0.00	0.06	0.24	1.00

5.4 Optimización de la red de monitoreo

El filtro de Kalman estático necesita de estimaciones a priori del vector de estado y de la matriz de covarianza del error en la estimación. La estimación a priori del vector de estado se integró para cada año, con el promedio de las simulaciones de la carga hidráulica correspondientes. La matriz de covarianza espacio-temporal a priori fue calculada del modelo de semivariograma espacio-temporal, aplicando la ecuación 3.15; se construyó para 109 posiciones de monitoreo/estimación y 24 tiempos, de tal manera que su dimensión fue de 2616 columnas (109×24) por 2616 renglones. Por otro lado, se supone que no hay error de medición en los datos generados por la red de monitoreo.

Orden de prioridad

Cuando se aplica el método de optimización de redes de monitoreo propuesto en esta tesis, el orden en el cual son seleccionadas las posiciones de monitoreo espacio-temporal de forma natural representa un orden de prioridad, porque en cada paso se selecciona la posición que reduce al máximo la varianza (en otras palabras, la posición que ofrece el máximo de información). De esta manera, el orden de selección es un indicador de la importancia de los datos obtenidos en esas posiciones, en la reducción de la varianza total: si el orden de selección es pequeño, la prioridad es grande. Esto puede verse claramente en la Fig. 5.17, la varianza total se reduce grandemente al elegir las primeras posiciones de monitoreo espacio-temporal, pero conforme se incrementa el número de posiciones seleccionadas, la reducción de la varianza total decrece.

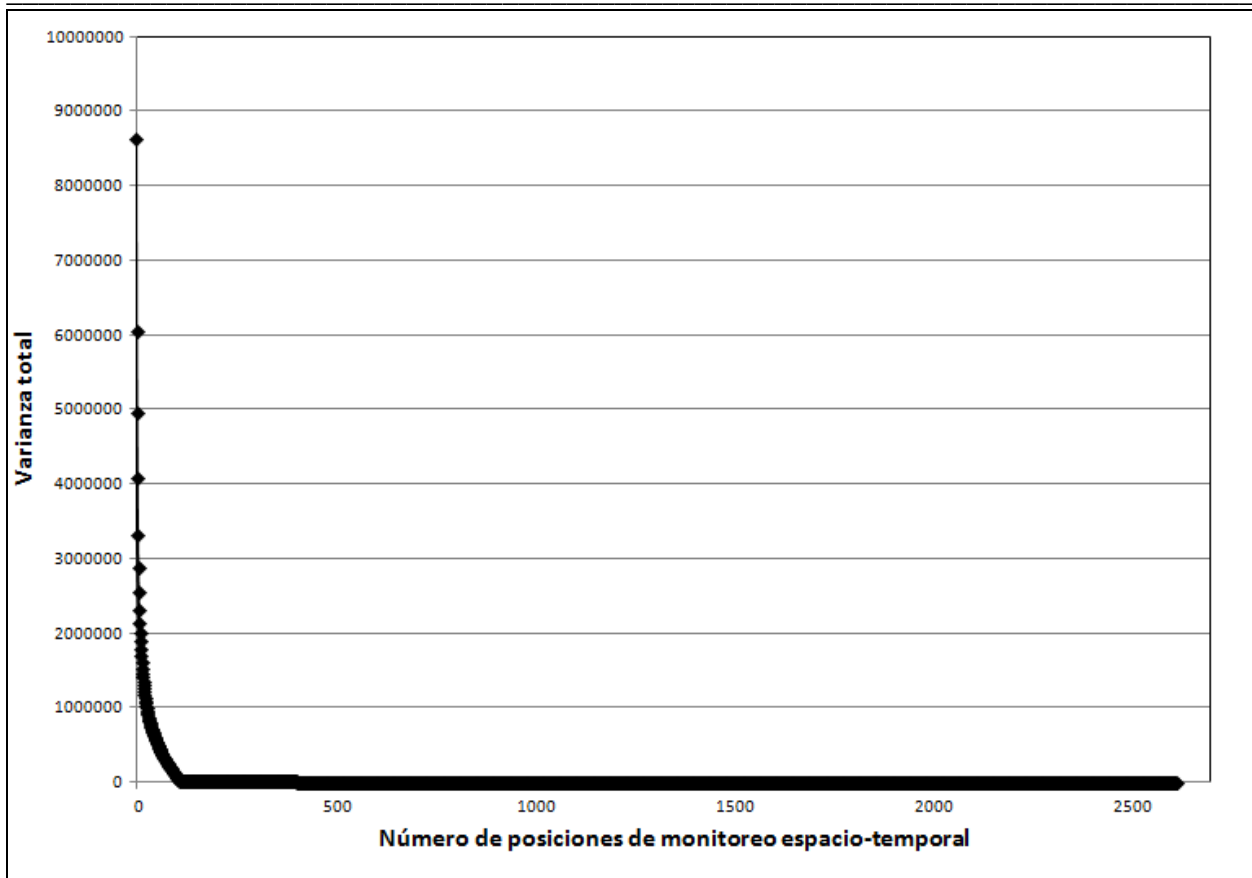


Fig. 5.17 Varianza Total vs Número de posiciones de monitoreo espacio-temporal

5.5 Determinación del número total de posiciones de monitoreo

En la Fig. 5.17 puede verse que después de monitorear en las 100 posiciones con la mayor prioridad, la varianza total permanece prácticamente constante (en comparación con la reducción total). Esto sugiere que las 2516 posiciones restantes pudieran ser redundantes, esto es, proveen poca información y no contribuyen mucho en la reducción de la varianza total.

Para definir el número de posiciones de la red de monitoreo óptimo espacio-temporal, se empleó el criterio descrito en el capítulo cuatro. Los valores calculados de la raíz cuadrada ($\sqrt{\text{Varianza total} / \text{Posiciones de estimación espacio-temporal}}$) contra el número de posiciones de monitoreo espacio-temporal se muestran en la Fig. 5.18.

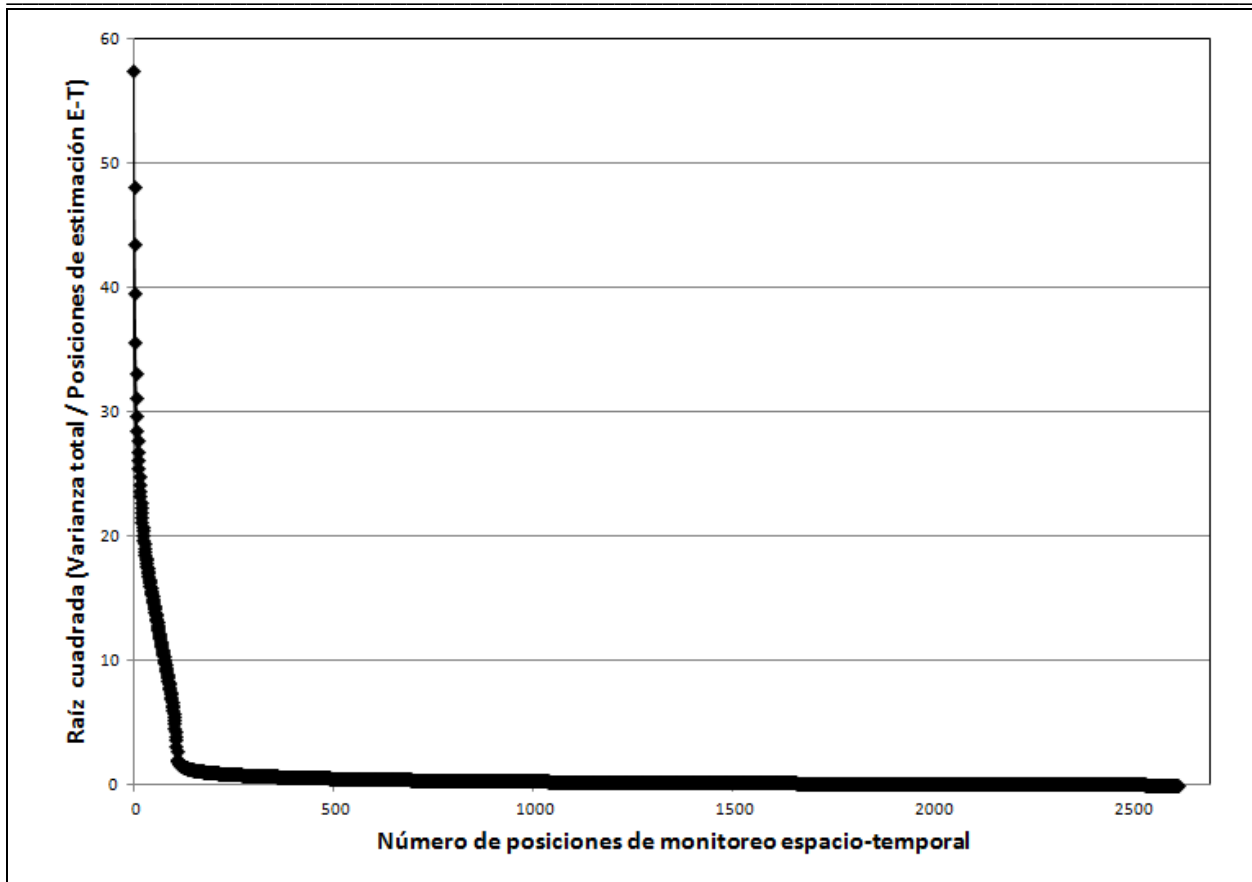


Fig. 5.18 Raíz cuadrada (Varianza total/Posiciones de estimación espacio-temporal) vs Número de posiciones de monitoreo espacio-temporal

El 99% del valor de la MRP (ecuación 4.3) se alcanza cuando se han seleccionado 477 posiciones de monitoreo espacio-temporal y se espera una incertidumbre en la estimación promedio de 0.58 m en la malla de estimación espacio-temporal. Este valor es muy bueno, como se esperaba, debido a que la red de monitoreo se encuentra muy bien distribuida sobre todo el acuífero. Para el caso de estudio, las posiciones de monitoreo espacio-temporal coinciden con las de estimación, entonces es posible reducir la desviación estándar del error en la estimación promedio a 0.00 m si se monitorean todas las posiciones; sin embargo, debe señalarse que antes de elegir alguna posición de monitoreo se tiene que la raíz cuadrada de la varianza promedio del error en la estimación de la carga hidráulica es de 57.44 m, esto significa que las primeras 477 posiciones la reducen en 56.86 m, mientras que las últimas 2139 posiciones únicamente la reducen 0.58 m más. Estos resultados muestran que el método propuesto es exitoso en propagar la información en el espacio y en el tiempo.

5.6 Red de monitoreo óptima

En este ejemplo y de acuerdo a lo expuesto anteriormente, 477 posiciones fueron seleccionadas para constituir la red de monitoreo óptima y su programa de monitoreo. La Fig. 5.19 muestra el número de pozos en los que debe medirse la carga hidráulica para cada tiempo de monitoreo. Puede verse que la campaña más intensa de monitoreo se propone para el primer mes del período de diseño, lo cual es consistente con el criterio de optimización elegido. Todas las posiciones del

primer mes del período de diseño son las primeras posiciones en ser seleccionadas por el método; esto se explica por dos razones: la primera, porque en la reducción de la varianza total para ese mes no participan ni monitoreos previos (porque aquí inicia el monitoreo), ni tampoco participan los monitoreos posteriores (al estar definida la función objetivo para una estimación en tiempo real), y la segunda, porque la varianza temporal es muy pequeña con respecto a la espacial (Tabla 5.5). Por lo tanto, resulta prioritario capturar primero, de manera adecuada, la distribución espacial de la carga hidráulica; en este sentido, el monitorear todas las posiciones del primer mes nos revela el patrón que seguirá la distribución espacial de la carga hidráulica para todo el periodo, y las posiciones seleccionadas en cada mes posterior adecuarán esta distribución en su correspondiente fecha de monitoreo.

En la Fig. 5.19 se aprecia que después del monitoreo exhaustivo en el primer mes, son necesarias pocas posiciones en los meses posteriores (la información recabada en el primer mes aporta mucha información sobre todo el periodo de monitoreo); del segundo al séptimo mes se incrementa la necesidad de información, pero a partir de allí, se presenta un descenso periódico gradual que responde a la influencia que tiene el monitoreo en cada mes sobre la estimación en tiempos futuros.

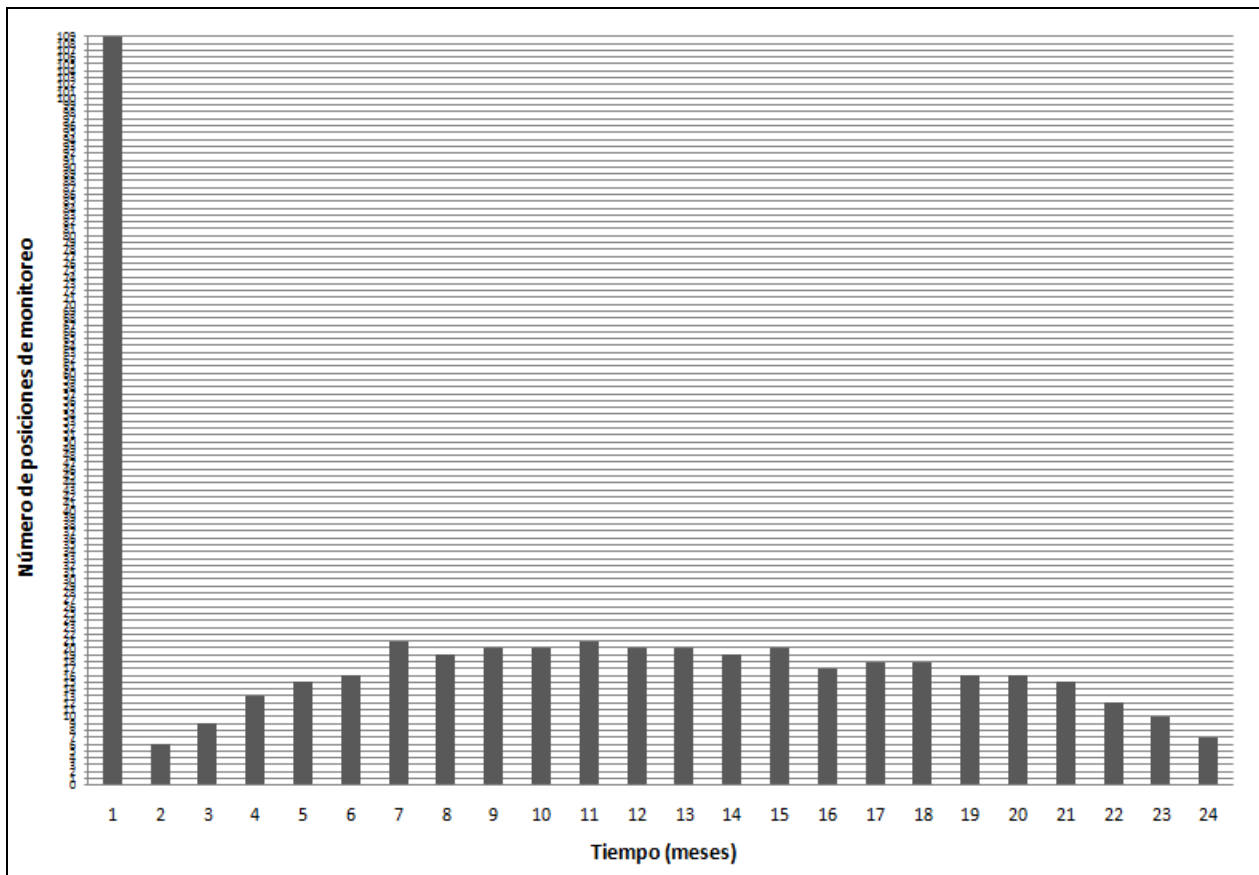


Fig. 5.19 Posiciones de monitoreo para cada mes

La Fig. 5.20 muestra el número de veces que debe ser monitoreado cada pozo. Como puede verse, la posición que necesita ser monitoreada con mayor frecuencia (9 veces) está localizada justo en la parte central del acuífero (pozo 55), esto se explica porque esa posición es la que tiene influencia sobre un mayor número de nodos de la malla de estimación.



Fig. 5.20 Número de veces de monitoreo para cada pozo en el programa de monitoreo óptimo espacio-temporal (arriba) y nombre del pozo (abajo)

5.7 Estimaciones y varianzas del error en la estimación para el período de diseño

Como una ilustración del nivel de información que proporciona el programa de monitoreo propuesto para el periodo de diseño, se obtuvieron estimaciones de la carga hidráulica y de la varianza del error de estimación sobre las posiciones y tiempos de interés. Éstas se obtuvieron aplicando el filtro de Kalman estático.

Estimación de la carga hidráulica

En las Figuras 5.21, 5.23 y 5.25 se presentan estimaciones en tiempo real de la carga hidráulica, obtenidas para algunos meses con la red de monitoreo propuesta.

Estas figuras muestran los cambios espaciales y temporales de la carga hidráulica estimados en el acuífero. Las zonas de recarga corresponden a las zonas con el máximo nivel de la carga

hidráulica, las cuales coinciden con los máximos niveles topográficos, al norte, noreste, este y suroeste del acuífero. El flujo del agua subterránea es radial, con una dirección preferencial noreste-suroeste (los niveles máximos en el rango 1880-1990 msnm), y otra viniendo del sur con dirección hacia el centro (con niveles máximos en el rango 1800-1820 msnm); la parte central del acuífero presenta los niveles más bajos con el cono de abatimiento más grande localizado al noroeste de la ciudad (por debajo de 1725 msnm y alcanzando en algunas temporadas valores mínimos de 1716 msnm). A través del tiempo, niveles entre 1730 y 1740 msnm se extienden desde la parte central hacia los límites del acuífero.

Varianzas del error en la estimación

En las Figuras 5.22, 5.24 y 5.26 se presentan las varianzas del error en la estimación obtenidas en tiempo real para algunos meses, con la red de monitoreo y el programa de monitoreo propuestos. El orden de prioridad se incluye también.

Para el primer mes la varianza es cero en todas las posiciones de estimación porque todas ellas son monitoreadas en esa ocasión. Para los otros meses, los valores mínimos se sitúan alrededor de las posiciones de monitoreo, y se incrementan en las zonas pobremente cubiertas por éstas. Se puede ver que para todo el periodo de estimación, las varianzas se encuentran por debajo de 0.6 m^2 , para casi toda el área de estudio; lo cual nos habla de la adecuada cobertura de la red propuesta. Estos mapas de varianza junto con criterios prácticos e hidrogeológicos pueden ser muy útiles en la elección de nuevas posiciones de monitoreo en las zonas con las máximas varianzas.

DISEÑO DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON ALTO NIVEL DE INFORMACIÓN ESPACIO-TEMPORAL

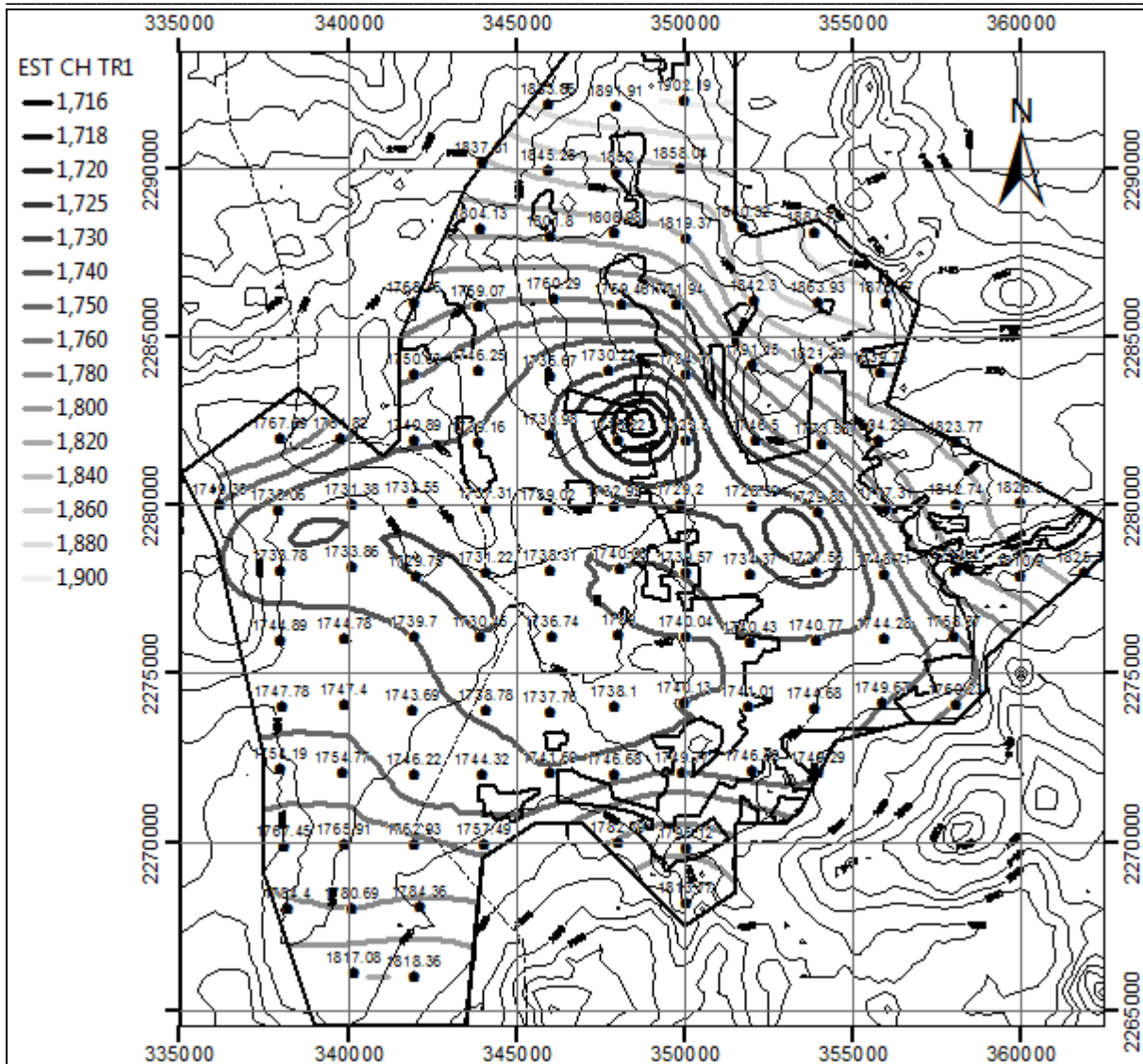


Fig. 5.21 Estimación de la carga hidráulica para agosto del primer año

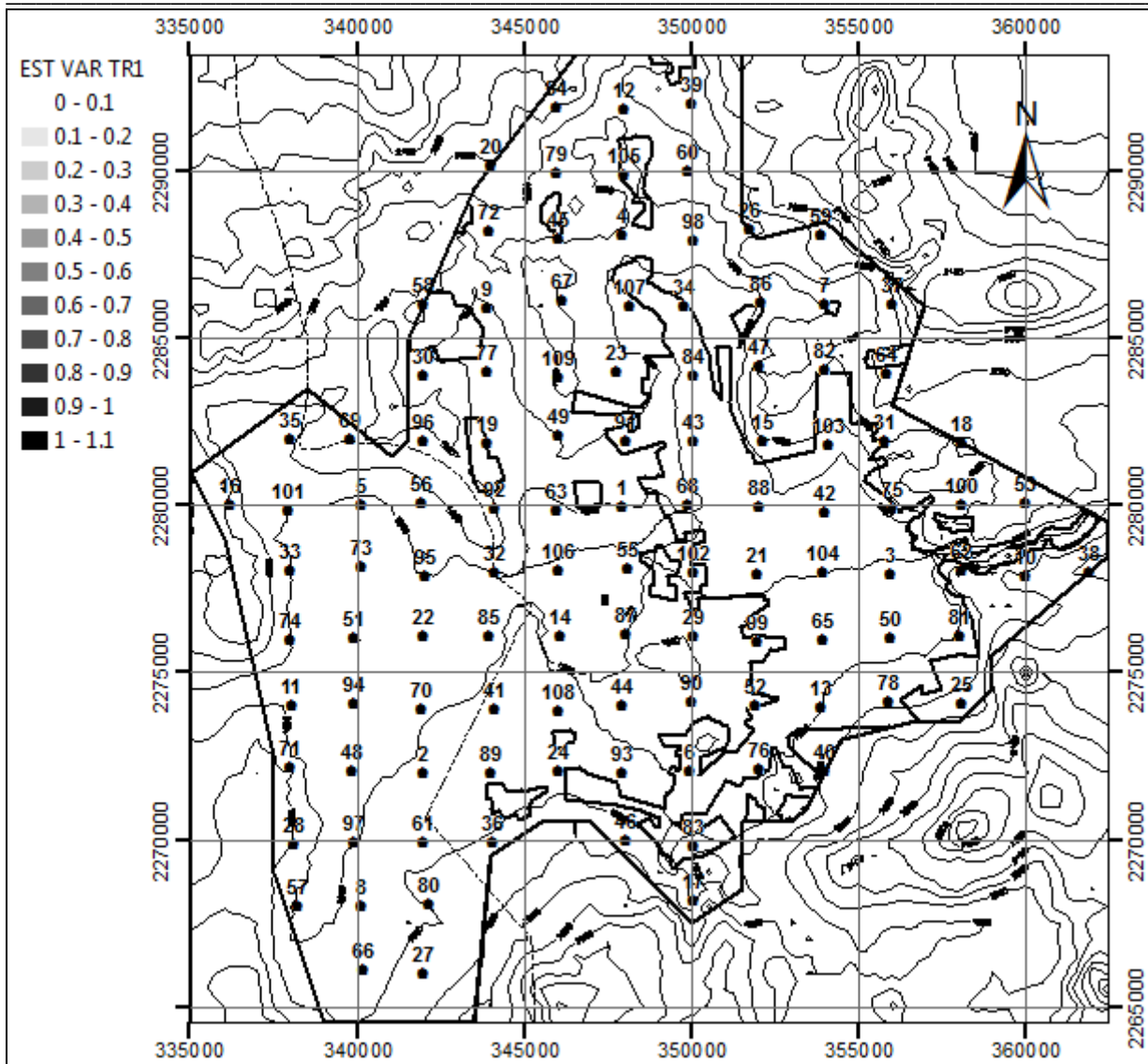


Fig. 5.22 Varianzas del error en la estimación y orden de prioridad de los pozos para agosto del primer año

DISEÑO DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON ALTO NIVEL DE INFORMACIÓN ESPACIO-TEMPORAL

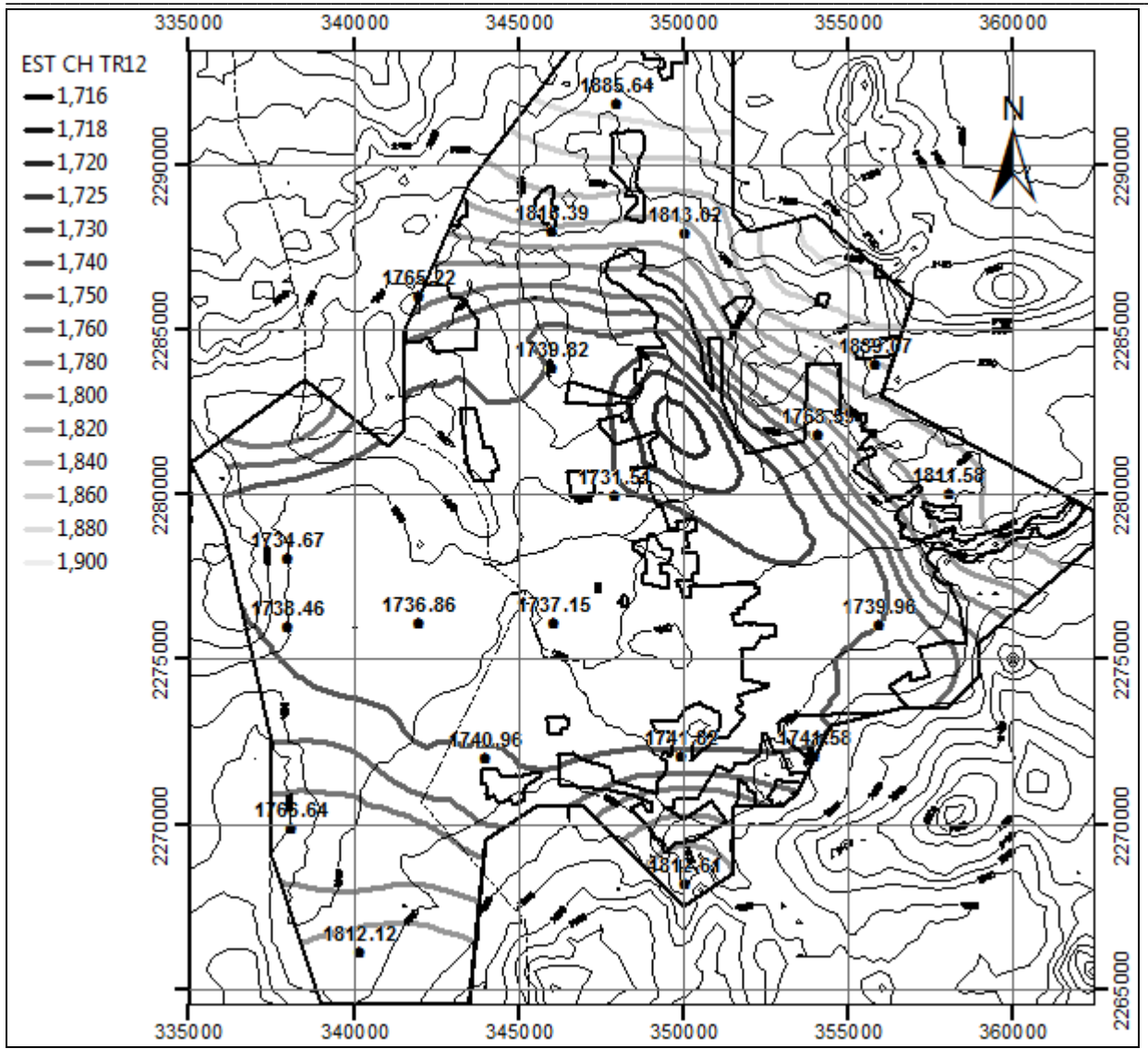


Fig. 5.23 Estimación de la carga hidráulica para julio del segundo año

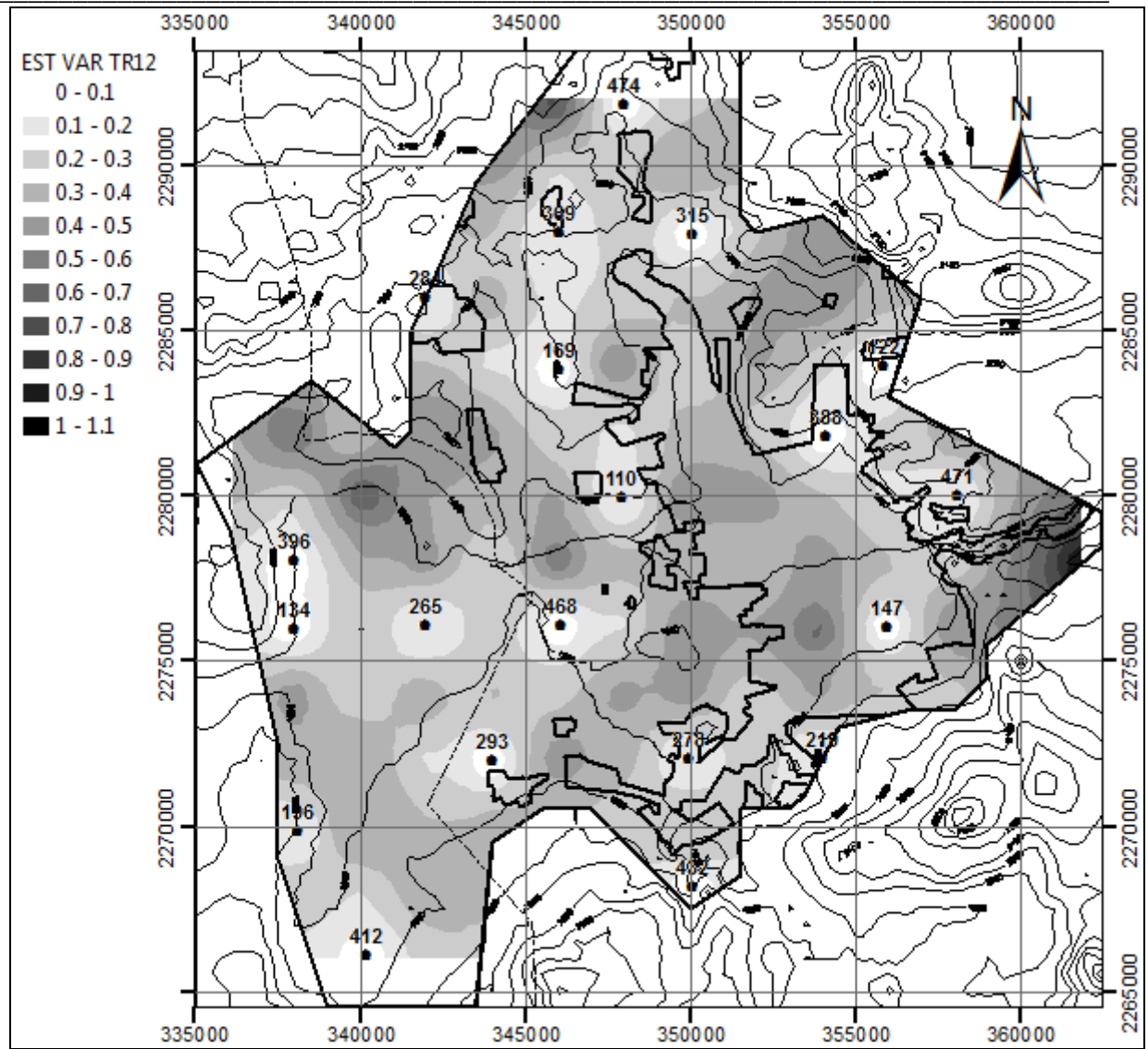


Fig. 5.24 Varianzas del error en la estimación y orden de prioridad de los pozos para julio del segundo año

DISEÑO DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON ALTO NIVEL DE INFORMACIÓN ESPACIO-TEMPORAL

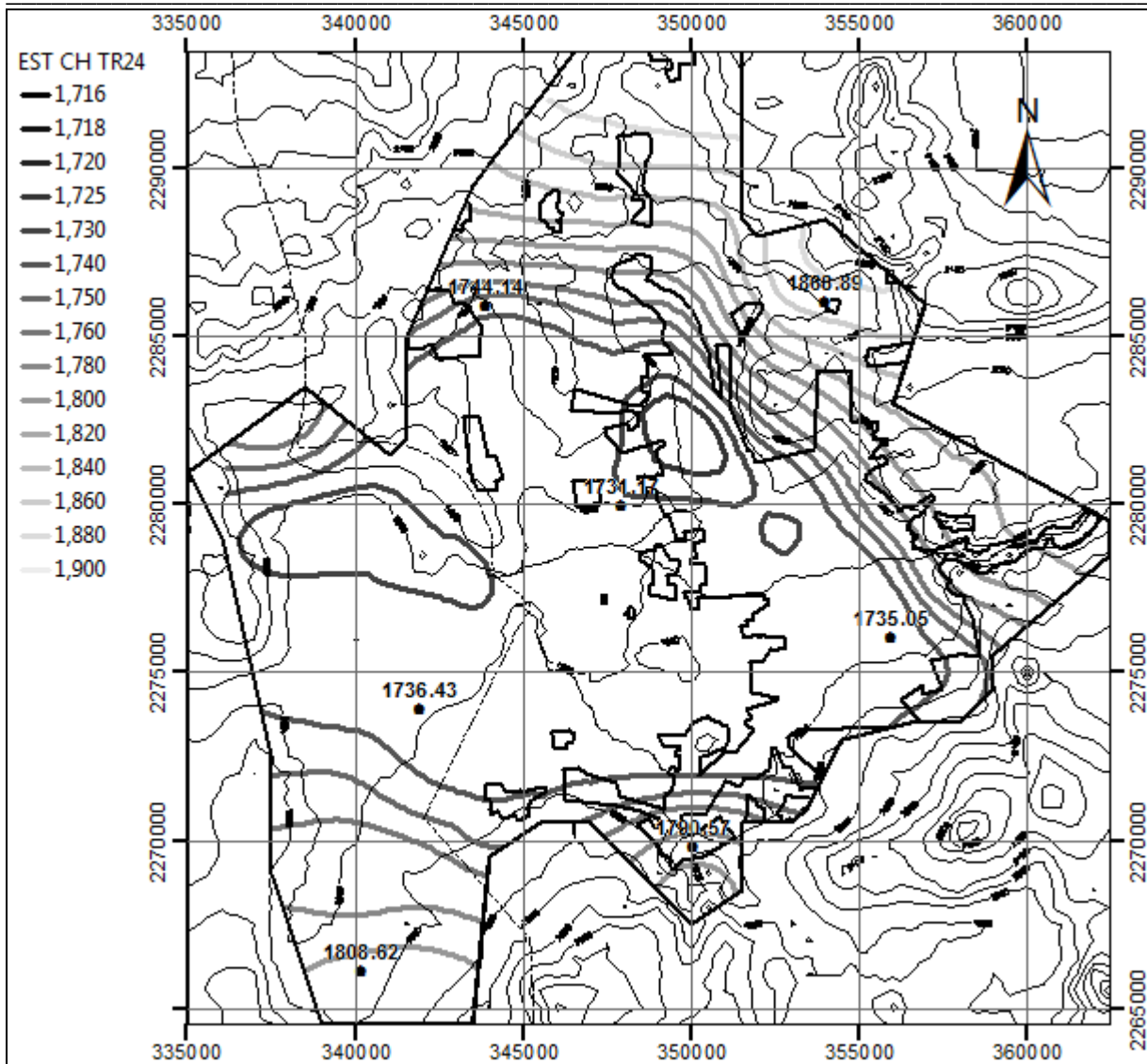


Fig. 5.25 Estimación de la carga hidráulica para julio del tercer año

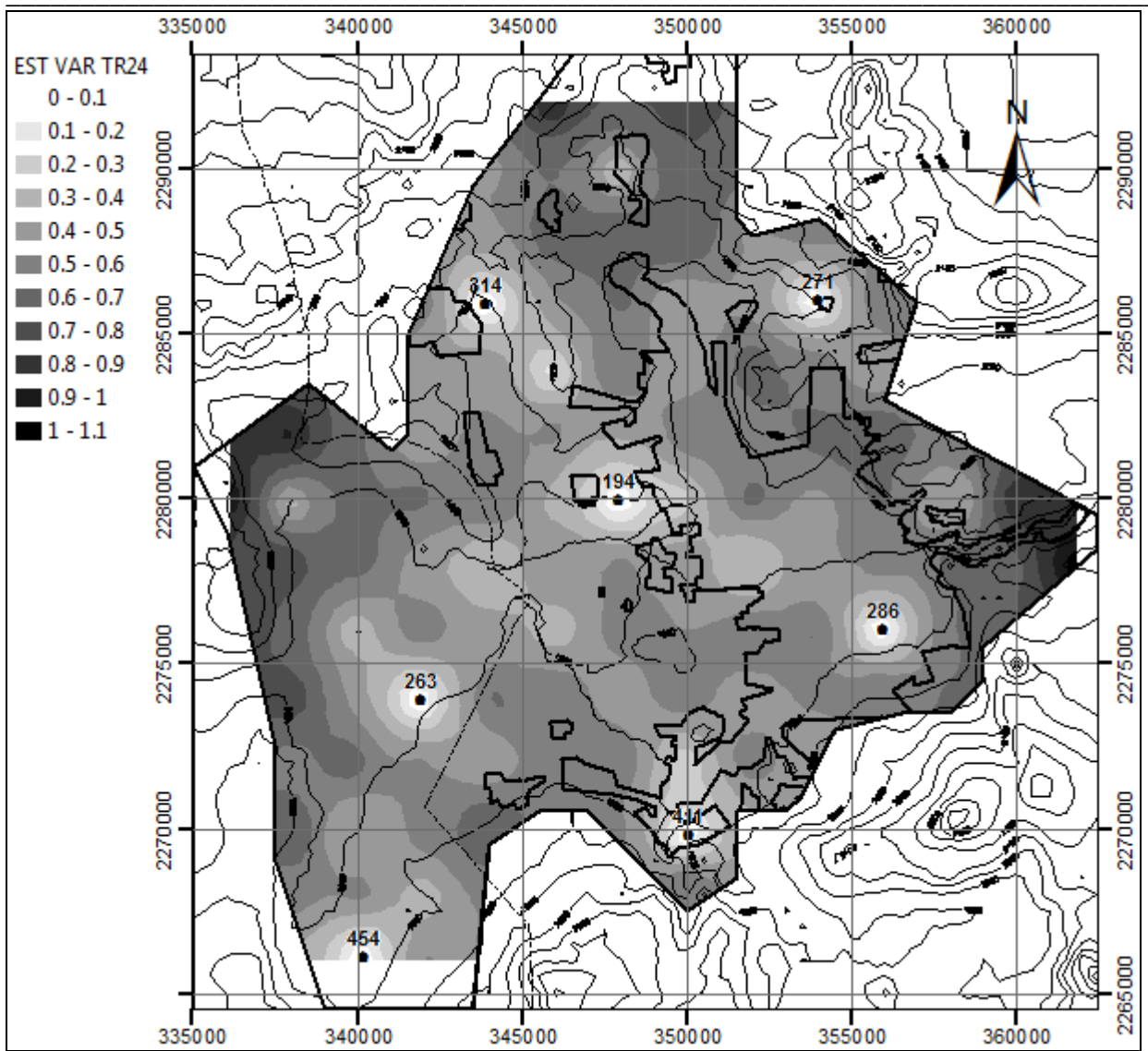


Fig. 5.26 Varianzas del error en la estimación y orden de prioridad de los pozos para julio del tercer año

5.8 Evaluación del desempeño de la red de monitoreo en el futuro

Como en este caso se tienen datos “futuros” de la carga hidráulica (simulaciones del modelo numérico de flujo de agosto del tercer año a julio del quinto año) en todas las posiciones de monitoreo/estimación espacio-temporal, entonces es posible evaluar el desempeño en el futuro de la red de monitoreo. Como el programa de monitoreo seleccionado en el paso previo fue diseñado para un periodo de dos años (agosto del primer año a julio del tercer año), entonces se aplicará el mismo programa para un periodo igual en el futuro (agosto del tercer año a julio del quinto año), considerando que no cambiarán mucho las condiciones de funcionamiento del acuífero. Para verificar el desempeño de la red, se compararon las estimaciones obtenidas con la red de monitoreo propuesta contra los datos de la carga hidráulica en las posiciones de estimación espacio-temporal (obtenidos del modelo numérico de flujo), a la diferencia entre la estimación y el dato se le llamó error en la estimación. Como puede verse en la Tabla 5.7, el

desempeño de la red de monitoreo en el futuro es adecuado; de hecho, los errores en las estimaciones del futuro son menores con respecto a los que se presentan en el periodo de diseño, debido a que conforme se avanza en el tiempo se tiene mayor información contribuyendo en las estimaciones. Un estadístico seleccionado para esta evaluación fue el error cuadrático medio estándar (ECME), éste mide la consistencia entre las varianzas y el error en la estimación. Se define como el promedio del cuadrado de las diferencias entre los valores medidos y los estimados sobre la varianza del error en la estimación, en las posiciones de estimación espacio-temporal.

$$ECME = \frac{1}{N} \sum_{j=1}^{Npe} \sum_{q=1}^{Nte} \left\{ \frac{[Z^*(x_j, t_q) - Z(x_j, t_q)]^2}{\sigma_{j,q}^2} \right\}$$

donde $N = Npe * Nte$ = Número de posiciones de estimación espacio-temporal, y $Z^*(x_j, t_q)$, $Z(x_j, t_q)$ y $\sigma_{j,q}^2$ son respectivamente los valores de la estimación, del dato medido, y de la varianza del error en la estimación, en el punto x_j y en el tiempo t_q .

