

UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

01159
UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE INGENIERÍA

DIVISIÓN DE ESTUDIOS DE POSGRADO

DISEÑO DE UNA RED DE MONITOREO PARA EVALUAR EL
COMPORTAMIENTO DE LOS NIVELES DEL AGUA EN EL
ACUÍFERO DEL VALLE DE QUERÉTARO DURANTE LA
OPERACIÓN DEL PROYECTO DE LA PRESA EXTORAZ.

ING. JOSÉ AVIDÁN BRAVO JÁCOME

T E S I S

PRESENTADA COMO REQUISITO PARA
OBTENER EL GRADO
MAESTRO EN INGENIERÍA CIVIL
(HIDRÁULICA)

DIRIGIDA POR:

Dra. Graciela Herrera Zamarrón

JIUTEPEC, MORELOS
2005

m346668



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

JURADO ASIGNADO

PRESIDENTE: DR. FRANCISCO JAVIER APARICIO MIJARES

VOCAL: DRA. GRACIELA HERRERA ZAMARRON

SECRETARIO: DR. ALDO IVAN RAMIREZ OROZCO

SUPLENTE: M. EN C. CARLOS GUTIERREZ OJEDA

SUPLENTE: MIGUEL ANGEL MEJIA GONZALEZ

LUGAR DONDE SE DESARROLLÓ LA INVESTIGACIÓN:

INSTITUTO MEXICANO DE TECNOLOGÍA DEL AGUA

DIRECTOR DE TESIS:

DRA. GRACIELA HERRERA ZAMARRON

SUSTENTANTE:

JOSÉ AVIDÁN BRAVO JÁCOME

Autorizo a la Dirección General de Bibliotecas de la UNAM a difundir en formato electrónico e impreso el contenido de mi trabajo recepcional.

NOMBRE: JOSE ANIBAL BRAVO

JACOME

FECHA: 7 AGOSTO 2005

FIRMA: _____



DEDICATORIAS:

A DIOS. Sin ti, la no existiría.

A mi PADRE. Que con su experiencia y sabiduría me ha conducido a lograr el éxito, llevando como emblema la HONESTIDAD, PROFESIONALISMO Y SINCERIDAD.

A mi MADRE. Esto es lo menos que te puedo ofrecer, sin tu amor, esfuerzo y sacrificio no hubiera logrado nada de lo que soy ahora. Algún día estaré en tu misma posición, y a través de ellos, te pagaré todo lo que me haz dado.

A mi HERMANA. Gracias por hacerme compañía desde el día en que nació, pero sobre todo, por darme el privilegio de conocer a dos hermosos seres.

A mi FAMILIA. Ustedes han sido el ejemplo que me ha hecho crecer.

A mis AMIGOS. Sin ustedes la soledad se hubiera hecho presente, gracias por su compañía comprensión y apoyo.

*A lo largo de mi vida he aprendido que el principal
motivo por el cual se creó este mundo es el amor.*

Gracias a ti, Dorys, he vuelto a creer en él.

AGRADECIMIENTOS:

A MI DIRECTORA DE TESIS. Por su paciencia, dedicación y por compartirme su conocimiento, pero principalmente por permitirme conocerla como persona.

A LA UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO. Por brindarme un espacio en la máxima casa de estudios de nuestro país y forjarme como profesionista.

AL INSTITUTO MEXICANO DE TECNOLOGÍA DEL AGUA. Por resguardarme en sus instalaciones y otorgarme los datos necesarios para llevar a cabo mi investigación.

A CONACYT. Por apoyarme con la beca de estudios y así poder lograr el objetivo deseado.

*A la vida, por enseñarme que a
pesar de las adversidades, con esfuerzo,
se puede llegar a la cima.*

INDICE

CAPÍTULO 1

1. GENERALIDADES.....	1
1.1 Antecedentes.....	2
1.2 Objetivos.....	5
1.2.1 Contribuciones esperadas.....	5
1.3 Metodología.....	6
1.4 Contenido.....	7

CAPÍTULO 2

2. REVISIÓN DEL ESTADO DEL ARTE EN EL DISEÑO DE REDES DE MONITOREO PIEZOMÉTRICAS.....	8
2.1 Métodos basados en reglas para identificar planes de monitoreo a largo plazo..	9
2.1.1 Método hidrogeológico.....	9
2.1.2 Métodos de tendencia basados en reglas.....	9
2.2 Métodos estadísticos.....	10
2.2.1 Métodos geoestadísticos.....	10
2.2.1.1 Método geoestadístico basado en la varianza.....	10
2.2.1.2 Método geoestadístico para reducir redundancia espacial y temporal...	11
2.3 Métodos probabilísticos.....	11
2.3.1 Método de filtro de Kalman.....	12
2.3.2 Simulación probabilística.....	14
2.4 Métodos de optimización matemática.....	15
2.4.1 Programación entera.....	15
2.4.2 Recocido simulado.....	16
2.4.3 Algoritmo genético.....	17
2.5 Conclusiones.....	17
2.6 Método utilizado en esta tesis.....	18

CAPÍTULO 3

3. MARCO FÍSICO REGIONAL Y MODELO NUMÉRICO DE FLUJO

PARA EL ACUÍFERO	20
3.1 Marco físico regional.....	20
3.1.1 Localización.....	20
3.1.2 Extensión.....	22
3.1.3 Vías de Comunicación.....	23
3.1.4 Población y Actividad Económica.....	23
3.1.5 Descripción de la Zona de Estudio.....	24
3.1.5.1 Provincia Fisiográfica.....	24
3.1.5.2 Hidrografía.....	25
3.1.5.3 Clima.....	26
3.1.5.4 Suelo y Vegetación.....	28
3.1.6 Características del Acuífero.....	29
3.1.6.1 Geología.....	29
3.1.6.2 Modelo conceptual de la dinámica del flujo del acuífero.....	36
3.2 Modelo dinámico de flujo.....	41
3.2.1 Princeton Transport Code (PTC).....	41
3.2.2 Área modelada.....	43
3.2.3 Discretización espacial y temporal.....	44
3.2.4 Condiciones iniciales y de frontera.....	46
3.2.5 Recarga vertical.....	49
3.2.6 Conductividad hidráulica.....	51
3.2.7 Coeficiente de almacenamiento.....	51
3.2.8 Calibración.....	52
3.2.9. Análisis sensitivo.....	54
3.3 Predicción.....	55
3.4 Conclusiones.....	59

CAPÍTULO 4

4. DISEÑO ÓPTIMO DE LA RED DE MONITOREO	61
4.1 Análisis preliminar.....	62
4.2 Análisis geoestadístico.....	66
4.2.1 Análisis estructural.....	68
4.2.2 Estimación.....	72
4.2.3 Validación cruzada.....	77
4.2.4 Semivariograma promedio teórico.....	81
4.3 Matriz de covarianza a priori.....	84
4.3.1 Pozos de monitoreo base y puntos de estimación.....	84
4.3.2 Cálculo de la matriz de covarianza a priori.....	88
4.4 Propuesta de red de monitoreo.....	89
4.5 Estimación y validación cruzada.....	94
4.6 Periodicidad del monitoreo.....	113

CAPÍTULO 5

5. CONCLUSIONES	114
------------------------------	------------

BIBLIOGRAFÍA	117
---------------------------	------------

ANEXO I TEORÍA PROBABILIDAD Y ESTADÍSTICA, GEOESTADÍSTICA Y FILTRO DE KALMAN	124
---	------------

ANEXO II TABLAS	163
------------------------------	------------

ÍNDICE DE FIGURAS

Figura 3.1 Localización del área de estudio.....	21
Figura 3.2 Topografía del Valle de Querétaro.....	24
Figura 3.3 Escorrentía Superficial.....	26
Figura 3.4 Zonificación de climas.....	28
Figura 3.5 Geología del área de estudio (Simuta, 2005).....	31
Figura 3.6 Plano Hidrogeológico Querétaro.....	35
Figura 3.7 Flujo Subterráneo (Guysa, 1996).....	37
Figura 3.8 Distribución de las captaciones en el área de estudio (tomado de CEAQ 2005).....	40
Figura 3.9 Área del modelo y configuración piezométrica promedio para los años 1981 a 2003.....	43
Figura 3.10 Malla triangular del modelo.....	44
Figura 3.11a Topografía.....	45
Figura 3.11b Basamento.....	45
Figura 3.12 Condiciones de frontera (tomado de CEAQ, 2005). En color azul se muestran las fronteras de carga asignada con su respectiva numeración de celda y en color rojo las posiciones de los pozos.....	47
Figura 3.13 Áreas y porcentajes de infiltración de lámina precipitable por el método de los cloruros (CEAQ, 2005).....	49
Figura 3.14 Zonas de retorno por riego y filtración de agua potable.....	50
Figura. 3.15 Conductividad hidráulica de entrada al modelo (m/d).....	51
Figura 3.16 Coeficiente de Almacenamiento.....	52
Figura 3.17. Valores de conductividad hidráulica calibrados en m/día.....	53
Figura 3.18. Valores de coeficiente de almacenamiento calibrados en m/día.....	53
Figura 3.19. Análisis de sensibilidad para el año 1999.....	54
Figura 3.20. Análisis de sensibilidad para el año 2030.....	55
Figura 3.21. Piezometría simulada para el año 2010, en msnm (tomado de CEAQ, 2005).....	57
Figura 3.22. Piezometría simulada para el año 2020, en msnm (tomado de CEAQ, 2005).....	58

Figuras 3.23. Piezometría simulada para el año 2030, en msnm (tomado de CEAQ, 2005).....	59
Figura 4.1 Localización de pozos en el periodo 1993 a 1995.....	63
Figura 4.2 Localización de pozos para el periodo 1995 a 1999.....	64
Figura 4.3 Estadígrafos de abatimientos anuales (1993-1995).....	65
Figura 4.4 Estadígrafos de abatimientos anuales (1995-1999).....	65
Figura 4.5 Ajuste de modelo esférico a semivariograma muestral (1993-1995).....	71
Figura 4.6 Ajuste de modelo esférico a semivariograma muestral (1995-1999).....	71
Figura 4.7 Estimación con kriging ordinario de abatimientos anuales (1993-1995).....	73
Figura 4.8 Estimación con kriging ordinario de abatimientos anuales (1995-1999).....	74
Figura 4.9 Varianza de estimación de abatimientos anuales (1993-1995).....	75
Figura 4.10 Varianza de estimación de abatimientos anuales (1995-1999).....	76
Figura 4.11 Histograma de frecuencias y errores de estimación de abatimientos anuales (1993-1995).....	78
Figura 4.12 Histograma de frecuencias y errores de estimación de abatimientos anuales (1995-1999).....	78
Figura 4.13 Distribución espacial de errores de estimación de abatimientos anuales (1993-1995).....	80
Figura 4.14 Distribución espacial de errores de estimación de abatimientos anuales (1995-1999).....	81
Figura 4.15 Semivariograma promedio teórico de abatimientos anuales (1993-1999)....	83
Figura 4.16 Comparación de semivariogramas muestrales y promedio teórico.....	83
Figura 4.17 Comparación de semivariogramas muestrales y promedio teórico adecuado.....	84
Figura 4.18 Red de monitoreo base.....	87
Figura 4.19 Puntos de estimación (los puntos rojos son los adicionales).....	88
Figura 4.20 Reducción de la varianza total vs. número de pozos en la red de monitoreo.	90
Figura 4.21 Varianza del error de la estimación máxima.....	91
Figura 4.22 Orden de selección de los pozos de monitoreo.....	92
Figura 4.23 Errores (estimación vs simulado) 2010-2020, 29 pozos.....	95
Figura 4.24 Errores (estimación vs simulado) 2010-2020, 36 pozos.....	95

Figura 4.25 Errores (estimación vs simulado) 2020-2030, 29 pozos.....	96
Figura 4.26 Errores (estimación vs simulado) 2020-2030, 36 pozos.....	96
Figura 4.27 Distribución espacial de errores (estimación vs simulado) 2010-2020, 29 pozos.	97
Figura 4.28 Distribución espacial de errores (estimación vs simulado) 2010-2020, 36 pozos.....	98
Figura 4.29 Distribución espacial de errores (estimación vs simulado) 2020-2030, 29 pozos.....	99
Figura 4.30 Distribución espacial de errores (estimación vs simulado) 2020-2030, 36 pozos.	100
Figura 4.31 Estimación de abatimientos con 29 pozos (2010-2020).....	101
Figura 4.32 Estimación de abatimientos con 36 pozos (2010-2020).....	102
Figura 4.33 Estimación de abatimientos con 29 pozos (2020-2030).....	103
Figura 4.34 Estimación de abatimientos con 36 pozos (2020-2030).....	104
Figura 4.35 Orden de selección de pozos de red de monitoreo optima.....	107
Figura 4.36 Estadígrafos de errores con 32 pozos (2010-2020).....	108
Figura 4.37 Estadígrafos de errores con 32 pozos (2020-2030).....	108
Figura 4.38 Distribución espacial de errores (2010-2020).....	109
Figura 4.39 Distribución espacial de errores (2020-2030).....	110
Figura 4.40 Errores porcentuales (2010-2020).....	111
Figura 4.41 Errores porcentuales (2020-2030).....	112

ÍNDICE DE TABLAS

Tabla 2.1 Métodos para el diseño de redes de monitoreo.....	17
Tabla 3.1 Climas.....	27
Tabla 3.2 Suelos y Vegetación en el valle de Querétaro (INEGI, 2004).....	29
Tabla 3.3 Periodos de simulación.....	45
Tabla 3.4 Cargas asignadas a las condiciones de frontera del tipo I (tomado de CEAQ, 2005).....	48
Tabla 3.5 Porcentajes de infiltración para cada zona por cada periodo.....	50
Tabla 3.6. Errores de calibración y verificación.....	54

Tabla 4.1 Semivarianza para cada periodo y promedio.....	82
Tabla 4.2 Pozos que conforman red de monitoreo base.....	86
Tabla 4.3. Puntos adicionales.....	86
Tabla 4.4 Red de monitoreo preliminar.....	93
Tabla 4.5 Diferencias para cada periodo (valores simulados vs valores de estimación)...	105
Tabla 4.6 Lista de pozos de la red de monitoreo óptima.....	106
Tabla II.1 Abatimiento anual para los periodos 93-95 y 95-99.....	163
Tabla II.2 Valores de los pozos elegidos por filtro de Kalman extraídos del modelo de flujo.....	167

RESUMEN

El presente trabajo consistió en el diseño de una red de monitoreo para medir los niveles del agua del acuífero del Valle de Querétaro, para esto se utilizó una combinación de un método geoestadístico aplicado a datos históricos y “datos” obtenidos de un modelo de flujo del agua subterránea, un filtro de Kalman y un método de optimización heurístico.

Para hacer el diseño de la red de monitoreo se conjuntaron una serie de pozos a los cuales se les llamó red de monitoreo base. De la red base se eligieron los pozos que proporcionaban información no redundante utilizando el filtro de Kalman y optimización heurística. A la red de monitoreo resultante se le llamó red preliminar.

Para introducir información al filtro de Kalman fue necesario realizar un análisis geoestadístico de datos históricos. Los datos para hacer este análisis fueron los abatimientos anuales del acuífero de los periodos 1993-1995 y 1995-1999.

De la red de monitoreo preliminar se tomó un conjunto de pozos para determinar la cantidad de información que aportaban, esto se hizo por medio de una estimación de los errores obtenidos en el periodo 2010 al 2030. Para esto se utilizó como sustituto de los datos de abatimientos los resultados de un modelo 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, y como parte final se propuso una frecuencia de monitoreo.

1. GENERALIDADES

El crecimiento demográfico en los últimos tiempos está siendo un factor importante para que existan mayores demandas de los recursos naturales y esto ha generado la necesidad de tener una mejor visión sobre el manejo y uso de éstos, tal es el caso del agua. Lo anterior se ha acentuado en la región del Altiplano, principalmente en donde se localizan las grandes concentraciones urbanas, esto es por el motivo de que los recursos superficiales se han estado agotando o su calidad ha sido afectada.

El agua subterránea se convirtió en la mejor alternativa de abasto de los últimos 30 años, pero actualmente se ha registrado una sobre explotación importante, que en algunos casos es severa.

Los efectos que provoca el sobrebombeo se manifiestan con el descenso acelerado de los niveles de agua subterránea, lo que tiene como consecuencia que los costos por bombeo se incrementen año con año; el terreno sufre asentamientos que provoca daños irreversibles ya que el subsuelo jamás recupera su estado natural y se pierde el espacio de almacenamiento, además de que provoca daños en las estructuras superficiales. Otro problema importante que se presenta es el de agrietamientos en el terreno, que permiten la entrada de contaminantes con mayor facilidad.

Es de urgencia dar solución a esta situación tan grave, una manera de hacerlo es evaluar diferentes escenarios de operación, es decir, analizar cuál es la forma más conveniente de operar un acuífero sin agotar su capacidad y sin dejar de proporcionar la cantidad suficiente del vital líquido a las poblaciones. Esto se puede hacer con ayuda de un modelo de flujo del agua subterránea que represente de manera idónea un acuífero. Pero no termina aquí, sabiendo que el agua es parte fundamental de la vida, su calidad cada vez es menor y que en la época actual ha disminuido su disponibilidad, es necesario monitorearla para conocer temporal y espacialmente cantidad y calidad.

Una red de monitoreo bien diseñada ayuda a determinar, en cualquier tiempo y espacio, el comportamiento de niveles del agua subterránea, así como las condiciones de calidad del agua. Tal conocimiento apoyaría por ejemplo, a proteger el recurso, manejar de manera más eficiente la explotación del agua subterránea, identificar oportunamente zonas que presenten mayores abatimientos y riesgos consecuentes, cantidad disponible para su uso, características hidrodinámicas del acuífero y además ayudaría a profesionales en planeación para que puedan ordenar el crecimiento urbano.

Los elementos principales para el diseño de la red de monitoreo se fundamentan en los productos del análisis de los datos de profundidad y elevación del nivel estático del acuífero del Valle de Querétaro, en donde se trata de diseñar una red de monitoreo que permita obtener información representativa del funcionamiento del mismo con el empleo de la menor cantidad de recursos posibles; es decir, que la red que se obtenga sea óptima en el sentido de mínimo costo (mínima cantidad de pozos) y máxima calidad de la información (mínima Varianza del error).

El costo de una red de monitoreo es un factor importante ya que a mayor cantidad de pozos a monitorear es mayor el costo, debido a que es necesario mayor personal para obtener mediciones, el tiempo que se lleva a cabo es mayor al cubrir una gran cantidad de pozos, en algunos casos no se tiene fácil acceso al lugar donde se localizan los pozos y también existe la posibilidad de recorrer grandes distancias para medir cada uno de ellos.

1.1. Antecedentes

En México existen zonas importantes en donde se almacenan grandes cantidades de agua en mantos subterráneos, con capacidad de proveer el vital líquido a grandes núcleos de población. Tal es el caso de la ciudad de Querétaro, cuyo principal abasto es el agua subterránea proveniente del acuífero en donde está asentada debido a que el agua superficial no es suficiente y su calidad no es la aceptable.

La ciudad de Querétaro, capital del estado de Querétaro, es la que presenta mayor crecimiento poblacional en el estado. En el periodo de 1990-2000 la ciudad creció un 40% con tasas cercanas al 4 % anual (INEGI, 2000).

Actualmente la zona metropolitana de Querétaro concentra el 52% de la población estatal y las proyecciones del crecimiento poblacional indican que para el año 2005, la región estará ocupada por un millón de habitantes y con ello el aumento en la demanda del preciado líquido; adicionalmente, los planes de crecimiento y desarrollo de las actividades industriales, comerciales y agropecuarias que están enmarcados dentro del programa de gobierno incrementa significativamente la demanda de este recurso.

Del balance hidrológico subterráneo para el valle de Querétaro, estimado por la SARH en el año de 1987, se determinó un déficit de 79 millones de m³ anuales. En el año 2002, el déficit es de 181.8 millones de m³ anuales (CEAQ, 2002) y el descenso de niveles freáticos se ha incrementado hasta 3.5 m/año en la zona de la mancha urbana (CEA-UAQ, 2002).

En el año 2002 el acuífero de Querétaro presenta un déficit de 130% mayor que el año 1987, situación que marca precisamente las condiciones cambiantes que se observan en la respuesta del subsuelo y que pueden afectar el modelo conceptual de funcionamiento geohidrológico.

Como parte de los estudios realizados para conocer la situación del acuífero y con la finalidad de utilizarlo en la planeación de la explotación del recurso subterráneo, se ha desarrollado un modelo de flujo para el acuífero de Querétaro que en diferentes años ha sido actualizado y modificado. En 1991 la Cía Geofísica de Exploraciones Guysa S. A. de C.V., realizó el estudio geohidrológico integral del Valle de Querétaro y sus alrededores para el manejo automatizado de los recursos hidráulicos subterráneos, cuyos objetivos fueron conocer la situación hidrodinámica que guardaba el Valle de Querétaro en cuanto a sus relaciones de descarga y recarga de agua subterránea; estimar los volúmenes de reserva susceptibles de ser explotados sin incrementar drásticamente el deterioro del balance volumétrico; y mediante el modelo matemático calcular la posición de los niveles de agua subterránea para diferentes tiempos con distintas políticas de extracción. De 1992 a 1996 la Comisión Estatal de Aguas de Querétaro realizó una actualización del modelo del acuífero del Valle de Querétaro en donde el mismo fue calibrado para cada año. En 1994 el modelo fue utilizado para desarrollar un modelo de optimización, elaborado por la Comisión Estatal de Aguas de Querétaro, que como objeto tuvo la optimización del uso del agua bajo diferentes políticas de manejo. En el año 2005 el Ing. Roel Simuta Champo de la Universidad Nacional Autónoma de México (UNAM)

campus Morelos. elaboró un modelo de flujo del acuífero de Querétaro para analizar las condiciones actuales que presenta y predecir su comportamiento en el futuro, esto con fines de obtener el grado de Maestría en Ingeniería Hidráulica.

Estudios anteriores reflejan claramente los problemas que se presentaban en la ciudad de Querétaro y casi una década después los procesos continúan. Como se comentó anteriormente, la situación actual del acuífero del Valle de Querétaro presenta un déficit anual semejante al del 2002 de 181.8 Mm³ (CEAQ, 2002).

El abastecimiento de agua potable se hace en base al agua subterránea del acuífero del Valle de Querétaro, la extracción es por medio de 63 pozos que suministran 2,279.5 l. p. s. Para el año 2025 se espera una demanda de 4,276 litros por segundo (CEAQ, 2005). Obviamente, con la oferta actual de agua en todas las zonas existirá deficiencia. Debido a ello, la Comisión Estatal de Aguas analizó algunas fuentes de abastecimiento potenciales para la Ciudad de Querétaro y su zona Conurbada. Una de estas es la explotación de las aguas superficiales del Río Extóraz mediante la construcción de una presa.

Por otra parte, este valle cuenta con un historial de niveles de agua subterránea que inicia desde el año de 1970 con 6 pozos piloto monitoreados, en 1980 sólo se tiene registro de 7 pozos. Para 1990 se cuenta con 20 registros de piezometría, en el año 2000 aumentó la cantidad a 48 y para el año 2003 se tienen 112 pozos. Cabe aclarar que no existe una frecuencia de monitoreo y que las mediciones no se han efectuado sistemáticamente.

Esta red de monitoreo se ha venido adecuando año con año, a veces porque ha resultado casi imposible medir en los pozos de algunas localidades; en ocasiones, por el cambio de equipamiento que llega a obturar el orificio usado para colocar la sonda de medición. Otras irregularidades se deben a la operación continua de los equipos de bombeo (específicamente los destinados al abasto de agua potable) o a la continua profundización de los pozos.

Por desgracia, todo ello ocasiona que el historial piezométrico de cierto lugar tenga que ser interrumpido lo cual, de hecho, ha ocurrido en la mayoría de los acuíferos estatales.

Por todo lo anterior, el uso del modelo de flujo del acuífero de Valle de Querétaro, realizado por el Instituto Mexicano de Tecnología del Agua contratado por la Comisión Estatal de Agua de Querétaro (CEAQ), en el diseño de una red de monitoreo puede ser de gran utilidad para evaluar las medidas que se están contemplando para la protección del acuífero.

Bajo el contexto anterior se aplicará un análisis estadístico y geoestadístico de los datos de piezometría para cada periodo, se utilizarán los resultados de un modelo numérico de flujo para sustituir datos del periodo 2003-2030, contemplado en la planeación, y posteriormente el uso de un filtro de Kalman y un método de optimización heurística para la determinación de la red de monitoreo.

El diseño óptimo de la red de monitoreo tomará como base lo antes descrito y considerará aspectos de minimización del costo de operación, y elección de cantidad de pozos de observación y frecuencia de medición.

1.2. Objetivos

El objetivo general de esta tesis es proponer un programa de monitoreo para evaluar el abatimiento del nivel del agua en el acuífero del Valle de Querétaro durante la operación del proyecto de la presa Extóraz.

Un objetivo particular de este trabajo es probar la metodología antes descrita en la evaluación del abatimiento del nivel del agua en acuíferos.

1.2.1. Contribuciones esperadas

Una red de monitoreo que contribuya a evaluar el comportamiento del acuífero del Valle de Querétaro durante la operación del proyecto de la presa Extóraz.

Una metodología para evaluar el abatimiento del nivel del agua en acuíferos.

1.3. Metodología

El método utilizado para cumplir con el objetivo de este trabajo es una combinación de un método geoestadístico aplicado a datos históricos y “datos” obtenidos de un modelo de flujo del agua subterránea, un filtro de Kalman y un método de optimización heurístico. El método es una modificación al método propuesto por Herrera (1998), explicado en la sección 2.4.1.

Para hacer el diseño de la red de monitoreo, la CEAQ propuso un grupo de 33 pozos como posibles pozos para ser incluidos en la red de monitoreo. A éstos se añadieron 7 posiciones adicionales en zonas en las que se detectó, mediante inspección visual, que podía hacer falta información. Se le llamará red de monitoreo base al conjunto de pozos propuestos por la CEAQ y los siete puntos adicionales. De entre las posiciones incluidas en la red base se eligieron las que proporcionaban información no redundante utilizando el filtro de Kalman y optimización heurística. A la red de monitoreo resultante le llamamos red preliminar.

En este caso, la matriz de covarianza a priori utilizada por el filtro de Kalman como dato de entrada se calculó mediante un análisis geoestadístico de datos históricos. Como se requería que la matriz de covarianza fuera representativa de las características de los abatimientos del agua en el acuífero en un periodo de tiempo, se calculó esta matriz 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 una estimación de los errores obtenidos en el periodo 2010 al 2030 si éstos no eran incluidos en la red de monitoreo. Para esto se utilizó como sustituto de los datos de abatimientos los resultados del modelo 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.

Por último se hizo una propuesta de frecuencia de monitoreo basada en criterios geohidrológicos.

1.4. Contenido

Este trabajo cuenta con cinco capítulos y dos anexos, los cuales se explican en breve a continuación.

Como ya se vio, en este capítulo se describieron los antecedentes, se identificó el problema que se vive actualmente en la zona de estudio, se enunciaron los objetivos de este trabajo y la metodología empleada en el mismo.

En el capítulo 2, se hace una descripción de los diferentes métodos que han sido utilizados en el diseño de redes de monitoreo para acuíferos, esto para evaluar las ventajas y desventajas de cada uno de ellos, y ubicar el método propuesto en este trabajo en el contexto de los métodos ya existentes.

En el capítulo 3 se describen las características del acuífero con base en información proporcionada en estudios previos y también se describen las características del modelo de flujo utilizado para el diseño de la red de monitoreo, el cual fue elaborado por el Instituto Mexicano de Tecnología del Agua (IMTA).

En el capítulo 4 se explica la forma en la que se diseñó la red de monitoreo y se presentan los resultados.

Por último se analizan los resultados y se presentan las conclusiones correspondientes en el capítulo 5.

En lo que corresponde a los anexos, en el anexo I se tiene información básica para llevar a cabo la red de monitoreo. es decir, se tiene teoría sobre estadística, probabilidad, geoestadística y filtro de Kalman. Como parte final, en el anexo II se presentan tablas de datos utilizados para realizar el diseño de la red de monitoreo.

2. REVISIÓN DEL ESTADO DEL ARTE EN EL DISEÑO DE REDES DE MONITOREO PIEZOMÉTRICAS

Monitoreo a largo plazo se define como la prueba de las condiciones de agua subterránea en un periodo de tiempo ordenado, incluye la colección de datos químicos, concentración de contaminantes, y datos físicos, tal como niveles de agua o temperatura. La longitud de operación de estos monitoreos está definida por objetivos de ejecución, si ciertas condiciones son alcanzadas, la operación del sistema puede ser terminada, (Enviromental and Water Resources Institute, 2003).

Las redes de monitoreo en aguas subterráneas son indispensables para el estudio y solución de los problemas de acuíferos. El diseño de estas redes consiste en determinar las posiciones de los pozos de monitoreo y el programa de monitoreo de cada pozo. Para obtener una adecuada red de monitoreo se deben de tomar en cuenta varios factores: (1) la escala del programa de monitoreo, (2) el objetivo del programa de monitoreo, (3) el tipo de datos, (4) si existen problemas de contaminación la naturaleza del proceso de los contaminantes, (5) la naturaleza de estado estacionario vs transitorio de las propiedades del agua subterránea, y (6) las metas cambiantes del programa de monitoreo a largo plazo. En este capítulo se mencionan algunos métodos para el diseño de redes de monitoreo y trabajos elaborados en base a ellos.

A continuación se hace una revisión de métodos para el diseño de redes de monitoreo piezométricas. Se incluye además un método para el diseño de redes de monitoreo de la calidad del agua porque en esta tesis éste se modifica para su uso en el diseño de la red de monitoreo para el acuífero de Querétaro.

2.1 Métodos basados en reglas para identificar planes de monitoreo a largo plazo

2.1.1 Método hidrogeológico

Este método se usa comúnmente para programas de monitoreo en agua subterránea. Se alimenta de información hidrogeológica cualitativa y cuantitativa, esta disciplina se utiliza paralelamente con todos los métodos que tratan con agua subterránea. El programa de monitoreo en base a esta metodología es determinado por medio de cálculos y criterios hidrogeológicos sin el recurso de técnicas estadísticas y/o probabilísticas (ASCE, 2003).

La distribución horizontal y vertical de localización de muestras está dada por la geología, forma, y los patrones de flujo regional y local.

El método hidrogeológico también puede utilizar simulaciones numéricas y analíticas del flujo del agua subterránea, para definir probables zonas de impacto y llevar a cabo pruebas a priori de la efectividad de las redes de muestreo de agua subterránea.

2.1.2 Métodos de tendencia basados en reglas

Otra clase de métodos de monitoreo se basa en la combinación del criterio profesional y la experiencia, con métodos estadísticos. Este método sigue una serie de etapas de decisiones lógicas que guían a especificar decisiones de cómo, dónde y cuándo monitorear el agua subterránea. Un árbol de decisiones es una forma conveniente de organizar un sistema secuencial de decisiones de problemas que involucran incertidumbre, un número de puntos de decisión y escenarios alternativos (ASCE, 2003).

2.2 Métodos estadísticos

2.2.1 Métodos geoestadísticos

2.2.1.1 Método geoestadístico basado en la varianza

Rohuani (1985) implementó un método de reducción de varianza que utiliza aproximaciones geoestadísticas, el cual busca el número y localización de sitios de monitoreo que minimizan la varianza del error de estimación de una muestra dada de agua subterránea en uno o más sitios. La búsqueda de una red de sitios para el monitoreo de agua subterránea inicia con un número de pozos de monitoreo existente al cual se agregan pozos, uno a la vez, de un conjunto de sitios potenciales. Cada sitio de monitoreo adicional es seleccionado en base a la mayor reducción de la varianza del error de estimación dados los pozos de monitoreo existentes. Más sitios de monitoreo son agregados hasta que no haya cambios significativos de la varianza o cuando las ganancias marginales en la exactitud estadística son excedidas por algún criterio de competencia, tal como criterios de costo en la construcción de nuevos pozos. Si originalmente no existen pozos, se puede seleccionar un sistema de localización de monitoreo basado en un método hidrogeológico o cualquier otra consideración. En este artículo el método se aplicó en el diseño de redes de monitoreo de la calidad del agua pero también se puede aplicar en el diseño de redes de monitoreo piezométricas.

Algunos investigadores se han basado en la minimización de una función global de la varianza del error de estimación, esto lo toman como base para seleccionar los sitios de monitoreo (ASCE, 1990a, b). En estos trabajos se busca el mejor patrón (cuadrado o triangular) y la mejor densidad (número de sitios de monitoreo por unidad de área del acuífero) que minimiza un criterio de representación global. Este método es el más adecuado para la selección preliminar de sitios de monitoreo.

Conwell, Silliman y Zheng (1997) diseñaron una red de piezometría para la estimación del semivariograma muestral del gradiente hidráulico. Utilizaron el proceso del diseño para ilustrar una estrategia para el diseño de una red de monitoreo la cual es apropiada para la estimación de estadísticos espaciales y eficiente en términos de instrumentos de medición.

2.2.1.2 Método geoestadístico para reducir redundancia espacial y temporal

Cameron y Hunter (2000) desarrollaron varios métodos basados en geoestadística para reducir la redundancia temporal y espacial. El algoritmo temporal consiste en construir un semivariograma compuesto temporal que combina datos de series de tiempo de muchos pozos y usa el semivariograma para identificar frecuencias de monitoreo apropiadas. Se identifica el intervalo de tiempo más pequeño en el que aproximadamente se alcanza la cima del semivariograma, el que representa el mínimo intervalo de monitoreo que provee datos que prácticamente no tienen correlación temporal. Las frecuencias de muestreo se ajustan tomando este intervalo mínimo como criterio. El algoritmo espacial usa kriging para identificar pozos redundantes que no necesitan ser muestreados. Un mapa inicial es generado usando kriging con todas las posibles localizaciones de pozos. Se asignan pesos numéricos a cada ubicación de pozo para medir su importancia relativa en el mapa inicial; los pesos se obtienen del kriging y son llamados pesos globales de kriging. Se eliminan los pozos con los menores pesos y se vuelve a regenerar el mapa. Se calcula la varianza del nuevo mapa y se compara con la varianza inicial para ver si se ha incrementado sustancialmente la incertidumbre espacial. Si no, entonces se quitan más pozos y así sucesivamente.

2.3 Métodos probabilísticos

Debido a la incertidumbre substancial con respecto a condiciones futuras, los métodos probabilísticos usan aproximaciones probabilísticas para identificar futuros diseños con la mayor confianza. Algunos de estos métodos usan modelos numéricos para predecir futuras condiciones (ASCE, 2003).

Dagan (1985) utilizó un modelo estocástico del flujo de agua subterránea para probabilidad condicional e incondicional (problema inverso). La solución al problema inverso es la función de densidad de probabilidad (fdp) normal condicional de Y (log-transmisividad), condicionado sobre medidas de H (carga hidráulica) y Y , las cuales son expresadas en términos de la fdp conjunta incondicional de Y y H . El problema es reducido a determinar la covarianza de carga-log-transmisividad incondicional y el semivariograma de las cargas para una covarianza seleccionada de Y la cual depende sobre algunos parámetros desconocidos. Esto es alcanzado resolviendo una aproximación de primer orden de la ecuación de flujo.

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 deseadas. Una acción de control obtenida en la forma de retroalimentación se muestra en este trabajo 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.

2.3.1 Método de filtro de Kalman

Van Geer, et al. (1991) utilizaron el filtro de Kalman para mejorar y cuantificar la incertidumbre de simulación numérica de agua subterránea, para posteriormente diseñar una red de monitoreo. Esto con ayuda de un modelo numérico de flujo para obtener la matriz de covarianza y posteriormente su calibración. Usaron el código de modelación MODFLOW con la combinación del filtro de Kalman. Trabajaron en el diseño de muestreo para monitorear cambios espacio-temporales de cargas de agua subterráneas, causadas 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 determinísticos) 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 estáticas se satisfacen y en el cual todas las matrices requeridas el algoritmo de filtro de Kalman se suponen fijas (no cambian con el tiempo). El método estima los parámetros secuencialmente: primero una estimación de la covarianza del error del modelo es propuesta y la primera calibración de parámetros determinísticos es realizada, entonces estos parámetros se mantienen constantes y los primeros parámetros estocásticos son obtenidos. Estas dos etapas se repiten varias veces bajo un criterio de error preestablecido para ambos sistemas de parámetros que los satisfagan.

Yangxiao et al. (1991) analizaron dos ejemplos de campo. En ambos la frecuencia de muestreo se mantuvo fija y las densidades de la red fueron minimizadas bajo la restricción de un valor

límite dado de la desviación estándar del error de estimación. Varias alternativas fueron analizadas y la mejor fue elegida por inspección. Los autores analizaron la relativa importancia de la densidad de la red espacial y frecuencias de muestreo relacionándolo 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, la optimización temporal y espacial es importante.

Herrera (1998) desarrolló un método geoestadístico que combina un modelo de flujo y transporte estocástico con un filtro de Kalman. El modelo numérico es usado para calcular concentraciones de contaminante espacio-temporales y por medio de una simulación estocástica la matriz de covarianza espacio-temporal. El filtro de Kalman es usado para predecir la incertidumbre que la estimación de concentración podría exhibir si datos de concentración de muestras tomadas de localizaciones calculadas en diferentes tiempos fueron usados para actualizar las estimaciones previas. Es similar al utilizar el método de kriging en espacio-tiempo para predecir la incertidumbre de la estimación, pero con el filtro de Kalman, en lugar de utilizar un semivariograma espacio-tiempo producto de un análisis geoestadístico, se usa la matriz de covarianza espacio-temporal calculada por medio de un modelo estocástico de flujo y transporte. Una función de la varianza de la estimación es usada como un criterio para elegir los tiempos y posiciones de muestreo de la red de monitoreo mediante un método de optimización heurístico.

Como se explica en el anexo I, el filtro de Kalman puede calcular la varianza del error de estimación con base en la posición de los sitios de muestreo, sin necesidad de conocer el dato medido en campo. De esta manera este filtro se puede utilizar para determinar, dados n sitios de muestreo, cómo afecta añadir un sitio de muestreo adicional a la varianza del error de la estimación resultante.

Para construir la red de monitoreo se examinan uno por uno todos los pozos que se pueden incluir en la red y se escoge el que reduce al máximo la suma de la varianza del error de la estimación en todos los puntos de estimación (a lo que llamamos varianza total), esto es, cada nuevo pozo que se añade a la red es aquel que resulta en una varianza total menor al aplicar el filtro de Kalman. Este procedimiento se repite nuevamente para escoger de entre los pozos que

aún no se han seleccionado, el que reduce al máximo la varianza total, y así sucesivamente. A este método es al que nos referimos aquí como un método de optimización heurístico.

Júnez (2005) modificó el método de Herrera (1998), para su utilización en el diseño de una red de monitoreo óptima para varios parámetros de calidad del agua. Substituyó la matriz de covarianza obtenida por medio del modelo con matriz de covarianza calculada a través de un análisis geoestadístico de datos de calidad del agua y por otro lado utilizó el método de optimización heurístico para minimizar la varianza normalizada de todos los parámetros de forma conjunta.

2.3.2 Simulación probabilística

Cuando un modelo es calibrado por regresión no lineal, la estadística de diagnóstico calculada y medidas de incertidumbre proporcionan bastante información sobre muchos aspectos del sistema. Ely et al. (2000) y Hill et al. (2000) presentan un método de medir la importancia de las probables localizaciones existentes de observación usando medidas de incertidumbre de la predicción. Esto sugiere que se supervise continuamente en las localizaciones más importantes y una menor supervisión en las localizaciones menos importantes. En este informe se desarrolla la metodología para demostrar lo anterior usando las localizaciones de observación de las cargas hidráulicas de un modelo de tres capas del sistema de flujo regional de Death Valley. Las predicciones de interés son el transporte subsuperficial debajo de la montaña Yucca y 14 zonas de prueba subterráneas. El componente advectivo de transporte es considerado debido a que este es el componente más afectado por el sistema dinámico representado por el modelo a escala que se usó. El problema es tratado usando las capacidades del programa de computadora de U. S Geological Survey MODFLOW-2000, con su paquete ADVective-Travel Observation (ADV), y un adicional programa de computadora desarrollado para su trabajo.

Los métodos presentados en este informe son usados de tres formas. (1) Los grados para las observaciones individuales son obtenidos manipulando las medidas de incertidumbre de la predicción, y no implica el volver a calibrar el modelo. En este análisis, localizaciones de observación son omitidas individualmente y el aumento en incertidumbre resultante en las predicciones son calculadas. La incertidumbre es medido como desviación estándar sobre el

transporte advectivo simulado. El incremento en la incertidumbre es medido como el incremento porcentual en la desviación estándar causada por omitir una de las localizaciones de observación del cálculo de la desviación estándar. En general, las localizaciones de observación asociadas con un mayor incremento son tomadas como más importantes. (2) Los grados para los grupos formados en gran parte geográficamente son obtenidos usando una extensión directa del método usado por localizaciones de observaciones individuales. Este análisis es necesario donde las observaciones se conglomeran para determinar si el área es importante para las predicciones de interés. (3) Finalmente, el método es usado para la evaluación de omitir un sistema de 100 localizaciones de observación. Estas localizaciones fueron seleccionadas porque tenían grados individuales bajos y no fueron una de las pocas localizaciones en la cuales las cargas hidráulicas fueran medidas.

2.4 Métodos de optimización matemática

Los métodos de optimización matemática son algoritmos de búsqueda numérica que pueden ser usados para encontrar automáticamente un diseño de monitoreo óptimo que satisfaga objetivos y limitaciones de monitoreo especificados por el usuario. El uso de estos métodos son más apropiados cuando el número de posibles diseños hacen incómodo la búsqueda manual, tal como: (1) numerosos pozos de monitoreo existentes o potenciales, (2) numerosos constituyentes que necesitan ser considerados en el diseño de muestreo, o (3) diferentes tipos de muestras con niveles variantes de exactitud y costo. Bajo estas condiciones, la identificación de planes de muestreo óptimo con métodos heurísticos o esquemas enumerativos pueden llegar a consumir demasiado tiempo (ASCE, 2003).

Para usar métodos de optimización, primero se identifica el objetivo de monitoreo deseado. Los objetivos son plasmados en una función objetivo que cuantifica el valor o costo de varios planes de muestreo.

2.4.1 Programación entera

La programación entera ha sido el método más extensamente usado para resolver problemas en el diseño de monitoreo con optimización matemática (Mayer y Brill, 1988; Loaiciga, 1989; Hudak y Loaiciga, 1993; Wagner, 1995; Storck et al., 1997; Mahar y Datta, 1997). Puede ser

aplicado para problemas de optimización donde las variables de decisión toman un único valor entero.

2.4.2 Recocido Simulado

Varios algoritmos de optimización matemática, heurísticos o probabilísticos pueden ser usados para buscar diseños de monitoreo óptimos, de los cuales recocido simulado y algoritmos genéticos son ejemplos típicos. Las ventajas de este algoritmo son (1) pueden ser aplicados a cualquier tipo de problema de optimización; (2) pueden fácilmente ser combinados con simulación existente o rutinas de interpolación de la pluma sin mayor recodificación; (3) son particularmente efectivos para encontrar buenos, pero no necesariamente óptimos, la solución al gran problema de la optimización, substancialmente más grande que la programación entera pueda manejar; (4) pueden ser aplicados por personas sin excepción de expertos, aunque los usuarios experimentados pueden obtener mejoras significativas en funcionamiento y resultados.

Dougherty y Marryott (1991), Marryott et al. (1993), y Rizzo y Dougherty (1996) aplicaron métodos de recocido simulado para diseños de remediación de agua subterránea, pero su método puede ser fácilmente adaptado al diseño de monitoreo. El método de recocido simulado se originó de la analogía entre la optimización de una función y el comportamiento de un sistema termodinámico.

Nunes, et al. (2004) proponen un método para la optimización de una red de monitoreo de nitrato de agua subterránea con ausencia de datos. El método fue desarrollado para reducir redes existentes donde los registros de series de tiempo son escasos. Un algoritmo de optimización recocido simulado es usado para minimizar la varianza del error de estimación obtenida por kriging en problemas de combinatoria, creado seleccionando un subconjunto óptimo de estaciones del sistema original. La optimización es realizada para varios tiempos de muestreo, obteniendo un número igual de redes de dimensiones pequeñas optimizadas; las estaciones que se repiten más frecuentemente en estas redes son seleccionadas para formar parte de la red final.

2.4.3 Algoritmo genético

Otro método de optimización que ha sido implementado para el diseño de monitoreo es el algoritmo genético (Wegner, 1995; Reed et al., 2000, 2001a, b). El algoritmo genético, el cual fue desarrollado por Holland (1975), simula el mecanismo de la selección natural en busca de una solución óptima. Sobre el uso de algoritmo genético, “series” son formadas que son representaciones digitales (binario o decimal) de decisiones diferentes. Para el diseño de monitoreo, cada serie (llamada un cromosoma) puede representar un plan de muestreo. En una representación binaria, una serie puede ser 01110, el cual puede representar un no muestreo del primer pozo, muestreo del segundo a través de cuatro pozos, y un no muestreo del quinto pozo.

El algoritmo genético comienza con una “población” inicial aleatoria de series y evoluciona usando tres básicas operaciones: (1) selección, (2) sobre mezcla (unión), y (3) mutación (Goldberg, 1989).

2.5 Conclusiones

El método que se utilice para el diseño de una red de monitoreo en particular, depende de la cantidad de datos e información, además del objetivo al que se pretende llegar. En la tabla 2.1 se tiene un resumen de los métodos vistos anteriormente y su uso con respecto a la cantidad de información y datos disponibles.

Método	Cantidad de datos e información		
	Pocos	Intermedios	Muchos
Métodos basados en reglas			
Método hidrogeológico	-		
Métodos de tendencia basados en reglas	-	-	
Métodos estadísticos			
Métodos geoestadísticos		-	-
Métodos probabilísticos			
Métodos del filtro de Kalman			-
Métodos de simulación probabilística		-	-
Métodos de optimización matemática		-	-

Tabla 2.1 Métodos para el diseño de redes de monitoreo

2.6 Método utilizado en esta tesis

Como se mencionó en el capítulo 1, el objetivo general de esta tesis es proponer un programa de monitoreo para evaluar el abatimiento del nivel del agua en el acuífero del Valle de Querétaro durante la operación del proyecto de la presa Extóraz.

El método utilizado para cumplir con el objetivo de este trabajo es una combinación de varios de los métodos incluidos en la tabla 2.1. específicamente, se utiliza una combinación de un método geoestadístico aplicado a datos históricos y “datos” obtenidos de un modelo de flujo del agua subterránea, un filtro de Kalman y un método de optimización heurístico. El método es una modificación al método propuesto por Herrera (1998), explicado en la sección 2.4.1.

Para hacer el diseño de la red de monitoreo, la CEAQ propuso un grupo de 33 pozos como posibles pozos para ser incluidos en la red de monitoreo. A éstos se añadieron 7 posiciones adicionales en zonas en las que se detectó, mediante inspección visual, que podía hacer falta información. Se le llamará red de monitoreo base al conjunto de pozos propuestos por la CEAQ y los siete puntos adicionales. De entre las posiciones incluidas en la red base se eligieron las que proporcionaban información no redundante utilizando el filtro de Kalman y optimización heurística. A la red de monitoreo resultante le llamamos red preliminar.

En este caso, la matriz de covarianza a priori utilizada por el filtro de Kalman como dato de entrada se calculó, al igual que en el trabajo de Junez (2005), mediante un análisis geoestadístico de datos históricos. Pero a diferencia de este último trabajo, en esta tesis se requería que la matriz de covarianza fuera representativa de las características de los abatimientos del agua en el acuífero en un periodo de tiempo, por lo que se calculó esta matriz 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 del modelo 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.

Por último se hace una propuesta de frecuencia de monitoreo basada en criterios geohidrológicos.

3. MARCO FÍSICO REGIONAL Y MODELO NUMÉRICO DE FLUJO PARA EL ACUÍFERO

3.1 Marco físico regional

3.1.1 Localización

La zona de estudio se ubica en la porción sur-oriental del Estado de Querétaro y el estado de Guanajuato a una distancia de 210 km de la ciudad de México; sus coordenadas geográficas van de los 20° 28' 07'' a 20° 45' 03'' de latitud Norte y de 100° 19' 14'' a 100° 34' 20'' de longitud Oeste, (CEAQ, 2005). El límite que se encuentra al norte es un parte aguas natural que divide al Valle de Querétaro con el de Buenavista, cuya línea pasa por la población Santa Rosa de Jáuregui. El lado sur se limita con una línea imaginaria que pasa por el cerro El Cimatorio y el cerro Buenavista; el lado oriental se limita con una línea imaginaria que pasa por la localidad de Villa del Marqués y el occidental con la divisoria natural de aguas superficiales que divide al valle de Los Apaseos con el valle de Obrajuelo (figura 3.1).

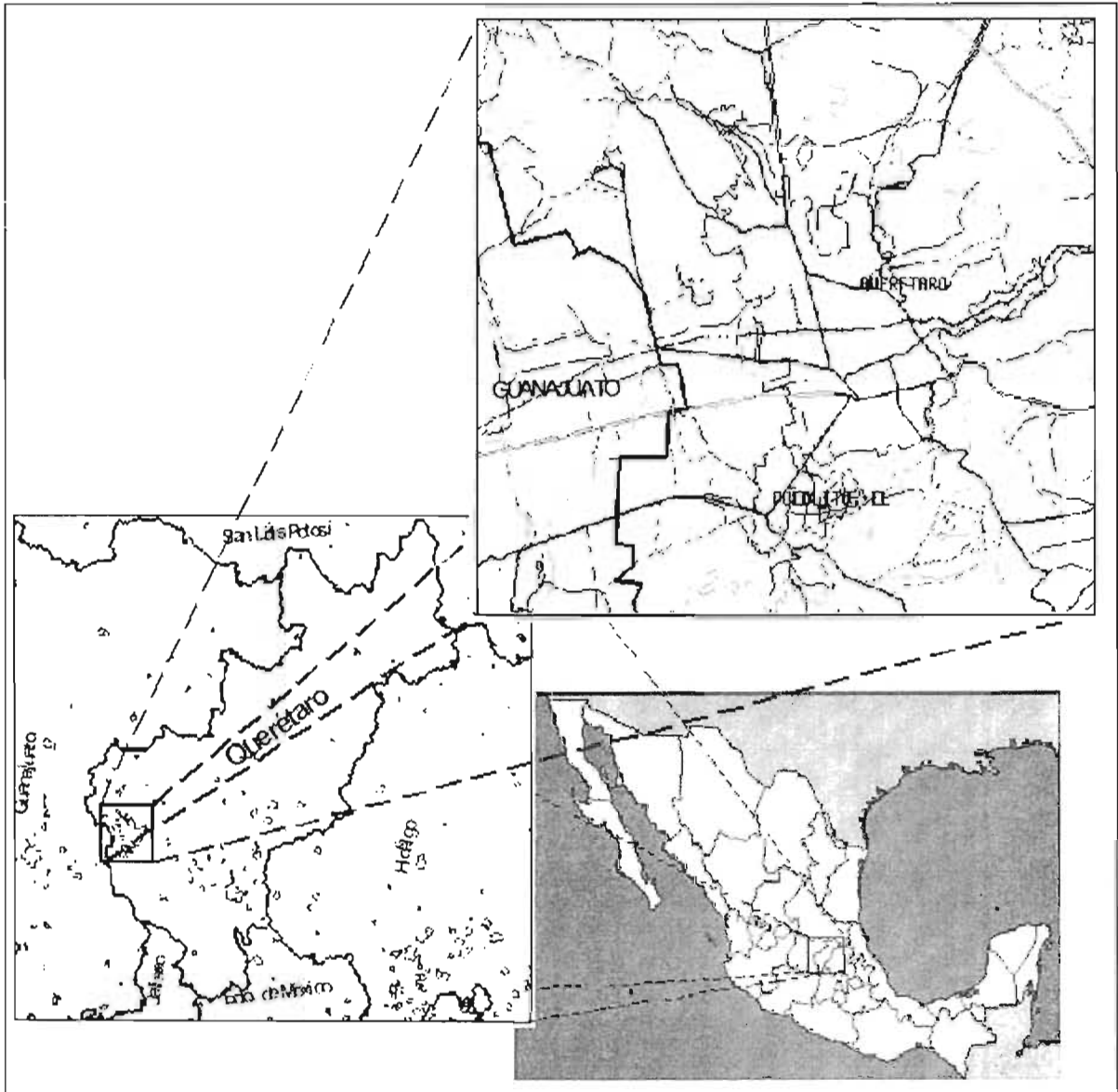


Figura 3.1 Localización del área de estudio.

