



Título del Trabajo Fin de Máster:

***EVALUACIÓN DE LA INCERTIDUMBRE  
DE LAS PREDICCIONES DE LOS  
NIVELES PIEZOMÉTRICOS DEL  
ACUÍFERO DE LA FÁBRICA DE  
ANDÚJAR MEDIANTE EL USO DE  
FILTROS DE KALMAN DE CONJUNTOS***

Intensificación:

***RECURSOS HÍDRICOS***

Autor:

***GUTIÉRREZ ESPARZA, JULIO CÉSAR***

Director/es:

***DR. GÓMEZ HERNÁNDEZ, J. JAIME***

Fecha: **ABRIL 2012**



<b>TITULO:</b>	<b>EVALUACIÓN DE LA INCERTIDUMBRE DE LAS PREDICCIONES DE LOS NIVELES PIEZOMÉTRICOS DEL ACUÍFERO DE LA FÁBRICA DE ANDÚJAR MEDIANTE EL USO DE FILTROS DE KALMAN DE CONJUNTOS</b>		
<b>ALUMNO:</b>	<b>JULIO CÉSAR GUTIÉRREZ ESPARZA</b>	<b>LUGAR DE REALIZACION</b>	<b>VALENCIA</b>
<b>DIRECTOR:</b>	<b>J. JAIME GÓMEZ HERNÁNDEZ</b>	<b>FECHA DE LECTURA</b>	<b>MAYO, 2012</b>
<b>RESUMEN:</b>	<p><i>Este trabajo se ha centrado en el área que es inmediata a la Fábrica de Uranio de Andújar (FUA), más concretamente en el acuífero que ahí se localiza, en el cual, tras la clausura de la fábrica se observó la presencia de lixiviados, que se desplazan lentamente desde los diques de estériles. Posterior al descubrimiento se han creado varios modelos matemáticos de flujo y de transporte de solutos, seleccionándose para este estudio el último de los modelos mencionados.</i></p> <p><i>Para evaluar la incertidumbre en las predicciones de los niveles piezométricos se hizo un análisis geostatístico de las conductividades de la zona a estudiar, corroborando los patrones observados en trabajos anteriores; a partir del análisis, se realizaron diversas simulaciones obteniéndose mapas de la conductividad hidráulica de los diferentes materiales existentes en la zona.</i></p> <p><i>Obtenidos los mapas se desarrolló un código numérico que acopla la metodología de los filtros de Kalman de conjuntos al código CORE (Código para la simulación numérica de procesos de flujo de agua, transferencia de calor y transporte de solutos REactivos). Este código utiliza los valores medidos de piezometría para condicionar los campos de conductividad. El análisis de los campos de piezometría resultantes nos permiten hacer una evaluación de la incertidumbre en las predicciones obtenidas con el código CORE.</i></p>		
<b>ABSTRACT:</b>	<p><i>This work focuses on the aquifer underlying the Andújar Uranium Mill (FUA). After the mill shutdown, this aquifer shows presence of contamination from the mill's tailings. Several mathematical groundwater flow and solute transport models have been created, choosing for this study the latest one.</i></p> <p><i>In order to evaluate the piezometric head's uncertainty, a geostatistical analysis of the conductivity data in the studied area has been done, verifying the observed patterns in earlier assessments; from this analysis, different simulations were made, obtaining hydraulic conductivity maps for the study area.</i></p> <p><i>After obtaining the maps, a numerical code has been written that couples CORE (COde for flow, heat transfer and solute transport numerical simulation of REactive solutes) with the ensemble Kalman Filter</i></p>		

	<p><i>methodology. This program generates conductivity maps conditioned to the observed piezometric heads. The analysis of the ensemble of piezometric head maps allows analyzing their uncertainty</i></p>
<p><b>RESUM:</b></p>	<p><i>Aquest treball s'ha centrat en l'àrea que és immediata a la Fabrica d'Urani de Andújar (FUA), més concretament en l'aqüífer que ací es localitza, en el qual, després de la clausura de la fàbrica es va observar la presència de lixiviats, que es desplacen lentament des dels dics d'estèrils. Posterior al descobriment s'han creat diversos models matemàtics de flux i de transport de soluts, seleccionant-se per a aquest estudi l'últim dels models esmentats.</i></p> <p><i>Per a avaluar la incertesa en les prediccions dels nivells piezomètrics es va fer una anàlisi geoestadística de les conductivitats de la zona a estudiar, corroborant els patrons observats en treballs anteriors; a partir de l'anàlisi, es van realitzar diverses simulacions obtenint-se mapes de la conductivitat hidràulica dels diferents materials existents en la zona.</i></p> <p><i>Obtinguts els mapes es va desenvolupar un codi numèric que acobla la metodologia dels filtres de Kalman de conjunts al codi CORE (Codi per a la simulació numèrica de processos de flux d'aigua, transferència de calor i transport de soluts REactius). Aquest codi utilitza els valors mesurats de piezometria per a condicionar els camps de conductivitat. L'anàlisi dels camps de piezometria resultants ens permeten fer una avaluació de la incertesa en les prediccions obtingudes amb el codi CORE.</i></p>
<p><b>PALABRAS CLAVE:</b></p>	<p><i>Incertidumbre/ Analisis Geoestadístico/ Filtro de Kalman de Conjuntos/ Conductividad Hidraulica/ Nivel Piezométrico/</i></p>