Valores del ECME cercanos a uno indican que los errores son consistentes con las varianzas estimadas. Para este caso, se observa en la Tabla 5.7 que se obtuvieron valores altos del ECME debido a que la varianza se redujo más rápidamente que los errores, alcanzando en muchas posiciones de estimación valores cercanos a cero, por la buena distribución de la red de monitoreo espacio-temporal.

Tabla 5.7 Desempeño de la red de monitoreo óptima

Periodo de monitoreo (años)	Posiciones de monitoreo espacio-temporal	Error medio (m)	Error mínimo (m)	Error máximo (m)	RECM (m)	ECME
1-4	954	-0.01	-15.14	12.21	2.04	17.95
1-2	477	-0.01	-15.14	12.21	2.39	13.79
3-4	477	-0.01	-11.45	10.00	1.61	22.11

Estimaciones futuras de la carga hidráulica

En las Figuras 5.27 y 5.29 se presentan estimaciones futuras en tiempo real de la carga hidráulica, obtenidas para algunos meses con la red de monitoreo propuesta.

Las principales direcciones de flujo se mantienen para los próximos dos años. La parte central del acuífero continúa con niveles por debajo de 1740 msnm, el mayor cono de abatimiento ubicado al noroeste de la ciudad parece estabilizarse, pero ahora, en una porción grande de la ciudad se presentan niveles por debajo de los 1730 msnm, y el cono de abatimiento que se ubica al oeste del acuífero se ha incrementado para alcanzar niveles por debajo de los 1725 msnm.

Varianzas del error en la estimación futuras

En las Figuras 5.28 y 5.30 se presentan las varianzas del error en la estimación obtenidas en tiempo real para algunos meses del futuro, con los datos de la red de monitoreo y su programa de monitoreo propuesto. El orden de prioridad se incluye también.

Durante el periodo de estimación futuro, las varianzas del error en la estimación se encuentran por debajo de 0.5 m^2 , este valor es más bajo que el obtenido para las estimaciones en el periodo de diseño, porque como es explicó anteriormente, la información previa contribuye en reducir la incertidumbre en la estimación.

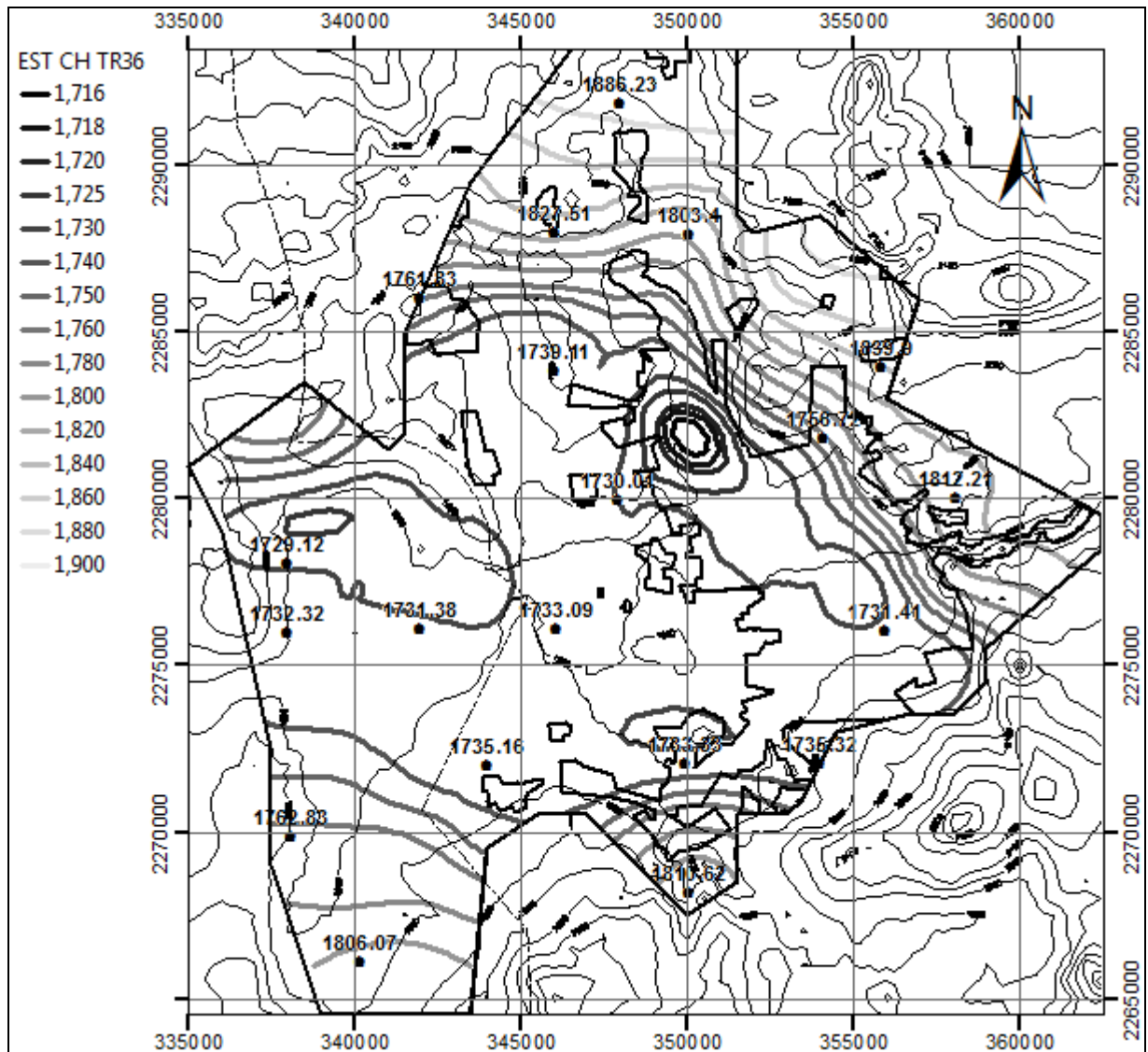


Fig. 5.27 Estimación de la carga hidráulica para julio del cuarto año

DISEÑO DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON ALTO NIVEL DE INFORMACIÓN ESPACIO-TEMPORAL

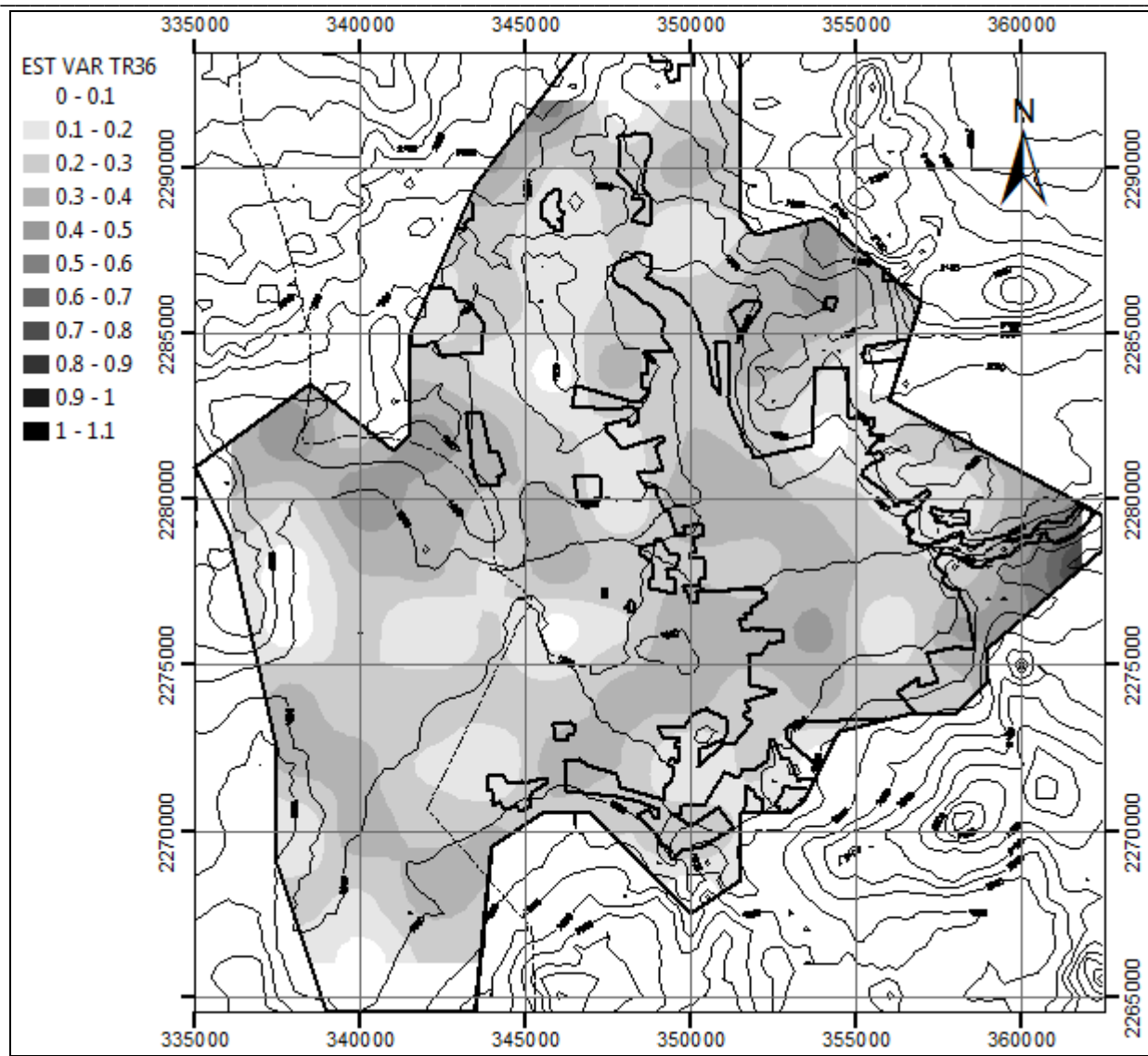


Fig. 5.28 Varianzas del error en la estimación y orden de prioridad de los pozos para julio del cuarto año

DISEÑO DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON ALTO NIVEL DE INFORMACIÓN ESPACIO-TEMPORAL

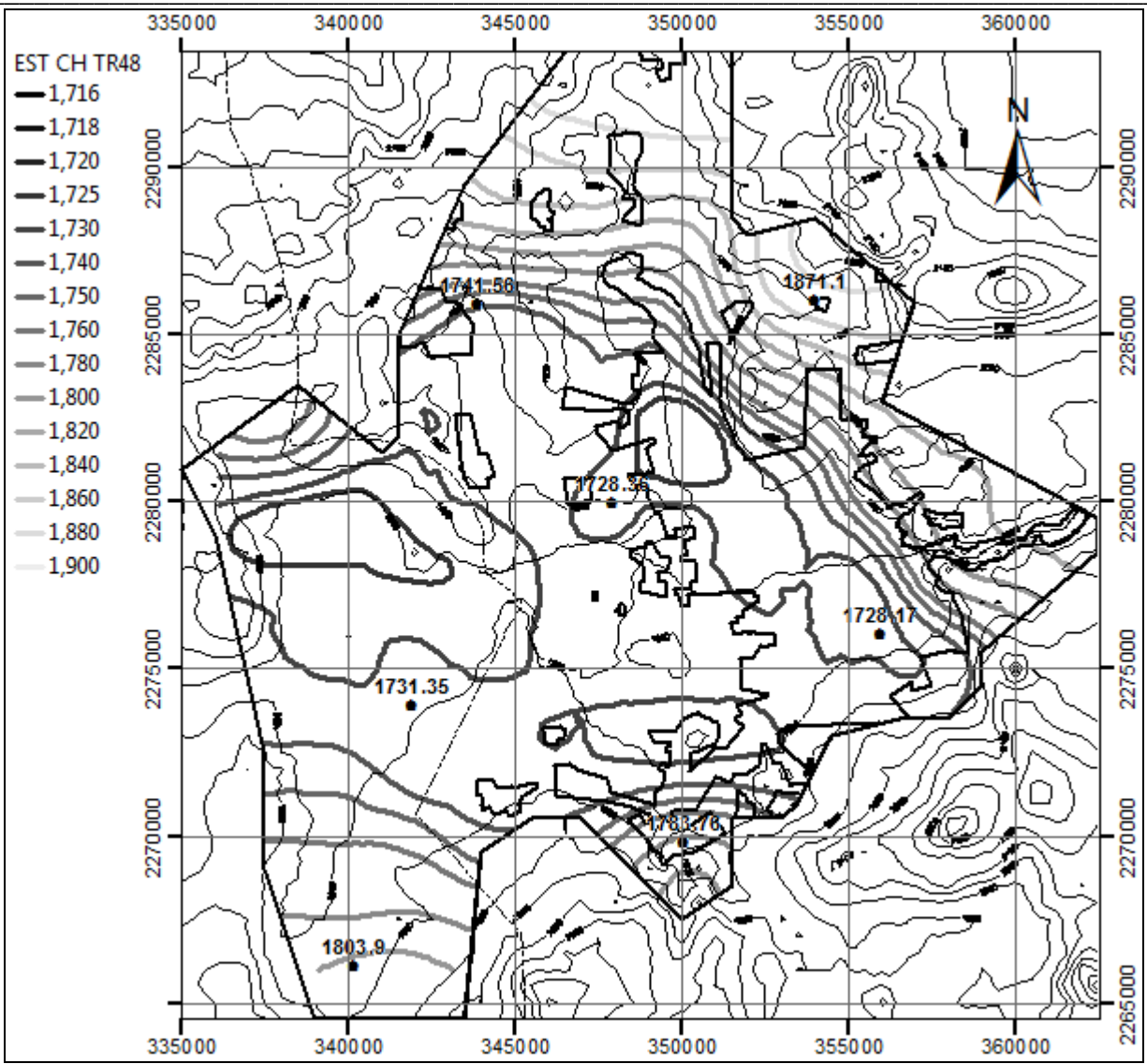


Fig. 5.29 Estimación de la carga hidráulica para julio del quinto año

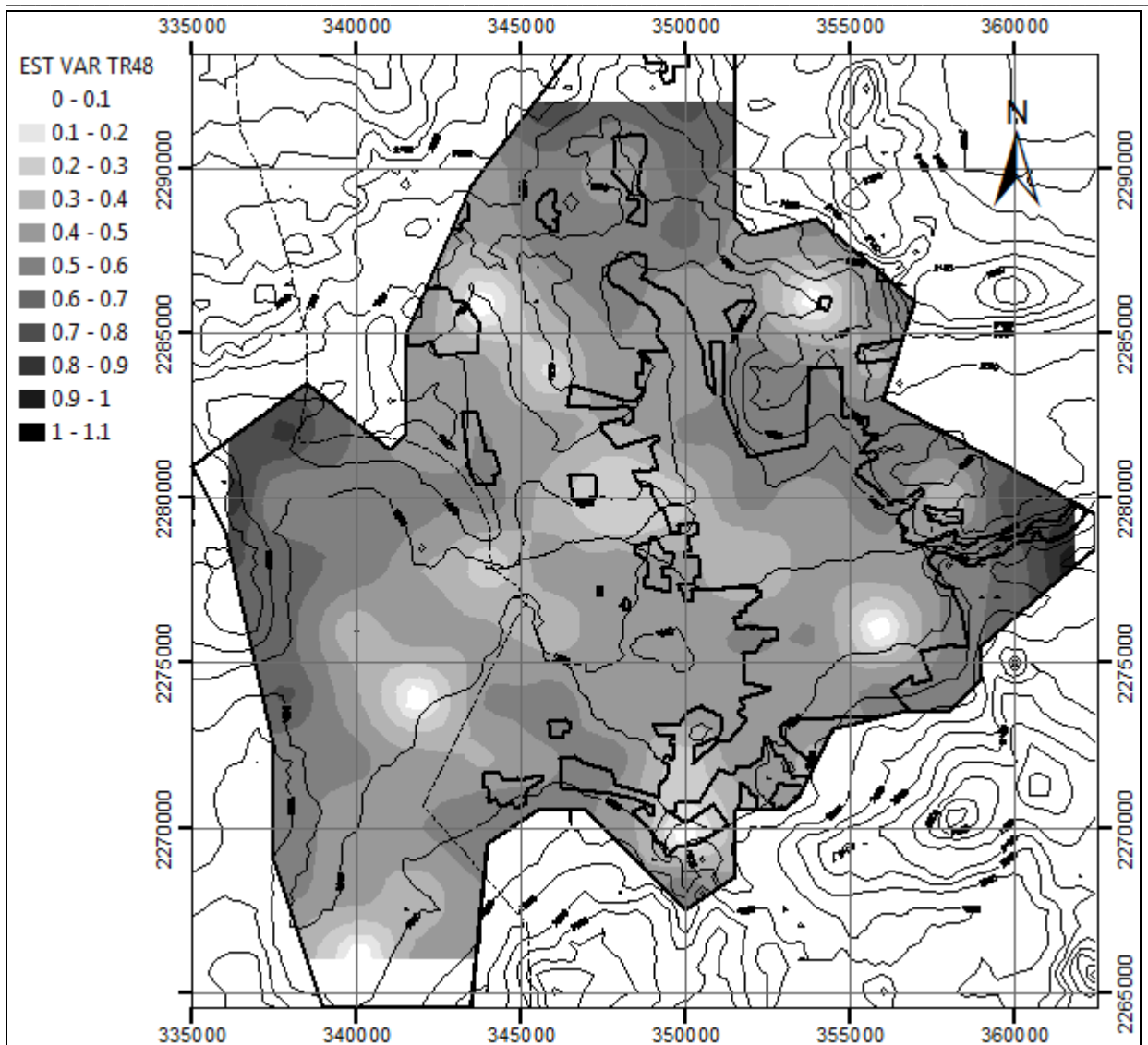


Fig. 5.30 Varianzas del error en la estimación y orden de prioridad de los pozos para julio del quinto año

5.9 Evaluación del desempeño de la red de monitoreo a través del tiempo

Para evaluar la forma en que la red de monitoreo propuesta captura el comportamiento temporal de la carga hidráulica dentro del periodo de agosto del primer año a julio del quinto año, se seleccionaron los pozos L-246, 644-A, L-107, 2082, 1792, 752 y 976, distribuidos de manera más o menos uniforme en el acuífero (Fig. 5.4). Las Figuras 5.5 - 5.11 muestran las estimaciones de la carga hidráulica para cada uno de estos pozos durante el periodo de monitoreo de la red y las simulaciones obtenidas del modelo numérico de flujo. Se observa que para los pozos en los que se tiene un descenso sostenido de los niveles (L-246, 644-A, L-107, 2082) la red de monitoreo captura muy bien el comportamiento de la carga hidráulica con un error mínimo en la estimación de -1.66 m y un máximo de 4.26 m, esta estimación adecuada de los niveles se debe, de manera local, a los datos que aportan las posiciones que son monitoreadas a través del tiempo, y de manera global al descenso sostenido de los niveles en las estimaciones debido a la forma en

que fue integrado el vector de estado a priori, en donde se propuso este comportamiento en el acuífero a través del tiempo; sin embargo, esto último propicia que en los sitios donde existe una recuperación de los niveles a través del tiempo (pozos 1792, 752 y 976) se obtengan los mayores errores en la estimación de la carga hidráulica. La estructura del vector de estado obliga a que los niveles descendan conforme se avanza en el tiempo, únicamente el monitoreo logra corregir las estimaciones de manera local. Las simulaciones obtenidas para el pozo 752 muestran recuperación de los niveles en los primeros meses (agosto del primer año a septiembre del segundo año) y después parece que éstos se estabilizan en el periodo de octubre del segundo año a julio del quinto año, allí se presenta el máximo error de estimación para los pozos analizados (-10.55 m). En la Tabla 5.8 se presenta un resumen de algunos estadígrafos de los errores en la estimación para los pozos analizados.

Tabla 5.8 Errores en la estimación a través del tiempo para algunos pozos utilizando la red de monitoreo propuesta

Pozo	Error mín (m)	Error máx (m)	Error medio (m)	RECM (m)
L-246	-1.44	1.34	0.27	0.59
0644-A	-0.43	4.26	1.27	1.76
L-107	-1.66	1.66	-0.16	0.62
2082	-1.59	3.62	0.17	0.92
1792	-5.89	0.87	-1.21	1.75
752	-10.55	0.00	-3.08	3.85
976	-4.36	0.00	-1.03	1.51

5.10 Evaluación del desempeño de la red de monitoreo en la captura de la distribución espacial

Se consideró importante evaluar las estimaciones de la carga hidráulica en el corte A-A' de la Fig. 5.4 porque atraviesa la zona donde se ubican los abatimientos de niveles más significativos y además comprende otros sitios importantes como son zonas de recarga del acuífero. Como se observa en la Fig. 5.12, la red de monitoreo estima de forma adecuada los comportamientos para el corte mostrado para las fechas inicial (agosto del primer año) y final (julio del quinto año) del periodo de monitoreo. Debe señalarse que para este caso, los niveles de las simulaciones y las estimaciones del primer mes de monitoreo coinciden porque en esa fecha se monitorean todas las posiciones de estimación. En la Tabla 5.9 se presentan estadígrafos de los errores en la estimación de la carga hidráulica en el corte A-A' para el periodo de monitoreo, se aprecia que el error mínimo es de -4.92 m para marzo del cuarto año en la latitud 2290041.983, el error máximo es de 10.03 m en junio del quinto año en la latitud 2287964.109, el error medio más grande es de 1.72 m para noviembre del primer año, mientras que la RECM máxima es de 3.53 m para junio del tercer año, este valor se encuentra prácticamente a la mitad del periodo de monitoreo. Estos resultados son satisfactorios para una configuración tan complicada de capturar a través del tiempo.

Tabla 5.9 Errores en la estimación de la carga hidráulica para el corte A-A' de la Fig. 5.4 utilizando la red de monitoreo propuesta

Fecha de monitoreo	Error mín (m)	Error máx (m)	Error medio (m)	RECM (m)
ago-01	0.00	0.00	0.00	0.00
sep-01	-0.31	2.25	0.56	0.96
oct-01	-1.23	3.54	0.36	1.26
nov-01	-0.95	7.33	1.72	3.09
dic-01	-4.27	6.46	0.40	2.83
ene-02	-1.49	6.73	1.36	2.71
feb-02	-3.95	5.17	0.70	2.54
mar-02	-4.28	3.68	-0.13	2.44
abr-02	-2.63	3.70	0.04	1.60
may-02	-3.28	4.26	0.09	2.10
jun-02	-2.40	3.77	0.44	2.08
jul-02	-3.31	1.30	-1.06	1.91
ago-02	-1.63	4.63	0.67	1.90
sep-02	-0.39	6.04	1.16	2.15
oct-02	-1.47	5.62	0.61	2.04
nov-02	-3.70	3.16	-0.81	2.09
dic-02	-2.07	3.71	0.27	1.61
ene-03	-3.17	4.85	-0.29	1.97
feb-03	-4.53	5.57	-0.10	2.63
mar-03	-4.54	7.31	0.02	3.03
abr-03	-1.50	6.41	0.96	2.64
may-03	-2.55	5.54	0.39	2.45
jun-03	-3.71	9.07	0.97	3.53
jul-03	-2.34	8.71	1.12	3.26
ago-03	0.00	0.00	0.00	0.00
sep-03	-0.47	0.74	0.09	0.36
oct-03	-1.18	2.30	-0.10	0.91
nov-03	-1.37	4.09	0.53	1.50
dic-03	-4.24	3.94	-0.28	1.98
ene-04	-2.45	6.43	1.15	2.48
feb-04	-4.69	5.88	0.30	2.42
mar-04	-4.92	2.01	-0.84	2.24
abr-04	-3.27	3.24	0.17	1.61
may-04	-3.66	4.55	0.20	1.99
jun-04	-3.33	4.30	0.06	2.04
jul-04	-4.66	0.66	-0.97	1.76
ago-04	-2.55	2.83	0.39	1.47
sep-04	-1.09	3.93	0.72	1.70
oct-04	-1.95	5.18	0.52	1.95
nov-04	-3.10	4.17	-0.65	1.95
dic-04	-1.69	5.09	0.38	1.77
ene-05	-2.87	6.34	-0.09	2.11
feb-05	-4.19	7.23	-0.04	2.65
mar-05	-4.51	8.71	0.09	2.97
abr-05	-3.13	7.11	0.51	2.49
may-05	-3.27	6.80	0.24	2.45
jun-05	-4.71	10.03	0.79	3.39
jul-05	-4.32	9.87	0.75	3.25

5.11 Conclusiones y recomendaciones

La metodología propuesta para el diseño óptimo de redes de monitoreo de la carga hidráulica incorpora la correlación espacio-temporal entre los datos históricos medidos de esta variable en un acuífero, usando un modelo de semivariograma espacio-temporal obtenido de un análisis geoestadístico. Una ventaja importante del método es que considera la correlación cruzada entre las cargas hidráulicas en diferentes posiciones y tiempos, lo que permite una evaluación más completa de las posiciones y/o tiempos de monitoreo redundantes.

Los resultados muestran que la raíz cuadrada de la varianza promedio del error en la estimación de la carga hidráulica en el área de estudio, cuando se monitorean las 477 posiciones con la más alta prioridad es menor a 0.58 m. Para el caso de estudio, antes de elegir alguna posición de monitoreo se tiene un valor de la raíz cuadrada de la varianza promedio del error en la estimación de 57.44 m, esto significa que las primeras 477 posiciones reducen este valor en 56.86 m, mientras que las últimas 2139 posiciones únicamente lo reducen 0.58 m más. Estos resultados muestran que el método propuesto es exitoso en propagar la información en el espacio y en el tiempo, lo que se traduce en una reducción significativa del costo asociado al monitoreo si consideramos que éste es proporcional al número de posiciones monitoreadas.

La fecha de mayor monitoreo para la red óptima seleccionada fue el primer mes disponible dentro del periodo de diseño, esto se explica por dos razones: la primera, porque en la reducción de la varianza total del error en la estimación para ese mes no contribuyen ni datos de monitoreos previos, ni de monitoreos posteriores, y la segunda, porque la varianza temporal es muy pequeña con respecto a la espacial. Por lo tanto, resulta prioritario estimar adecuadamente primero la distribución espacial de la carga hidráulica; en este sentido, el monitorear todas las posiciones del primer mes nos revela el patrón que seguirá la distribución espacial de la carga hidráulica para todo el periodo analizado, y las posiciones seleccionadas en cada mes posterior adecuarán esta distribución en su correspondiente fecha de monitoreo.

La evaluación en el futuro confirma que la red de monitoreo y el programa de monitoreo propuesto permiten estimar de manera adecuada la evolución de la carga hidráulica en un periodo de tiempo igual al periodo de diseño, mientras las condiciones del acuífero no cambien significativamente, es decir, la red y su programa tendrán que rediseñarse para los casos donde se presenten cambios de uso de suelo o de los volúmenes de extracción de agua subterránea que propicien cambios radicales en la configuración de la distribución de la carga hidráulica en el acuífero.

La raíz del error cuadrático medio para el periodo de diseño es de 2.39 m mientras que para el periodo de evaluación en el futuro es de 1.61 m; esto se debe a que a que la información obtenida durante el periodo de diseño colabora para mejorar las estimaciones de la carga hidráulica en el futuro.

Se observó que la red de monitoreo captura la distribución espacial de los niveles de carga hidráulica en el acuífero de manera aceptable a pesar de presentarse configuraciones muy complicadas de capturar como en donde se presentan abatimientos locales significativos.

Los valores de la varianza del error en la estimación se encuentran por debajo de 0.6 m^2 para el periodo de diseño, y por debajo de 0.5 m^2 para el periodo de evaluación en el futuro, para casi toda el área de estudio. Esto demuestra que la cobertura de la red propuesta es adecuada. Estos mapas de varianza junto con criterios prácticos e hidrogeológicos pueden ser muy útiles en la elección de nuevas posiciones de monitoreo en las zonas con las máximas varianzas.

Para este caso en particular, el método geoestadístico espacio-temporal empleado probó ser muy eficiente al obtenerse estimaciones de los niveles de la carga hidráulica muy cercanas a las simulaciones del modelo numérico de flujo, a pesar de no emplear las ecuaciones que describen la física del problema.

6. EVALUACIÓN DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON BAJO NIVEL DE INFORMACION ESPACIO-TEMPORAL

Es muy probable que una red de monitoreo sea insuficiente para describir el comportamiento espacio-temporal de la carga hidráulica en zonas de interés de un acuífero y/o que las posiciones que conforman la red de monitoreo provean información redundante (que se traduce en costos innecesarios), si no se ha considerado un criterio de optimización para su diseño.

En este capítulo se presenta una evaluación de una red de monitoreo a través de la metodología para el diseño espacio-temporal óptimo de redes de monitoreo de la carga hidráulica introducida en el capítulo cuatro. Se evalúa el desempeño de una red que emula un caso real y se compara con un caso cercano a lo ideal (la presentada en el capítulo cinco). También se analiza el efecto en el análisis geoestadístico de huecos en la información espacio-temporal de la carga hidráulica, ya que al tratar de emular un caso real, en el ejemplo de este capítulo se emplea sólo una parte de los datos utilizados en el análisis presentado en el capítulo cinco.

6.1 Base de datos de la carga hidráulica

Para la evaluación del desempeño de la red de monitoreo se seleccionó información mensual de la carga hidráulica en 109 posiciones (Fig. 6.1) que cubren de manera homogénea el área del acuífero dentro de un periodo de cuatro años (agosto del primer año a julio del quinto año); estas posiciones corresponden a nodos de la discretización mediante elemento finito del modelo. La información de los dos primeros años se empleó para la selección del programa óptimo de monitoreo (agosto del primer año a julio del tercer año) y la información de los dos años siguientes (agosto del tercer año a julio del quinto año) se usó para la evaluación de su desempeño en el futuro.

Para el análisis geoestadístico espacio-temporal únicamente fueron empleadas simulaciones de la carga hidráulica en 38 posiciones que fueron consideradas como pozos de la red de monitoreo (círculos de la Fig. 6.1); siguiendo la idea de emular un caso real, se consideró que esos pozos sólo habían sido monitoreados en agosto, septiembre, octubre, y diciembre, del primer año; abril, mayo, y octubre, del segundo año; y mayo del tercer año. Para la evaluación de la red en total se emplearon 304 datos de la carga hidráulica.

5.2 Objetivo del monitoreo

Se planteó como objetivo del diseño estimar adecuadamente las variaciones mensuales de la carga hidráulica en el acuífero. Con fines de comparación, se definieron las mismas posiciones y tiempos de estimación del ejemplo del capítulo cinco. Las 109 posiciones mostradas en la Fig. 6.1 son de estimación y 38 de ellas (círculos) coinciden con los pozos que conforman la red de monitoreo. Un posible tiempo de estimación fue seleccionado para cada uno de los meses dentro del periodo agosto del primer año a julio del tercer año. De esta manera, 109 puntos y 24 tiempos fueron seleccionados para conformar la malla de estimación espacio-temporal; al contar la red con 38 pozos, se tienen 912 posibles posiciones de monitoreo espacio-temporal. La evaluación se dividió en dos fases: 1. Seleccionar un programa de monitoreo óptimo para la red de monitoreo y 2. Evaluar su desempeño en el futuro.

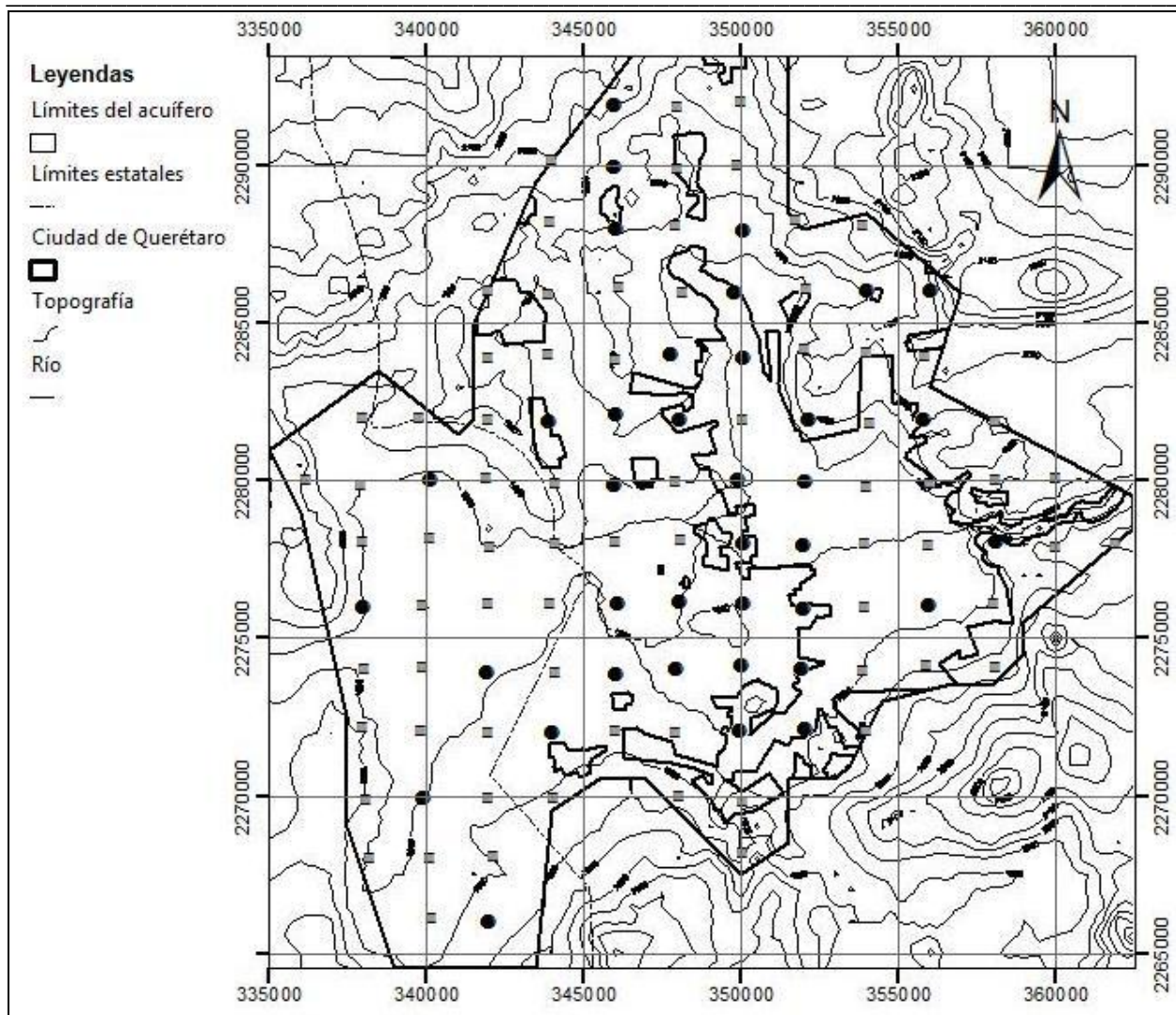


Fig. 6.1 Área de estudio y localización de los puntos de monitoreo/estimación (círculos) y puntos únicamente de estimación (cuadros)

6.3 Análisis geostatístico espacio-temporal de la carga hidráulica

El análisis geostatístico se basó en la construcción del semivariograma espacio-temporal muestral con los 304 datos piezométricos señalados anteriormente (Fig. 6.2). Para realizar el análisis, se usó la rutina GAMV del GSLIB (Deutsch *et al.*, 1997), modificada para modelar semivariogramas espacio-temporales, desarrollada por De Cesare *et al.* (2002). Se usaron 20 intervalos espaciales, con 1312.15 m de separación, y 20 intervalos temporales con 0.085 de año como separación.

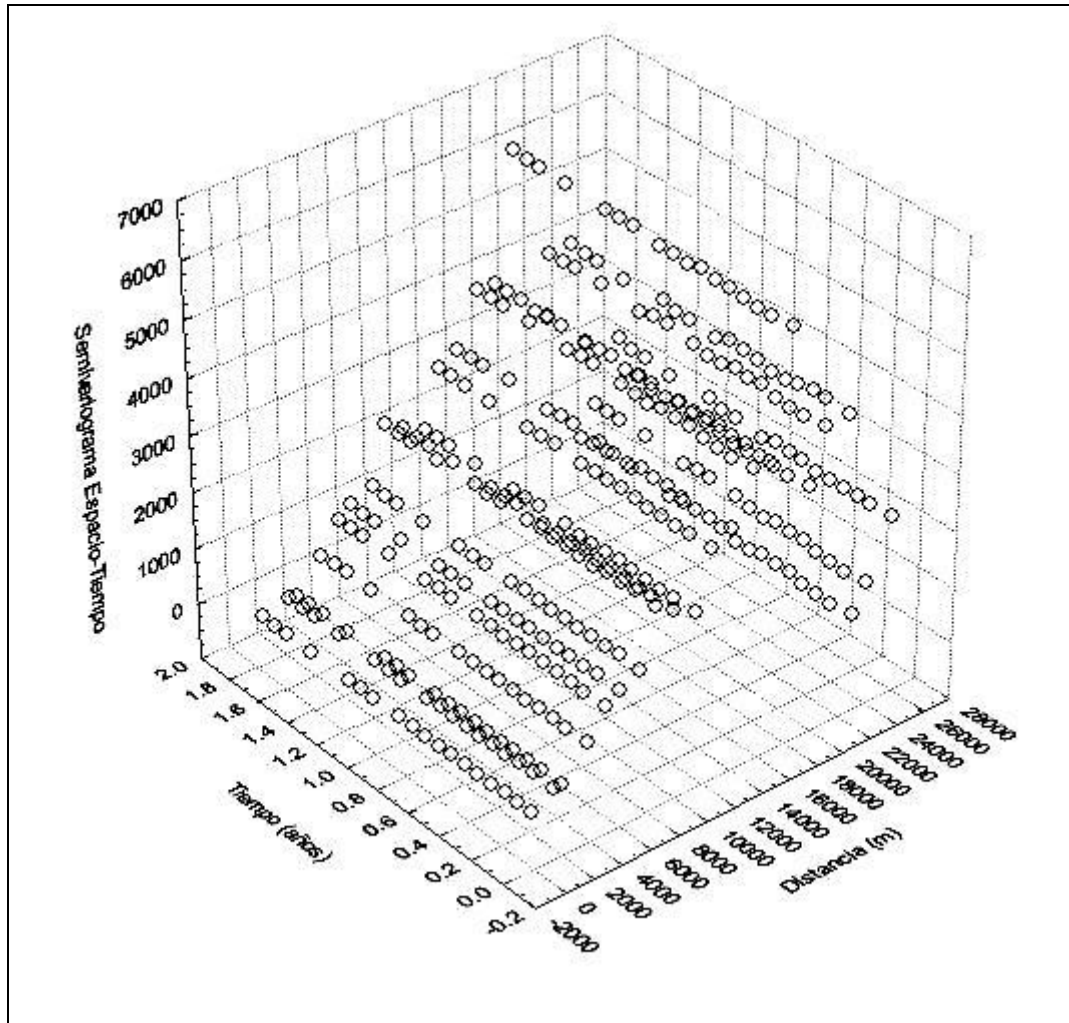


Fig. 6.2 Semivariograma espacio-temporal muestral de la carga hidráulica

Para analizar mejor el comportamiento del semivariograma espacio-temporal muestral, presentamos su proyección sobre el plano semivariograma-intervalos espaciales (Fig. 6.3) y sobre el plano semivariograma-intervalos temporales (Fig. 6.4). La proyección espacial muestra que el semivariograma aumenta hasta los 18,369.68 metros y a partir de ese punto decrece, fluctuando y alcanzando su mínimo valor a los 23,618.16 m, este comportamiento refleja los huecos en la información en espacio-tiempo que por lo que el semivariograma no queda tan bien definido como el obtenido en el capítulo cinco. La proyección temporal muestra un comportamiento homogéneo donde se advierte que la componente de varianza temporal tiene poca influencia en la forma del semivariograma espacio-temporal muestral. No se advirtió la presencia de tendencia en los datos.