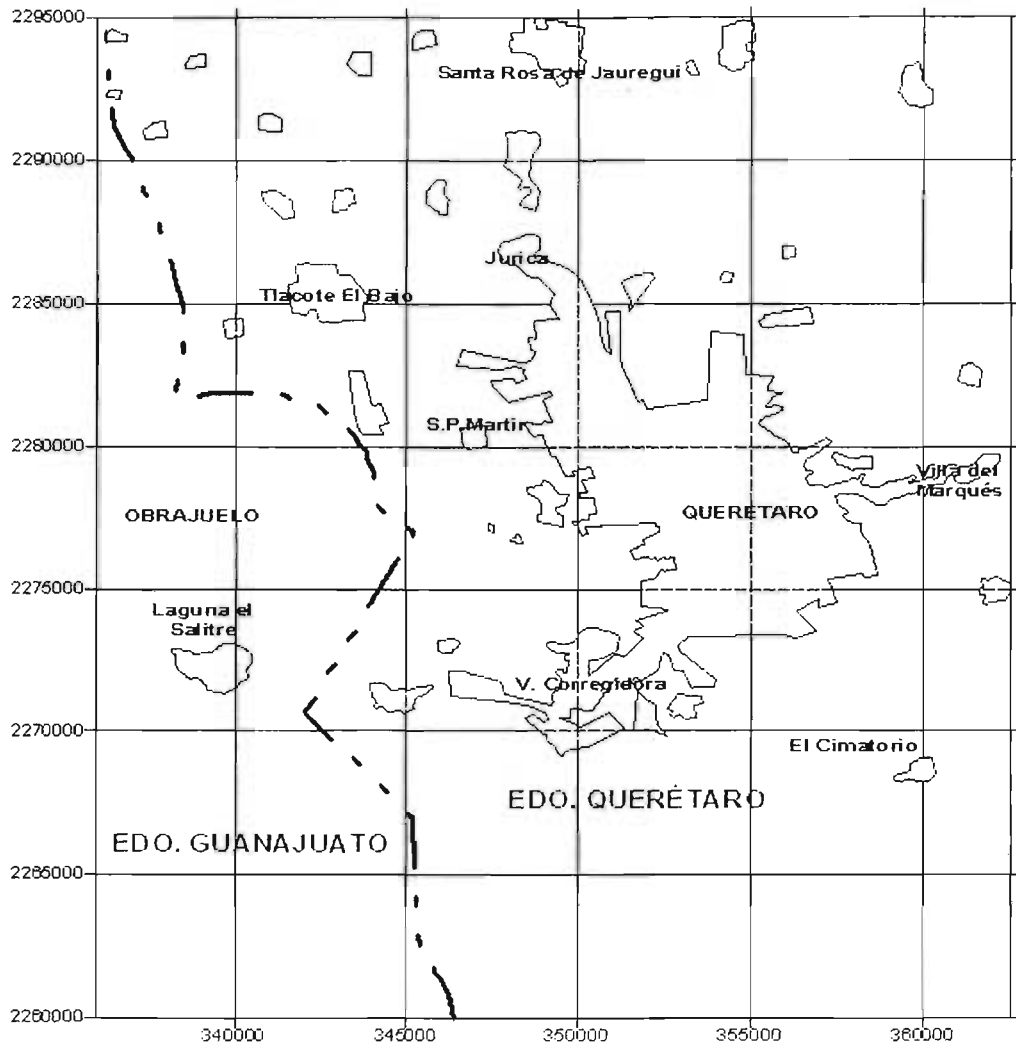


Figura 3.1 Continuación.

3.1.2 Extensión

La zona de estudio tiene una forma rectangular de 35 km de largo por 27.5 km de ancho con un área total de 962.5 km² (CEAQ, 2005). Dentro de esta superficie queda incluida la ciudad de Querétaro, Corregidora, Jurica, Tlacote El Bajo y el corredor industrial Benito Juárez. En la parte occidental la zona agrícola de los municipios de Querétaro y Obrajuelo en el estado de Guanajuato.

3.1.3 Vías de Comunicación

La zona cuenta con una extensa red de carreteras; entre ellas se encuentran las autopistas: México-Querétaro (No. 57) que cruza de sureste a oeste del área y Querétaro-San Luis Potosí (No. 57) que atraviesa de centro hacia el norte, además de la carretera federal Querétaro-Celaya (No. 45). (INEGI, 2004). Existen también las carreteras estatales pavimentadas y de terracería como los son de: Huimilpan, Tlacote el Bajo, Colón, Amascala, Villa de Marqués, La Griega, El Nabo, La Solana, Chichimequillas; Tequisquiapan y Bernal entre otros no menos importantes.

También cuenta con un sistema ferroviario con doble vía que va hacia la ciudad de México y comunica a las ciudades de Guadalajara y Nuevo Laredo.

La terminal aérea en la capital del estado ofrece un servicio nacional limitado.

3.1.4 Población y Actividad Económica

El 50% de la población a nivel estatal se concentran en los municipios de Querétaro y Corregidora con un número de 715,944 habitantes (INEGI, 2000).

La principal actividad económica es la de servicios, desarrollada por un 38.9 % del total de su población, seguida por la industria de la transformación (24.5%), el comercio con un 18.0%, la construcción (7.8 %) y actividades de gobierno (4.6 %). Por último es el sector agropecuario que ocupa sólo el 2.0% de la población asentada en el área de estudio. El 4.2 % restante lo ocupan en minería, electricidad y agua, y otras actividades no identificadas (INEGI, 2000).

En esta porción del estado se concentra parte de la industria manufacturera, productos alimenticios, textil, metal-mecánicos, eléctricos y químicos más importantes del país, así como una de las actividades más importantes como el turismo, considerándose un patrimonio nacional por su valor histórico (INEGI, 2000).

3.1.5 Descripción de la Zona de Estudio

3.1.5.1 Provincia Fisiográfica

La superficie del área presenta una topografía accidentada donde predominan procesos fluviales erosivos controlados por estructuras tectónicas (INEGI, 2004). Esta zona corresponde a la parte sur, norte y este del área, que se encuentra constituida por sierras de laderas tendidas con lomeríos de forma alargada con orientación norte-sur desde Santa Rosa de Jáuregui hasta el poblado de Huimilpan, constituida por rocas volcánicas de edad cuaternaria, se extienden desde el Cerro El Cimatario al sur hasta Obrajuelo y Mompani al norte. Las principales elevaciones dentro del área, corresponden al Cerro Cimatario (2340 msnm).

La parte oeste del área está constituida por la planicie, la cual presenta una altitud máxima de 1800 msnm y se encuentra parcialmente urbanizada. La ciudad de Querétaro se desarrolló principalmente dentro de la planicie (figura 3.2).

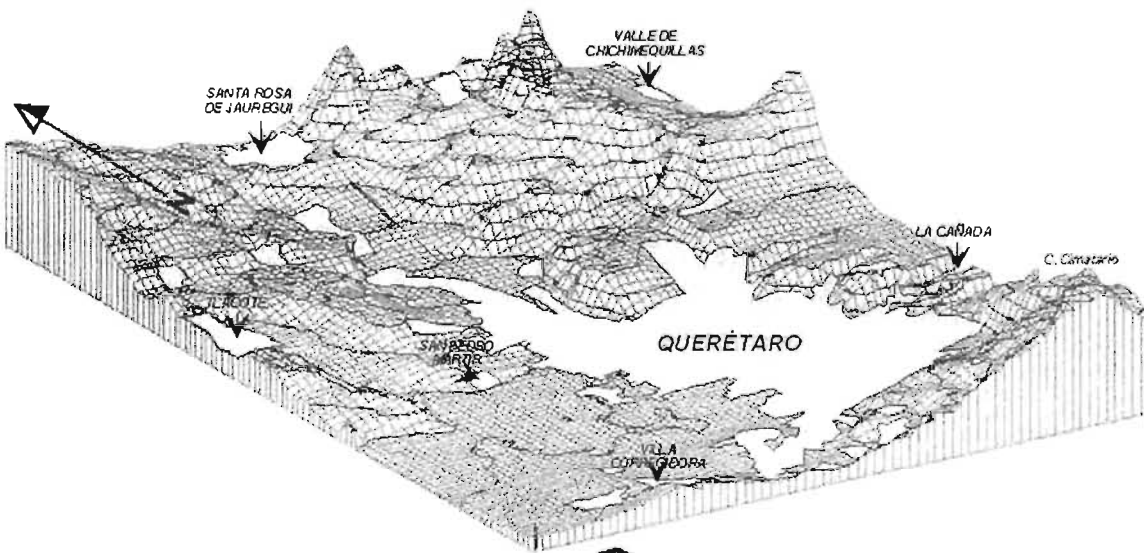


Figura 3.2 Topografía del Valle de Querétaro.

3.1.5.2 Hidrografía

En esta zona los factores climáticos son quienes controlan los recursos hidráulicos superficiales. Dentro del área de estudio se presentan las corrientes poco caudalosas de los ríos Querétaro y El Pueblito, cuya característica esencial es la de pertenecer a una zona semiárida, por tal razón en el estado de Querétaro el único recurso suficientemente aprovechable es el subterráneo (INEGI, 2004).

La corriente principal es el río Querétaro, a este cauce lo forma el tributario principal o río Chichimequillas que recoge volúmenes de la cuenca situada al noreste del Valle de Querétaro, fluye en dirección este-oeste desde La Cañada hasta Santa María Magdalena donde se intercepta con el río El Pueblito. Además existe otra escorrentía importante, esta es el dren industrial que fluye de norte a sur hasta interceptar con el río Querétaro (figura 3.3).

En las partes altas el drenaje superficial es denso con disección profunda y fuerte pendiente, en la zona de lomeríos es menos denso y poca disección en las rocas. En la zona plana del valle el drenaje es escaso con poca disección debido a la poca pendiente y reducida capacidad de infiltración, lo que genera graves problemas de infiltración en la zona urbana.

Las presas más importantes distribuidas en la periferia de las sierras son El Batán localizada al sur del valle de Querétaro (que es la de mayor capacidad), y El Cajón situada al norte de Jurica, ambos fuera del área de estudio. Los bordos más importantes son Benito Juárez, Britania, Azteca y Bolaños. Estas presas y bordos fueron construidos con la finalidad de funcionar como obras de desvío, uso agrícola y control de avenidas.

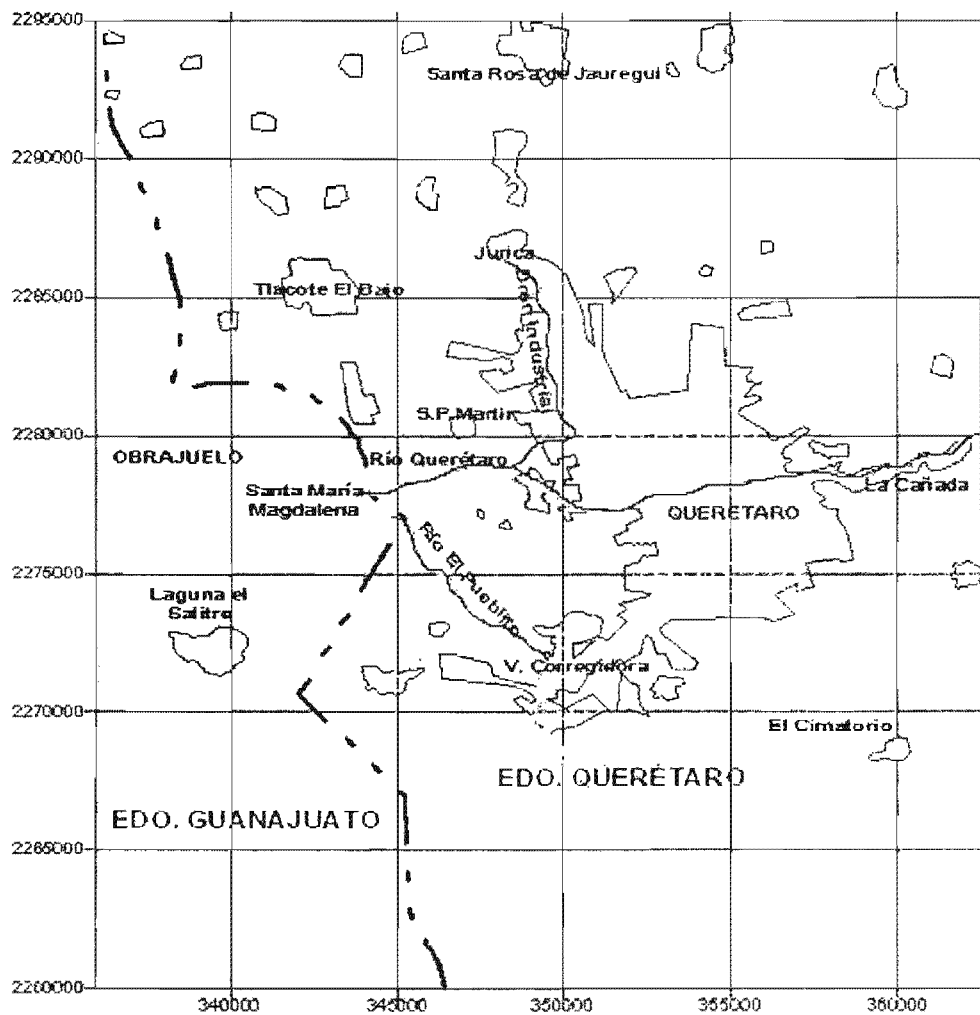


Figura 3.3 Escorrentía Superficial.

3.1.5.3 Clima

El clima del Valle de Querétaro está condicionado a la topografía y factores geográficos definidos por las diferentes altitudes de la Sierra Madre Oriental, que funciona como barrera e impide el paso de los vientos marítimos provenientes del Golfo de México (INEGI, 2004).

En la región del Valle de Querétaro se presenta un solo tipo de clima semiseco Bs_1 (tabla 3.1, figura 3.4) el cual se divide en dos subtipos:

a) $Bs_1hw(w)$ Semiseco Semicálido. La temperatura media anual es de $17.4^{\circ}C$ presentándose la máxima en mayo con $22.6^{\circ}C$ y la mínima en enero con $14.2^{\circ}C$. La precipitación de la región

es insuficiente, la media anual en la zona es de 533.87 mm, con un periodo de lluvias entre los meses de mayo a octubre y de secas de noviembre a abril. La evaporación potencial media anual es de 1689.43 mm con base en los registros de las estaciones climatológicas de Juriquilla y El Pueblito correspondientes al periodo 1979-1991 y 1960-1991 respectivamente.

b) $Bs_1kw(w)$ Semiseco templado con verano cálido. La temperatura media anual es de 17.2° C, la máxima en el mes de mayo con 20.4° C y la mínima en enero con 12.8° C. La precipitación media anual oscila entre 600 y 700 mm, presentándose el periodo de lluvias de mayo a octubre y secas de noviembre a abril con base en el registro de la estación climatológica La Palma correspondiente al periodo (1949-1993).

Estación	Altitud	Clima	Precipitación Media Anual (Mm)	Evaporación Potencial Media Anual (Mm)	Temperatura Media Anual (°C)	Región Donde Predomina
Juriquilla	1860	$Bs_1hw(W)$	508.6	1508.6	17.4	Zona central, sur y occidente del VQ
El Pueblito	1800	$Bs_1hw(W)$	566	1870.2	17.5	
La Palma	1962	$Bs_1kw(W)$	613.73	2040.98	17.02	Al este de la ciudad de Querétaro, cerro del Cimatorio y al norte de Santa Rosa de Jauregui
NOTA Clasificación tomada del Anexo Cartográfico de INEGI, 2004.						

Tabla 3.1 Climas.

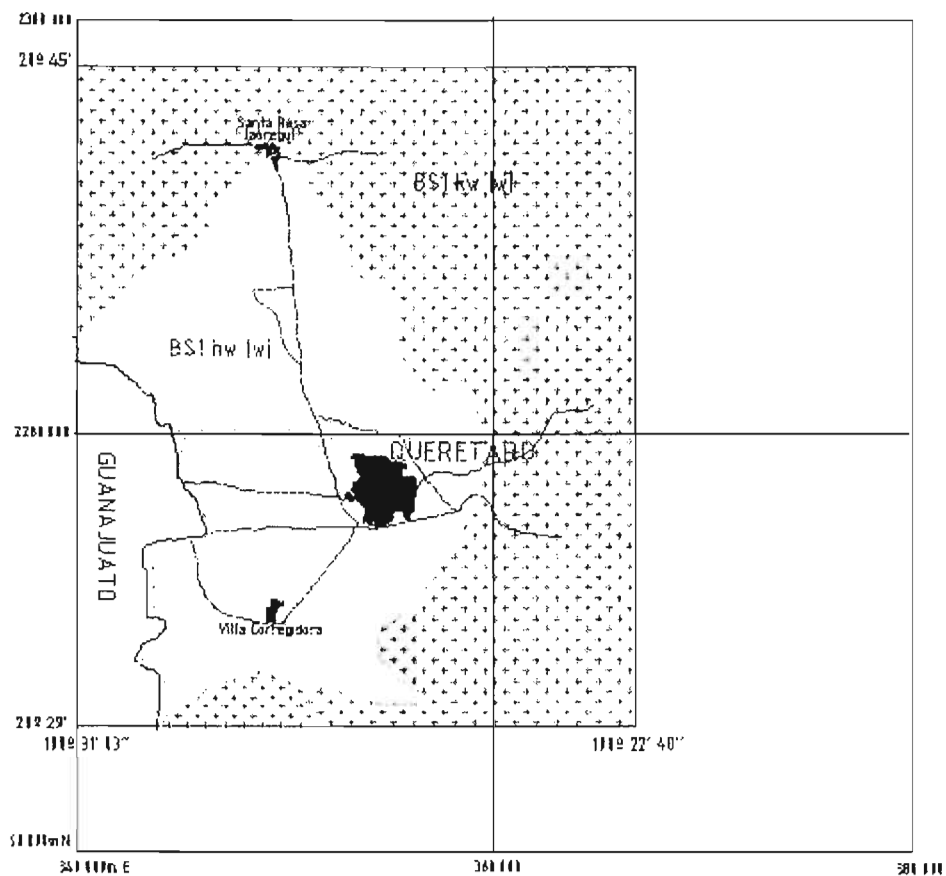


Figura 3.4 Zonificación de climas.

3.1.5.4 Suelo y Vegetación

En el área de estudio predomina la superficie urbanizada, aunque existen áreas destinadas a la agricultura de riego (con aprovechamiento del agua subterránea y residual) así como extensiones importantes de matorral y otras menores de pastizal y bosque (tabla 3.2).

Sistema de Topoformas	Suelos	Vegetación
Sierra de laderas tendidas con lomeríos.	Feozem háplicos asociados con vertisoles opelicos y litosoles. Limitados por roca y teotate.	Matorral crasicuale y subtropical, pastizal inducido y chaparral.
Pequeño llano aislado.	Dominan los suelos negros arcillosos y profundos como los vertisoles pélicos asociados con feozem háplicos y litosoles.	Matorral crasicuale (de carácter secundario con fisonomía de subinerme) y subtropical.
Lomerío con colinas redondeadas con llanuras.	Vertisoles pélicos asociados con feozem háplicos y litosoles.	Matorral crasicuale (de carácter secundario con fisonomía de subinerme, nopalera y cardonal) y subtropical pastizal inducido y chaparral.

Tabla 3.2 Suelos y Vegetación en el valle de Querétaro (INEGI, 2004).

3.1.6 Características del Acuífero

3.1.6.1 Geología

La información presentada aquí, de las características del acuífero fue tomada del estudio realizado por la empresa Guysa en el año 1996, contratada por la Comisión Estatal de Aguas de Querétaro.

El área de estudio comprende al valle de Querétaro e incluye una pequeña porción del estado de Guanajuato, conocida como el acuífero de Obrajuelo. De acuerdo con el estudio de Guysa, 1996, el valle de Querétaro se aloja en una gran fosa tectónica limitada por las fallas Querétaro y Tlacote de rumbo noroeste-sureste situadas al oriente y poniente respectivamente; a profundidad, el graben se divide en un sistema de fosas escalonadas sepultadas por una potente acumulación de sedimentos vulcanolacustres con derrames intercalados de lavas ácidas y básicas que forman un paquete acuífero de mas de 600 metros de espesor extendiéndose con características similares hasta la falla de Obrajuelo.

La parte central de la zona de estudio comprende una planicie en donde se asienta la ciudad de Querétaro con elevaciones del relieve considerables hacia la parte norte y sur del valle.

En la geología superficial se observan (figura 3.5): calizas y lutitas, ubicadas en la zona noroccidental en el poblado Juriquilla; andesitas con afloramientos en toda el área estudiada (esta unidad se considera como el basamento del acuífero en el presente estudio); ignibrita, de poca extensión en las localidades El Nabo, sobre el arroyo Mompani y en La Cañada; Basalto lajeado que aflora en diversos lugares y diversas dimensiones en toda el área de estudio; tobas líticas que afloran en mínima área en las localidades de Menchaca, en La Cañada sobre el cauce del río Querétaro, al suroeste de El Nabo sobre el arroyo Mompani; sedimentos vulcanolacustres, existen prácticamente en toda el área de estudio y rellenan el graben de Querétaro hasta formar la planicie, formando el acuífero granular del valle; basalto, esta unidad conjuntamente con los depósitos aluviales conforman la mayor parte de los afloramientos presentes en la zona de estudio, se observan en los cerros: La Cruz, Las Campanas, Gordo y El Tambor entre otros; brecha volcánica, esta unidad se observa sobre la cima del cerro Cimatorio, el cerro El Nabo, sobre las márgenes del río Querétaro, sobre el corte de la carretera al Penal, en el extremo noroccidental del área estudiada entre otros lugares; Aluvión y/o suelo residual que constituyen propiamente el valle de Querétaro formando la planicie del centro; y rocas ígneas intrusivas, localizadas en las inmediaciones del poblado Juriquilla al norte de la ciudad de Querétaro.

En el acuífero del valle de Querétaro se tienen una serie de estructuras originadas por el tectonismo ocurridos en el terciario. Dentro de estas estructuras se encuentra el graben de Querétaro, que es un extenso valle de forma poco alargada en dirección noroeste-sureste, relleno de sedimentos aluviales, flujos de lava y depósitos vulcanoclásticos lacustres potentes. Estos materiales están parcialmente saturados de agua y forman el acuífero del valle de Querétaro. De acuerdo con estudios geofísicos la parte más profunda se encuentra entre las localidades de Santa María Magdalena y Corregidora, y es de aproximadamente 600 m. Otras estructuras son una serie de fallas que se encuentran dentro del área de estudio, dentro de las fallas se pueden mencionar: falla Querétaro, al este; falla El Tlacote, al oeste, en los límites con el estado de Guanajuato; Falla Obrajuelo, en el estado de Guanajuato; falla La Cañada en la parte este y falla Menchaca en la parte noreste.

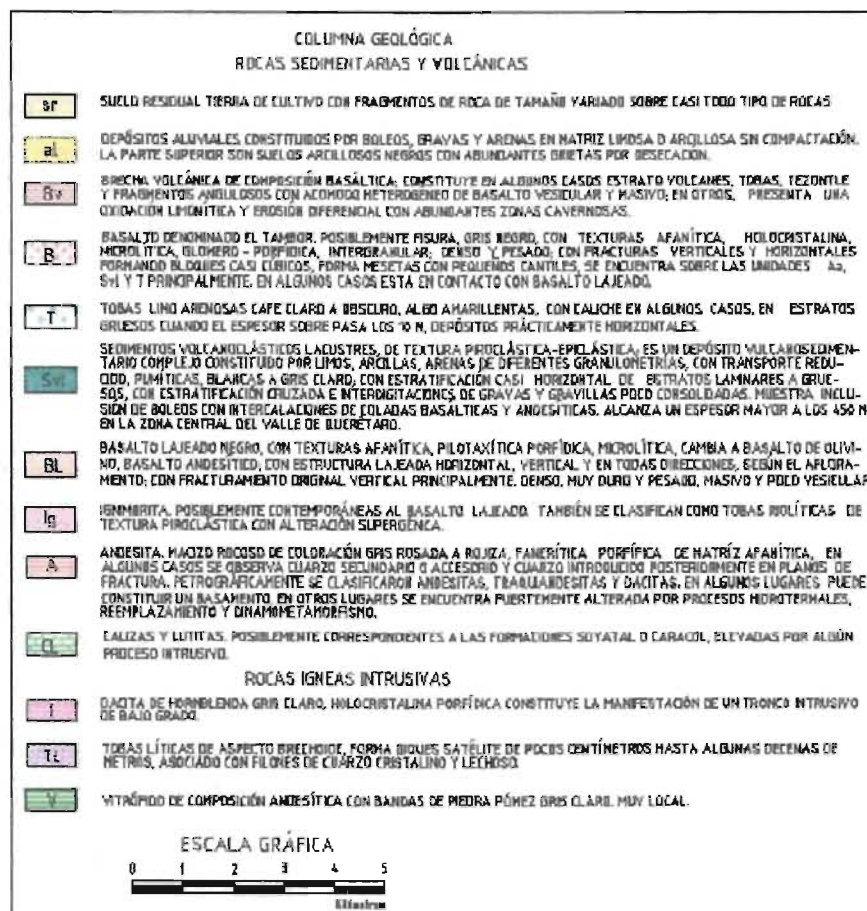
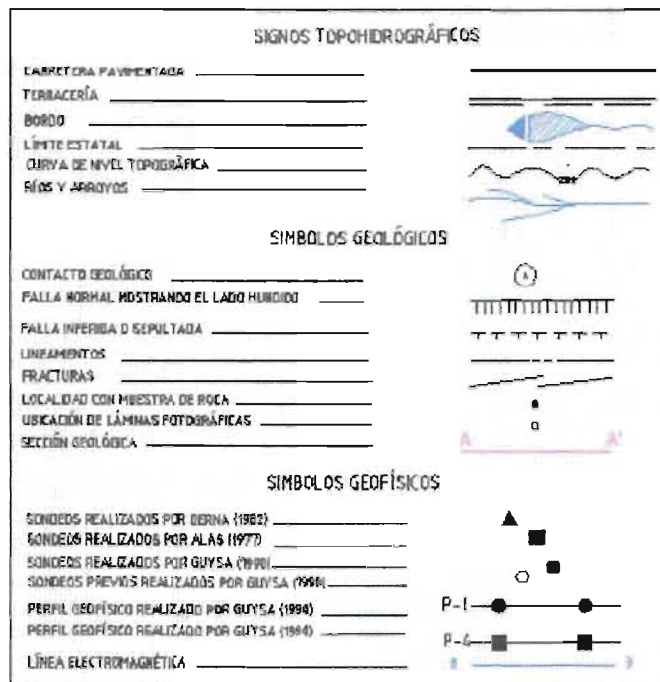


Figura 3.5 Continuación

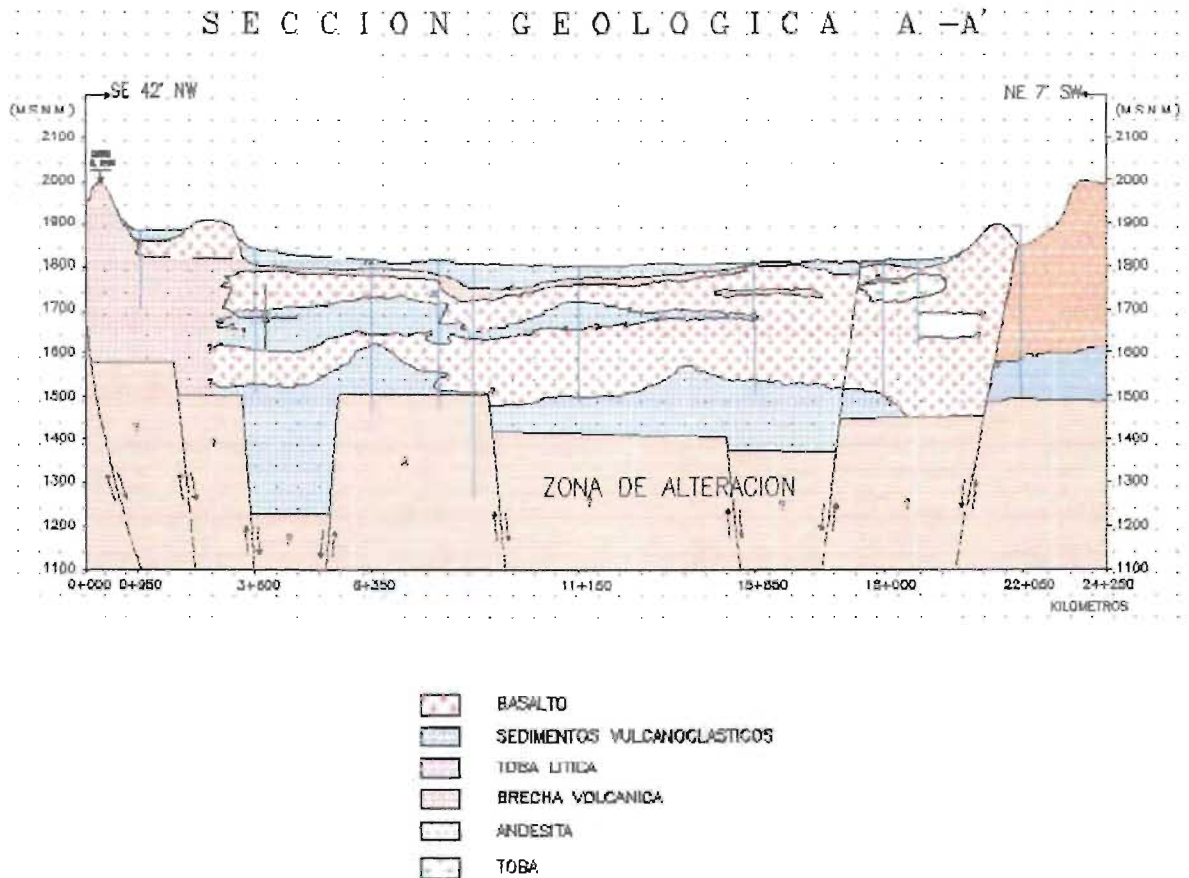


Figura 3.5 Continuación

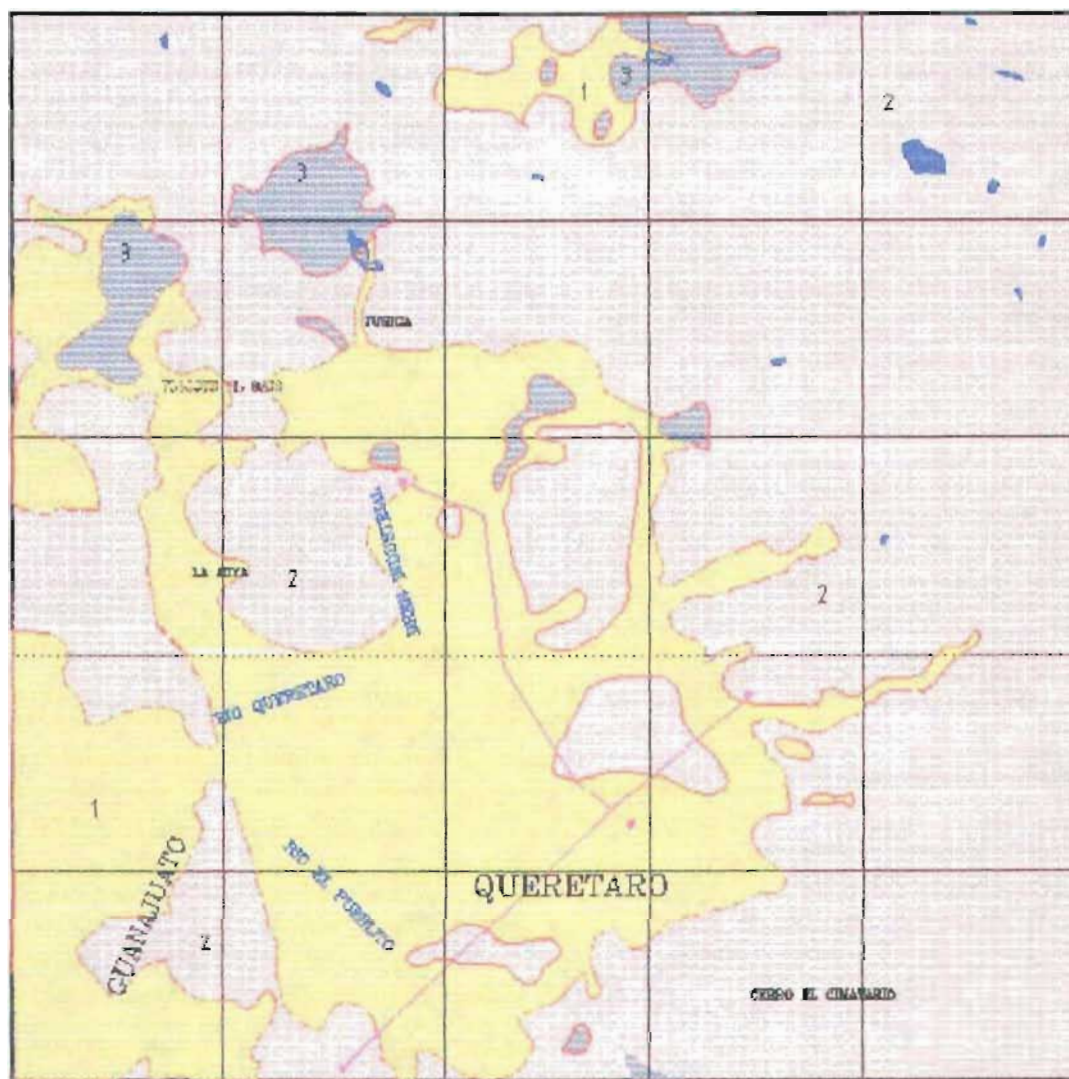
Entre los efectos del sistema tectónico de las fallas Tlacote, Obrajuelo y Querétaro, se manifiesta a profundidad y localmente, la elevación de la temperatura del agua subterránea, a 29 °C en la mayor parte del valle pero es posible encontrar algunas captaciones de hasta 43 °C y en la zona de la laguna El Salitre de 90 °C.

La geología del subsuelo de acuerdo con los cortes litológicos de pozos por los que el agua circula está constituida por los medios poroso y fracturado (figura 3.6), los materiales de estos medios se encuentran intercalados e interdigitados entre sí y son los siguientes:

- Medio poroso: Gravas, arenas, arcillas, tepetate arenoso, material aluvial de gravas y arcillas, aglomerado volcánico, toba lacustre, boleos inestables, conglomerados de clásticos medios y finos, barro, tobas arcillo arenosas, arena arcillosa, arcilla arenosa, gravas con arcilla, limo y tezontle.

- Medio fracturado: andesitas, andesitas alteradas, ignibritas, riolitas del terciario y basaltos, brechas volcánicas y tobas líticas y vítreas del cuaternario.

Cabe además mencionar que de los materiales anteriores se tiene la presencia considerable de arcillas y andesitas en toda el área de estudio. Las arcillas se encuentran principalmente en la zona de la ciudad de Querétaro y la zona industrial Benito Juárez. Los estratos de arcilla y andesita se encuentran intercalados e interdigitados con otros materiales de propiedades hidráulicas más favorables para la aportación de agua subterránea.



- 1**

MEDIO POROSO
LITOLOGÍA: DEPOSITOS ALUVIALES, TOBAS ARENOSAS Y SEDIMENTOS VULCANOLACUSTRES.
DISTRIBUCION Y ESPESOR: SE ENCUENTRA RELLENANDO LA FOSA TECTONICA DE QUERETARO, EN LA ZONA CENTRO SE HAN ENCONTRADO ESPESORES DE 350 METROS.
FUNCIONAMIENTO HIDROGEOLOGICA: ACUIFERO DE LIBRE A SEMICONFINADO SE PRESENTA COMO ZONA DE RECARGA.
- 2**

MEDIO FRACTURADO
LITOLOGÍA: BASALTO Y BRECCIAS VOLCANICAS
DISTRIBUCION Y ESPESOR: SE ENCUENTRA INTERCALADA E INTERDIGITADA CON EL MEDIO POROSO, LAS AREAS DE AFLORAMIENTO SON LAS PARTES MAS ALTAS TOPOGRAFICAMENTE SU ESPESOR PROMEDIO ESTIMADO DENTRO DE LA FOSA DE QUERETARO ES DE 300 A 350 METROS.
FUNCIONAMIENTO HIDROGEOLOGICO: ACUIFERO LIBRE, ZONA DE RECARGA DESDE LA SUPERFICIE HACIA EL ACUIFERO PRINCIPAL.
- 3**

UNIDAD IMPERMEABLE
LITOLOGÍA: ANDESITA, ANDESITA ALTERADA, IGNIEMITA, RIOLITA, CALIZA, TOBA LITICA Y TOBA VITREA.
FUNCIONAMIENTO HIDROGEOLOGICO: IMPERMEABLE TANTO EN SUPERFICIE COMO EN LOS POCOS POZOS QUE LA HAN ATRAVESADO.

Figura 3.6 Plano Hidrogeológico Querétaro.

3.1.6.2 Modelo conceptual de la dinámica del flujo del acuífero

La información de esta sección es tomada de CEAQ (2005).

La disposición natural de las unidades litológicas agrupadas en los medios poroso y fracturado, conforman el sistema que define el flujo del agua subterránea en la zona de estudio, lo que constituye un acuífero del tipo libre, comportándose como confinado en algunas localidades por la presencia de arcilla.

El acuífero es heterogéneo ya que está integrado por variaciones litológicas importantes, que se presentan en sentido horizontal y vertical.

En condiciones iniciales la fuente principal de recarga del agua al acuífero del valle de Querétaro provenía de la infiltración de la lluvia en las superficies permeables, por fracturamiento local de las sierras periféricas, en los lineamientos estructurales mayores. Otro mecanismo de recarga se realizaba por la infiltración en los cauces de las principales corrientes superficiales como son el río Querétaro, el arroyo El Pueblito y los arroyos La Gallina y Jurica.

En cuanto a las recargas por entradas subterráneas horizontales, se tienen los corredores tectonoestratigráficos de La Cañada, Jurica, Tlacote y El Pueblito, mas la proveniente de la periferia semipermeable, estas zonas de recarga estaban controladas por las variaciones estacionales y del nivel del río Querétaro que funcionaba como efluente natural, figura 3.7.

Las descargas de la cuenca tenían lugar a través de manantiales en las partes altas, y también a la evapotranspiración debido al nivel somero que tenía el acuífero.

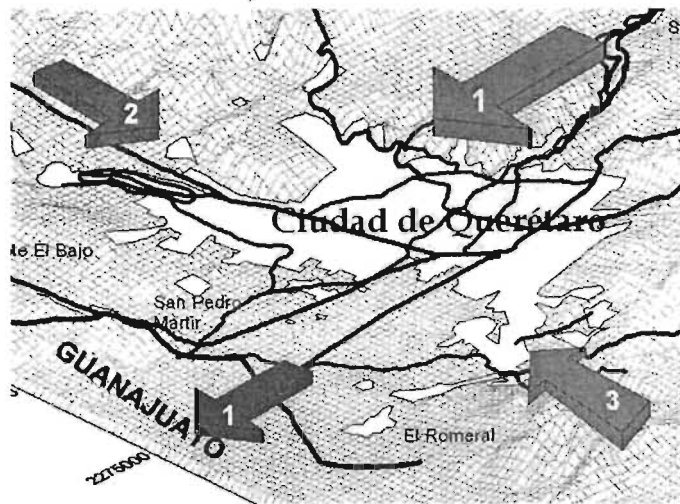


Figura 3.7 Flujo Subterráneo (Guysa, 1996).

Por otra parte se tenía un drenado subterráneo hacia el acuífero de Los Apaseos como respuesta a cargas hidráulicas definidas por un sistema de flujo intermedio en un acuífero preponderantemente libre. El sentido del flujo en el acuífero era preferencialmente de oriente a poniente.

En esas condiciones originales, la recarga de agua total era igual a la descarga del sistema, condición que empezó a modificarse con la explotación del acuífero a mediados de la década de los 70's.

El incremento de pozos y la continua explotación del acuífero hicieron que los niveles del acuífero disminuyeran de forma considerable. La continua disminución de los niveles de saturación originó en primer lugar la cancelación de los volúmenes descargados a través de manantiales y poco después el cambio de régimen de escurrimiento permanente en los cauces. Los cuerpos de agua disminuyeron su tirante, por lo que la tasa de recarga procedente del escurrimiento disminuyó. Estos fenómenos incidieron también en el fenómeno de la evapotranspiración por lo que la humedad relativa del ambiente también decreció considerablemente.

Cuando el nivel freático disminuyó de tal forma que alcanzó las capas de baja conductividad hidráulica el acuífero variaba en algunas localidades de libre a semiconfinado y se definieron dos sistemas de flujo subterráneo. Un sistema lo forma una zona con elevaciones del nivel freático que presentan un nivel colgado o local de poca profundidad en la zona central del valle, principalmente a lo largo del cauce del río Querétaro, identificándose este comportamiento en los pozos 959, 992, 1434, 595 y 743. Por otro lado se presentó un sistema de flujo intermedio cuyas líneas de flujo responden a cargas hidráulicas menores respecto del flujo local. Cabe destacar que este sistema sustenta casi el total de extracción en el valle y que esta formado también por el medio poroso y fracturado.

En el sistema local los niveles se encontraban para 1996 entre 15 y 40 m de profundidad, y se infiere que está limitado en su parte inferior por una capa semiconfinante que en algunas localidades se presenta en medios porosos y en otras fracturados. En el sistema intermedio el nivel del agua subterránea variaba de 90 a 110 metros de profundidad en la mayor parte del valle y de 170 metros en las zonas menos permeables como el subsuelo de la zona industrial Benito Juárez.

De acuerdo con la temperatura del agua subterránea se tienen registros que indican la presencia de termalismo principalmente en las localidades de San Bartolo Agua Caliente y la laguna El Salitre; ambas alineadas con la falla El Salitre.

Para el periodo de diciembre de 1999 – diciembre de 2003, se observan abatimientos de 30 m en la zona sur de la ciudad de Querétaro, un abatimiento de 25 m en el valle de Obrajuelo, otro de 20 m en la zona de Villa Corregidora y en general se observan abatimientos de 10 m, en toda el área de estudio.

Dentro del área de estudio se observan dos conos de abatimientos, el primero ubicado en la parte central de la ciudad de Querétaro con una elevación piezométrica mínima de 1630 msnm, y el segundo en la zona de San Pedro Mártir con una elevación piezométrica mínima de 1640 msnm. En general en el resto del valle se observan elevaciones piezométricas que van de los 1690 a los 1670 msnm. La dirección del flujo va de las zonas de recarga hacia los conos de abatimiento.

En la zona de descarga horizontal del acuífero hacia el acuífero de Obrajuelo los gradientes del agua subterránea se han invertido modificando el sistema de flujo, de tal forma que se ha formado un parteaguas subterráneo, producto de la explotación que se tienen tanto en el acuífero de Los Apaseos como en el de Obrajuelo, teniéndose un pequeño canal de descarga en la parte superior de esa frontera con gradientes hidráulicos muy pequeños.

La empresa Gondwana en el año 2000 actualizó el censo general de captaciones en ese mismo año, se recorrieron un total de 263 aprovechamiento, de los cuales 38 estaban inactivos y 225 activos, de estos últimos, 8 se encontraban momentáneamente inactivos. Dentro del área de estudio se cuenta también con 110 pozos que corresponden al acuífero de Obrajuelo, la información de estos pozos se obtuvo de la base de datos proporcionada por la Comisión Estatal de Agua y Saneamiento de Guanajuato (CEASG, 1998), de ellos 86 se encuentran actualmente activos y 24 se encuentran inactivos. En la figura 3.8 se presenta la distribución de las captaciones en el área de estudio.

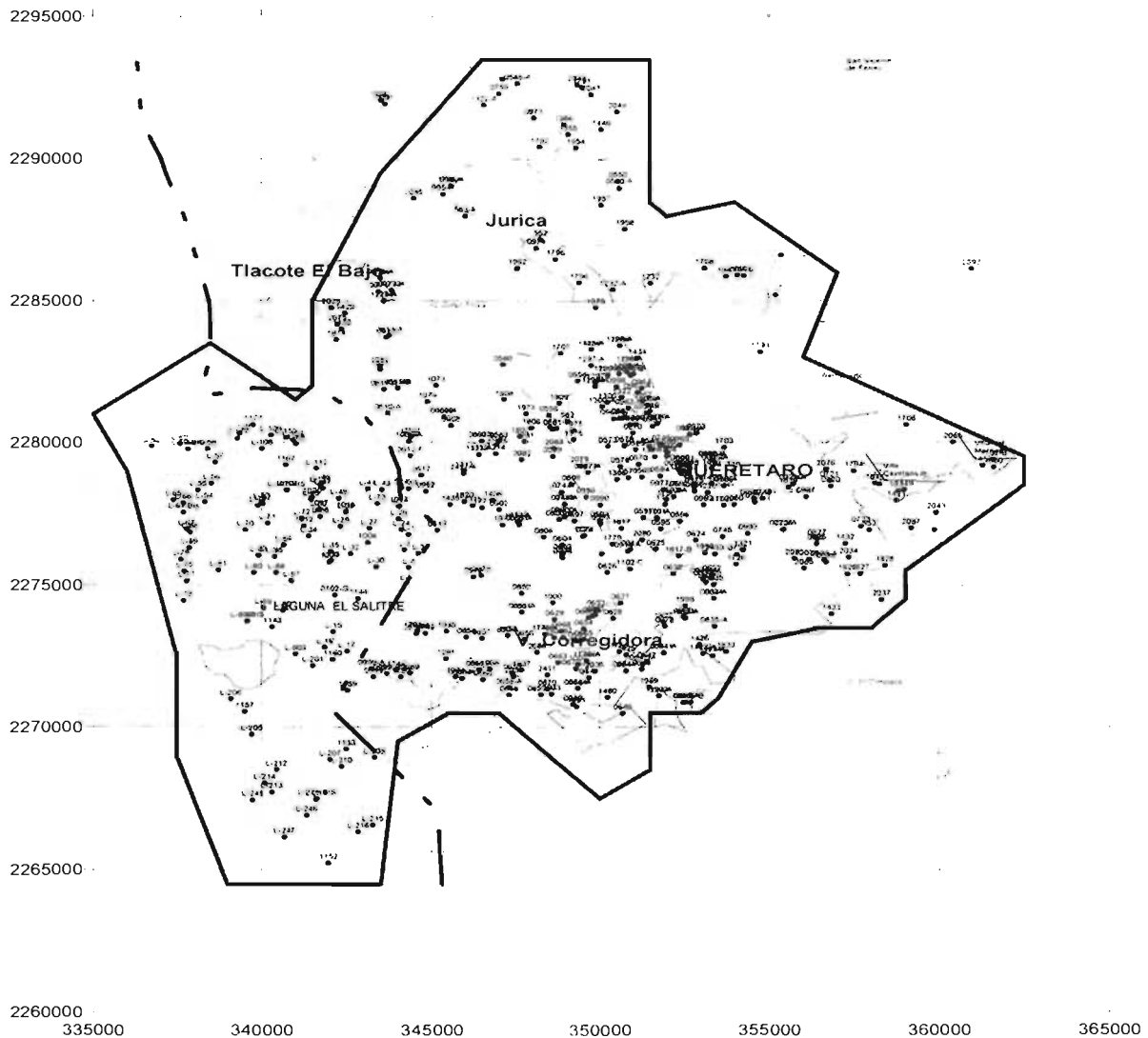


Figura 3.8 Distribución de las captaciones en el área de estudio (tomado de CEAQ 2005).

De acuerdo con las evoluciones de los niveles estáticos se tiene que para el periodo 81-99 se tiene un abatimiento promedio de 2.59 m por año, teniéndose un déficit en el acuífero de $101.36 \text{ Mm}^3/\text{año}$ (Simuta 2005).

3.2 Modelo dinámico de flujo

El objetivo de esta tesis es la de realizar el diseño de la red de monitoreo en el acuífero del Valle de Querétaro para evaluar el comportamiento de sus niveles piezométricos durante la operación del proyecto de la presa Extóraz. Tal presa se encuentra en etapa de proyecto, por consiguiente es necesario realizar una predicción del comportamiento del acuífero para los años en el cual ya se encuentre en funcionamiento y posteriores.

Una herramienta útil para evaluar el comportamiento de un acuífero es un modelo. Un modelo es un dispositivo que representa una aproximación de una situación en campo. Para el caso de la geohidrología existen dos tipos de modelos: modelos físicos y modelos matemáticos (numéricos y analíticos). Por la gran heterogeneidad que presenta un acuífero en todas sus partes es imposible caracterizarlo por medio de un modelo físico. Un modelo matemático es el más adecuado, ya que resuelve ecuaciones que representan los procesos físicos y describe los flujos que ocurren en el sistema. En un modelo matemático analítico resultaría una solución fuera de la realidad, la solución analítica está hecha en base a simplificaciones y no refleja el verdadero comportamiento del sistema. El modelo matemático numérico permite la entrada de sistemas complejos (de acuerdo a las características del acuífero a modelar), además de que se cuenta en la actualidad con herramientas computacionales que permiten la solución de estas ecuaciones.

Para poder construir la red de monitoreo en base a los niveles piezométricos que presentará el acuífero cuando se encuentre en operación la presa Extóraz, fue necesario utilizar un modelo numérico, el cual fue desarrollado por el Instituto Mexicano de Tecnología del Agua (CEAQ, 2005) y que se da a continuación una breve descripción de sus características.

3.2.1 Princeton Transport Code (PTC)

El modelo de flujo del acuífero del Valle de Querétaro utiliza un programa de cómputo o código que se emplea para resolver la ecuación gobernante de flujo en medios porosos, resultante de la ley de Darcy. Este programa es llamado simulador Princeton Transport Code (PTC) el cual usa la siguiente ecuación diferencial:

$$\frac{\partial}{\partial x} \left(K_{xx} \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_{yy} \frac{\partial h}{\partial y} \right) + \frac{\partial}{\partial z} \left(K_{zz} \frac{\partial h}{\partial z} \right) = S \frac{\partial h}{\partial t} - Q$$

donde:

h es la carga hidráulica [L]

K_{xx} es la conductividad hidráulica en la dirección horizontal x [LT^{-1}]

K_{yy} es la conductividad hidráulica en la dirección horizontal y [LT^{-1}]

K_{zz} es la conductividad hidráulica en la dirección vertical z [LT^{-1}]

S es el coeficiente de almacenamiento específico, [L^{-1}]

Q es el término fuente/sumidero [L^3T^{-1}] (e.g. bombeo; se representa con valores negativos)

PTC utiliza un método de solución llamado elemento finito, tiene la ventaja de ser capaz de representar con mayor exactitud la geometría irregular de un acuífero por medio de un acomodo triangular e incluso con mallas rectangulares deformadas. Para información más detallada sobre el código de solución Pinceton Transport Code (PTC), se puede consultar en Babu et al. (1997). Además la interfaz gráfica para PTC (Olivares, 2001) es desarrollada en el programa Argus ONE (Argus Interwere, 1997). que permite configurar gráficamente el modelo.

3.2.2 Área modelada

La determinación del área del modelo se basó en la presencia del parte aguas subterráneo entre el acuífero de Obrajuelo y el de los Apaseos, mencionado en la sección 3.1; entradas y salidas de flujo subterráneo; geología, barreras impermeables y topografía del terreno. El área del modelo cuenta con una extensión de 426 km² (figura 3.9).