## **Agradecimientos**

Quiero agradecer en primera instancia a la Conselleria de Educación, Formación y Empleo, en particular a Vicente José Ruiz Sánchez por su apoyo brindado en los diversos procesos que ha conllevado el ser becario del programa Santiago Grisolia. A Mi director de Tesis J. Jaime Gómez Hernández por el apoyo y por brindarme parte de su experiencia cuando he tenido dudas.

También quiero agradecer a todos los miembros del grupo de Hidrogeología de la Universidad Politécnica de Valencia que, de una u otra manera me han aconsejado o ayudado a lo largo de mi estancia en España, nombrando de manera particular a Oscar, quien ha sido un buen compañero y amigo tanto en lo laboral como fuera de la oficina.

Quiero agradecer a mis padres por darme una buena educación, a mi hermana por darme buenos consejos, quiero también agradecer a mi esposa, quien al igual que yo ha estudiado en esta universidad y que me da su punto de vista sobre el trabajo realizado, pero sobre todo, este documento se lo dedico a Natalia mi hija quien nacerá en este verano.



## **Índice general**

### Índice

<u>general</u> .....	V
<u>Índice de figuras</u> .....	VII
<u>Índice de tablas</u> .....	XI
1 Introducción.....	1
1.1 Objetivos.....	5
1.2 Organización del documento .....	7
2 Estado del arte .....	9
2.1 El uso de los filtros de Kalman.....	9
2.2 Aplicación de los filtros de Kalman de conjuntos en hidrogeología .....	11
2.3 Filtros de Kalman de conjuntos .....	13
3 Modelo del FUA 2004 .....	17
3.1 Gestión de datos del GIS.....	19
4 Análisis geoestadístico .....	23
4.1 Modelo de continuidad espacial .....	23
5 Estimación.....	29
5.1 Kriging ordinario.....	29
6 Simulación.....	33
6.1 Simulación gaussiana secuencial.....	33
7 CORE.....	41
7.1 Aspectos teóricos del código.....	41
7.1.1 CORE acuífero confinado.....	41

7.1.2	CORE acuífero libre .....	43
7.1.3	CORE zona no saturada .....	44
7.2	Solución numérica del flujo en CORE .....	47
7.2.1	Solución del flujo en el acuífero .....	47
7.2.2	Solución del flujo en la zona de transición.....	55
7.3	Datos de entrada al modelo CORE .....	59
7.4	Simulación de alturas piezométricas modelo CORE.....	61
8	Filtro de Kalman de conjuntos .....	67
8.1	Filtro de Kalman de conjuntos aplicado al FUA.....	67
8.2	Descripción del programa.....	69
9	Resultados y conclusiones.....	71
9.1	Resultados .....	71
9.2	Análisis de sensibilidad .....	79
9.2	Análisis de incertidumbre .....	83
9.3	Conclusiones.....	87
10	Bibliografía.....	89

## Índice de figuras

Figura 3.1 Malla de elementos triangulares creada a partir de la información de entrada utilizada por el código CORE.....	20
Figura 3.2 Cuadrícula de 887040 elementos que envuelve el área de estudio.....	20
Figura 3.3 Mapa de localización de los 154 datos de Conductividades.....	21
Figura 4.1 Distribución espacial de los valores de conductividad hidráulica.....	23
Figura 4.2 Histograma y curva de probabilidad acumulada.....	24
Figura 4.3 Variograma omnidireccional.....	25
Figura 4.4 Variograma unidireccional en la dirección de los 30 grados azimut.....	26
Figura 4.5 Variograma unidireccional en la dirección de los 120 grados azimut.....	26
Figura 4.6 Ejes de máxima y mínima continuidad de los datos de conductividad hidráulica..	27
Figura 5.1 Mapa de conductividades obtenido mediante el Krigeado Ordinario.....	31
Figura 5.2 Mapa de la varianza en la conductividad hidráulica de la zona de estudio.....	32
Figura 6.1 Mapa de conductividades hidráulicas (m/d) resultado de la simulación 22.....	38
Figura 6.2 Mapa de conductividades hidráulicas (m/d) resultado de la simulación 35.....	38
Figura 6.3 Mapa de conductividades hidráulicas (m/d) resultado de la simulación 43.....	39
Figura 6.4 Mapa de los valores esperados de las conductividades hidráulicas (m/d).....	39
Figuras 6.5 Varianza Condicional en del Área de estudio .....	40
Figura 7.41.1 Mapa con 12254 zonas de conductividades a partir de la simulación 11.....	62
Figura 7.41.2 Mapa con 12254 zonas de conductividades a partir de la simulación 61.....	63
Figura 7.41.3 Mapa de valores esperados de las 12254 zonas de conductividades a partir de el conjunto de simulaciones.....	63