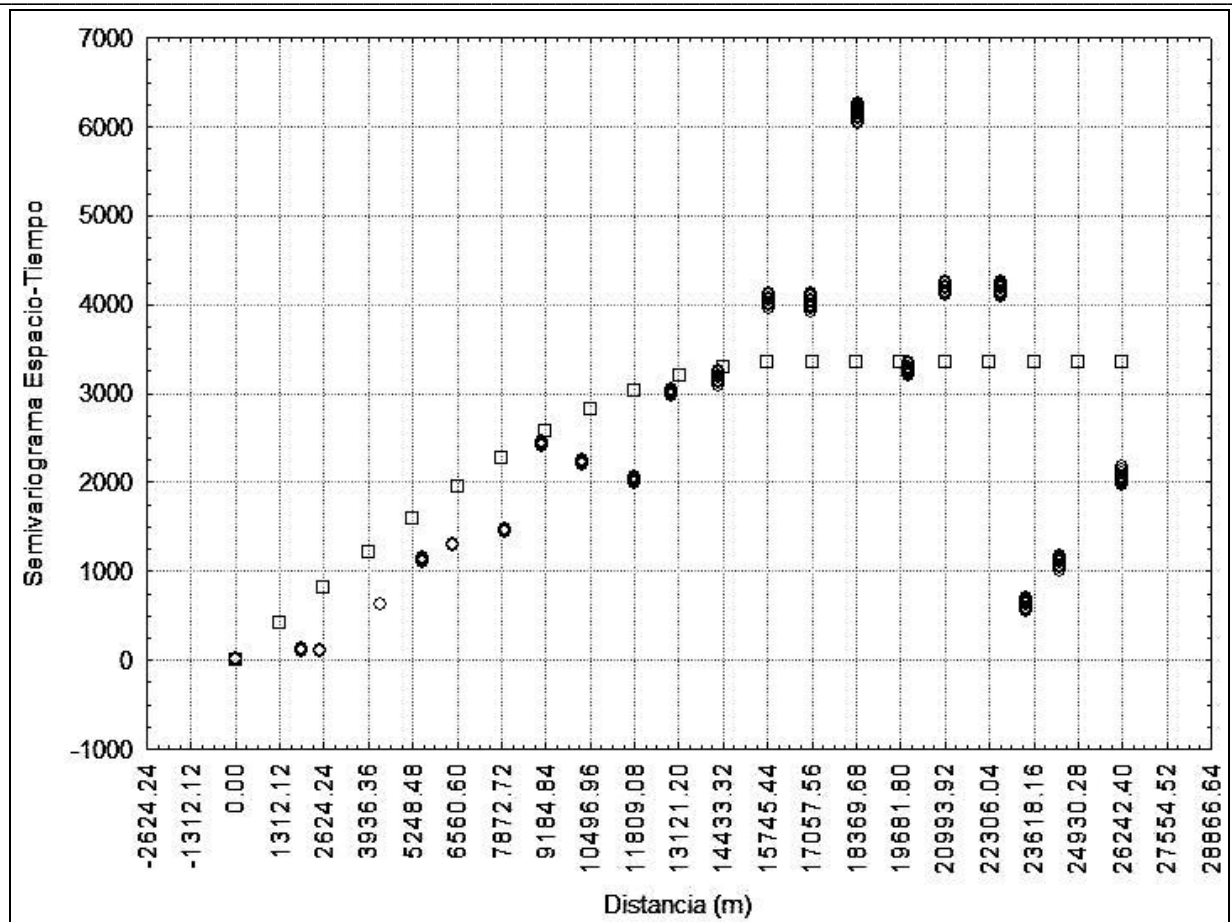


Fig. 6.3 Proyección del semivariograma espacio-temporal muestral de la carga hidráulica sobre el plano semivariograma-intervalos espaciales (círculos) y modelo de semivariograma ajustado (cuadros)

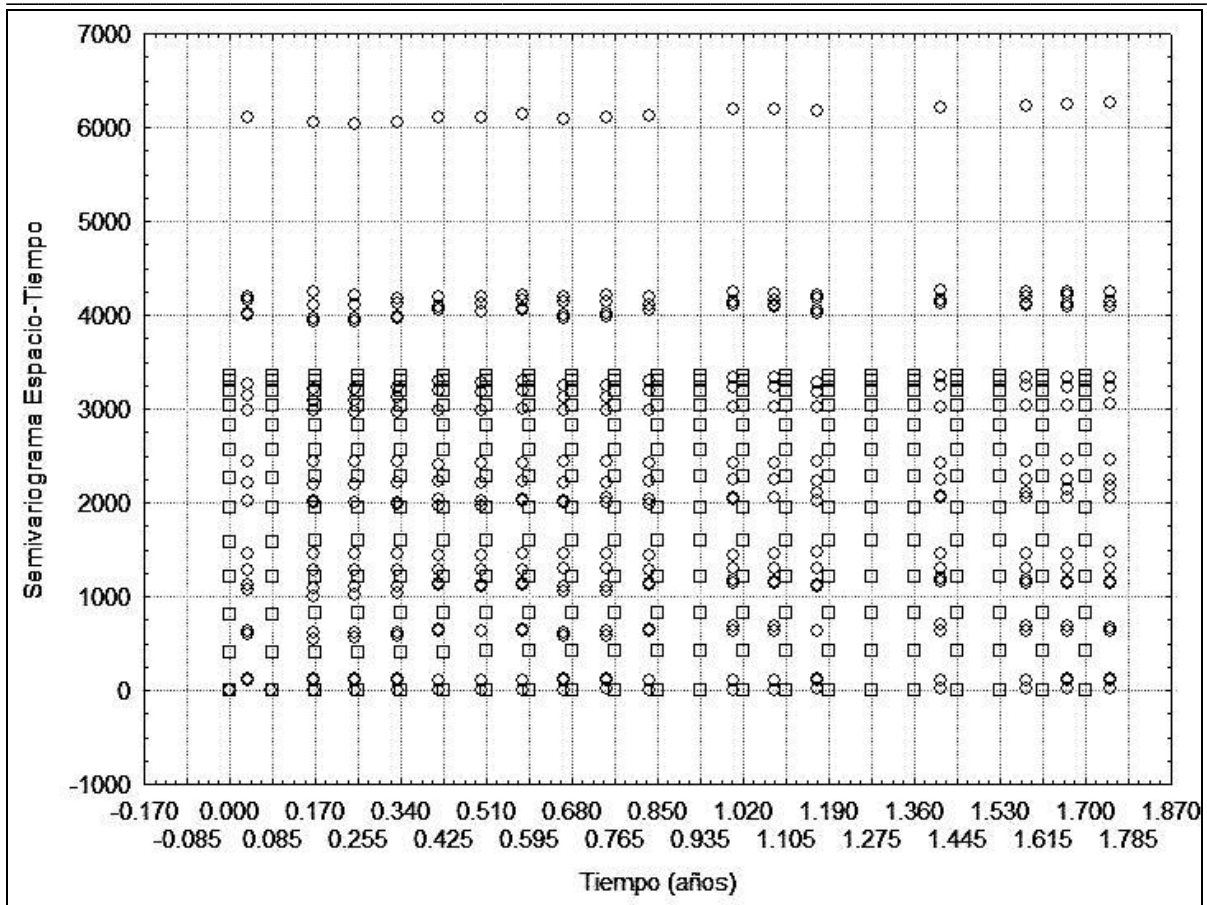


Fig. 6.4 Proyección del semivariograma espacio-temporal muestral de la carga hidráulica sobre el plano semivariograma-intervalos temporales (círculos) y modelo de semivariograma ajustado (cuadros)

Ajuste del modelo de semivariograma

Siguiendo a De Iaco *et al.* (2002), un modelo producto-suma se ajustó manualmente al semivariograma espacio-temporal muestral (esto quiere decir que diferentes modelos fueron propuestos “a ojo” y evaluados para seleccionar aquél con los mejores resultados en la validación cruzada para el periodo agosto del primer año a julio del tercer año). Los parámetros del modelo seleccionado se presentan en la Tabla 6.1. Las diferencias más significativas con respecto del modelo ajustado en el capítulo cinco, son la reducción en un 25% del alcance de la componente espacial y el aumento en más del 300% del valor de la meseta temporal, esto sugiere que al emplear un menor número de datos en el análisis geoestadístico se obtiene un modelo más limitado.

Tabla 6.1 Parámetros del modelo de semivariograma espacio-temporal de la carga hidráulica

Modelo	Componente	Pepita	Meseta	Alcance
Esférico	Espacial	0	3350	16000
Esférico	Temporal	0	10.4	1.615
-	Global	-	3350	-

Los resultados de la validación cruzada para el periodo agosto del primer año a julio del tercer año se presentan en la Tabla 6.2; como puede verse, los errores máximos y mínimos son pequeños, considerando la alta variabilidad de los datos.

Tabla 6.2 Resultados de la validación cruzada para el periodo de agosto del primer año a julio del tercer año

Error mín (m)	Error máx (m)	Error medio (m)	ECM (m²)	RECM (m)	ECME
-3.82	3.69	0.01	0.64	0.80	1.00

6.4 Optimización de la red de monitoreo

El filtro de Kalman estático necesita de estimaciones a priori del vector de estado y de la matriz de covarianza del error en la estimación. La estimación a priori del vector de estado se integró para cada año, con el promedio de las simulaciones de la carga hidráulica correspondientes consideradas en el análisis geoestadístico. La matriz de covarianza espacio-temporal a priori fue calculada del modelo de semivariograma espacio-temporal, aplicando la ecuación 3.15; se construyó para 109 posiciones y 24 tiempos, de tal manera que su dimensión fue de 2616 columnas (109×24) por 2616 renglones. Por otro lado, se supone que no hay error de medición en los datos generados por la red de monitoreo.

Orden de prioridad

Cuando se aplica el método de optimización de redes de monitoreo propuesto en esta tesis, el orden de selección es un indicador de la importancia de los datos obtenidos en esas posiciones, en la reducción de la varianza total. Esto puede verse claramente en la Fig. 6.5, la varianza total se reduce grandemente al elegir las primeras posiciones de monitoreo espacio-temporal, pero conforme se incrementa el número de posiciones seleccionadas, la reducción de la varianza total decrece.

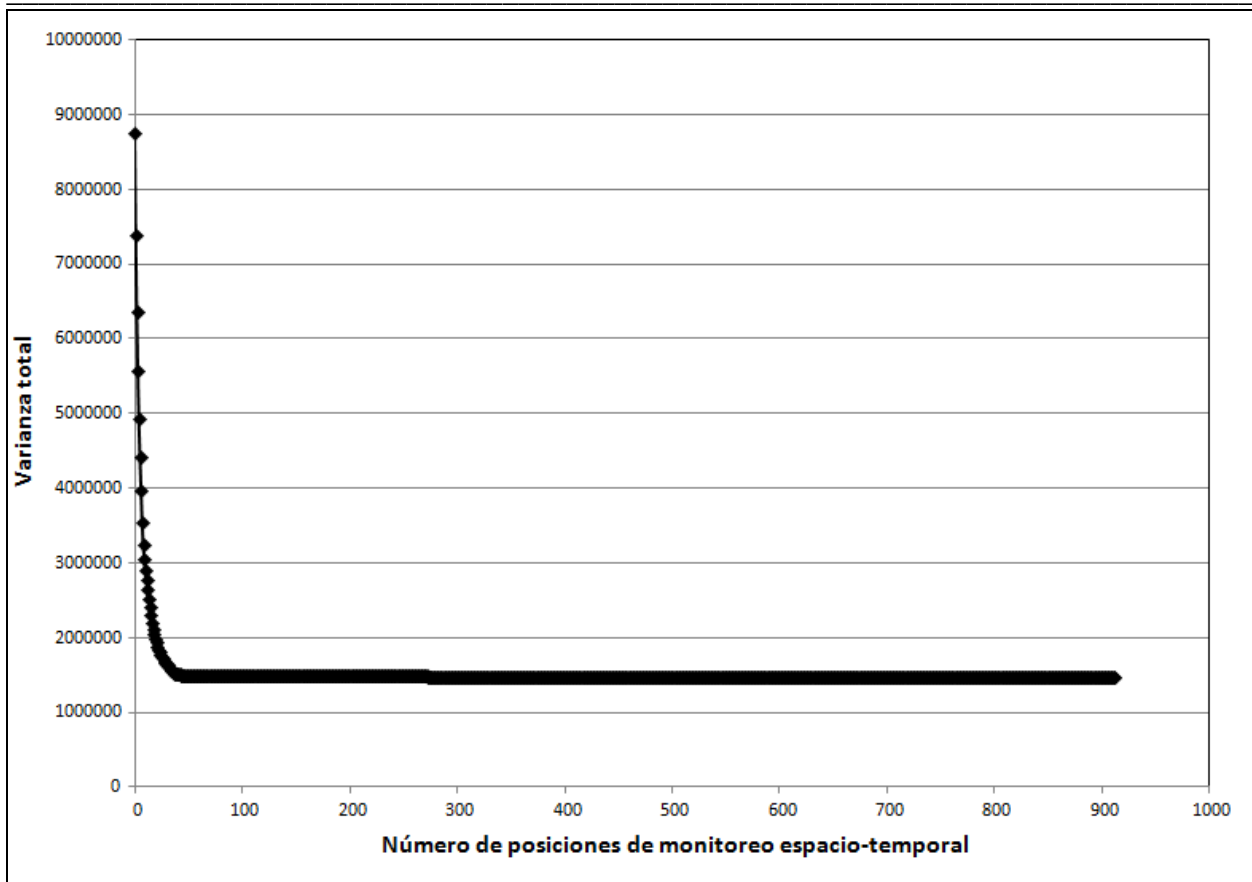


Fig. 6.5 Varianza Total vs Número de posiciones de monitoreo espacio-temporal

6.5 Determinación del número total de posiciones de monitoreo

Puede verse en la Fig. 6.5, después de monitorear en aproximadamente las 30 posiciones con la mayor prioridad, la varianza total permanece prácticamente constante, alrededor de 1'480,000 m². Esto sugiere que las 2586 posiciones restantes pudieran ser redundantes, esto es, proveen poca información y no contribuyen mucho en la reducción de la varianza total.

Para definir el número de posiciones de la red de monitoreo espacio-temporal, se utilizó el mismo criterio que en el capítulo cinco. Los valores obtenidos de la raíz cuadrada (Varianza total/Posiciones de estimación espacio-temporal) contra el número de posiciones de monitoreo espacio-temporal se muestran en la Fig. 6.6.

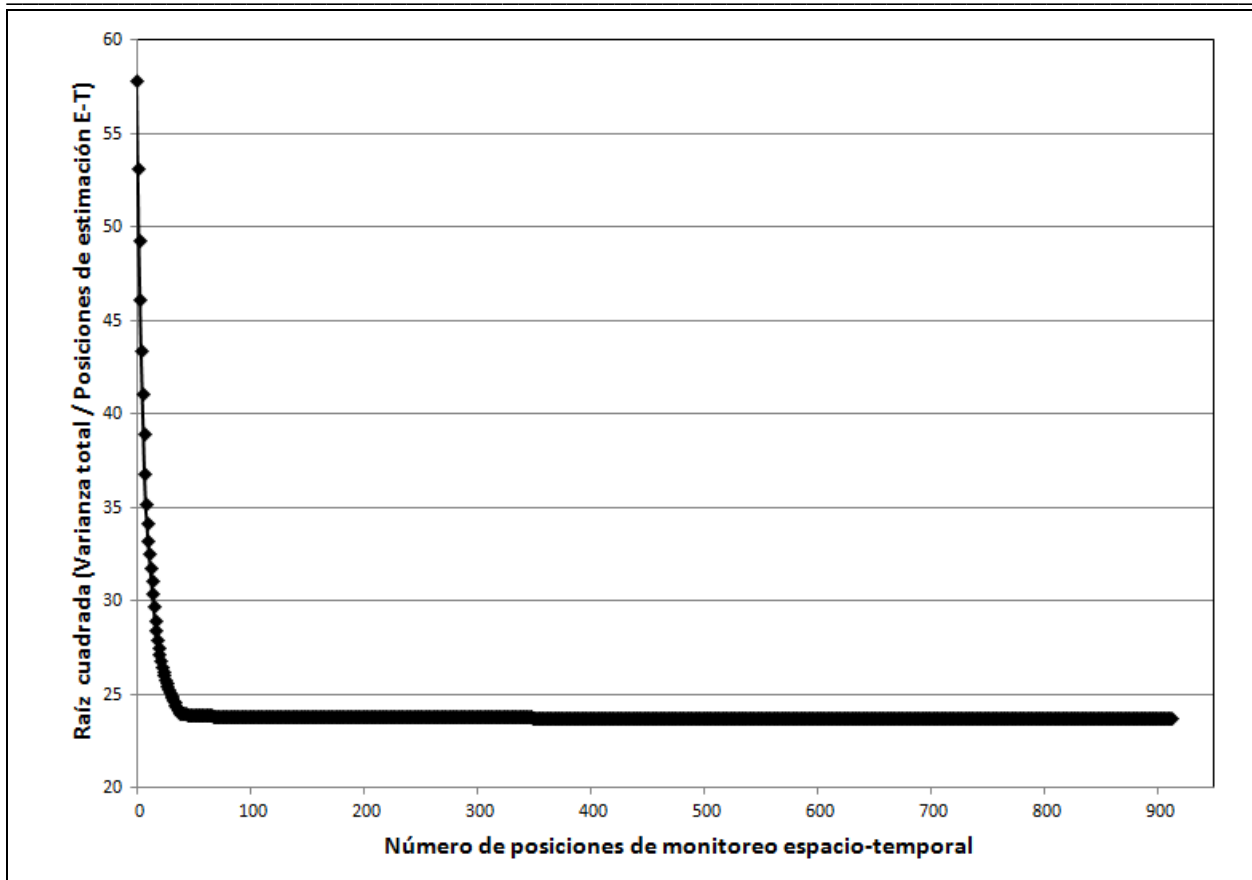


Fig. 6.6 Raíz cuadrada (Varianza total/Posiciones de estimación espacio-temporal) vs Número de posiciones de monitoreo espacio-temporal

El 99% del valor de la MRP se alcanza cuando 37 posiciones de monitoreo espacio-temporal se han seleccionado y una incertidumbre en la estimación promedio de 24.12 m es esperada en la malla de estimación espacio-temporal. Este valor es alto pero es muy cercano al obtenido usando todas las posiciones de monitoreo espacio-temporal disponibles (23.78 m). Debe señalarse que antes de elegir alguna posición de monitoreo se tiene una raíz cuadrada de la varianza promedio de 57.88 m, esto significa que las primeras 37 posiciones reducen este valor en 33.76 m, mientras que las últimas 875 posiciones únicamente lo reducen 0.34 m más. Estos resultados muestran que el método propuesto es exitoso en propagar la información en el espacio y en el tiempo. Sin embargo, los valores obtenidos se encuentran muy por encima de lo deseado en un monitoreo de la carga hidráulica, esto se debe a que la red de monitoreo no se encuentra adecuadamente distribuida en el espacio y por lo tanto no permite reducir lo suficiente la incertidumbre en algunas zonas; es clara entonces la necesidad de monitorear posiciones adicionales.

6.6 Programa de monitoreo óptimo

En este ejemplo y de acuerdo a lo expuesto anteriormente, 37 posiciones fueron seleccionadas para constituir el programa de monitoreo óptimo de la red evaluada. Todas las posiciones seleccionadas por el método propuesto para el diseño óptimo de redes de monitoreo espacio-temporal de la carga hidráulica corresponden al primer mes del periodo de diseño; esto se explica

por dos razones: 1) en la reducción de la varianza total para el primer mes no participan ni monitoreos previos (porque aquí inicia el monitoreo), ni tampoco participan los monitoreos posteriores (al estar definida la función objetivo para una estimación en tiempo real), y 2) la varianza temporal es muy pequeña con respecto a la espacial (Tabla 6.1); esto quiere decir que al monitorear 37 de las 38 posiciones disponibles en el primer mes obtenemos el patrón de la distribución espacial de la carga hidráulica que nos puede ofrecer la red de monitoreo (en el periodo de diseño).

La Fig. 6.7 muestra el orden de prioridad (arriba) y el nombre (abajo) para cada pozo. Como puede verse, el pozo 1, localizado en la parte central del acuífero, es el primero en ser elegido, esto se explica porque esa posición influye sobre el mayor número de nodos de la malla de estimación.

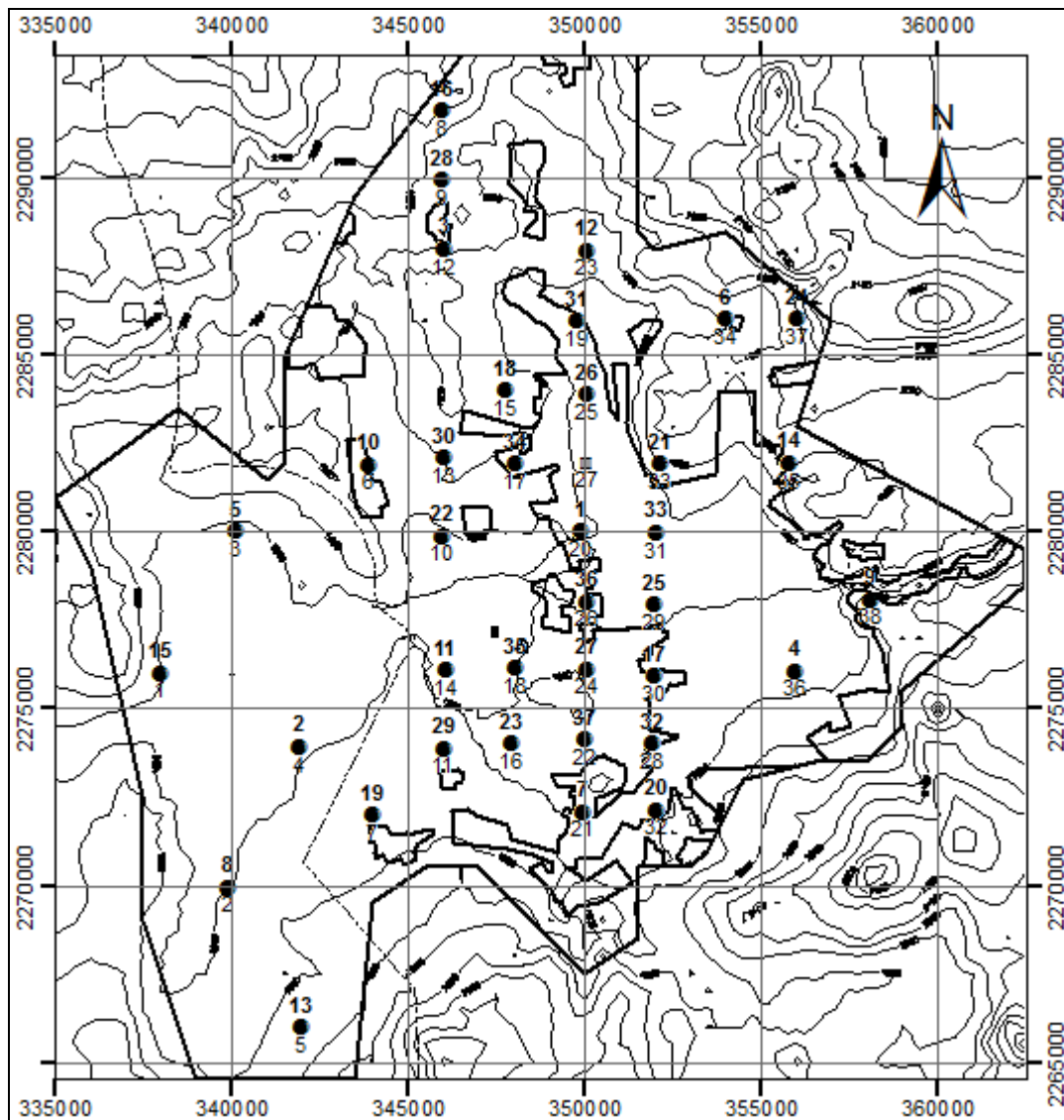


Fig. 6.7 Número de veces de monitoreo para cada pozo en el programa de monitoreo óptimo espacio-temporal (arriba) y nombre del pozo (abajo)

6.7 Estimaciones y varianzas del error en la estimación para el período de diseño

Como una ilustración del nivel de información que proporciona el programa de monitoreo propuesto para el periodo de diseño, se obtuvieron estimaciones de la carga hidráulica y de la varianza del error de estimación sobre las posiciones y tiempos de interés. Éstas se obtuvieron aplicando el filtro de Kalman estático.

Estimación de la carga hidráulica

En las Figuras 6.8, 6.10 y 6.12 se presentan estimaciones en tiempo real de la carga hidráulica, obtenidas para algunos meses con el programa de monitoreo propuesto.

Estas figuras muestran los cambios espaciales y temporales de la carga hidráulica estimada en el acuífero. Las principales zonas de recarga corresponden a donde han sido monitoreados los niveles máximos de carga hidráulica, y en este caso coinciden con los niveles topográficos máximos, al norte, noreste, este y suroeste del acuífero. El flujo del agua subterránea es radial, con una dirección preferencial de flujo noreste-suroeste (niveles máximos en el rango 1880-1900 msnm), y otra proveniente del sur con dirección hacia el centro (niveles máximos en el rango 1800-1820 msnm); la parte central del acuífero presenta los niveles mínimos con el mayor cono de abatimiento al noroeste de la ciudad (por debajo de los 1730 msnm y alcanzando en algunos meses valores mínimos de 1715 msnm). En el curso del tiempo, los niveles mayores a 1740 msnm se mantienen prácticamente invariantes pero los principales abatimientos ubicados al noroeste de la ciudad y al oeste del acuífero aumentan.

Varianzas del error en la estimación

En las Figuras 6.9, 6.11 y 6.13 se presentan las varianzas del error en la estimación obtenidas en tiempo real para algunos meses, con la red de monitoreo y el programa de monitoreo propuesto.

Para todo el periodo de monitoreo, los valores mínimos de la varianza se sitúan alrededor de las posiciones de monitoreo (entre 0 y 50 m²), y se incrementan en las zonas pobremente cubiertas por las posiciones de monitoreo de la red, alcanzando valores mayores a 1600 m² en la periferia del acuífero. Durante el periodo de monitoreo, se puede ver que las varianzas no se incrementan en el curso del tiempo, debido a la baja varianza temporal en comparación con la espacial, pero podemos apreciar que cerca del 75% del área de estudio presenta varianzas entre 50 y 1000 m², que nos hablan del gran nivel de incertidumbre en la estimación obtenido con el arreglo actual de la red de monitoreo. A diferencia de la red de monitoreo con una buena cobertura espacio-temporal, como la mostrada en el capítulo cinco, la red analizada en este capítulo, que simula un caso real, es muy deficiente en proveer una estimación espacio-temporal de la carga hidráulica con baja incertidumbre. Estos mapas de varianza junto con criterios prácticos e hidrogeológicos pueden ser muy útiles en la elección de nuevas posiciones de monitoreo en las zonas con las máximas varianzas.