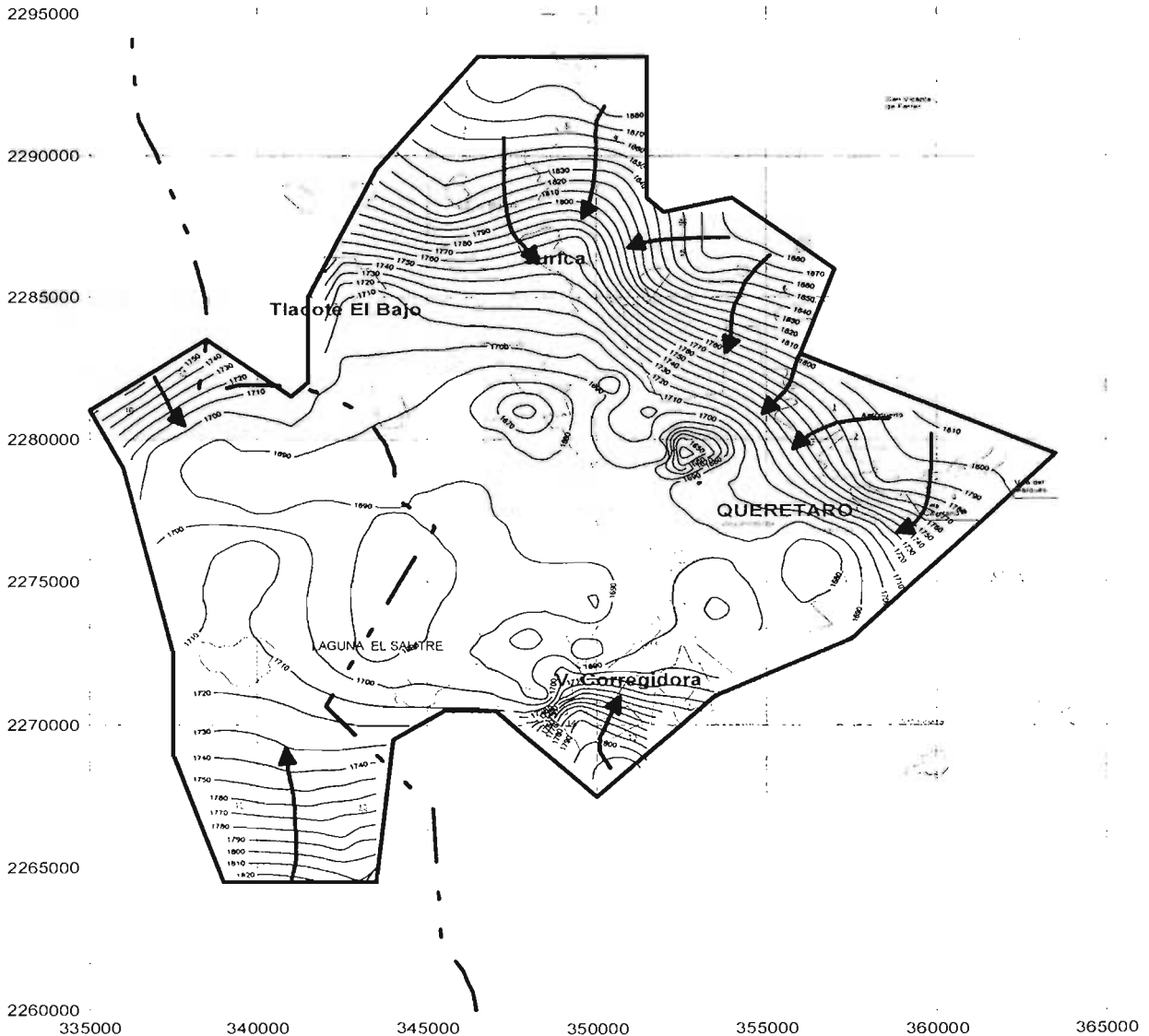


Figura 3.9 Área del modelo y configuración piezométrica promedio para los años 1981 a 2003.

La superficie del modelo considera principalmente la zona plana del Valle de Querétaro que se extiende de este a oeste, en donde se localizan la ciudad de Querétaro, la zona industrial Benito Juárez, la zona agrícola del valle y la de Obrajuelo; al sur el poblado de Villa Corregidora y una pequeña porción cerro El Cimatorio; al norte, las zonas montañosas del valle de Buenavista que se consideran como zonas de recarga, Jurica y Tlacote el Bajo (CEAQ, 2005)

3.2.3 Discretización espacial y temporal

El modelo cuenta con una malla triangular que en zonas donde se tiene mayor densidad de pozos es más refinada. Tiene 6676 nodos y 13066 elementos (figura. 3.10).

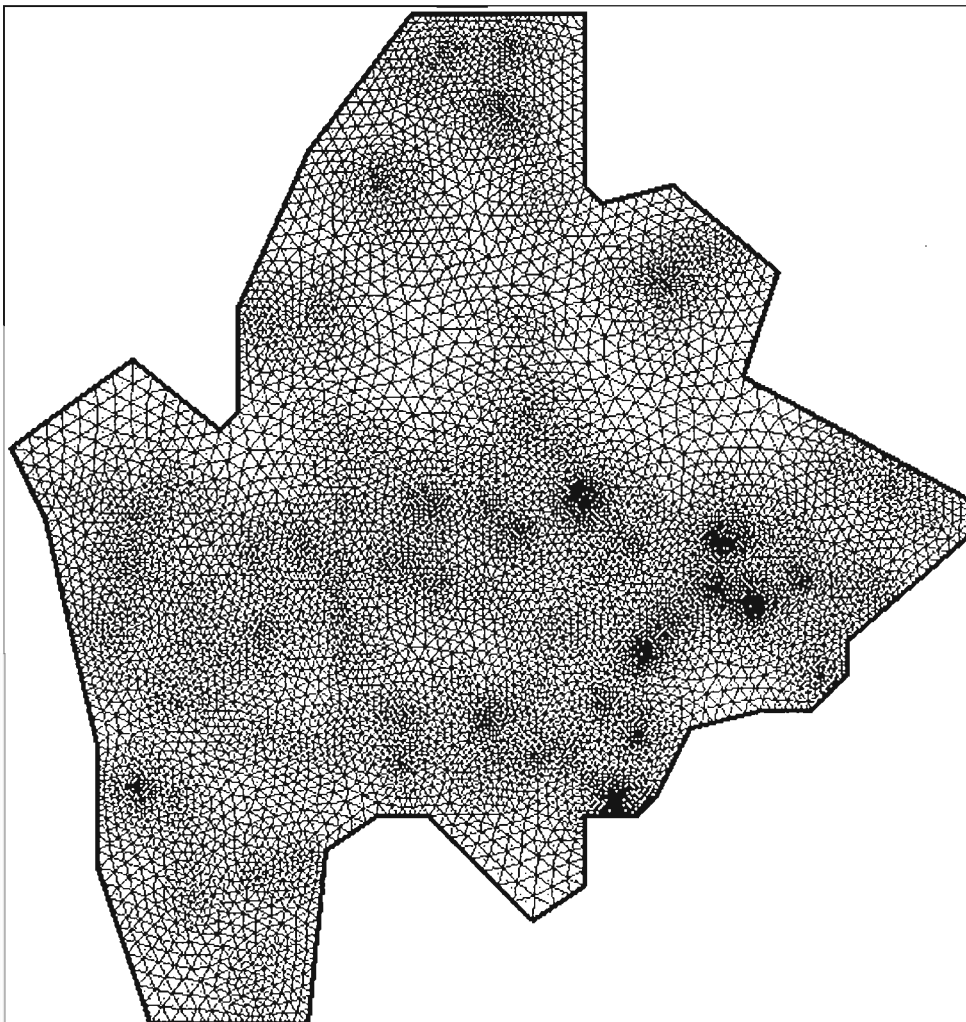


Figura 3.10 Malla triangular del modelo.

El modelo es de una sola capa. El límite inferior llamado basamento se definió por medio de las secciones geológicas proporcionadas por la empresa Guysa (1996), el límite superior es en base a la topografía del terreno proporcionada por la Comisión Estatal de Aguas del Estado de Querétaro, figuras 3.11a y 3.11b.

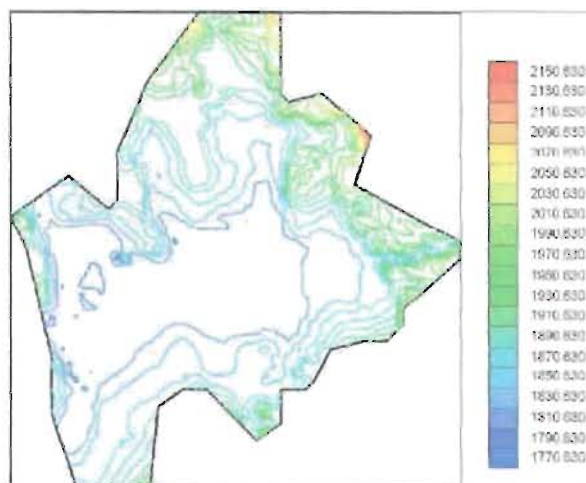


Figura 3.11a Topografía

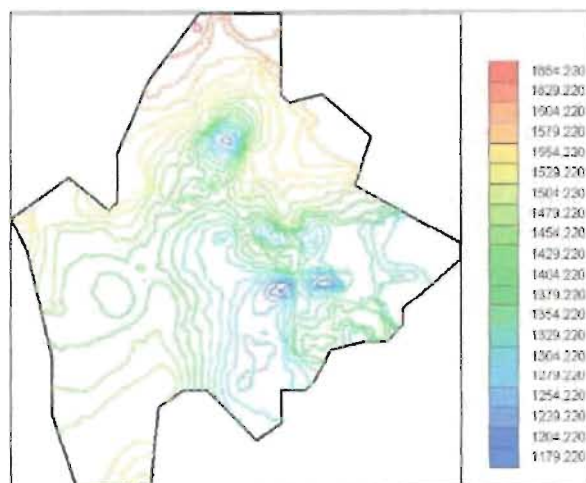


Figura 3.11b Basamento

En cuanto a la discretización temporal el modelo se simuló en estado transitorio, a partir de 1985 hasta 1999, dividido en 4 periodos, contados en días. En la tabla 3.3 se muestran los periodos de simulación.

Periodo	Inicia día	Termina día	Total de días
Ago 1985 – ago 1990	1	1825	1825
Ago 1990 – dic 1993	1826	3042	1217
Dic 1993 – dic 1995	3043	3772	730
Dic 1995 – dic 1999	3773	5232	1460

Tabla 3.3 Periodos de simulación.

3.2.4 Condiciones iniciales y de frontera

Para resolver la ecuación de flujo en medios porosos es necesario contar con condiciones iniciales y de frontera. Como condiciones iniciales se tienen los niveles estáticos para el año de 1985.

Las condiciones de frontera se definieron tomando en cuenta el modelo conceptual del funcionamiento hidrogeológico. Se utilizaron fronteras de carga asignada y sin flujo (figura 3.12). Para la etapa de calibración, en las zonas de recarga y descarga del modelo se asignaron fronteras de carga asignada variable, con los valores asignados para cada periodo que se muestran en la tabla 3.4. El valor de la carga asignada depende del nivel estático de las configuraciones piezométricas para cada periodo de calibración. En el resto de la frontera no se tiene flujo.

Los pozos de extracción son considerados como fronteras de flujo asignado. La empresa Guysa realizó una reconstrucción del bombeo en el valle de Querétaro de 1938 a 1996 (Guysa, 1996), que se actualizó con datos proporcionados por la CEAQ y por la Comisión Estatal de Agua y Saneamiento de Guanajuato (CEASG), la localización de estos pozos se muestran en la figura 3.12.

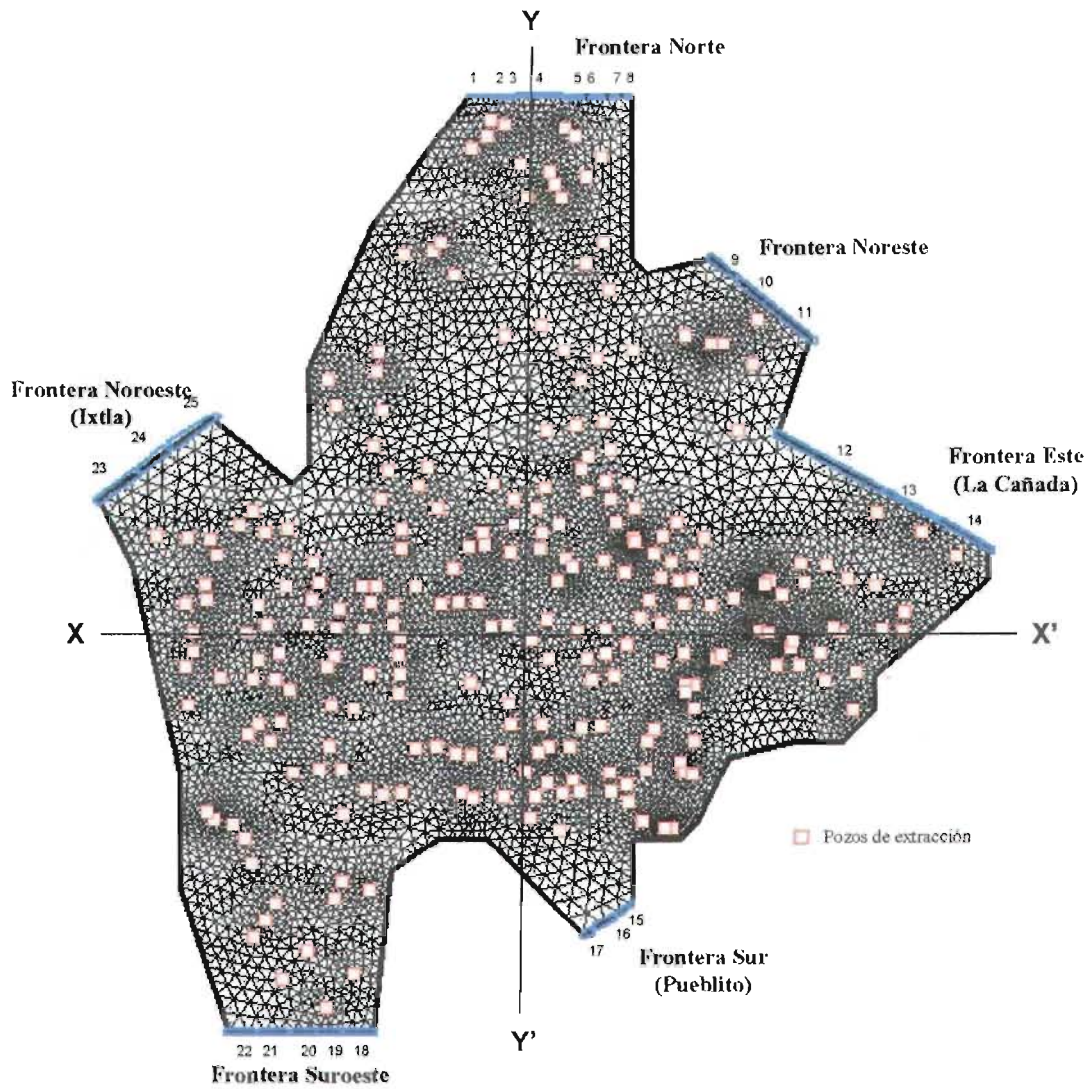


Figura 3.12 Condiciones de frontera (tomado de CEAQ, 2005). En color azul se muestran las fronteras de carga asignada con su respectiva numeración de celda y en color rojo las posiciones de los pozos.

No. (Misma numeración que se muestra en la figura 3.12)	Frontera	Carga asignada (msnm)			
		Periodo 1	Periodo 2	Periodo 3	Periodo 4
1	Norte	1902.04	1896.72	1893.91	1888.45
2		1900.62	1895.50	1892.54	1886.74
3		1901.13	1895.30	1891.61	1885.69
4		1902.39	1895.49	1890.82	1884.9
5		1903.94	1896.53	1891.39	1885.63
6		1905.17	1897.96	1893.00	1887.65
7		1905.11	1898.30	1893.62	1888.48
8		1904.81	1898.21	1893.68	1888.63
9	Noreste	1888.27	1887.90	1887.51	1886.79
10		1888.62	1888.54	1888.44	1888.19
11		1880.48	1879.87	1879.81	1879.65
12	Este (La Cañada)	1823.77	1813.05	1806.89	1806.00
13		1829.71	1818.59	1812.75	1811.37
14	Sur (Pueblito)	1829.10	1819.39	1816.13	1812.15
15		1814.86	1813.47	1812.07	1811.45
16		1818.06	1817.50	1816.77	1816.49
17	Suroeste	1816.27	1815.45	1814.26	1813.71
18		1833.39	1826.26	1821.77	1816.79
19		1837.18	1829.72	1825.11	1820.42
20		1842.50	1835.12	1829.91	1825.73
21		1843.66	1836.74	1831.59	1827.85
22		1841.30	1834.49	1829.76	1826.31
23	Noroeste (Ixtla)	1778.13	1770.64	1764.79	1759.9
24		1779.84	1772.16	1766.19	1760.7
25		1776.99	1769.05	1762.9	1756.77

Tabla 3.4 Cargas asignadas a las condiciones de frontera del tipo I (tomado de CEAQ, 2005).

3.2.5 Recarga Vertical

En cuanto a la recarga vertical esta se presenta en dos formas: recarga natural (lluvia) y recarga inducida (retorno por riego y fugas en red de agua potable). La recarga por lluvia no es constante en toda el área a modelar, la Universidad Autónoma de Querétaro en convenio con la Comisión Estatal de Aguas de Querétaro realizó un estudio en el 2002 en el cuál se obtuvieron los porcentajes y áreas de infiltración de la cantidad de lámina precipitable promedio para cada periodo, por medio del método de los cloruros Simuta (2005) utilizó esta información para determinar el coeficiente de infiltración que se presenta en la figura 3.13 y las láminas de precipitación al aplicar estos coeficientes se presentan en la tabla 3.5

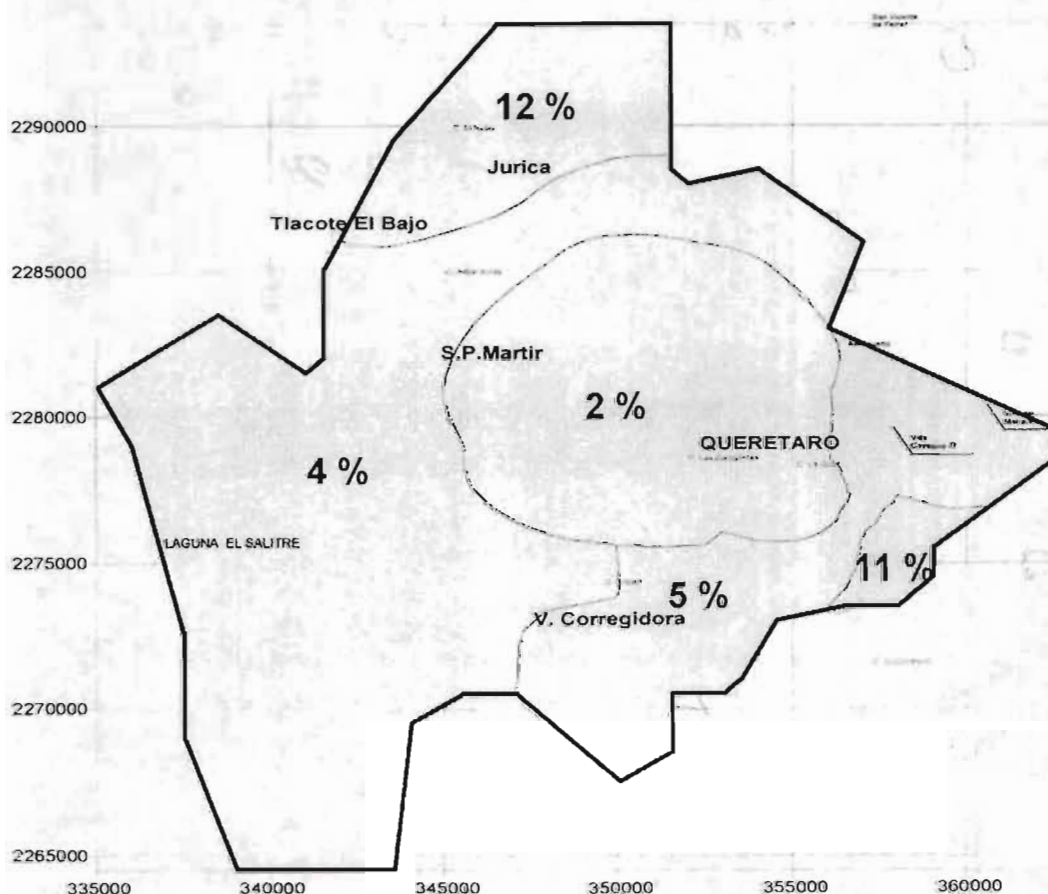


Figura 3.13 Áreas y porcentajes de infiltración de lámina precipitable por el método de los cloruros (CEAQ, 2005)

Zona	Lámina (m/día)			
	Periodo 1	Periodo 2	Periodo 3	Periodo 4
2%	0.00003007	0.00002685	0.00003562	0.00003123
4%	0.00006014	0.00005370	0.00007123	0.00006247
5%	0.00007518	0.00006712	0.00008904	0.00007808
11%	0.00016539	0.00006548	0.00019589	0.00017178
12%	0.00018043	0.00016110	0.00021370	0.00018740

Tabla 3.5 Porcentajes de infiltración para cada zona por cada periodo.

La recarga inducida está asignada de acuerdo al coeficiente de infiltración estimado en el balance de masa realizado por Simuta (2005) es constante para todos los periodos de modelación (figura 3.14).

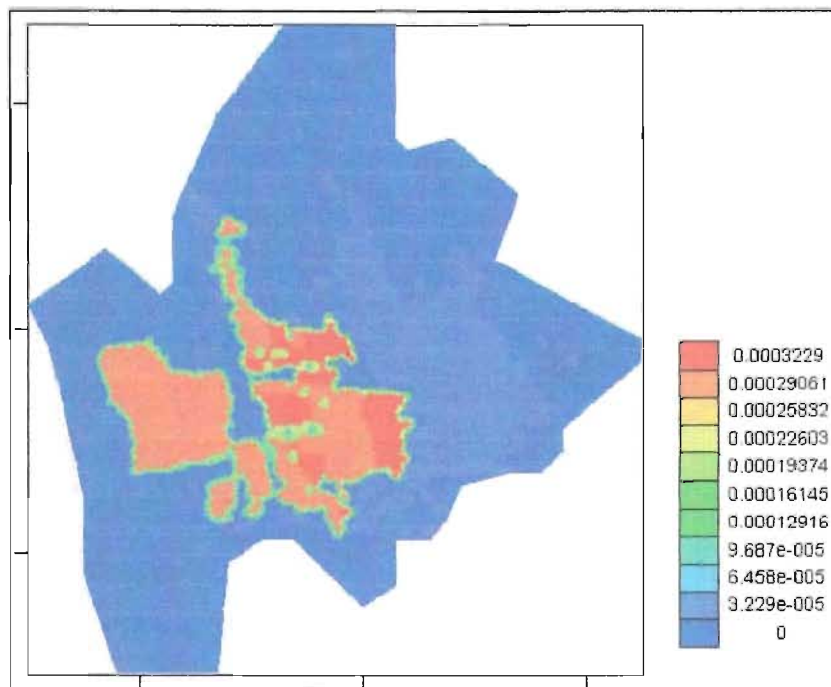


Figura 3.14 Zonas de retorno por riego y filtración de agua potable.

3.2.6 Conductividad hidráulica

La conductividad hidráulica inicial utilizada en la calibración del modelo numérico fue tomada de Simuta (figura 3.15).

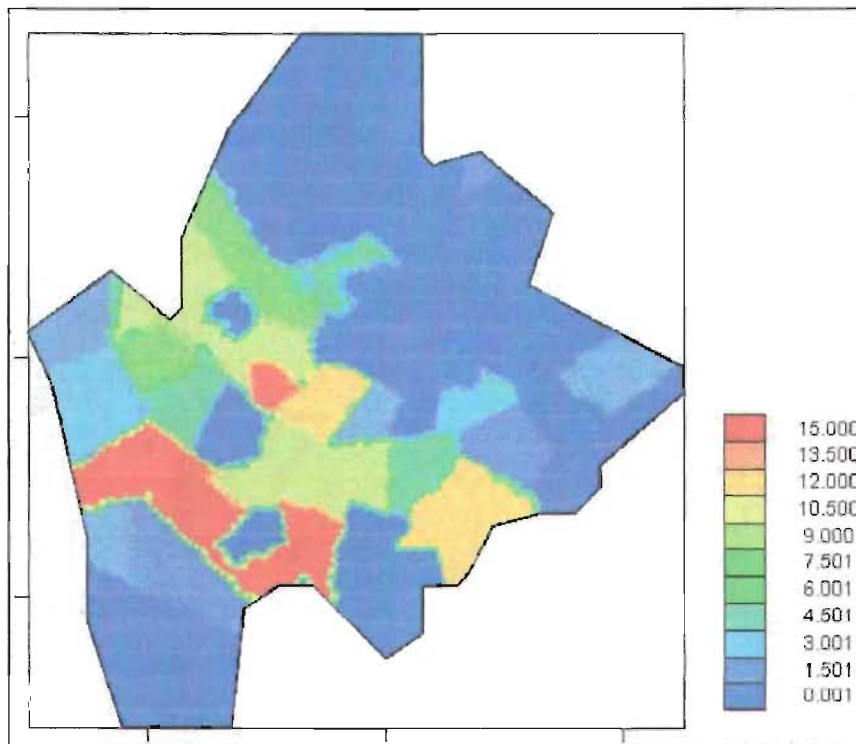


Figura. 3.15 Conductividad hidráulica de entrada al modelo (m/d).

3.2.7 Coeficiente de almacenamiento

En base al rendimiento específico calculado del balance de masas por Simuta (2005) se tiene que el coeficiente de almacenamiento para la zona de medio poroso (zona del valle) es de 0.10 y para el medio fracturado (zona montañosa) de 0.08 (figura 3.16). Debido a la gran cantidad de depósitos de arcilla en la cuenca del valle de Querétaro los valores del coeficiente de almacenamiento son bajos ya que este material impide el paso de agua en el subsuelo.

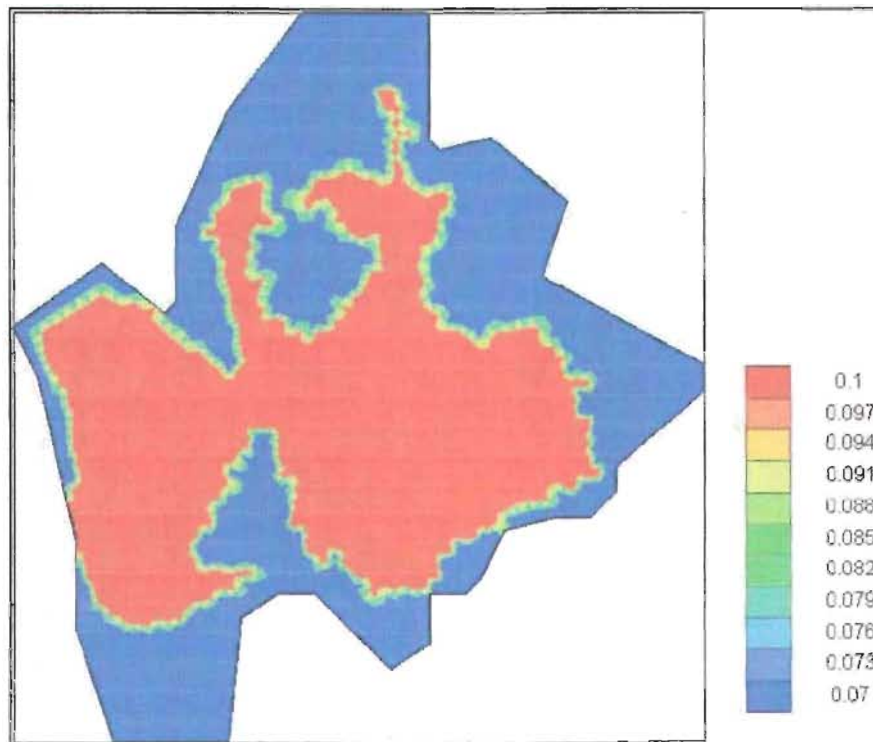


Figura 3.16 Coeficiente de Almacenamiento.

3.2.8 Calibración

El modelo se calibró en estado transitorio a partir de 1985 utilizando cuatro periodos de simulación: 1985-1990, 1990-1993, 1993-1995 y 1995-1999, el periodo de 95-99 fue utilizados para verificar el modelo. El procedimiento de calibración fue a prueba y error, y la carga inicial utilizada es el esquema de 1985, obtenido de la interpolación de los datos existentes para ese año. En la calibración del modelo los parámetros ajustados fueron la conductividad hidráulica (figura 3.17) y el coeficiente de almacenamiento (figura 3.18). Los errores para los periodos de calibración y verificación de los años para los que se utilizan datos en el diseño de la red de monitoreo se muestran en la tabla 3.6.

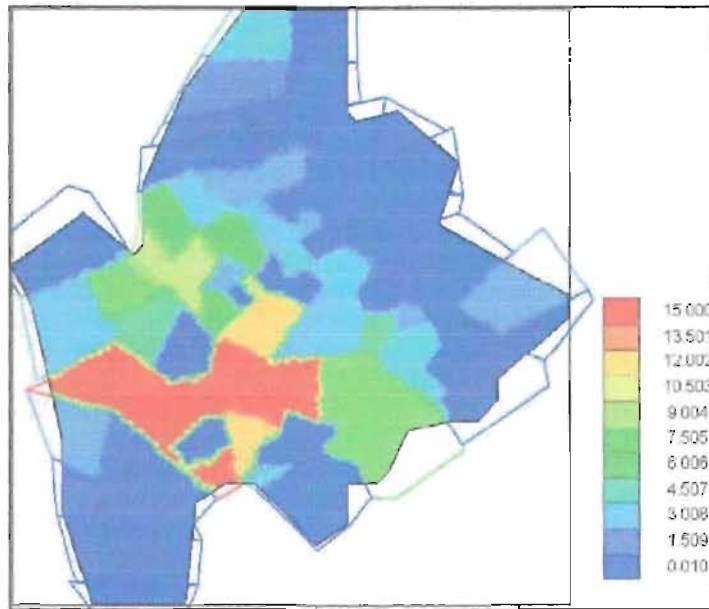


Figura 3.17. Valores de conductividad hidráulica calibrados en m/día.

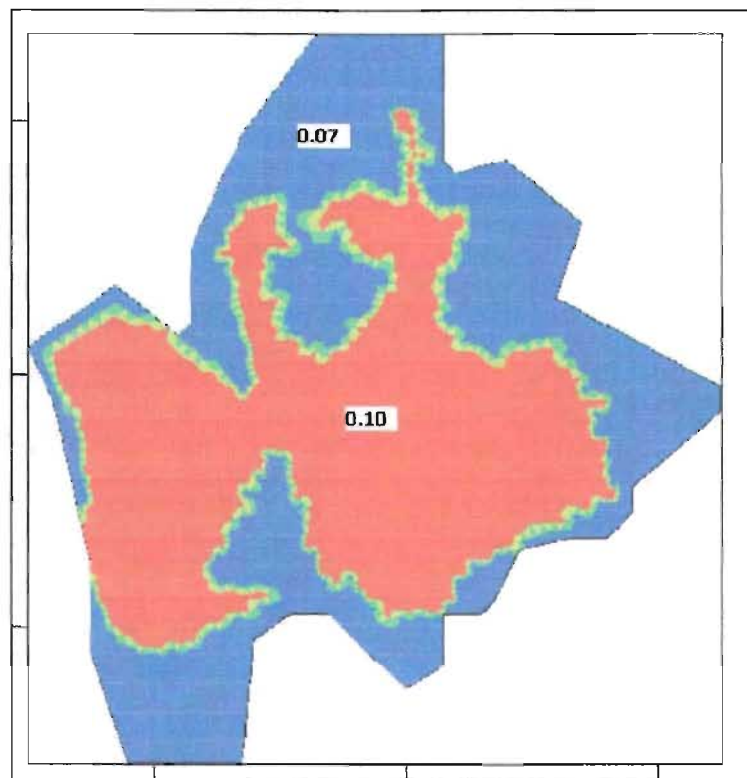


Figura 3.18. Valores de coeficiente de almacenamiento calibrados en m/día.

Medida del error	Periodo		
	90-93	93-95	95-99
Error Cuadrático Medio (m):	5.04	5.35	5.19
Error medio (m)	-0.64	-1.04	-0.80
Error absoluto medio (m)	4.01	4.37	3.89

Tabla 3.6. Errores de calibración y verificación.

3.2.9 Análisis sensitivo

Los parámetros que variaron en este caso fueron la conductividad hidráulica (K), el rendimiento específico (Sy) y la recarga vertical (Rv) por ser los parámetros en los que más incertidumbre se tiene. Estos parámetros cambiaron en un -10 %, -20 %, +10% y +20%. En la figuras 3.19 y 3.20 se muestran los resultados del análisis de sensibilidad de cada parámetro para el año de 1999 y 2030 respectivamente. El año de 1999 fue utilizado en el análisis de sensibilidad porque es el último año considerado en la calibración del modelo y el año 2030 por ser un escenario de predicción que se explica en la sección siguiente..

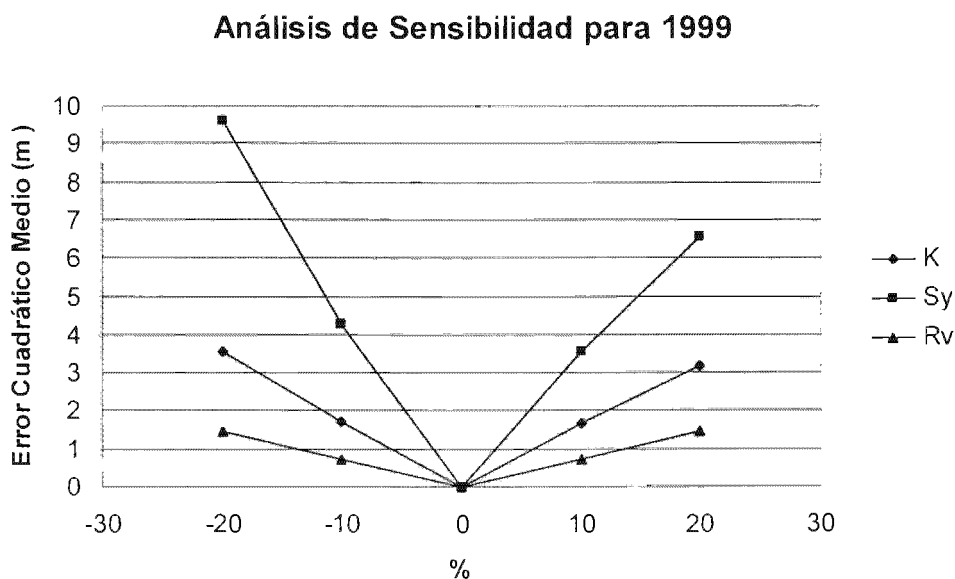


Figura 3.19. Análisis de sensibilidad para el año 1999.

Análisis de Sensibilidad para 2030

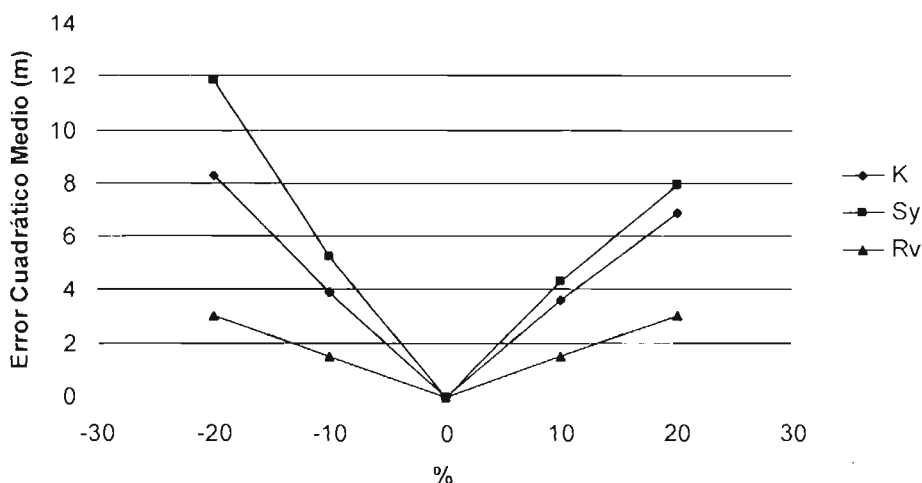


Figura 3.20. Análisis de sensibilidad para el año 2030.

Del análisis de sensibilidad se observa que para el año 1999 y 2030 hay más sensibilidad al disminuir los valores de S_y y K que al aumentarlos, no así para la recarga vertical, para el que se observa que mantiene la misma sensibilidad al aumentar o disminuir el valor del parámetro, aunque a este parámetro es poco sensible el modelo. El parámetro al que es más sensible el modelo es el rendimiento específico. Para todos los parámetros es siempre mayor la sensibilidad en el año 2030.

3.3 Predicción

De un conjunto de predicciones, en los cuales se hacen suposiciones a cerca del comportamiento que tendrá el acuífero del Valle de Querétaro en los años 2010, 2020 y 2030 en base a diferentes escenarios (políticas de operación), se tiene que el escenario que arroja mejores resultados, es el que considera el volumen de extracción que corresponde a la demanda de agua menos el volumen generado por distintas fuentes alternas al acuífero y por el reuso del agua, además de la presa Extóraz aportando 1250 lps en el 2009, 625 lps más en el 2013 y 625 lps más en el 2017. Estas consideraciones fueron hechas por la Comisión Estatal de Aguas de Querétaro, que además decidieron que es la mejor manera de operar al acuífero para su recuperación.

De estas fuentes de agua potable alternas, se consideraron en la simulación predictiva, además de la presa Extóraz, únicamente las que se encuentran ya en operación, las que ya están en proceso de construcción y a las que se les está haciendo el proyecto operativo actualmente. Éstas son:

1. Sistema Cimatario Sur. Este sistema se compone de cinco pozos ubicados al sur de la ciudad, en las inmediaciones del parque nacional Cimatario, aproximadamente a 15 km de la zona conurbada de la ciudad de Querétaro. Este sistema entró en operación en el 2004 y aporta un gasto de 150 lps.
2. Sistema Poniente II. Propocionará un gasto de 500 lps y se tiene planeado que entre en operación en el año 2005.
3. Sistema Norte III. Propocionará un gasto de 90 lps y se tiene planeado que entre en operación en el año 2005.
4. Sistema Batán-Quiotillos. Propocionará un gasto de 400 lps y se tiene planeado que entre en operación en el año 2005.
5. Sistema Paraiso-La Noria. Propocionará un gasto de 250 lps a partir del año 2006 y 150 lps más en el año 2007.
6. Presa Batán. Aportará un gasto de 80 lps a partir del 2008.
7. Presa Extóraz. El proyecto tiene como objetivo la construcción de una presa de almacenamiento en la subcuenca del río Extóraz, que permita el abastecimiento de agua potable de la ciudad de Querétaro. La entrada en operación de esta presa se tiene planeada para el año 2009, aportaría 1250 lps en este año, 625 lps más en el 2013 y 625 lps más en el 2017.

Además, la CEAQ tiene contemplado iniciar el reuso del agua en el año 2008, de lo que se obtendrían 150 lps de agua potable.

Por el motivo de que la CEAQ planea realizar este escenario de operación para la recuperación del acuífero, se decidió que los resultados de esta predicción se utilizarán para el diseño de la red de monitoreo, que es el objetivo de este trabajo de tesis.

Las configuraciones piezométricas de los resultados de este escenario muestran que para el año 2010, el nivel representativo en el valle de Querétaro y Obrajuelo es de 1660 m.s.n.m.,

teniéndose un abatimiento en el centro de la ciudad de Querétaro que llega a los 1630 m.s.n.m. Para el año 2020 los resultados muestran niveles que van de los 1660 m.s.n.m., en el valle de Querétaro; a los 1630 m.s.n.m., en el valle de Obrajuelo; teniéndose una dirección del flujo preferencial en sentido Este-Oeste. Y para el año 2030 el nivel piezométrico va de los 1660 m.s.n.m. en el valle de Querétaro a los 1610 m.s.n.m en el valle de Obrajuelo; teniéndose una dirección del flujo preferencial en sentido Este-Oeste. Se observa además un abatimiento en el centro del valle de Querétaro con un nivel piezométrico de 1600 m.s.n.m. En las figuras 3.21, 3.22 y 3.23 se muestran la piezometría simulada para el año 2010, 2020 y 2030, respectivamente.

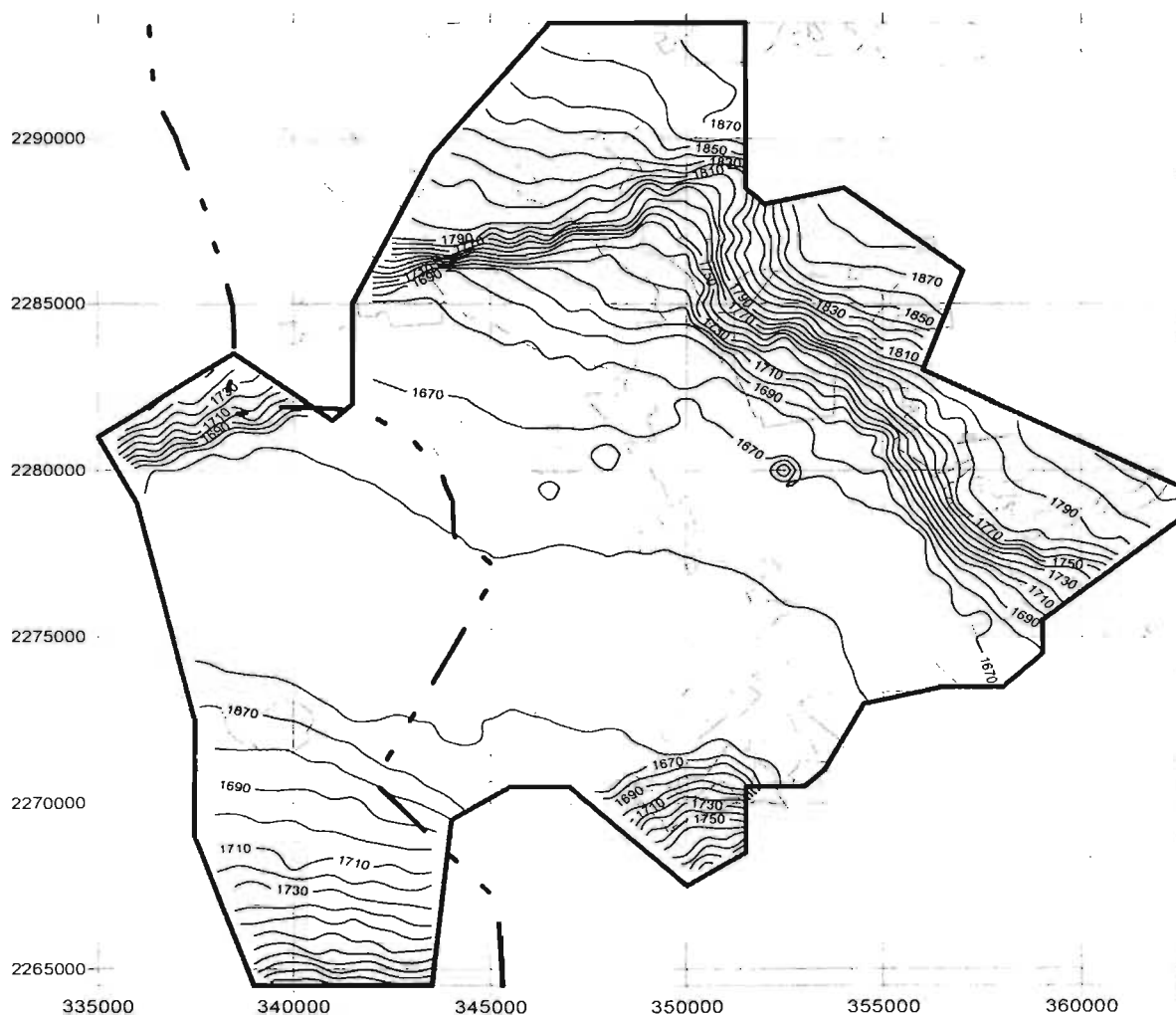


Figura 3.21. Piezometría simulada para el año 2010, en msnm (tomado de CEAQ, 2005).

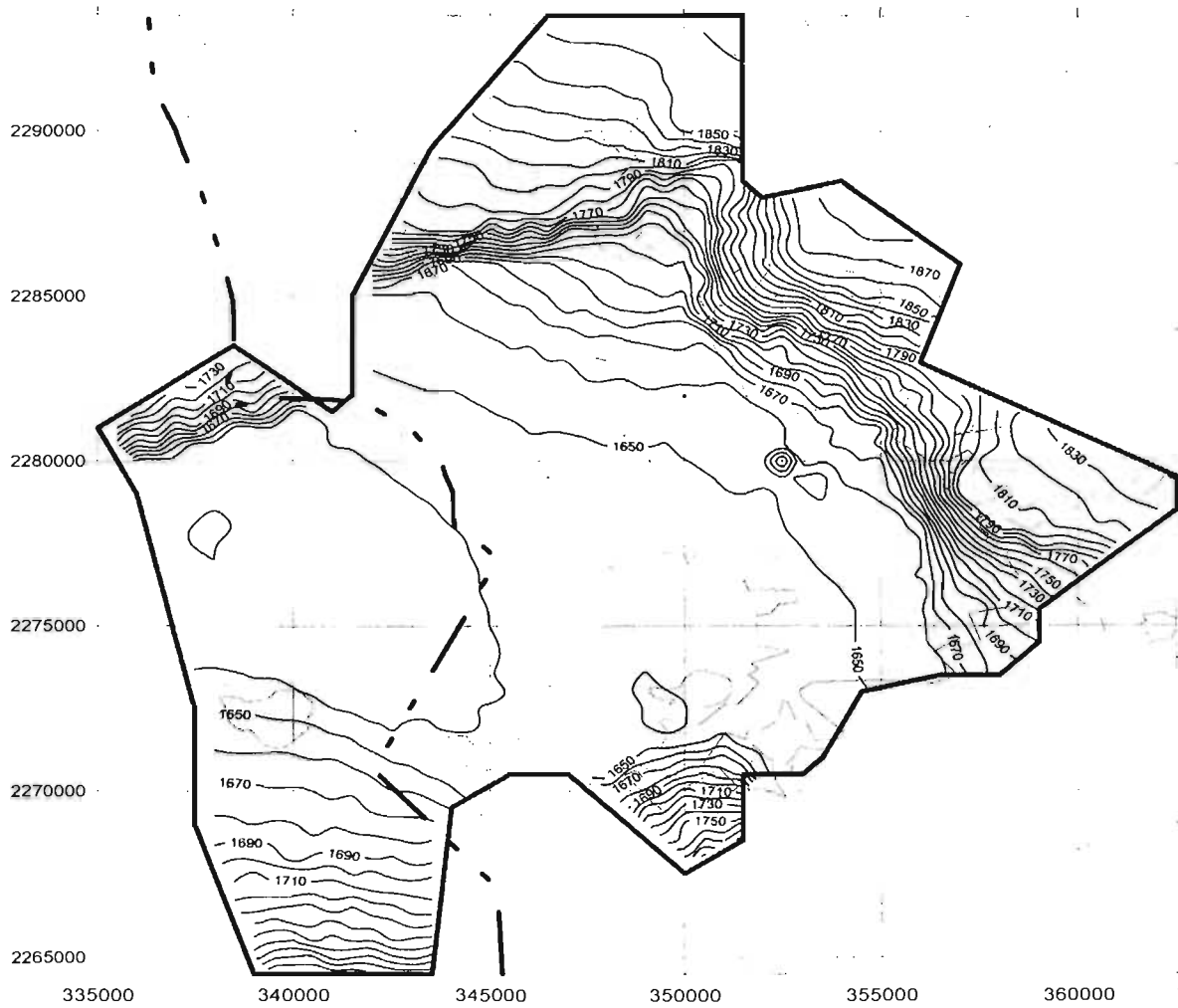
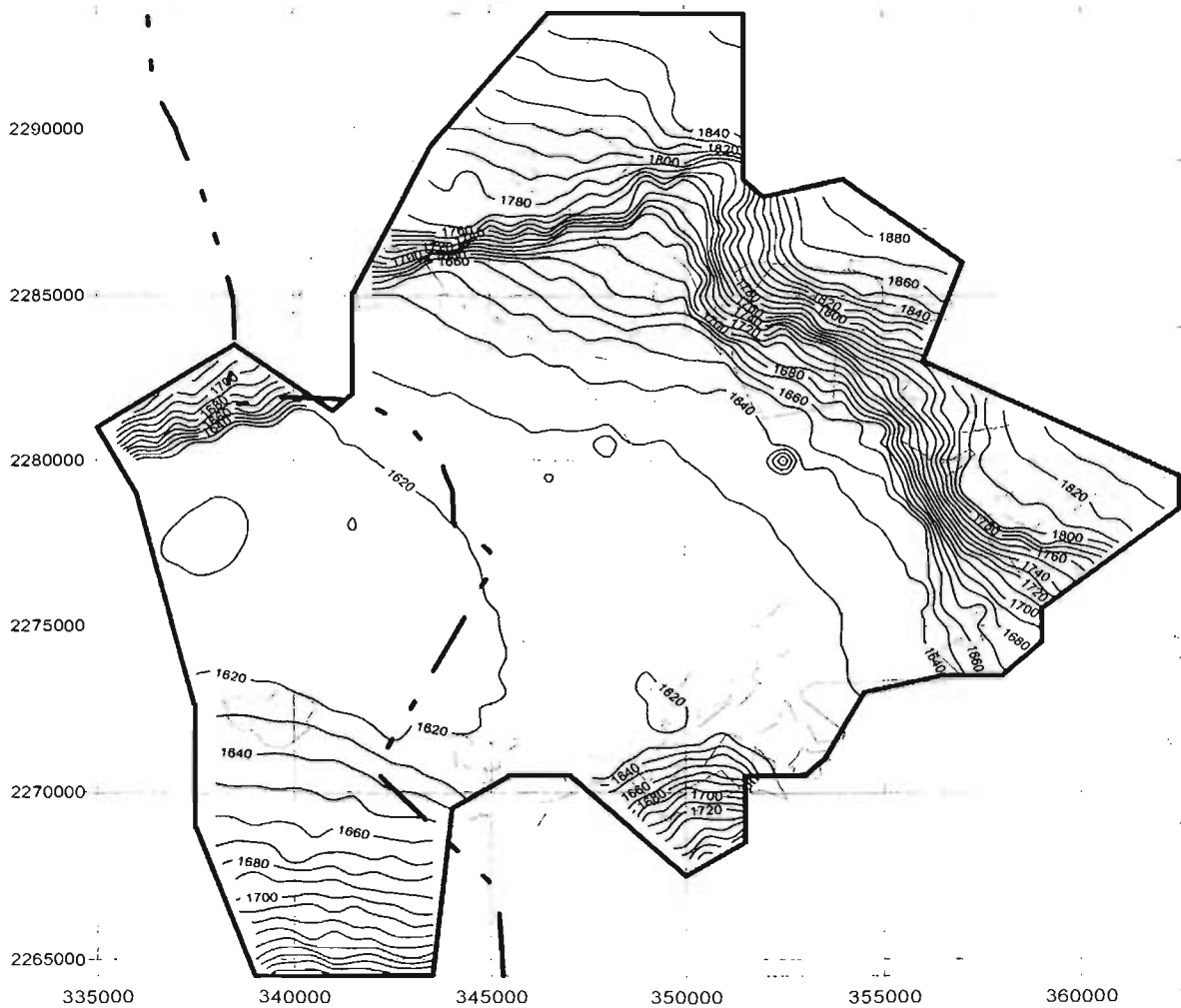


Figura 3.22. Piezometría simulada para el año 2020, en msnm (tomado de CEAQ, 2005).