Figura 7.41.4 Mapa de varianza condicional de las 12254 zonas de conductividades a partir de el conjunto de simulaciones.....	64
Figura 7.4.1.5 Alturas Piezométricas observadas en el PC1 comparadas con los valores que reproduce el CORE para el periodo 2003-2020.....	64
Figura 7.4.1.6 Alturas Piezométricas observadas en el PC5 comparadas con los valores que reproduce el CORE para el periodo 2003-2020.....	65
Figura 7.4.1.7 Alturas Piezométricas observadas en el pozo 611 comparadas con los valores que reproduce el CORE para el periodo 2003-2020.....	65
Figura 7.4.1.8 Alturas Piezométricas observadas en el pozo 472 comparadas con los valores que reproduce el CORE para el periodo 2003-2020.....	66
Figura 8.2.1 Diagrama de flujo del programa de Filtros de Kalman de Conjuntos añadido al código CORE.....	70
Figura 9.1.1 Ajuste de conductividades realizado mediante Filtros de Kalman de conjuntos para el pozo 611 realización 2.....	71
Figura 9.1.2 Ajuste de conductividades realizado mediante Filtros de Kalman de conjuntos para el pozo 505 realización 2. ....	72
Figura 9.1.3 Ajuste de conductividades realizado mediante Filtros de Kalman de conjuntos para el PC6 realización 2. ....	72
Figura 9.1.4 Ajuste de conductividades realizado mediante Filtros de Kalman de conjuntos para el PC5 realización 2. ....	73
Figura 9.1.5 Ajuste de conductividades realizado mediante Filtros de Kalman de conjuntos para el PC3 realización 2. ....	73
Figura 9.1.6 Ajuste de conductividades realizado mediante Filtros de Kalman de conjuntos para el PC2 realización 2. ....	74
Figura 9.1.7 Ajuste de conductividades realizado mediante Filtros de Kalman de conjuntos para el PC1 realización 2.....	74

Figuras 9.1.8 a 9.1.13 Ajuste de conductividades realizado mediante Filtros de Kalman de conjuntos con la realización 3.....	75
Figuras 9.1.14 a 9.1.19 Ajuste de conductividades realizado mediante Filtros de Kalman de conjuntos con la realización 2.....	76
Figura 9.1.20 Campo de conductividades obtenido a partir de los filtros de Kalman de Conjuntos para la Realización 2.....	77
Figura 9.1.21 Campo de conductividades obtenido a partir de los filtros de Kalman de Conjuntos para la Realización 1.....	77
Figura 9.1.22 Campo de conductividades obtenido a partir de los filtros de Kalman de Conjuntos para la Realización 0.....	78
Figura 9.2.1 Comparación de alturas piezométricas obtenidas a partir de 10 campos de conductividades distintos para el PC1.....	79
Figura 9.2.2 Comparación de alturas piezométricas obtenidas a partir de 10 campos de conductividades distintos para el PC2.....	80
Figura 9.2.3 Comparación de alturas piezométricas obtenidas a partir de 10 campos de conductividades distintos para el PC3.....	80
Figura 9.2.4 Comparación de alturas piezométricas obtenidas a partir de 10 campos de conductividades distintos para el PC5.....	81



## **Índice de tablas**

Tabla 6.1 Modelos de interpolación o extrapolación utilizados al transformar los valores..	37
Tabla 9.2.1 Pozos de control y zonas correspondientes para la realización 0.....	83
Tabla 9.2.2 Pozos de control y zonas correspondientes para la realización 1.....	83
Tabla 9.2.3 Pozos de control y zonas correspondientes para la realización 2.....	84
Tabla 9.2.4 Pozos de control y zonas correspondientes para la realización 3.....	84
Tabla 9.2.5 Pozos de control y zonas correspondientes para la realización 4.....	84
Tabla 9.2.6 Pozos de control y zonas correspondientes para la realización 5.....	85
Tabla 9.2.7 Pozos de control y zonas correspondientes para la realización 6.....	85
Tabla 9.2.8 Pozos de control y zonas correspondientes para la realización 7.....	85
Tabla 9.2.9 Pozos de control y zonas correspondientes para la realización 8.....	86
Tabla 9.2.10 Pozos de control y zonas correspondientes para la realización 9.....	86



## **1 Introducción**

La Fábrica de Uranio de Andújar (F.U.A.) ha sido monitorizada desde su clausura en el año de 1981 por medio de diversos modelos para la simulación del transporte de contaminantes en el agua subterránea, debido a la presencia de contaminantes, que por razones adversas fue lixiviando lentamente desde los diques de estériles.