EVALUACIÓN DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON BAJO NIVEL DE INFORMACIÓN ESPACIO-TEMPORAL

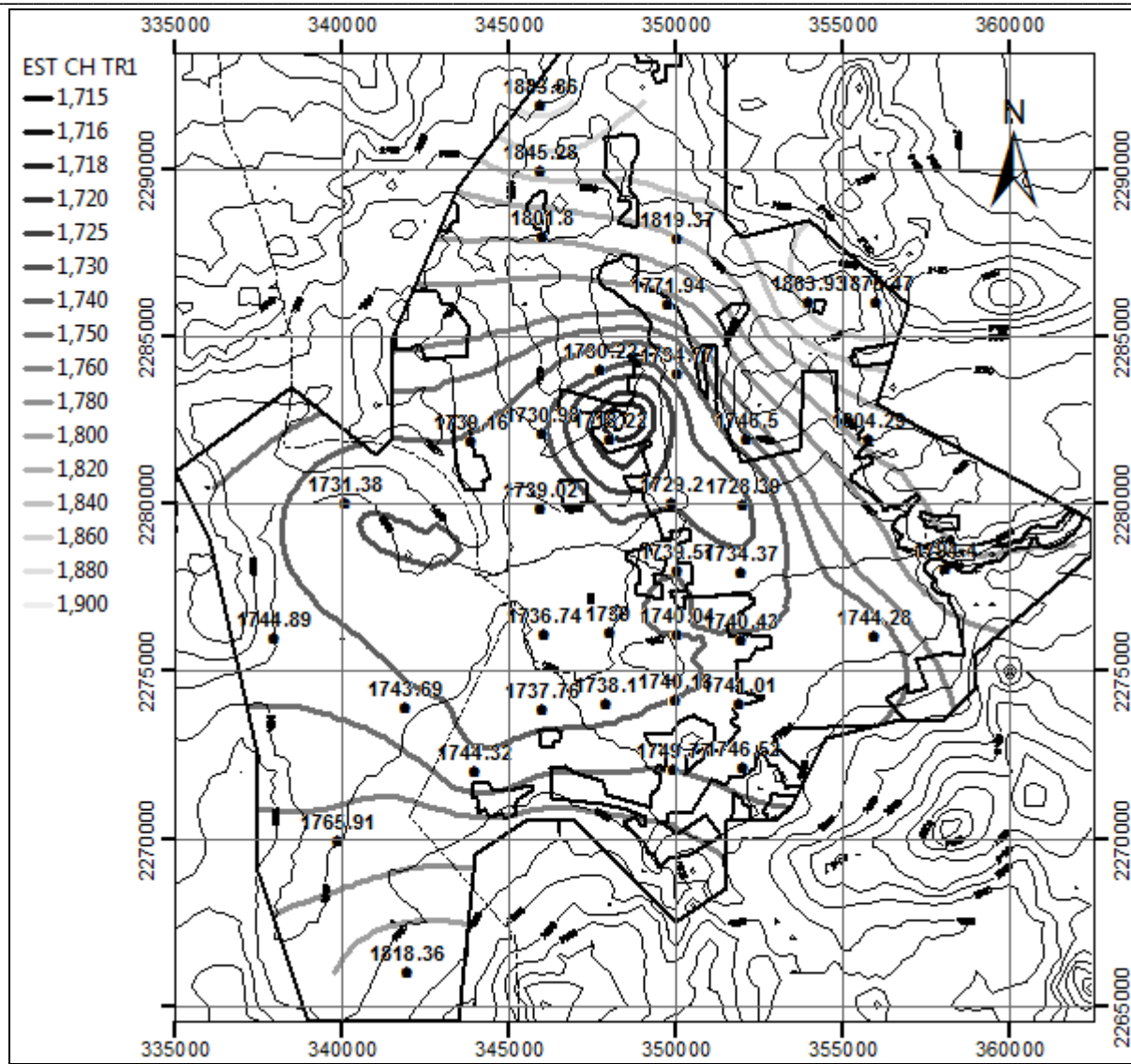


Fig. 6.8 Estimación de la carga hidráulica para agosto del primer año

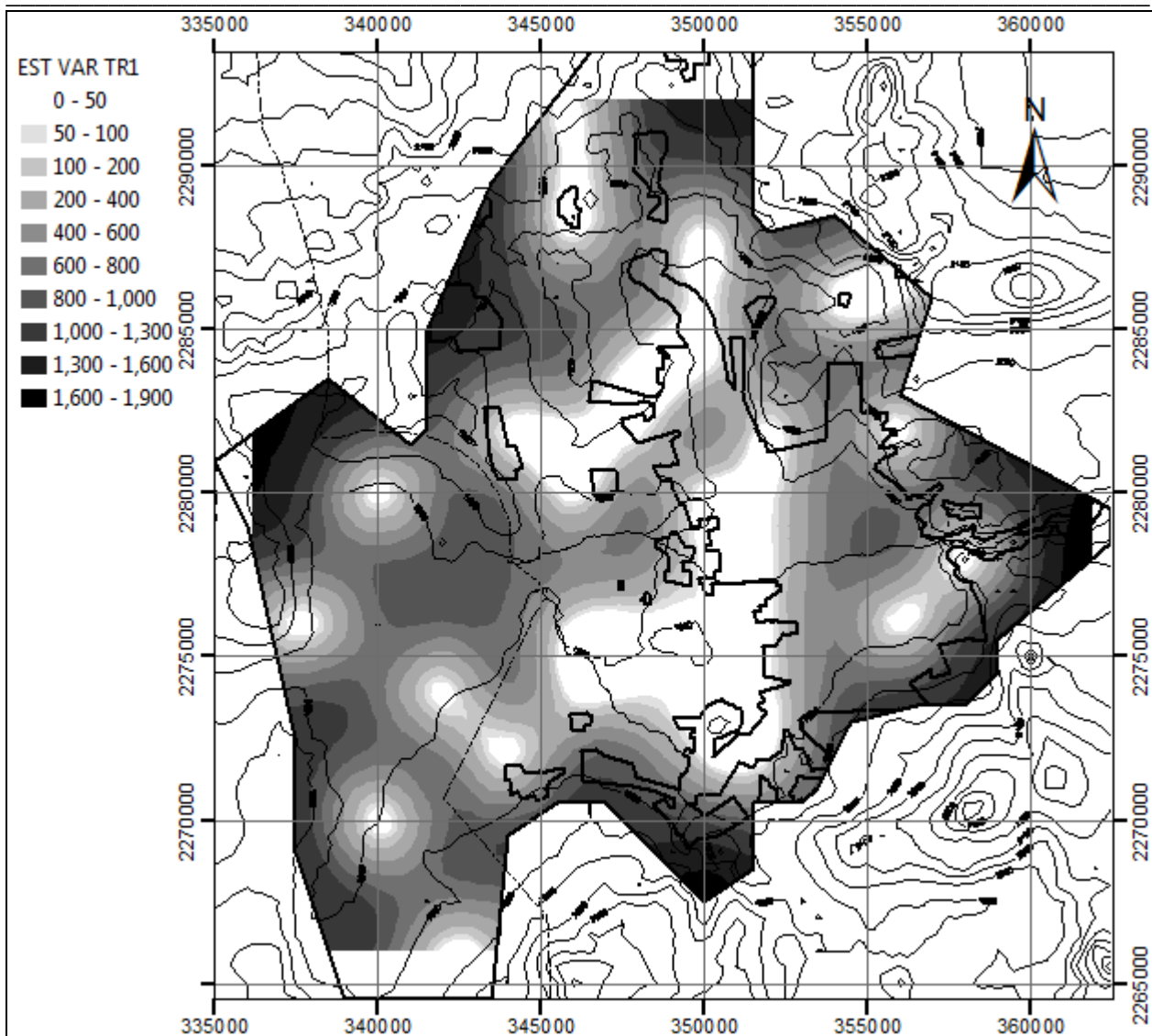


Fig. 6.9 Varianzas del error en la estimación para agosto del primer año

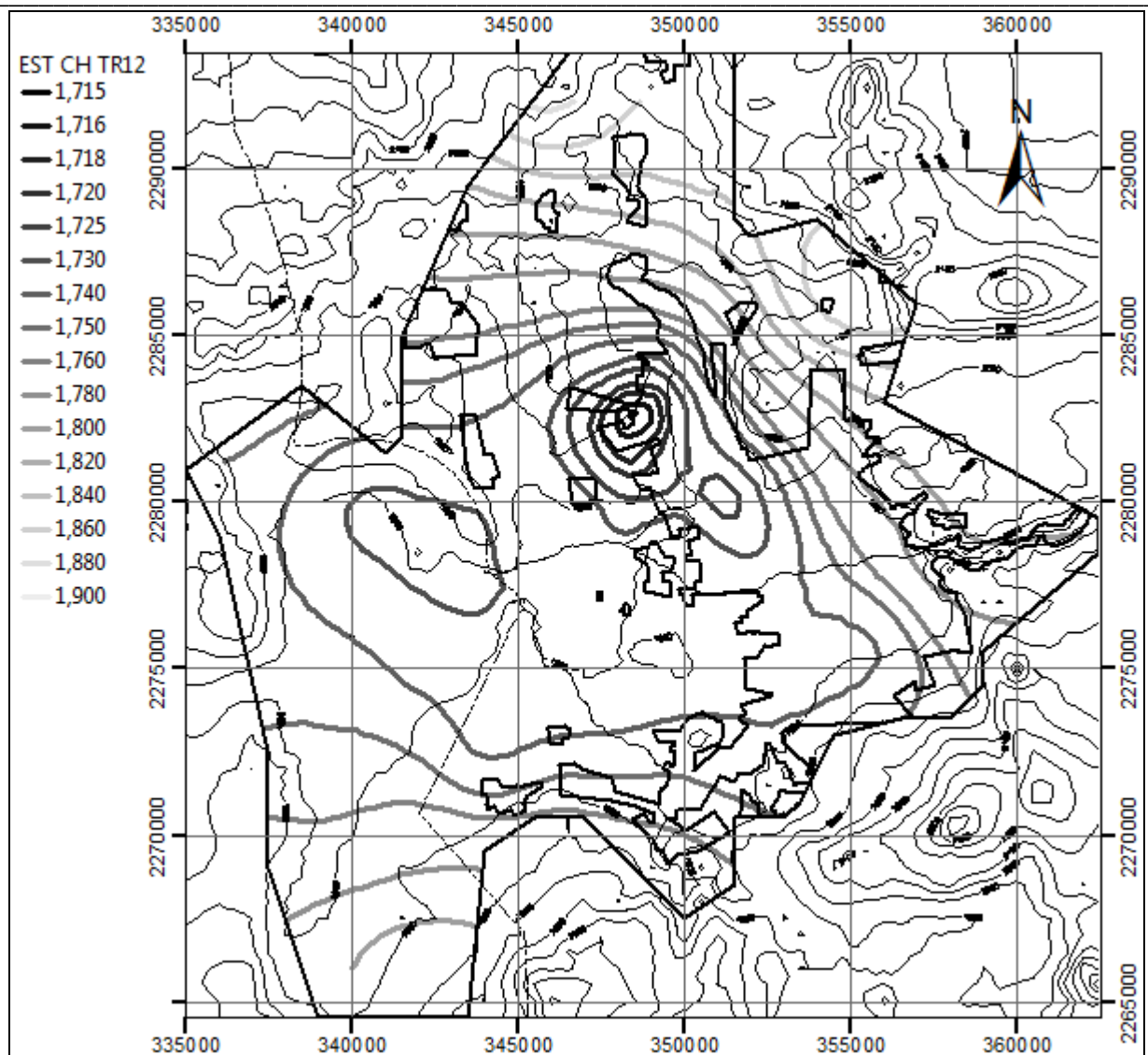


Fig. 6.10 Estimación de la carga hidráulica para julio del segundo año

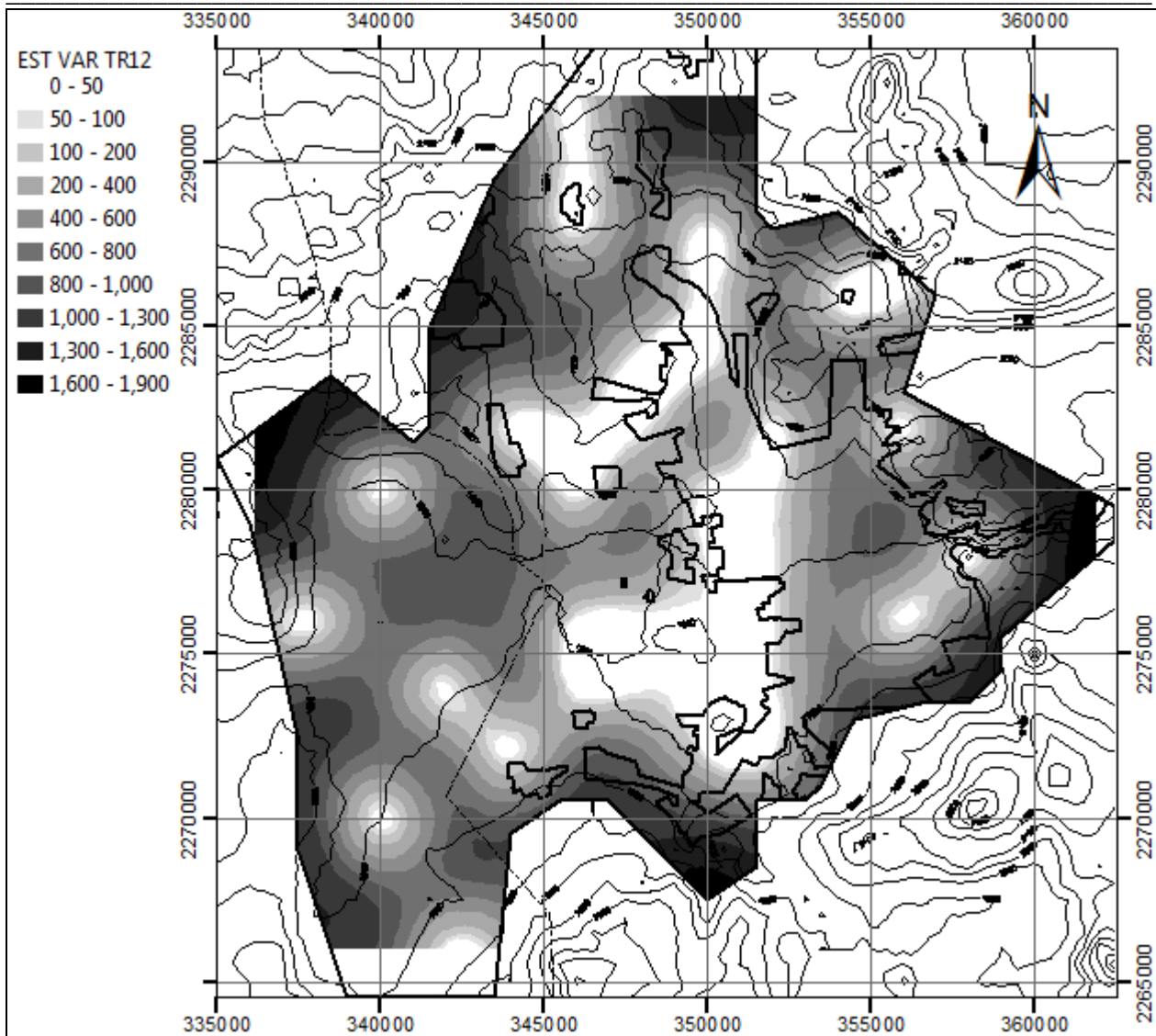


Fig. 6.11 Varianzas del error en la estimación para julio del segundo año

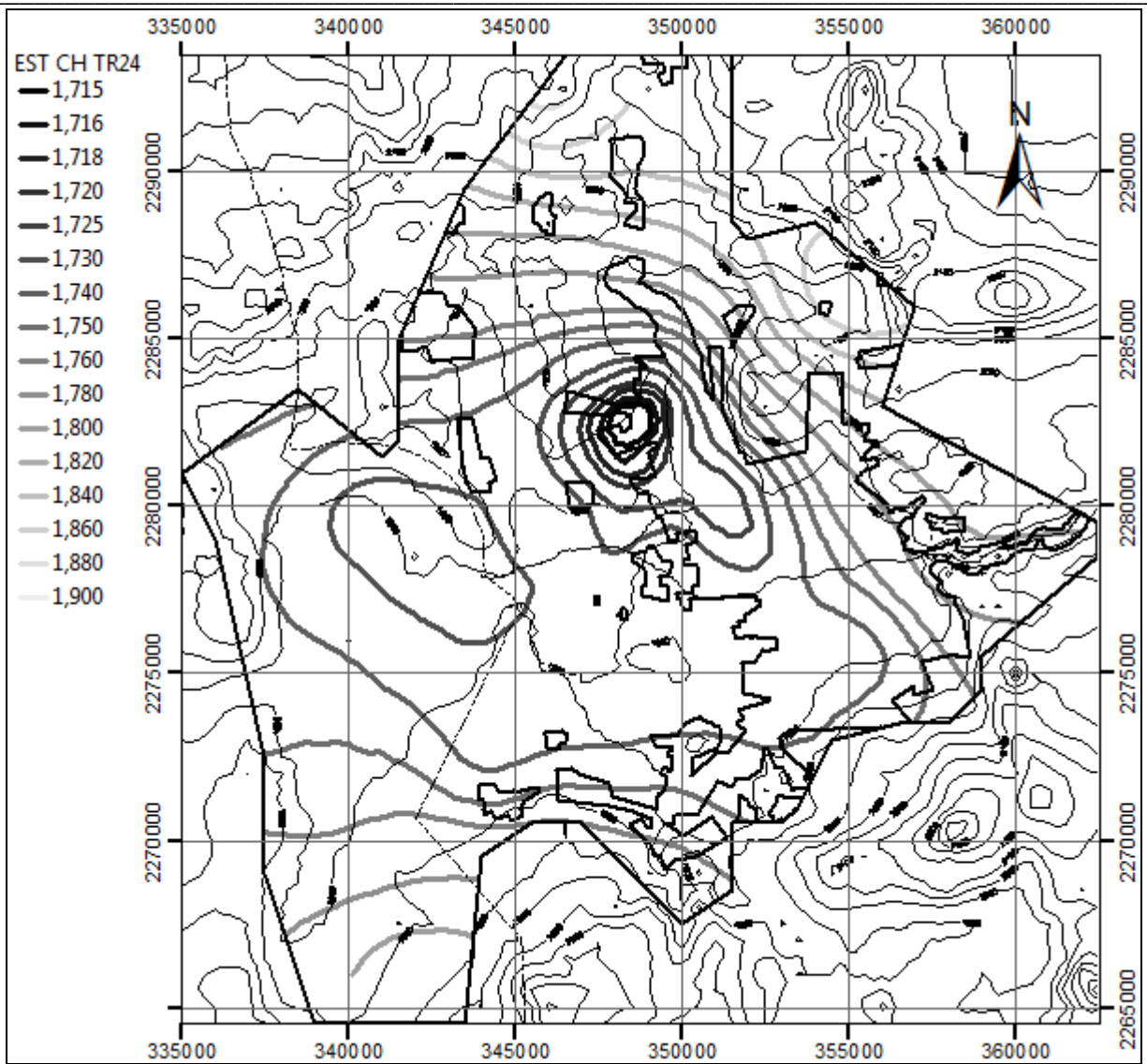


Fig. 6.12 Estimación de la carga hidráulica para julio del tercer año

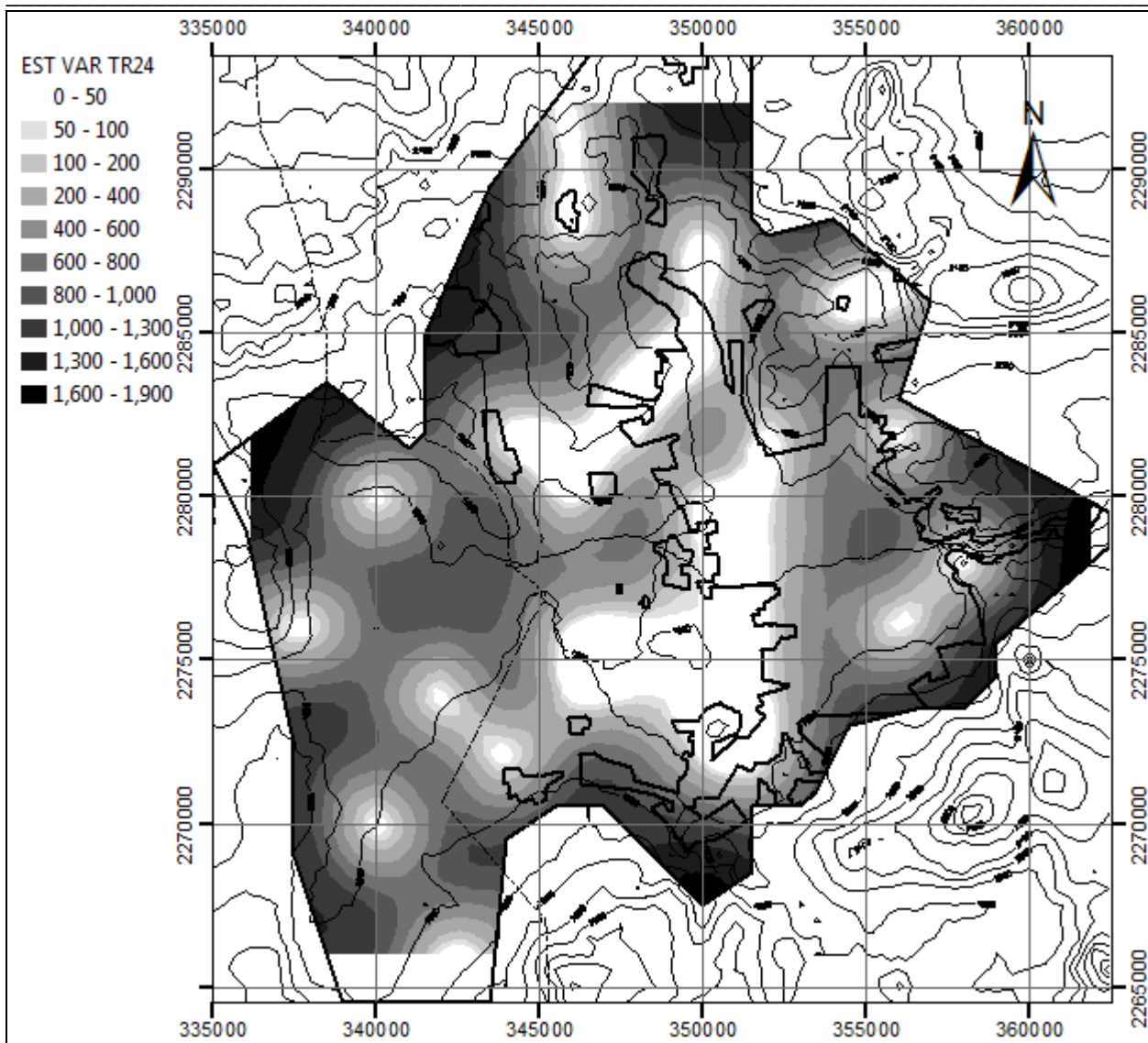


Fig. 6.13 Varianzas del error en la estimación para julio del tercer año

6.8 Evaluación del desempeño de la red de monitoreo en el futuro

Como en este caso se tienen datos “futuros” de la carga hidráulica (simulaciones del modelo numérico de flujo) en todas las posiciones de monitoreo/estimación espacio-temporal, entonces podemos evaluar el desempeño de la red de monitoreo en el futuro. Para la selección del programa de monitoreo óptimo se utilizó un periodo de dos años (agosto del primer año a julio del tercer año), por lo que se aplicará el mismo programa para un periodo igual en el futuro (agosto del tercer año a julio del quinto año), considerando que las condiciones de funcionamiento del acuífero se mantienen. Para verificar el desempeño de la red, se compararon las estimaciones obtenidas con la red de monitoreo propuesta contra los datos de la carga hidráulica en las posiciones de estimación espacio-temporal (obtenidos del modelo numérico de flujo), a la diferencia entre la estimación y el dato se le llamó error en la estimación.

Como puede verse en la Tabla 6.3, y como era de esperarse, el desempeño de la red de monitoreo en el futuro tampoco es adecuado; de hecho, los errores en las estimaciones del futuro

son mayores con respecto a los que se presentan en el periodo de diseño, debido a que conforme se avanza en el tiempo, el comportamiento en algunas posiciones de monitoreo, es contrario a lo que sucede en los nodos de estimación vecinos, esto se explicará un poco más a detalle a continuación. Los valores del ECME de la Tabla 6.3 demuestran que los errores cuadráticos son mayores a las varianzas calculadas, pero no se obtienen valores tan altos como en el diseño del capítulo cinco (Tabla 5.7); esto se debe a que las varianzas no se reducen mucho en ciertas zonas, por la alta incertidumbre en las estimaciones, con lo que se confirma la insuficiencia de la red de monitoreo para estimar la variabilidad espacial de la carga hidráulica en el acuífero.

Tabla 6.3 Desempeño de la red de monitoreo

Período de monitoreo (años)	Posiciones de monitoreo espacio-temporal	Error medio (m)	Error mínimo (m)	Error máximo (m)	RECM (m)	ECME
1-4	74	-1.49	-51.45	40.44	12.20	1.89
1-2	37	-1.92	-44.22	25.38	11.38	2.03
3-4	37	-1.06	-51.45	40.44	12.97	1.75
1-4	37	-1.91	-45.10	25.45	12.40	1.76

Como puede verse en la Tabla 6.3, para un período de cuatro años, los errores máximo y mínimo en la estimación de la carga hidráulica se obtienen cuando se realiza una segunda campaña de monitoreo en agosto del tercer año. Para explicar este fenómeno, hemos elegido el nodo 63 (donde más se incrementan los errores al monitorear en el futuro). En la Fig. 6.14, puede verse que los pozos 6 y 12 contribuyen fuertemente en las estimaciones del nodo 63 (localizado en la parte noroeste del acuífero).

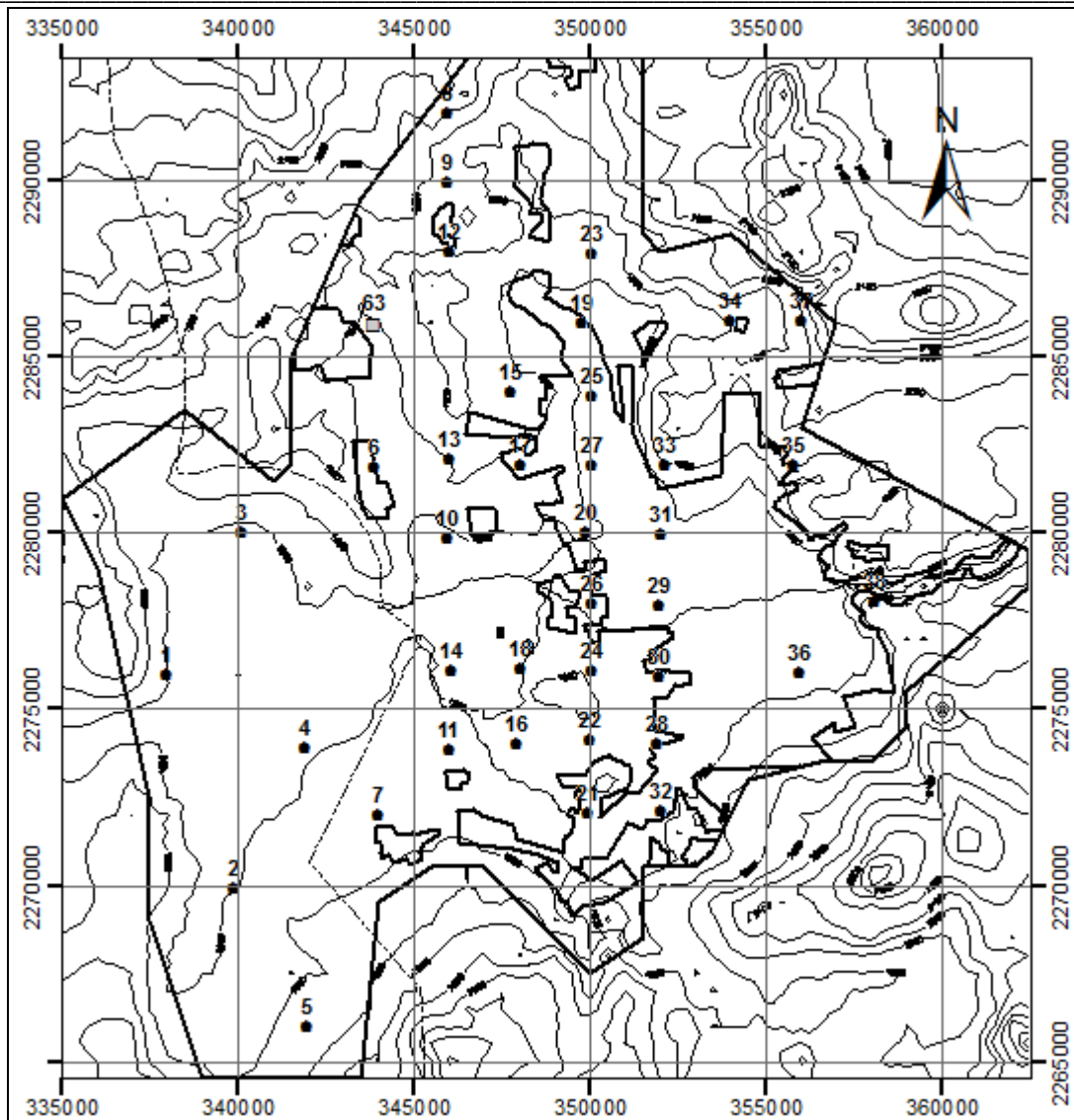


Fig. 6.14 Red de monitoreo (círculos) y nodo de estimación con los errores más altos (cuadro)

En la Fig. 6.15 puede verse que al usar las primeras 37 posiciones de monitoreo espacio-temporal, la estimación del primer mes en el nodo 63 es aproximadamente el promedio de los datos en los pozos 6 y 12 (que son los que más influencia tienen sobre esta estimación), y después se mantiene una tendencia de ligero descenso lineal de la carga hidráulica, este descenso es aproximadamente paralelo al comportamiento real de la carga hidráulica en el nodo 63. Sin embargo, al usar 74 posiciones de monitoreo espacio-temporal, la estimación del primer mes en el futuro para el nodo 63 aumenta sustancialmente, volviendo a ser aproximadamente el promedio de los datos en los pozos 6 y 12, alejándose este valor del comportamiento en el nodo 63, debido a la significativa recuperación en el pozo 12, contraria al comportamiento descendente en el nodo 63.

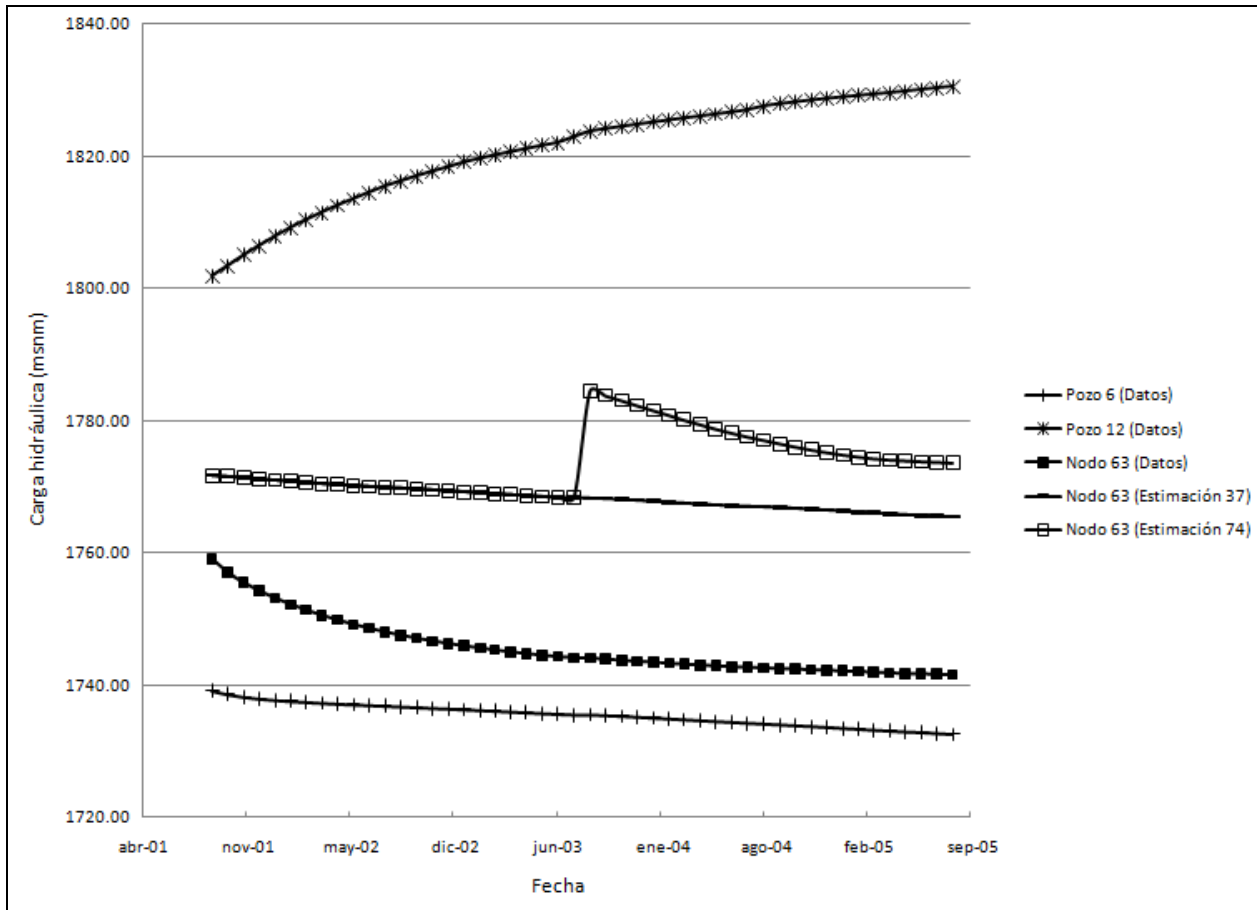


Fig. 6.15 Datos en los pozos 6 y 12, y datos y estimaciones en el nodo 63

Los resultados previos muestran que la variabilidad espacial de la carga hidráulica en la zona noroeste no es capturada adecuadamente por el arreglo de la red de monitoreo, y por lo tanto otra(s) posición(es) se necesita(n) para obtener estimaciones más precisas en el nodo 63. Lo anterior demuestra que un exhaustivo monitoreo en el tiempo es inútil si la variabilidad espacial no es capturada adecuadamente por la red de monitoreo. En este caso, al monitorear en las 37 posiciones antes mencionadas (todas ellas correspondientes a la primera fecha de monitoreo) se obtiene prácticamente el mismo nivel de información de la evolución espacio-temporal de la carga hidráulica en el acuífero que al monitorear en la red de monitoreo mensualmente durante un periodo de cuatro años, esto ocurre porque las estimaciones en algunas zonas se van deteriorando al emplear pozos con un comportamiento contrario a las posiciones donde se desea estimar, se trata por lo tanto de una red de monitoreo pobremente distribuida en el espacio.

Estimaciones futuras de la carga hidráulica

En las Figuras 6.16 y 6.18 se presentan estimaciones futuras en tiempo real de la carga hidráulica, obtenidas para algunos meses empleando los 37 datos monitoreados en agosto del primer año y otros 37 datos monitoreados en las mismas posiciones en agosto del tercer año.

Las principales direcciones de flujo se mantienen para los próximos dos años. La parte central del acuífero continúa con niveles por debajo de 1740 msnm, pero ahora, el mayor cono de abatimiento ubicado al noroeste de la ciudad y el localizado en la parte oeste del acuífero se han expandido y unido en un enorme cono con niveles por debajo de los 1730 msnm.

Varianzas del error en la estimación futuras

En las Figuras 6.17 y 6.19 se presentan las varianzas del error en la estimación obtenidas en tiempo real para algunos meses del futuro, con los datos de la red de monitoreo y su programa de monitoreo propuesto.

Durante el periodo de estimación futura, las varianzas del error en la estimación se comportan de la misma manera que en los dos primeros años de monitoreo.

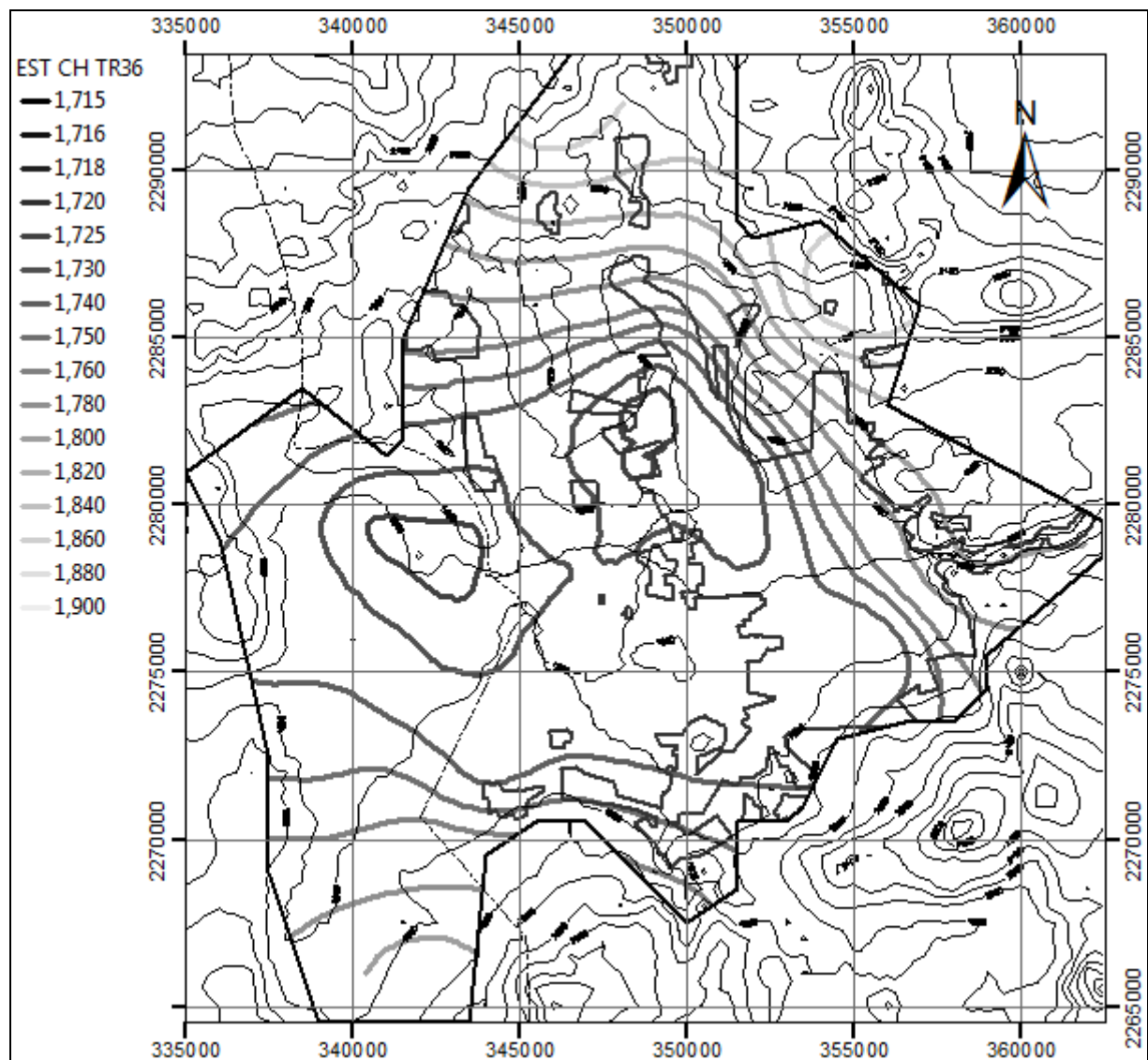


Fig. 6.16 Estimación de la carga hidráulica para julio del cuarto año

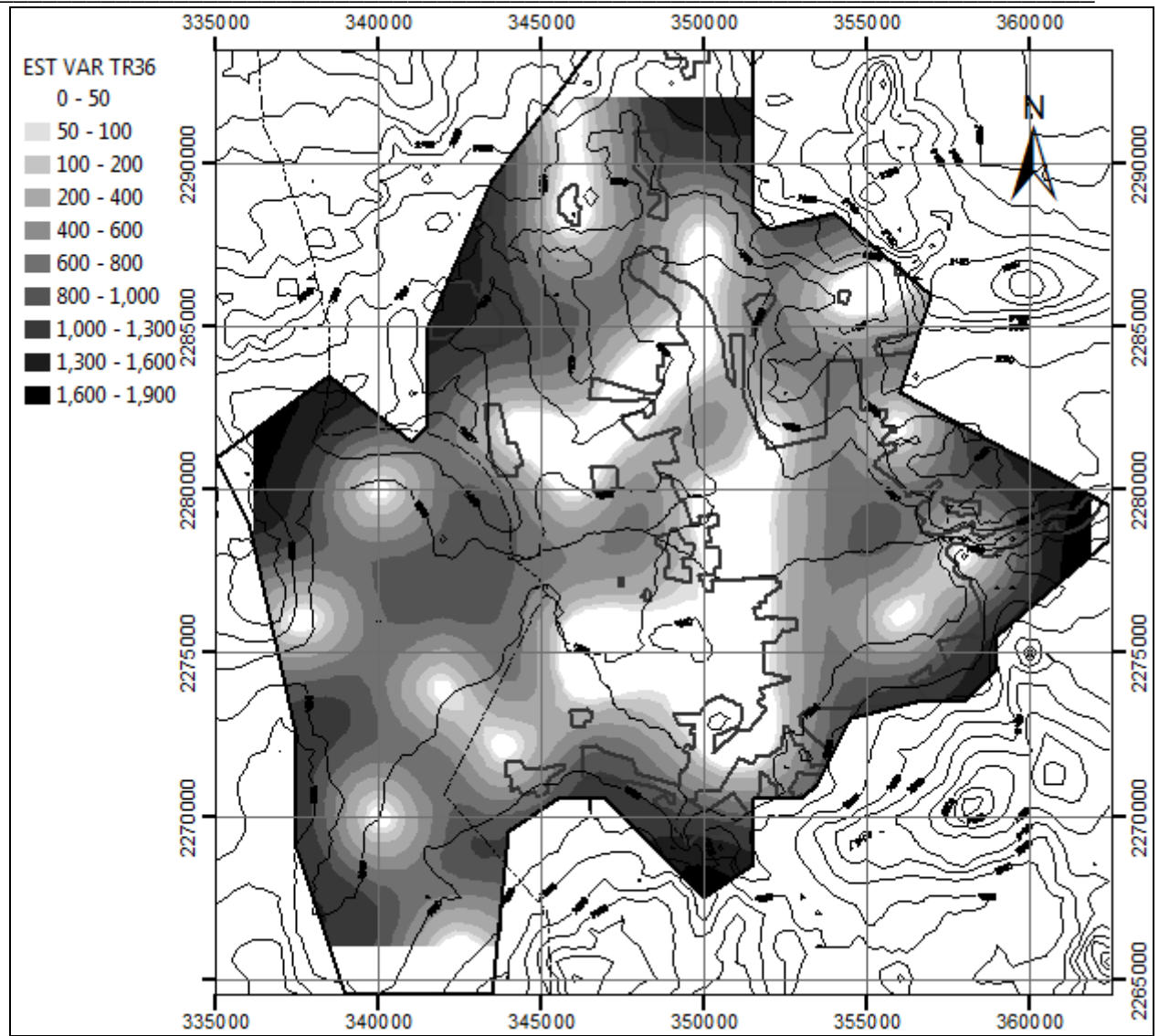


Fig. 6.17 Varianzas del error en la estimación para julio del cuarto año

EVALUACIÓN DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON
BAJO NIVEL DE INFORMACIÓN ESPACIO-TEMPORAL

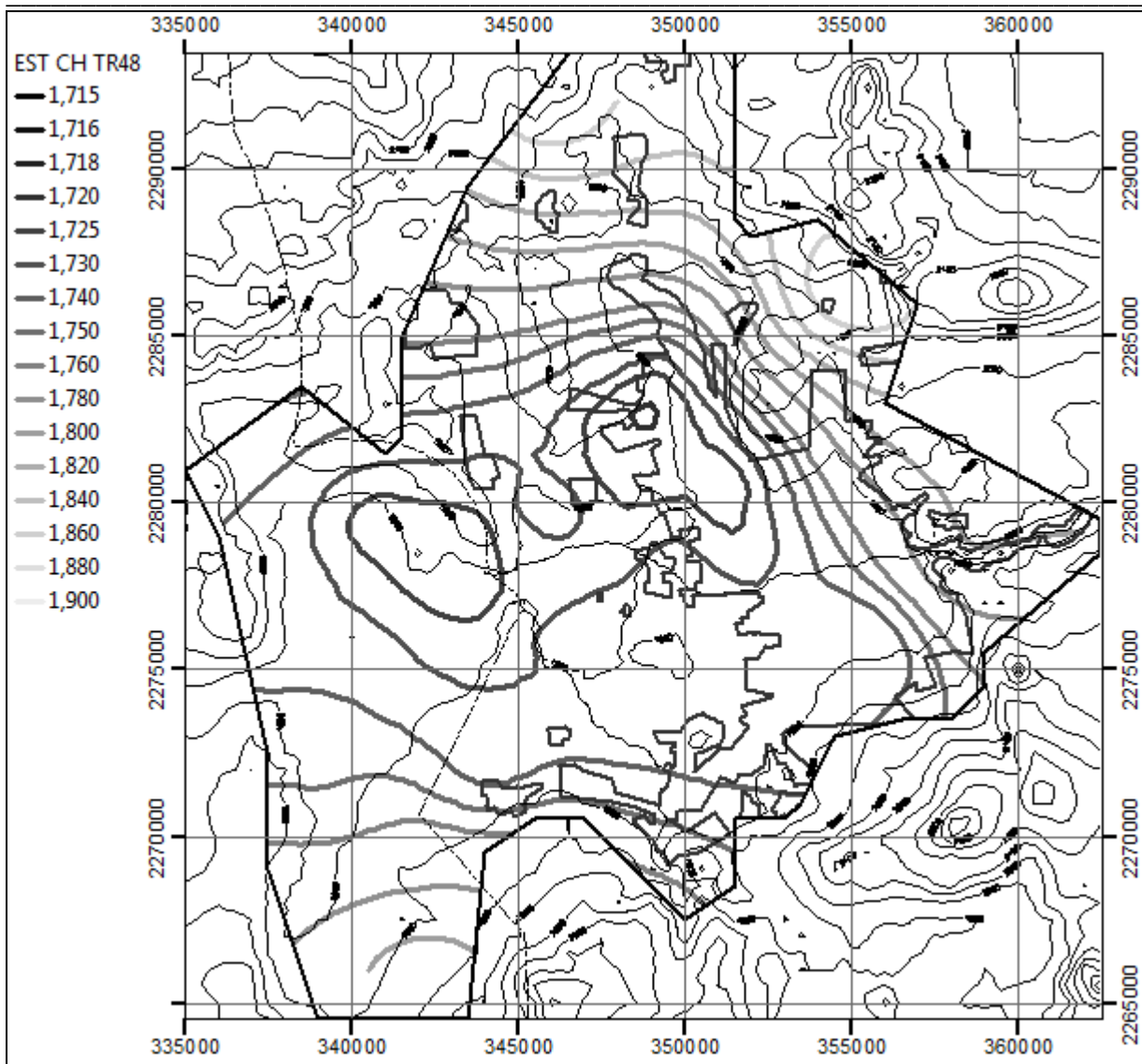


Fig. 6.18 Estimación de la carga hidráulica para julio del quinto año

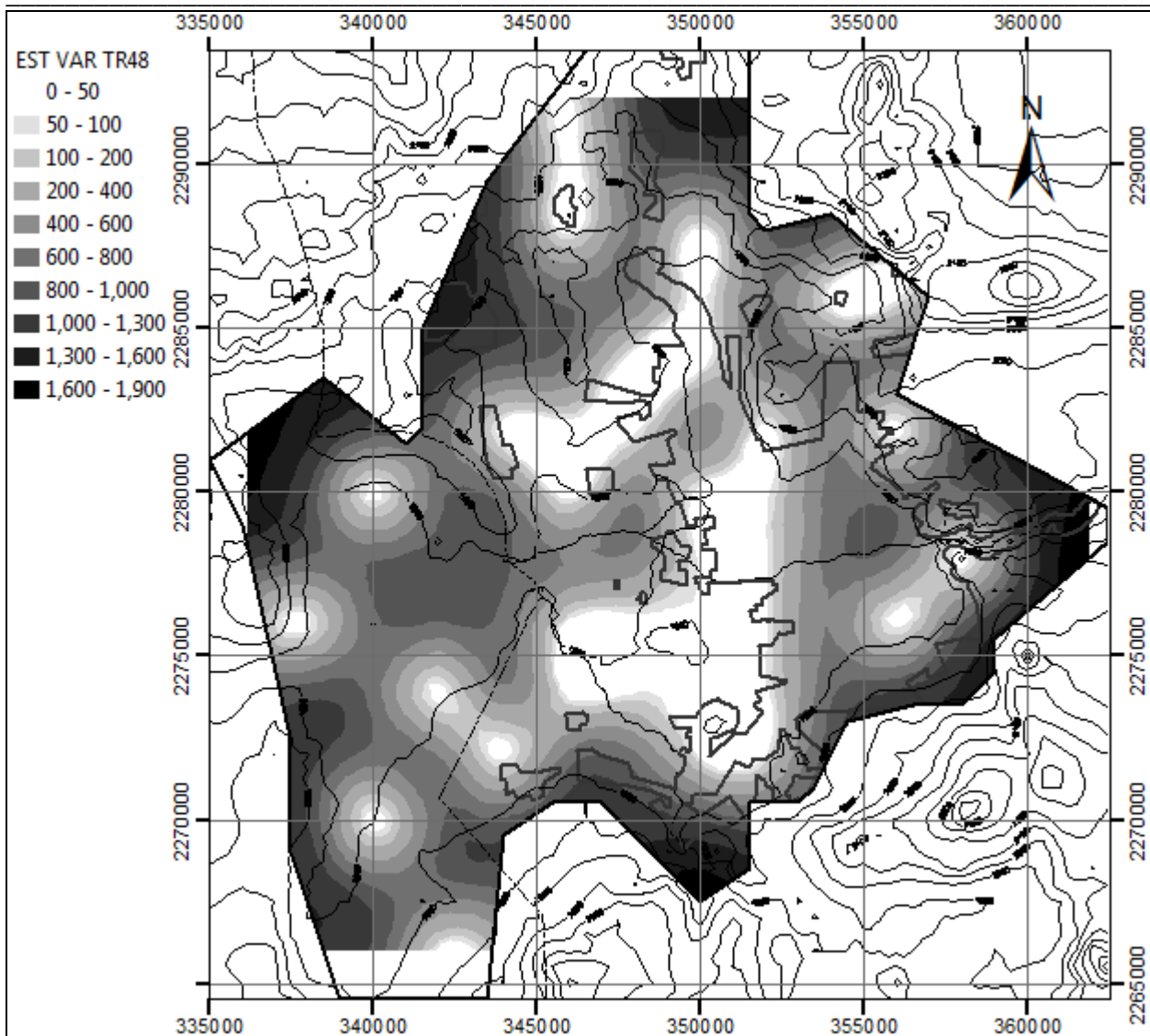


Fig. 6.19 Varianzas del error en la estimación para julio del quinto año

6.9 Evaluación del desempeño de la red de monitoreo a través del tiempo

Para evaluar la forma en que la red de monitoreo captura el comportamiento temporal de la carga hidráulica dentro del periodo de agosto del primer año a julio del quinto año, se seleccionaron los pozos L-246, 644-A, L-107, 2082, 1792, 752 y 976, distribuidos de manera más o menos uniforme en el acuífero (Fig. 6.4). Las Figuras 6.20 - 6.26 muestran las estimaciones de la carga hidráulica para cada uno de estos pozos durante el periodo de monitoreo de la red y las simulaciones obtenidas del modelo numérico de flujo. Se observa que para los pozos en los que se tiene un descenso sostenido de los niveles (salvo el pozo 644-A que se explicará con más detalle un poco más adelante) la red de monitoreo captura un poco mejor el comportamiento de la carga hidráulica que para los casos en los que los niveles se recuperan a través del tiempo (pozos 1792, 752 y 976), esto se debe al descenso sostenido de los niveles en las estimaciones debido a la forma en que fue integrado el vector de estado a priori, en donde se propuso este comportamiento en el acuífero a través del tiempo; la estructura del vector de

estado obliga a que los niveles descieran conforme se avanza en el tiempo, únicamente el monitoreo en esa posición o en posiciones cercanas podría modificar las estimaciones de manera local. Las estimaciones en el pozo 644-A (Fig. 6.21) muestran una aparente recuperación de los niveles a partir de la segunda campaña de monitoreo (julio del tercer año), esto puede asociarse a que los niveles en pozos vecinos son mayores al dato monitoreado en ese pozo en julio del tercer año, por lo que los niveles de las estimaciones posteriores en ese pozo aumentan ligeramente hasta que se encuentran en valores similares a los de los vecinos para después reiniciar con el descenso a partir de octubre del cuarto año. Los mayores errores en la estimación de la carga hidráulica se encontraron para el pozo 1792 (Fig. 6.24), esto se explica porque el comportamiento en ese pozo es contrario al de sus vecinos que forman parte de la red de monitoreo.