Figuras 3.23. Piezometría simulada para el año 2030, en msnm (tomado de CEAQ, 2005).

3.4 Conclusiones

Con el propósito de estudiar los cambios en los niveles del agua en el acuífero del Valle de Querétaro, bajo políticas de explotación muy diferentes a las actuales como se plantea en este estudio, se estudió en forma conjunta este acuífero y el del valle de Obrajuelo, Guanajuato, ya que no posible establecer una frontera física o natural entre estos dos acuíferos y, por lo tanto, cambios importantes del nivel del agua en uno de los acuíferos afecta en forma importante los niveles en el otro acuífero.

Los errores cuadráticos medios en la calibración para los años 1993, 1995 y 1999, que son los más recientes para los cuáles se cuenta con datos suficientes, son del orden de 5 metros. Por otro lado, se confirmó que los gradientes verticales de la carga hidráulica no son despreciables por lo que es probable que estos errores no se pudieran reducir por este motivo. Por otro lado, el modelo reproduce en forma adecuada las tendencias históricas de la piezometría en la mayoría de los pozos y por tanto es útil para predecir los abatimientos promedio en el acuífero bajo diferentes políticas de extracción.

4. DISEÑO ÓPTIMO DE LA RED DE MONITOREO.

En capítulos anteriores se vio la problemática que presenta el acuífero, se hizo una revisión del marco físico regional y del modelo dinámico de flujo elaborado por el Instituto Mexicano de Tecnología del Agua por contrato para la CEAQ (2004). Con todo esto en conjunto se diseñó una red de monitoreo para evaluar el comportamiento de los niveles del agua en el acuífero del Valle de Querétaro durante la operación del proyecto de la presa Extóraz. lo cual es el objetivo de esta tesis. En la sección 2.6 se explica el método utilizado en esta tesis para proponer la red de monitoreo óptima. Los pasos que se siguieron para obtenerla, que se describen a continuación, fueron los siguientes:

- a) Análisis preliminar.
- b) Análisis geoestadístico.
- c) Construcción de la matriz de covarianza.
- d) Selección de pozos a incluir en la red de monitoreo con base en el método propuesto por Herrera.
- e) Análisis de resultados.
- f) Estimación de errores en puntos de monitoreo.
- g) Propuesta final de la red de monitoreo.

Para el diseño de la red de monitoreo en este trabajo es necesario que el semivariograma que se utilice sea adecuado para hacer estimaciones en los años en que opere la presa. Esto se hizo aquí, al igual que en la simulación numérica, para el periodo 2003 al 2030. Se presenta a continuación un análisis geoestadístico de los abatimientos anuales en los periodos 1993-1995 y 1995-1999, esto es, de los valores de piezometría para cada uno de estos años se obtuvo el abatimiento por medio de sus diferencias. Se eligieron estos años porque son los más recientes en que se cuenta con mayor número de datos y que tienen un comportamiento parecido entre ellos, probablemente porque se tomaron los datos en el mismo periodo del año en todos ellos.

En el modelo numérico de flujo se incluye una zona del acuífero de Obrajuelo (Guanajuato) esto por razones del comportamiento del flujo subterráneo, pero en el diseño de la red de monitoreo sólo es de interés la parte del acuífero del Valle de Querétaro, ya que es la zona que

la CEAQ monitoreará, por tal motivo los pozos del acuífero de Obrajuelo no son incluidos en este análisis.

La teoría de probabilidad y estadística, geoestadística, y filtro de Kalman se presentan en el anexo I.

4.1 Análisis preliminar.

El estudio conceptual del fenómeno implica entender sus causas y presupone un conocimiento de la ciencia, en este caso de la geohidrología. Puede requerir una interpretación preliminar de los datos al objeto de eliminar aquellos que sean claramente erróneos y confirmar o desechar alguna hipótesis.

Para validar un análisis geoestadístico de datos se deben cumplir los siguientes requisitos:

- Que su distribución de probabilidad sea parecida a la normal.
- No exista tendencia.
- No se vea afectada por valores atípicos (outliers), tanto distribucionales como espaciales; los primeros son los que tienen valores que se alejan significativamente del valor medio de la muestra y los segundos son los que tienen un valor diferente al de sus vecinos más cercanos.
- Distribución espacial homogénea.

La base de datos (tabla II.1, anexo II) proporcionada por la CEAQ, cuenta con datos de piezometría en 139 pozos en 1993, 164 en 1995 y 135 en 1999. De estos valores piezométricos se obtuvieron los abatimientos anuales para cada periodo, es decir, se tiene el abatimiento por periodo de 1993 a 1995, 1995 a 1999 y se divide entre el número de años, con esto se obtiene el abatimiento anual para cada periodo. El resultado es de 105 datos de abatimiento anual para el periodo 1993 a 1995 y de 92 para 1995 a 1999. La cantidad de datos para el segundo periodo es menor debido a que en algunos pozos no se tienen datos en alguno de los tres años analizados. La localización de los pozos para los abatimientos anuales del periodo 1993 a 1995 se ve en la figura 4.1 y para el periodo 1995 a 1999 en la figura 4.2.

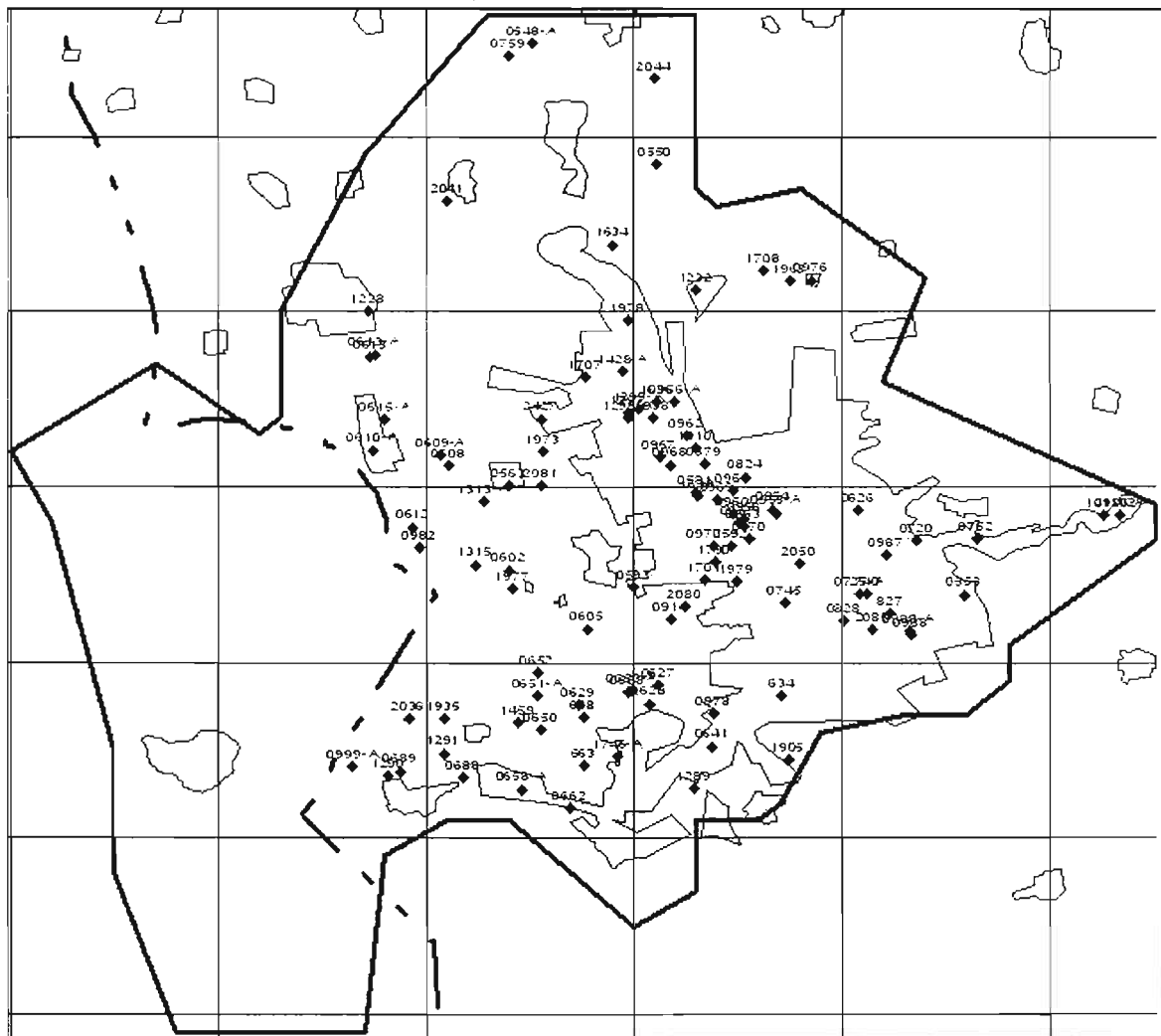


Figura 4.1 Localización de pozos en el periodo 1993 a 1995.

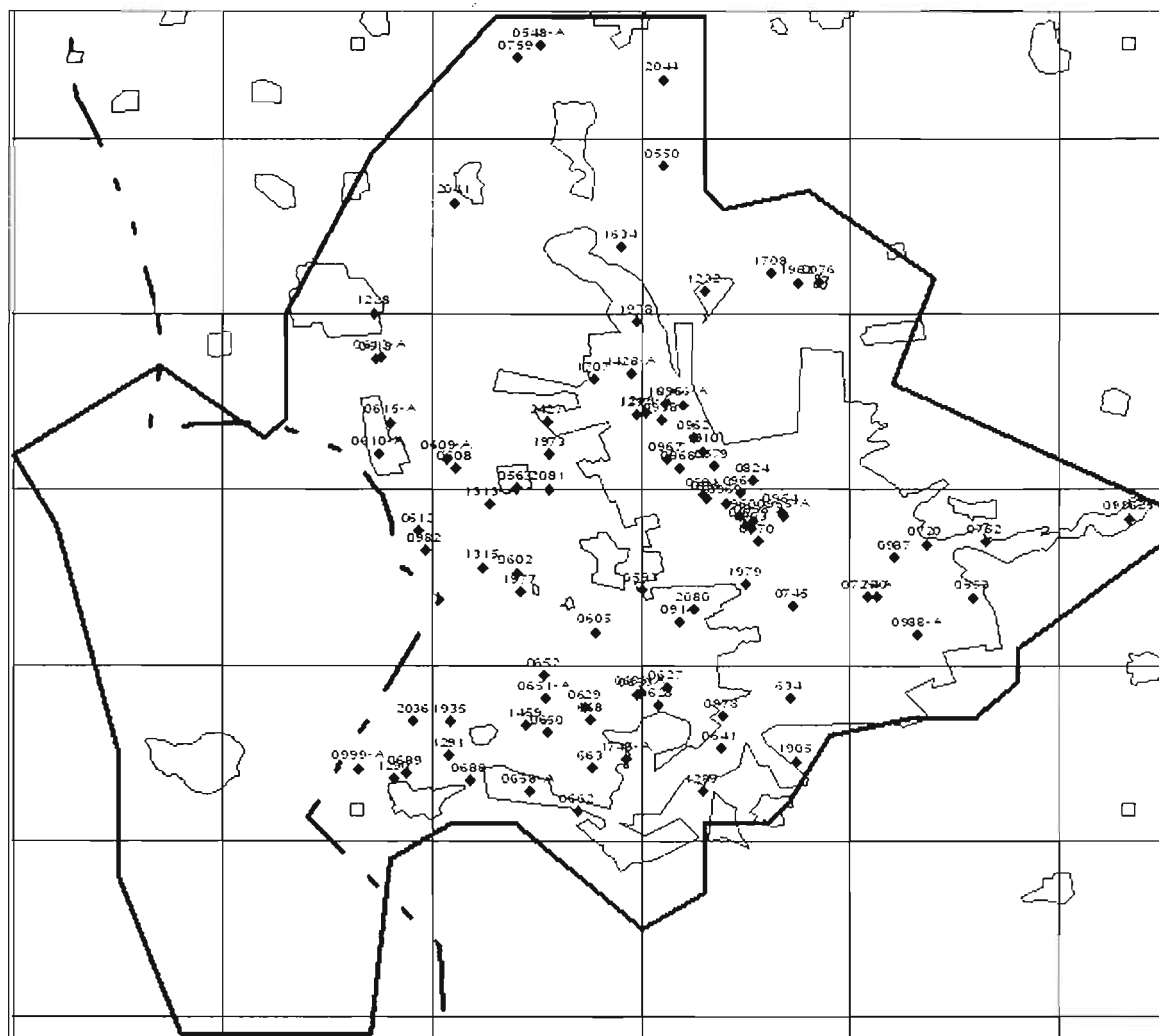


Figura 4.2 Localización de pozos para el periodo 1995 a 1999.

El estudio de los datos requiere el cálculo de estadísticos (media, desviación estándar, etc.) y la representación mediante histogramas de frecuencias. El objetivo de esta fase es definir su distribución de probabilidad, estudiar si los datos provienen de una o varias poblaciones, revisar los datos extremos para evaluar si son erróneos, y por lo tanto eliminarlos o adecuarlos. Los estadígrafos de los abatimientos anuales de los dos periodos analizados se muestran en la figuras 4.3 y 4.4.

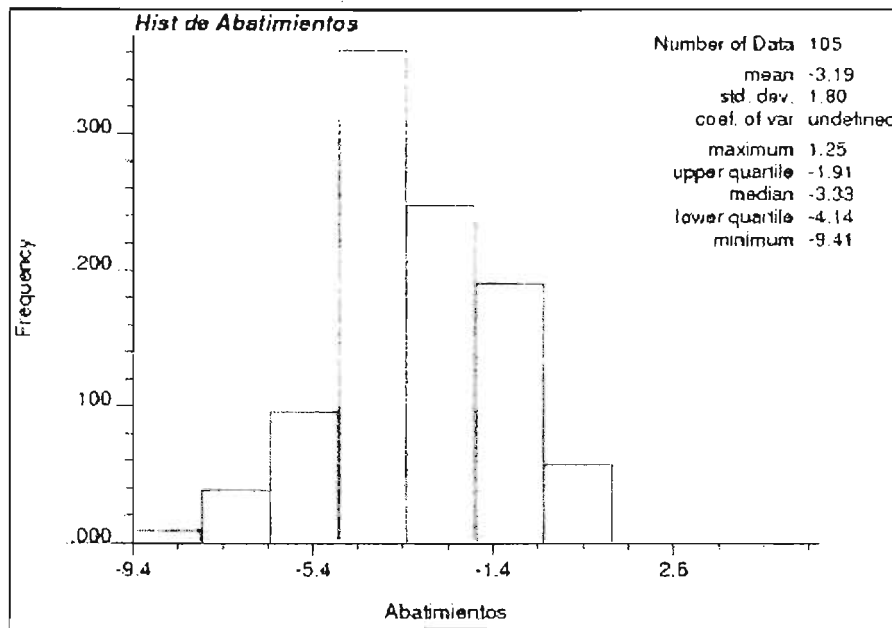


Figura 4.3 Estadígrafos de abatimientos anuales (1993-1995).

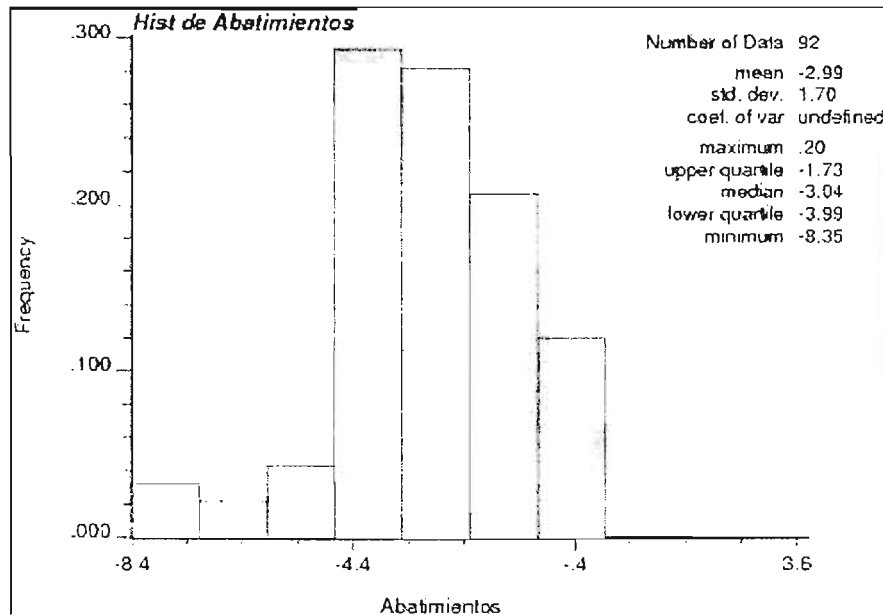


Figura 4.4 Estadígrafos de abatimientos anuales (1995-1999).

En términos generales se tiene que los dos periodos cuentan con una buena distribución, la asimetría es leve, y la media y mediana son muy parecidas, esto indica que el semivariograma será válido sin necesidad de eliminar datos atípicos (*outlayers*) o hacer alguna transformación de los datos.

4.2 Análisis geoestadístico.

Isaaks y Srivastava (1989) definen a la geoestadística como una herramienta que “ofrece una forma para describir la continuidad espacial que es una característica esencial de muchos fenómenos naturales y provee adaptaciones de técnicas de regresión clásicas para aprovechar esta continuidad”. La variable que se analiza puede visualizarse como una variable aleatoria definida en todos los puntos del espacio y con correlaciones espaciales. La geoestadística se ha empleado en la mayoría de las ciencias de la tierra: geología, geotecnia, minería, edafología, hidrología, meteorología, etc.

Las aplicaciones de la geoestadística a la geohidrología se remontan a los años setenta; este enfoque ha supuesto un revulsivo ya que el tratamiento de las variables geohidrológicas como funciones aleatorias ha conducido a una revisión de los conceptos clásicos de geohidrología cuantitativa.

La mayoría de los modelos matemáticos usados en geohidrología se basan en la teoría clásica del flujo en medios porosos. Los métodos de análisis se basan en estas ecuaciones en derivadas parciales considerando que sus parámetros son funciones determinísticas en el espacio. En la práctica, la solución de estas ecuaciones se obtiene mediante modelos matemáticos numéricos en los que los parámetros se representan mediante funciones continuas o funciones continuas por zonas o simplemente discontinuas. Dada la escasez de información sobre los parámetros, estas funciones tienden a representar la variación espacial de los parámetros de una forma uniforme y suavizada. Los parámetros hidrogeológicos muestran una variabilidad espacial errática que difícilmente puede ser representada mediante funciones determinísticas. a esto hay que añadir la variabilidad inducida por los métodos empleados para su determinación.

Una forma de tener en cuenta la aleatoriedad de las variaciones de los parámetros geohidrológicos y la incertidumbre asociada con la insuficiente información acerca de su distribución espacial consiste en adoptar una interpretación probabilística. La interpretación estadística de los parámetros geohidrológicos tiene la clara ventaja de que permite estudiar su incertidumbre y por tanto la de las predicciones obtenidas en base a los mismos.

En la solución de la ecuación del flujo subterráneo mediante métodos numéricos, los parámetros suelen representar valores medios en ciertas zonas, dada la incertidumbre asociada a este proceso de estimación, los parámetros estimados se convierten en variables aleatorias, el modelo numérico es estocástico y sus predicciones son por tanto inciertas. Tanto los modelos determinísticos como los estocásticos requieren conocer los estadísticos de la variación espacial de los datos. el papel de la geoestadística es precisamente proporcionar herramientas para analizar la variabilidad espacial de los parámetros y variables geohidrológicas.

Los fenómenos distribuidos en el espacio presentan un carácter mixto, un comportamiento caótico o aleatorio a nivel local; pero a la vez estructural a gran escala, de donde se puede sugerir la idea de interpretar este fenómeno en términos de una función aleatoria (FA). es decir, a cada punto x del espacio se le asocia una variable aleatoria (VA) $Z(x)$, para dos puntos diferentes x e y , se tendrán dos VAs $Z(x)$ y $Z(y)$ diferentes pero no independientes. y es precisamente su grado de correlación el encargado de reflejar la continuidad de cualquier fenómeno en estudio (Matheron, 1975).

De forma general, la geoestadística, se construye suponiendo condiciones de estacionaridad y se define como el estudio de las variables numéricas distribuidas en el espacio (Chauvet, 1994), o variables regionalizadas, porque a cada valor observado o desconocido está asociada una posición en el espacio. Su estimador, el kriging, tiene como objetivo encontrar el mejor estimador lineal insesgado a partir de la información disponible (Chica-Olmo, 1997), y en efecto, el valor estimado obtenido $Z^*(x)$ de un valor real y desconocido $Z(x)$, consiste en una combinación lineal de pesos asociados a cada localización donde fue muestreado un valor $Z(x_i)$ ($i = 1, \dots, n$) del fenómeno estudiado, observando dos condiciones fundamentales:

- Que el estimador sea insesgado. $E[Z^* - Z] = 0$, y
- Que la varianza $\text{Var}[Z^* - Z]$ sea mínima, consiguiéndose de este modo minimizar la varianza de error de estimación.

El kriging, utiliza en la estimación las características de variabilidad y correlación espacial del fenómeno estudiado, por lo que su uso implica un análisis previo de la información, con el objetivo de definir o extraer de esta información inicial un modelo que represente su

continuidad espacial, una vez logrado esto, estamos en condiciones de obtener a través del kriging el mejor valor posible en cada localización o bloque a estimar, acompañadas de la varianza kriging como medida del error de la estimación realizada (Armstrong y Carignan, 1997).

4.2.1 Análisis estructural.

Después de realizar el análisis estadístico y determinar el tipo de distribución de probabilidad a la que obedece la variable en estudio, se procede a realizar el análisis estructural, es decir, a la construcción e interpretación de semivariogramas.

Esta etapa consiste esencialmente en el ajuste de la función semivariograma a la variabilidad espacial *in-situ* que caracteriza al fenómeno en estudio.

El semivariograma es la herramienta fundamental de la geoestadística, pues aunque no describe la totalidad de las características del fenómeno en estudio, sí describe en forma sintética sus características más importantes; de esta manera, mediante el semivariograma se pueden cuantificar e interpretar importantes rasgos geológicos, tales como:

1. El área de influencia de una muestra respecto de su vecindad.
2. La medida de la continuidad del fenómeno en estudio.
3. Conocer el comportamiento del depósito de acuerdo con sus tendencias direccionales, o dicho en otras palabras, obtener una medida de la isotropía o anisotropía del fenómeno.

Siendo el semivariograma una función representativa de la variabilidad espacial del fenómeno, se define como la varianza de la diferencia de dos variables aleatorias regionalizadas, separadas una de la otra por una distancia h .

La expresión matemática del semivariograma experimental:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 = \frac{1}{2} E \left[\{Z(x_i) - Z(x_i + h)\}^2 \right]$$

en donde:

$Z(x_i)$ es el valor de la variable aleatoria regionalizada en la posición x_i .

$Z(x+h)$ es el valor de la variable aleatoria regionalizada en la posición $x_i + h$.

$N(h)$ es el número de pares formados a una distancia h .

h es el intervalo o distancia de separación entre cada par de variables aleatorias regionalizadas.

A efecto de obtener los conceptos anteriores, después de construido el semivariograma se ajusta un modelo matemático. El objetivo de dicho ajuste es tener el semivariograma expresado mediante una fórmula matemática representativa del fenómeno o depósito en estudio. Consiste en buscar una función analítica que represente adecuadamente los valores estimados del semivariograma. Los modelos teóricos que pueden ser usados son: el esférico, el gaussiano, el exponencial, el lineal con meseta y sin meseta, y el de potencias. Estos modelos se caracterizan fundamentalmente por tres parámetros: el alcance o radio de correlación espacial, la meseta (sill) y la microvarianza o varianza nugget, y en algunos casos la potencia y pendiente (para una descripción más detallada de estos modelos ver el anexo I).

Este ajuste de la fórmula matemática representativa del fenómeno en estudio se hace a prueba y error de manera visual, modificando los parámetros del modelo hasta obtener una buena modelación, o en su defecto, para un mejor ajuste se utiliza el criterio de selección de modelos, esto lo ofrece el Criterio de Información de Akaike (CIA) en donde intervienen el número de valores estimados del variograma muestral, la suma residual de los cuadrados de las diferencias entre los valores experimentales y los del modelo ajustado, además del número de parámetros del modelo ajustado (la fórmula matemática para el CIA se presenta en el anexo I). Cuando se encuentra el menor CIA se considera que se obtiene el mejor modelo.

Para construir el semivariograma muestral para el abatimiento anual del periodo 1993-1995 se determinó la distancia máxima que hay entre los pozos más alejados (21821.89 m), con esta distancia se definen los intervalos entre cada par de variables aleatorias regionalizadas, es decir, la mitad de la distancia máxima dividida entre 10 que es el número de intervalos que se

desea tener (estas consideraciones se tomaron de Díaz Martín, 2002). La distancia de cada intervalo es de 1091.09 m con una tolerancia de 545.55 m. Una vez construido el semivariograma experimental se procedió a ajustarle un modelo, este modelo es del tipo esférico con un nugget de 2.1, meseta de 1.25 y alcance de 4100 m; el Criterio de Información de Akaike es de -134.06. La figura 4.5 se tiene el semivariograma muestral y el ajuste del modelo esférico.

La ecuación del modelo ajustado del tipo esférico es la siguiente:

$$\gamma(h) = a_0 + \begin{cases} c \left[3 \left(\frac{h}{a} \right) - \left(\frac{h}{a} \right)^3 \right] & h \leq a \\ c & h > a \end{cases}$$

donde

a_0 = nugget (efecto pepita)

c = varianza (nugget + meseta)

h = intervalo de distancia

a = alcance (distancia a la cual la varianza es constante)

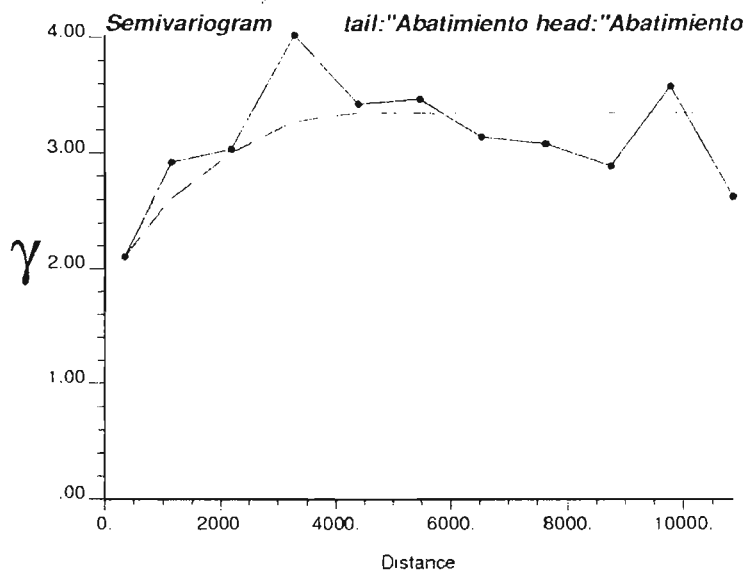


Figura 4.5 Ajuste de modelo esférico a semivariograma muestral (1993-1995).

Siguiendo con las mismas consideraciones para construir el semivariograma muestral de los abatimientos anuales del periodo 1993-1995 se realizó el semivariograma muestral del periodo 1995-1999. La distancia máxima y la distancia de cada intervalo es la misma que para el periodo anterior. El modelo ajustado es esférico con un nugget de 1.22, meseta de 1.8 y alcance de 7000 m; el Criterio de Información de Akaike da como resultado -138.64. El semivariograma muestral y modelo ajustado se muestran en la figura 4.6.

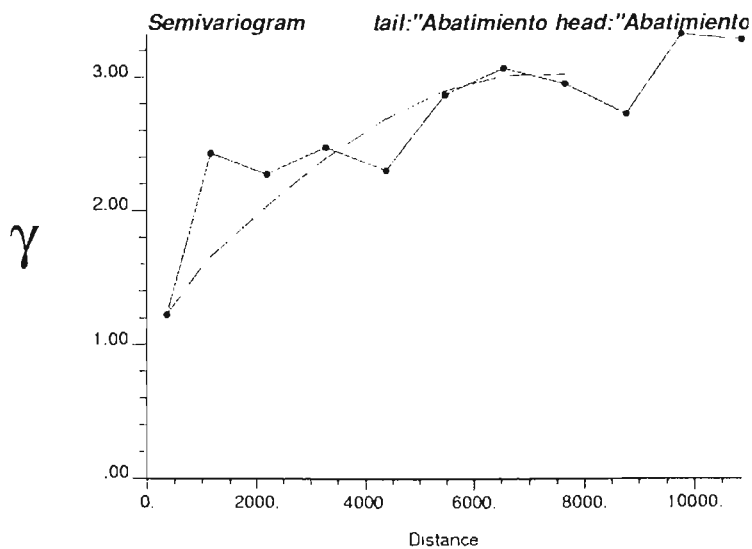


Figura 4.6 Ajuste de modelo esférico a semivariograma muestral (1995-1999).

4.2.2 Estimación.

Esta parte del análisis geoestadístico comprende el uso de la información muestreada para predecir valores en áreas que no se tienen muestreadas. El mejor estimador lineal insesgado de un punto y el mejor promedio lineal móvil ponderado de un bloque es el kriging, llamado así por Matheron alrededor de 1960. Matheron creó una técnica en Francia a partir de los trabajos de D. G. Krige quien probablemente fue el primero en hacer uso de la correlación espacial y del mejor estimador lineal insesgado en el campo de la evaluación de yacimientos minerales.

El kriging es una técnica de estimación local que a partir de una característica desconocida en estudio ofrece el mejor estimador lineal insesgado. Solamente se requiere el conocimiento de segundo orden de la función aleatoria, ya sea el variograma o la covarianza.

Una de las características que ofrece el kriging es que la varianza de la estimación no depende del valor medio de la variable. es decir, que el análisis tiene como finalidad localizar los puntos de medición de tal manera que se minimice la incertidumbre o varianza.

El objetivo es obtener una estimación $Z^*(x)$, de $Z(x)$, como una combinación lineal de los valores medidos Z_i en los puntos de observación x_i , además de obtener la varianza del error de estimación. El planteamiento del problema en general es el siguiente:

$$Z^* = \sum_{i=1}^n \lambda_i Z_i$$

λ_i se denomina coeficiente de ponderación o coeficientes de kriging. Z^* debe ser insesgado, es decir, $E(Z^* - Z) = 0$. Si se conoce el semivariograma o bien la función de covarianza, entonces basta obtener λ_i ($i = 1 \dots n$, número de puntos de observación), para que Z^* y su varianza queden perfectamente determinados. Los coeficientes λ_i se obtienen imponiendo que la varianza de estimación sea mínima. Esto da lugar a un sistema lineal de ecuaciones que se conoce con el nombre de ecuaciones de kriging.

Los datos de los modelos ajustados a cada uno de los semivariogramas permite realizar la estimación, esta se hizo por medio de kriging ordinario (las ecuaciones del kriging ordinario se

presentan en el anexo I). Se utilizó una malla de 151 celdas de longitud 122.998 m en el eje X y 151 celdas de 145.35 m en el eje Y. Para el periodo de 1993-1995 el radio de búsqueda es de 4000 m y un número de vecinos de 8. En el periodo de 1995-1999 el radio de búsqueda es de 7000 m y el número de vecinos de 8. Con esto se llegó a la solución del sistema lineal de ecuaciones, para así obtener la estimación de los abatimientos anuales por periodo con sus respectivas varianzas de los errores de estimación. La estimación de los dos periodos se muestra en las figuras 4.7 y 4.8. Las varianzas de estimación en la figuras 4.9 y 4.10.

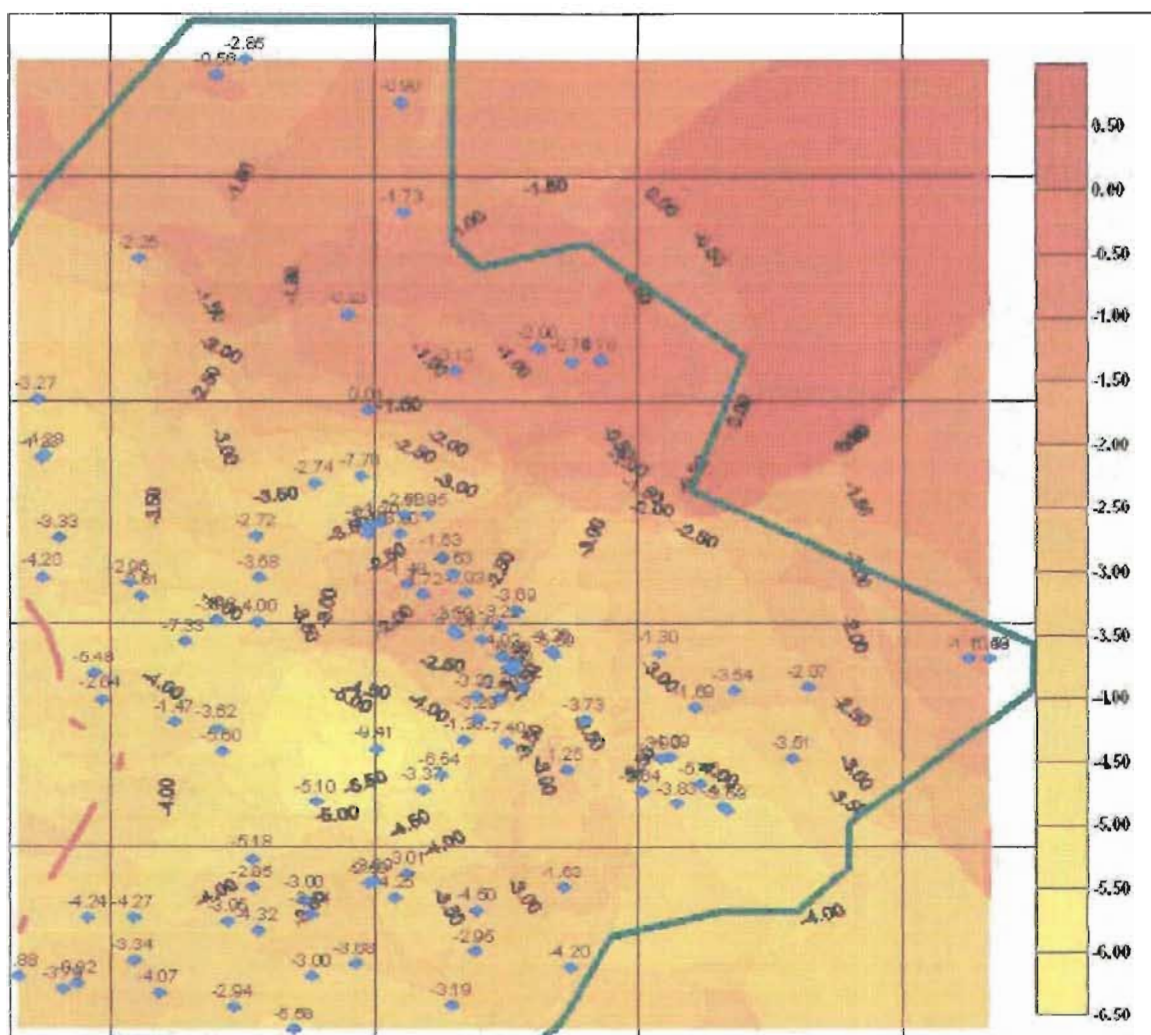


Figura 4.7 Estimación con kriging ordinario de abatimientos anuales (1993-1995).

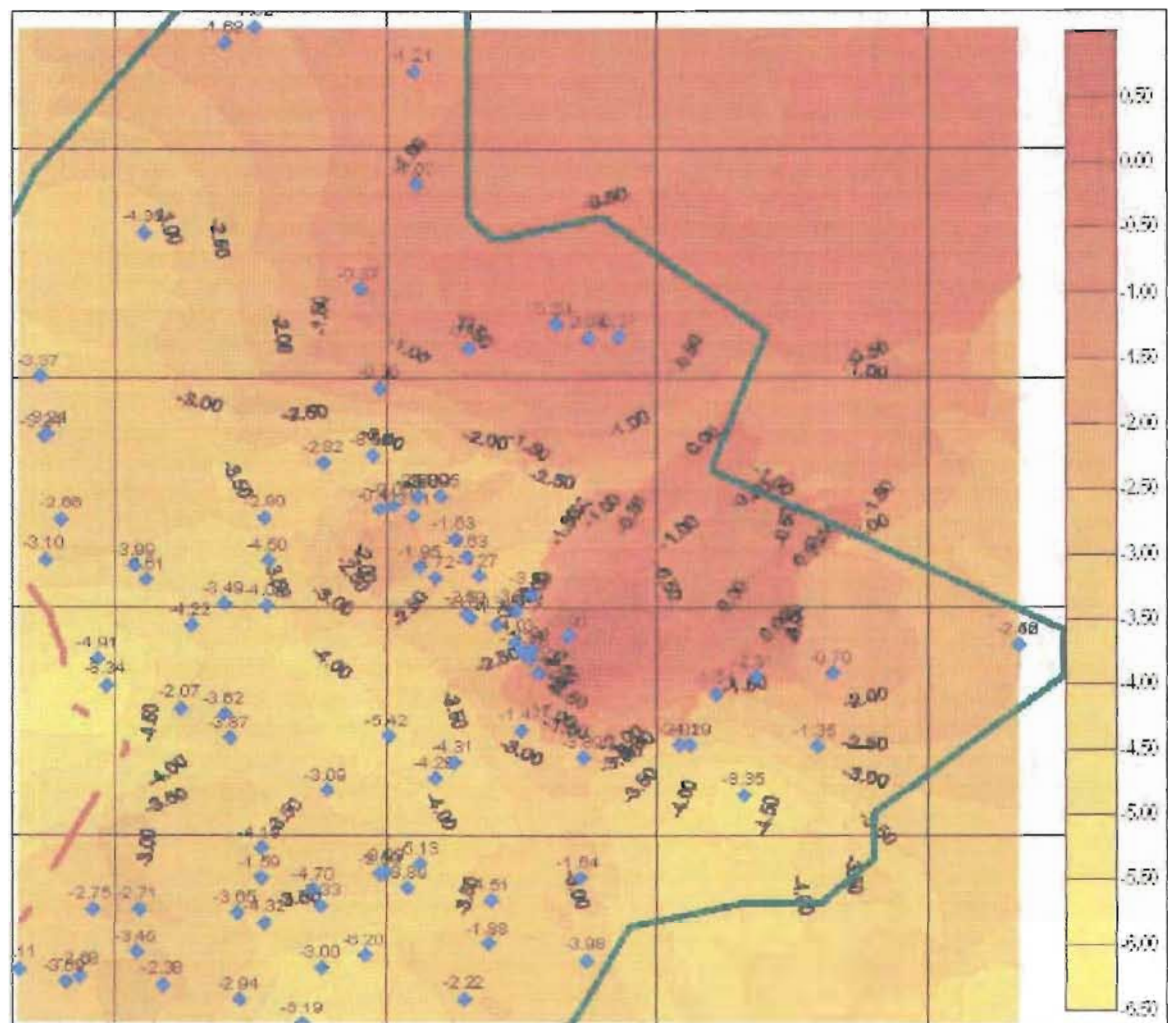


Figura 4.8 Estimación con kriging ordinario de abatimientos anuales (1995-1999).

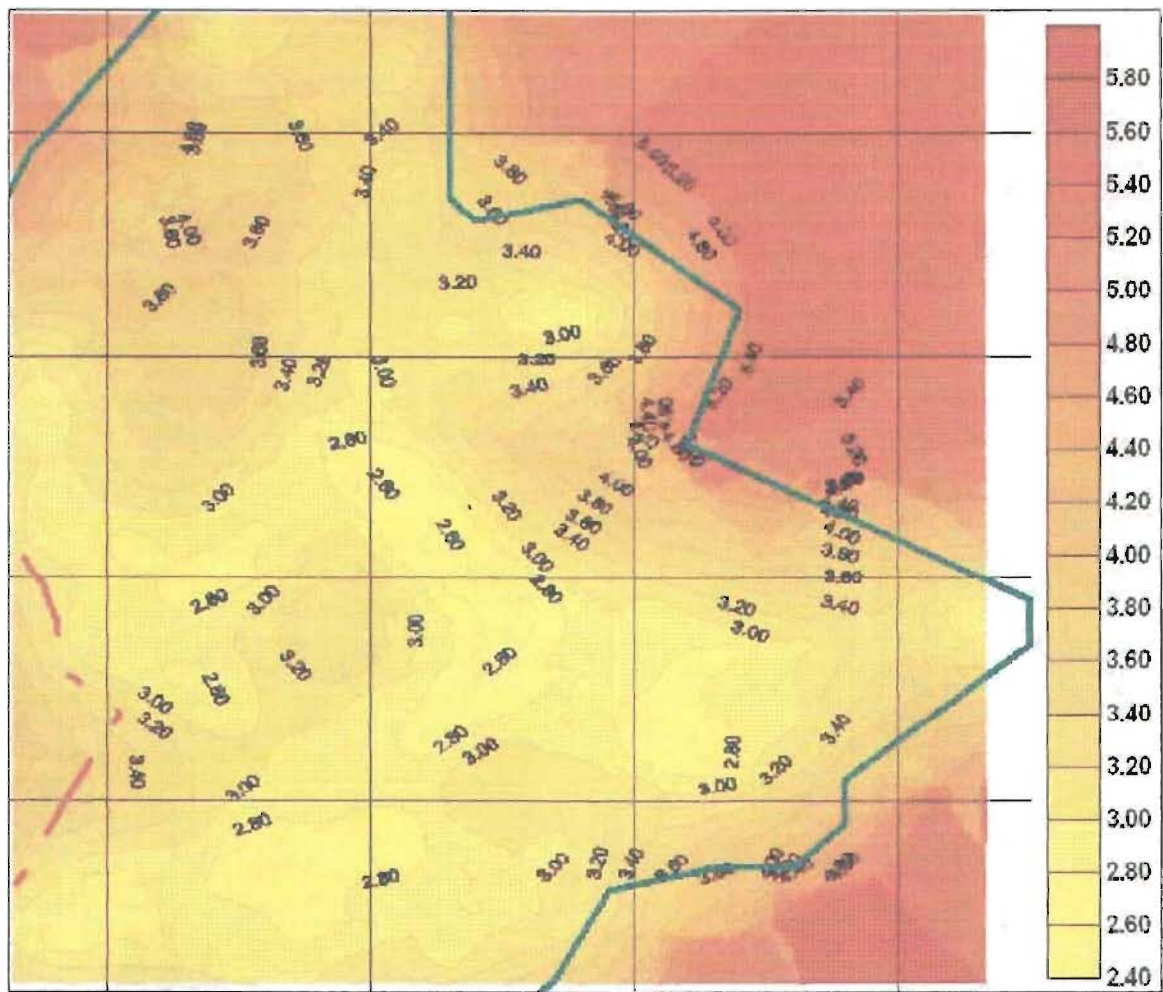


Figura 4.9 Varianza de estimación de abatimientos anuales (1993-1995).

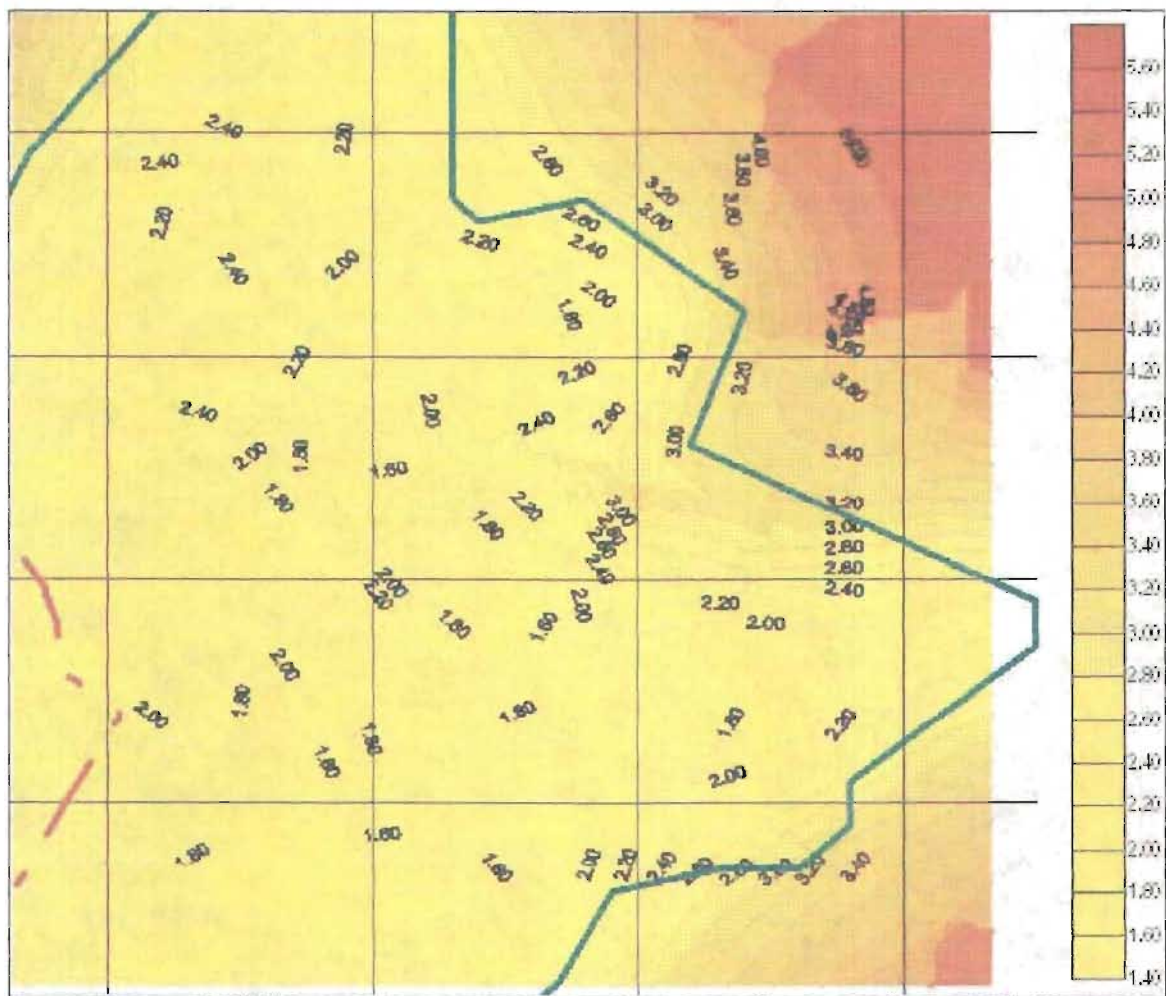


Figura 4.10 Varianza de estimación de abatimientos anuales (1995-1999).

En la estimación de los abatimientos anuales para los dos periodos se observa que los abatimientos mayores se presentan en la zona central del acuífero del Valle de Querétaro y disminuyen hacia las zonas en donde existe entrada de flujo subterráneo.

La varianza de estimación en el primer periodo es ligeramente mayor que el segundo, en el primero van de 2.4 hasta 4.00, esto dentro de las fronteras del acuífero. En el segundo periodo la varianza es de 1.40 a 3.00 manteniéndose en promedio de 2.00 en toda el área analizada.

4.2.3 Validación cruzada.

La aplicación del semivariograma es la estimación espacial de una variable regionalizada $Z(x)$, por tal motivo se plantea un estimador del semivariograma que se basa en la idea de minimizar los errores de estimación de $Z(x)$.

La validación cruzada es un método para la evaluación de los errores de la estimación (Efron y Gong, 1983).

Sea $Z(x)$ una función aleatoria estacionaria de segundo orden con semivariograma $\gamma(h)$ y función de covarianza $C(h)$, así mismo, sea Z_1, Z_2, \dots, Z_n los valores de $Z(x)$ en n puntos de medida. La validación cruzada consiste en suprimir el i -ésimo valor medido Z_i y en estimarlo a partir del resto de los datos. El valor estimado Z_i^* se calcula por krigado ordinario mediante la expresión

$$Z_i^* = \sum_{m \in N_i} \lambda_{im} Z_m$$

donde N_i es el conjunto de puntos utilizados para estimar Z_i (N_i no incluye el punto i). Los coeficientes de krigado λ_{im} representan el peso dado al dato Z_m en la estimación de Z_i .

En la validación cruzada se utilizaron los siguientes criterios para los errores de estimación y sus (ver descripción más detallada de los criterios que se pueden utilizar en el anexo I):

- a) La media de los errores debe ser próxima a cero.
- b) Mínimo error cuadrático medio.

Para verificar que la estimación obtenida sea aceptable, se realizó la validación cruzada. En las figuras 4.11 y 4.12 se muestran los histogramas de frecuencias y estadígrafos de los errores de estimación.

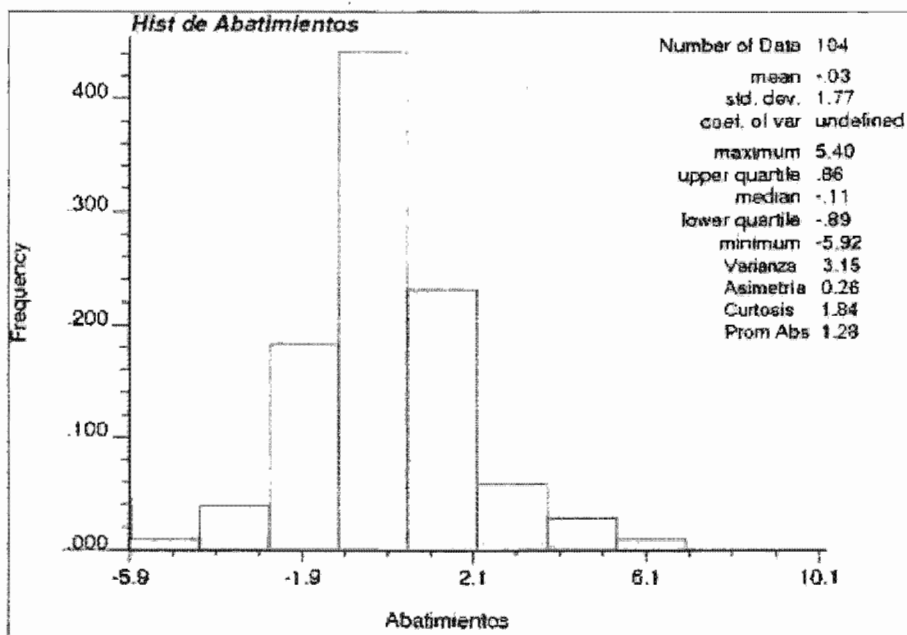


Figura 4.11 Histograma de frecuencias y errores de estimación de abatimientos anuales (1993-1995).

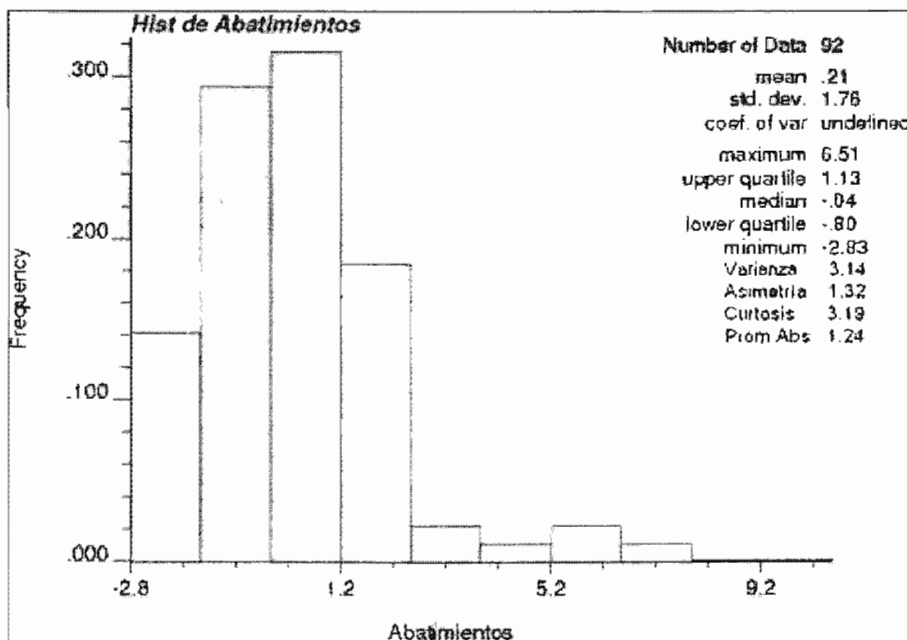


Figura 4.12 Histograma de frecuencias y errores de estimación de abatimientos anuales (1995-1999).