Según el reporte de ENRESA del año 2004 se construyeron un total de 5 diques, de los cuales 3 fueron construidos desde 1960, y además, posteriormente se añadieron 2 diques más que comenzaron a construirse en el año de 1966. Cabe señalar que antes de ser vertido en los diques, los estériles eran neutralizados con objeto de precipitar el uranio, y existió muy poca movilidad de lixiviados durante la fase de funcionamiento de la fábrica, no obstante y debido a valores bajos de pH como consecuencia de la oxidación de piritas existentes en los estériles, algunas sustancias fueron movilizándose poco a poco (ENRESA, 2004).

La posibilidad de que el uranio pudiera desplazarse verticalmente desde los diques al acuífero subyacente fue confirmada por la presencia de uranio en las aguas subterráneas del acuífero aluvial en algunos de los pozos próximos a la FUA en la década de los 80 (ENRESA, 2004).

Posteriormente a estos acontecimientos y debido a la necesidad de evaluar las concentraciones de solutos en el entorno de la FUA, ENRESA creó en 1988 un modelo de transporte de solutos que proporcionara información acerca del estado actual del área de estudio, dicho modelo también fue utilizado para predecir posibles efectos debidos al Plan de Clausura (Samper et al., 1988) que se tenía previsto para el FUA. Estos estudios fueron centrados en analizar la migración desde la escombrera al acuífero, y al movimiento a través del mismo.

El modelo creado en 1988 fue contrastado con las mediciones obtenidas durante el periodo comprendido entre mayo de 1988 y agosto de 1991, y a partir de esta información, se revisó y posteriormente se calibró el modelo (Samper et al., 1991).

Habiéndose obtenido una calibración satisfactoria, se realizaron predicciones para el periodo comprendido entre 1993 y 2008, estas predicciones abarcaban un radio de alrededor de 1km de distancia de los límites de propiedad de la FUA (Samper et al., 1994). Este trabajo sirvió para crear una red de vigilancia con la cual monitorizar la evolución del área en estudio.

A partir de esta red, que fue monitorizada con una frecuencia trimestral, se realizó una segunda calibración del modelo en el año de 2003 con la cual se añadió la información obtenida a partir de la red desde el año 1994 hasta el 2003. (Enresa, 2004). Dicha calibración añadió una malla con mayor resolución en el área de estudio, a partir de la cual se podían considerar variaciones significativas encontradas en los trabajos realizados entre 1992 y 1994.

Además de esta mejora, también se utilizó el código CORE, una herramienta más potente para el análisis de las concentraciones logrando así un mejor ajuste de las mediciones, cabe mencionar también que en esta revisión se incluyó la posibilidad de tomar en cuenta un modelo con espesor variable en el tiempo (acuífero libre), así como utilizar modelos de transporte reactivo para modelizar la evolución geoquímica del término fuente.

El reporte de la actualización del modelo de flujo y transporte de Andújar, realizado para Enresa en agosto de 2004 concluye que se sigue una tendencia general similar a la existente en 1994, solo que con una disminución de la concentración en la mayoría de los puntos de la red.

Existen además diferencias entre los valores observados y las predicciones realizadas. Estas diferencias son más marcadas en las cercanías del término fuente, así como en algunas zonas al norte y sur del término fuente (Enresa 2004).

Se atribuyó como causantes de estas discrepancias en primer lugar a las concentraciones del término fuente, por lo que fueron revisadas las funciones de concentración de las vías de entrada de uranio al acuífero desde los diques. Debido a la gran cantidad de parámetros necesarios para representar la función utilizada en 1994, se decidió optar por una campana de Gauss definida mediante 2 parámetros de forma y tres puntos, los cuales son la concentración en el tiempo inicial a partir del cual aumentan las concentraciones

observadas, el valor máximo de la concentración y el valor de estabilización definido por el año en el que alcanza el valor asintótico de la concentración (Enresa 2004).

Se obtuvieron varias hipótesis a partir de los resultados obtenidos, entre las que se mencionan que existe la posibilidad de que el uranio quede retenido o precipite debido a procesos geoquímicos en el acuífero, la posible salida de uranio a través del río debido a vías no consideradas en el modelo, o la posibilidad de que el uranio quede atrapado en la zona no saturada del aluvial debido a procesos hidrodinámicos.

También se descartaron algunas hipótesis debido a que los resultados obtenidos al tomar en cuenta ciertos factores no aportaba una base clara de las diferencias tan marcadas entre las mediciones y el modelo. Entre estas hipótesis se encuentra por ejemplo la precipitación del uranio, la cual no guardaba una relación clara con las concentraciones de uranio.

Los resultados mostraron que existía una fuerte relación entre los procesos de dilución causados hidrológicamente, y los valores medidos en los pozos de observación, por lo que se ajustó el modelo para reproducir este efecto.