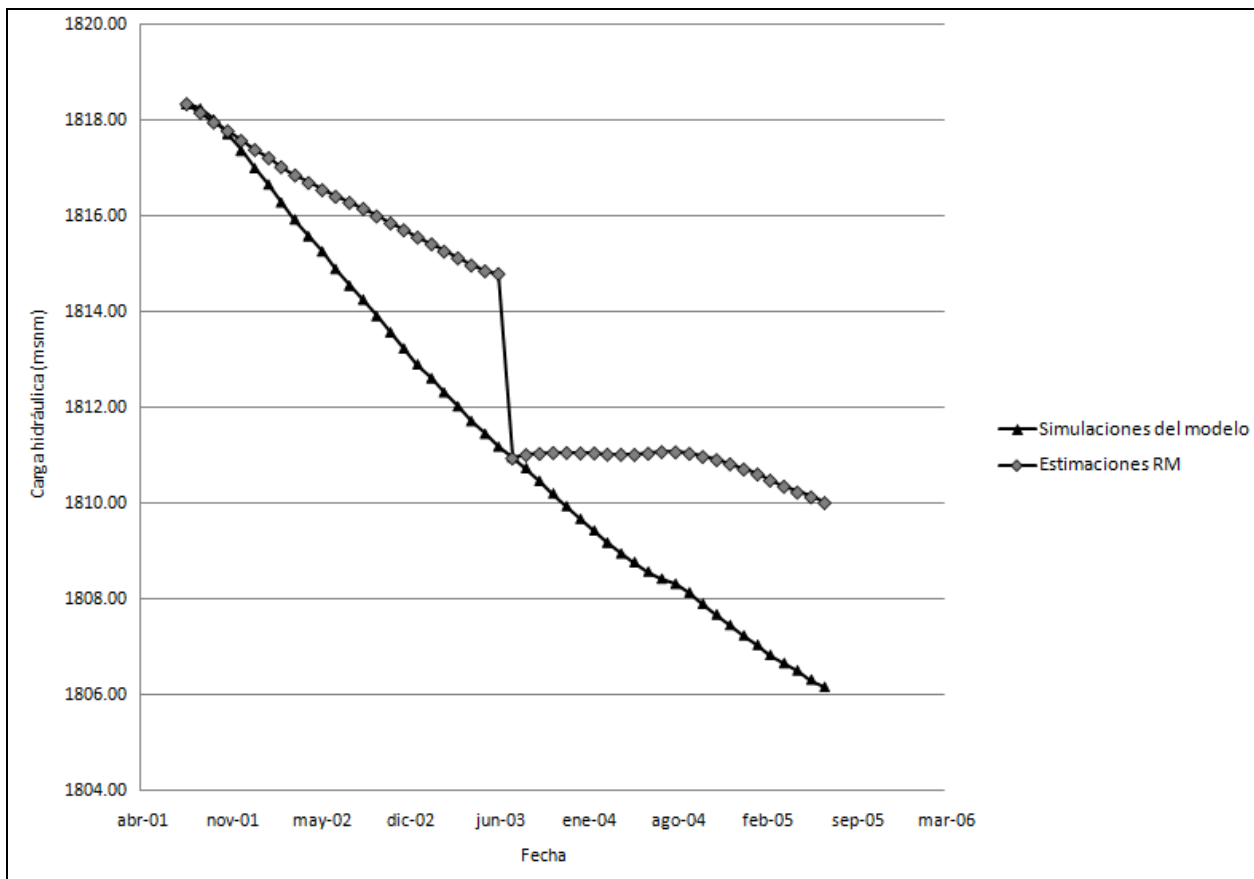


Fig. 6.20 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo L-246

EVALUACIÓN DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON
BAJO NIVEL DE INFORMACIÓN ESPACIO-TEMPORAL

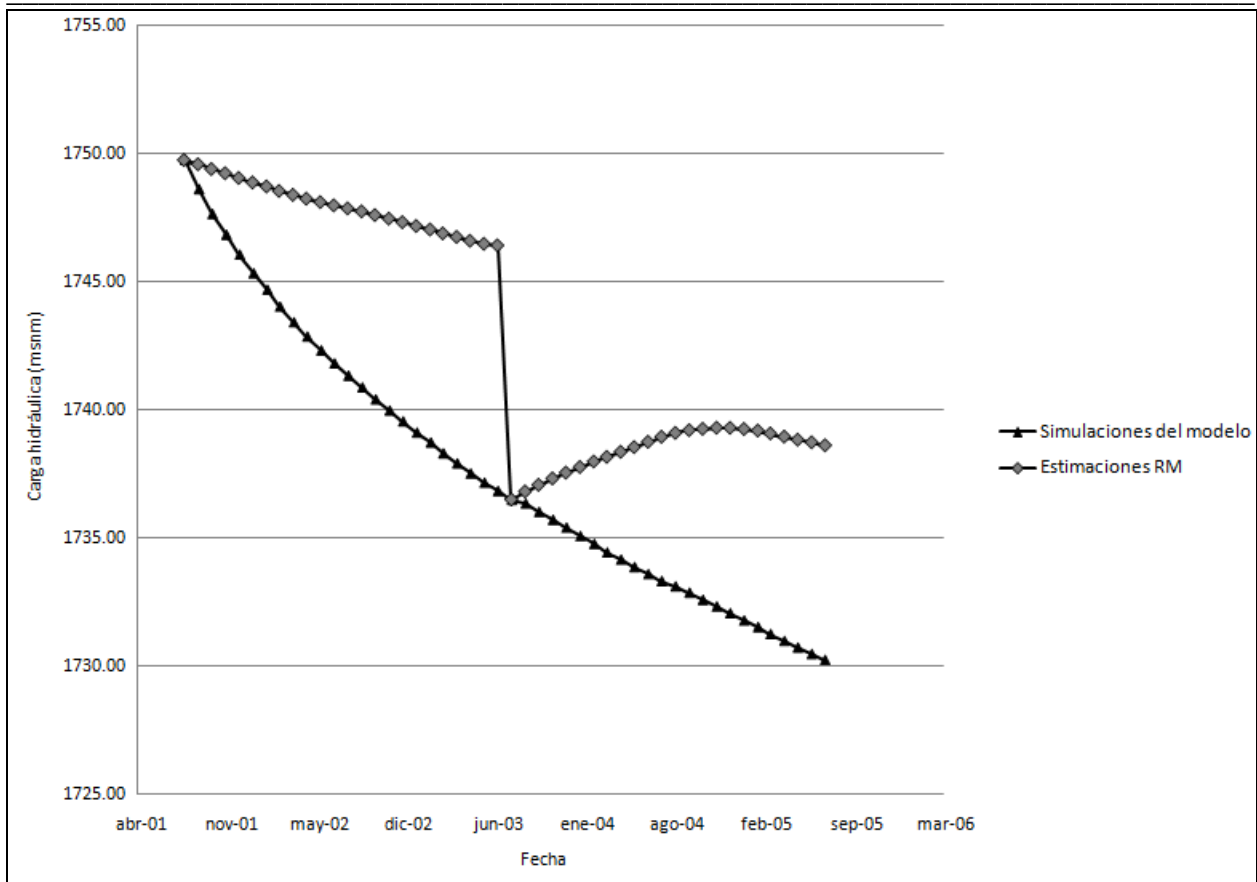


Fig. 6.21 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo 644-A

EVALUACIÓN DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON
BAJO NIVEL DE INFORMACIÓN ESPACIO-TEMPORAL

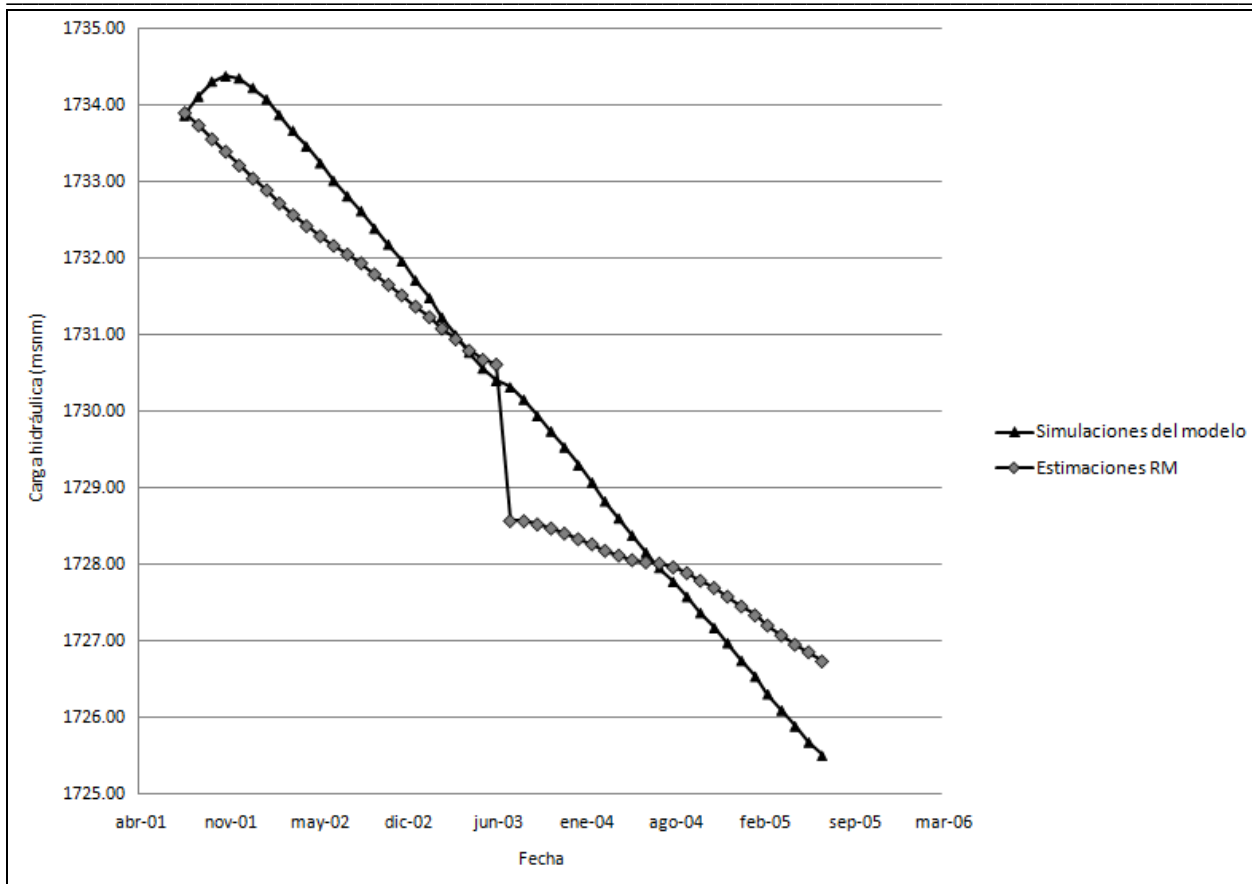


Fig. 6.22 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo L-107

EVALUACIÓN DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON
BAJO NIVEL DE INFORMACIÓN ESPACIO-TEMPORAL

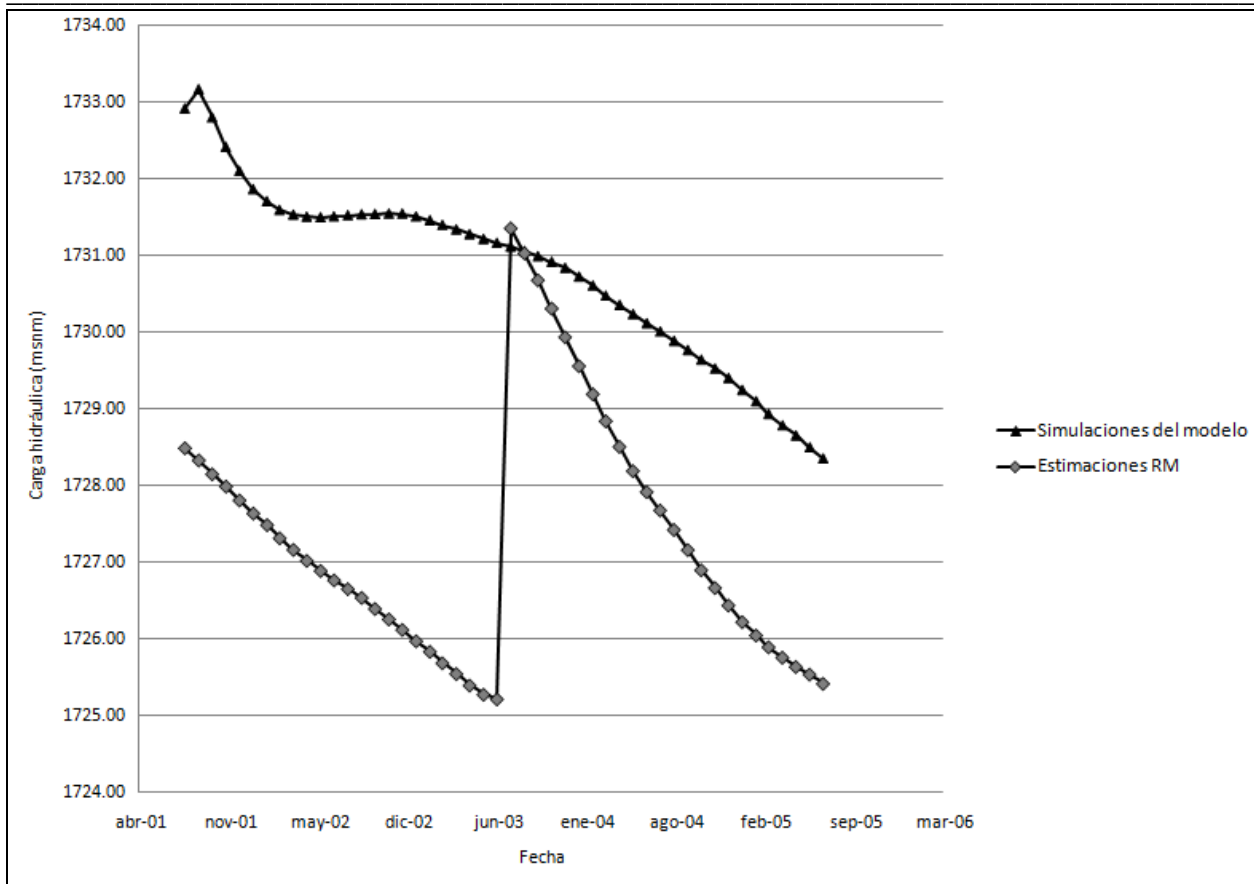


Fig. 6.23 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo 2082

EVALUACIÓN DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON
BAJO NIVEL DE INFORMACIÓN ESPACIO-TEMPORAL

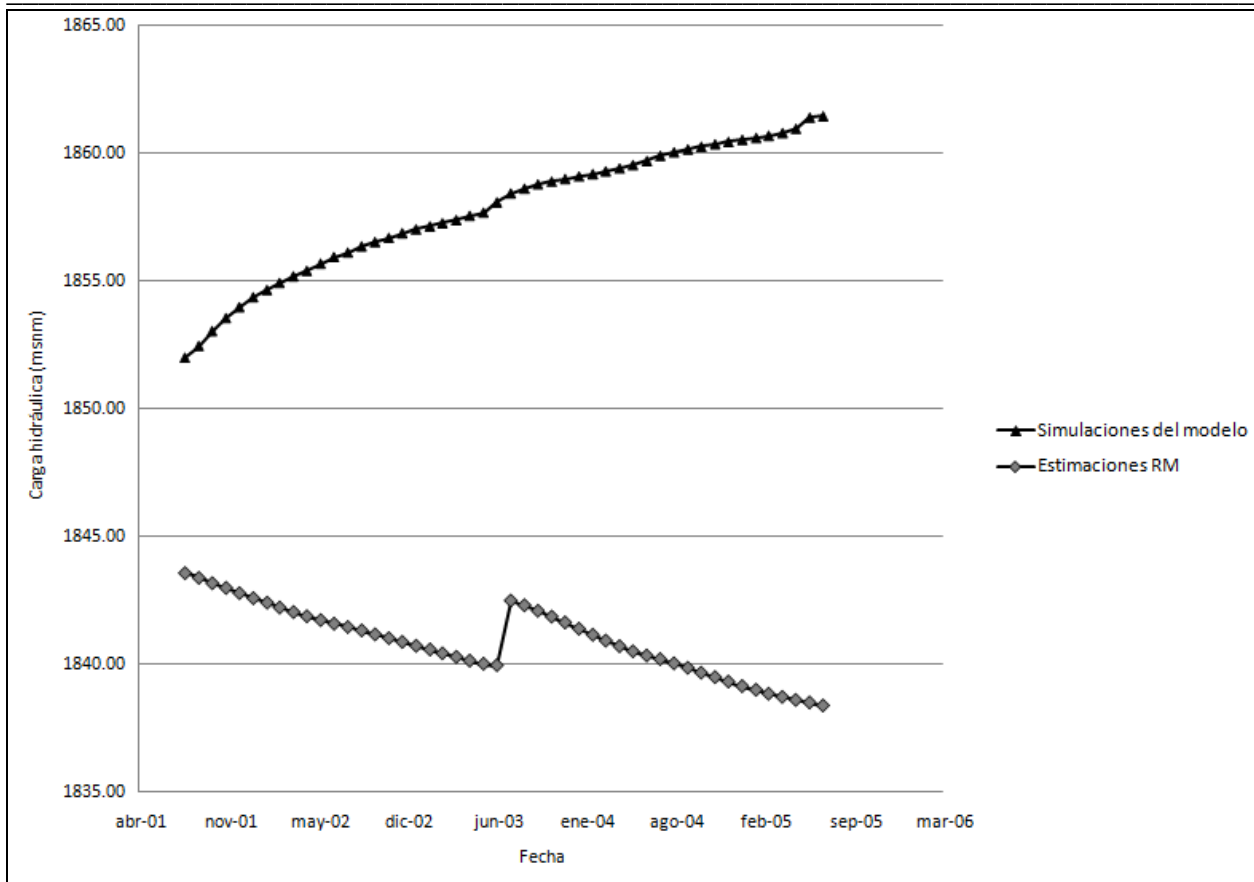


Fig. 6.24 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo 1792

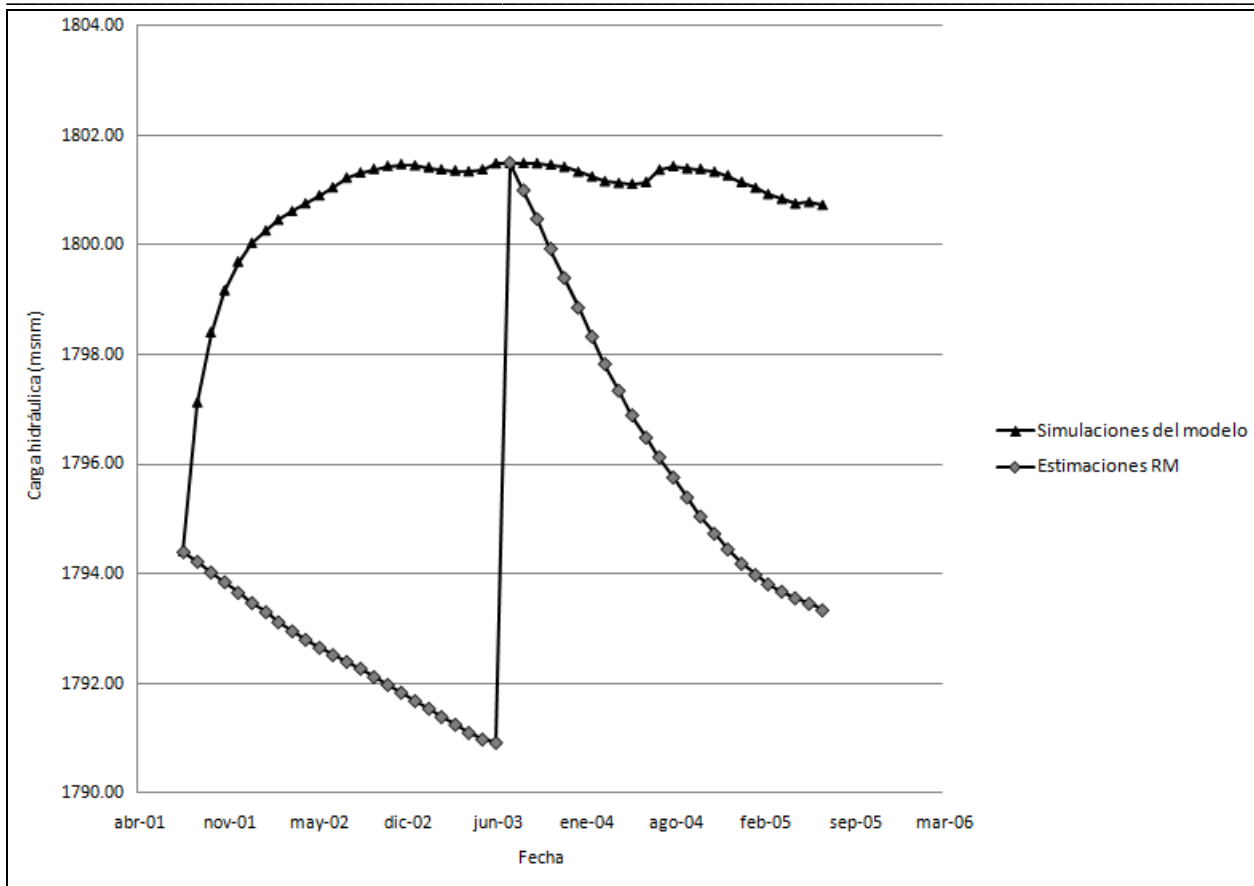


Fig. 6.25 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo 752

EVALUACIÓN DE UNA RED DE MONITOREO DE LA CARGA HIDRÁULICA EN UN ACUÍFERO CON BAJO NIVEL DE INFORMACIÓN ESPACIO-TEMPORAL

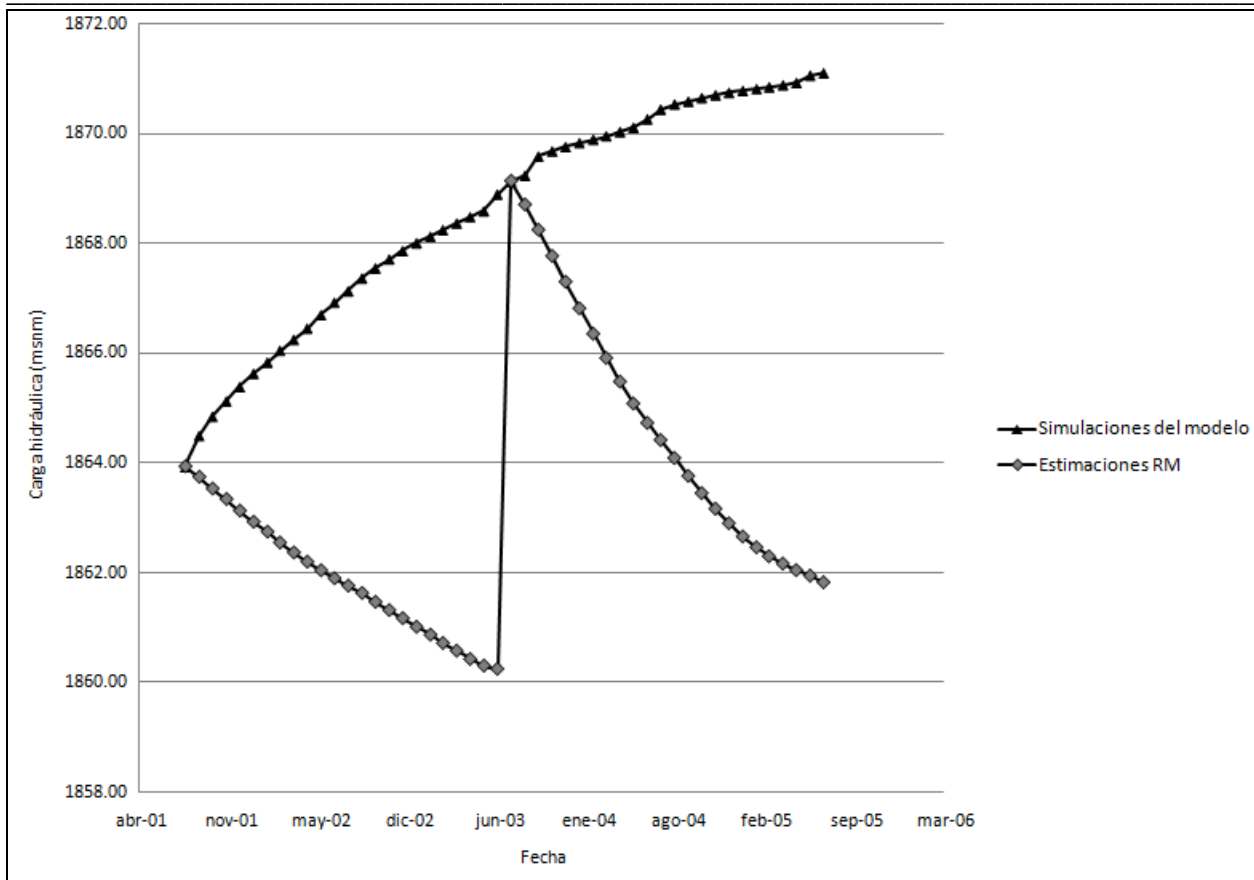


Fig. 6.26 Simulación y estimación de la carga hidráulica empleando la red de monitoreo propuesta para el pozo 976

La forma tan pobre en que la red de monitoreo captura la evolución de la carga hidráulica a través del tiempo confirma su deficiente distribución en el espacio. En la Tabla 6.4 se presenta un resumen de algunos estadígrafos de los errores en la estimación para los pozos analizados.

Tabla 6.4 Errores en la estimación a través del tiempo para algunos pozos utilizando la red de monitoreo y el programa propuesto

Pozo	Error mín (m)	Error máx (m)	Error medio (m)	RECM (m)
L-246	-0.08	3.87	2.04	2.40
0644-A	0.00	9.59	5.49	6.13
L-107	-1.75	1.23	-0.33	0.84
2082	-5.94	0.24	-3.51	3.90
1792	-23.11	-8.42	-16.89	17.30
752	-10.56	0.00	-6.30	6.95
976	-9.29	0.00	-5.28	5.93

6.10 Evaluación del desempeño de la red de monitoreo en la captura de la distribución espacial

Se consideró importante evaluar las estimaciones de la carga hidráulica en el corte A-A' de la Fig. 6.4 porque atraviesa la zona donde se ubican los abatimientos de niveles más significativos y además comprende otros sitios importantes como son zonas de recarga del acuífero. Como se observa en la Fig. 6.27, los mayores errores en las estimación de la carga hidráulica se presentan en la periferia del acuífero, en las zonas considerada como de recarga; esto se debe a que la red de monitoreo no tiene ningún pozo en esos sitios que aporte información de los niveles.

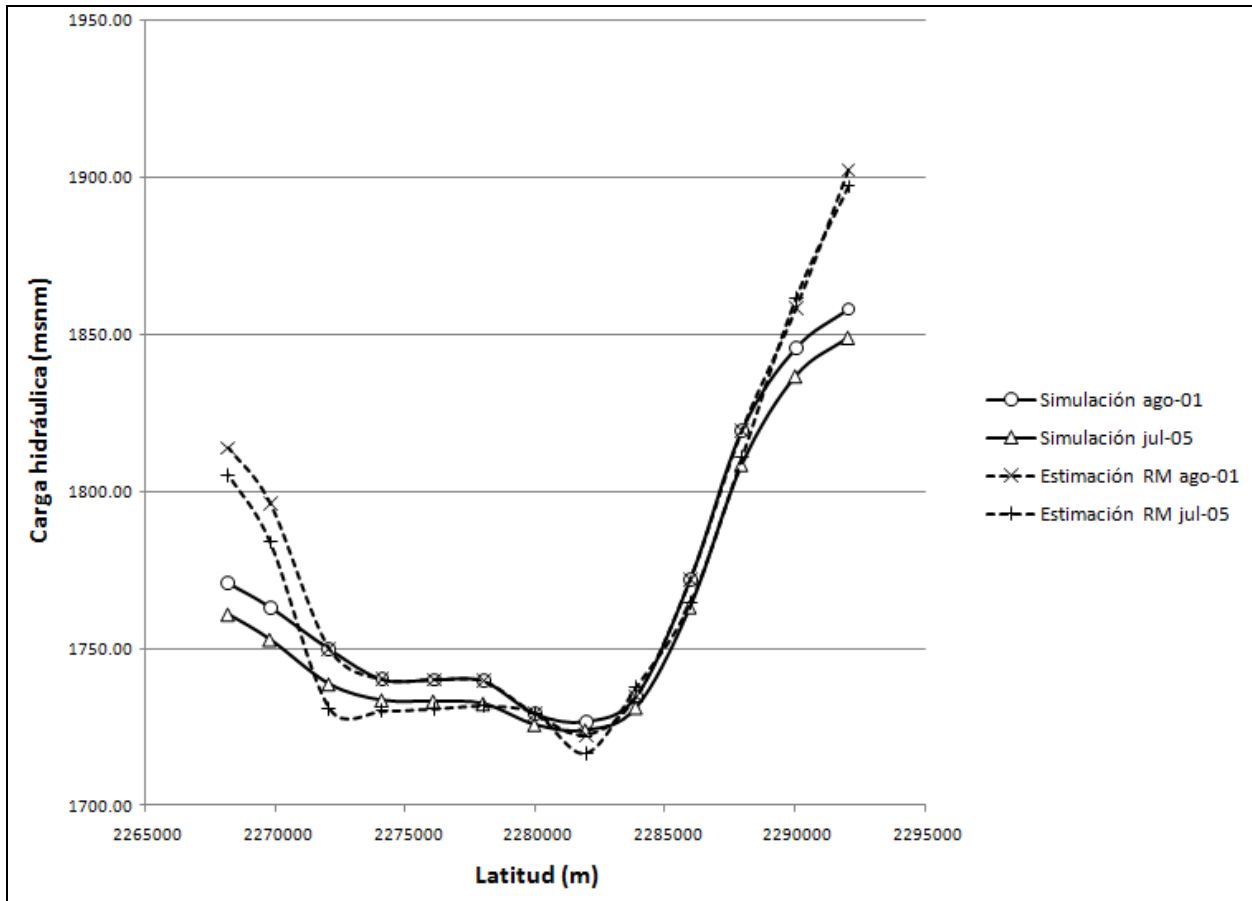


Fig. 6.27 Simulación y estimación de la carga hidráulica para el corte A-A' de la Fig. 5.4

En la Tabla 6.5 se presentan estadígrafos de los errores en la estimación de la carga hidráulica en el corte A-A' para el periodo de monitoreo, se aprecia que el error mínimo (-16.25 m) y el error máximo (51.45 m) se presentan para agosto del tercer año, justo al iniciar la segunda campaña de monitoreo; estos errores ocurren en las latitudes 2281952.746 y 2268187.854, respectivamente. El error medio más grande es de 10.32 m para enero del cuarto año, mientras que la RECM máxima es de 23.33 m para agosto del tercer año, esto confirma que los datos recabados por la segunda campaña de monitoreo perjudican las estimaciones en las zonas que no son cubiertas por la red y que tienen comportamiento contrario al de los pozos monitoreados, como se había señalado anteriormente.

Tabla 6.5 Errores en la estimación de la carga hidráulica para el corte A-A' de la Fig. 5.4 utilizando la red de monitoreo propuesta

Fecha de monitoreo	Error mín	Error máx	Error medio	RECM (m)
ago-01	-4.28	44.22	8.57	19.76
sep-01	-3.34	44.12	8.90	19.75
oct-01	-4.38	43.36	8.52	19.65
nov-01	-1.44	42.93	9.72	19.44
dic-01	-4.00	43.00	8.27	18.47
ene-02	-3.89	44.55	9.12	19.75
feb-02	-3.47	41.95	8.36	18.61
mar-02	-3.84	40.09	7.45	18.24
abr-02	-6.93	43.57	7.55	19.61
may-02	-7.84	43.97	7.54	20.10
jun-02	-4.01	41.80	7.85	18.80
jul-02	-6.77	43.71	6.30	19.20
ago-02	-5.11	44.13	7.99	19.53
sep-02	-5.34	44.86	8.44	20.08
oct-02	-5.71	43.24	7.87	19.62
nov-02	-9.06	42.77	6.43	19.44
dic-02	-7.66	43.94	7.49	19.55
ene-03	-8.26	42.62	6.92	19.42
feb-03	-8.53	42.36	7.11	18.79
mar-03	-8.59	40.96	7.22	18.46
abr-03	-6.40	42.63	8.15	19.13
may-03	-7.02	42.24	7.59	18.94
jun-03	-9.33	43.15	8.16	19.31
jul-03	-8.31	42.92	8.32	19.60
ago-03	-16.25	51.45	9.53	23.33
sep-03	-15.50	51.15	9.57	23.16
oct-03	-15.65	50.10	9.32	22.93
nov-03	-13.99	49.44	9.86	22.72
dic-03	-14.30	46.45	8.97	21.79
ene-04	-12.09	48.03	10.32	22.61
feb-04	-11.84	45.91	9.40	21.48
mar-04	-12.48	45.66	8.19	21.02
abr-04	-13.62	49.72	9.13	22.60
may-04	-13.44	49.26	9.09	22.67
jun-04	-10.37	46.07	8.88	21.14
jul-04	-12.15	49.17	7.76	21.74
ago-04	-9.90	48.97	9.07	22.04
sep-04	-9.10	48.41	9.37	22.28
oct-04	-8.54	47.19	9.14	21.80
nov-04	-11.80	46.61	7.93	21.43
dic-04	-10.60	48.70	8.95	21.91
ene-05	-11.25	47.18	8.46	21.81
feb-05	-9.34	47.03	8.51	21.16
mar-05	-9.07	45.81	8.65	20.92
abr-05	-7.96	45.53	9.08	21.00
may-05	-8.29	45.42	8.81	20.97
jun-05	-8.25	48.29	9.38	21.54
jul-05	-7.67	48.08	9.35	21.56

6.11 Conclusiones y recomendaciones

El semivariograma espacio-temporal muestral queda mejor definido conforme se cuenta con una mayor cantidad de información espacio-temporal de la variable de estudio (como en el caso presentado en el capítulo cinco), mientras que uno estimado con huecos en la información espacio-temporal, como el del presente capítulo (en el que se emplea un subconjunto de los datos empleados en el análisis del capítulo cinco) presentará fluctuaciones. Las diferencias más significativas entre los modelos de semivariograma ajustados son la reducción en un 25% del alcance de la componente espacial y el aumento en más del 300% del valor de la meseta temporal para el caso del análisis con menos datos, esto sugiere que al emplear un menor número de datos en el análisis geostatístico se obtiene un modelo más limitado.

Los resultados muestran que la raíz cuadrada de la varianza promedio del error en la estimación de la carga hidráulica en el área de estudio, cuando se monitorean las 912 posiciones de monitoreo espacio-temporal disponibles, es menor a 23.78 m, y que cuando se monitorean las 37 posiciones con la más alta prioridad es menor a 24.12 m. Esto significa que las últimas 875 posiciones de monitoreo espacio-temporal son redundantes ya que la información que ellas proveen reduce muy poco el valor de la raíz cuadrada de la varianza promedio. Para el caso de estudio, antes de elegir alguna posición de monitoreo se tiene un valor de la raíz cuadrada de la varianza promedio del error en la estimación de 57.88 m, esto significa que las primeras 37 posiciones reducen este valor en 33.76 m, mientras que las últimas 875 posiciones únicamente lo reducen 0.34 m más. Estos resultados muestran que el método propuesto es exitoso en propagar la información en el espacio y en el tiempo.

Todas las posiciones seleccionadas por el método propuesto para el diseño óptimo de redes de monitoreo espacio-temporal de la carga hidráulica corresponden al primer mes del periodo de diseño; esto se explica por dos razones: 1) en la reducción de la varianza total para el primer mes no contribuyen ni datos de monitoreos previos (porque aquí inicia el monitoreo), ni de monitoreos posteriores (al estar definida la función objetivo para una estimación en tiempo real), y 2) la varianza temporal es muy pequeña con respecto a la espacial.

A diferencia de una red de monitoreo con una buena cobertura espacial, como la mostrada en el capítulo cinco, para la red analizada en este capítulo, que simula un caso real, se obtienen valores grandes de la raíz cuadrada de la varianza promedio mínima posible, esto indica que la actual red de monitoreo es muy deficiente en proveer la información suficiente para representar con alta certidumbre la evolución espacio-temporal de la carga hidráulica en el acuífero; el monitoreo a través del tiempo resulta innecesario si no se captura primero de forma adecuada el patrón de la distribución espacial de la carga hidráulica en todo el acuífero. Esto indica que nuevos pozos deben ser agregados a la red de monitoreo óptima, en zonas con los valores más altos de la varianza del error en la estimación. Para proponer la ubicación de los nuevos pozos, también deben usarse criterios prácticos e hidrogeológicos.

Como era de esperarse, empleando la red y el programa de monitoreo propuesto no es posible capturar de manera adecuada la evolución de la carga hidráulica en el futuro, de hecho los errores aumentan. La raíz del error cuadrático medio para el periodo de diseño es de 11.38 m mientras que para el periodo de evaluación en el futuro es de 12.97 m. En redes de monitoreo pobremente distribuidas en el espacio, las estimaciones obtenidas en la malla de estimación espacio-temporal pueden deteriorarse a través del tiempo si se monitorea en pozos cuyo comportamiento es contrario al de las posiciones de estimación espacio-temporal vecinas.

7. DISEÑO DE UNA RED DE MONITOREO DEL NIVEL DEL AGUA PARA EL ACUÍFERO VALLE DE QUERÉTARO

La metodología propuesta es demostrada con datos históricos de elevaciones del nivel del agua medidos en un acuífero ubicado en la parte central de México. Se trata de un acuífero libre que se encuentra localmente confinado en algunas regiones (Guysa, 1996). Cubre un área de 379.25 km² y las zonas más profundas son cercanas a los 600 m. Este acuífero se encuentra en dos estados mexicanos: Querétaro y Guanajuato. Presentamos el diseño para la parte este del acuífero, localizada en el estado de Querétaro (Fig. 7.1).

7.1 Base de datos del nivel del agua subterránea

La base de datos para el análisis espacio-temporal incluye información de niveles del agua subterránea en 273 pozos localizados en el acuífero Valle de Querétaro para el periodo comprendido entre agosto de 1970 y noviembre de 2007 (Fig. 7.1). No existen registros de niveles del agua subterránea anuales completos para ninguno de los pozos durante este periodo. El pozo 612-F tiene el mayor número de datos (32), por otro lado, diciembre de 1995 es el mes con más información (60 datos). En total, contamos con 1435 datos del nivel del agua subterránea.

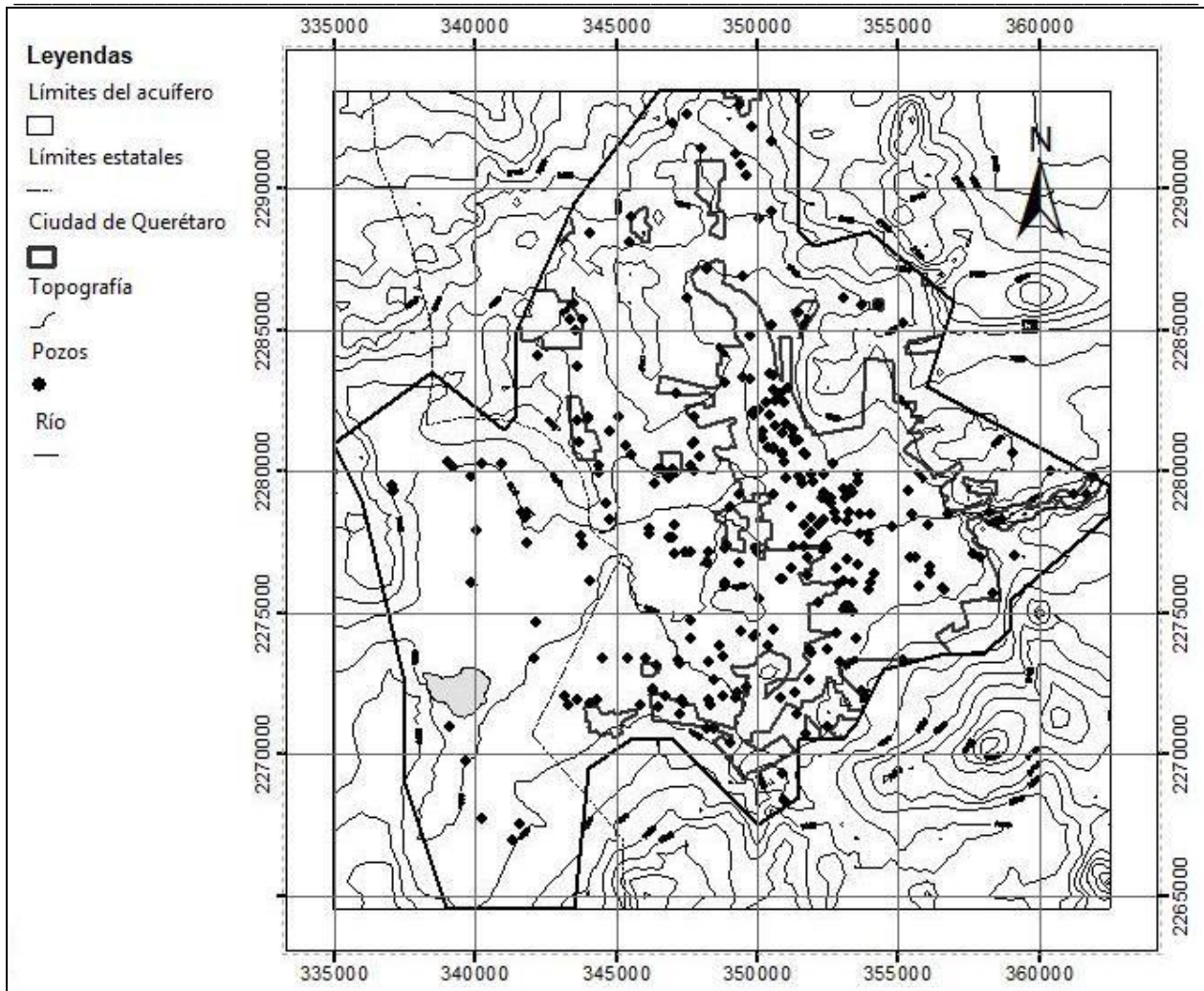


Fig. 7.1 Área de estudio y ubicación de pozos con datos del nivel del agua subterránea

7.2 Objetivo del monitoreo

La red de monitoreo fue diseñada para el periodo 1997-2007. El objetivo del diseño es seleccionar las posiciones y tiempos de monitoreo necesarios para obtener una buena estimación de la variable en cada año para todo el acuífero. Como se asignó el mismo peso a la incertidumbre en todo el acuífero, los puntos de estimación espacial fueron ubicados en una malla regular, con elementos cuadrados de 2 km de lado. Estos nodos (82) constituyen lo que llamamos la malla de estimación (Fig. 7.2). Un posible tiempo de estimación fue seleccionado para diciembre (el mes más monitoreado dentro del periodo de diseño) de cada año, con excepción del 2007, para el cual se usó noviembre porque sólo se tenía información disponible de ese mes. De esta manera, 82 posiciones y 11 tiempos fueron seleccionados para la malla de estimación espacio-temporal.



Fig. 7.2 Malla de estimación (cuadros) y posibles puntos de monitoreo espacial (círculos)

Se seleccionaron treinta y ocho posiciones como posibles posiciones de monitoreo espaciales (Fig. 7.2). Se seleccionaron estos pozos porque fueron monitoreados en las campañas de monitoreo más recientes de niveles del agua subterránea, en diciembre de 2006 y/o noviembre de 2007. Cada diciembre del periodo 1997-2006 y noviembre de 2007 fueron seleccionados como posibles tiempos de monitoreo.

7.3 Análisis geoestadístico espacio-temporal del nivel del agua subterránea

El análisis se basó en la construcción del semivariograma espacio-temporal muestral con los datos del nivel del agua subterránea disponibles (Fig. 7.3). Para llevar a cabo este análisis utilizamos una rutina modificada del GAMV (Deutsch *et al.*, 1997) por De Cesare *et al.* (2002) para la modelación espacio-temporal de semivariogramas. Empleamos 20 intervalos espaciales, con 1362.4 m de separación de intervalo espacial, y 37 intervalos temporales con una separación de intervalo de un año.