El promedio de los errores de estimación es pequeño para los dos periodos, sin embargo la desviación estándar es de 1.77 que es aproximadamente la mitad de los abatimientos anuales simulados por el modelo numérico de flujo. Sin embargo, como se verá más adelante, el semivariograma “promedio” ajustado a los datos de ambos periodos resultó adecuado para estimar las abatimientos en el periodo en el que se pretende hacer uso de la red de monitoreo. En las figuras 4.13 y 4.14 se tienen las distribuciones espaciales de los errores de estimación para cada periodo, en las cuales se tiene que hay una distribución heterogénea y no se observa que en alguna determinada zona los errores sean menores o mayores para los dos periodos, pero también podemos ver, de acuerdo a los estadígrafos de los errores (figuras 4.11 y 4.12) que en promedio son de 1.26, y aproximadamente 50 pozos presentan errores, para cada periodo, entre cero y la unidad.

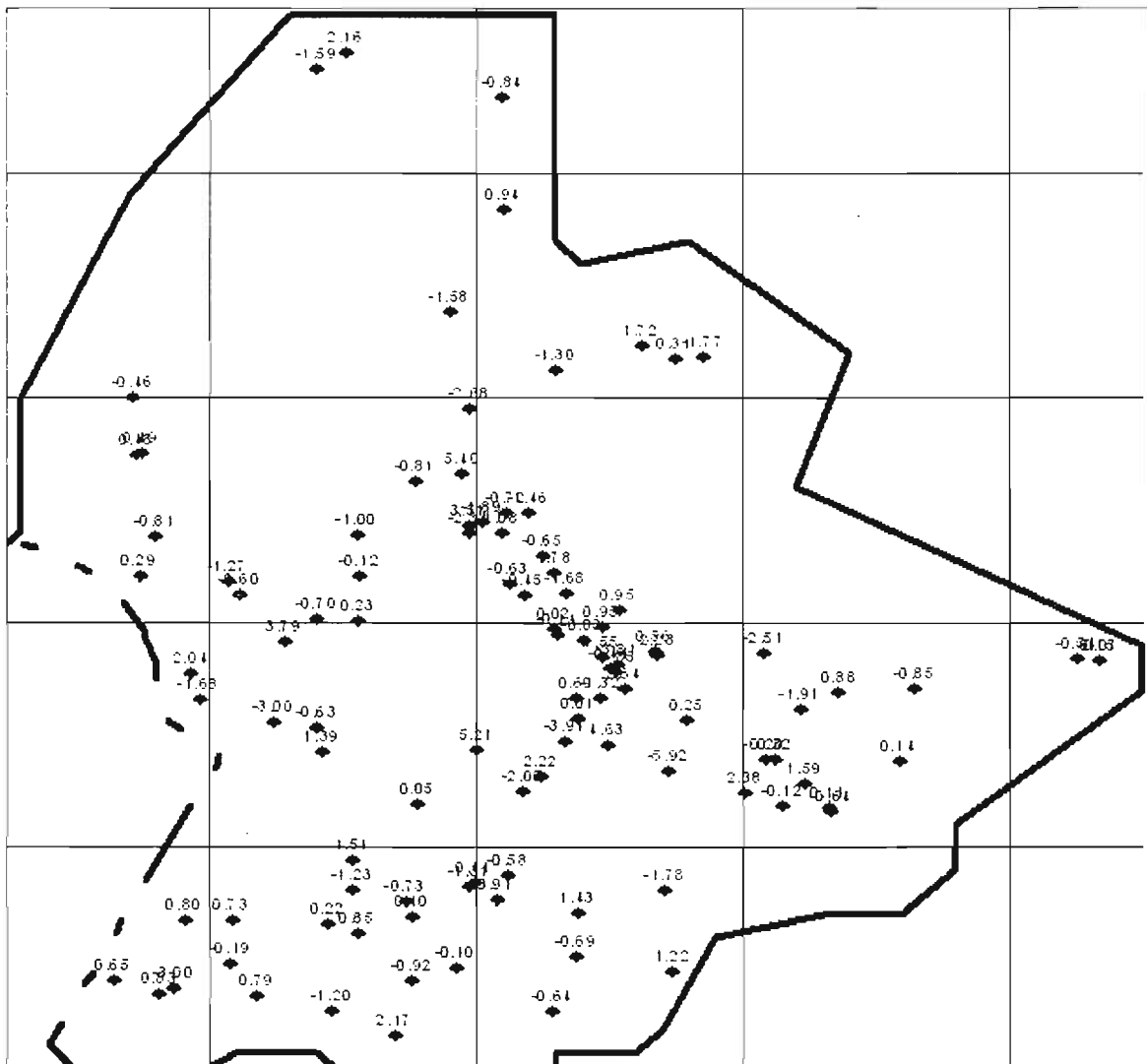


Figura 4.13 Distribución espacial de errores de estimación de abatimientos anuales (1993-1995).

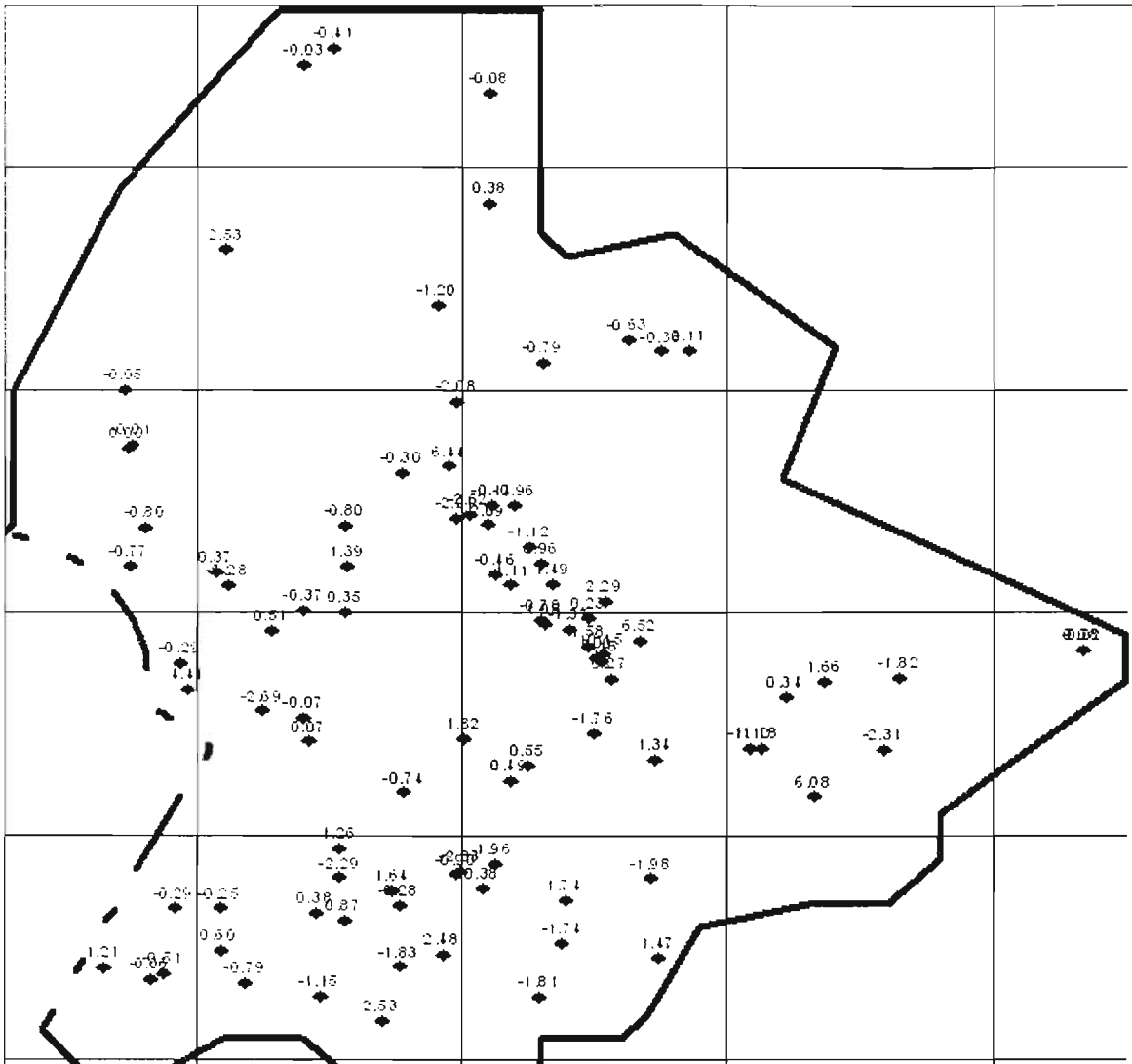


Figura 4.14 Distribución espacial de errores de estimación de abatimientos anuales (1995-1999).

4.2.4 Semivariograma promedio teórico.

En base al análisis estructural obtenido de los abatimientos anuales para los periodos 1993-1995 y 1995-1999 se realizó un semivariograma promedio teórico. El propósito de utilizar estos dos periodos en el análisis es que el semivariograma resultante sea representativo de todo el periodo de tiempo que se está analizando, ya que las características de los datos pueden cambiar con el tiempo. La idea de que el semivariograma resultante se pueda utilizar en el

análisis de los datos que genere la red de monitoreo en el periodo de tiempo en que opere la presa Extóraz.

De los modelos ajustados se obtiene un semivariograma promedio teórico calculando el promedio de las semivarianzas de los dos modelos esféricos, esto es, dado un intervalo de distancia h se calculó las semivarianzas en dicho punto con ayuda de las ecuaciones que definen al modelo esférico ajustado para cada periodo (tabla 4.1 y figura 4.15). A los puntos resultantes se les ajustó un modelo esférico con los siguientes parámetros: nugget de 1.66, varianza de 3.19 y alcance 7949.9.

Lags	Dist	γ		
		93-95	95-99	Prom
0	0	2.10	1.22	1.66
1	1325	2.68	1.72	2.20
2	2650	3.14	2.19	2.67
3	3974.9	3.35	2.59	2.97
4	5299.9	3.35	2.87	3.11
5	6624.9	3.35	3.01	3.18
6	7949.9	3.35	3.02	3.19
7	9274.9	3.35	3.02	3.19
8	10600	3.35	3.02	3.19
9	11925	3.35	3.02	3.19
10	13250	3.35	3.02	3.19

Tabla 4.1 Semivarianza para cada periodo y promedio.

Semivariogramas Teóricos y Semivariograma Promedio con Fórmula

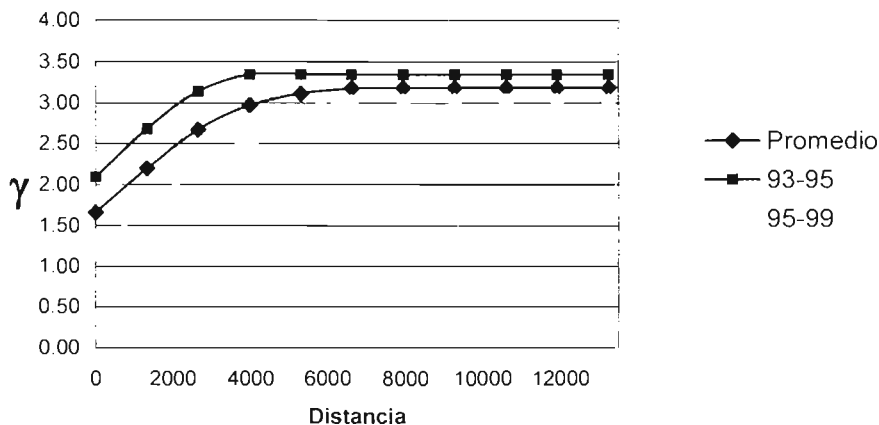


Figura 4.15 Semivariograma promedio teórico de abatimientos anuales (1993-1999).

A manera de comparación se graficaron los semivariogramas muestrales y se incluyó el semivariograma promedio teórico (figura 4.16), para verificar que el semivariograma promedio es representativo de los dos semivariogramas muestrales.

Semivariogramas Muestrales y Semivariograma Promedio con Fórmula

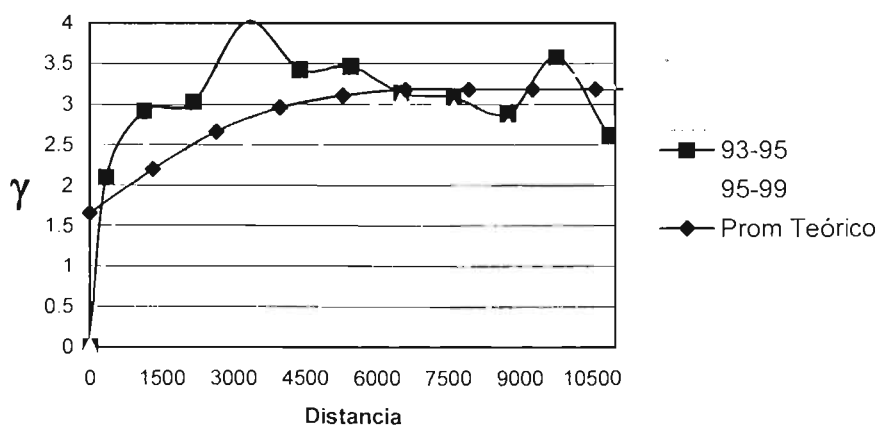


Figura 4.16 Comparación de semivariogramas muestrales y promedio teórico.

De la comparación visual entre estos semivariogramas se concluyó que el semivariograma promedio no caracteriza de buena manera en forma conjunta a los dos semivariogramas muestrales, ya que haciendo un análisis de las diferencias entre el semivariograma promedio con los muestrales se tiene que el error cuadrático medio conjunto de ambos semivariogramas es de 6.00 y el criterio de información de Akaike de -120.82. Por consecuencia de esto se procedió a ajustar otro modelo teórico a los dos semivariogramas muestrales, el resultado es un modelo esférico (figura 4.17) con nugget de 1.66, meseta de 1.53 y alcance de 4200 (menor que el anterior), con esto, el error cuadrático medio resultó de 4.36 y el CIA de -183.77. Este modelo es el que se utilizará para el diseño de la red de monitoreo.

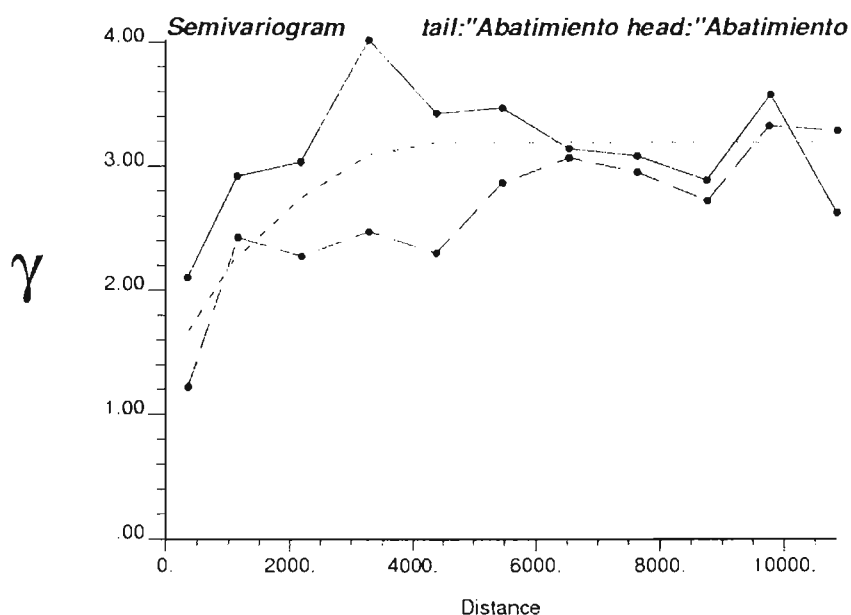


Figura 4.17 Comparación de semivariogramas muestrales y promedio teórico adecuado.

4.3 Matriz de covarianza a priori

4.3.1 Pozos de monitoreo base y puntos de estimación

En el método utilizado se eligen los pozos que conformarán la red de monitoreo óptima entre un grupo de pozos de observación propuestos de antemano a los que, como ya se dijo, llamamos red de monitoreo preliminar. La selección de estos pozos se basa en la cantidad de información que aportan (medida con la reducción en varianza del error de la estimación).

En este caso, la CEAQ propuso un grupo de 33 pozos como posibles pozos para ser incluidos en la red de monitoreo. La lista de estos pozos, incluyendo el número asignado por la CNA para ellos, su nombre de referencia o titular del mismo y sus coordenadas en unidades UTM, se presenta en la tabla 4.2. Por otro lado, la distribución geográfica de los mismos se muestra en la figura 4.18. Como se puede observar en la figura, la distribución geográfica de los pozos propuestos por la CEAQ tiene conglomeraciones en algunas zonas y ausencia de pozos en otras, por tal razón en las zonas con ausencia de pozos se propusieron 7 puntos adicionales para evaluar si es necesario incluir pozos en las zonas donde se ubican. Se le llamó red de monitoreo base al conjunto de pozos propuestos por la CEAQ y los siete puntos adicionales, ésta se muestra también en la figura 4.18. Las coordenadas de los puntos adicionales se pueden consultar en la tabla 4.3.

En el método propuesto por Herrera se eligen un grupo de puntos, a los que llamamos de estimación, en los cuales se estiman los abatimientos anuales utilizando los datos obtenidos de la red de monitoreo. Sobre estos puntos es que se miden las reducciones de la varianza del error de la estimación.

Con la idea de que la red de monitoreo permita evaluar los cambios en los niveles del agua de todo el acuífero, en este caso, estos puntos se distribuyeron sobre toda el área del acuífero del Valle de Querétaro, con una distribución espacial relativamente homogénea. Por simplicidad, estos puntos se eligieron tratando que sus posiciones coincidieran en lo posible con las de los pozos de la red de monitoreo base. En las zonas donde existen conglomeraciones se quitaron 4 pozos que no se utilizaron como puntos de estimación. Esto es, de los 40 pozos que se tienen como red base, 36 son puntos de estimación (figura 4.19).

CENSO_CNA	TITULAR	UTM_X	UTM_Y
0600-A	RANCHO JAN	348956.97	2277407.64
0605	RANCHO SANTA MARIA DEL RETABLO	348898.90	2275968.32
0628	RANCHO VANEGAS	350387.28	2273836.93
0629	EL CERRITO FRACCIONAMIENTO 2	348669.68	2273800.62
0651-A	EJIDO EL CERRITO	347686.71	2274063.18
0652	EJIDO EL CERRITO #4	347678.01	2274721.55
0725-A	(CEA) ALAMEDA # 2	355420.65	2276956.78
0752	ALBERCA LETICIA	358234.78	2278528.43
0759	EJIDO EL NABO	347008.51	2292307.80
0870	FABRICA STA.MARIA DE GUADALUPE	352798.70	2278543.00
0914	(CEA) CAPILLA	350902.69	2276249.77
0955-A	GERBER	353419.00	2279259.85
0976	HACIENDA SAN PEDRITO	354275.73	2285906.88
0982	RANCHO LAS ADJUNTAS	344823.34	2278283.85
0987	INSTITUTO SAN JAVIER	356080.19	2278086.66
1010	UNIROYAL # 2	351497.30	2281097.79
1434	BACHOCO	351114.99	2282977.75
1638	VIDRIERA QUERETARO (NORTE)	351546.77	2279745.58
1935	EJIDO LOS ANGELES	345429.42	2273385.79
1963	RANCHO SAN PEDRITO	353745.17	2285875.04
1977	RANCHO JAN	347109.18	2277110.43
1978	(CEA) LOMA BONITA	349883.83	2284770.73
2041	(CEA) EL NABO #2	345539.10	2288162.60
1313-A	GRANJA EL ROSARIO	346400.99	2279589.92
1428-A	BLACK AND DECKER, S.A.	349750.91	2283293.98
1746-A	(CEA) PUEBLITO 2A	349621.29	2272343.73
0609-A	RANCHO EL RINCON	345355.38	2280897.86
0613-A	EJIDO TLACOTE	343748	2283778
0653-A	EJIDO SANTA VIRGINIA	349969.00	2274204.00
0988-A	(CEA) CENTRO EXPOSITOR #2	356609.73	2275882.74
0999-A	EJIDO EL ROMERAL	343223.08	2272053.10
0608	GRANJA LUCHIS	345571.54	2280599.00
1102	(CEA) CAÑADA #2	361672.79	2279169.53

Tabla 4.2 Pozos que conforman red de monitoreo base.

Punto	X	Y
1	349679.5	2289686
2	346653.8	2284625
3	354527.4	2282421
4	348923.1	2280462
5	353702.2	2276258
6	352773.9	2272462
7	350057.7	2269238

Tabla 4.3. Puntos adicionales.

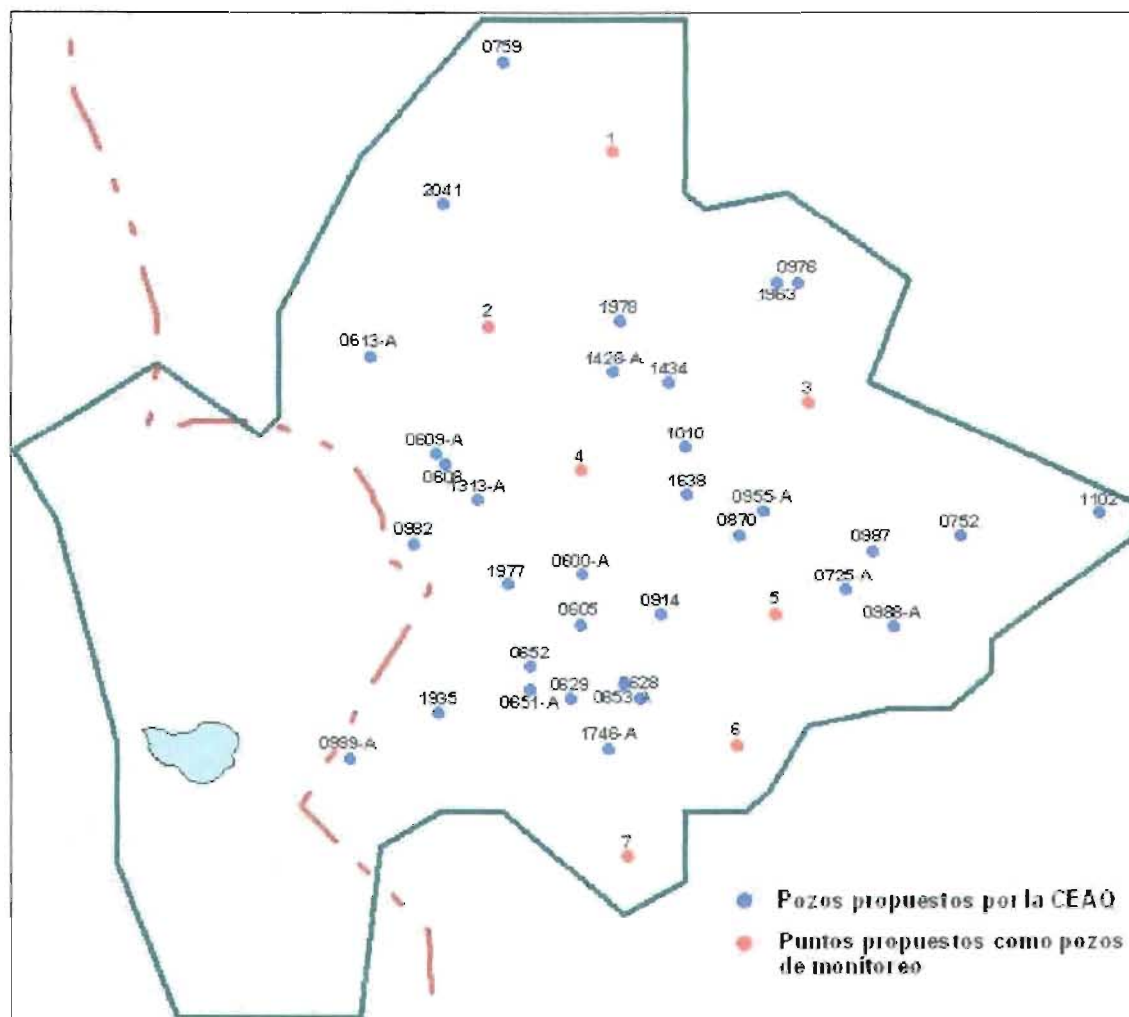


Figura 4.18 Red de monitoreo base.

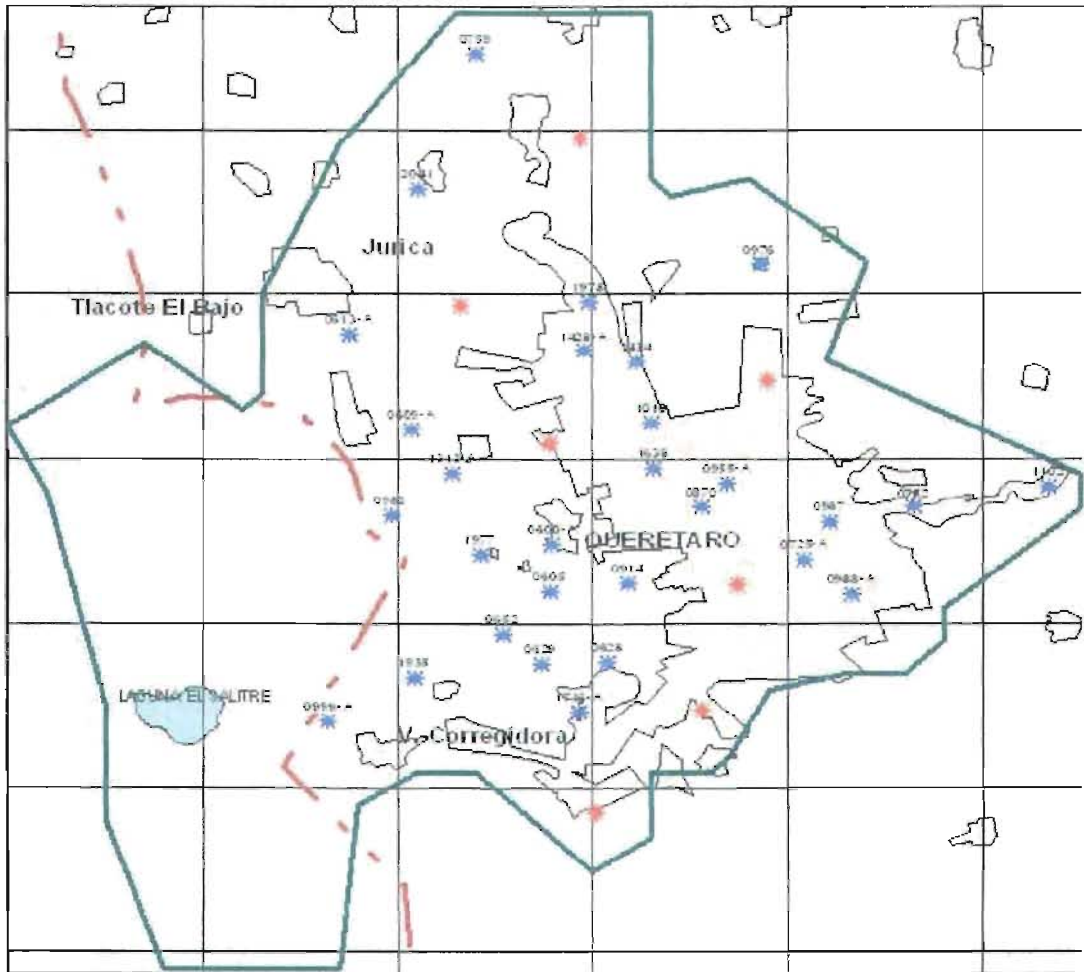


Figura 4.19 Puntos de estimación (los puntos rojos son los adicionales).

4.3.2 Cálculo de la matriz de covarianza a priori

La fórmula que se utilizó para hacer el cálculo de la matriz de covarianza fue:

$$C(h) = C(0) - \gamma(h)$$

donde $C(0)$ es la varianza de las cargas, tomado igual a la meseta del semivariograma más el efecto nugget o pepita, y $\gamma(h)$ es el modelo de semivariograma promedio obtenido. Para hacer este cálculo es necesario que el semivariograma sea acotado, para poder obtener la meseta. Al incluir la ecuación del modelo esférico se tiene la siguiente expresión:

$$C(h) = \frac{C(0)}{2} \left[2 - 3 \left(\frac{h}{a} \right) + \left(\frac{h}{a} \right)^3 \right] \quad \text{para} \quad 0 \leq h \leq a,$$

$$C(h) = 0 \quad \text{para} \quad h > a,$$

donde a es el alcance del modelo de semivariograma y h la distancia entre dos posiciones en las que se quiere calcular la covarianza. La matriz de covarianza se calcula para todos los pozos de la red de monitoreo base, por lo que se utilizó una matriz de tamaño 40 x 40 y las distancias h son las distancias entre todos los pozos. El tamaño de la matriz es debido al número de pozos propuestos por la CEAQ más los 7 puntos adicionales propuestos.

4.4 Propuesta de red de monitoreo

El filtro de Kalman puede calcular la varianza del error de estimación con base en la posición de los sitios de muestreo, sin necesidad de conocer el dato medido en campo. De esta manera este filtro se puede utilizar para determinar, dados n sitios de muestreo, cómo afecta añadir un sitio de muestreo adicional a la varianza del error de la estimación resultante.

Como se mencionó antes, el filtro de Kalman utiliza como dato de entrada la matriz de covarianza del error de la estimación a priori o inicial. Los parámetros para el cálculo de la matriz de covarianza son la varianza (3.19 m^2) y el alcance (4200 m) del modelo ajustado a los dos semivariogramas muestrales obtenidos en el análisis geoestadístico presentado anteriormente. Se definieron los puntos que son de estimación-muestreo y muestreo, los 33 pozos propuestos por la CEAQ son tomados en cuenta para ser pozos de monitoreo, pero además como ya se había mencionado, los 7 puntos de estimación adicionales son propuestos para ser pozos de monitoreo. los únicos puntos que no son considerados como de estimación son los 4 pozos que causan conglomeraciones. Esto es, se tienen 36 pozos de estimación-muestreo y 40 de muestreo.

Después de correr el programa de cómputo que aplica el método propuesto, se obtuvieron las varianzas totales que se muestran en la figuras 4.20 al añadir uno a uno los pozos de la red de monitoreo base.

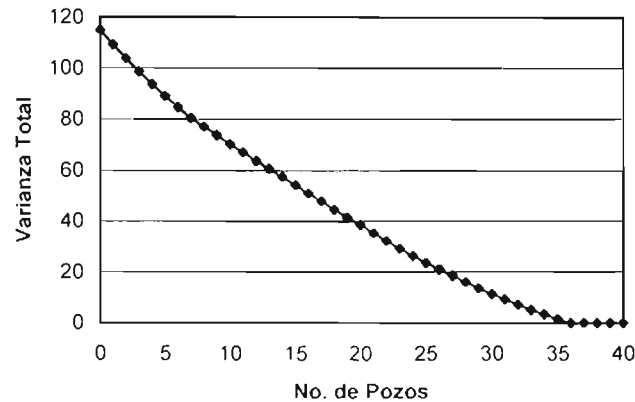


Figura 4.20 Reducción de la varianza total vs. número de pozos en la red de monitoreo.

Como se observa, los primeros 36 pozos elegidos proporcionan mucha información y la varianza total se reduce a cero, y conforme se van añadiendo pozos a la red de monitoreo, la información obtenida por cada pozo va disminuyendo. También se puede observar que los últimos cuatro pozos que se agregan ya no reducen la varianza total.

La figura 4.21 muestra la varianza máxima del error la estimación, esto es, al obtener la estimación con el número de pozos que se indican en el eje horizontal cuál fue la varianza del error de la estimación mayor para todos los puntos de estimación. La varianza máxima nos proporciona una medida de la máxima incertidumbre o máxima falta de información en todos los puntos de estimación. Como se puede observar la varianza máxima se mantiene constante hasta que se han agregado 17 pozos y posteriormente comienza a disminuir hasta llegar a cero.

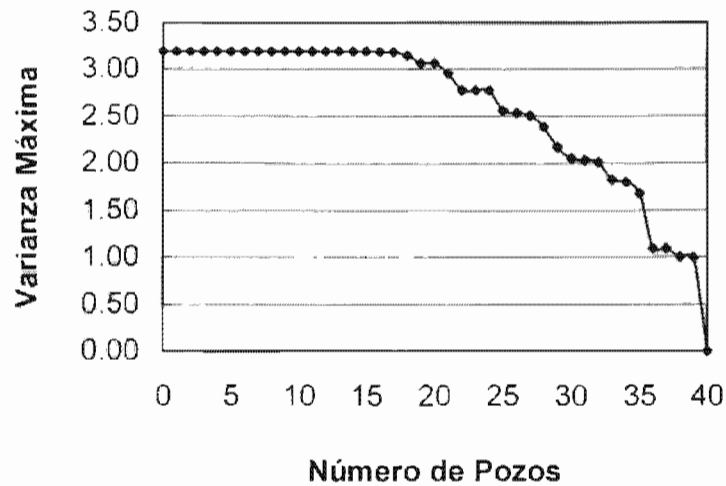


Figura 4.21 Varianza del error de la estimación máxima.

En la figura 4.22 se puede ver el orden en que se seleccionaron los pozos para ser incluidos en la red de monitoreo. El orden en que se seleccionan indica la importancia que tiene cada pozo en la reducción de la varianza total, los que tienen un número menor en la figura son los que aportan más información.

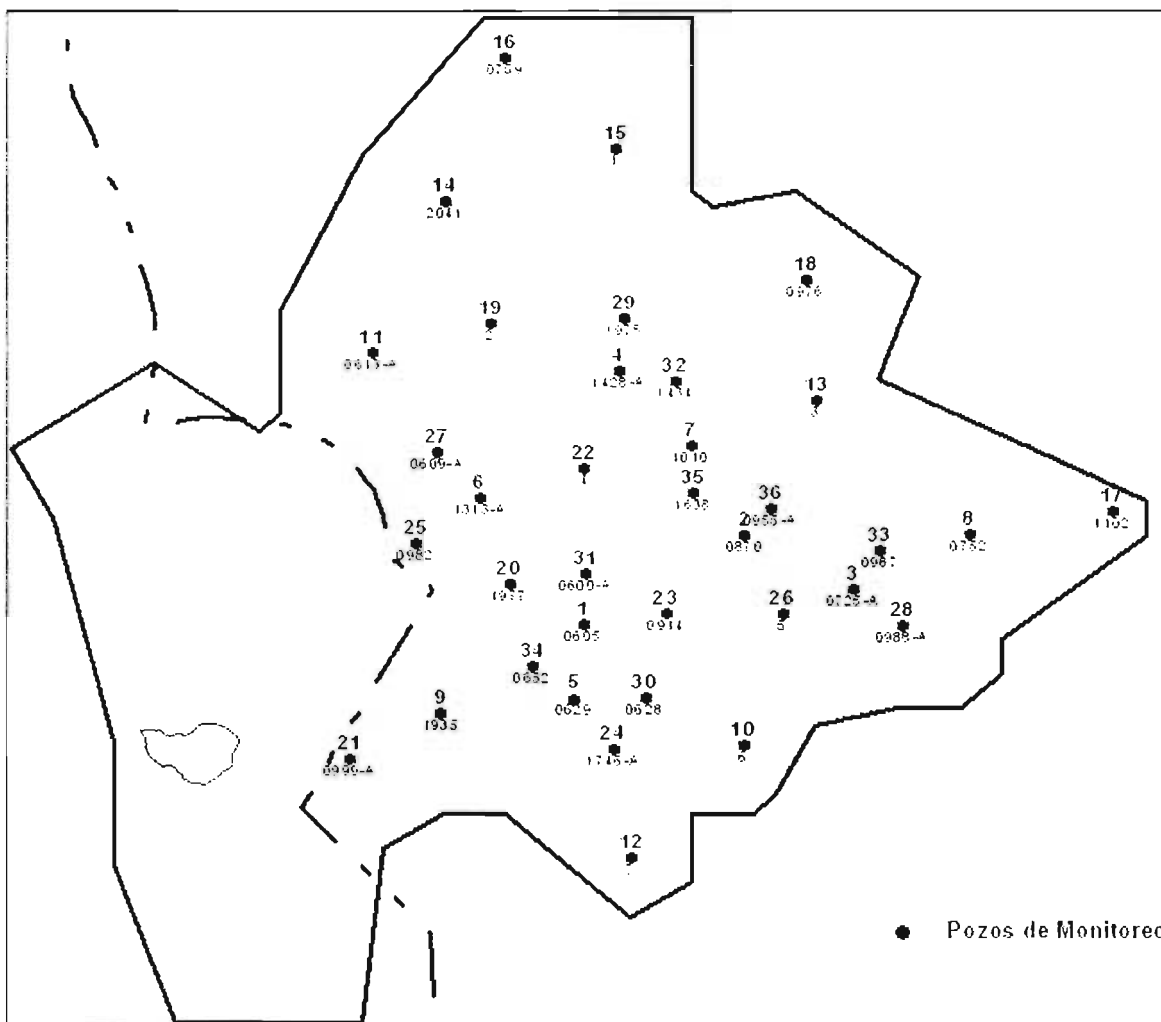


Figura 4.22 Orden de selección de los pozos de monitoreo.

En conclusión, los primeros 36 puntos que se eligieron son los que no proporcionan información redundante, la varianza total disminuye hasta cero al igual que la varianza máxima, de los 33 pozos propuestos por la CEAQ, 29 fueron elegidos al igual que los 7 puntos adicionales propuestos aquí, esto puede ser indicio de que es necesario incluir en la red de monitoreo óptima los 29 pozos propuestos por la CEAQ pero además en cada una de las zonas cercanas a los 7 puntos adicionales se debe de incluir un pozo en la red de monitoreo, es decir, localizar algún pozo cercano a estos puntos o, si fuera necesario construir un pozo. Estos pozos se listan en la tabla 4.4.

Orden de Selección	No. Pozo	TITULAR	X	Y
1	0605	RANCHO SANTA MARIA DEL RETABLO	348899	2275968
2	0870	FABRICA STA.MARIA DE GUADALUPE	352799	2278543
3	0725-A	(CEA) ALAMEDA # 2	355421	2276957
4	1428-A	BLACK AND DECKER, S.A.	349751	2283294
5	0629	EL CERRITO FRACCIONAMIENTO 2	348670	2273801
6	1313-A	GRANJA EL ROSARIO	346401	2279590
7	1010	UNIROYAL # 2	351497	2281098
8	0752	ALBERCA LETICIA	358235	2278528
9	1935	EJIDO LOS ANGELES	345429	2273386
10	6	0	352774	2272462
11	0613-A	EJIDO TLACOTE	343748	2283778
12	7	0	350058	2269238
13	3	0	354527	2282421
14	2041	(CEA) EL NABO #2	345539	2288163
15	1	0	349679	2289686
16	0759	EJIDO EL NABO	347009	2292308
17	1102	(CEA) CAÑADA #2	361673	2279170
18	0976	HACIENDA SAN PEDRITO	354276	2285907
19	2	0	346654	2284625
20	1977	RANCHO JAN	347109	2277110
21	0999-A	EJIDO EL ROMERAL	343223	2272053
22	4	0	348923	2280462
23	0914	(CEA) CAPILLA	350903	2276250
24	1746-A	(CEA) PUEBLITO 2A	349621	2272344
25	0982	RANCHO LAS ADJUNTAS	344823	2278284
26	5	0	353702	2276258
27	0609-A	RANCHO EL RINCON	345355	2280898
28	0988-A	(CEA) CENTRO EXPOSITOR #2	356610	2275883
29	1978	(CEA) LOMA BONITA	349884	2284771
30	0628	RANCHO VANEGAS	350387	2273837
31	0600-A	RANCHO JAN	348957	2277408
32	1434	BACHOCO	351115	2282978
33	0987	INSTITUTO SAN JAVIER	356080	2278087
34	0652	EJIDO EL CERRITO #4	347678	2274722
35	1638	VIDRIERA QUERETARO (NORTE)	351547	2279746
36	0955-A	GERBER	353419	2279260

Tabla 4.4 Red de monitoreo preliminar.

4.5 Estimación y validación cruzada

Se realizó una validación cruzada y estimación de abatimientos por medio de kriging ordinario con la finalidad de estimar los posibles errores que se tengan al comparar los valores estimados, a través del modelo ajustado a los dos semivariogramas muestrales, en los puntos propuestos a formar la red de monitoreo (tabla 4.4) con los valores simulados por el modelo numérico de flujo en los periodos de predicción cuando entra en operación la presa Extóraz, esto en los años 2010, 2020 y 2030.

El criterio que se sigue es realizar una validación cruzada y estimación de abatimientos utilizando los 29 pozos propuestos por la CEAQ elegidos por filtro de Kalman y también con los 36 pozos propuestos para formar la red de monitoreo, esto dará una idea de cómo serán los errores al incluir o eliminar los 7 pozos adicionales. Las estimaciones realizadas con los 29 pozos se comparan con cada uno de los valores simulados por el modelo de flujo en los 7 pozos adicionales y se determina cuáles de estos puntos (pozos) adicionales se agregan a los 29 pozos antes elegidos para formar la red de monitoreo óptima.

Del modelo numérico del acuífero de Querétaro se extrajeron los valores simulados para el año 2010, 2020 y 2030, pero cabe aclarar que estos valores fueron obtenidos en base al escenario de predicción explicado en la sección 3.3. Los valores extraídos del modelo fueron aquellos que se encontraron en las mismas localizaciones de los pozos elegidos por filtro de Kalman. De los valores simulados para cada año se obtuvieron los abatimientos anuales por periodo (2010-2020 y 2020-2030), es decir, de la piezometría presentada para los tres años se determinó el abatimiento anual por medio de las diferencias y divididas en el número de años entre periodos. Esto se hace porque el análisis geoestadístico y el diseño de la red de monitoreo se realizan en base a los abatimientos anuales de 1993 a 1999. Estos datos se muestran en la tabla II.2 (Anexo II).

Los estadígrafos de los errores de estimación de la validación cruzada con 29 pozos se muestran en las figuras 4.23 y 4.25, su distribución espacial de los errores en las figuras 4.27 y 4.29. Los estadígrafos de los errores de estimación con 36 datos en las figuras 4.24 y 4.26, su distribución espacial en las figuras 4.28 y 4.30.

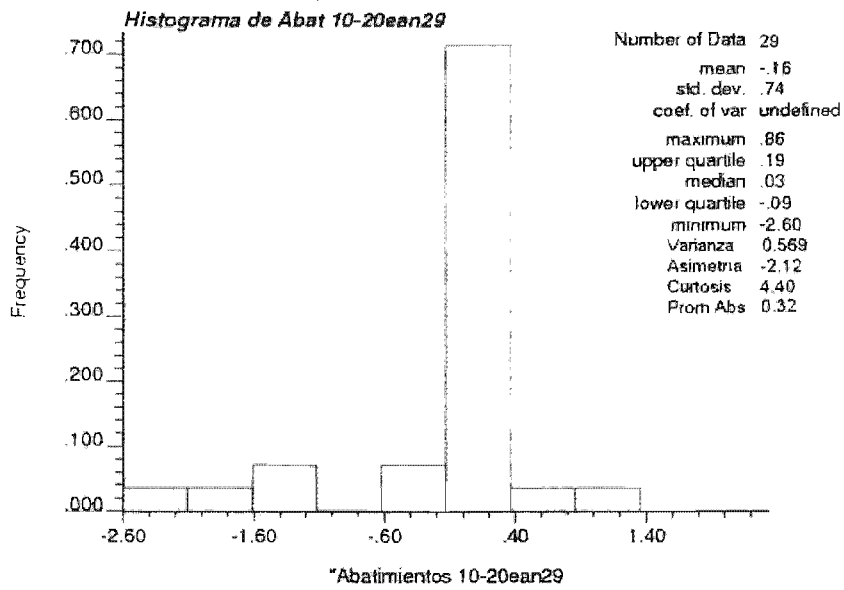


Figura 4.23 Errores (estimación vs simulado) 2010-2020, 29 pozos.

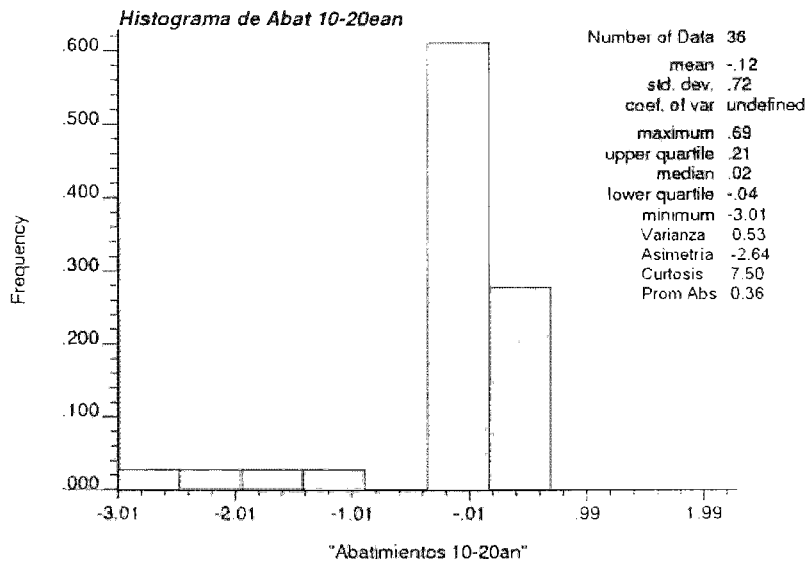


Figura 4.24 Errores (estimación vs simulado) 2010-2020, 36 pozos.

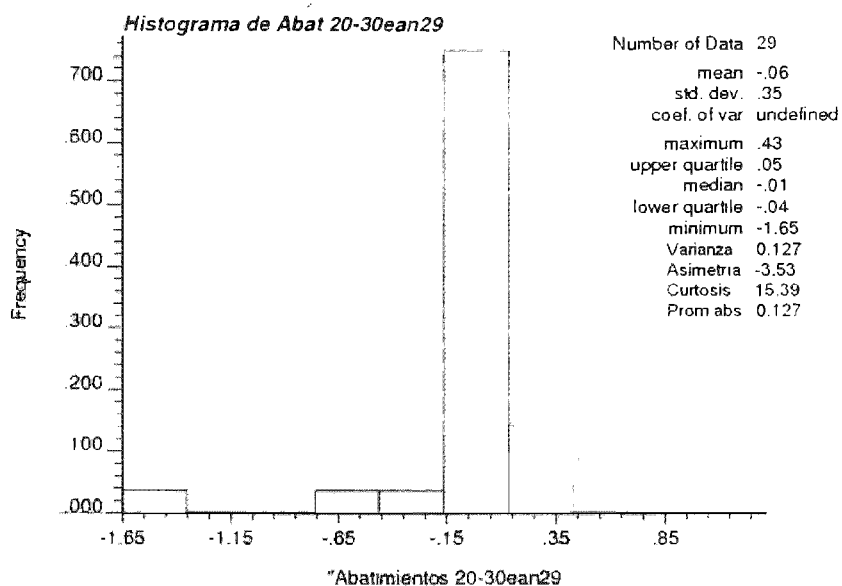


Figura 4.25 Errores (estimación vs simulado) 2020-2030, 29 pozos.

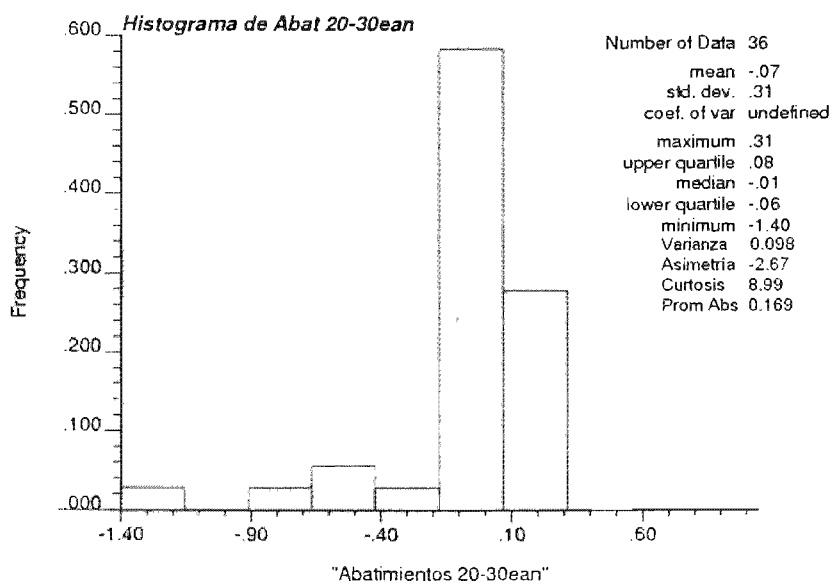


Figura 4.26 Errores (estimación vs simulado) 2020-2030, 36 pozos.

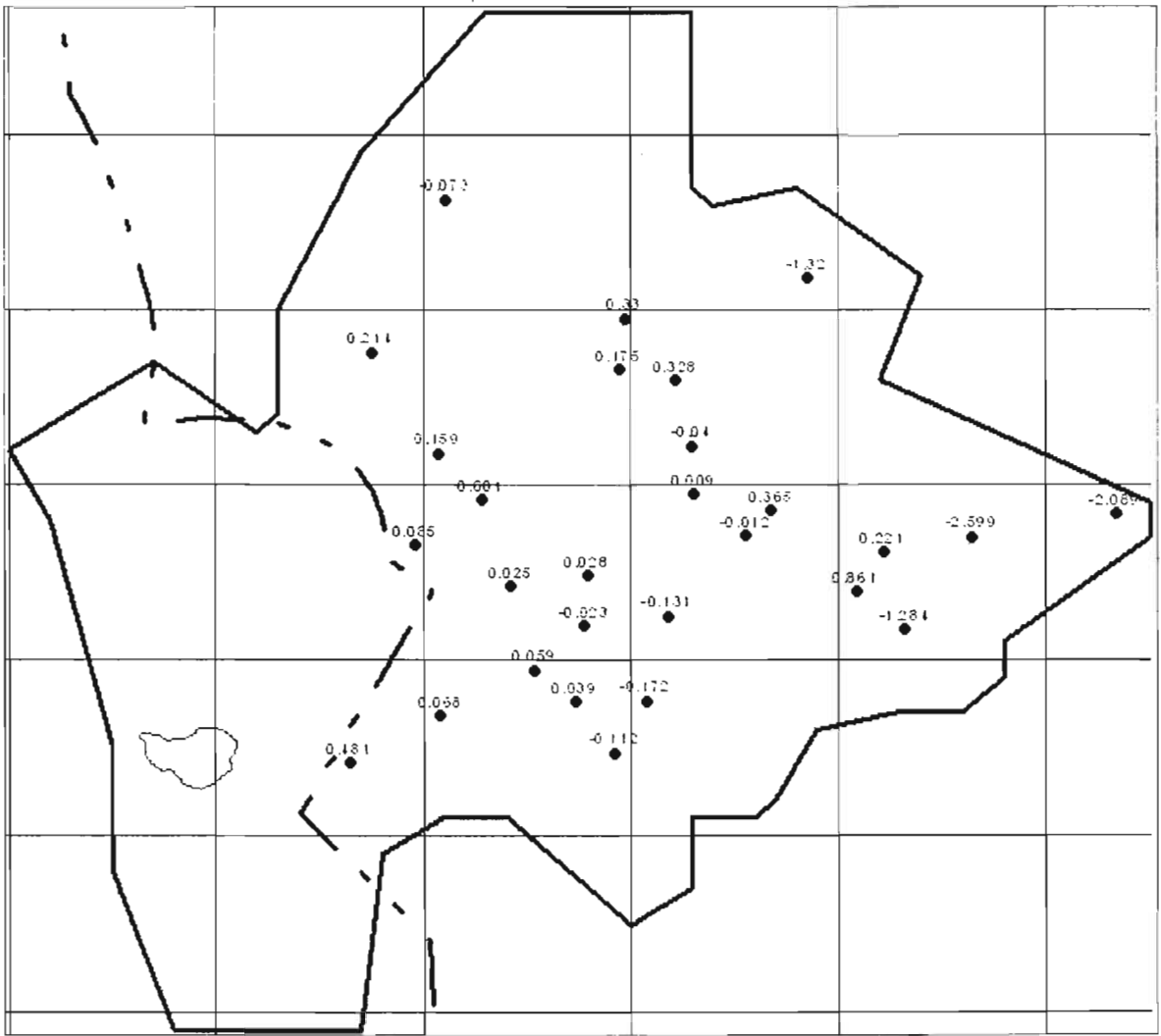


Figura 4.27 Distribución espacial de errores (estimación vs simulado) 2010-2020, 29 pozos.