El hecho de que no existiese una razón clara para que se reprodujesen de manera congruente los valores llevó a la modificación del término fuente, empleando algunas hipótesis hidrogeoquímicas, de las cuales solamente se trataron a grandes rasgos algunas posibilidades debido a la gran exigencia de cálculo y la complejidad que se tenía.

En cuanto a la heterogeneidad espacial de la zona, el estudio introdujo cambios ligeros en la estructura del modelo de flujo, para representar adecuadamente los datos de las zonas próximas a la instalación; también se recalibraron los parámetros del modelo de flujo y transporte siendo similares a los obtenidos en el modelo de 1994.

Posterior a este reporte, en 2011, se realizó un análisis de incertidumbre del modelo FUA04 variando la conductividad hidráulica que se introduce en los ficheros de entrada del modelo. Para ello se llevó a cabo un análisis estocástico por medio del uso de técnicas de simulación de Monte Carlo (Gómez, 2011).

El análisis requirió generar un modelo de continuidad de las variables aleatorias mediante una función multigaussiana a fin de crear múltiples realizaciones del campo de

conductividades que estuviesen espacialmente correlacionadas, posteriormente el modelo del FUA04 era dotado del nuevo campo heterogéneo de conductividades.

Los resultados obtenidos por este análisis muestran que, en promedio, los niveles piezométricos no se ven alterados en gran medida al introducir heterogeneidades a pequeña escala, no obstante existe una gran discrepancia entre las observaciones y el modelo tanto al tomar en cuenta la heterogeneidad como cuando no es tomada en cuenta. El análisis señala que aun teniendo en cuenta la heterogeneidad del medio a pequeña escala, no se logra explicar en su totalidad la gran variación de concentraciones que se tiene en el medio, en parte por las pocas realizaciones que se llevaron a cabo, las cuales no logran ajustar adecuadamente las concentraciones.

Es a partir de estos resultados que se decide emplear los filtros de Kalman de conjuntos para, intentar reducir esa discrepancia entre los valores observados y obtenidos, ajustando primeramente la piezometría para después utilizar las herramientas creadas a fin de acoplarlas al problema de transporte en un futuro estudio.

## 1.1 Objetivos

El presente estudio plantea observar la capacidad que tienen los filtros de Kalman de conjuntos para asimilar información existente a fin de ajustar las predicciones del modelo del FUA. Simultáneamente preparar el modelo que utiliza el código CORE para reproducir la piezometría y posteriormente en trabajos subsecuentes al alcance de este estudio, intentar reproducir las concentraciones de contaminantes observadas en los distintos puntos de la red de vigilancia.

En resumen los objetivos del trabajo son:

- Aplicación de los filtros de Kalman de Conjuntos a un caso real complejo.
- Generar una serie de campos heterogéneos de conductividades que sean representativos de las zonas de materiales utilizadas en el anterior modelo del FUA, a fin de reproducir la heterogeneidad del medio a pequeña escala.
- Modificar el CORE para poder utilizar un mayor número de zonas de material y así tomar en cuenta la heterogeneidad del terreno a pequeña escala, que, aunque para resolver el problema de flujo no es tan significativo, sí lo es para resolver el problema de transporte de un contaminante reactivo.
- Hacer uso de la metodología de los filtros de Kalman de conjuntos para ajustar las conductividades hidráulicas utilizando conjuntos de campos de conductividades que, una vez ajustados por el filtro reproduzcan satisfactoriamente la piezometría de la zona.
- Finalmente obtenidos los campos ajustados medir la incertidumbre que introduce el uso de diferentes conductividades hidráulicas al modelo.

Entre los objetivos secundarios de este trabajo están:

- Crear subrutinas para mayor facilidad en el manejo de los datos.
- Crear o modificar según sea el caso los archivos de entrada para cada programa utilizado.
- Generar los formatos necesarios para introducir los resultados y observarlos visualmente en el programa SGEMS.



## 1.2 Organización del documento

Este trabajo trata de tomar en cuenta la heterogeneidad a pequeña escala para poder reproducir adecuadamente los valores observados en campo en el área circundante a la FUA. Para esto se han empleado filtros de Kalman de conjunto, una herramienta que ha probado ser robusta en el campo de la ingeniería del petróleo.

Por tanto se realiza una revisión del estado del arte de los filtros de Kalman de conjuntos, sus aplicaciones en diversos campos, los estudios realizados en hidrogeología más actuales que usan esta metodología, así como un ejemplo ilustrativo que intenta describir su funcionamiento de manera sencilla al lector de este documento, todo esto se verá más a fondo en el capítulo 2.

El capítulo 3 nos habla acerca del modelo del FUA, la extracción de datos realizada y sus diversas modificaciones.

El capítulo 4 hace hincapié al análisis geoestadístico realizado. Para ello se hizo uso de un modelo de continuidad espacial a fin de poder representar de manera adecuada la variación de valores que tienen el área de estudio en cuanto a conductividades se refiere.