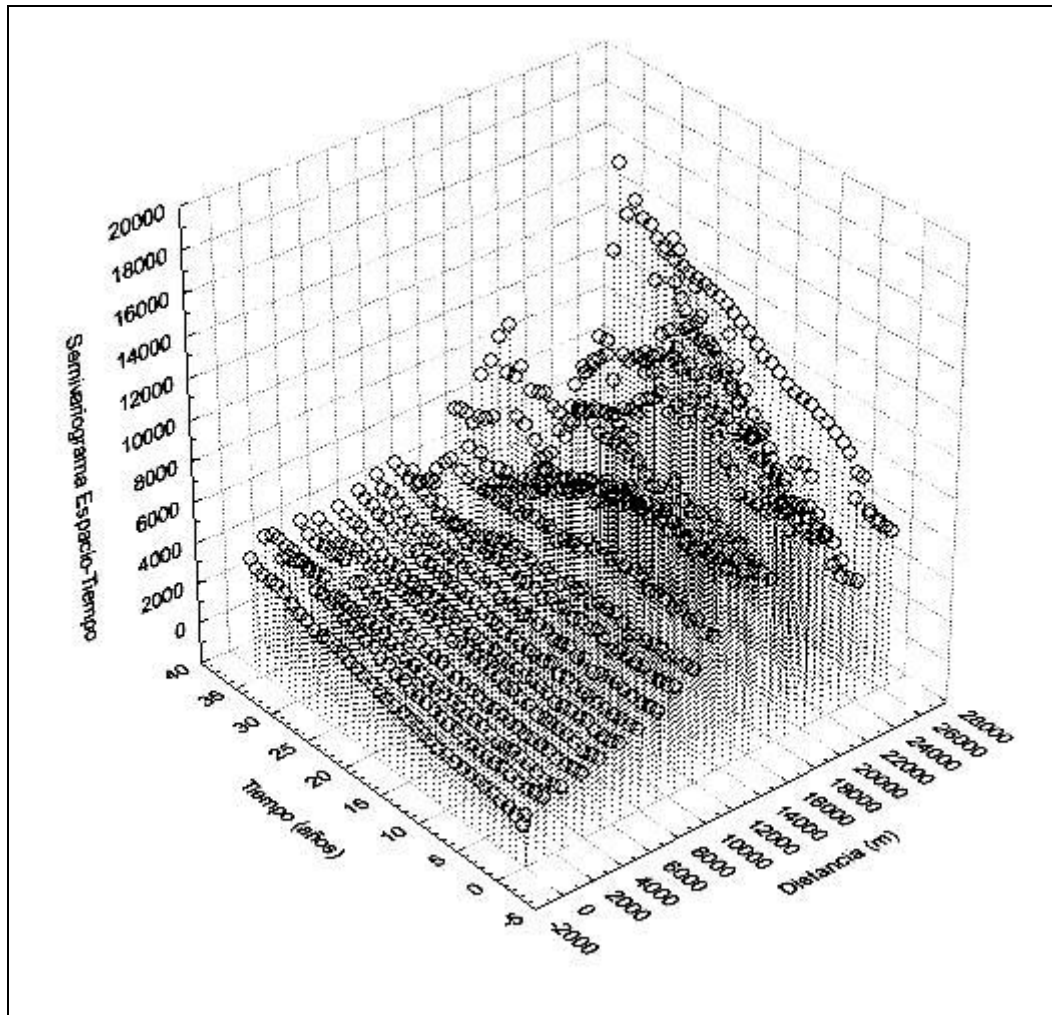


Fig. 7.3 Semivariograma espacio-temporal muestral de niveles del agua subterránea

Para analizar mejor el comportamiento del semivariograma espacio-temporal muestral, presentamos su proyección sobre el plano semivariograma-intervalos espaciales (Fig. 7.4) y sobre el plano semivariograma-intervalos temporales (Fig. 7.5). La proyección espacial muestra al inicio una curvatura cercana a una función cuadrática, lo que indica la posible presencia de tendencia (Journel y Huijbregts 1978). Sin embargo, el semivariograma decrece a partir del intervalo de 22,000 metros. Puede observarse un comportamiento más disperso a partir del intervalo de 13,000 metros. La proyección temporal muestra que los valores más altos del semivariograma tienen mucha mayor variabilidad que los valores pequeños del semivariograma.

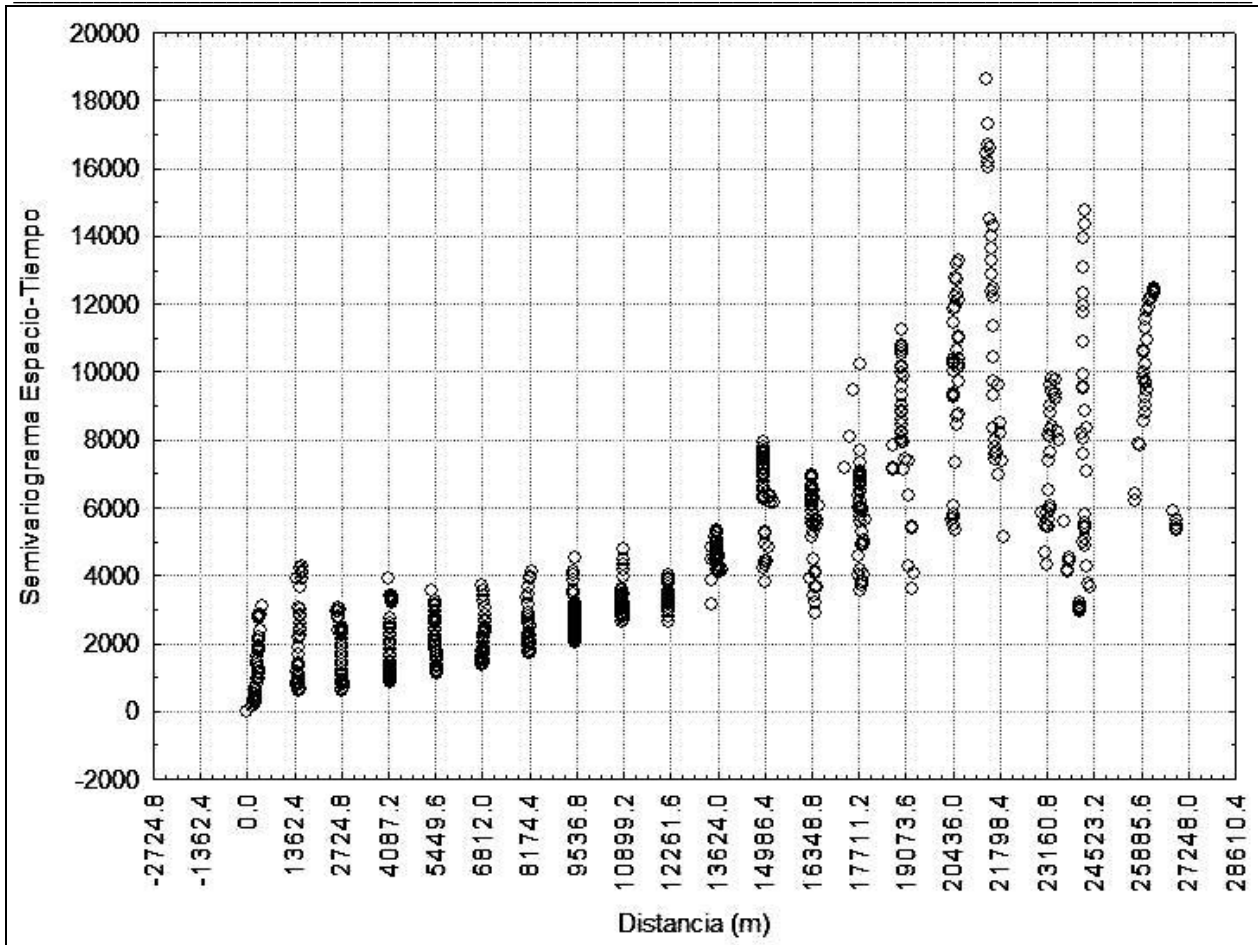


Fig. 7.4 Proyección del semivariograma espacio-temporal muestral de niveles del agua subterránea sobre el plano semivariograma-intervalos espaciales

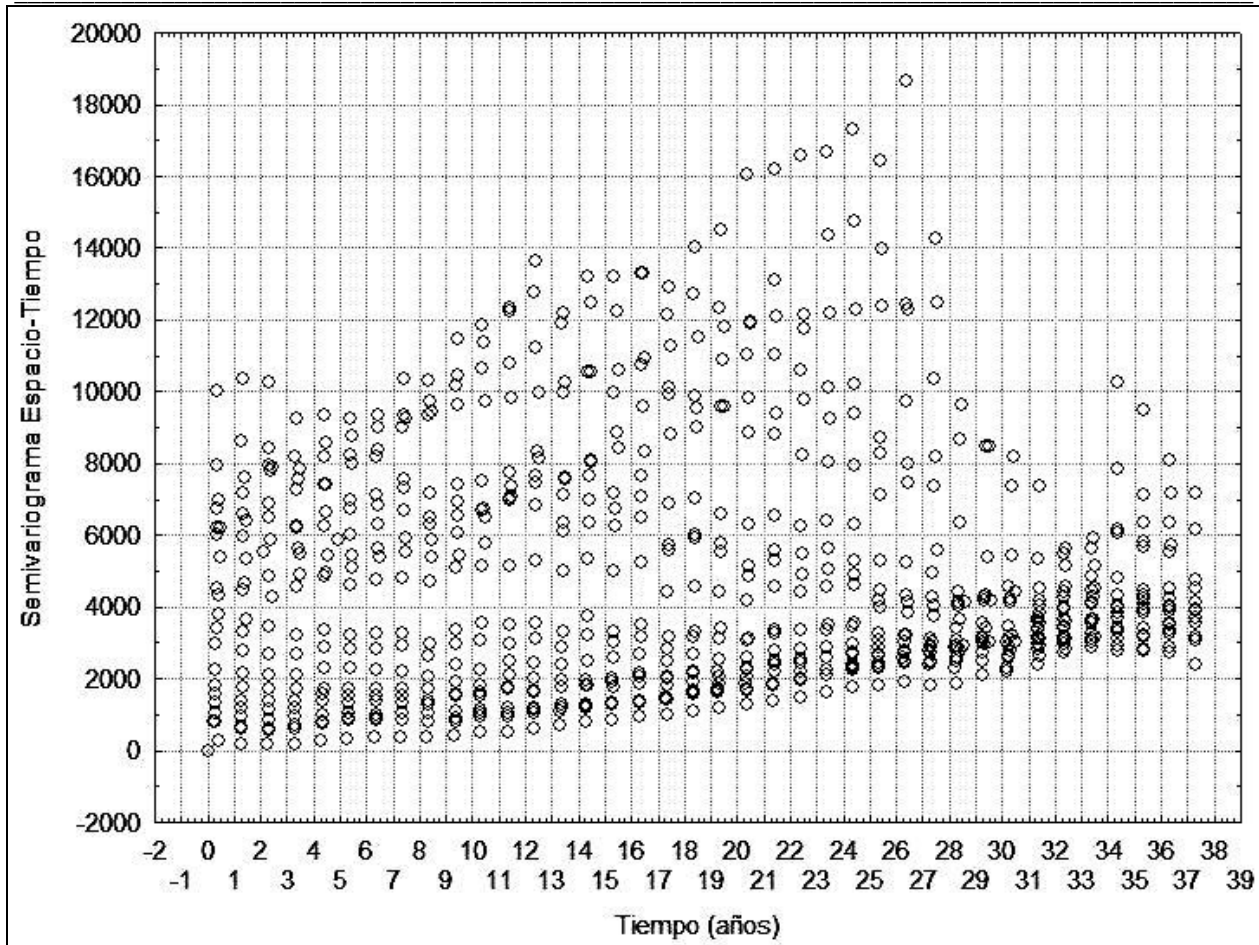


Fig. 7.5 Proyección del semivariograma espacio-temporal muestral de niveles del agua subterránea sobre el plano semivariograma-intervalos temporales

Para continuar con el análisis espacio-temporal, la tendencia fue removida mediante el ajuste de una función polinomial a través de mínimos cuadrados.

Remoción de la tendencia

Se probaron polinomios de primer y segundo grado. El ajuste del polinomio de primer grado fue elegido porque el semivariograma espacio-temporal calculado con los residuos obtenidos está mejor definido (Fig. 7.6). Se usaron el mismo número de intervalos espaciales y temporales, y la separación de intervalos, que en el análisis de niveles del agua subterránea.

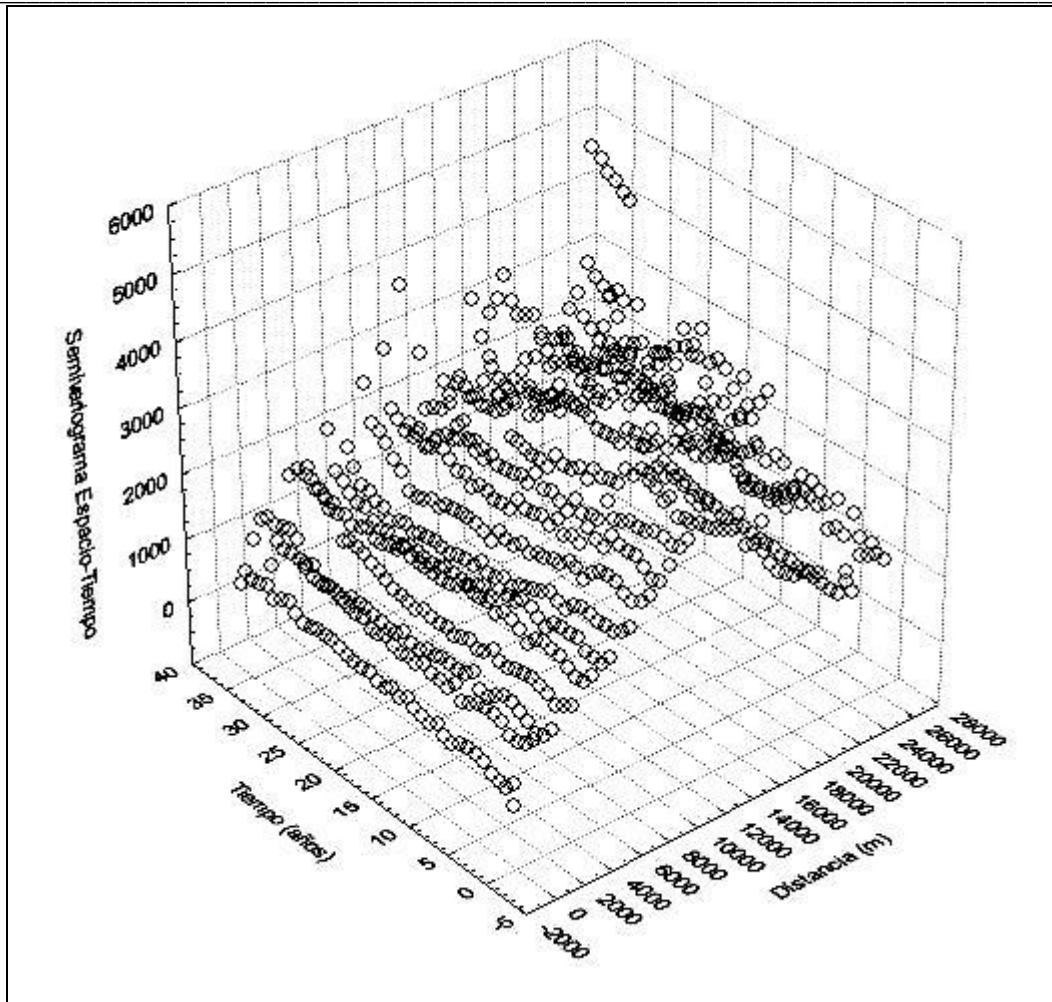


Fig. 7.6 Semivariograma espacio-temporal muestral de los residuos de un ajuste de un polinomio de primer grado

La Fig. 7.7 muestra la proyección espacial. El comportamiento de función cuadrática ya no se presenta en el semivariograma, éste se acota a los 16,000 m. Sin embargo, el comportamiento decreciente en la parte final del semivariograma aún está presente, pero ahora ocurre después del intervalo de 17,000 m. La Fig. 7.8 muestra la proyección temporal. Se encuentra un comportamiento más homogéneo que antes de remover la tendencia, aunque se sigue presentando una variabilidad mayor en los valores altos del semivariograma.

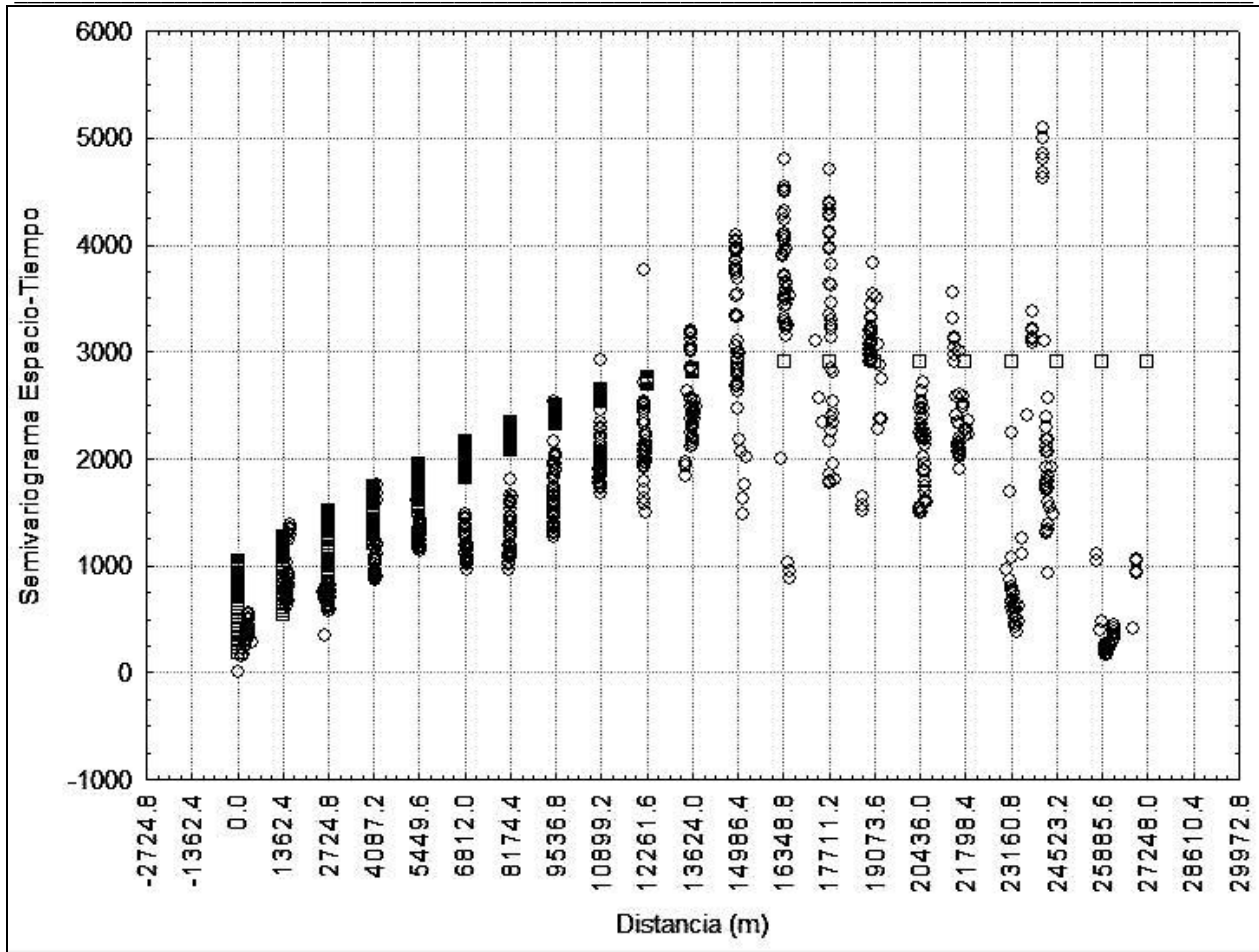


Fig. 7.7 Proyección del semivariograma espacio-temporal muestral de los residuos de un ajuste de un polinomio de primer grado sobre el plano semivariograma-intervalos espaciales (círculos) y modelo de semivariograma ajustado (cuadros)

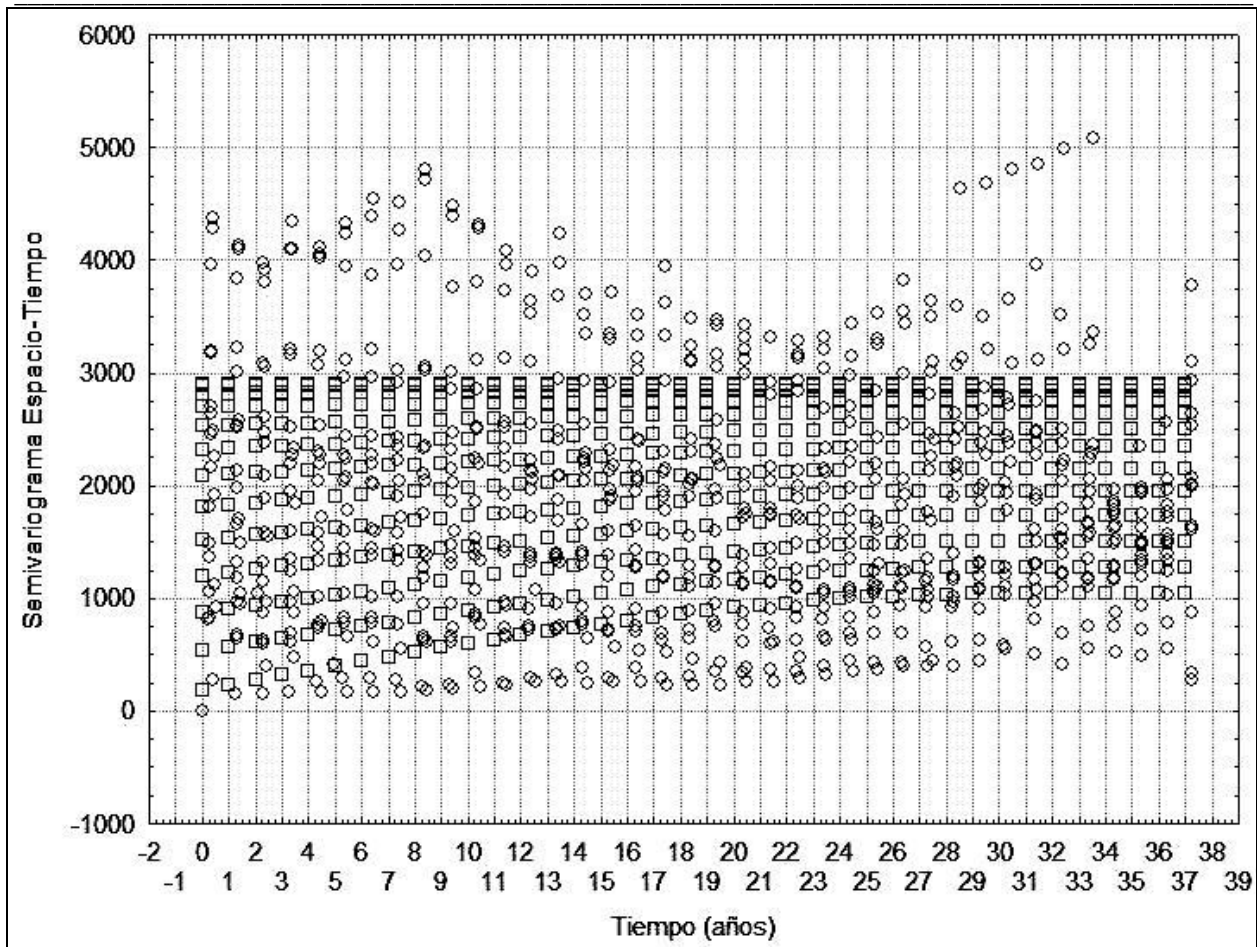


Fig. 7.8 Proyección del semivariograma espacio-temporal muestral de los residuos de un ajuste de un polinomio de primer grado sobre el plano semivariograma-intervalos temporales (círculos) y modelo de semivariograma ajustado (cuadros)

Ajuste del modelo de semivariograma

Para el ajuste del modelo de semivariograma, se emplearon los datos del nivel del agua subterránea de los últimos 10 años (diciembre de 1997-noviembre de 2007) porque las condiciones del acuífero no han cambiado drásticamente desde ese periodo de tiempo, y por lo tanto un semivariograma bien ajustado para esos años será útil para estimar los niveles del agua subterránea con la red diseñada.

Siguiendo a De Iaco *et al.* (2002), un modelo producto-suma fue ajustado manualmente al semivariograma espacio-temporal muestral (esto quiere decir que diferentes modelos fueron propuestos “a ojo” y evaluados para seleccionar aquél con los mejores resultados en la validación cruzada). Los parámetros del modelo seleccionado se presentan en la Tabla 7.1.

Tabla 7.1 Parámetros del modelo de semivariograma espacio-temporal de los residuos de un ajuste de un polinomio de primer grado

Modelo	Componente	Pepita	Meseta	Alcance
Esférico	Espacial	150	2900	16000
Esférico	Temporal	40	940	30
-	Global	-	2900	-

Los resultados de la validación cruzada para los últimos 10 años con información se presentan en la Tabla 7.2; como puede verse, los errores máximos y mínimos parecen ser grandes, sin embargo, estos valores corresponden al 8.4% y 6%, respectivamente, del valor del nivel del agua subterránea medido en campo en esas posiciones, siendo además estos los mayores errores porcentuales de la validación, por lo que consideramos que se obtiene un buen ajuste tomando en cuenta la alta variabilidad de los datos.

Tabla 7.2 Resultados de la validación cruzada para los últimos 10 años con información

Error mín (m)	Error máx (m)	Error medio (m)	ECM (m ²)	RECM (m)	ECME
-163.91	101.89	-0.87	422.24	20.55	1.01

7.4 Optimización de la red de monitoreo

Como el modelo de semivariograma espacio-temporal fue ajustado a residuos del ajuste de un polinomio de primer grado, el filtro de Kalman estático fue aplicado para estimar esta variable. Una vez que los residuos eran estimados, los valores del nivel del agua subterránea en cada nodo de estimación espacio-temporal fueron obtenidos sumando su estimación residual y el valor de la función polinomial en la correspondiente posición espacio-temporal. Como se mencionó anteriormente, el filtro de Kalman estático necesita de estimaciones a priori del vector de estado y de la matriz de covarianza del error en la estimación. Para cada año, el vector de estado a priori se conformó con el promedio de los residuos correspondientes a las posiciones de los datos de niveles del agua subterránea medidos en campo. Al no haber datos para diciembre de 2001 y diciembre de 2002, los componentes el vector de estado para esos años fueron obtenidos mediante una interpolación lineal de los promedios de años adyacentes. La matriz de covarianza espacio-temporal a priori fue calculada del modelo de semivariograma espacio-temporal, aplicando la ecuación 3.15; fue construida para 82 (posiciones de estimación) + 38 (posiciones de monitoreo) + 29 (posiciones de apoyo adicionales para estimar datos faltantes) = 149 posiciones y 11 tiempos, de tal manera que su dimensión fue de 1639 columnas (149×11) por 1639 renglones.

Orden de prioridad

Una vez aplicado el método de optimización de redes de monitoreo propuesto en esta tesis, el orden de selección es un indicador de la importancia de los datos obtenidos en esas posiciones, en la reducción de la varianza total. Puede verse en la Fig. 7.9 que la varianza total se reduce grandemente al elegir las primeras posiciones de monitoreo espacio-temporal, pero conforme se incrementa el número de posiciones seleccionadas, la reducción de la varianza total decrece.

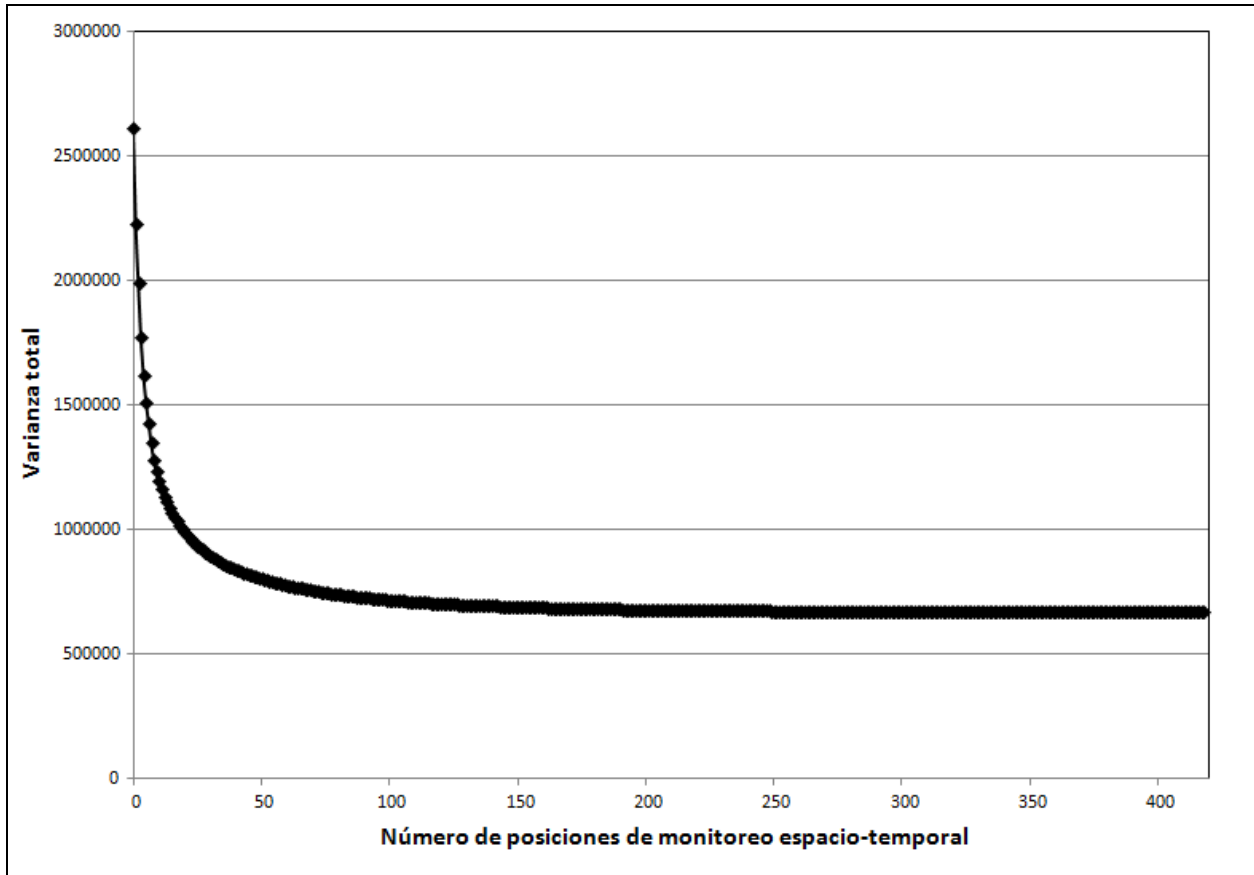


Fig. 7.9 Varianza Total vs Número de posiciones de monitoreo espacio-temporal

7.5 Determinación del número total de posiciones de monitoreo

Puede verse en la Fig. 7.9 que después de monitorear en las 200 posiciones con la mayor prioridad, la varianza total permanece prácticamente constante (alrededor de 675,000 m²). Esto sugiere que las 218 posiciones restantes pudieran ser redundantes, esto es, proveen poca información y no contribuyen mucho en la reducción de la varianza total.

Para definir el número de posiciones de la red de monitoreo espacio-temporal, se utilizó el mismo criterio que en los capítulos cinco y seis. Los valores calculados de la raíz cuadrada (Varianza total/Posiciones de estimación espacio-temporal) contra el número de posiciones de monitoreo espacio-temporal se muestran en la Fig. 7. 10.

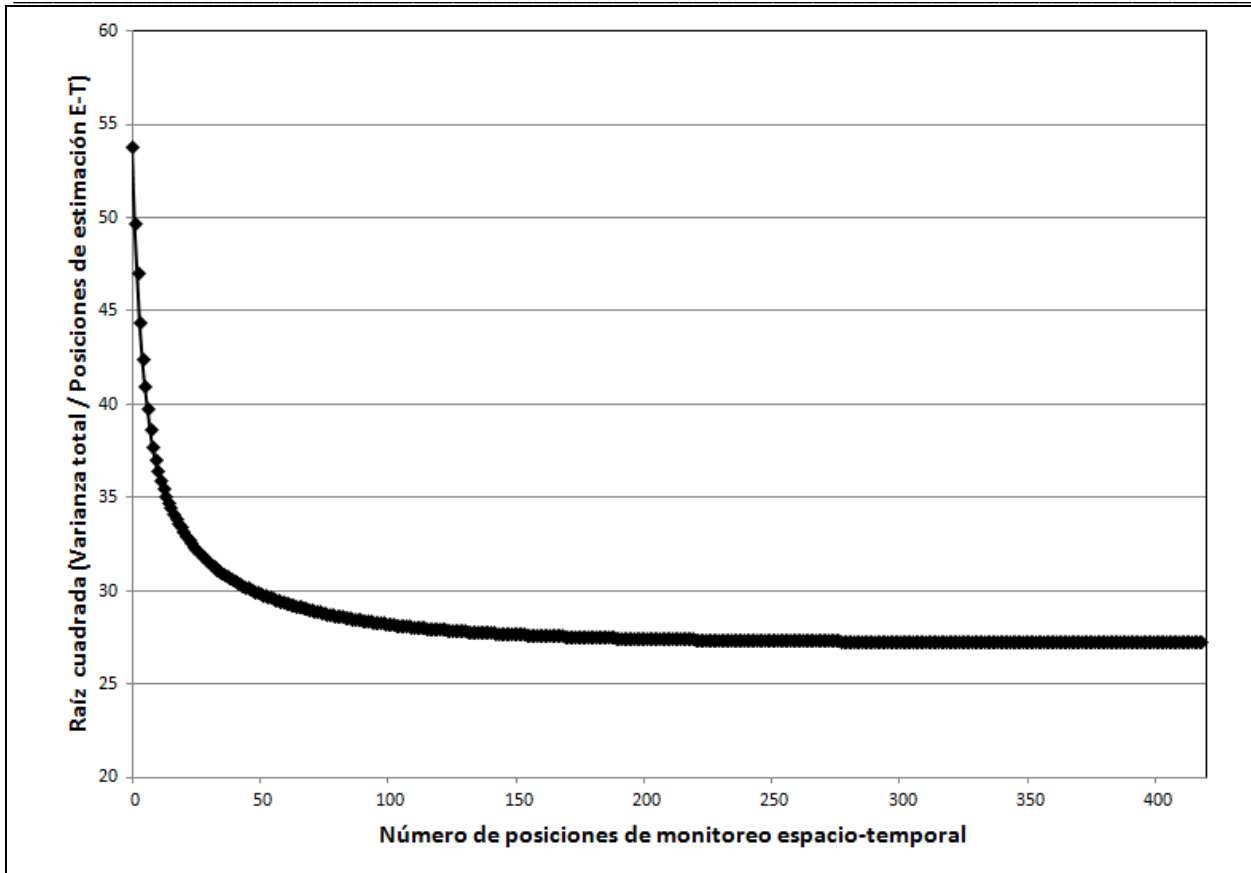


Fig. 7.10 Raíz cuadrada (Varianza total/Posiciones de estimación espacio-temporal) vs número de posiciones de monitoreo espacio-temporal

Se alcanza un 99% del valor de la MRP cuando 178 posiciones de monitoreo espacio-temporal son seleccionadas y una incertidumbre en la estimación promedio de 27.53 m es esperada en la malla de estimación espacio-temporal. Este valor es alto pero es extremadamente cercano al obtenido usando todas las posiciones de monitoreo espacio-temporal disponibles (27.27 m). Esto significa que la red de monitoreo existente no permite reducir lo suficiente la incertidumbre en algunas zonas, y es necesario monitorear otras posiciones.

7.6 Red de monitoreo óptima

En este ejemplo y de acuerdo a lo expuesto anteriormente, 178 posiciones de monitoreo espacio-temporal fueron seleccionados para constituir la red de monitoreo óptima. La Fig. 7.11 muestra el número de posiciones a ser monitoreadas en cada uno de los tiempos propuestos. Puede verse que la campaña más intensa de monitoreo es propuesta para el primer año del período de diseño, lo cual es consistente con el criterio de optimización elegido. Las posiciones del primer año tienen la más alta prioridad porque no hay información disponible (en el periodo de diseño) de años previos para reducir la varianza total del error en la estimación de ese año, y también, porque cada pozo elegido en el primer año provee información para todo el periodo (por el alcance del modelo seleccionado). Cuando avanzamos en el tiempo, menos pozos son necesitados, pero periódicamente un pequeño incremento ocurre porque decrece la influencia de

años previos altamente monitoreados, por lo que se va necesitando el monitoreo en nuevos pozos para mejorar las estimaciones. En comparación con los casos presentados en los capítulos cinco y seis, se necesita un mayor monitoreo a través del tiempo porque la meseta temporal del modelo ajustado es mucho mayor para el caso analizado en este capítulo; esto se explica porque en esta aplicación se analiza un período de tiempo mucho más grande, en el que se presenta una mayor variabilidad de la información a través del tiempo.

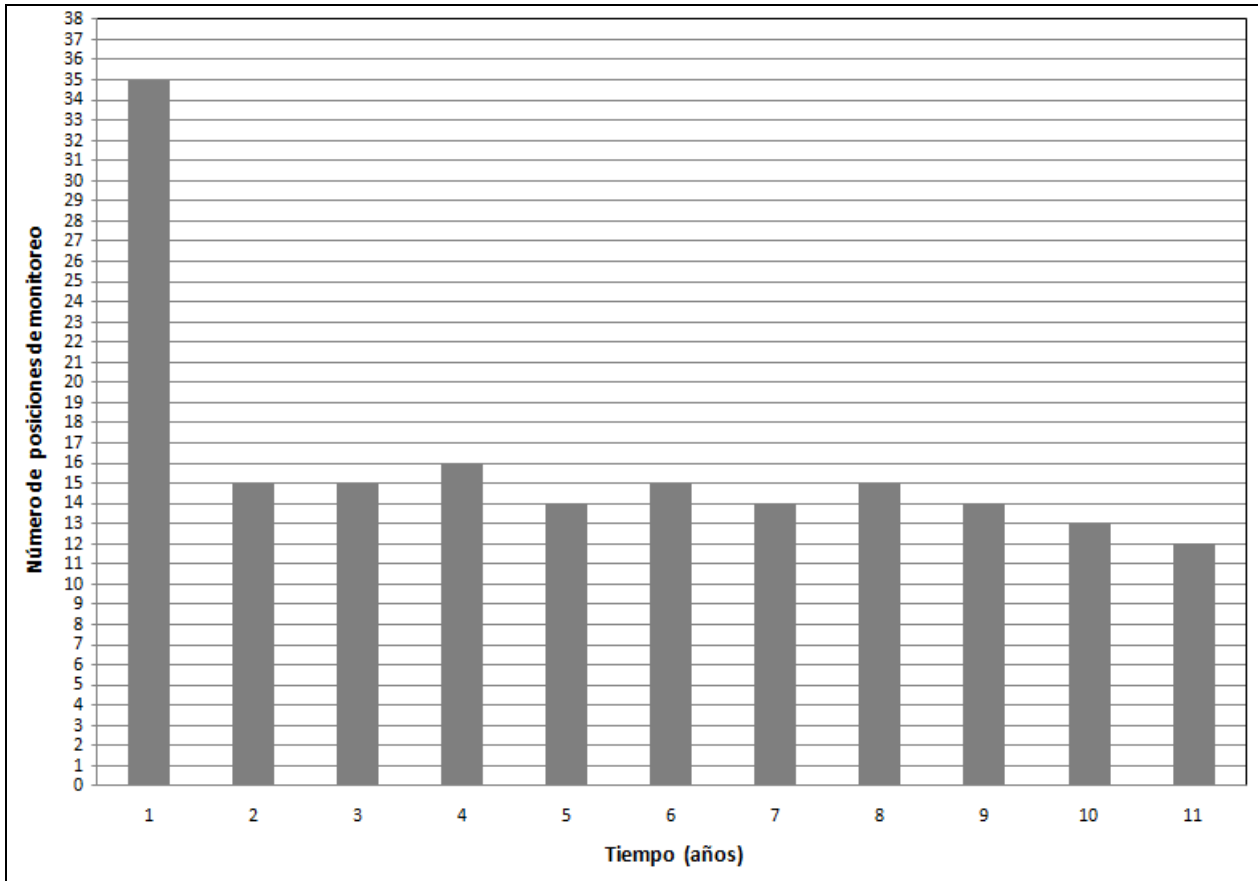


Fig. 7.11 Posiciones de monitoreo para cada año

La Fig. 7.12 muestra el número de veces que debe ser monitoreado cada pozo. Como puede verse, las posiciones que necesitan ser monitoreadas con mayor frecuencia se encuentran ubicadas en zonas pobremente cubiertas por el arreglo espacial actual de la red de monitoreo.