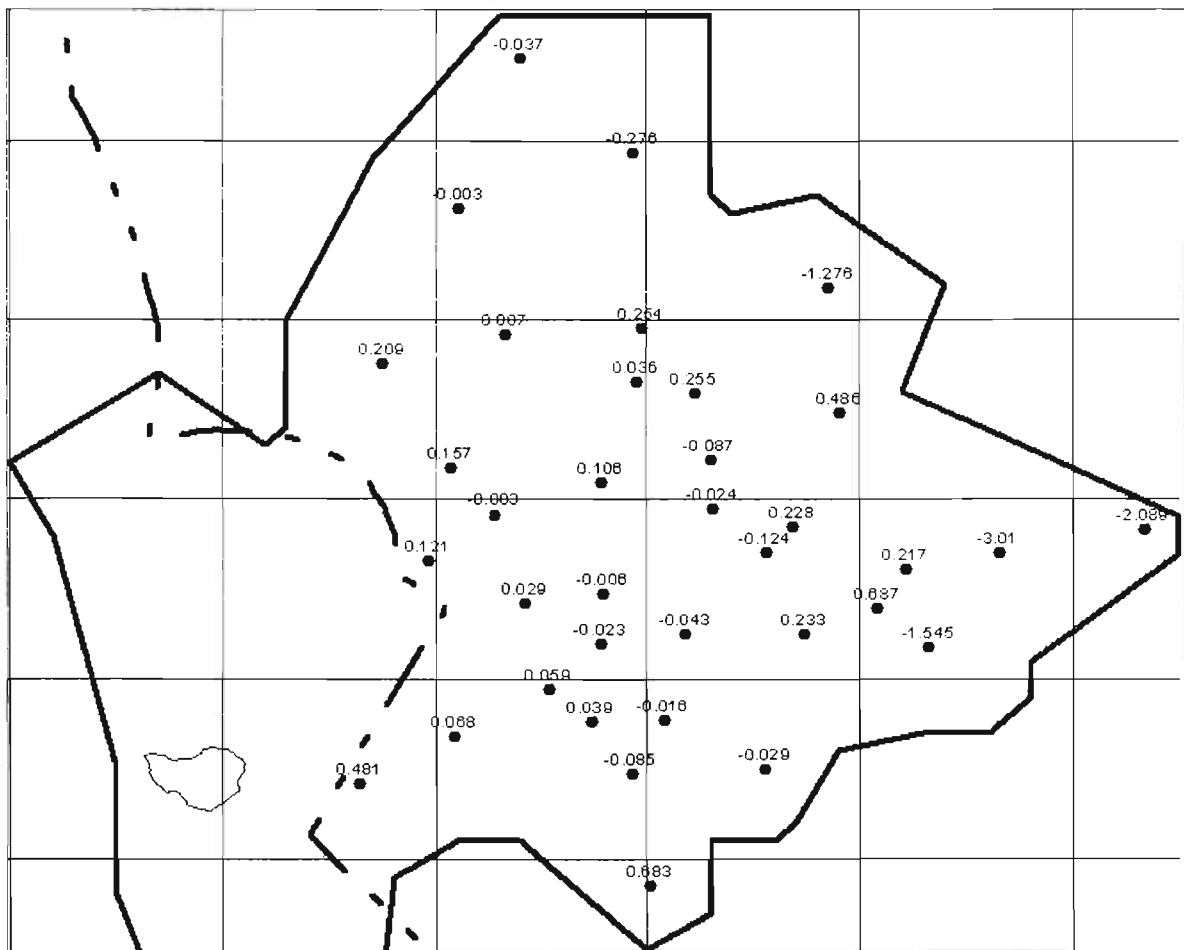


Figura 4.28 Distribución espacial de errores (estimación vs simulado) 2010-2020, 36 pozos.

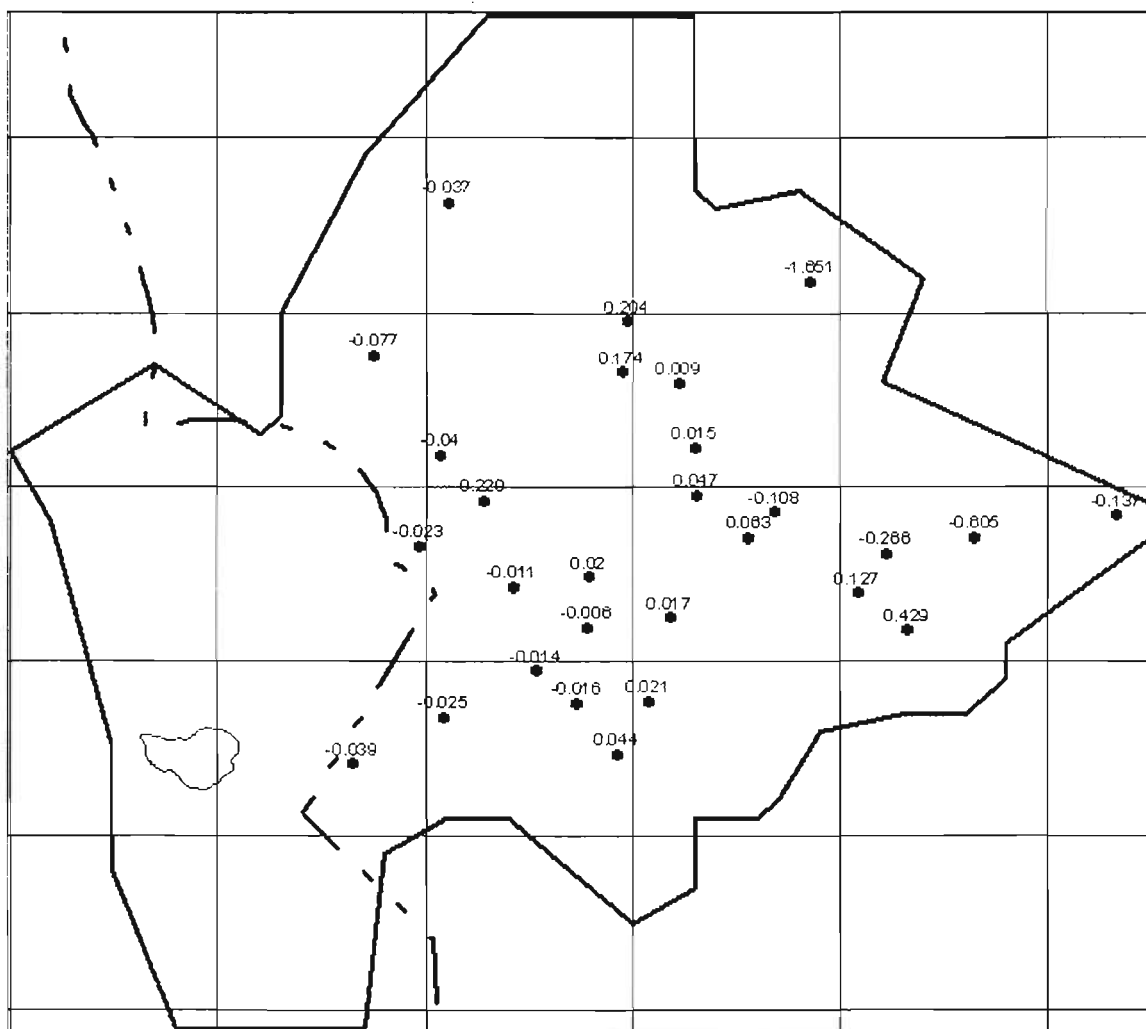


Figura 4.29 Distribución espacial de errores (estimación vs simulado) 2020-2030, 29 pozos.

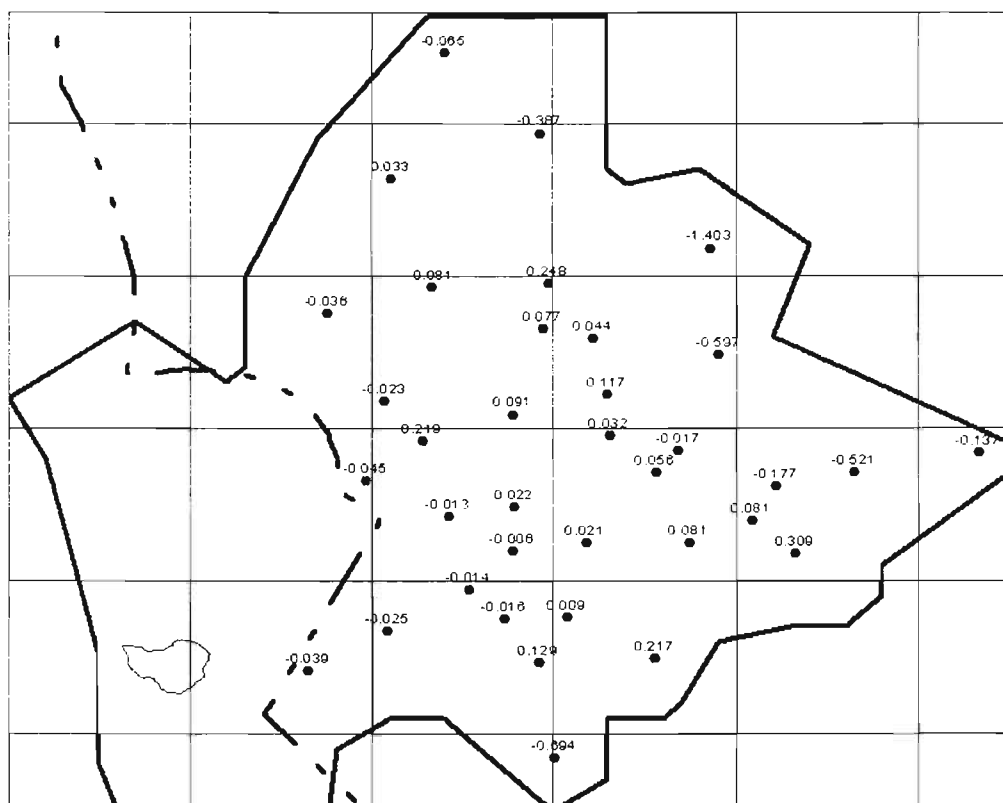


Figura 4.30 Distribución espacial de errores (estimación vs simulado) 2020-2030, 36 pozos.

En la validación cruzada tanto para 29 pozos como para 36 los resultados muestran que estos estadígrafos son bastante aceptables, los errores son muy pequeños, ya que la media y mediana se encuentran cerca de cero, la desviación estándar es pequeña y la curtosis al ser del tipo leptocúrtica indica que la mayoría de los errores son aproximadamente cero. Cabe hacer mención que los errores del periodo 2020-2030 son menores que los del periodo 2010-2020. Al hacer esta validación cruzada para 29 y 36 pozos se ve que los errores en general son similares, no hay una diferencia significativa, por tal motivo no se obtiene una conclusión de cuántos pozos son necesarios para formar la red de monitoreo óptima.

Como ya se había mencionado, también se realizó la estimación de abatimientos para los dos periodos (2010-2020 y 2020-2030) en toda el área que ocupan los pozos propuestos a formar la red de monitoreo, esto es, para 29 datos y para 36. Las gráficas de estimaciones de abatimientos para 29 datos se muestran en la figura 4.31 y para 36 en la 4.32 para los periodos 2010-2020 y en las figuras 4.33 y 4.34 para el periodo 2020-2030.

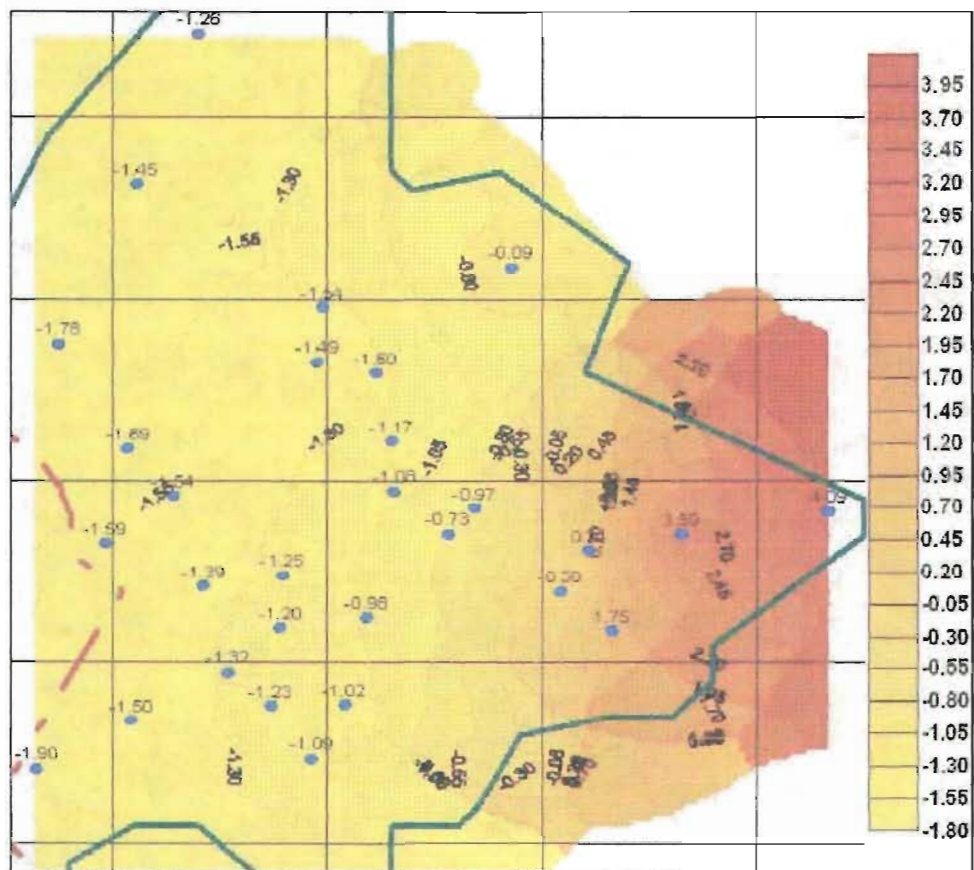


Figura 4.31 Estimación de abatimientos con 29 pozos (2010-2020).

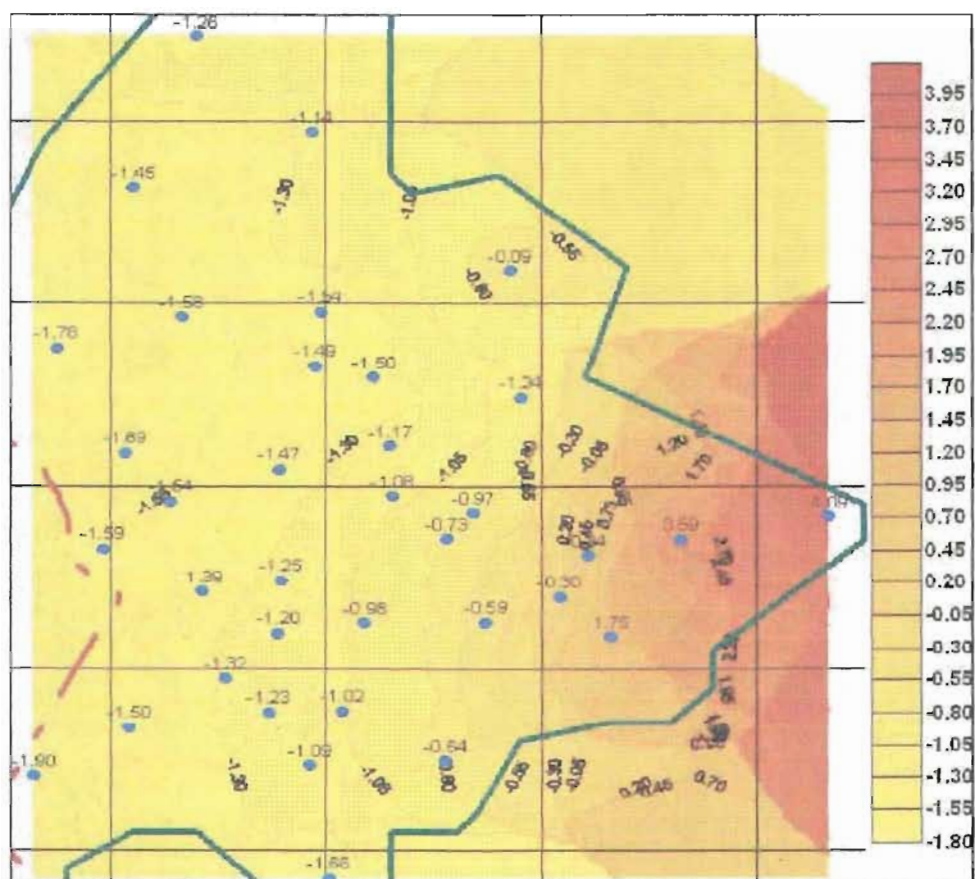


Figura 4.32 Estimación de abatimientos con 36 pozos (2010-2020).

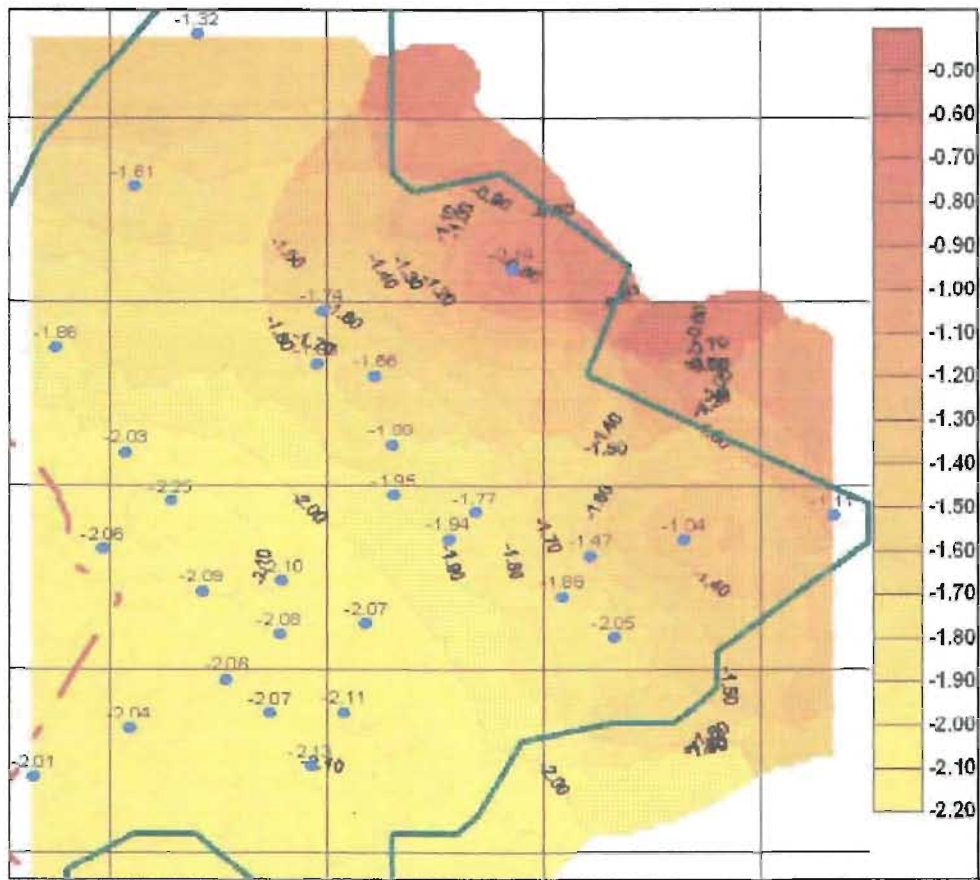


Figura 4.33 Estimación de abatimientos con 29 pozos (2020-2030).

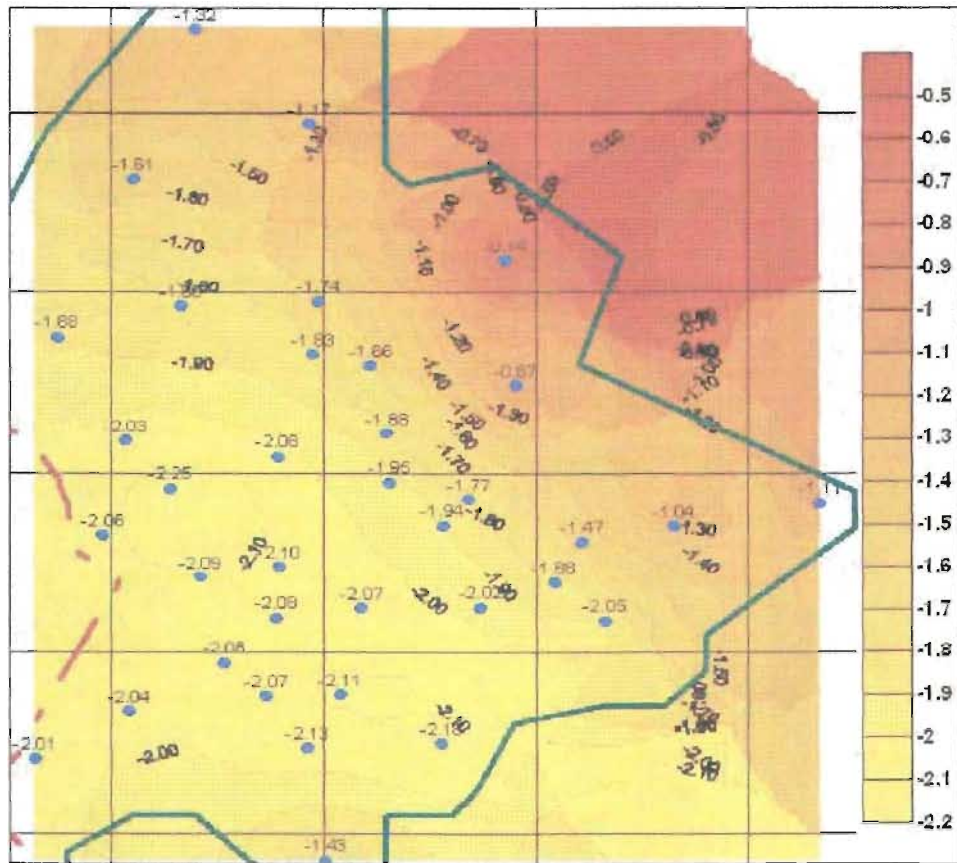


Figura 4.34 Estimación de abatimientos con 36 pozos (2020-2030).

Se observa que en forma conjunta las estimaciones de abatimientos anuales para cada periodo son semejantes, presentándose abatimientos máximos de 1.8 m en el periodo 2010-2020 en el oeste del acuífero y abatimientos menores hacia el centro. También se ve que en este periodo existe una recuperación de 4.00 m en la carga hidráulica en el este del acuífero, hacia la zona de la Cañada. Para el 2020-2030 los abatimientos máximos anuales son de 2.20 m en la zona centro del acuífero y menores cercanos a cero hacia la zona noreste, en este caso no se observa recuperación en los niveles de agua, esto debido a que en este periodo se vuelve a iniciar la extracción por bombeo al acuífero.

Para determinar la cantidad de pozos a formar la red de monitoreo óptima se comparan los valores de abatimientos anuales para cada periodo simulados por el modelo de flujo en cada uno de los 7 pozos propuestos adicionales con los valores estimados por el kriging en la misma localización, esto para obtener sus respectivos errores, estas diferencias se presentan en la tabla 4.5.

Pozo adic.	X	Y	Sim 10-20	Sim 20-30	Est 10-20	Est 20-30	Dif Est-Sim 2010-2020	Dif Est-Sim 2020-2030
1	349679	2289686	-1.140	-1.171	-1.416	-1.557	0.276	0.386
2	346654	2284625	-1.578	-1.797	-1.581	-1.852	0.003	0.055
3	354527	2282421	-1.243	-0.874	-0.754	-1.467	-0.489	0.593
4	348923	2280462	-1.474	-2.084	-1.371	-1.995	-0.103	-0.089
5	353702	2276258	-0.589	-2.018	-0.388	-1.923	-0.201	-0.095
6	352774	2272462	-0.643	-2.181	-0.605	-2.057	-0.038	-0.124
7	350058	2269238	-1.659	-1.433	-1.113	-2.104	-0.546	0.671

Tabla 4.5 Diferencias para cada periodo (valores simulados vs valores de estimación).

Los resultados muestran que hay mayor diferencia en los pozos 1, 3 y 7 para los dos periodos, con esto se puede concluir que a los 29 pozos elegidos por filtro de Kalman propuestos por la CEAQ se recomienda agregar los pozos 1, 3 y 7 propuestos aquí, la red de monitoreo óptima se compone de 32 pozos (tabla 4.6 y figura 4.35).

Orden de Selección	No. Pozo	TITULAR	X	Y
1	0605	RANCHO SANTA MARIA DEL RETABLO	348898.90	2275968.32
2	0870	FABRICA STA.MARIA DE GUADALUPE	352798.70	2278543.00
3	0725-A	(CEA) ALAMEDA # 2	355420.65	2276956.78
4	1428-A	BLACK AND DECKER, S.A.	349750.91	2283293.98
5	0629	EL CERRITO FRACCIONAMIENTO 2	348669.68	2273800.62
6	1313-A	GRANJA EL ROSARIO	346400.99	2279589.92
7	1010	UNIROYAL # 2	351497.30	2281097.79
8	0752	ALBERCA LETICIA	358234.78	2278528.43
9	1935	EJIDO LOS ANGELES	345429.42	2273385.79
10	0613-A	EJIDO TLACOTE	343747.63	2283777.68
11	7	0	350057.67	2269237.77
12	3	0	354527.37	2282420.72
13	2041	(CEA) EL NABO #2	345539.10	2288162.60
14	1	0	349679.47	2289685.63
15	0759	EJIDO EL NABO	347008.51	2292307.80
16	1102	(CEA) CAÑADA #2	361672.79	2279169.53
17	0976	HACIENDA SAN PEDRITO	354275.73	2285906.88
18	1977	RANCHO JAN	347109.18	2277110.43
19	0999-A	EJIDO EL ROMERAL	343223.08	2272053.10
20	0914	(CEA) CAPILLA	350902.69	2276249.77
21	1746-A	(CEA) PUEBLITO 2A	349621.29	2272343.73
22	0982	RANCHO LAS ADJUNTAS	344823.34	2278283.85
23	0609-A	RANCHO EL RINCON	345355.38	2280897.86
24	0988-A	(CEA) CENTRO EXPOSITOR #2	356609.73	2275882.74
25	1978	(CEA) LOMA BONITA	349883.83	2284770.73
26	0628	RANCHO VANEGAS	350387.28	2273836.93
27	0600-A	RANCHO JAN	348956.97	2277407.64
28	1434	BACHOCO	351114.99	2282977.75
29	0987	INSTITUTO SAN JAVIER	356080.19	2278086.66
30	0652	EJIDO EL CERRITO #4	347678.01	2274721.55
31	1638	VIDRIERA QUERETARO (NORTE)	351546.77	2279745.58
32	0955-A	GERBER	353419.00	2279259.85

Tabla 4.6 Lista de pozos de la red de monitoreo óptima.

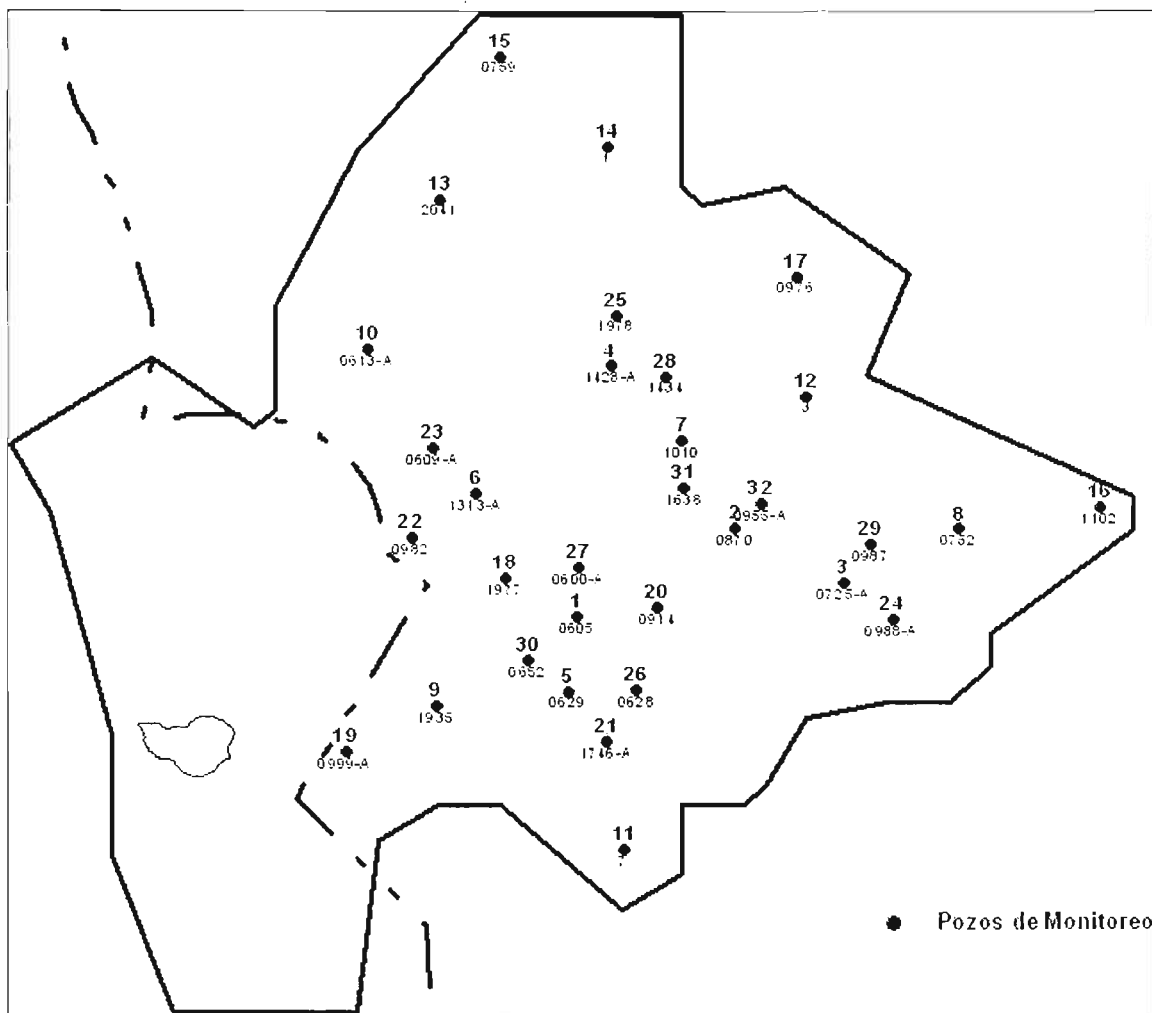


Figura 4.35 Orden de selección de pozos de red de monitoreo óptima.

Para verificar que al utilizar estos pozos como red de monitoreo óptima se obtengan estimaciones adecuadas en los años de predicción se realizó otra validación cruzada con 32 pozos, los estadígrafos se muestran en la figura 4.36 y 4.37 para el periodo 2010-2020 y 2020-2030 respectivamente. la distribución espacial de los errores en las figuras 4.38 y 4.39, y errores porcentuales en las figuras 4.40 y 4.41 para cada uno de los periodos. Los errores porcentuales se muestran con el fin de conocer la dimensión en forma porcentual de los mismos. Estos se obtuvieron con la relación de cada error en valor absoluto con los valores obtenidos del modelo de flujo en cada pozo.

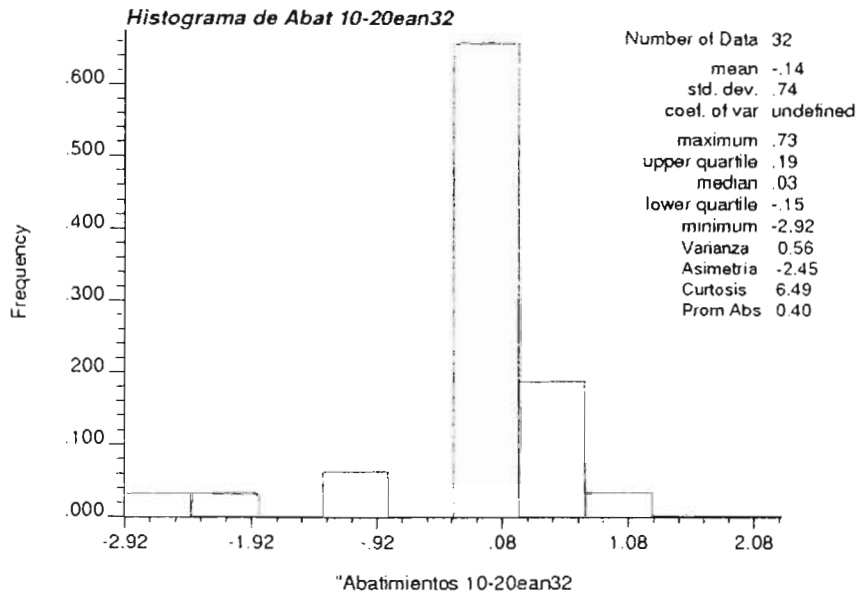


Figura 4.36 Estadígrafos de errores con 32 pozos (2010-2020).

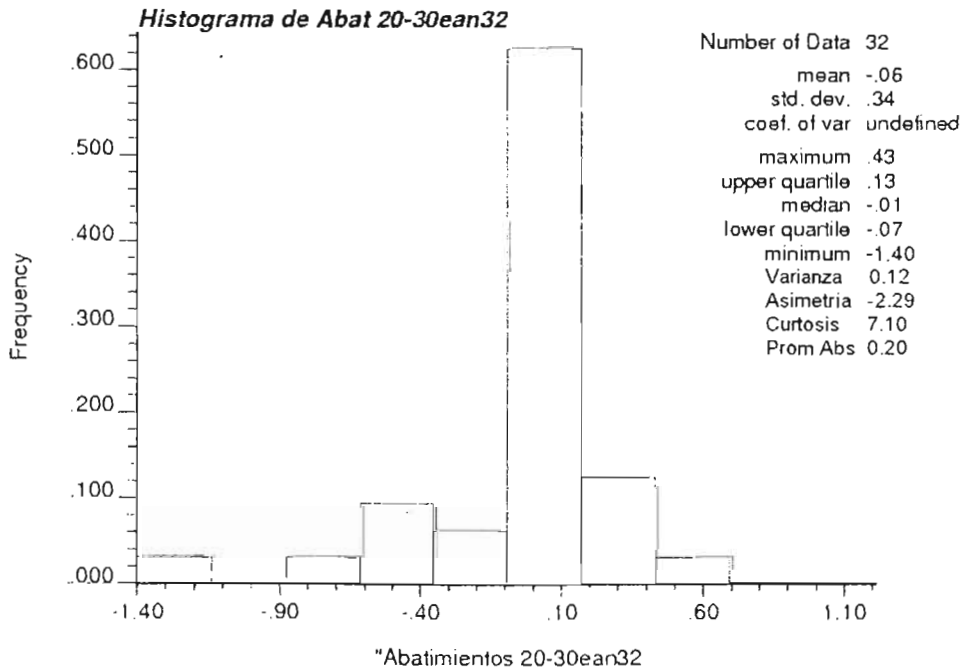


Figura 4.37 Estadígrafos de errores con 32 pozos (2020-2030).

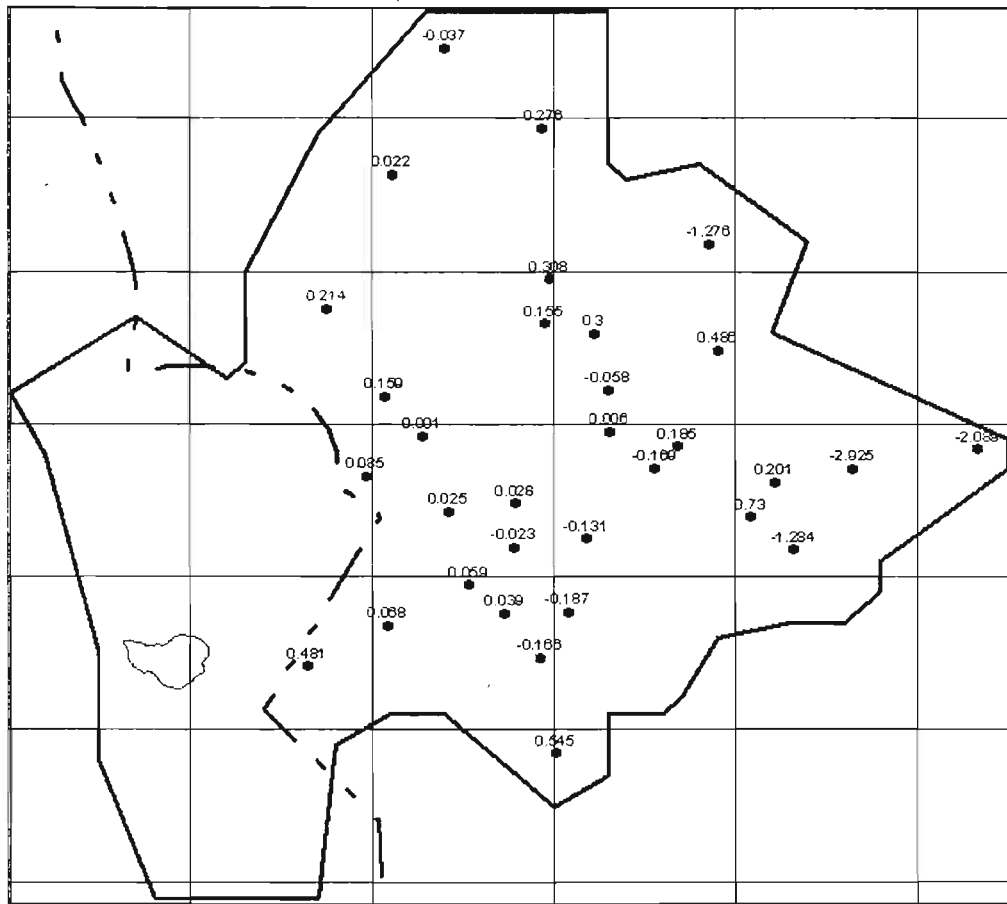


Figura 4.38 Distribución espacial de errores (2010-2020).

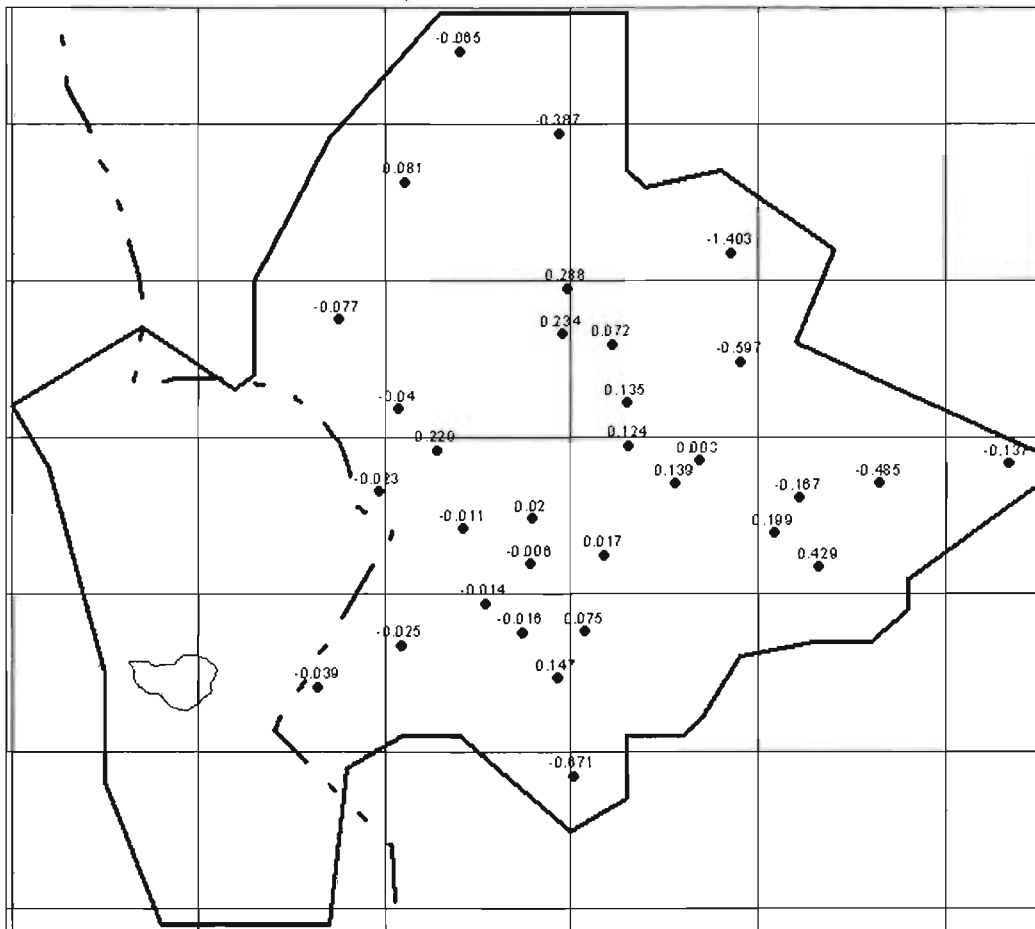


Figura 4.39 Distribución espacial de errores (2020-2030).

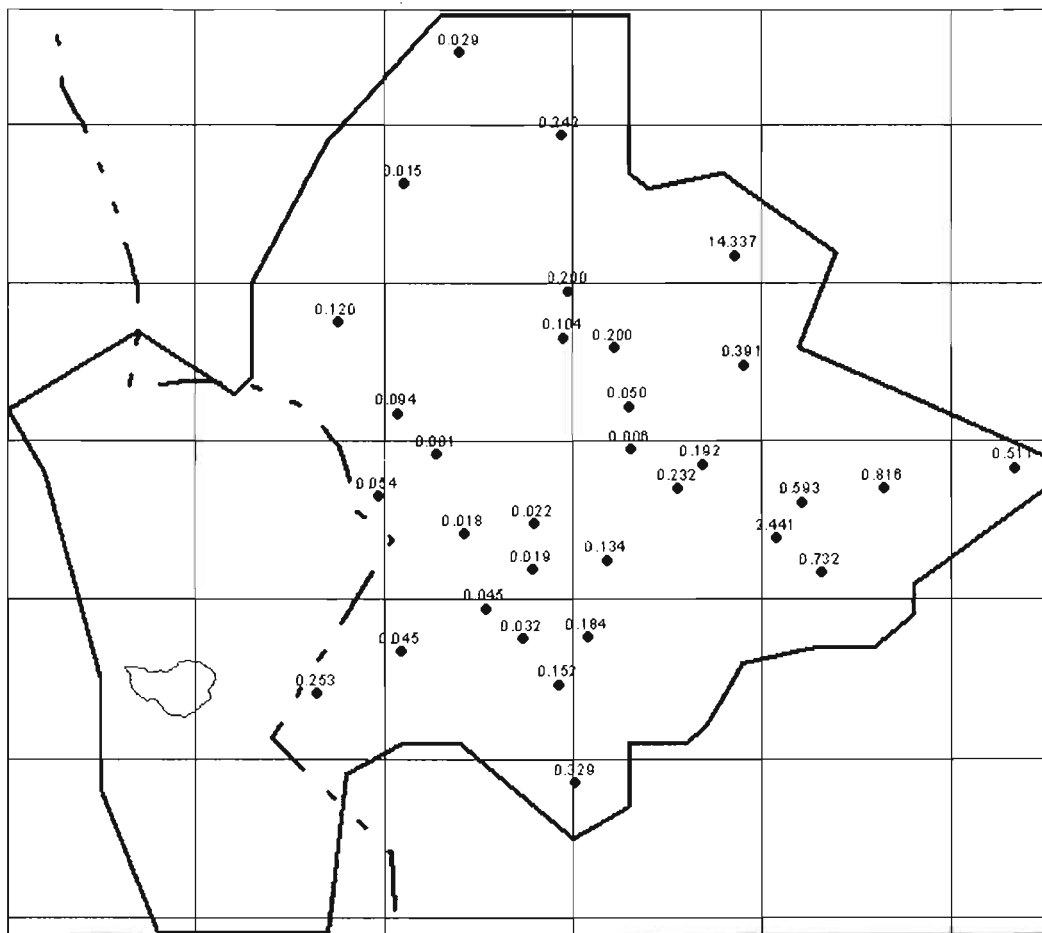


Figura 4.40 Errores porcentuales (2010-2020).

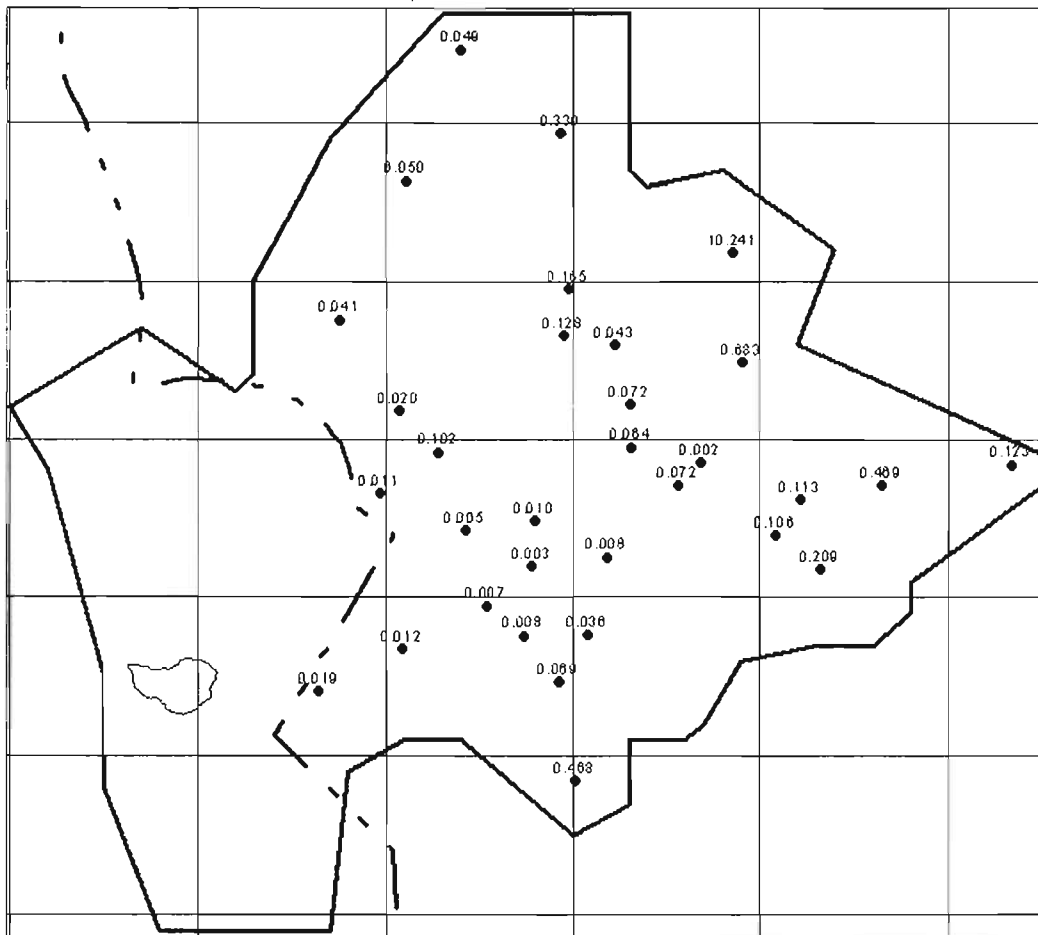


Figura 4.41 Errores porcentuales (2020-2030).

La validación cruzada para los 32 pozos arroja como resultado en los dos periodos que los estadígrafos son muy aceptables, teniendo una media y mediana semejantes, la varianza es pequeña, así como la desviación estándar y la curtosis indica que la distribución es leptocúrtica, lo que significa que la mayoría de los errores se encuentran cerca del cero. También se tiene en cuanto a la distribución espacial que en la zona de mayor abatimiento los errores son mínimos haciéndose más grandes hacia la periferia del acuífero. En cuanto a los errores porcentuales, los valores más pequeños son los localizados al centro del valle. En los dos periodos existe un porcentaje bastante alto en la parte noreste, esto por motivo que en esa zona los abatimientos son casi nulos y al dividir entre estos valores el error porcentual se incrementa significativamente. A pesar de que este error porcentual es grande no se considera que esto sea un problema debido a que, como ya se dijo, en esa zona los abatimientos son muy pequeños. Por otro lado, también se tienen valores altos, en comparación con los del centro del

valle, en la zona este del acuífero, sobre todo para el periodo 2010-2020. Esto se debe a que en esa zona los gradientes obtenidos por el modelo para ese periodo son muy pronunciados (ver figuras 3.21 y 3.22 del capítulo 3) y esto hace que al obtener la validación cruzada quitando alguno de los pozos los errores se incrementen significativamente ya que las diferencias de los niveles del agua en esa zona pueden ser muy importantes. Se consideró la posibilidad de incorporar una densidad de pozos mayor en esa zona para disminuir estos errores pero se vio que en la validación cruzada los errores no disminuían.

Con los resultados finales ya vistos anteriormente la red de monitoreo óptima está conformada por 29 pozos propuestos por la CEAQ y 3 pozos ubicados en las zonas propuestas aquí, haciendo una cantidad de 32 pozos de monitoreo.

4.6 Periodicidad del monitoreo.

Para el objetivo de la red de monitoreo, que es medir los cambios en los niveles del agua subterránea durante la operación de la presa Extóraz, es suficiente con medir los niveles del agua una vez al año. Sin embargo, se recomienda que una vez cada cinco años se realice un monitoreo más denso, de entre 80 y 100 pozos, que permita actualizar el semivariograma que se utiliza para obtener las estimaciones de los abatimientos, ya que la calidad de las estimaciones obtenidas por kriging depende en gran medida de que el semivariograma que se utilice refleje en forma adecuada las correlaciones espaciales de los abatimientos. Hay que recordar que el semivariograma utilizado en el análisis presentado se obtuvo con datos de abatimientos anuales de los periodos 1993-1995 y 1995-1999, y ya que si la presa Extóraz es construida y se deja de extraer agua del acuífero conforme lo tiene planeado la CEAQ, el comportamiento del acuífero cambiará en forma abrupta por lo que un semivariograma obtenido con datos más recientes puede mejorar las estimaciones obtenidas con el kriging.