Luego de haber obtenido el modelo de continuidad espacial, este es usado en el capítulo 5 para realizar una estimación del campo de conductividades hidráulicas mediante un krigeado ordinario. En el mismo capítulo se describe como se realiza el krigeado, así como los resultados del mismo.

El capítulo 6 muestra cómo, a partir del modelo de continuidad, se pueden recrear diversas realizaciones de campos equiprobables, con los cuales poder realizar el análisis de sensibilidad del modelo posteriormente. En este capítulo también se realiza un desarrollo del método de simulación que ha sido realizado con el programa SGEMS (Stanford Geostatistical Modeling Software); por último se observan algunas de las simulaciones creadas, así como los valores esperados del conjunto de simulaciones.

El capítulo 7 hace una revisión al CORE, código que resuelve las ecuaciones de flujo, transporte de solutos y transferencia de calor en contornos irregulares. Se detallan los aspectos teóricos del CORE, la resolución de las ecuaciones de flujo en acuífero confinado, acuífero libre, así como en la zona con saturación variable. También en el capítulo 7 se describe la metodología que utiliza CORE para aproximar la solución real por medio de la solución numérica del flujo en CORE. Esta se describe para acuífero confinado y libre, y además la metodología iterativa que utiliza CORE para resolver el problema numérico en la zona de transición. Finalmente se describen los datos de entrada al modelo, así como la respuesta del modelo antes de introducir el uso de los filtros de Kalman de conjuntos.

El capítulo 8 nos habla de los filtros de Kalman de conjuntos, como fueron aplicados al FUA, así como el programa que fue utilizado para ajustar las conductividades por medio de los archivos de salida provenientes de CORE.

El capítulo 9 nos habla de los resultados y conclusiones que se obtuvieron al utilizar los filtros de Kalman de conjuntos.

## 2 Estado del arte

### 2.1 El uso de los filtros de Kalman

El filtro de Kalman fue introducido en 1960 por Rudolph Kalman quien en su publicación (Kalman, 1960) hizo una gran contribución a la teoría de control estocástica, hoy conocida como el filtro de Kalman.

Es un método bayesiano secuencial de asimilación de datos, que en cada intervalo de asimilación combina una estimación previa con la medida observada para dar un estimado posterior, condicionado a que toda la información disponible en ese momento sea utilizada. Ambos el valor estimado previo y el posterior, como también la medición, contienen cierta incertidumbre. La solución posterior es generalmente considerada la solución óptima para el caso de procesos lineales con variables Gaussianas (Thulin, Skaug, Aanonsen, Nævdal, 2012).

El filtro fue originalmente diseñado para solucionar el problema lineal clásico de la estimación de mínimos cuadrados en el procesamiento de señales y la teoría de control. Posteriormente el filtro de Kalman extendido (EKF) fue introducido para estimar el estado de un sistema en un proceso no lineal.

Básicamente el filtro de Kalman extendido lineariza la función para acercarse a la media estimada actual y calcula la matriz Jacobiana del estado de transición y de la función de observación en cada intervalo de cálculo. Cabe enfatizar que aun las aplicaciones tempranas del filtro de Kalman extendido introdujeron la estimación de parámetros del modelo desconocidos (Cox,1964;Kopp y Orford, 1963) añadiéndolos a la estimación del estado de la variable dinámica. Esto es conocido como la *estimación dual*.

Sin embargo el EKF no es apropiado para modelos muy grandes o con no linealidades importantes. El filtro de Kalman de conjunto (EnKF) fue introducido como una alternativa por Evensen (1994). El EnKF usa un conjunto de realizaciones para representar los estadísticos del estimador actual, y todas las realizaciones son propagadas hacia adelante y analizadas de acuerdo a las ecuaciones del filtro de Kalman.

Este método ha probado tener aproximaciones prometedoras, además de ser eficiente y robusto aun cuando se utiliza en modelos no lineales muy grandes y se ha vuelto muy popular entre aplicaciones como los modelos oceánicos (Evensen 2006), modelos de reservorios (Haugen et al., 2008; Skjervheim et al., 2007; Nævdal et al. 2005) y modelos atmosféricos (Sun et al. 2009, Kepert, Sun y Steinle, 2003).

## 2.2 Aplicación de los filtros de Kalman de conjuntos en hidrogeología

Como se comentó, el uso de los filtros de Kalman de conjuntos ha ido en aumento en las ciencias atmosféricas, en el estudio de la interacción suelo-atmosfera, en la ingeniería del petróleo y en hidrogeología. Mientras que en las ciencias atmosféricas y en modelos que estudian la interacción suelo-atmosfera son usados generalmente solo para modelar el estado del sistema y actualizarlo, en la ingeniería del petróleo y en la ingeniería se han utilizado para actualizar tanto las variables de estado como los parámetros del sistema (Naevdal et al., 2005). Los filtros de Kalman de conjuntos se han aplicado satisfactoriamente para asimilar la altura piezométrica y así mejorar la capacidad de predicción de modelos de manera dinámica. (Chen y Zhang, 2006; Hendricks Franssen y Kinzelbach, 2008; Li et al., 2011; Zhou et al., 2011).