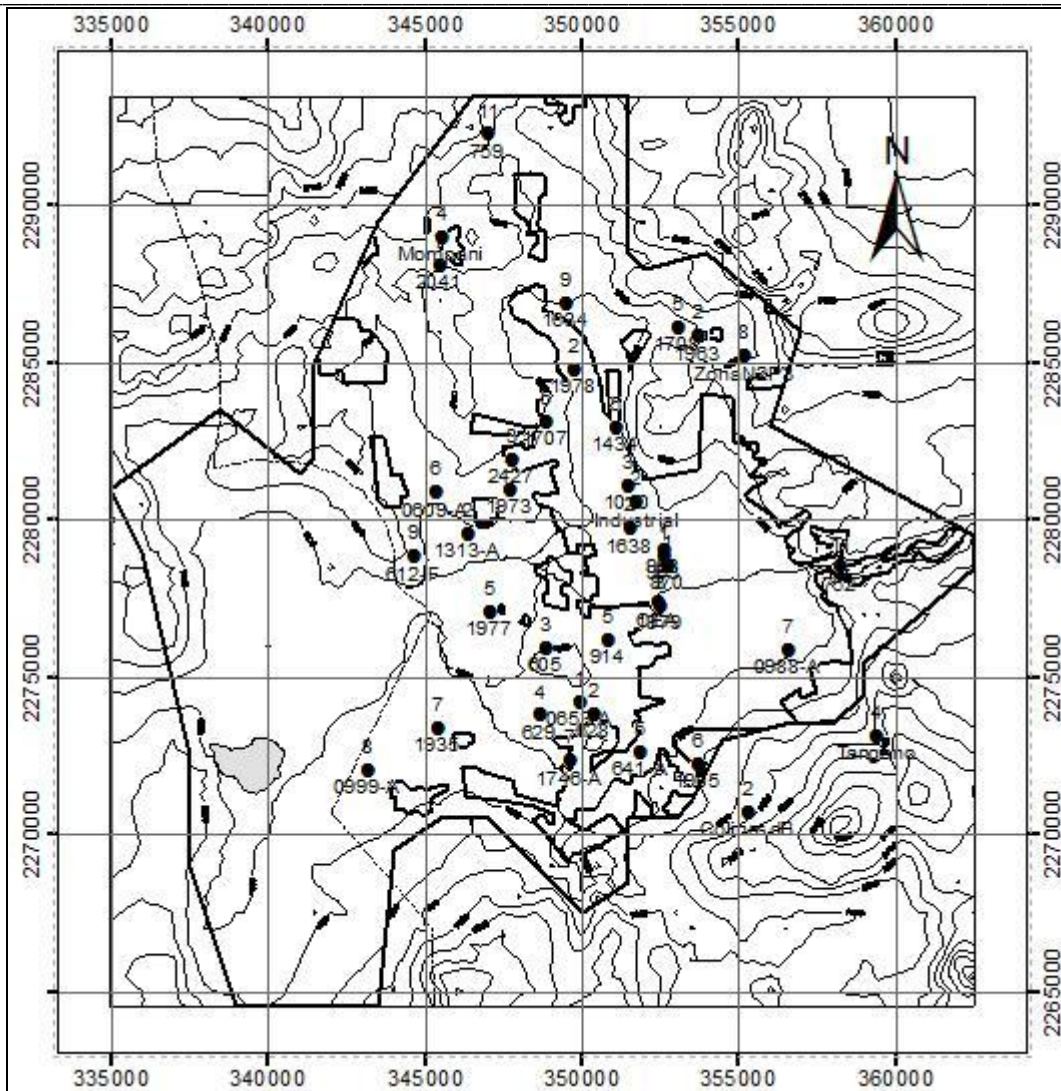


Fig. 7.12 Número de veces de monitoreo para cada pozo en el programa de monitoreo óptimo espacio-temporal (arriba) y nombre del pozo (abajo)

7.7 Estimaciones y varianzas del error en la estimación para el período de diseño

Como una ilustración del nivel de información que proporciona el programa de monitoreo propuesto para el periodo de diseño, se obtuvieron estimaciones del nivel del agua subterránea y de la varianza del error de estimación sobre las posiciones y tiempos de interés. Éstas se obtuvieron aplicando el filtro de Kalman estático.

Como para este caso no hay información en todas las posiciones de monitoreo para todo el periodo de diseño, fue necesario estimar la información faltante. Los valores faltantes fueron obtenidos mediante estimaciones con el filtro de Kalman estático, usando el modelo de semivariograma espacio-temporal seleccionado previamente y los residuos de los datos existentes en las 38 posiciones de monitoreo y las mediciones de apoyo en 29 pozos adicionales (en total, 255 datos espacio-temporales fueron usados para la estimación). Entonces fueron estimados residuos de los datos de niveles del agua subterránea en todas las posiciones de

monitoreo espacio-temporal sin información medida en campo. Mediante este procedimiento fue posible entonces contar con información de los niveles del agua subterránea en las 418 posibles posiciones de monitoreo espacio-temporal.

Estimación del nivel del agua subterránea

En los números impares de las Figuras 7.13 - 7.20 se presentan estimaciones en tiempo real del nivel del agua subterránea, obtenidas para cada año con la red de monitoreo propuesta.

Estas figuras muestran los cambios espaciales y temporales estimados en el acuífero. Las zonas identificadas como de recarga corresponden a las zonas con la máxima elevación del nivel del agua subterránea, las cuales coinciden con los máximos niveles topográficos. En la Fig. 13 podemos ver que para diciembre de 1997, la dirección del flujo es noreste-suroeste, los niveles del agua subterránea en la entrada norte son mayores a 1860 msnm; otra entrada se sitúa al noreste donde el flujo se dirige hacia el centro del acuífero. Los niveles mínimos del agua subterránea se presentan en las zonas central y sur del acuífero (niveles por debajo de 1700 msnm). El cono de abatimiento localizado en la zona central se aprecia más extenso en diciembre de 2001 (Fig. 7.15), los niveles más bajos en esta zona se encuentran por debajo de 1680 msnm. En esos años, los niveles son menores a 1670 msnm en la zona suroeste. A partir de este momento, los niveles empiezan a recuperarse en las zonas con los abatimientos más pronunciados, pero desde el 2003 un muy importante cono de abatimiento se forma en la zona sur del acuífero con niveles por debajo de 1670 msnm.

Varianzas del error en la estimación

En los números pares de las Figuras 7.13 - 7.20 se presentan las varianzas del error en la estimación obtenidas en tiempo real para cada año, con la red de monitoreo propuesta. El orden de prioridad se incluye también.

Podemos ver que las varianzas tienen prácticamente la misma distribución espacial para todos los tiempos de monitoreo. Los valores mínimos se sitúan alrededor del centro del acuífero donde la mayor parte de los pozos se encuentran ubicados, y las máximas varianzas se localizan en el margen del acuífero porque allí solo se encuentran pocos pozos de monitoreo. Sin embargo, las varianzas tienen valores altos (entre 300 y 1700 m²). Es importante mencionar que estos valores son muy cercanos a los valores mínimos posibles que pueden obtenerse con la actual posición de los pozos de monitoreo. Estos mapas de varianza junto con criterios prácticos e hidrogeológicos pueden ser muy útiles en la elección de nuevas posiciones de monitoreo en las zonas con las máximas varianzas.

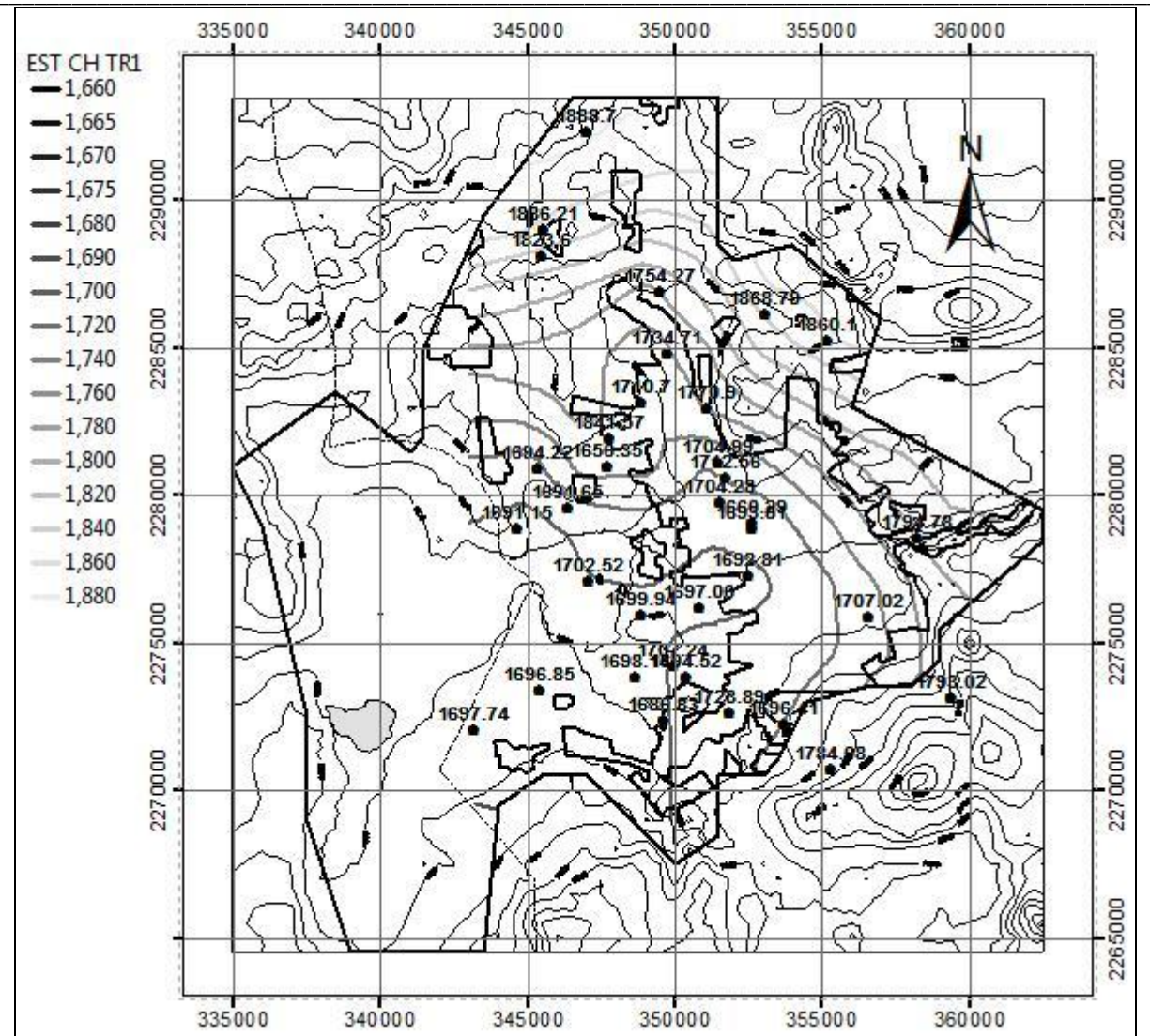


Fig. 7.13 Estimación del nivel del agua subterránea para diciembre de 1997

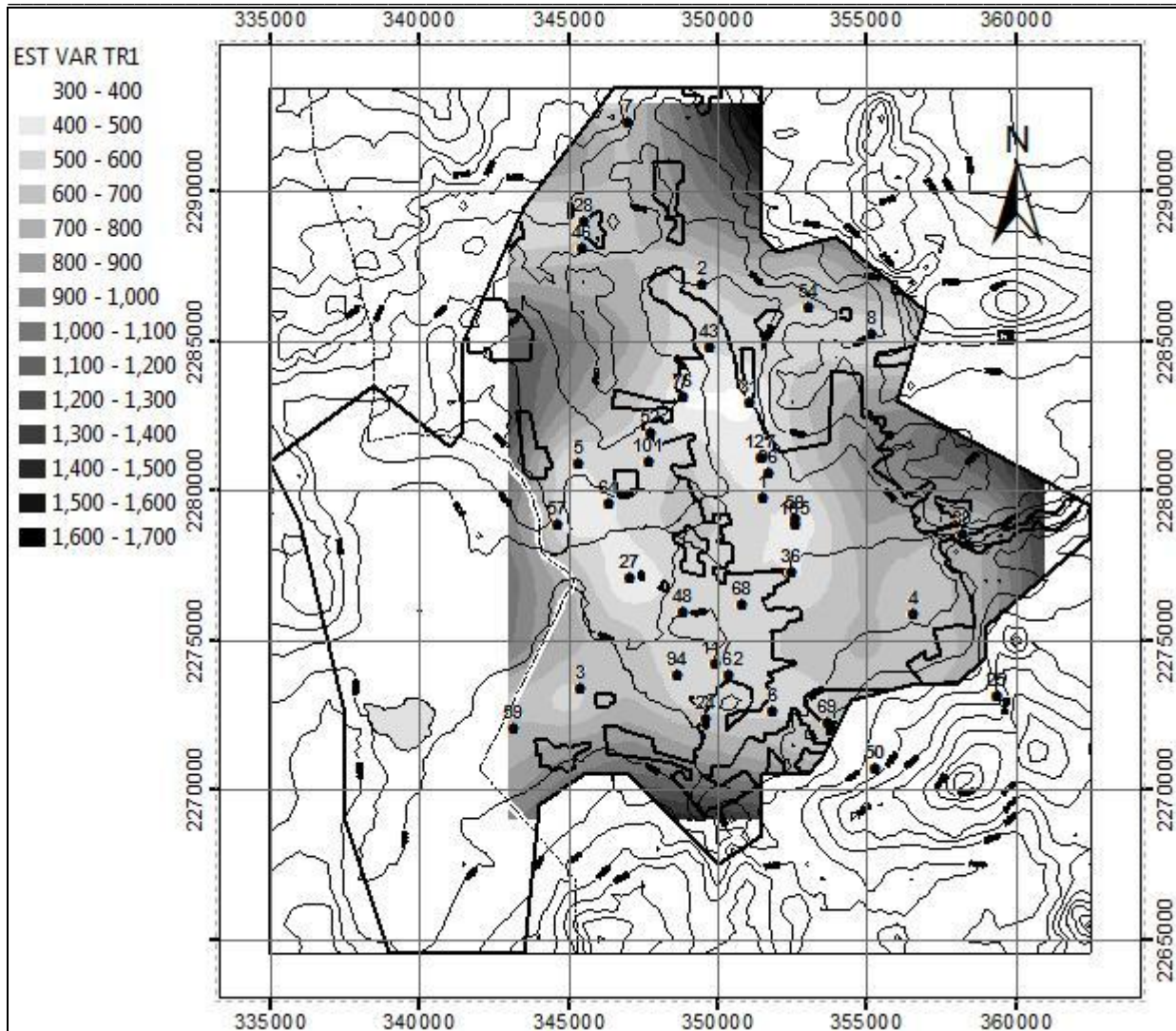


Fig. 7.14 Varianzas del error en la estimación y orden de prioridad de los pozos para diciembre de 1997

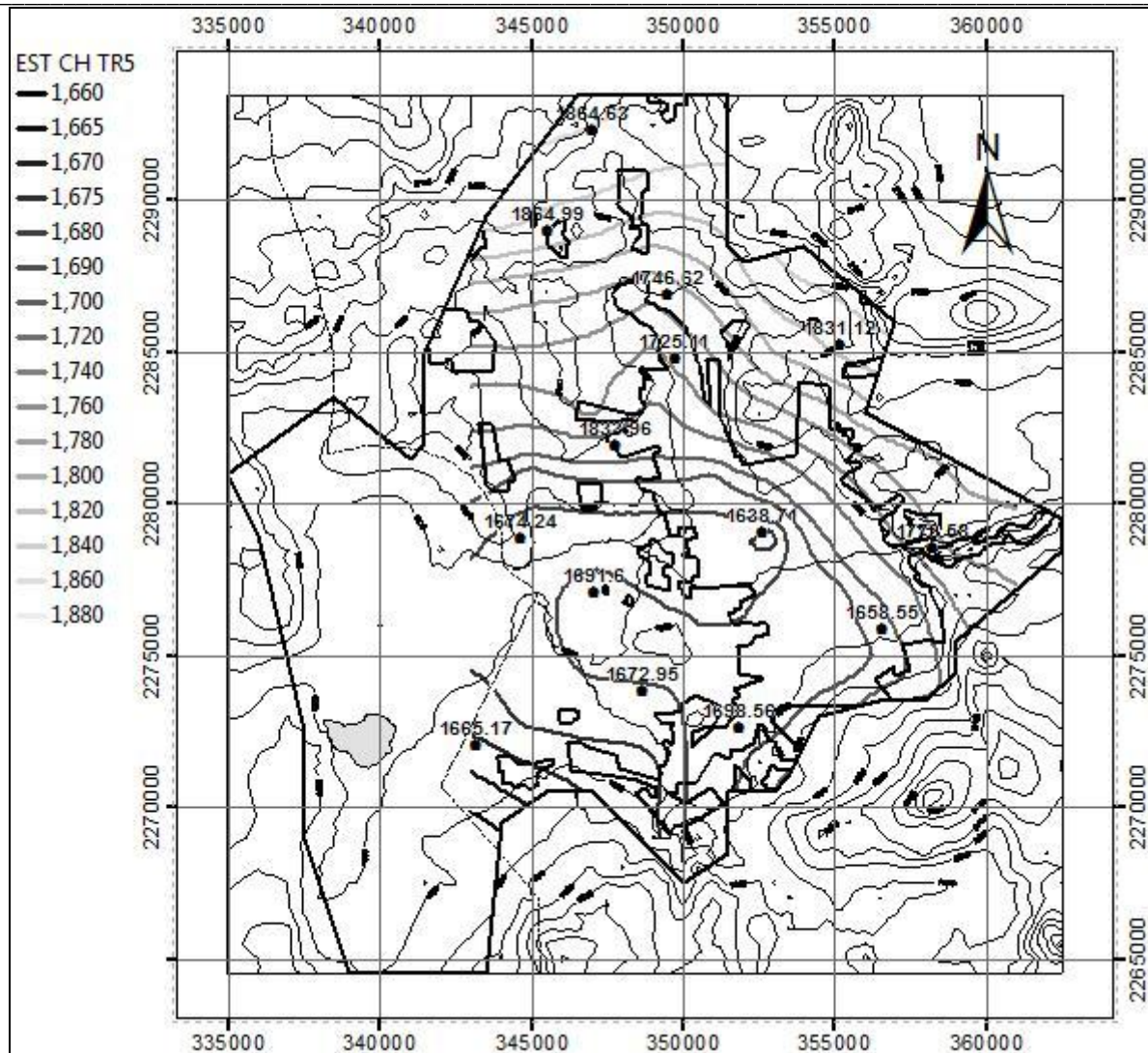
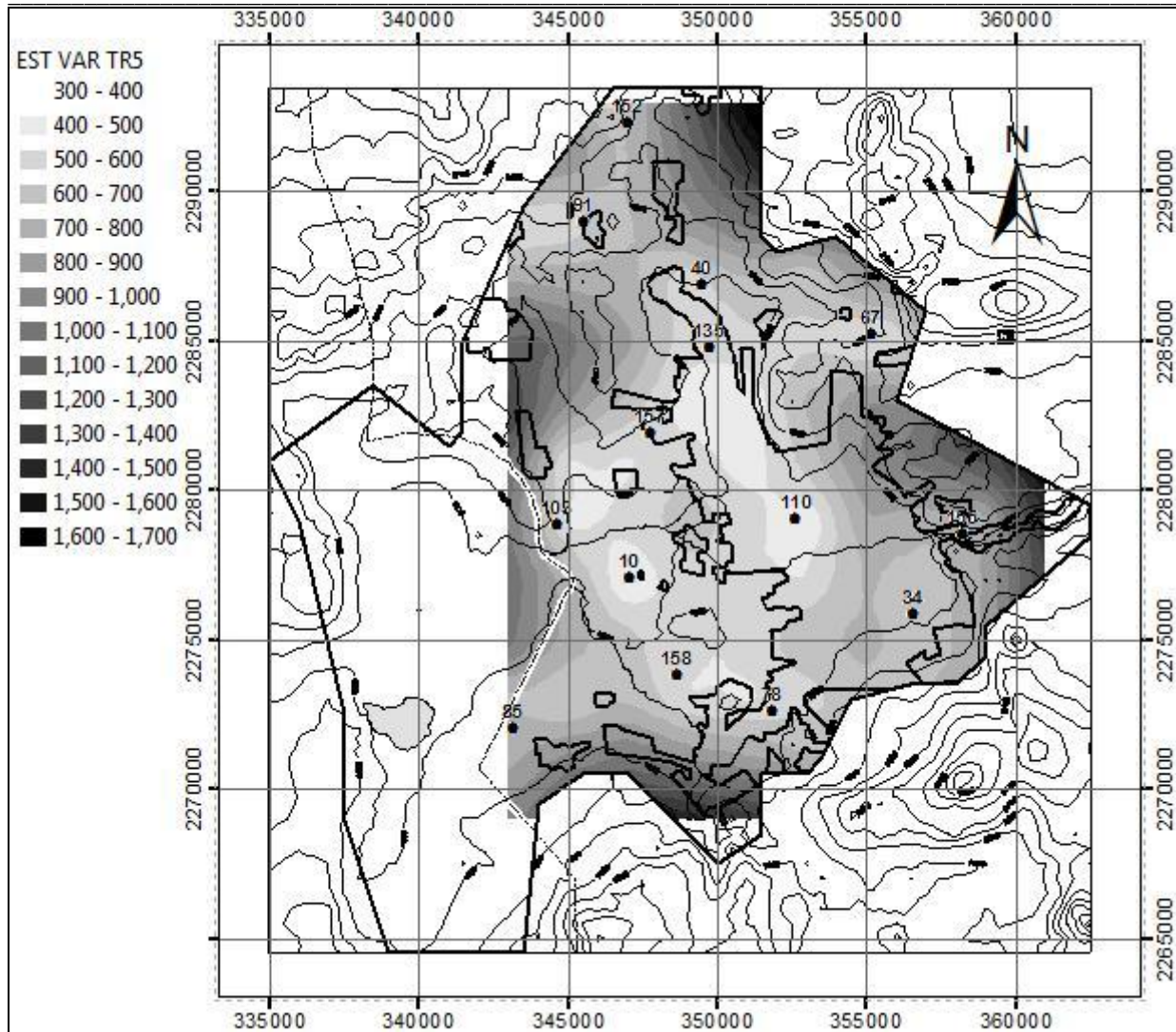


Fig. 7.15 Estimación del nivel del agua subterránea para diciembre de 2001



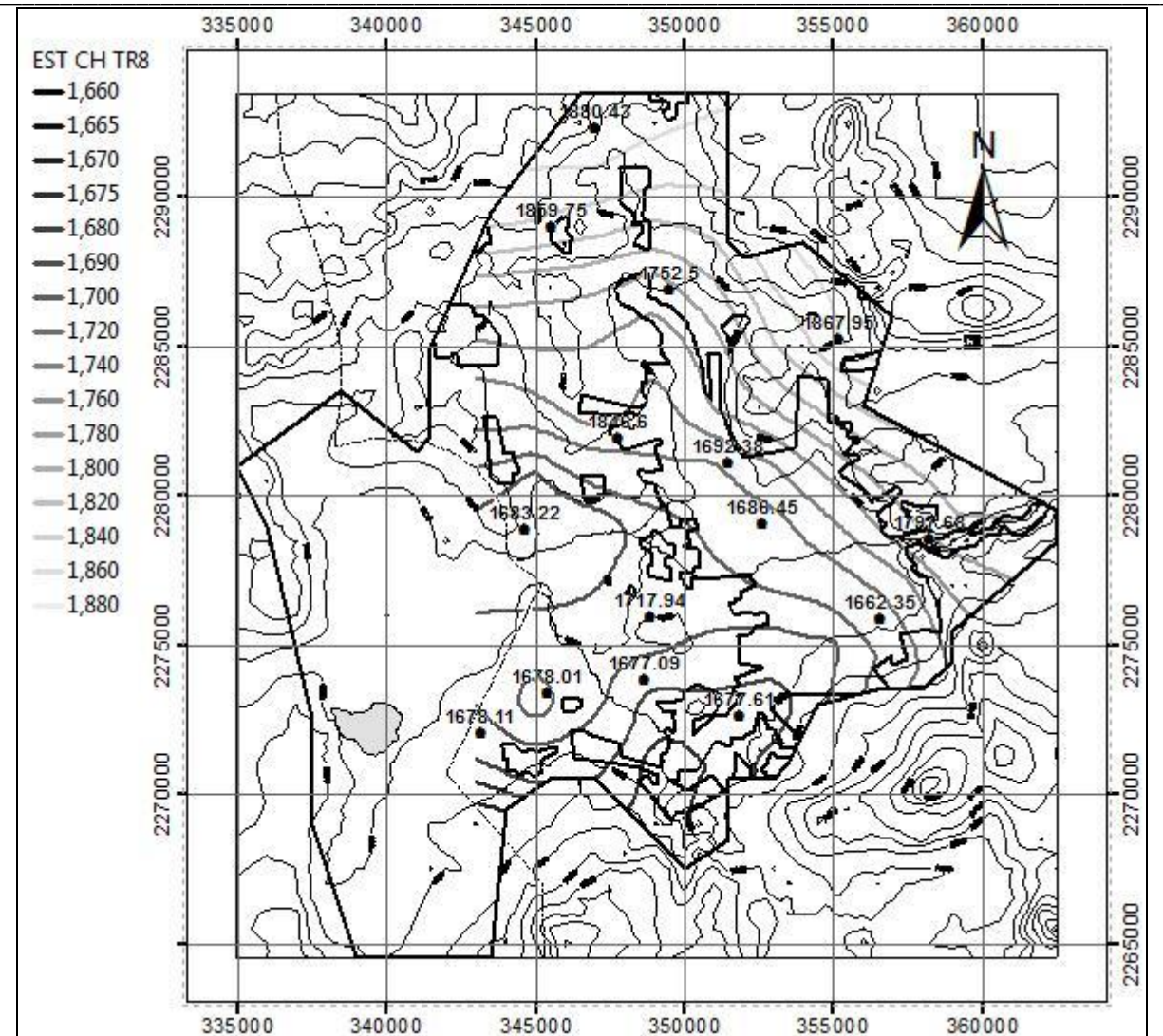


Fig. 7.17 Estimación del nivel del agua subterránea para diciembre de 2004

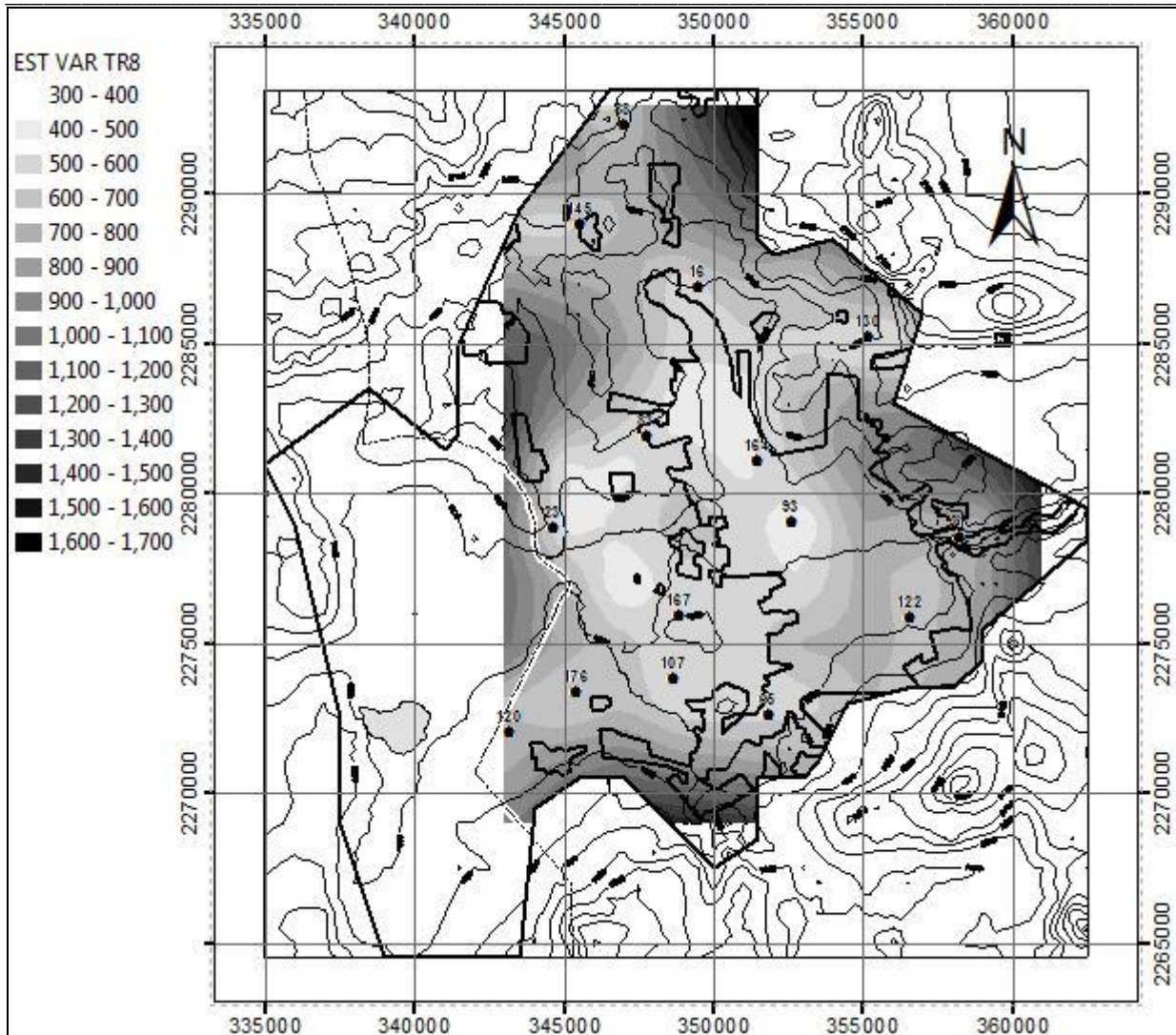


Fig. 7.18 Varianzas del error en la estimación y orden de prioridad de los pozos para diciembre de 2004

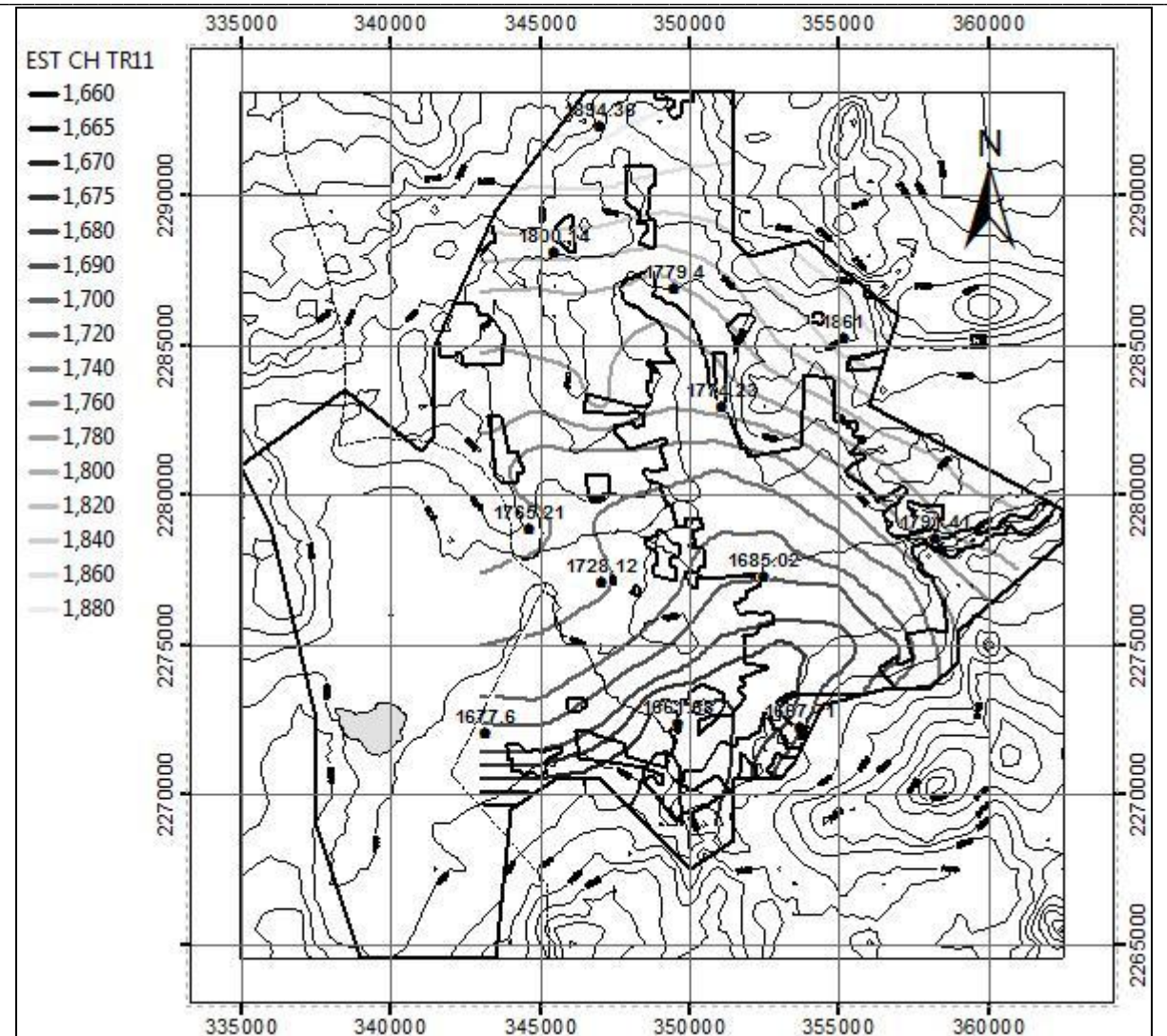


Fig. 7.19 Estimación del nivel del agua subterránea para noviembre 2007

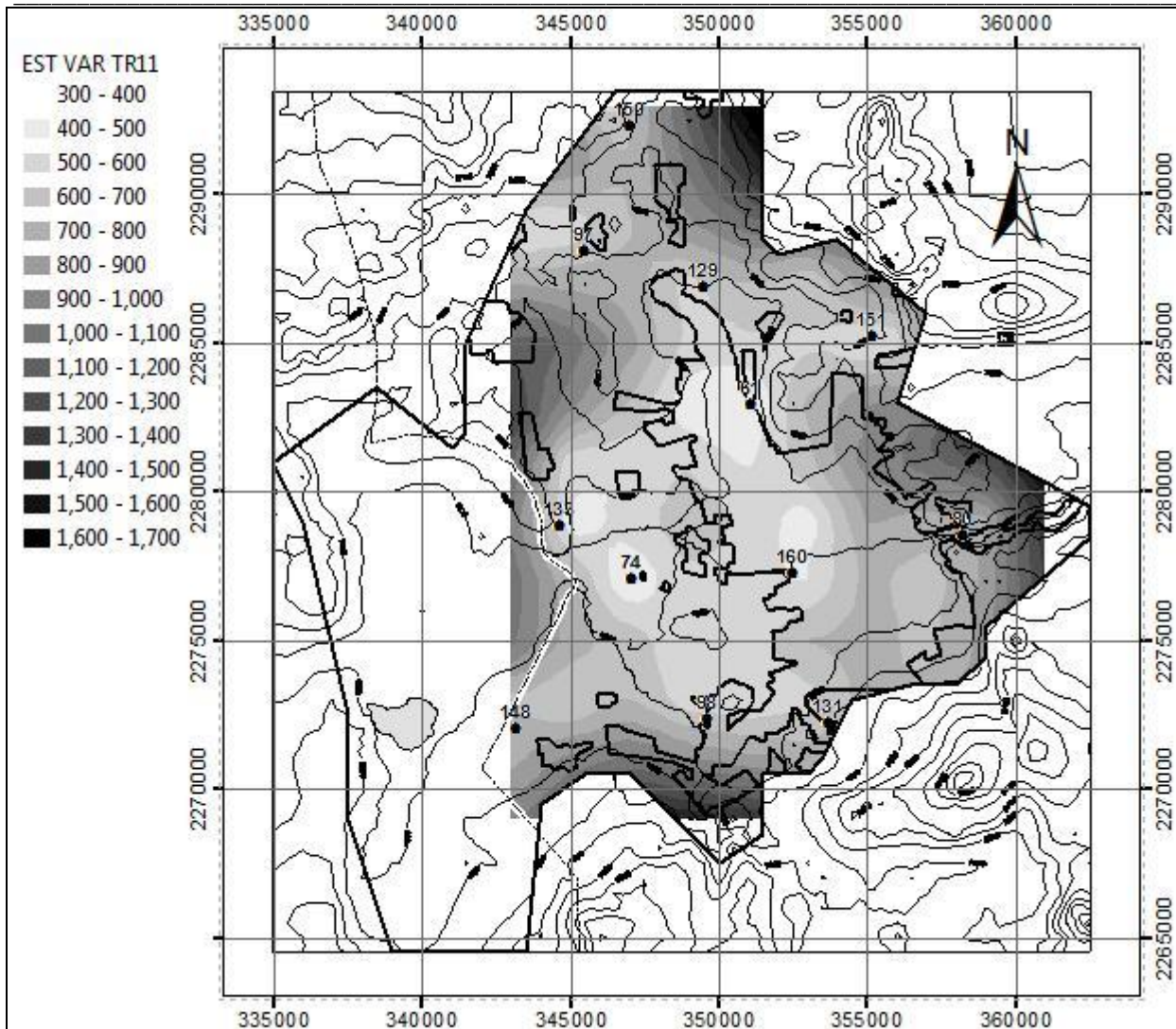


Fig. 7.20 Varianzas del error en la estimación y orden de prioridad de los pozos para noviembre de 2007

7.8 Conclusiones y recomendaciones

Los resultados muestran que la raíz cuadrada de la varianza promedio del error en la estimación del nivel del agua subterránea en el área de estudio, cuando se monitorean las 418 posiciones de monitoreo espacio-temporal disponibles, es menor a 27.27 m, y que cuando se monitorean las 178 posiciones con la más alta prioridad es menor a 27.53 m. Esto significa que las últimas 240 posiciones de monitoreo espacio-temporal son redundantes ya que la información que ellas proveen reduce muy poco el valor de la raíz cuadrada de la varianza promedio. Estos resultados muestran que el método propuesto es exitoso en propagar la información en el espacio y en el tiempo.

La mayoría de las posiciones seleccionadas por el método propuesto para el diseño óptimo de redes de monitoreo espacio-temporal corresponden al primer tiempo disponible dentro del periodo de diseño, porque en el proceso de optimización el efecto del monitoreo en un pozo se

considera únicamente para tiempos presentes y futuros, lo que hace que el primer tiempo de monitoreo provea la mayor cantidad de información del nivel del agua subterránea en el periodo de monitoreo.

Los grandes valores de la varianza total del error en la estimación del nivel del agua subterránea después de medir en todas las posiciones de monitoreo espacio-temporal indican que la actual red de monitoreo es deficiente para proveer la información que permita representar con alta certidumbre la evolución espacio-temporal del nivel del agua subterránea en el acuífero. Esto quiere decir que nuevos pozos deben ser agregados a la red de monitoreo actual, en zonas con los valores más altos de la varianza del error en la estimación. Para proponer la ubicación de nuevos pozos, también deben usarse criterios prácticos e hidrogeológicos.

La tendencia se encuentra usualmente presente en datos del nivel del agua subterránea por los gradientes hidráulicos que participan en el fenómeno del flujo del agua subterránea; en este trabajo, como en muchos anteriores (Samper y Carrera, 1990; Kumar *et al.*, 2005; Ahmadi y Sedghamiz, 2007; Mendoza y Herrera, 2007; Mendoza, 2008), para eliminar la tendencia se trabajó con la técnica llamada kriging residual. Algunos autores (ver por ejemplo Webster y Oliver, 2007) discuten que existen dos desventajas en el uso del kriging residual. Primero, la tendencia es usualmente estimada mediante mínimos cuadrados ordinarios (MCO), con los cuales se obtienen estimaciones insesgadas, pero no con varianza mínima a menos que los sitios de muestreo hayan sido seleccionados a través de un plan de muestreo. La segunda desventaja es que las semivarianzas calculadas con los residuos son sesgadas. En trabajo futuro, debe realizarse una evaluación del efecto de este problema en las estimaciones del filtro de Kalman estático, y si es necesario, usar un método diferente para remover la tendencia de información del nivel del agua subterránea.

8. CONCLUSIONES Y RECOMENDACIONES

La metodología propuesta para el diseño óptimo de redes de monitoreo de la carga hidráulica incorpora la correlación espacio-temporal entre los datos históricos medidos de esta variable en un acuífero, usando un modelo de semivariograma espacio-temporal obtenido de un análisis geoestadístico. Una ventaja importante del método es que considera la correlación cruzada entre las cargas hidráulicas en diferentes posiciones y tiempos, lo que permite una evaluación más completa de las posiciones y/o tiempos de monitoreo redundantes.

Las estimaciones obtenidas con las redes de monitoreo y sus programas de monitoreo propuestos ofrecen niveles de información muy cercanos al que proporciona el monitoreo en todas las posiciones de monitoreo espacio-temporal disponibles, por lo que el método propuesto es exitoso en propagar la información en el espacio y en el tiempo. Esto se traduce en una reducción significativa del costo asociado al monitoreo si consideramos que éste es proporcional al número de posiciones monitoreadas.

El semivariograma espacio-temporal muestral queda mejor definido conforme se cuenta con una mayor cantidad de información espacio-temporal de la variable de estudio. Al reducir el número de datos en el análisis geoestadístico se obtiene un modelo de semivariograma espacio-temporal aparentemente más limitado.

La fecha en que se propone el mayor monitoreo para cada una de los casos presentados corresponde al primer tiempo disponible dentro de su correspondiente periodo de diseño, esto se debe principalmente a que en la reducción de la varianza total del error en la estimación para ese tiempo no participan ni datos de monitoreos previos, ni de monitoreos posteriores. En todos los casos resulta prioritario estimar adecuadamente primero la distribución espacial de la carga hidráulica, las estimaciones se irán adecuando a través del tiempo conforme se vaya monitoreando en las posiciones seleccionadas de fechas posteriores; cuando la varianza temporal es muy pequeña con respecto a la espacial se requerirá monitorear pocos pozos a través del tiempo.