5. CONCLUSIONES

- Para el diseño de la red de monitoreo fue necesario proponer un semivariograma que fuera adecuado para obtener estimaciones de los abatimientos en el período en el que operaría la presa Extóraz. Con este fin se propusieron dos formas alternativas para obtenerlo con base en datos históricos, la primera fue hacer el cálculo promediando el mejor ajuste de semivariograma teórico para los datos de los periodos 1993-1995 y 1995-1999, y la segunda fue hacer un ajuste conjunto a los semivariogramas maestres. De estos dos métodos, resultó más adecuado el segundo, ya que proporcionó el error cuadrático medio menor y el mejor criterio de información de Akaike. Al hacer la validación cruzada de este modelo “promedio” con los abatimientos resultantes del modelo de flujo para el acuífero para los periodos 2010-2020, 2020-2030, en los pozos elegidos para conformar la red de monitoreo, se obtuvieron resultados satisfactorios.
- La CEAQ propuso un grupo de pozos existentes como posibles para formar parte de la red de monitoreo, esto por motivo de que construir pozos exclusivos de monitoreo conllevaría a asumir grandes costos, estos pozos no presentaron una distribución espacial homogénea, se tuvieron que incluir pozos para que se tuviera la distribución deseada. En base a esto y a los resultados obtenidos en el análisis geoestadístico se procedió a obtener los pozos que disminuyeran la varianza al máximo, con ayuda del programa filtro de Kalman y de la estimación de abatimientos para los años de predicción (2010-2030) por medio de kriging ordinario. La red de monitoreo óptima quedó conformada por 29 pozos propuestos por la CEAQ y 3 pozos ubicados en las zonas propuestas aquí, haciendo una cantidad de 32 pozos de monitoreo.
- Es suficiente con medir los niveles del agua una vez al año. Sin embargo, se recomienda que una vez cada cinco años se realice un monitoreo más denso, de entre 80 y 100 pozos, que haga posible actualizar el semivariograma que se utiliza en las estimaciones, ya que las condiciones climáticas y el cambio en la naturaleza por los seres humanos son un factor que no se puede controlar y pueden cambiar en el futuro, y con esto las condiciones del acuífero.

- Se observó que no siempre aumentar la cantidad de pozos disminuye la varianza en los errores de la estimación obtenidos en la validación cruzada, el principal factor es la información que proporciona cada pozo, que depende en parte de la estructura o superficie que tengan los abatimientos, es decir que, la heterogeneidad, gradientes horizontales y verticales, fracturas, propiedades del suelo e hidráulicas, aunado a lo anterior la exactitud del modelo numérico de flujo, y sobre todo, el grado de explotación de los niveles del agua del acuífero son los principales factores que influyen en la obtención de una estimación adecuada en base a los pozos seleccionados para formar la red de monitoreo.
- Anteriormente a este diseño de la red de monitoreo para medir los abatimientos que presentará el acuífero en años de predicción, se realizó un análisis geoestadístico a los niveles piezométricos, los resultados obtenidos no fueron satisfactorios porque los estadígrafos de los datos, análisis estructural y estadígrafos de errores no se apegaron a los requerimientos marcados para validar el análisis geoestadístico. Los intervalos de los valores menores a mayores en la piezometría son altos, errores grandes en la validación cruzada, y por resultado las estimaciones de los niveles piezométricos presentaron discrepancias con respecto a los niveles simulados por medio del modelo numérico de flujo. Se concluye que esto fue por que los niveles piezométricos en el acuífero presentan una tendencia que no se puede remover en forma adecuada a través de un polinomio, debido a la presencia de abatimientos locales.
- Cada uno de los métodos vistos en el capítulo 2, aplicados de manera independiente, puede no ser satisfactorio para el propósito que se quiera alcanzar. Es necesario hacer una combinación de los métodos y aplicar el conocimiento tanto empírico como científico, combinar tanto el criterio del especialista como las ciencias matemáticas. Para lograr el objetivo de esta tesis fue necesario (como ya se mencionó en la metodología utilizada en este trabajo) llevar a cabo una combinación de un método geoestadístico, del filtro de Kalman, un método de optimización heurístico y criterios basado en conocimientos geohidrológicos. Con el método geoestadístico y el del filtro de Kalman se llegó a una serie de pozos que efectivamente redujeron al máximo la varianza, pero no así los costos ó cantidad mínima de pozos; en el método heurístico se analizaron los pozos que se podían eliminar reduciendo costos pero sin aumentar la

varianza. Finalmente con el criterio geohidrológico se identificaron las zonas que se necesitaba incluir pozos, zonas que por el comportamiento del flujo subterráneo no permitían reducir las varianzas y las frecuencias de monitoreo. La ventaja de la metodología que se utilizó en esta tesis, es el hecho de poder combinar las diversas metodologías, y tener la certeza de la exactitud que proporciona las ciencias matemáticas aplicando también criterios geohidrológico.

- Una desventaja que se tiene en este método es el procesamiento de datos, sobre todo en la parte de geoestadística, ya que no es posible calcular en forma automática con un programa de cómputo los datos de entrada necesarios para aplicar el programa del filtro de Kalman. Pero en todo estudio, y como ya se dijo anteriormente, es indispensable combinar el criterio humano con la tecnología para obtener resultados satisfactorios.
- Para finalizar es muy importante continuar con los estudios de diseños de redes de monitoreo para los niveles del agua, de especial atención lo que respecta a la parte de los gradientes verticales, ya que estos provocan obtener resultados erróneos en las estimaciones.

BIBLIOGRAFÍA

1. Armstrong, M., et Carignan, J. (1997). Géostatistique Linéaire, Application au Domaine Minier. École de Mines de Paris, France, 112 p.
2. Andricevic, R. (1990). "A Real-Time Approach to Management and monitoring of Groundwater Hydraulics". Water Resources Research, 26(11), 2747-2755.
3. ASCE Task Committee on Geostatistical Techniques. (1990a). "Review of geostatistics in geohydrology I: Basic concepts". Journal of Hydraulic Engineering, 116(5), 612-632.
4. ASCE Task Committee on Geostatistical Techniques. (1990b). Review of geostatistics in geohydrology II: Applications". Journal of Hydraulic Engineering, 116(5), 633-658.
5. Buller, R. D., Gradet, A., and Reed, V. S. (1984). "Ground water monitoring system design in a complex geologic setting." In Proceedings of the Fourth National Symposium on Aquifer Restoration and Ground Water Monitoring, National Water Well Association, Worthington, OH, 186-194.
6. Cameron, K., and Hunter, P. (2000). Optimization of LTM networks using GTS: Statistical approaches to spatial and temporal redundancy, Air Force Center for Environmental Excellence, Brooks AFB, TX. (Online at <http://www.afcee.brooks.af.mil/er/rpo/GTSOptPeper.pdf>).
7. CEA-UAQ. (2002). Convenio de colaboración, "Estudio Integral Del Recurso Agua En Los Acuíferos Del Estado De Querétaro", México.
8. Chauvet, P. (1994). Aide-Memoire de Géostatistique Minière, École des mines de Paris, France 210p.
9. Chica-Olmo, M. (1997). Análisis Geoestadístico en el Estudio de la Explotación de Recursos Minerales, Tesis Doctoral, Universidad de Granada, España, 387 p.
10. Comisión Estatal de Aguas de Querétaro. (2002). Plan Hidráulico del Estado de Querétaro.
11. Comisión Estatal de Aguas de Querétaro. (2005). Evaluación del impacto regional de la recuperación del acuífero del valle de Querétaro debido a la puesta en marcha de la presa Extóraz.
12. Comisión Estatal de Agua y Saneamiento de Guanajuato. (1998). Estudio hidrogeológico y modelo matemático del acuífero del valle de los Apaseos, Guanajuato.

-
-
13. Conwell, P. M., Silliman, S. E., and Zheng, L. (1997). "Design of a piezometer network for estimation of the sample variogram of the hydraulic gradient: The role of the instrument". *Water Resources Research*, 33(11), 2489–2494.
 14. Dagan, G. (1985). "Stochastic Modeling Of Groundwater Flow by Unconditional and Conditional Probabilities: The Inverse Problem". *Water Resources Research*, 21(1), 65-72.
 15. Davis, C. B., and McNichols, R. J. (1994a). Ground water monitoring statistics update: part I: Progress since 1988. *Ground Water* (fall), 148-158.
 16. Davis, C. B., and McNichols, R. J. (1994b). Ground water monitoring statistics update: part II: Nonparametric prediction limits. *Ground Water* (fall), 159-175.
 17. Davis, C. B. (1994). Environmental regulatory statistics. In *Handbook of Statistics*, Volume 12, Chapter 26, p. 817-865, Patil, G. P. and Rao, C. R., eds., Elsevier Science, New York.
 18. Díaz, V. Martín A. (2002). "Geoestadística Aplicada", Instituto de Geofísica, UNAM, Instituto de Geofísica y Astronomía, CITMA, Cuba.
 19. Dougherty, D. E., and Marryott, R. A. (1991). "Optimal groundwater management: 1. Simulated annealing." *Water Resources Research*. 27 (10), 2493-2508.
 20. Efron, B. y G. Gong. (1983). "A leisurely look at the bootstrap, the jack-knife, and cross-validation", *The American Statistician*, 37(1).
 21. Ely, D. M., Hill, M. C., Tiedeman, C. R., and 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.
 22. Geofísica de Exploraciones Guysa, S.A. de C.V. (1996). Estudio de simulación hidrodinámica y diseño óptimo de la red de observación en los acuíferos de Aguascalientes y Querétaro. tomo III. Comisión Nacional del Agua. Gerencia de Aguas Subterráneas, Querétaro, GAS-008-PR-96.
 23. Herrera Graciela. (1998). *Cost Effective Groundwater Quality Sampling Network Design*, Ph.D. Thesis. University of Vermont.

-
-
24. Herrera Graciela, Joseph Guarnaccia and George F. Pinder, 2000. A methodology for the design of space-time groundwater quality sampling networks, Proceedings of the Conference Computational Methods in Water resources XIII, vol. 1, Computational methods, subsurface and transport, Eds. L. R. Bentley, et al., Balkema, Rotterdam, 579-585.
 25. Herrera, G. S., and Pinder G. F. (2003). Cost Effective Groundwater Quality Sampling Network Design. Reporte de la Universidad de Vermont, Burlington. Vermont.
 26. Hill, M. C., Ely, M. D., Tiedeman, C. R., D'Agnesse, F. A., Faunt, C. C., and 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://water.usgs.gov/pubs/wri/wri004282/>.
 27. Holland, J. H. (1975). Adaptation in natural artificial systems. University of Michigan Press, Ann Arbor, MI.
 28. Hudak, P. F. and Loáiciga, H. A. (1993). "An optimization method for monitoring network design in multilayered ground-water flow systems." Water Resources Research, 29(8), 2835-2845.
 29. Hudak, P. F. Loáiciga, H. A., and Marino, M. A. (1995). "Regional-scale ground-water quality monitoring via integer programming". Journal of Hydrology, 164. 153-170.
 30. Hoeksema, R. J., and Kitanidis, P. K. (1984). "An Application of the Geostatistical Approach to the Inverse Problem in Two-Dimensional Groundwater Modeling". Water Resources Research, 20(7), 1003-1020.
 31. Hudak, P. F., and Loaiciga. H. A. (1993). "An Optimization Method for Monitoring Network Design in Multilayered Groundwater Flow Systems". Water Resources Research, 29(8), 2835-2845.
 32. INEGI, XII Censo General de Población y Vivienda, 2000 http://www.inegi.gob.mx/est/librerias/tabulados.asp?tabulado=tab_po04a&c=708&e=

-
-
33. INEGI, Vías de comunicación, _____, 2004
<http://mapserver.inegi.gob.mx/geografia/espanol/estados/qro/viascomunicacion.cfm?c=458&e=22&CFID=260381&CFTOKEN=21021484>.
 34. INEGI, Regiones y cuencas hidrológicas, 2004
<http://mapserver.inegi.gob.mx/geografia/espanol/estados/qro/rh.cfm>.
 35. INEGI, Temperatura media mensual, 2004.
<http://mapserver.inegi.gob.mx/geografia/espanol/estados/qro/temper-tot-mens.cfm?c=451&e=22&CFID=260381&CFTOKEN=21021484>.
 36. Isaaks, E. H., and Srivastava, R. M. (1989). *An introduction to Applied Geostatistics*. Oxford University Press.
 37. JÚnez Ferreira Hugo. (2005). *Diseño de una red de monitoreo de la calidad del agua para el acuífero Irapuato-Valle, Guanajuato*, Tesis de maestría, Universidad Nacional Autónoma de México.
 38. Johnson, V. M., Ridley, M. N. Tuckfield, R. C., and Anderson, R. A. (1996). "Reducing the sampling frequency of ground water monitoring wells," *Environmental Science and Technology*. 30, 355-358.
 39. Kitanidis P. K., and Vomvoris E. G. (1983). "A Geostatistical Approach to the Inverse Problem In Groundwater Modeling (Steady State) and One-Dimensional Simulations". *Water Resources Research*, 19(3), 677-690.
 40. Loáiciga, H. A. (1989). "An optimization approach for ground-water quality monitoring network design." *Water Resources Research*, 25(8), 1771-1780
 41. Lucas, J. M. (1982). "Combined Shewhart-CUSUM quality control schemes". *Journal of Quality Technology*, 14, 51-59.
 42. Mahar, P. S. and Datta, B. (1997). "Optimal monitoring network and ground-water-pollution source identification." *Journal of Water Resources Planning and Management*, 123(4), 199-207.
 43. Marryott, R. A., Dougherty, D. E., and Stollar, R. L. (1993). "Optimal groundwater management: 2. Application of simulated annealing to a field-scale contamination site." *Water Resources Research*, 29, (4) 847-860.
 44. Matheron. G., 1975, "Les Concepts de Base et L'Evolution de la Geostatistique Miniere.", NATO A.S.I. Rome, Italie, 13-25 Octobre. 8 p.

-
-
45. Massmann, J., and Freeze, R. A. (1987). "Ground-water contamination from waste-management sites: The interaction between risk-based engineering design and regulatory policy, 1. Methodology." *Water Resources Research*, 23, (4) 847-860.
 46. Meyer, P. D., and Brill, E. D. (1988). "Method for locating wells in a ground-water monitoring network under conditions of uncertainty." *Water Resources Research*, 24 (8), 1277-1282.
 47. Meyer P. D., and Valocchi, A. J., and Eheart, J. W. (1994). "Monitoring network design to provide initial detection of groundwater contamination." *Water Resources Research*, 30(9), 2467-2659.
 48. Naber, S., Buxton, B., McMillan, N., and Soares, A. (1997). "Statistical method for assessing the effectiveness of intrinsic remediation." *Proceeding of the Fourth International Symposium on In Situ and On-Site Bioremediation*, 5, 348-354, Battelle Press, Columbus, OH.
 49. Nunes, L. M., Paralta, E., Cunha, M. C., and Ribeiro, L. (2004). "Groundwater nitrate monitoring network optimization with missing data". *Water Resources Research*, Vol 40.
 50. Olivares, José-Luis. (2001). *Argus ONE-PTC interface, v.2.2, User Guide*, documentación de la interfaz gráfica distribuida por Research Center for Groundwater Remediation Design (RCGRD), Universidad de Vermont, E.U.A.
 51. Reed, P., Minsker, B., and Valocchi, A. J. (2000). "Cost-effective long-term ground-water monitoring design using a genetic algorithm and global mass interpolation." *Water Resources Research*, 36(12), 3731-3741.
 52. Reed, P., Minsker, B., and Valocchi, A. J. (2001a). "Why optimize long term groundwater monitoring design? A multiobjective case study of Hill Air Force Base." *Proceeding of the ASCE World Water And Environmental Resources Congress Orlando Florida, FL*.
 53. Reed, P., Minsker, B., and Goldberg, D. (2001b). "A multiobjective approach to cost effective long-term groundwater monitoring using an elitist nondominated sorted genetic algorithm with historical data." *Journal of Hydroinformatics*, 3(2), 71-89.
 54. Ridley, M., and MacQueen, D. (1995). "Cost-effective sampling of groundwater monitoring wells: A data review & well frequency evaluation." *Proceedings of the Hazardous Materials Management Conference and Exhibition*, April 4-6, 1995, San Jose, California, 14-21. Updated version at <http://www-erd/library/>.

-
-
55. Rizzo, D. M., Dougherty, D. E. (1996). "Design optimization for multiple management period groundwater remediation." *Water Resources Research*, 32(8), 2549-2561.
 56. Rizzo, D. M., Dougherty, D. E., and Yu, M. (2000). "An adaptive long-term monitoring and operations system (aLTMOs™) for optimization in environmental management," *Building Partnerships: Proceeding of 2000 ASCE Joint Conference on Water Resources Engineering and Water Resources Planning and Management*. Minniapolis. MN (ISBN 0-7844-0517-4), ed. R. H. Hotchkiss and M. Glade, American Society of Civil Engineers, Reston, VA.
 57. Rizzo, D. M., Dougherty, D. E., and Yu, M. (2000). "An adaptive monitoring and operations system (aLTMOs™) for environmental management." In *Building Partnerships: Proc. of 2000 ASCE Joint Conference on Water Resources Engineering and Water Resources Planning & Management*, R. H. Hotchkiss and M. Glade (eds.), American Society of Civil Engineers, Reston, VA. CD-ROM (ISBN 0-7844-0517-4).
 58. Rizzo, D. M., Dougherty, D. E., and Yu, M. (2001). "Devising groundwater LTM strategies for different objectives using optimization," *Bringing the Gap, Proceeding of 2001 EWRI/ASCE World Water and Environmental Resources Congress*, Orlando, FL (ISBN 0-7844-9569-7) ed. D. Phelps and G. Sehlke, American Society of Civil Engineers, Reston, VA.
 59. Rouhani, S. (1985). "Variance reduction analysis." *Water Resources Research*, 21(6), 837-846.
 60. Rouhani, S. and Hall, T. J. (1988). "Geostatistical schemes for ground-water sampling." *Journal of Hydrology*, 81(1), 85-102.
 61. Samper Calvete, F. Javier y Carrera Ramírez, Jesús. (1990). "Aplicaciones a la hidrología subterránea", *Geoestadística*. Centro Internacional de Métodos Numéricos en Ingeniería, Universidad Politécnica de Catalunya, Barcelona.
 62. Scheibe, T. D., and Letternmaier, D. P. (1989). "Risk-based selection of monitoring wells for assesseing agricultural chemical contamination of ground water." *Ground Water Monitoring Review*, 9(4), 98-105.
 63. Simuta Champo, Roel. (2005). *Modelo en elemento finito para el flujo del acuífero del valle de Querétaro*, Tesis de maestría, Universidad Nacional Autónoma de México.

-
-
64. Storck, P. J., Eheart, J. W., and Valocchi, A. J. (1997). "A method for the optimal location of monitoring wells for detection of ground-water contamination in three-dimensional aquifers." *Water Resources Research*, 33(9), 2081-2088.
 65. Tuckfield, R. C., Shine, E. P., Hiersgesell, R. A., Denham, M. E., Beardsley, C., and Reboul, S. (2001). "Using geoscience and geostatistics to optimize ground-water monitoring networks at the Savannah River Site." *Proceeding of the Water and Environment Congress*, 1-17, ASCE Press, Reston, Virginia.
 66. U.S. EPA. (1989). *Statistical analysis of ground-water monitoring data at RCRA (Resource Conservation and Recovery Act) facilities, interim final guidance*. U.S. EPA Report No. EPA/530-SW-89-026, Office of Solid Waste, USEPA, Washington, D. C.
 67. Van Geer, F. C., Te Stroet, C. B. M., and 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.
 68. Wagner, B. (1995). "Sampling design methods for groundwater modeling under uncertainty." *Water Resource Research*, 31(10), 2581-2591.
 69. Yeh, T. C. J. "Stochastic modeling of groundwater flow and solute transport in aquifers." *Hydrological Processes*, 6, 369-395.
 70. Yangxiao, Z., C. B. Te Stroet, and F. Van Geer, Using Kalman filtering to improve and quantify the uncertainty of numerical groundwater simulations. 2. Application to monitoring network design. *Water Resour. Res.*, 27 (1991), 1999.

ANEXO I

Introducción

En este capítulo se presentan conceptos básicos que se utilizaron en esta tesis, ya que en toda ella se van a usar continuamente una serie de términos estadísticos, estos son: Estadística Clásica y Probabilidad, Geoestadística, y Filtro de Kalman.

En base a esta teoría es posible formular las ecuaciones necesarias para alcanzar el objetivo que se persigue, es decir, el diseño de una red de monitoreo.

Estos conceptos se han obtenido de tres referencias, la sección 1 se obtuvo de Samper y Carrera (1990), la sección 2 de Díaz Martín (2002), y la sección 3 de Herrera y Pinder (2003).

1. Estadística clásica y probabilidad

1.1 Funciones de densidad y distribución

Estos conceptos se van a explicar para funciones continuas únicamente.

La función de distribución $F(x)$ de una variable aleatoria X establece la probabilidad de que X sea igual o menor que x , es decir:

$$F(x) = Prob(X \leq x). \quad (1.1)$$

La función de densidad $f(t)$ está relacionada con la distribución mediante

$$F(x) = \int_x^{\infty} f(t) dt. \quad (1.2)$$

Dadas n variables aleatorias X_1, X_2, \dots, X_n , la función de distribución n -dimensional se define de forma similar a (1.1) como:

$$F(x_1, x_2, \dots, x_n) = Prob(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n). \quad (1.3)$$

Para el caso de dos variables aleatorias (X, Y) la función de distribución conjunta bidimensional se denota con F_{xy} . La distribución marginal de X, $F_x(x)$ se define como

$$F_x(x) = \lim_{y \rightarrow \infty} F_{xy}(x, y). \quad (1.4)$$

Si $F_{xy}(x, y) = F_x(x) F_y(y)$ para todo (x, y) se dice que las variables aleatorias X e Y son independientes. Si existe una función $f_{xy}(u, v)$ no negativa tal que

$$F_{xy}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{xy}(u, v) du dv, \quad (1.5)$$

se dice que F_{xy} tiene función de densidad bidimensional f_{xy} .

Si X e Y son independientes, $f_{xy}(x, y)$ se puede expresar como el producto $f_x(x)f_y(y)$ donde $f_x(x)$ y $f_y(y)$ son las funciones de densidad de las distribuciones marginales de X e Y respectivamente. La función de distribución condicional de X dado $Y=y$ se define como:

$$F_{x|y}(x|y) = \frac{\text{Prob}(X \leq x, Y = y)}{\text{Prob}(Y = y)}, \quad (1.6)$$

si $\text{Prob}(Y = y) > 0$. Algunas propiedades de $F_{x|y}$ son:

- $F_{x|y}(x|y)$ es una función de distribución de x para un y dado.
- $F_{x|y}(x|y)$ es una función de y para un x fijo.
- Para cualesquiera valores de x e y se cumple:

$$\text{Prob}(X \leq x, Y \leq y) = \int_{-\infty}^y F_{x|y}(x|u) f_y(u) du. \quad (1.7)$$

- El teorema de las probabilidades compuestas establece que para una familia de sucesos totalmente excluyentes A_i y exhaustivos (su unión es igual al espacio muestral) se cumple que:

$$\text{Prob}(B) = \sum_i \text{Prob}(A_i) \text{Prob}(B|A_i), \quad (1.8a)$$

una expresión semejante es:

$$Prob(X \leq x) = \int_{-\infty}^{\infty} F_{x,y}(x,y) f_y(y) dy. \quad (1.8b)$$

La función de densidad condicionada de X dado Y=y se define como:

$$f_{x,y}(x,y) = \frac{f_{xy}(x,y)}{f_y(y)}. \quad (1.9)$$

Para toda función g(X) cuya esperanza matemática es finita, $E[g(X)|Y=y]$ es una función de y. Otras propiedades interesantes acerca de la esperanza matemática condicionada son:

(a) $E[g(X)|Y] = E[g(X)]$ si X e Y son independientes

(b) $E[g(X)h(Y)|Y] = h(Y)E[g(X)|Y]$

(c) $E[g(X)|X] = g(X)$

(d) $E[g(X)] = E\{E[g(X)|Y]\}$

Para el caso particular g(X) = X, (d) establece que

$$E\{E[X|Y]\} = E(X). \quad (1.10)$$

Se puede demostrar que

$$Var\{E[X|Y]\} \leq Var(X). \quad (1.11)$$

En esta última propiedad se basan los estimadores condicionados.

La covarianza Cov(X, Y) o σ_{XY} de X e Y define como:

$$Cov(X, Y) = E\{(X - E(X))(Y - E(Y))\}. \quad (1.12)$$

Nótese que $\text{Cov}(X, Y) = \text{Cov}(Y, X)$ y $\text{Cov}(X, X) = \sigma_X^2$ es la varianza de X. Si $\text{Cov}(X, Y) = 0$, se dice que X e Y son variables aleatorias no correlacionadas.

El coeficiente de correlación ρ_{XY} se define como:

$$\rho_{XY} = \frac{\text{Cov}(x, y)}{\sigma_X \sigma_Y} \quad (1.13)$$

y mide el grado de dependencia lineal entre X e Y. Por ejemplo, si $Y = aX + b$, se cumple que:

$$\rho_{XY} = \begin{cases} 1 & \text{si } a > 0 \\ -1 & \text{si } a < 0 \end{cases}$$

Se puede demostrar que $-1 \leq \rho_{XY} \leq 1$. Nótese que si X e Y son independientes, $\rho_{XY} = 0$. El recíproco, sin embargo, no siempre es cierto. Un caso en el que sí lo es correspondiente al caso en que X e Y tienen una distribución normal bidimensional.

1.2 Propiedades de la distribución normal

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

$$f(x) = (2\pi)^{-n/2} C^{-1/2} \exp\left\{-\frac{1}{2}(x-\mu)' C^{-1}(x-\mu)\right\}, \quad (1.14)$$

donde:

x = vector n-dimensional (x_1, x_2, \dots, x_n)

μ = vector n-dimensional de las medias, $\mu_i = E(X_i)$.

C = Matriz $n \times n$ de covarianza de las n variables aleatorias, definida como $C_{ij} = \text{Cov}(X_i, X_j)$.

C^{-1} = Matriz inversa de C, y $|C^{-1}|$: su determinante.

El superíndice t quiere decir transpuesta. Para el caso de una sola variable, (1.14) se reduce a la conocida fórmula:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]. \quad (1.15)$$

Algunas propiedades interesantes de la distribución normal son:

- (a) (X_1, X_2, \dots, X_n) tienen una distribución normal n-dimensional si y sólo si cualquier combinación lineal $Y = \sum_i a_i X_i$ tiene también una distribución normal 1-dimensional.
- (b) Si $(X_1, X_2, \dots, X_n, Y)$ tienen una distribución normal (n+1)-dimensional se cumple que:
 - (b.1) La distribución de Y condicionada a $X_i = x_i$ para todo $i = 1, 2, \dots, n$ es también normal.
 - (b.2) La varianza de la distribución condicionada de Y no depende de los x_i .
 - (b.3) Existe un conjunto de números reales b_i tales que para todo x_1, x_2, \dots, x_n :

$$E[Y | X_1 = x_1, X_2 = x_2, \dots, X_n = x_n] = E[Y] + \sum_{i=1}^n b_i x_i. \quad (1.16)$$

1.3 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 P).

Formalmente, $\{e_n\}$ es un proceso de ruido blanco si $E(e_n) = 0$ para toda n. $E(e_i^T e_i) = P$ y $E(e_i e_j^T) = 0$ para i diferente de j.

1.4 Muestreo aleatorio

Sean X_1, X_2, \dots, X_n los valores de las n observaciones obtenidas en un muestreo aleatorio de una variable aleatoria X cuya media y varianza son μ y σ^2 respectivamente. La media muestral \bar{X}_n se calcula mediante

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i, \quad (1.17)$$

y la varianza muestral S^2 a partir de

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2. \quad (1.18)$$

La media y la varianza muestral, \bar{X}_n y S^2 , satisfacen las siguientes propiedades:

$$E(\bar{X}_n) = \mu,$$

$$E(S^2) = \sigma^2,$$

$$\text{Var } \bar{X}_n = \frac{\sigma^2}{n}, \quad (1.19)$$

De acuerdo con el teorema central del límite, cuando n es suficientemente grande la distribución de \bar{X}_n tiende a una distribución normal $N(\mu, \sigma / \sqrt{n})$ de media μ y varianza σ^2/n . Puesto que σ^2 es en general desconocida y sólo se dispone de su estimador S^2 , sería deseable obtener un resultado análogo al anterior pero en términos de S^2 . De hecho es posible, ya que cuando n es suficientemente grande $\sqrt{n}(\bar{X}_n - \mu)/S$ tiende a una distribución t de Student con $(n-1)$ grados de libertad.

1.5 Estimación e inferencia estadística

El problema de la inferencia estadística consiste en la estimación de los parámetros de una población o variable aleatoria X basada en la información contenida en una muestra de dicha población (varias observaciones de la variable aleatoria.) Un estimador θ^* es un estadístico (es

decir, una función de las observaciones) que nos permite obtener un valor aproximado de alguna característica de la distribución de una variable aleatoria. Puesto que es una función de las n observaciones X_1, \dots, X_n , el estimador es otra variable aleatoria que tiene su propia función de distribución $g(\theta^*, \theta)$ y su correspondiente ley de probabilidad. Se pueden distinguir estimadores por punto y estimadores por intervalo que tienen asociado un nivel de confianza.

1.5.1 Estimadores por punto

En lo que sigue θ denota el parámetro a estimar y θ^* su estimador que en general es una función $\theta^*(X_1, X_2, \dots, X_n)$ de las observaciones.

Se dice que θ^* es un estimador insesgado de θ si:

$$E(\theta^*) = \theta. \quad (1.20)$$

en caso contrario se verifica que

$$E(\theta^*) = \theta + b(\theta),$$

donde $b(\theta)$ es el sesgo. La media y varianza muestrales son estimadores insesgados de la media y de la varianza.

El error cuadrático medio, ECM, de un estimador θ^* se define como $E[(\theta^* - \theta)^2]$ y es igual a la varianza más el cuadrado del sesgo:

$$ECM(\theta^*) = Var \theta^* + b(\theta)^2$$

Nótese que si el estimador es sesgado, el estimador de mínima varianza no tiene porqué coincidir con el estimador de mínimo ECM.

Se dice que θ^* es un estimador consistente de θ , si θ^* converge en probabilidad a θ . En el caso particular en el que θ^* es insesgado, θ^* es consistente si

$$\lim_{n \rightarrow \infty} \text{Var } \theta^* = 0.$$

La media muestral \bar{X}_n es un ejemplo de estimador consistente.

1.5.2 Estimadores por intervalo

El objeto de los estimadores por intervalo es construir alrededor de θ^* un intervalo de forma que sepamos, con una probabilidad dada, que θ (desconocido) está contenido en él. Este intervalo ($\theta^* - h, \theta^* + h$) se denomina intervalo de confianza y se define como:

$$\text{Prob}[\theta^* - \theta \leq h] = 1 - \alpha, \quad (1.21)$$

donde α es el nivel de confianza. Para la media muestral de tamaño n , un intervalo aproximado de confianza del $100(1 - \alpha)\%$ viene dado por $\bar{X}_n \pm t_{\alpha/2} S / \sqrt{n}$ donde $t_{\alpha/2}$ es el valor de la t de Student con $(n-1)$ grados de libertad para $\alpha/2$.

1.6 Regresión lineal. Mínimos cuadrados

Sea Y una variable aleatoria que puede expresarse como

$$Y = m + e, \quad (1.22)$$

donde e representa el error de fluctuación de Y en torno a su valor esperado m . Generalmente m se expresa como una función lineal de p parámetros b_1, b_2, \dots, b_p y p variables $(1, X_1, X_2, \dots, X_{p-1})$. Dadas n observaciones (Y_1, Y_2, \dots, Y_n) correspondientes a n experimentos, el problema que se plantea es la estimación de los parámetros b_1, b_2, \dots, b_p , de forma que minimice la suma de los cuadrados de los errores

$$Se = \sum_{i=1}^n e_i^2.$$

Para acortar la notación, llamamos Y al vector columna de los n valores observados de Y , b al vector columna de los p parámetros y X a la matriz $(n \times p)$ cuyos elementos X_{ij} corresponden

al valor de la variable X_{j-1} en la observación i . Nótese que todos los elementos de la primera columna de X son iguales a 1. El vector de errores e tiene esperanza matemática nula, $E(e) = 0$ y una matriz de covarianza que denotaremos $\sigma^2 C_e$ siendo C_e una matriz definida positiva.

1.6.1 Mínimos cuadrados ordinarios

Si los errores e_i son independientes, su matriz de covarianza puede expresarse como $\sigma^2 \mathbf{I}$, donde \mathbf{I} es la matriz de identidad de orden n . El estimador de b por mínimos cuadrados ordinarios b^* viene dado por:

$$b^* = (X'X)^{-1} X'Y, \quad (1.23)$$

y la suma de los cuadrados de los errores S_e es:

$$S_e = Y'Y - b^{*'} X'Xb^*. \quad (1.24)$$

Un estimador insesgado de σ^2 viene dado por S^2 :

$$S^2 = \frac{1}{(n-p)} S_e. \quad (1.25)$$

Finalmente, b^* es un estimador insesgado de b cuya matriz de covarianza C_b viene dada por:

$$C_b = \sigma^2 (X'X)^{-1}. \quad (1.26)$$

1.6.2 Mínimos cuadrados generalizados

Los resultados anteriores pueden generalizarse al caso en que los errores e_i no son independientes. En este caso se tiene:

$$b^* = (X' C_e^{-1} X)^{-1} X' C_e^{-1} Y, \quad (1.27)$$

$$S_e = (Y - Xb^*)' C^{-1} (Y - Xb^*). \quad (1.28)$$

$$C_b = \sigma^2 (X' C_e^{-1} X)^{-1}, \quad (1.29)$$

Los mínimos cuadrados ponderados son un caso particular de mínimos cuadrados generalizados en el que la matriz C_e es diagonal (no necesariamente igual a la matriz identidad).

2. Geoestadística

2.1 Variables regionalizadas

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

Una definición más rigurosa matemáticamente equivalente consistiría en decir que una variable regionalizada es una variable aleatoria z definida en un punto del espacio \underline{x} . Donde en el caso más general \underline{x} es un punto en el espacio tridimensional, es decir, $\underline{x} = (x_1, x_2, x_3)$.

2.2 Función aleatoria

Si a cada punto \underline{x} que pertenece a un dominio en el espacio le hacemos corresponder una variable aleatoria $z(\underline{x})$, que en sentido general pueden ser dependientes, entonces el conjunto de variables aleatorias espacialmente distribuidas $\{z(\underline{x}), \underline{x} \in \Omega\}$ será una función aleatoria $Z(\underline{x})$.

Al tomar una muestra de una función aleatoria $Z(\underline{x})$, a la que llamaremos realización, se obtendrá una función espacial discreta $Z' = \{Z(\underline{x}_i), \underline{x}_i \in \Omega, i = 1, \dots, n\}$ la cual constituye una variable regionalizada. Es decir una realización de una función aleatoria $Z(\underline{x})$ es una variable regionalizada Z' .

2.2.1 Momentos de una función aleatoria

- El momento de primer orden de $Z(\underline{x})$ es la esperanza matemática definida como:

$$m(x) = E[Z(x)], \quad (2.1)$$

- Los momentos de segundo orden considerados en geoestadística son:

i) La varianza de $Z(\underline{x})$

$$\sigma^2(x) = \text{Var}[Z(x)] = E\left[\{Z(x) - m(x)\}^2\right], \quad (2.2)$$

ii) La covarianza de dos variables aleatorias $Z(\underline{x}_i)$ y $Z(\underline{x}_j)$ definida como:

$$C(x_i, x_j) = E\left[\{Z(x_i) - m(x_i)\}\{Z(x_j) - m(x_j)\}\right]. \quad (2.3)$$

Esta función es conocida como función de autocovarianza.

iii) El semivariograma γ que se define como:

$$2\gamma(x_i, x_j) = \text{Var}[Z(x_i) - Z(x_j)], \quad (2.4)$$

$$\gamma(x_i, x_j) = \frac{1}{2} E\left[\{Z(x_i) - Z(x_j)\}^2\right]. \quad (2.5)$$

También conocido como función de semivarianza. Además el semivariograma es $2\gamma(\underline{x}_i, \underline{x}_j)$ pero con frecuencia se usa el término indistintamente para designar a $\gamma(\underline{x}_i, \underline{x}_j)$.

2.2.2 Funciones aleatorias estacionarias

Se dice que una función aleatoria es estrictamente estacionaria si su función de distribución es invariante a cualquier traslación respecto a un vector \underline{h} o lo que es equivalente, la función de distribución del vector aleatorio $\{Z(\underline{x}_1), Z(\underline{x}_2), \dots, Z(\underline{x}_n)\}$ es idéntica a la del vector $\{Z(\underline{x}_1 + \underline{h}), Z(\underline{x}_2 + \underline{h}), \dots, Z(\underline{x}_n + \underline{h})\}$ para cualquier \underline{h} .

Usualmente sólo se trabaja con los momentos hasta de segundo orden resulta práctico limitar la hipótesis de estacionaridad a estos primeros momentos. Una función aleatoria es estacionaria de segundo orden si se cumple que:

i) su valor esperado existe y no depende de \underline{x} .

$$E[Z(x)] = m; \quad \forall x. \quad (2.6)$$

ii) para cualquier par de variables aleatorias $Z(\underline{x})$ y $Z(\underline{x} + \underline{h})$, su covarianza existe y sólo depende del vector de separación \underline{h} ,

$$C(\underline{h}) \equiv C(\underline{x} + \underline{h}, \underline{x}) = E[Z(\underline{x} + \underline{h})Z(\underline{x})] - m^2. \quad (2.7)$$

La estacionaridad de la varianza implica que la varianza existe, es finita y no depende de \underline{x} , es decir,

$$\sigma^2 = C(0) = Var[Z(\underline{x})]. \quad (2.8)$$

Así mismo bajo esta hipótesis el semivariograma también es estacionario y se cumple que:

$$\gamma(\underline{h}) \equiv \gamma(\underline{x} + \underline{h}, \underline{x}) = \frac{1}{2} E[\{Z(\underline{x} + \underline{h}) - Z(\underline{x})\}^2]. \quad (2.9)$$

Además existe una relación directa entre el semivariograma y la función de covarianza

$$\gamma(\underline{h}) = C(0) - C(\underline{h}). \quad (2.10)$$

Es suficiente usar una de las dos funciones para caracterizar la dependencia espacial.

2.2.3 Funciones aleatorias intrínsecas

Existen funciones aleatorias $Z(\underline{x})$ que representan a fenómenos físicos que muestran una capacidad casi limitada de variación, por lo que para estas funciones no están definidas la varianza ni la covarianza. Sin embargo existen casos en que sus incrementos o diferencias $Z(\underline{x} + \underline{h}) - Z(\underline{x})$ tienen una varianza finita. En otras palabras, esto quiere decir que las diferencias son estacionarias de segundo orden.

Por lo tanto las funciones aleatorias intrínsecas son aquellas que cumplen las siguientes condiciones:

i) El valor esperado de las diferencias es

$$E[Z(\underline{x} + \underline{h}) - Z(\underline{x})] = 0, \quad (2.11)$$

ii) La varianza de las diferencias es

$$\text{Var}[Z(x+h) - Z(x)] = 2\gamma(h). \quad (2.12)$$

Estas condiciones se conocen como *Hipótesis Intrínseca*.

Es evidente que una función aleatoria estacionaria de segundo orden es siempre intrínseca. Lo contrario no se cumple. A las funciones que cumplen con la hipótesis intrínseca se les considera como *débilmente estacionarias*.

2.2.4 Funciones aleatorias no estacionarias

Las *funciones aleatorias no estacionarias* son aquellas cuya esperanza matemática depende de \underline{x} :

$$E[Z(\underline{x})] = m(\underline{x}). \quad (2.13)$$

A $m(\underline{x})$ se le conoce como función de deriva o tendencia.

Si consideramos a la función aleatoria $Z(\underline{x})$ como la suma de una componente determinística $m(\underline{x})$ y de un residuo $R(\underline{x})$ estacionario con media nula, es decir:

$$Z(\underline{x}) = m(\underline{x}) + R(\underline{x}). \quad (2.14)$$

Entonces vemos que el semivariograma de $Z(\underline{x})$ depende de \underline{x}

$$\gamma(\underline{x} + h, \underline{x}) = \gamma_R(h) + \frac{1}{2} \{m(\underline{x} + h) - m(\underline{x})\}^2. \quad (2.15)$$

En el caso en que la deriva sea lineal $m(\underline{x}) = m_0 + m_1 \cdot x$ el semivariograma no depende de \underline{x}

$$\gamma(h) = \gamma_R(h) + \frac{1}{2} (m_1 \cdot h)^2, \quad (2.16)$$

pero crece con el cuadrado de \underline{h} , lo cual puede servir de indicador para la detección de la existencia de la tendencia.

Existe un enfoque que considera a las funciones aleatorias no estacionarias como *intrínsecas de orden k* . Esto quiere decir que si se toman las diferencias de un orden k apropiado éstas resultan ser estacionarias.

2.3 Análisis estructural

El análisis estructural es uno de los tópicos más importantes de la geoestadística puesto que se encarga de la caracterización de la estructura espacial de una propiedad o fenómeno regionalizado. Es el proceso en el marco del cual se obtiene un modelo geoestadístico para la función aleatoria que se estudia.

El análisis estructural consiste en estimar y modelar una función que refleje la correlación espacial de la variable regionalizada a partir de la adopción razonada de la hipótesis más adecuada acerca de su variabilidad. Esto quiere decir, que en la dependencia de las características de estacionaridad del fenómeno se modelará la función de covarianzas o la de semivarianzas.

2.3.1 Estimadores del semivariograma

La forma de estimación más común del semivariograma viene dada por

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i + h) - Z(x_i)]^2, \quad (2.17)$$

donde $N(\underline{h})$ es el número de pares $Z(\underline{x}_i)$ y $Z(\underline{x}_i + \underline{h})$ separados a una distancia $h = \underline{h}$.

Debido a que $\gamma(h)$ es esencialmente una media muestral, tiene todas las desventajas comúnmente asociadas a este tipo de estimador como es la no robustez.

Sus características son las siguientes:

- Es un estimador no paramétrico.
- Es óptimo cuando se dispone de una malla regular de muestreo que sea representativa y la distribución sea normal. En estas condiciones el sesgo es mínimo posible.

En la práctica a menudo el empleo de este estimador produce semivariogramas experimentales erráticos, lo cuál se debe a desviaciones del caso ideal para la aplicación del mismo.

Estas desviaciones pueden ser enumeradas por su importancia en el orden siguiente:

- **Desviación en la distribución.**

$Z(\underline{x})$ y $Z(\underline{x} + \underline{h})$ no pueden ser representadas apropiadamente por una distribución binormal. Con frecuencia $Z(\underline{x})$ es sesgada y con grandes colas.

- **Heterocedasticidad.**

Se dice que los datos son heterocedásticos, si el diagrama del valor medio m contra la desviación estándar σ , calculadas en vecindades móviles, muestra que la dispersión de los valores (medidos por σ) está relacionada con su magnitud (medida por m). Por lo tanto, por ser el semivariograma una medida de la dispersión, su magnitud está ligada a la magnitud de los valores de los datos. Cuando se da este tipo de fenómeno, se dice que existe un efecto proporcional.

- **Desviaciones en el muestreo.**

Las observaciones $\{Z(\underline{x}_i), i = 1: \dots, n\}$ de la función aleatoria $Z(\underline{x})$ no están distribuidas espacialmente en forma aleatoria. Consecuentemente se produce un sesgo en el muestreo, es decir, en la elección de las localizaciones de las observaciones, tendientes a estar agrupadas en

áreas de altos o bajos valores. Estas desviaciones pueden ser detectadas mediante un gráfico del número de muestras n contra el valor medio m para vecindades móviles de un tamaño determinado.

- **Existencia de valores atípicos (*outliers*).**

Normalmente las observaciones son correctas, pero a veces suelen aparecer observaciones discordantes que pueden ser debidas a errores humanos y/o mal funcionamiento de los instrumentos de medición. Si es posible constatar que un valor de estos es erróneo, lo mejor es eliminarlo, aunque esto no siempre es factible. De ahí que la eliminación de estas observaciones pueda causar una subestimación del potencial de un área.

2.3.2 Consideraciones prácticas para el cómputo del semivariograma

Para el cómputo del semivariograma es necesario tener en cuenta algunas reglas prácticas que permiten elevar la eficiencia y la calidad de la estimación, independientemente del tipo de estimador que se utilice.

Estas son las reglas siguientes:

- En la estimación del semivariograma los pares de las observaciones se agrupan según la distancia dentro de un intervalo $h = h$, conocido como “lag”, con una tolerancia $\pm \Delta h/2$ y dentro de una dirección θ con una tolerancia $\pm \Delta \theta/2$. El semivariograma así estimado es considerado suavizado o regularizado.
- El semivariograma muestral debe ser considerado solamente para pequeñas distancias por lo que generalmente, se estima para valores de h menores que la mitad de la distancia máxima ($h < d_{\max} / 2$).
- La elección del número de intervalos es arbitraria. No obstante se considera que un número máximo de 25 intervalos es suficiente para cualquier propósito, y un mínimo de 10 debe de ser usado para determinar con precisión el rango y la meseta del semivariograma.

- El largo de los intervalos debe ser elegido de forma tal que el número de pares en cada intervalo sea lo suficientemente grande para que el estimado del semivariograma sea relativamente estable. Se considera que entre 30 y 50 pares satisfacen este requerimiento.
- Los valores estimados para cada intervalo se deben graficar contra la distancia promedio de todos los pares que se encuentran dentro de dicho intervalo.

2.3.3 Formas generales del semivariograma muestral

En sentido amplio se considera por su forma que hay dos tipos principales de semivariogramas. En el primer tipo, la semivarianza se incrementa con el incremento del valor absoluto del intervalo h hasta alcanzar un valor máximo a partir del cuál se mantiene relativamente constante y oscila alrededor del mismo. Estos semivariogramas son conocidos como de tipo transitivo. 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 $Z(\underline{x})$.

Una variable con este tipo de semivariograma no sólo cumple la hipótesis intrínseca, sino también es estacionaria de segundo orden. Esto significa que su valor esperado es constante

$$E[Z(x)] = m; \quad \forall x. \quad (2.18)$$

y la función de covarianza existe y está definida por

$$\begin{aligned} C(h) &= Cov[Z(x+h), Z(x)] \\ &\equiv E[\{Z(x) - m\}\{Z(x+h) - m\}] \end{aligned} \quad (2.19)$$

y el semivariograma está relacionado con las funciones de covarianza y de autocorrelación como sigue:

$$\gamma(h) = C(0) - C(h) \quad (2.20)$$

$$\rho(h) = C(h) / C(0) = 1 - \gamma(h) / C(0), \quad (2.21)$$

donde $\sigma^2 = C(0)$ es la varianza de la muestra, también conocida como varianza *a priori*, de la función aleatoria $Z(\underline{x})$.

El segundo tipo de semivariograma aparenta un incremento sin límites, es decir, son no acotados, por esto no presentan una varianza *a priori* finita.

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 “nugget” o pepita, y el valor del semivariograma en cero $\gamma(0)$ es conocido como la varianza “nugget” o microvarianza. En principio esto puede ocurrir solamente si existen discontinuidades en la función aleatoria. En la práctica si existencia se debe a la variación espacial que no puede explicar el semivariograma debido a la escala del muestreo. Para un material continuamente variable el efecto nugget se produce a partir de la contribución de los errores de medición y la variación a distancias mucho menores que el intervalo de muestreo más pequeño.

2.3.4 Condiciones que deben cumplir los modelos del semivariograma

En la selección de una función adecuada para ser ajustada a un semivariograma muestral, se debe tener en cuenta en la mayoría de los casos hasta tres elementos: un intercepto con la ordenada, una sección monótonamente creciente y una meseta. Sin embargo, no servirá cualquier modelo que aparente ajustar a los valores empíricos debido a la siguiente razón.

Supongamos que $Z(\underline{x})$ es una función aleatoria estacionaria de segundo orden de la cual obtenemos la variable regionalizada $\{Z(\underline{x}_i), i = 1, \dots, n\}$, donde su semivariograma y su función de covarianzas son $\gamma(h)$ y $C(h)$ respectivamente.

Sea Y una combinación lineal de $Z(\underline{x}_i)$ tal que

$$Y = \sum_{i=1}^n \lambda_i Z(x_i). \quad (2.22)$$

La magnitud Y es también una variable regionalizada con varianza:

$$\text{Var}[Y] = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C(x_i, x_j). \quad (2.23)$$

La varianza de Y puede ser positiva o cero, pero no negativa y la función de covarianza en la parte derecha de la expresión anterior debe asegurar que esta condición sea satisfecha.

La matriz de covarianzas

$$\begin{bmatrix} C(x_1, x_1) & C(x_1, x_2) & \dots & C(x_1, x_n) \\ C(x_2, x_1) & C(x_2, x_2) & \dots & C(x_2, x_n) \\ \dots & \dots & \dots & \dots \\ C(x_n, x_1) & C(x_n, x_2) & \dots & C(x_n, x_n) \end{bmatrix}, \quad (2.24)$$

debe ser definida positiva, es decir, el determinante y todos sus menores principales son positivos o cero. Aquí las covarianzas están definidas como $C(x_i, x_j) = \text{Cov}[Z(x_i), Z(x_j)]$; $i, j = 1, \dots, n$.

La función de covarianzas $C(h)$ si existe debe ser positiva semidefinida por lo que serán permisibles funciones que cumplan este criterio. La definición de función positiva semidefinida es la siguiente:

Sea $q(x_1, \dots, x_n)$ una forma cuadrática diremos que q es semidefinida positiva *para todo* $\bar{x} \in R^n, q(\bar{x}) \geq 0$ y existe algún vector no nulo \bar{x}^0 tal que $q(\bar{x}^0) = 0$.

Las propiedades que no tengan covarianzas a priori finitas no tendrán definida la función de covarianzas. En el caso de que cumplan la hipótesis intrínseca, entonces estará definido el semivariograma y se cumple que:

$$Var[Y] = C(0) \sum_{i=1}^n \lambda_i \sum_{j=1}^n \lambda_j - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma(x_i, x_j). \quad (2.25)$$

Si los pesos λ_i suman 0, es decir $\sum_{i=1}^n \lambda_i = 0$ entonces:

$$Var[Y] = - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma(x_i, x_j). \quad (2.26)$$

Esto implica que $-\gamma(\underline{h})$ debe ser condicionalmente positiva semidefinida, con la condición $\sum_{i=1}^n \lambda_i = 0$. Esto es equivalente a decir que $\gamma(\underline{h})$ debe ser condicionalmente negativa semidefinida.

Como consecuencia de esta propiedad, se puede demostrar que el semivariograma debe tener un ritmo de crecimiento inferior a h^2 , es decir se debe cumplir que:

$$\lim_{h \rightarrow \infty} \frac{\gamma(h)}{h^2} = 0. \quad (2.27)$$

Cuando no se satisface esta condición puede ser un indicador de que la función aleatoria no sea estacionaria.

Por otro lado de la definición de semivariograma se infiere que éste debe anularse en el origen, es decir $\gamma(0) = 0$.

Las funciones que cumplen las condiciones anteriores son conocidas como modelos autorizados del semivariograma.

Como una propiedad importante se debe destacar que cualquier combinación lineal de modelos autorizados es un modelo autorizado.

2.3.5 Modelos autorizados para el semivariograma

Según la forma del semivariograma se pueden dividir en transitivos y no acotados.

2.3.5.1 Modelos transitivos o acotados

Este grupo de modelos se deriva a partir de la noción de autocorrelación entre los valores promedios dentro de los bloques. La idea es que la función aleatoria, de la cual la propiedad medida es una realización, depende del grado de solapamiento de los bloques, es decir, una zona de transición.

- Modelo lineal con meseta.

En una dimensión los bloques simplemente son líneas. Supongamos que estos son de longitud a y la distancia entre sus centros es h . Su solapamiento es entonces $a - h$, el cual, cuando se expresa como una proporción de a es $1 - (h/a)$, siendo $h \leq a$.