Hendricks Franssen et al (2011) en su artículo publicado en Water Resources Reserch (WRR), utilizó información y modelos en tiempo real de la ciudad de Zurich (Suiza) para simular el flujo dando como resultado que la asimilación diaria de datos de alturas piezométricas por medio de filtros de Kalman de conjuntos logra una mejor caracterización de las alturas piezométricas que la obtenida por medio de calibración inversa con datos históricos pero sin asimilar nueva información.

Li et al. (2011) mostró que los filtros de Kalman de Conjuntos utilizando variables normalizadas NS-EnKF (Zhou et al. 2011) pueden caracterizar distribuciones hidráulicas no multigaussianas aun y cuando se tiene un modelo de distribución erróneo como distribución de partida. Nowak y Hendricks Franssen (2012) utilizan los filtros de Kalman de Conjuntos para estimar parámetros a partir de una tomografía hidráulica de un campo log-multigaussiano de conductividades en 3D. Combinando los filtros de Kalman de conjuntos con la anamorfosis gaussiana (GA). Gráficamente, la anamorfosis consiste en deformar el histograma de los datos en un histograma no gaussiano, de modo que la variable transformada, denotada  $Y(x)$ , tenga una distribución gaussiana estándar (media 0 y varianza 1). Los resultados obtenidos logran un mejor acercamiento en la predicción de flujo y transporte.



## 2.3 Filtros de Kalman de conjuntos

Para ilustrar el funcionamiento de los Filtros de Kalman se ha creado un ejemplo ilustrativo:

Supóngase que se ha comprado un horno el cual ha sido mal calibrado y tiene un termostato con un cierto error de medición desconocido. Se desea conocer el tiempo ideal que logra cocinar un pastel  $x$ , a la temperatura máxima. Para esto se van a cocinar una serie de pasteles de prueba a fin de conocer el tiempo ideal de cocción. Por suerte se cuenta con la ayuda de la abuela, la cual tiene más experiencia que tú en la elaboración de pasteles.

Se comienza con el primer pastel al cual tú le asignas un tiempo de cocción de  $z_1$  minutos, el cual tiene una cierta incertidumbre  $\sigma_1$ . Al terminar los  $z_1$  minutos y sacar el pastel del horno, tu abuela te dice que el tiempo de cocción debiera de ser  $z_2$  minutos. Tu abuela, aunque más experimentada que tú, también tiene cierta incertidumbre en el tiempo de cocción, por lo que cuenta con una incertidumbre  $\sigma_2$  menor que  $\sigma_1$ .

En términos matemáticos el problema se traduce a lo siguiente:

Se desea estimar la variable  $x^* = x|z_1$  asumiendo que  $x^* \sim N(z_1, \sigma_1)$ . Esto es llamado estimación previa y provee una información a-priori del tiempo exacto. La función de la estimación dada por la abuela es dada por  $z_2|x^* \sim N(x^*, \sigma_2)$ , esto crea el efecto de añadir nueva información, a la cual se le da cierto peso.

Por lo tanto nos interesan los momentos  $\mu$  y  $\sigma$  de  $f_x|z_1, z_2$ .

El teorema bayesiano señala que la función  $f_{x^*|z_2}$ . Es proporcional a  $f_{z_2|x^*} f_{x^*}$ .

Se puede escribir lo anterior de la siguiente manera:

$$\frac{(x^* - \mu)^2}{2\sigma^2} = \frac{(z_2 - x^*)^2}{2\sigma_2^2} + \frac{(x^* - z_1)^2}{2\sigma_1^2} + C$$

(2.3.1)

Donde  $C$  es una constante. Las siguientes formulas siguen la comparación de los coeficientes en  $(x^*)^2$  y  $x^*$  respectivamente:

$$\mu = \frac{\sigma_2^2}{(\sigma_1^2 + \sigma_2^2)z_1} + \frac{\sigma_1^2}{(\sigma_1^2 + \sigma_2^2)z_2} \quad (2.3.2)$$

$$\frac{1}{\sigma^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \quad (2.3.3)$$

El tiempo óptimo  $\mu$  es un promedio ponderado entre la cantidad de tiempo dadas por ti y por tu abuela. Como es de esperar, la estimación de tu abuela tiene mayor peso que tu estimación. Por otra parte la varianza  $\sigma^2$  es proporcional a la media armónica de las varianzas en las observaciones. En particular la nueva desviación estándar es menor que las dos anteriores. Por tanto la varianza decrece en cada intervalo.

Luego:

$x^f = E[x|z_1] = z_1$  El estado promedio del sistema, dada la información existente. Esto corresponde al tiempo que Tú le das de cocción al pastel.