Valores grandes de la varianza total después de medir la carga hidráulica en todas las posiciones de monitoreo espacio-temporal disponibles indican que la red de monitoreo es muy deficiente en proveer la información suficiente para representar con alta certidumbre la evolución espacio-temporal de la carga hidráulica en el acuífero. Esto quiere decir que nuevos pozos deben ser agregados a la red de monitoreo, en zonas con los valores más altos. Para proponer la ubicación de los nuevos pozos, también deben usarse criterios prácticos e hidrogeológicos, éstos dependerán de los objetivos del monitoreo pudiendo darse mayor prioridad según sea el caso a: zonas identificadas como de recarga y descarga, zonas donde se presenten cambios de uso de suelo, donde se tengan consideradas fronteras hidrológicas o físicas dentro de un modelo conceptual (como puede ser alguna corriente superficial, fallas, etc.), zonas de mayor explotación, etc.

La evaluación en el futuro de las estimaciones de la carga hidráulica obtenidas empleando las redes de monitoreo y los programas de monitoreo óptimos propuestos muestran que una red de monitoreo bien distribuida en el espacio, captura adecuadamente la evolución de la carga hidráulica en un periodo de tiempo igual al periodo de diseño; mientras que las estimaciones obtenidas por redes de monitoreo pobremente distribuidas en el espacio pueden deteriorarse a través del tiempo, aún si se monitorea con una alta frecuencia en las posiciones disponibles, si el comportamiento en zonas no cubiertas por la red es contrario al de la mayoría de los pozos que

tienen datos de la carga hidráulica (por ejemplo, recuperación del nivel en un caso, y descenso en el otro).

Una red de monitoreo bien distribuida en el espacio (en la que se tienen pozos que cubren toda la zona de estudio) permite estimar de manera precisa la distribución espacial del nivel del agua subterránea en la parte de la unidad acuífera estudiada a pesar de tenerse configuraciones muy complicadas de capturar, como el caso en el que se presentan abatimientos locales significativos.

Si el nivel de información que proporciona la red de monitoreo óptima es el deseado, se propone aplicar el programa de monitoreo óptimo propuesto en el futuro para un periodo de tiempo igual al del diseño, siempre y cuando las condiciones del acuífero no cambien significativamente, es decir, la red y su programa tendrán que rediseñarse para los casos donde se presenten cambios de uso de suelo o de los volúmenes de extracción de agua subterránea que propicien cambios radicales en la configuración de la distribución de la carga hidráulica en el acuífero.

En el caso de que se tengan identificados diferentes sistemas de flujo en un acuífero, es necesario el diseño de una red de monitoreo para cada uno de éstos, como lo hace Hergt (2009) para el diseño de redes de monitoreo de la calidad del agua subterránea.

La correcta interpretación de mediciones de la carga hidráulica involucra el conocimiento de las variaciones en las características fisicoquímicas del agua en las tres dimensiones, para estandarizar las mediciones del nivel del agua subterránea. Esto se debe a que en el caso de tenerse densidades diferentes en el agua en el plano vertical, los niveles medidos no corresponden a la carga hidráulica.

Se demostró que las técnicas geoestadísticas combinadas con estimación espacio-temporal empleando el filtro de Kalman son muy eficientes para los casos en que se cuenta con un alto nivel de información espacio-temporal de la variable de interés ya que las estimaciones de la carga hidráulica resultaron ser muy cercanas a las simulaciones de un modelo numérico de flujo, a pesar de no emplear las ecuaciones que describen la física del problema. Este resultado confirma su potencial en el diseño y evaluación del desempeño de una red.

Es posible extender la metodología al diseño de redes de monitoreo en 3D, sería necesario contar con información multinivel de los niveles potenciométricos en la vertical y de las características fisicoquímicas del agua para una adecuada estimación de la carga hidráulica.

La tendencia se encuentra usualmente presente en la carga hidráulica porque los gradientes hidráulicos participan en el fenómeno del flujo del agua subterránea; en este trabajo, como en muchos anteriores (Samper y Carrera, 1990; Kumar *et al.*, 2005; Ahmadi y Sedghamiz, 2007; Mendoza y Herrera, 2007; Mendoza, 2008), la tendencia fue removida a través de la técnica llamada kriging residual. Algunos autores (ver por ejemplo Webster y Oliver, 2007) discuten que existen dos desventajas en el uso del kriging residual. Primero, la tendencia es usualmente estimada mediante mínimos cuadrados ordinarios (MCO), con los cuales se obtienen estimaciones insesgadas, pero no con varianza mínima a menos que los sitios de muestreo hayan sido seleccionados a través de un plan de muestreo. La segunda desventaja es que las semivarianzas calculadas con los residuos son sesgadas. En trabajo futuro, debe realizarse una evaluación del efecto de este problema en las estimaciones del filtro de Kalman, y si es necesario, usar un método diferente para remover la tendencia de los datos de la carga hidráulica.

9. BIBLIOGRAFÍA

Ahmadi, S. H., & Sedghamiz, A. (2007). Geostatistical analysis of spatial and temporal variations of groundwater level. *Environmental Monitoring and Assessment*, 129, 277-294.

Ahmadi, S. H., & Sedghamiz, A. (2008). Application and evaluation of kriging and cokriging methods on groundwater depth mapping. *Environmental Monitoring and Assessment*, 138, 357-368.

Ahmed, S., & De Marsily, G. (1993). Cokriged estimation of aquifer transmissivity as an indirect solution of the inverse problem: A practical approach. *Water Resources Research*, 29(2), 521-530.

Andricevic, R. (1990). A Real-Time Approach to Management and monitoring of Groundwater Hydraulics. *Water Resources Research*, 26(11), 2747-2755.

Babu, D. K., Pinder, G. F., Niemi, A., Ahlfeld, D. P., & Stothoff, S. A. (1997). *Chemical transport by three-dimensional groundwater flows*, Princeton, NJ, USA.

Bear, J. (1979). *Hydraulics of groundwater*, Ed. McGraw-Hill, NY, USA.

Becerra-Soriano, L. (2009). *Estimación de lluvia en el Distrito Federal utilizando datos de pluviógrafos y de radar meteorológico*, Tesis de Maestría, UNAM, México.

Bierkens, M., Knotters, M., & Hoogland, T. (2001). Space-time modeling of water table depth using a regionalized time series model and the Kalman filter. *Water Resources Research*, 37(5), 1277-1290.

Bilonick, R. A. (1985). The space-time distribution of sulfates deposition in the northeastern United States. *Atmospheric Environment*, 9(11), 1829-1985.

Bravo, J. A. (2005). *Diseño de una red de monitoreo para evaluar el comportamiento del acuífero del Valle de Querétaro durante la operación de la presa Extóraz*. Tesis de maestría, UNAM, México.

Cameron, K., & Hunter, P. (2000). *Optimization of LTM Networks Using GTS: Statistical Approaches to Spatial and Temporal Redundancy*. Tech. Rep., Air Force Center for environmental excellence, Brooks AFB, TX, USA.

Carabias, J., Landa, R., Collado, J., & Martínez, P. (2005). *Agua, medio ambiente y sociedad. Hacia la gestión integral de los recursos hídricos en México*. Universidad Nacional Autónoma de México. El Colegio de México A. C., Fundación Gonzalo Río Arronte, I. A. P., México.

- Carrillo-Rivera, J. J. (2000). Application of the groundwater-balance equation to indicate interbasin and vertical flow in two drainage basins, Mexico. *Hydrogeology Journal*, 8(5), 503-520.
- Carrillo-Rivera, J. J., Cardona, A., & Moss, D. (1996). Importance of the vertical component of groundwater hydrogeochemical approach in the valley of San Luis Potosí, México. *Journal of hydrology*, 185, 23-44.
- CFE (Comisión Federal de Electricidad), Cruickshank, C., Berezowsky, M., Domínguez, M., & Padilla, J. (1983). *Manual de diseño de obras civiles, Hidrotecnia-Hidrología-Geohidrología*, CFE, México.
- Chávez, R. (1986). *La necesidad de plantear el aprovechamiento de las aguas subterráneas en México*, inédito, México.
- Christakos, G. (2000). *Modern Spatiotemporal Geostatistics*, Oxford, UK.
- Cressie, N. A. (1991). *Statistics for spatial data*, Wiley, NY, USA.
- De Cesare, L., Myers, D. E. & Posa, D. (1997). Spatial-temporal modeling of SO₂ in the Milan district. *Geostatistics Wollong '96*, Baafi, E. Y., & Schofield, N. A. (eds), Kluwer academic publishers.
- De Cesare, L., Myers, D. E., & Posa, D. (2001). Estimating and modeling space-time correlation structures. *Statistics and Probability Letters*, 51, 9-14.
- De Cesare, L., Myers, D. E., & Posa, D. (2002). Fortran programs for space-time modeling. *Computers & Geosciences*, 28, 205-212.
- De Iaco, S., Myers, D. E., & Posa, D. (2001). Space-time analysis using a general product-sum model. *Statistics & Probability Letters*, 52, 21-28.
- De Iaco, S., Myers, D. E., & Posa, D. (2002). Space-time variograms and a functional form of total air pollution measurements. *Computational Statistics & Data Analysis*, 41, 311-328.
- De Iaco, S., Myers, D. E., & Posa, D. (2003). The linear coregionalization model and the product-sum space-time variogram. *Mathematical Geology*, 35(1), 25-38.
- Delhomme, J.P. (1978). Kriging in the hydrosiences. *Advances in Water Resources*, 1(5), 251-266.
- Deutsch, C. V., & Journel, A. G. (1997). *GSLIB: Geostatistical Software Library and User's Guide*, Second Edition, Oxford University Press, NY, USA.
- Díaz, M. A. (2002). *Geoestadística Aplicada*, Instituto de Geofísica, UNAM, Instituto de Geofísica y Astronomía, CITMA, Cuba.

- Ely, D. M., Hill, M. C., Tiedeman, C. R., & O'Brien, G. M. (2000). Evaluating observations in the context of predictions for the Death Valley regional groundwater system. *Proceedings of the 2000 Joint Conference on Water Resources Engineering and Water Resources Planning and Management, Minneapolis, MN, compact disk, American Society of Civil Engineers, Washington, DC, USA.*
- Faisal, K. Z., Ahmed, S., Dewandel, B., & Maréchal, J. (2007). Optimizing a piezometric network in the estimation of the groundwater budget: a case study from a crystalline-rock watershed in southern India. *Hydrogeology Journal*, 15(6), 1131-1145.
- Fetter, C. W. (1994). *Applied hydrogeology*, Tercera edición, Ed. Prentice-Hall, NJ, USA.
- Freeze, R. A., & Cherry, J. A. (1979). *Groundwater*, Prentice-Hall Inc., Englewood Cliffs, NJ, USA.
- Gambolati, G., & Volpi, G. (1979). Groundwater contour mapping in Venice by stochastic interpolators: 1. Theory. *Water Resources Research*, 15(2), 281-290.
- Gangopadhyay, S., Gupta, A. S., & Nachabe, M. H. (2001). Evaluation of Ground Water Monitoring Network by Principal Component Analysis. *Ground Water*, 39(2), 181-191.
- Gelb, A., Kasper, J. F., Nash, R. A., Price, C. F., & Sutherland, A. A. (1996). *Applied Optimal Estimation*, THE M.I.T. PRESS, Massachusetts, USA.
- Gleick, P. H. (1996). *Encyclopedia of Climate and Weather*, Oxford University Press, NY, USA.
- Hergt, T. (2009). *Diseño optimizado de redes de monitoreo de la calidad del agua de los sistemas de flujo subterráneo en el acuífero 2411 "San Luis Potosí": Hacia un manejo sustentable*, Tesis de doctorado, Universidad Autónoma de San Luis Potosí, 2009.
- Herrera, G. S. (1998). *Cost Effective Groundwater Quality Sampling Network Design*, Ph. D. Dissertation, University of Vermont, USA.
- Herrera, G.S., Guarnaccia, J., Pinder, G.F., & Simuta, R. (2001). Design of Efficient Space-Time Groundwater Quality Sampling Networks. *Proceedings of the 2001 International Symposium on Environmental Hydraulics, ISEH*.
- Herrera, G. S., Júnez-Ferreira, H. E., González, L., & Cardona, A. (2004). Diseño de una red de monitoreo de la calidad del agua para el acuífero Irapuato-Valle, Guanajuato. *Memorias del XVIII Congreso Nacional de Hidráulica, AMH, SLP, México.*
- Herrera, G. S., & Pinder, G. F. (2005). Space-time optimization of groundwater quality sampling networks. *Water Resources Research*, 41, W12407, 15 pp.

- Herrera, G. S., Simuta, R., & Bravo, J. A. (2007) Modelo de flujo del agua subterránea y diseño de la red de monitoreo para el acuífero del valle de Querétaro. *Memorias del 4to Congreso Internacional, 2do Congreso Nacional de Métodos Numéricos en Ingeniería y Ciencias Aplicadas*, UMSNH, México.
- Hill, M. C., Ely, M. D., Tiedeman, C. R., D'Agnesse, F. A., Faunt, C. C., & O'Brien, B. A. (2000). *Preliminary evaluation of the importance of existing hydraulic-head observation locations to advective-transport prediction, Death Valley regional flow system, California and Nevada. U.S. Geological Survey Water-Resources Investigations Report 00-4282*, <http://pubs.usgs.gov/wri/wri004282/book/wri004282.pdf>. *accesado el 3 de septiembre de 2009*.
- IMTA (Instituto Mexicano de Tecnología del Agua), Herrera, G., Martínez, M., Flores, G., Simuta, R., Bravo, A. & Mendoza, E.Y. (2005). *Evaluación del impacto regional en los niveles del agua del acuífero del valle de Querétaro debido a la puesta en marcha de la presa Extóraz*, Comisión Nacional del Agua, México.
- Journel, A. G., & Huijbregts, C. J. (1978). *Mining geostatistics*, Academic Press, London, UK.
- Júnez-Ferreira, H. E. (2005). *Diseño de una red de monitoreo de la calidad del agua para el acuífero Irapuato-Valle, Guanajuato*, Tesis de Maestría, UNAM, México.
- Kashyap, D., & Vakkalagadda, R. (2009). New Model of Variogram of Groundwater Hydraulic Heads. *Journal of Hydrologic Engineering*, 14(8), 872-875.
- Kawecki, M. W. (1995). Correction for temperature effect in the recovery of a pumped well. *Ground Water*, 33(6), 917-926.
- Kemal, S. G., & Guney, I. (2007). Spatial analyses of groundwater levels using universal kriging. *Journal of Earth System Science*, 116(1), 49-55.
- Kresic, N. (2007). *Hydrogeology and groundwater modeling*, Segunda edición, Ed. Taylor and Francis Group, Boca Raton, FL, USA.
- Kumar, S., Sondhi, S. K., & Phogat, V. (2005). Network design for groundwater level monitoring in Upper Bari Doab canal tract, Punjab, India. *Irrigation and Drainage*, 54, 431-442.
- Kyriakidis, P. C., & Journel, A. G. (1999). Geostatistical Space-Time Models: A Review. *Mathematical Geology*, 31 (6), 651-684.
- Lin, Y., & Rouhani, S. (2001). Multiple-Point Variance Analysis for optimal adjustment of a monitoring network. *Environmental Monitoring and Assessment*, 69, 239-266.
- Loaiciga, H. A. (1989). An optimization approach for groundwater quality monitoring network design. *Water Resources Research*, 25(8), 1771-1782.

- Loaiciga, H. A., Charbeneau, R. J., Everett, L. G., Fogg, G. E., & Hobbs, B. F. (1992). Review of groundwater quality monitoring network design. *Journal of Hydraulic Engineering*, 118(1), 11-32.
- López, C. y J. Samper. (1997). Numerical aspects of the universal kriging method for hydrological applications. *GeoENV I - Geostatistics for environmental applications. Proceedings*, 65-76, Lisboa, Portugal.
- Maderey-Rascón, L. E., & Carrillo-Rivera, J. J. (2005). *El Recurso Agua en México: Un análisis geográfico*, Instituto de Geografía, UNAM, México.
- Masoumi, F., & Kerachian, R. (2010). Optimal redesign of groundwater quality monitoring networks: a case study. *Environmental Monitoring and Assessment*, 161, 247-257.
- Mendoza, E. Y., & Herrera, G. S. (2007). Estimación multivariada espacio-tiempo de la carga hidráulica en el Valle de Querétaro-Obrajuelo. *Ingeniería hidráulica en México*, 22(1), 63-80.
- Mendoza, E. Y. (2008). *Análisis de alternativas para la estimación de la carga hidráulica utilizando métodos geoestadísticos en espacio y espacio-tiempo*, Tesis de doctorado, UNAM, México.
- Mogheir, Y., Singh, V. P., & de Lima, J. L. M. P. (2006). Spatial assessment and redesign of a groundwater quality monitoring network using entropy theory, Gaza Strip, Palestine. *Hydrogeology Journal*, 14, 700-712.
- Myers, D. E., & Journel, A. (1990). Variograms with zonal anisotropies and noninvertible kriging systems. *Mathematical Geology*, 22(7), 779-785.
- Nejadkoorki, F., Nicholson, K., & Hadad, K. (2011). The design of long-term air quality monitoring networks in urban areas using a spatiotemporal approach. *Environmental Monitoring and Assessment*, 172, 215-223.
- Nunes, L. M., Cunha, M. C., & Ribeiro, L. (2004a). Groundwater monitoring network optimization with redundancy reduction. *Journal of Water Resources Planning and Management*, 1(33), 33-43.
- Nunes, L. M., Cunha, M. C., & Ribeiro, L. (2004b). Optimal space-time coverage and exploration costs in groundwater monitoring networks. *Environmental Monitoring and Assessment*, 93, 103-124.
- Rodríguez-Iturbe, I., & Mejía, J. M. (1974). The design of rainfall networks in time and space. *Water Resources Research*, 10(4), 713-728.
- Rouhani, S. (1985). Variance reduction analysis. *Water Resources Research*, 21(6), 837-846.

- Rouhani, S., & Hall, T. (1989). *Space-time kriging of groundwater data*, *Geostatistics*, Ed. M. Armstrong, vol. 2: Kluwer academic publ., Dordrecht, Holanda.
- Rouhani, S., & Myers, D.E. (1990). Problems in space-time kriging of geohydrological data, *Mathematical Geology*, 22(5), 611-623.
- Samper, F. J., & Carrera, J. (1990). *Geoestadística, aplicaciones a la hidrogeología subterránea*, Centro Internacional de Métodos Numéricos en Ingeniería, UPC, Barcelona, España.
- Stein, M. L. (1986). A simple model for spatial-temporal processes. *Water resources research*, 22(13), 2107-2110.
- Stein, M. L. (1999). *Interpolation of spatial data: Some theory for Kriging*, Springer-Verlag, NY, USA.
- Storck, P., Valocchi, A. J. & Eheart, J. W. (1995). Optimal location of monitoring wells for detection of groundwater contamination in three-dimensional heterogeneous aquifers. *Models for Assessing and monitoring groundwater quality (Proceedings of a Boulder Symposium. IAHS)*, 227, 39-47.
- Szucs, A. & Jordan, G. (1994). Analysis of sampling frequency in groundwater quality monitoring systems: a case study. *Water Science Technology*, 30(10), 73-78.
- Ta'any, R. A. & Tahboub, A. B. (2009). Geostatistical analysis of spatiotemporal variability of groundwater level fluctuations in Amman-Zarqa basin, Jordan: a case study. *Environmental Geology*, 57, 525-535.
- Tonkin, M. J. & Larson, S. P. (2002). Kriging water levels with a regional-linear and point-logarithmic drift. *Ground Water*, 40(2), 185-193.
- Van Geer, F. C., Te Stroet, C. B. M., & Yangxiao, Z. (1991). Using Kalman Filtering to Improve and Quantify the Uncertainty of Numerical Groundwater Simulations: 1. The Role of System Noise and Its Calibration. *Water Resources Research*, 27(8), 1987-1994.
- Volpi, G., Gambolati, G., Carbognin, L., & Mozzi, G. (1979). Groundwater contour mapping in Venice by stochastic interpolators: 2. Results. *Water Resources Research*, 15(2), 291-297.
- Webster, R., & Oliver, M. (2007). *Geostatistics for environmental scientists*, Wiley, UK.
- Wu, Y. (2003). Optimal design of a groundwater monitoring network in Daqing, China. *Environmental Geology*, 45, 527-535.
- Yangxiao, Z., C. B. Te Stroet, & Van Geer, F. (1991). Using Kalman filtering to improve and quantify the uncertainty of numerical groundwater simulations: 2. Application to monitoring network design. *Water Resources Research*, 27(8), 1995-2006.

Zhang, Y., Pinder, G. F., & Herrera, G. S. (2005). Least cost design of groundwater quality monitoring networks. *Water Resources Research*, 41, W08412, 1.

ANEXO A. Revisión del estado del arte del diseño de redes de monitoreo de la calidad del agua subterránea

A continuación se presentan distintos trabajos para el diseño de redes de monitoreo de la calidad del agua subterránea revisados para esta investigación. Se presentan de acuerdo a su fecha de publicación sin considerar el marco de diseño al que pertenecen. Se incluye además un trabajo que propone una metodología para el diseño de una red de monitoreo de la calidad del agua que pudiera adaptarse para el monitoreo del agua subterránea.

Rouhani (1985) propuso un método de reducción de la varianza para aumentar una red de monitoreo existente. La red de monitoreo se aumenta agregando muestras de pozos en lugares que minimizan la varianza del error en la estimación de la concentración de un contaminante, calculada por medio de kriging. Este trabajo no involucra la frecuencia temporal de muestreo en el diseño.

Al parecer Loaiciga (1989) fue el primer autor que propuso un método para el diseño óptimo espacio-temporal de redes de monitoreo de la calidad del agua subterránea en el que la redundancia espacial y temporal de una red de monitoreo se analizan en conjunto. Para predecir la incertidumbre del error de la estimación de las concentraciones del contaminante, el método combina kriging con una ecuación de transporte estocástica. Este autor usó el método de kriging generalizado para incluir el dominio en espacio y tiempo, y de esa manera evaluar cómo una muestra tomada de un pozo dado y en un tiempo dado reduce la incertidumbre de todas las posiciones y tiempos en los que se obtendrían estimaciones. Para poder resolver el sistema de kriging es necesario contar con la matriz de covarianza espacio-temporal del error de la estimación en las posibles posiciones y tiempos de muestreo. Loaiciga deriva los elementos de la matriz de covarianza de la ecuación de advección-dispersión que gobierna el transporte de masa discretizada numéricamente. Usó un término aditivo de error aleatorio en la ecuación de transporte que considera la incertidumbre del modelo; este término cambia de un paso de tiempo a otro y en cada nuevo paso es independiente de todos los errores previos. Para la optimización de la red, Loaiciga busca determinar dónde y cuándo muestrear para minimizar la varianza del error de la concentración estimada (obtenida del sistema de kriging espacio-temporal) en las posiciones de interés, sujeto a restricciones de costo e insesgo. La determinación de un plan de muestreo óptimo se plantea como un problema de programación entera mixta.

Storck *et al.* (1995) presentaron un método de optimización estocástica multiobjetivo para el diseño de redes de monitoreo en la detección inicial de la contaminación del agua subterránea en instalaciones de disposición de desperdicios. Dada una estructura de correlación de la conductividad hidráulica, se generan realizaciones estocásticas de este parámetro para ser utilizadas en un modelo de transporte de un acuífero heterogéneo hipotético en 3D en el que se simula el fenómeno de advección de un contaminante, no así el de dispersión (esto se considera como un caso más conservador ya que al desprestigiar el fenómeno de dispersión resulta más difícil detectar el contaminante). La salida de una simulación Monte Carlo es una lista de los pozos que detectan la pluma de contaminante y el volumen contaminado del acuífero en el tiempo de la detección inicial en cada pozo. A mayor número de pozos que comprenden la red, es más fácil detectar la pluma contaminante y menor será el volumen de contaminante en el flujo subterráneo. El problema multiobjetivo es tratado usando los métodos de pesado y restricción. La función contiene 3 objetivos que son: 1. Maximizar la probabilidad de detección, 2. Costo mínimo (mínimo número de pozos de monitoreo) y 3. Minimizar el volumen de agua subterránea

contaminada al momento de ser detectada la fuga. La función objetivo formal usada en el modelo es una suma pesada de los primeros dos objetivos, mientras que el tercer objetivo es tratado como una restricción. Las estimaciones se obtienen usando un modelo numérico de las ecuaciones de flujo y transporte del agua subterránea y el método de optimización es recocado simulado. La dimensión del diseño de la red es tridimensional y además considera el tiempo. Se concluye que la práctica común de seleccionar un pozo aguas arriba y tres aguas abajo de un área contaminada resulta no ser suficiente para un adecuado monitoreo en el caso hipotético presentado en el trabajo; para tres pozos localizados aguas abajo de la fuente contaminante se tiene una probabilidad de detección menor al 50% mientras que cinco pozos localizados de manera óptima aguas abajo pueden lograr una probabilidad de detección del 80%, 10 pozos pueden llegar al 95%.

Cameron y Hunter (2000) propusieron un método para optimizar redes de monitoreo de la calidad del agua subterránea mediante la reducción de la redundancia espacial y temporal. Éste utiliza dos algoritmos: uno temporal y otro espacial. El algoritmo temporal combina series de tiempo de datos de muchos pozos para construir un semivariograma temporal compuesto con el que se definen frecuencias de muestreo. En el algoritmo espacial se asignan pesos numéricos (llamados pesos de kriging global) a la posición de los pozos para evaluar su contribución relativa a la estimación de la pluma contaminante. Se remueven los pozos que no influyen significativamente en la varianza de la nueva estimación obtenida del kriging. De esta manera, la estimación del contaminante se obtiene mediante kriging, la optimización temporal se obtiene del alcance del semivariograma temporal compuesto y la optimización espacial se obtiene mediante kriging. En este trabajo no se considera la correlación cruzada espacio-temporal de la concentración del contaminante.

Ely *et al.* (2000) y Hill *et al.* (2000) presentaron un método para medir la importancia de posiciones de observación existentes usando medidas de la incertidumbre de la predicción de contaminantes en el agua subterránea. La importancia de las posiciones de observación para las predicciones del modelo fue evaluada usando las desviaciones estándar en las predicciones, calculadas a través de una ecuación lineal de inferencia estadística, como función de la matriz de sensibilidades de las predicciones con respecto a los parámetros del modelo. Se sugiere que se supervisen con mayor frecuencia las posiciones detectadas como más importantes. En el informe de Hill *et al.* (2000) desarrollan la metodología para demostrar lo anterior usando las posiciones de observación de las cargas hidráulicas de un modelo de tres capas de flujo y transporte para el sistema acuífero de Death Valley. Las predicciones de interés son el transporte subsuperficial debajo de la montaña Yucca y 14 zonas subterráneas de prueba. Únicamente la componente advectiva de transporte se considera debido a que ésta es la componente más afectada por el sistema dinámico representado por el modelo a que se usó. El problema es tratado usando las capacidades del programa de cómputo del U.S Geological Survey, MODFLOW-2000, con su paquete ADVective-Travel Observation (ADV), y un programa de cómputo adicional desarrollado para ese trabajo. Los métodos presentados en el informe son usados de tres formas: (1) Las valoraciones de las observaciones individuales se obtienen mediante la manipulación de las medidas de incertidumbre en la predicción, y no implican volver a calibrar el modelo. En este análisis, las posiciones de observación se omiten de forma individual y se calcula el consiguiente aumento de la incertidumbre en las predicciones. La incertidumbre se cuantifica con las desviaciones estándar en el transporte advectivo simulado. El aumento de la incertidumbre se cuantifica como el porcentaje de aumento en las desviaciones estándar causado por la omisión de cada posición de observación en el cálculo de las desviaciones estándar. En general, los lugares

de observación asociados con mayores incrementos se valoraron como más importantes. (2) La valoración de un área geográfica se obtiene utilizando una extensión directa del método utilizado para la observación de lugares individuales. Este análisis es necesario en donde las observaciones se agrupan para determinar si el área es importante para las predicciones de interés. (3) Por último, el método se utiliza para evaluar la omisión de un conjunto de 100 posiciones de observación. Las posiciones fueron elegidas porque tenían bajas puntuaciones individuales, excluyendo los pocos lugares en que se midió la carga hidráulica en la parte más profunda del sistema.

Herrera *et al.* (2001) redujeron el número de muestras necesarias para tener menores gastos de operación de una red de calidad del agua subterránea, y por tanto, disminuir el costo final de un estudio dado. El objetivo de la red de monitoreo fue minimizar la varianza total del error en la estimación para un tiempo determinado en algunas zonas de riesgo. Aplicaron un filtro de Kalman en combinación con simulación estocástica del flujo subterráneo y del transporte de solutos. Con la simulación estocástica obtuvieron una primera estimación de la distribución espacio-temporal de las concentraciones del contaminante y de la varianza del error de esta estimación; luego usaron el filtro de Kalman para predecir las reducciones en la varianza de los errores resultantes de añadir sucesivamente posiciones y tiempos de muestreo. La estimación se realiza mediante el filtro de Kalman y el método de optimización es el de inclusiones sucesivas. La dimensión del diseño de la red es bidimensional y además incluyó el tiempo. Obtuvieron resultados bastante alentadores.

Lin y Rouhani (2001) presentaron un algoritmo que establece el nivel de insuficiencia y redundancia de la información de una red de monitoreo de la calidad del agua subterránea. Este algoritmo es denominado análisis de la varianza multi-punto (VMP). VMP incluye un análisis de reducción de la varianza multi-punto (RVMP) en el que se agregan nuevos sitios, y un análisis de incremento de la varianza multi-punto (IVMP) para eliminar sitios redundantes. Se determina simultáneamente la adición y la eliminación de grupos de puntos de muestreo. Estos procedimientos de optimización eliminan la necesidad de una discretización espacial o selección secuencial. La eficiencia se prueba bajo condiciones de campo. La ecuación general de RVMP se desarrolla inicialmente mediante el cómputo del impacto que tendrá la adición de un punto de muestreo sobre la estimación de la varianza (mediante kriging ordinario) en otra posición. Algo semejante se realiza para evaluar el incremento en la varianza al eliminar un punto (IVMP). Las variables utilizadas en el ejemplo de aplicación son tricloroetileno (TCE) y tetracloroetileno (PCE). El tipo de estimación es kriging y el método de optimización empleado es llamado Steepest Ascent (ascenso más pronunciado). Los resultados muestran que el método de VMP es una herramienta efectiva para el ajuste de redes de monitoreo existentes. Los autores concluyen que la RVMP óptima parece ser más eficiente que la producida por un análisis de reducción de la varianza secuencial.

Nunes *et al.* (2004a) optimizaron redes de monitoreo de la calidad del agua subterránea considerando una reducción en la redundancia espacial y/o temporal. Proponen tres modelos de optimización para seleccionar el mejor subconjunto de estaciones de una red de monitoreo grande del agua subterránea: 1. Uno que maximiza la certidumbre espacial, 2. Uno que minimiza la redundancia temporal y 3. Un modelo que maximiza la certidumbre espacial y minimiza la redundancia temporal. Los modelos de optimización propuestos se resuelven con recocido simulado, junto con un algoritmo de parametrización usando entropía estadística. Los tres modelos se derivan de una ecuación que considera dos términos: uno espacial y otro temporal, los modelos empleados en el artículo resultan de simplificaciones a la ecuación objetivo. El tipo

de estimación es kriging y el método de optimización es recocido simulado. El modelo que utiliza la información en espacio y tiempo es el que arroja los mejores resultados. La ecuación general de la que se derivan los modelos contiene un término de varianza y un término que considera series de tiempo, las series de tiempo son representadas por funciones matemáticas comunes basadas en el juicio empírico y la experiencia.

Herrera *et al.* (2004) y Júnez (2005) modificaron el método de Herrera (1998), para su utilización en el diseño de una red de monitoreo óptima para varios parámetros de calidad del agua. Sustituyeron la matriz de covarianza obtenida por medio del modelo numérico de flujo con una matriz de covarianza calculada a través de un análisis geoestadístico de datos de calidad del agua y por otro lado utilizaron un método de optimización de inclusiones sucesivas para minimizar la varianza normalizada de todos los parámetros de forma conjunta. Se incluyeron pozos adicionales en algunas zonas consideradas como prioritarias. En estos trabajos se propuso un muestreo semestral para los parámetros identificados como de alta toxicidad y los de origen antropogénico, mientras que a los parámetros que no representan problemas para la salud se les asignó una frecuencia anual de muestreo.

Herrera y Pinder (2005) propusieron y evaluaron un método para el diseño óptimo espacio-temporal de redes de monitoreo de la calidad del agua. Los ejemplos presentados muestran que para una pluma contaminante en movimiento el método permite obtener redes de monitoreo que reducen considerablemente costos. El método combina un filtro de Kalman estático con un modelo de transporte estocástico para predecir la varianza del error de las concentraciones estimadas. Los elementos de esta matriz de covarianza se derivan de la ecuación de un modelo numérico de transporte de masa con advección y dispersión que gobierna el transporte de masa numéricamente discretizada. Para optimizar la red, buscan determinar dónde y cuándo muestrear para minimizar una función de la varianza del error de la concentración estimada (obtenida del filtro de Kalman) en las posiciones y tiempos de interés. La determinación de un plan de muestreo óptimo se obtiene a través de un procedimiento secuencial que selecciona los puntos de muestreo en espacio-tiempo que minimizan la función en cada paso, y se detiene cuando la función alcanza un valor predeterminado. Las estimaciones se realizan utilizando un filtro de Kalman ensamblado espacio-temporal y el método de optimización es secuencial. La dimensión del diseño es bidimensional en el espacio e incluye al tiempo. Se obtuvieron 2 redes de monitoreo óptimas y su calendario de muestreo siguiendo la metodología propuesta. Este método es especialmente útil cuando se cuenta con un modelo numérico de transporte del acuífero estudiado.

Zhang *et al.* (2005) diseñaron una red de monitoreo óptima de la calidad del agua subterránea combinando un filtro de Kalman y un algoritmo genético para reducir al máximo el coeficiente de variación en posiciones y tiempos establecidos. La metodología propuesta para diseñar redes de monitoreo de la calidad del agua consiste en: 1. Generar realizaciones del logaritmo natural de la conductividad hidráulica ($\ln K$) usando la técnica Latyn Hypercube Sampling (LHS), suponiendo que K es una variable aleatoria, 2. Simular el flujo y transporte del agua subterránea para generar realizaciones de la concentración del contaminante, usando las realizaciones del $\ln K$ obtenidas en el paso anterior, 3. Computar una matriz de covarianza espacio-temporal de las concentraciones del contaminante, 4. Usar un algoritmo genético para seleccionar posiciones y tiempos de estimación que reduzcan al máximo el coeficiente de variación en posiciones y tiempos de muestreo. Se usa un filtro de Kalman para actualizar la matriz de covarianza espacio-temporal después de que se ha tomado una muestra, y 5. Se sigue con el proceso de selección hasta que se logra un valor objetivo del coeficiente de variación y se

alcanza un escenario óptimo. La estimación se realiza mediante filtro de Kalman y el método de optimización es algoritmo genético. La dimensión del diseño de la red es bidimensional y considera el tiempo. El método de estimación es el empleado por Herrera (1998). La selección de los pesos usados en la reducción de la varianza se realizó mediante prueba y error. Lo mismo ocurre con la elección de los parámetros para el algoritmo genético. A diferencia de Herrera y Pinder (2005) las realizaciones de la conductividad hidráulica fueron generadas usando Latyn Hypercube.

Mogheir *et al.* (2006) propusieron una metodología para evaluar la eficiencia de una red de monitoreo de la calidad del agua subterránea, y su rediseño, usando la teoría de entropía. Mediante un ajuste por mínimos cuadrados obtuvieron un modelo de transinformación que indica la cantidad de transferencia y dependencia de información entre pozos como una función de la distancia. Los parámetros, como la mínima transinformación (elemento del modelo de transinformación que depende de la probabilidad conjunta entre la información y la distancia entre los pozos) y el alcance (distancia a la cual el modelo de transinformación se aproxima a un valor mínimo constante), se emplearon para evaluar la red de monitoreo, este análisis demostró que la mayoría de las distancias entre los pozos eran menores al alcance, es decir, se tiene un alto porcentaje de información redundante en la red. La red fue reducida sobreponiendo una malla regular de celdas cuadradas (la diagonal de cada celda igual a la mitad del alcance) sobre el área monitoreada y seleccionando un pozo por celda. No obtuvieron ningún tipo de estimación, la optimización se realizó mediante inspección. El diseño de la red es bidimensional. Esta metodología puede ser aplicada al análisis de cualquier tipo de red de monitoreo. La precisión y eficiencia de la metodología debe ser comparada con los métodos geoestadísticos que aparecen en la literatura. A pesar de que la información es espacio-temporal, únicamente se usa para definir el modelo de transinformación, y el diseño es únicamente espacial.

Hergt (2009) empleó un análisis de factores junto con un análisis de clústers de información de calidad del agua para clasificar algunos pozos del acuífero de San Luis Potosí, México, de acuerdo a su afiliación a un sistema de flujo intermedio o regional. Luego realizó el diseño de una red de monitoreo para cada sistema de flujo, siguiendo el método de optimización empleado por Herrera (1998), pero tomando como matriz de covarianza a priori, la matriz de los factores de carga correspondiente (flujo intermedio o regional), obtenida del análisis de factores. Realizó también otro diseño siguiendo el método propuesto por Júnez (2005) empleando las matrices de correlación obtenidas mediante técnicas geoestadísticas de los parámetros que resultaron más significativos del análisis de factores (F, Na, Li y Temp). Ambas metodologías resultaron ofrecer prácticamente los mismos resultados.

Masoumi y Kerachian (2010) propusieron una metodología basada en la transinformación espacial y temporal de datos para el rediseño de una red de monitoreo existente de la calidad del agua. Emplearon la medición de la transinformación basada en la teoría de entropía discreta para encontrar la distancia óptima entre los pozos de una red de monitoreo existente. La incertidumbre existente en la curva transinformación–distancia (T-D) es tratada mediante la teoría de conjuntos borrosos. La curva T-D borrosa es entonces usada en un modelo de optimización basado en algoritmos genéticos híbridos (AGH). Las funciones objetivo del modelo de optimización propuesto son: 1. Maximizar el nivel de satisfacción en la funciones de afiliación borrosas consideradas, 2. Minimizar la información redundante en el sistema, y 3. Maximizar la cobertura del sistema de monitoreo. Para mejorar la precisión de las curvas T-D, el método “C-means clustering” fue empleado para agrupar el área de estudio en zonas homogéneas, y diferentes curvas T-D fueron calculadas para diferentes zonas. La frecuencia de

muestreo de las variables de calidad del agua se definió usando curvas transinformación-tiempo (T-T). Para el diseño no se emplea un método de estimación y el método de optimización es el de algoritmos genéticos híbridos. El método de entropía es básicamente un procedimiento para establecer una red que se enfoca en la variabilidad de la calidad del agua en el tiempo y espacio. Sus ventajas básicas son que (1) provee una medida cuantitativa del contenido de información de un sitio de muestreo y de una serie de tiempo observada; (2) establece, en términos cuantitativos, la transferencia de información en el espacio; y (3) puede ser usada para determinar en forma conjunta muchas características de una red. En estudios futuros, esta metodología puede extenderse de tal manera que los sitios y las frecuencias de muestreo sean seleccionados de manera conjunta usando un modelo de optimización integrado. La metodología puede ser usada también para el diseño de una red de monitoreo integral del agua subterránea que cubra los objetivos relacionados a calidad y cantidad. En este diseño, el espacio y el tiempo se analizan por separado. Es necesario afinar el criterio para la determinación de frecuencias de muestreo mediante este método, ya que considerando la transinformación, para el caso de estudio, se recomienda el muestreo de calidad del agua de manera anual a pesar de los cambios estacionales de las variables analizadas.

Nejadkoorki *et al.* (2011) propusieron una metodología que consiste en la selección de posiciones que capturan la máxima varianza de la estimación de contaminantes del aire, mientras es minimizada la autocorrelación espacio-temporal entre los sitios seleccionados. Para la aplicación del método se obtienen estimaciones en el ARCGIS (mediante kriging) de las concentraciones del contaminante para las cuatro estaciones del año, dentro del área de estudio (usando los datos medidos disponibles), al mapa obtenido se le impone una malla regular; a cada celda se le asigna la concentración estimada por kriging y el centroide de cada una de ellas será una posible posición de monitoreo. Se calcula un semivariograma para cada temporada, y otro con el promedio anual de los datos. Para la optimización se emplea una función que minimiza lo que ellos llaman correlación espacio-temporal entre los datos; se prueban todas las posiciones que estén más allá de los alcances de los semivariogramas para cada época del año, y para el anual, y se eligen aquéllas con las que se obtiene una varianza máxima y que al mismo tiempo tengan la mínima correlación espacio-temporal (la componente espacial es considerada con el análisis para cada época, y la componente temporal con el análisis anual). La estimación se realiza con kriging pero no se especifica el método de optimización, solo se menciona que la función de optimización y los datos son procesados en MATLAB. Los autores sugieren que las autocorrelaciones espaciales y temporales son tan importantes que de alguna manera se reflejan en el número y la distribución óptima de los sitios ideales de monitoreo. Indican que la metodología puede implementarse para contaminación del agua, suelo y ruido. El método de estimación no incluye al espacio y al tiempo en forma conjunta y no se calcula una frecuencia de muestreo; únicamente se indican las posiciones que deben ser monitoreadas para cada época del año, y del análisis anual se determina el mínimo de posiciones a ser monitoreadas para ese período de tiempo, así como su ubicación.