La función de autocorrelación es entonces:

$$\rho(h) = 1 - \frac{h}{a}, \quad (2.28)$$

y el semivariograma se expresa como

$$\gamma(h) = \begin{cases} S(h/a) & \text{para } 0 \leq h \leq a \\ S & \text{para } h > a \end{cases} \quad (2.29)$$

- Modelo circular

En dos dimensiones, los bloques son discos de diámetro a . El área de intersección de éstos, separados a una distancia h de sus centros es:

$$A = \frac{a^2}{2} \cos^{-1}\left(\frac{h}{a}\right) - \frac{h}{2} \sqrt{a^2 - h^2}; \quad \text{para } h \leq a. \quad (2.30)$$

Expresado como una fracción del área del disco obtenemos la función de autocorrelación

$$\rho(h) = \frac{2}{\pi} \left\{ \cos^{-1} \left(\frac{h}{a} \right) - \frac{h}{a} \cdot 1 - (h/a)^2 \right\}, \quad (2.31)$$

y el semivariograma

$$\gamma(h) = \begin{cases} S \left[1 - \frac{2}{\pi} \cos^{-1} \left(\frac{h}{a} \right) - \frac{2h}{\pi a} \cdot 1 - (h/a)^2 \right] & \text{para } 0 \leq h \leq a; \\ S & \text{para } h > a \end{cases}, \quad (2.32)$$

$$\text{gradiente} = 4S/(\pi a)$$

Este modelo no ha encontrado mucha aplicación en las ciencias de la tierra.

- Modelo esférico

Por analogía podemos derivar este modelo considerando el solapamiento de los volúmenes de dos esferas de diámetro a y h la distancia que separa sus centros. El volumen de intersección es:

$$V = \frac{\pi}{4} \frac{2a^3}{3} - a^2 h + \frac{h^3}{3}; \quad \text{para } h \leq a. \quad (2.33)$$

Dividiendo por el volumen de la esfera ($\pi a^3/6$) se obtiene la función de autocorrelación

$$\rho(h) = 1 - \frac{3h}{2a} + \frac{1}{2} \left(\frac{h}{a} \right)^3, \quad (2.34)$$

y el semivariograma

$$\gamma(h) = \begin{cases} S \left[3 \left(\frac{h}{a} \right) - \left(\frac{h}{a} \right)^3 \right] & \text{para } 0 \leq h \leq a, \\ S & \text{para } h > a \end{cases}, \quad (2.35)$$

$$\text{gradiente} = 3S/(2a)$$

Debe señalarse que el modelo esférico es apropiado para el caso de tres dimensiones aunque se puede aplicar para casos de una y dos dimensiones.

El modelo circular sólo es apropiado para una y dos dimensiones. El lineal solamente para el caso unidimensional. Esto se debe a que solamente en estos casos estos modelos son funciones condicionalmente negativas semidefinidas.

- Modelo exponencial

Si el solapamiento de los bloques varía su tamaño de forma aleatoria, entonces el semivariograma resulta exponencial. En el caso isotrópico es:

$$\gamma(h) = S \left(1 - e^{-\frac{h}{r}} \right) \quad \text{para } h \geq 0, \quad (2.36)$$

$$\text{gradiente} = S/r$$

Se considera como rango efectivo $a = 3r$.

Los procesos autorregresivos de primer orden y de Markov producen modelos exponenciales.

- Modelo Gaussiano

Está dado por la expresión:

$$\gamma(h) = S \left(1 - e^{-\left(\frac{h}{r}\right)^2} \right) \quad \text{para } h \geq 0, \quad (2.37)$$

donde r es un parámetro no lineal que determina la escala espacial de la variación, como en el caso exponencial. El rango efectivo se considera $a = 3r$, que corresponde al valor $0.95S$ del semivariograma.

- Modelo de efecto agujero

El efecto agujero es indicativo de fenómenos con componentes periódicas. Las expresiones más comunes de modelos de semivariogramas son:

$$\gamma(h) = S \left(1 - \frac{\sin(h)}{h} \right) \quad \text{para } h > 0. \quad (2.38)$$

Esto puede ser usado para representar procesos regularmente continuos y que muestran un comportamiento periódico, el cual es frecuentemente encontrado. donde existe una sucesión de zonas ricas y pobres. Es un modelo negativo definido en tres dimensiones.

Otra alternativa es:

$$\gamma(h) = S(1 - \cos(h)) \quad \text{para } h \geq 0. \quad (2.39)$$

Si el efecto agujero es observado muy acentuado en cierta dirección, por ejemplo la vertical, cuando el fenómeno es una sucesión pseudoperiódica de estratificación horizontal. Este modelo es negativo definido en una dimensión.

2.3.5.2 Modelos no acotados

- Modelos potencia

Existen casos en que la varianza aparenta incrementarse indefinidamente. Así también si se toma cada vez un menor intervalo de muestreo, siempre existe alguna variación que queda sin resolver. Un punto de partida para entender esta variación es el movimiento Browniano en una dimensión. en el cual:

$$Z(x) - Z(x+h) = \varepsilon \quad (2.40)$$

es una variable aleatoria gaussiana, espacialmente independiente.

Su semivariograma es:

$$\gamma(h) = \frac{1}{2} h^\theta; \text{ para } 0 < \theta < 2. \quad (2.41)$$

Mandelbrot llamó al resultado de tales procesos fractales.

Existe una relación entre la dimensión fractal y el parámetro θ dada por la expresión

$$D = 2 - (\theta / 2), \quad (2.42)$$

donde podemos ver los siguientes casos extremos

- para $\theta = 2$ es una parábola y $D = 1$, no representa un proceso aleatorio, además

$$\gamma(h) = kh^2 \quad (2.43)$$

no es función condicionalmente negativa definida.

- para $\theta = 0$ representa el caso de ruido puro

Con frecuencia notamos que muchos semivariogramas aparentan ser de varianza nugget pura. Esto ocurre porque toda la varianza está dentro de un intervalo de muestreo más pequeño.

- Modelo de efecto pepita puro

Formalmente se puede definir como sigue:

$$\gamma(h) = S[1 - \delta(h)]. \quad (2.44)$$

- Modelo logarítmico (Modelo de Wjjsian)

Se define como:

$$\gamma(h) = k \log(h). \quad (2.45)$$

Puede ser de utilidad cuando el semivariograma experimental se comporta linealmente si se usa una escala logarítmica para las distancias.

2.3.6 Modelación del semivariograma experimental

2.3.6.1 Métodos de ajuste

Algunos geoestadísticos ajustan los modelos de forma visual, pero esta práctica no es fiable y es preferible usar algún procedimiento estadístico para estos fines. Con frecuencia es usada la aproximación por Mínimos Cuadrados.

- **Mínimos cuadrados ordinarios (MCO).**

Asume que los resultados están normalmente distribuidos y son independientes y que las semivarianzas estimadas poseen igual varianza..

- **Mínimos cuadrados generalizados (MCG).**

Se evitan muchos de los problemas de los MCO, pero requiere de la inversión de grandes matrices, lo que hace el procedimiento muy complicado.

- **Mínimos cuadrados ponderados.**

Se minimiza la siguiente expresión:

$$W(\tau, \gamma_k) = \sum_{j=1}^k \frac{[\gamma(h_j) - \gamma(h_j, \tau)]^2}{Var[\gamma(h_j)]}. \quad (2.46)$$

donde

τ es el vector de los parámetros del modelo,

$Var[\gamma(h_j)]$ es la varianza del valor estimado del semivariograma en h_j .

Desafortunadamente esta varianza es desconocida casi siempre. Con frecuencia se escoge en su lugar el número de pares de observaciones $m(h_j)$ con que se estima la misma, suponiendo que es inversamente proporcional a la varianza de los estimados.

Se ha demostrado teóricamente que un mejor procedimiento sería reemplazar $Var [\gamma(h_j)]$ por $[\gamma(h_j, \tau)]^2/m(h_j)$ dando un peso considerable mayor a las semivarianzas esperadas menores. La diferencia en los resultados de la aplicación de estos dos tipos de pesos es pequeña y se ha confirmado que cuando se utiliza $m(h_j)$ se le asignan pesos muy pequeños a los estimados de los intervalos h_j menores.

Laslett ha sugerido un mejoramiento. Sus pesos están dados por:

$$\frac{m(h_j) \gamma(h_j)}{[\gamma(h_j, \tau)]^3}, \quad (2.47)$$

y estos son actualizados en cada iteración del proceso cuando $\gamma(h_j, \tau)$ es reestimada. Teóricamente esta forma de ponderación debe dar mejor convergencia y en la práctica se ha encontrado que así es.

2.3.6.2 Criterios de selección de modelos

Un compromiso entre la bondad de ajuste y la complejidad del modelo puede ofrecerlo el Criterio de Información de Akaike (AIC), que se define como

$$AIC = -2 \ln(\text{máx. verosimilitud}) + 2(\text{núm. de parámetros}) \quad (2.48)$$

Y se puede estimar usando

$$(AIC) = \left\{ n \ln \left(\frac{2\pi}{n} \right) + n + 2 \right\} + n \ln(R) + 2p. \quad (2.49)$$

Debido a que la cantidad que se encuentra entre llaves en la expresión anterior (3.50) es constante ($n \ln(2\pi/n) + n + 2 = \text{const}$) independientemente del tipo de modelo, entonces para fines prácticos se calcula:

$$A = n \ln(R) + 2p, \quad (2.50)$$

que es un estimador simplificado del Criterio de Información de Akaike. Donde n es el número de valores estimados $\{\hat{\gamma}(h_i), i = 1, \dots, n\}$ del semivariograma muestral, R es la suma residual de los cuadrados de las diferencias entre los valores experimentales $\hat{\gamma}(h_i)$ y los del modelo ajustado $\gamma(h_i)$, es decir $R = \sum_{i=1}^n (\gamma(h_i) - \hat{\gamma}(h_i))^2$, mientras que p es el número de parámetros del modelo de semivariograma ajustado $\gamma(h)$.

Se considera que el modelo que presenta el menor AIC es el mejor. Este criterio es el que se utilizó en este trabajo.

2.3.7 Validación del modelo del semivariograma

Para validar el modelo obtenido de semivariograma se puede proceder de varias maneras. Un método que resulta atractivo por su sencillez y eficiencia es el leave one out (validación cruzada) que consiste en sacar un elemento de la muestra y estimar el valor en ese punto usando kriging con el modelo del semivariograma obtenido. De forma análoga se actúa para el resto de los elementos de la muestra. Como resultado se obtiene un mapa de las diferencias $Z(\underline{x}_i) - Z^*(\underline{x}_i), i = 1, \dots, n$ entre el valor real y el estimado.

De forma tal que si el método del semivariograma refleja adecuadamente la estructura espacial implícita en el conjunto de datos, entonces los valores estimados deben ser cercanos a los valores observados.

Esta “cercanía” puede ser caracterizada según los siguientes estadígrafos:

a) $\frac{1}{n} \sum_{i=1}^n \{Z(x_i) - Z^*(x_i)\}$ cercano a 0.

b) $\frac{1}{n} \sum_{i=1}^n \{Z(x_i) - Z^*(x_i)\}^2$ pequeño.

$$c) \frac{1}{n} \sum_{i=1}^n \left\{ \frac{Z(x_i) - Z^*(x_i)}{\sigma_i} \right\}^2 \text{ cercano a 1.}$$

d) La correlación muestral de $Z(\underline{x}_i)$, $Z^*(\underline{x}_i)$ sea cercana a 1.

e) La correlación muestral de $Z(\underline{x}_i)$, $\{Z(x_i) - Z^*(x_i)\} / \sigma_i$ es cercana a cero.

Siendo:

$Z(\underline{x}_i)$ los valores muestrales de la propiedad en x_i .

$Z^*(\underline{x}_i)$ los valores estimados de la propiedad en el punto x_i .

σ_i es la desviación estándar de la estimación en el punto x_i .

Idealmente todos los estadígrafos anteriores deben satisfacer simultáneamente, pero en la práctica una mejoría en uno de ellos puede degradar a otro. Por lo que es recomendable hacer un análisis integral de los estadígrafos de las diferencias.

Mediante el histograma de los errores normalizados se compila una lista con los puntos que poseen grandes errores. Esto es útil para la identificación de valores atípicos outliers, datos sospechosos o anomalías de otra naturaleza.

La eliminación de las localizaciones con elevados errores normalizados y el recálculo del semivariograma y su modelación puede producir mejorías significativas.

El procedimiento de leave one out es un caso particular del método conocido como Jackknifing, ya que de disponer de suficientes datos espacialmente distribuidos de forma homogénea de la función aleatoria $Z(\underline{x})$, éstos se podrían dividir en dos muestras: $Z_1(\underline{x})$ y $Z_2(\underline{x})$. La primera se usaría para estimar el semivariograma, mientras que la segunda serviría para validarlo. Es decir, se estimarían como en el método anterior los valores empleando kriging con el modelo de semivariograma obtenido usando la primera muestra $Z_1(\underline{x})$ en los puntos correspondientes a las observaciones que pertenecen a la segunda muestra $Z_2(\underline{x})$ y se evaluarían los estadígrafos de las diferencias $Z_1(\underline{x}_i) - Z_2^*(\underline{x}_i)$ de manera análoga.

2.4 Kriging

2.4.1 El mejor estimador lineal insesgado

El kriging es un término que ha sido acuñado para designar al “*mejor estimador lineal insesgado*” de un punto y al mejor promedio lineal móvil ponderado de un bloque.

Este nombre apareció alrededor de 1960 para nombrar una técnica creada en Francia por Matheron a partir de los trabajos de D. G. Krige quién fue probablemente el primero que hizo uso de la evaluación de yacimientos minerales.

El kriging es una técnica de estimación local que ofrece el mejor estimador lineal insesgado de una característica desconocida que se estudia. La limitación a la clase de estimadores lineales es bastante natural ya que esto significa que solamente se requiere el conocimiento del momento de segundo orden de la función aleatoria (la covarianza o el semivariograma) y que en general en la práctica es posible inferir a partir de una realización de la misma.

2.4.2 Clasificación de los diferentes tipos de kriging

I) Según la forma del estimador

- Lineales: Simple, Ordinario, Universal, Residual.
- No lineales: Disyuntivo, Indicador, Probabilístico.

II) Según el soporte de la medición de los datos

- Puntual
- En bloques

III) Kriging paramétrico y no paramétrico

- Paramétrico: Multigaussiano, Disyuntivo, Lognormal.
- No paramétrico: Simple, Ordinario, Universal, Residual, Indicador, Probabilístico.

2.4.3 Planteamiento del problema general

Sea $Z(\underline{x})$ una función aleatoria, la cual está definida en un soporte puntual y es estacionaria de segundo orden, con:

- un valor esperado.

$$E[Z(x)] = m; \quad \forall x. \quad (2.51)$$

donde m es una constante generalmente desconocida,

- una función de covarianza centrada

$$C(h) = E[Z(x+h)Z(x)] - m^2. \quad (2.52)$$

- un semivariograma

$$\text{Var}[Z(x+h) - Z(x)] = 2\gamma(h). \quad (2.53)$$

donde al menos uno de estos dos momentos de segundo orden se supone conocido.

Cuando solamente existe el semivariograma, entonces la función aleatoria $Z(\underline{x})$ se considera intrínseca.

Los valores experimentales consisten en un conjunto de valores discretos $\{Z(x_i), i = 1, \dots, n\}$. Con frecuencia estos valores están definidos en soportes puntuales o casi puntuales, en otros casos son los valores medios $Z_{V_i}(x_i)$ definidos en los soportes V_i , centrados en los puntos \underline{x}_i , donde los n soportes pueden ser todos diferentes.

La estimación del valor medio $Z_{V_i}(x_i)$ en el dominio de V_i se define como

$$Z_{V_i}(x_i) = \frac{1}{V_i} \int_{V_i} Z(x) dx.$$

Es bueno destacar que bajo la hipótesis de estacionaridad el valor esperado de cada uno de estos datos es $E[Z(x_i)] = m, \quad \forall i$.

El estimador lineal Z_k^* considerado es una combinación lineal de n valores de datos tal que:

$$Z^*(x_k) = \sum_{i=1}^n \lambda_i Z(x_i),$$

donde $Z_k^* = Z^*(x_k)$.

Los n coeficientes λ_i son calculadas de manera tal que el estimador sea insesgado y que la varianza de la estimación sea mínima. entonces se dice que el estimador Z_k^* es óptimo.

2.4.4 Ecuaciones del kriging ordinario

Como ya hemos planteado el estimador kriging se considera óptimo ya que es insesgado y es el que minimiza la varianza de la estimación. A partir de estos dos criterios se derivan las ecuaciones de kriging.

- La condición de insesgades.

Para obtener un valor esperado del error igual a cero resulta suficiente imponer la condición:

$$E[Z_k^*] = E\left[\sum_{i=1}^n \lambda_i Z(x_i)\right] = m; \quad (2.54)$$

donde m es el valor esperado de la función aleatoria $Z(\underline{x})$.

Esto implica que

$$\sum_{i=1}^n \lambda_i E[Z(x_i)] = E[Z_k^*], \quad (2.55)$$

entonces

$$\sum_{i=1}^n \lambda_i m = m, \quad (2.56)$$

y finalmente

$$\sum_{i=1}^n \lambda_i = 1, \quad (2.57)$$

- La condición de que la estimación sea de mínima varianza

Para satisfacer esta condición hay que minimizar la siguiente función:

$$F = \sigma_c^2 - 2\mu \left(\sum_{i=1}^n \lambda_i - 1 \right), \quad (2.58)$$

donde

σ_c^2 es la varianza de la estimación.

μ es un multiplicador de Lagrange.

Nótese que la función F a minimizar consta de la varianza de la estimación σ_c^2 e incluye la condición que garantiza el no sesgo de la estimación.

La varianza de la estimación se expresa de la siguiente manera:

$$\sigma_c^2 = \text{Var}[Z_k - Z_k^*] = E[(Z_k - Z_k^*)^2], \quad (2.59)$$

$$\sigma_c^2 = \text{Var}[Z_k] - \text{Cov}[Z_k, Z_k^*] + \text{Var}[Z_k^*]. \quad (2.60)$$

Sustituyendo en esta última fórmula la expresión del estimador Z_k^* se tiene:

$$\sigma_c^2 = Var[Z_k] - 2Cov\left[Z_k, \sum_{i=1}^n \lambda_i Z(x_i)\right] + Var\left[\sum_{i=1}^n \lambda_i Z(x_i)\right], \quad (2.61)$$

desarrollando se obtiene

$$\sigma_c^2 = \sigma_{z_k}^2 - 2\sum_{i=1}^n \lambda_i \sigma_{z_k z_i} + \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \sigma_{z_i z_j}. \quad (2.62)$$

Con las derivadas parciales de F respecto a los coeficientes desconocidos λ_i y con respecto a μ se obtiene el siguiente sistema de ecuaciones:

$$\begin{cases} \frac{\partial F}{\partial \lambda_i} = -2\sigma_{z_k z_i} + 2\sum_{j=1}^n \lambda_j \sigma_{z_i z_j} - 2\mu = 0, & i = 1, \dots, n \\ \frac{\partial F}{\partial \mu} = \sum_{i=1}^n \lambda_i - 1 = 0 \end{cases} \quad (2.63)$$

De una manera más usual se escribe como sigue:

$$\begin{cases} \sum_{j=1}^n \lambda_j \sigma_{z_i z_j} - \mu = \sigma_{z_k z_i}, & i = 1, \dots, n \\ \sum_{i=1}^n \lambda_i = 1 \end{cases} \quad (2.64)$$

El sistema de ecuaciones así obtenido sirve para el cálculo del kriging puntual. Y la varianza del error de la estimación se puede calcular de una manera más simple si se sustituye el valor de μ :

$$\sigma_c^2 = \sigma_{z_k}^2 - \sum_{i=1}^n \lambda_i \sigma_{z_k z_i} + \mu. \quad (2.65)$$

2.4.5 Ecuaciones del kriging ordinario en forma matricial

De forma análoga podemos representar el sistema de ecuaciones del kriging en forma matricial como sigue:

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} & 1 \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdots \\ \lambda_n \\ -\mu \end{bmatrix} = \begin{bmatrix} \sigma_{k1} \\ \sigma_{k2} \\ \cdots \\ \sigma_{kn} \\ 1 \end{bmatrix}, \quad (2.66)$$

donde $\sigma_{ij} = \sigma_{z,z}$.

Nótese que bajo la hipótesis intrínseca las covarianzas pueden ser reemplazadas por las semivarianzas, usando la expresión:

$$\sigma_{ij} = \sigma^2 - \gamma_{ij}, \quad (2.67)$$

donde $\sigma^2 = C(0)$, es la varianza total o a priori de la muestra.

Entonces sustituyendo la expresión de la covarianza en función de la semivarianza obtenemos el sistema del kriging en función de las semivarianzas:

$$\begin{bmatrix} 0 & \gamma_{12} & \cdots & \gamma_{1n} & 1 \\ \gamma_{21} & 0 & \cdots & \gamma_{2n} & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & 0 & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdots \\ \lambda_n \\ \mu \end{bmatrix} = \begin{bmatrix} \gamma_{k1} \\ \gamma_{k2} \\ \cdots \\ \gamma_{kn} \\ 1 \end{bmatrix}. \quad (2.68)$$

Y la varianza de la estimación se calcula como sigue:

$$\sigma_c^2 = \sum_{i=1}^n \lambda_i \gamma_{ki} + \mu. \quad (2.69)$$

Estas últimas expresiones resultan ser la formulación del kriging más común puesto que se aplica en el caso cuando la variable regionalizada cumple la hipótesis intrínseca la cual es una condición menos exigente que la estacionaridad de segundo orden.

3. Filtro de Kalman

El filtro de Kalman provee una estimación lineal insesgada de mínima varianza para el estado de un sistema, dados datos con ruido. También establece una forma para actualizar esta estimación cuando se obtiene un nuevo dato, sin necesidad de recurrir a datos pasados. Este filtro calcula las estimaciones del sistema secuencialmente, comenzando con una estimación a priori y agregando nuevos datos en cada etapa.

En su forma general el filtro de Kalman está basado en dos ecuaciones, una ecuación dinámica y una ecuación de medición. En este trabajo se aplica el filtro de Kalman usando únicamente la ecuación de medición, a lo que llamaremos *filtro de Kalman estático*, y es el que se explica a continuación.

Se tiene un vector estado C que expresa un modelo del sistema que describe la evolución en el tiempo de la cantidad que se desea estimar y $\{z_j, j = 1, 2, \dots\}$ es una secuencia de mediciones escalares. Estas mediciones son relacionadas a través de la ecuación de medición lineal,

$$z_j = H_j C + v_j, \quad (3.1)$$

donde H_j es el j -enésima matriz de medición. La matriz de medición de tamaño $1 \times N$ que es no cero únicamente en la posición correspondiente a la entrada de C a donde el j -enésima medición es tomada, y N es la dimensión del vector C . La secuencia $\{v_j, j = 1, 2, \dots\}$ represente error de medición. Este es un ruido blanco gaussiano, con media cero y varianza r_j . La secuencia del error de medición $\{v_j\}$ y el vector C son independientes.

La estimación insesgada de mínima varianza para el vector C , dada la secuencia de mediciones z_1, \dots, z_n , aquí denotado como \hat{C}^n , es el valor esperado de C dados los datos, esto es, $\hat{C}^n = E\{C | z_1, z_2, \dots, z_n\}$. El subíndice identifica el número de medidas que son usadas para obtener la estimación. La matriz de covarianza del error de estimación es

$$P^n = E\{(C - \hat{C}^n)(C - \hat{C}^n)^T\}, \quad (3.2)$$

donde T denota transpuesto.

Dada una estimación a priori sobre el vector estado, \hat{C}^0 , y la correspondiente matriz de covarianza, P^0 , la estimación lineal de mínima varianza para C se obtiene secuencialmente a través de las siguientes fórmulas:

$$\hat{C}^{n+1} = \hat{C}^n + K_{n+1} (z_{n+1} - H_{n+1} \hat{C}^n). \quad (3.3)$$

$$P^{n+1} = P^n - K_{n+1} H_{n+1} P^n. \quad (3.4)$$

donde

$$K_{n+1} = P^n H_{n+1}^T (H_{n+1} P^n H_{n+1}^T + r_{n+1})^{-1}. \quad (3.5)$$

Así, para calcular secuencialmente las estimaciones de \hat{C}^n y sus matrices de covarianzas P^n , se necesita proveer una estimación a priori de \hat{C}^0 , y su matriz de covarianza, P^0 .

De las fórmulas del filtro de Kalman (ecuaciones 3.3, 3.4 y 3.5) se obtiene el efecto sobre la varianza del error de estimación en la localización i de medición en la localización j . Si $P^k = \{P_{l,m}^k, l, m = 1, \dots, N\}$, entonces el producto $P^k H_{k-1}^T$ es

$$P^k H_{k-1}^T = (P_{1,j}^k, \dots, P_{N,j}^k)^T. \quad (3.6)$$

y

$$(H_{k+1} P_k H_{k+1}^T + r_{k+1})^{-1} = \frac{1}{P_{jj}^k + r_j}. \quad (3.7)$$

Substituyendo en la ecuación 3.5 tenemos

$$K_{k+1} = \frac{1}{P_{jj}^k + r_j} (P_{1,j}^k, \dots, P_{N,j}^k)^T. \quad (3.8)$$

Note que P^{k+1} en la ecuación 3.4 puede ser escrita como

$$P^{k+1} = (I - K_{k+1}H_{k+1})P^k, \quad (3.9)$$

donde I es una matriz de identidad de tamaño $N \times N$. La matriz $I - K_{k+1}H_{k+1}$ es

$$I - K_{k+1}H_{k+1} = \begin{pmatrix} 1 & 0 & \dots & 0 & -\frac{P_{1,i}^k}{P_{ii}^k + r_i} & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & -\frac{P_{2,i}^k}{P_{ii}^k + r_i} & & & \\ \vdots & & \ddots & & \vdots & & \vdots & \vdots \\ & & & 1 - \frac{P_{1,i}^k}{P_{ii}^k + r_i} & & & & \\ & & & -\frac{P_{ii}^k}{P_{ii}^k + r_i} & & \dots & & 0 \\ & & & \vdots & & & & \\ 0 & \dots & 0 & -\frac{P_{N,i}^k}{P_{ii}^k + r_i} & 0 & \dots & 1 \end{pmatrix}. \quad (3.10)$$

Entonces la matriz de covarianza P^k es simétrica. Por ejemplo, $P_{i,j}^k = P_{j,i}^k$, los elementos de la diagonal de la matriz $(I - K_{k+1}H_{k+1})P^k$ son

$$P_{ii}^{k+1} = P_{ii}^k - \frac{(P_{i,i}^k)^2}{r_i + P_{ii}^k}. \quad (3.11)$$

Por lo tanto, el cambio en la varianza de estimación P_{ii}^k en la localización x_i , debido a la medición de la localización x_j , se debe al término $\frac{(P_{i,i}^k)^2}{r_i + P_{ii}^k}$.

El efecto de tomar una nueva muestra en la localización x_j , sobre la varianza total del error está dada por

$$\sigma_r^2(k+1) = \sum_i P_{ii}^{k+1} = \sum_i P_{ii}^k - \frac{1}{r_j + P_{jj}^k} \sum_i (P_{i,i}^k)^2. \quad (3.12)$$

La primera suma en la pasada ecuación es la varianza total del error dadas k mediciones. Por lo tanto,

$$\sigma_i^2(k+1) = \sigma_i^2(k) - \frac{1}{r_i + P_{ii}^k} \sum_j (P_{i,j}^k)^2. \quad (3.13)$$

Entonces $\sigma_i^2(k+1)$ es mínimo si $\sum_j (P_{i,j}^k)^2$ es máximo. Cada nueva medición es elegida por inspección usando esta fórmula.

POZO	LONGITUD	LATTIUD	BROCAL	Dec-93	Dec-95	Abatimiento 93-95	Abatimiento Anual
0548-A	347560.43	2292670.01	1912.990	1894.43	1888.73	-5.70	-2.350
0590	350539.91	2289214.86	1907.270	1832.29	1828.83	-3.46	-1.730
0553	357933.15	2276926.84	1848.030	1721.94	1714.92	-7.02	-3.510
0563	347004.91	2280658.56	1807.685	1707.25	1700.29	-6.96	-3.480
0579	351714.42	2280664.50	1808.140	1722.66	1720.80	-1.86	-0.930
0581	351489.43	2279842.16	1807.250	1713.96	1708.96	-5.00	-2.500
0592	352363.24	2278314.18	1808.830	1713.84	1709.04	-4.80	-2.400
0593	350024.19	2277158.92	1801.678	1720.41	1701.59	-18.82	-9.410
0602	347007.08	2277644.72	1798.980	1709.76	1702.52	-7.24	-3.620
0603	348898.90	2275968.32	1801.133	1710.66	1700.46	-10.20	-5.101
0608	345571.54	2280599.00	1804.281	1704.26	1697.04	-7.22	-3.610
0609-A	345355.38	2280897.86	1806.281	1708.05	1702.12	-5.93	-2.965
0610-A	343706.15	2281037.05	1812.750	1713.24	1704.85	-8.39	-4.195
0612	344682.44	2278850.98	1803.720	1712.45	1701.49	-10.96	-5.480
0613	343659.10	2283712.65	1821.190	1724.02	1715.56	-8.46	-4.230
0613-A	343747.63	2283777.68	1821.190	1724.02	1715.56	-8.46	-4.230
0615-A	344001.13	2281913.19	1811.199	1714.50	1707.84	-6.66	-3.330
0627	350614.75	2274379.35	1804.070	1711.80	1705.77	-6.03	-3.015
0628	350387.28	2273836.93	1805.251	1711.31	1702.80	-8.51	-4.255
0629	348669.68	2273800.62	1807.540	1713.26	1707.25	-6.01	-3.005
634	353561.00	2274061.00	1836.000	1715.15	1711.88	-3.27	-1.635
0641	351806.93	2272614.79	1809.010	1709.78	1703.88	-5.90	-2.950
0650	347769.36	2273095.67	1802.000	1696.62	1687.98	-8.64	-4.320
0651-A	347686.71	2274063.18	1806.242	1707.11	1701.41	-5.70	-2.850
0652	347678.01	2274721.55	1804.394	1718.42	1708.05	-10.37	-5.183
0653	349887.92	2274146.77	1803.483	1720.79	1715.79	-5.00	-2.500
0653-A	349969.00	2274204.00	1803.541	1717.81	1710.03	-7.78	-3.890
0658-A	347301.68	2271399.15	1823.180	1711.90	1706.03	-5.87	-2.935
0662	348473.44	2270867.23	1836.386	1717.44	1706.28	-11.16	-5.578
663	348805.00	2272070.00	1847.000	1723.21	1717.22	-5.99	-2.995
668	348794.00	2273457.00	1823.000	1704.60	1697.71	-6.89	-3.445
0688	345896.56	2271720.79	1847.420	1711.73	1703.60	-8.13	-4.065
0689	344353.29	2271911.42	1833.333	1709.91	1708.26	-1.65	-0.823
0720	356810.75	2278459.13	1833.636	1763.63	1756.54	-7.09	-3.543
0725-A	355420.65	2276956.78	1816.530	1712.00	1704.19	-7.81	-3.905
740	355614.00	2276967.00	1844.000	1708.55	1699.97	-8.58	-4.290
0745	353629.52	2276694.18	1813.831	1701.69	1704.18	2.49	1.245
0752	358234.78	2278528.43	1840.130	1798.99	1794.85	-4.14	-2.070
0759	347068.51	2292307.80	1950.330	1893.61	1892.48	-1.13	-0.565
0824	352688.57	2280271.41	1815.641	1703.66	1696.28	-7.38	-3.690
0826	353398.86	2279318.77	1817.670	1736.00	1733.40	-2.60	-1.300
827	356168.00	2276397.00	1846.000	1699.95	1689.03	-10.92	-5.460
0828	355060.74	2276210.15	1818.000	1700.87	1690.58	-11.29	-5.645
0870	352798.70	2278543.00	1809.900	1716.83	1705.78	-11.05	-5.525
0914	350902.69	2276249.77	1803.060	1711.76	1705.03	-6.73	-3.365
0953	352644.95	2278899.61	1807.691	1706.55	1701.88	-4.67	-2.335
0954	353355.16	2279366.60	1810.488	1666.36	1657.85	-8.51	-4.256
0955-A	353419.00	2279259.85	1810.486	1640.42	1629.24	-11.18	-5.588
0957	352524.04	2278983.80	1807.690	1669.83	1665.11	-4.72	-2.360
0958	352671.87	2279083.17	1807.852	1658.08	1658.79	0.71	0.355
0960	352390.65	2279244.61	1807.624	1649.90	1641.85	-8.05	-4.025
0962	351278.86	2281478.74	1805.453	1718.12	1714.86	-3.26	-1.630
0965	352389.86	2279894.10	1807.945	1651.19	1644.75	-6.44	-3.220
0966-A	351003.74	2282424.04	1806.741	1722.42	1718.52	-3.90	-1.950
0967	350628.10	2280860.03	1803.450	1716.74	1713.78	-2.96	-1.480
0968	350910.31	2280626.98	1803.406	1717.69	1714.26	-3.43	-1.715
0969	352026.28	2279625.00	1806.655	1722.89	1719.31	-3.58	-1.790
0970	354275.73	2283906.88	1897.272	1857.15	1858.67	1.52	0.760
0977-A	351910.44	2278340.06	1803.107	1705.17	1698.46	-6.66	-3.330
0978	351934.35	2273559.45	1807.370	1712.44	1703.43	-9.01	-4.505
0982	344823.34	2278283.85	1798.972	1728.55	1723.27	-5.28	-2.640
0986-A	361672.81	2279168.25	1854.804	1814.68	1811.51	-3.17	-1.585
0987	356080.19	2278086.66	1824.622	1726.11	1722.72	-3.39	-1.694
0988	356665.53	2275806.63	1838.690	1711.20	1704.14	-7.06	-3.530
0988-A	356608.73	2275882.74	1818.003	1711.14	1702.90	-8.24	-4.120
0998	350490.57	2281979.86	1805.876	1702.62	1695.41	-7.21	-3.605
0999-A	343223.08	2272053.10	1829.110	1713.24	1705.48	-7.76	-3.880
1010	351497.30	2281097.79	1805.065	1717.32	1710.07	-7.25	-3.627
1102	361672.79	2279169.53	1853.900	1813.95	1811.09	-2.86	-1.430
1228	343597.12	2285012.22	1833.503	1722.19	1715.64	-6.55	-3.274
1232	351515.92	2285642.03	1830.860	1821.96	1821.69	-0.27	-0.135
1289	351458.86	2271409.77	1815.000	1720.24	1713.86	-6.38	-3.190
1290	344085.73	2271781.44	1831.258	1711.53	1703.95	-7.58	-3.790
1291	345413.42	2272418.40	1847.370	1711.58	1704.89	-6.69	-3.345
1297	349874.12	2282126.88	1806.647	1705.53	1693.15	-12.38	-6.188
1298-A	349869.59	2281976.67	1808.250	1697.56	1694.75	-2.81	-1.405
1299-A	350120.00	2282210.00	1856.000	1703.57	1700.17	-3.40	-1.700
1313-A	346400.99	2279589.92	1799.245	1714.09	1699.43	-14.66	-7.330
1315	346208.62	2277790.31	1798.410	1708.84	1705.89	-2.95	-1.475
1418	361264.69	2279188.32	1851.231	1809.40	1807.20	-2.20	-1.100
1428-A	349750.91	2283293.98	1812.771	1748.96	1733.56	-15.40	-7.700

Tabla II.1 Abatimiento anual para los periodos 93-95 y 95-99

POZO	LONGITUD	LATITUD	BROCAL	Dec.93	Dec.95	Abatimiento 93-95	Abatimiento Anual
1459	347225.00	2273291.00	1807.239	1711.97	1704.08	-7.89	-3.945
1634	349510.00	2286920.00	1830.000	1755.05	1754.60	-0.45	-0.225
1636	350573.72	2282427.69	1.808.35	1716.94	1711.88	-5.06	-2.530
1638	351546.77	2279745.58	1804.320	1714.83	1710.26	-4.57	-2.285
1701	351701.40	2277361.05	1804.570	1706.87	1704.26	-2.61	-1.305
1707	348851.88	2283152.82	1850.000	1723.12	1717.64	-5.48	-2.740
1708	353113.15	2286167.70	1896.057	1873.76	1869.76	-4.00	-2.000
1746-A	349621.29	2272343.73	1805.000	1702.16	1694.81	-7.35	-3.675
1790	351950.00	2277861.00	1804.911	1709.57	1702.99	-6.58	-3.290
1905	353700.00	2272250.00	1900.000	1713.20	1704.80	-8.40	-4.200
1935	345429.42	2273385.79	1837.700	1712.75	1704.20	-8.55	-4.275
1963	353745.17	2285875.04	1882.010	1860.86	1859.38	-1.48	-0.740
1973	347810.49	2281015.54	1800.358	1675.69	1668.52	-7.17	-3.584
1977	347109.18	2277119.43	1799.435	1717.70	1706.50	-11.20	-5.600
1978	349883.83	2284770.73	1816.517	1734.17	1734.20	0.03	0.014
1979	352492.44	2277297.49	1808.616	1707.76	1692.81	-14.95	-7.475
2036	344559.58	2273385.79	1824.518	1705.20	1697.22	-8.48	-4.239
2041	345539.10	2288162.60	1899.901	1854.78	1850.28	-4.50	-2.251
2044	350518.24	2291672.15	1981.472	1892.80	1890.99	-1.81	-0.905
2056	353976.89	2277817.26	1813.703	1712.46	1705.00	-7.46	-3.730
2080	351251.33	2276588.86	1805.000	1717.85	1704.77	-13.08	-6.540
2081	347781.09	2280029.00	1794.670	1699.86	1691.87	-7.99	-3.995
2085	355743.71	2275929.90	1818.000	1702.17	1694.52	-7.65	-3.825
2427	347765.00	2281950.00	1970.000	1696.64	1691.21	-5.43	-2.715
0548-A	347560.43	2292670.01	1912.99	1888.73	1883.46	-5.27	-1.3175
0550	356339.91	2289214.86	1907.27	1828.83	1824.55	-4.28	-1.0700
0553	357933.15	2276926.84	1848.03	1714.92	1709.54	-5.38	-1.3450
0563	347004.91	2280058.56	1807.69	1700.29	1686.33	-13.94	-3.4850
0579	351714.42	2280664.50	1808.14	1720.80	1703.72	-17.08	-4.2700
0581	351489.43	2279842.16	1807.25	1708.96	1698.96	-10.00	-2.5000
0593	350024.19	2277158.92	1801.68	1701.59	1679.89	-21.70	-5.4245
0602	347007.08	2277644.72	1798.98	1702.52	1688.04	-14.48	-3.6200
0605	348898.90	2275968.32	1801.13	1700.46	1688.10	-12.36	-3.0892
0608	345571.54	2280599.00	1804.28	1697.04	1682.59	-14.45	-3.6125
0609-A	345355.38	2280897.86	1806.28	1702.12	1686.15	-15.97	-3.9925
0610-A	343706.15	2281037.05	1812.75	1704.85	1692.43	-12.42	-3.1050
0612	344682.44	2278850.98	1803.72	1701.49	1681.85	-19.64	-4.9100
0613	343699.10	2283712.65	1821.19	1715.56	1705.61	-12.95	-3.2375
0613-A	343747.63	2283777.68	1821.19	1715.56	1702.61	-12.95	-3.2375
0615-A	344001.13	2281913.19	1811.20	1707.84	1697.19	-10.65	-2.6623
0627	350614.75	2274379.35	1804.07	1705.77	1685.26	-20.51	-5.1275
0628	350367.28	2273836.93	1805.25	1702.80	1687.22	-15.58	-3.8950
0629	348669.68	2273800.62	1807.54	1707.25	1688.44	-18.81	-4.7025
634	353561.00	2274061.00	1836.00	1711.88	1705.33	-6.55	-1.6375
0641	351896.93	2272614.79	1809.01	1703.88	1696.36	-7.52	-1.8800
0650	347769.36	2273095.67	1802.00	1687.98	1670.70	-17.28	-4.3200
0651-A	347686.71	2274063.18	1806.24	1701.41	1695.04	-6.37	-1.5925
0652	347678.01	2274721.55	1804.39	1708.05	1691.30	-16.75	-4.1885
0653	349887.92	2274146.77	1803.48	1713.79	1705.79	-10.00	-2.5000
0653-A	349969.00	2274204.00	1803.54	1710.03	1706.05	-3.98	-0.9948
0658-A	347301.68	2271399.15	1823.18	1706.03	1694.27	-11.76	-2.9400
0662	348473.44	2270867.23	1856.39	1706.28	1681.53	-24.75	-6.1875
663	348805.00	2272070.00	1847.00	1717.22	1705.24	-11.98	-2.9950
668	348794.00	2273457.00	1823.00	1697.71	1684.39	-13.32	-3.3300
0688	345896.56	2271720.79	1847.42	1703.60	1694.08	-9.52	-2.3800
0689	344353.29	2271911.42	1833.33	1708.26	1697.56	-10.70	-2.6758
0720	356810.75	2278459.13	1833.64	1756.54	1745.28	-11.26	-2.8150
0725-A	355420.65	2276956.78	1816.53	1704.19	1696.14	-8.05	-2.0125
740	355614.00	2276967.00	1844.00	1699.97	1682.80	-17.17	-4.2925
0745	353629.52	2276694.18	1813.83	1704.18	1688.62	-15.56	-3.8900
0752	358234.78	2278528.43	1840.13	1794.85	1792.06	-2.79	-0.6975
0759	347088.51	2292307.80	1950.33	1892.48	1885.77	-6.71	-1.6775
0824	352688.57	2280271.41	1815.64	1696.28	1681.53	-14.75	-3.6875
0870	352798.70	2278543.00	1809.90	1705.78	1702.10	-3.68	-0.9200
0914	350902.69	2276249.77	1803.06	1705.03	1687.86	-17.17	-4.2925
0953	352644.95	2278899.61	1807.69	1701.88	1688.84	-13.04	-3.2600
0954	353355.16	2279366.60	1810.49	1657.85	1642.24	-15.61	-3.9020
0957	352524.04	2278883.80	1807.69	1665.11	1655.68	-9.43	-2.3575
0958	352671.87	2279083.17	1807.85	1658.79	1654.65	-4.14	-1.0350
0960	352390.65	2279244.61	1807.62	1641.85	1625.73	-16.12	-4.0300
0962	351178.86	2281478.74	1805.45	1714.86	1708.33	-6.53	-1.6325
0965	352389.86	2279894.10	1807.95	1644.75	1631.86	-12.89	-3.2225
0966-A	351003.74	2282424.04	1806.74	1718.52	1710.71	-7.81	-1.9525
0967	350628.10	2280860.63	1803.45	1713.78	1705.96	-7.82	-1.9550
0968	350910.31	2280626.98	1803.41	1714.26	1707.39	-6.87	-1.7175
0969	352026.28	2279625.00	1806.66	1719.31	1712.33	-6.98	-1.7450
0976	354275.73	2285906.88	1897.27	1858.67	1857.21	-1.46	-0.3650
0978	351934.35	2273559.45	1807.37	1703.43	1685.40	-18.03	-4.5075
0982	344823.34	2278283.85	1798.97	1723.27	1689.92	-33.35	-8.3375
0986-A	361672.81	2279168.25	1854.80	1811.51	1801.66	-9.85	-2.4625
0987	356080.19	2278086.66	1824.62	1722.72	1716.57	-6.15	-1.5375

Tabla II.1 Abatimiento anual para los periodos 93-95 y 95-99

POZO	LONGITUD	LATITUD	BROCAL	Dec-93	Dec-95	Abatimiento 93-95	Abatimiento Anual
0988-A	356609.73	2275882.74	1818.00	1702.90	1669.50	-33.40	-8.3493
0998	350490.57	2281979.86	1805.88	1695.41	1680.98	-14.43	-3.6075
0999-A	343223.08	2272053.10	1829.11	1705.48	1689.05	-16.43	-4.1075
1010	351497.30	2281097.79	1805.07	1710.07	1695.54	-14.53	-3.6325
1102	361672.79	2279169.53	1853.90	1811.09	1801.00	-10.09	-2.5225
1228	343397.12	2285012.22	1833.50	1715.64	1702.16	-13.48	-3.3707
1232	351515.92	2285642.03	1836.86	1821.69	1821.19	-0.50	-0.1256
1289	351458.86	2271409.77	1815.00	1713.86	1704.97	-8.89	-2.2225
1290	344085.73	2271781.44	1831.26	1703.95	1691.61	-12.34	-3.0850
1291	345413.42	2272418.40	1847.37	1704.89	1691.08	-13.81	-3.4525
1297	348874.12	2282126.88	1806.65	1693.15	1691.51	-1.64	-0.4100
1299-A	350120.00	2282210.00	1856.00	1700.17	1699.96	-0.21	-0.0525
1313-A	346400.99	2279589.92	1799.25	1699.43	1682.55	-16.88	-4.2200
1315	346208.62	2277790.31	1798.41	1705.89	1697.61	-8.28	-2.0700
1428-A	349750.91	2283293.98	1812.77	1733.56	1701.49	-32.07	-8.0175
1459	347225.00	2273291.00	1807.24	1704.08	1689.46	-14.62	-3.6550
1634	349510.00	2286920.00	1850.00	1754.60	1753.11	-1.49	-0.3725
1636	350573.72	2282427.69	1808.35	1711.88	1701.56	-10.32	-2.5300
1638	351546.77	2279745.58	1804.32	1710.26	1695.06	-15.20	-3.8000
1707	348851.88	2283152.62	1850.00	1717.64	1706.37	-11.27	-2.8175
1708	353113.15	2286167.70	1896.06	1809.76	1870.50	0.80	0.2000
1746-A	349621.29	2272343.73	1805.00	1694.81	1670.00	-24.81	-6.2025
1905	353700.00	2272250.00	1900.00	1704.80	1688.88	-15.92	-3.9800
1935	345429.42	2273385.79	1837.70	1704.20	1693.35	-10.85	-2.7125
1963	353745.17	2285875.04	1882.01	1859.38	1859.50	0.12	0.0300
1973	347810.49	2281015.54	1800.36	1668.52	1650.50	-18.02	-4.5050
1977	347109.18	2277110.43	1799.44	1706.50	1691.02	-15.48	-3.8700
1978	348883.83	2284770.73	1816.52	1734.20	1732.76	-1.44	-0.3600
1979	352492.44	2277297.49	1808.62	1692.81	1686.93	-5.88	-1.4710
2036	344559.58	2273385.79	1824.52	1697.22	1686.21	-11.01	-2.7525
2041	345559.10	2288162.60	1899.90	1850.28	1832.88	-17.40	-4.3497
2044	350518.24	2291672.15	1981.47	1890.99	1886.16	-4.83	-1.2080
2080	351251.33	2276588.86	1805.00	1704.77	1687.52	-17.25	-4.3125
2081	347781.09	2280029.90	1794.67	1691.87	1675.70	-16.17	-4.0425
2427	347765.00	2281950.00	1970.00	1691.21	1680.03	-11.18	-2.7950

Tabla II.1 Abatimiento anual para los periodos 93-95 y 95-99

2010			2020			2030			Abat 10 20		Abat 20 30		Abat anual 10 30			
N	Pozo	PTC	N	PTC	Y	X	Y	PTC	N	Pozo	X	Y	PTC	Abat 10 20	Abat 20 30	Abat anual 10 30
1	7	350058	1	1751.63	2269238	350058	2269238	1751.64	1	7	350058	2269238	1751.64	-16.59	-14.33	-1.66
2	1102	361673	2	1796.25	2279170	361673	2279170	1840.11	2	1102	361673	2279170	1838.99	40.87	-11.12	4.69
3	752	358235	3	1779.88	2278528	358235	2278528	1815.33	3	752	358235	2278528	1804.88	35.85	-10.35	3.59
4	0988-A	356610	4	1665.30	2275883	356610	2275883	1682.84	4	0988-A	356610	2275883	1662.31	17.53	-20.53	1.75
5	759	347009	5	1857.64	2292308	347009	2292308	1845.66	5	759	347009	2292308	1831.82	-12.88	-13.25	-1.26
6	2041	345539	6	1809.10	2288163	345539	2288163	1794.57	6	2041	345539	2288163	1778.46	-14.53	-16.12	-1.45
7	1	349679	7	1849.93	2289686	349679	2289686	1838.53	7	1	349679	2289686	1828.82	-11.40	-11.71	-1.14
8	976	354276	8	1866.51	2283907	354276	2283907	1865.62	8	976	354276	2283907	1864.26	-0.89	-1.37	-0.09
9	3	354527	9	1746.55	2282421	354527	2282421	1754.12	9	3	354527	2282421	1755.38	-12.43	-8.74	-1.24
10	1978	349884	10	1704.60	2284771	349884	2284771	1689.20	10	1978	349884	2284771	1671.79	-18.41	-17.41	-1.54
11	0613-A	343748	11	1675.75	2283778	343748	2283778	1657.96	11	0613-A	343748	2283778	1639.15	-17.79	-18.81	-1.88
12	1434	351115	12	1688.98	2282978	351115	2282978	1674.01	12	1434	351115	2282978	1657.37	-14.98	-16.64	-1.66
13	1428-A	349751	13	1684.47	2283294	349751	2283294	1669.54	13	1428-A	349751	2283294	1651.19	-14.92	-18.35	-1.49
14	0929-A	343273	14	1661.75	2272053	343273	2272053	1642.72	14	0929-A	343273	2272053	1627.59	-19.02	-20.13	-1.99
15	987	350680	15	1679.20	2278087	350680	2278087	1682.59	15	987	350680	2278087	1667.86	3.39	-14.73	0.34
16	0725-A	355421	16	1664.07	2276957	355421	2276957	1661.08	16	0725-A	355421	2276957	1642.29	-2.99	-18.79	-0.50
17	0609-A	345355	17	1666.09	2280898	345355	2280898	1648.15	17	0609-A	345355	2280898	1628.90	-18.04	-20.25	-1.69
18	1010	351497	18	1666.82	2281998	351497	2281998	1658.14	18	1010	351497	2281998	1639.34	-11.68	-18.81	-1.17
19	982	344823	19	1666.05	2273284	344823	2273284	1644.19	19	982	344823	2273284	1623.56	-18.86	-20.63	-1.59
20	1313-A	346401	20	1650.70	2279590	346401	2279590	1635.34	20	1313-A	346401	2279590	1612.81	-15.56	-22.53	-1.54
21	1045	349396	21	1659.19	2279389	349396	2279389	1647.11	21	1045	349396	2279389	1635.22	-18.77	-20.71	-1.94
22	1045	349396	22	1659.19	2279389	349396	2279389	1647.11	22	1045	349396	2279389	1635.22	-18.77	-20.71	-1.94
23	870	352798	23	1664.56	2275413	352798	2275413	1651.24	23	870	352798	2275413	1634.66	-7.77	-10.37	-0.23
24	1638	351547	24	1665.05	2279746	351547	2279746	1654.24	24	1638	351547	2279746	1634.33	-16.81	-19.50	-1.95
25	652	347678	25	1657.39	2274722	347678	2274722	1644.19	25	652	347678	2274722	1623.55	-13.20	-20.65	-2.06
26	1977	347109	26	1658.50	2273110	347109	2273110	1644.65	26	1977	347109	2273110	1623.73	-13.85	-20.92	-1.99
27	0600-A	348957	27	1658.42	2277408	348957	2277408	1645.92	27	0600-A	348957	2277408	1624.88	-12.80	-21.04	-1.25
28	629	348670	28	1656.62	2273801	348670	2273801	1644.29	28	629	348670	2273801	1623.57	-13.32	-20.73	-1.23
29	1746-A	349621	29	1643.53	2272344	349621	2272344	1632.59	29	1746-A	349621	2272344	1611.32	-10.94	-21.77	-1.09
30	605	348899	30	1657.20	2275968	348899	2275968	1645.25	30	605	348899	2275968	1624.47	-11.95	-20.78	-2.08
31	628	350387	31	1656.03	2273837	350387	2273837	1645.87	31	628	350387	2273837	1624.80	-10.15	-21.07	-2.11
32	914	350903	32	1657.53	2276250	350903	2276250	1647.77	32	914	350903	2276250	1627.07	-9.76	-20.71	-0.98

Tabla II.2 Valores de los pozos elegidos por filtro de Kalman extraídos del modelo de flujo.