$x^a = E[x|z_1, z_2] = \mu$  El estado actualizado del proceso, dada la nueva información proporcionada por la abuela.

$(\sigma^f)^2 = var(x|z_1) = \sigma_1^2$  Es la varianza de los valores que asignas.

$(\sigma^a)^2 = var(x|z_1, z_2)$  Es la varianza actualizada.

La cantidad de tiempo actualizada es una combinación lineal de la media del estado anterior del sistema y la nueva medida:

$$x^a = x^f + K[z_2 - x^f] \quad (2.3.4)$$

Donde:

$$K = \frac{(\sigma^f)^2}{((\sigma^f)^2 + (\sigma^a)^2)} \quad (2.3.5)$$

Esto es comúnmente conocido como la ganancia de Kalman. La incertidumbre puede ser escrita como:

$$(\sigma^a)^2 = (\sigma^f)^2 - K(\sigma^f)^2 = (1 - K)(\sigma^f)^2 \quad (2.3.6)$$

La abuela “sabe” que la ecuación que controla la cocción del pastel es:

$$\frac{dx}{dT} = C_a + w \quad (2.3.7)$$

Donde  $\frac{dx}{dT}$  es diferencial del tiempo con respecto a la temperatura,  $C_a$  es la cantidad de masa de pastel, y  $w$  es un ruido blanco debido a la incertidumbre del termostato.

El papel principal del filtro es mejorar el tiempo que has asignado a la cocción mediante la estimación promedio de las observaciones de tu abuela.

Utilizando la ecuación 2.3.7 y reescribiéndola mediante el método de Euler la ecuación queda como sigue:

$$x^T = x_0 + \Delta T C_a + \Delta T w \quad (2.3.8)$$

Donde:

$$x_0 \sim N(z_0, \sigma_0)$$

$$w_0 \sim N(0, \sigma_w)$$

La ecuación 2.3.8 se va modificando a través del proceso, al igual que los momentos:

$$x^f = z_0 + \Delta T C_a \quad (2.3.9)$$

$$(\sigma^f)^2 = \sigma_0^2 + \Delta T \sigma_w^2 \quad (2.3.10)$$

Nótese que la varianza  $\sigma^f$  se incrementa. Sin embargo la información obtenida se vuelve la información a priori.

Si asumimos que  $C_a \sim N(0, \sigma)$ , la actualización de los tiempos será hecha de la siguiente manera:

$$x^a = x^f + K_1 [z_1 - x^f] \quad (2.3.11)$$

$$(\sigma^a)^2 = (1 - K_1)(\sigma^f)^2 \quad (2.3.12)$$

Donde:

$$K_1 = \frac{(\sigma^f)^2}{((\sigma^f)^2 + \sigma^2)} \quad (2.3.13)$$

Siendo  $K_1$  la ganancia de Kalman para la primer iteración. El proceso se reinicia con los momentos  $x^a$  y  $\sigma^a$  del estado inicial  $x_0$ .

### **3 Modelo del FUA 2004**

El modelo del FUA es un modelo horizontal bidimensional de flujo y transporte que simula un acuífero libre isotrópico con un régimen de flujo transitorio.

La fuente principal de recarga es resultado de la infiltración de la precipitación y del exceso del riego, esta recarga se discretiza en 21 zonas de recarga. Para cada zona de recarga se han empleado diferentes combinaciones de métodos hidrometeorológicos de calibración.

Para estimar la recarga se ha hecho un balance hídrico por medio de VISUAL-BALAN, el cual es un modelo agregado para el cálculo del balance hídrico. Se utilizaron datos hidrometeorológicos como base para la calibración posterior mediante la aplicación de INVERSE-CORE, una subrutina que se encuentra dentro de CORE 2.0.

La recarga luego se define en cada uno de los elementos triangulares con los que cuenta el modelo mediante la asignación de la función de recarga que le concierne a la zona donde se encuentra el elemento.

El modelo ha sido calibrado para los años 1977-2003 usando la aproximación mediante la resolución del problema inverso. Como ya se dijo el modelo utiliza la subrutina INVERSE-CORE la cual minimiza la función objetivo que depende de las medidas simuladas de piezometría y/o concentraciones adaptando los parámetros de flujo y transporte.



### 3.1 Gestión de datos del GIS

Se comenzó a trabajar con los datos del modelo FUA haciendo uso de los archivos de entrada que contienen la información necesaria para que el modelo estime las concentraciones y niveles piezométricos en los distintos nodos del modelo. Además se utilizó un proyecto *GIS* el cual contiene información útil del área de estudio, dicho proyecto fue creado en 1994, para *ARCVIEW3.2*, así que lo primero fue, a partir de la información existente, crear un proyecto semejante en *ArcGIS*, el cual se fue actualizando con la información más reciente y además se le fue añadiendo información paulatinamente a medida que se iba obteniendo.

Como se comentó en la introducción del documento se trabajó con el modelo del FUA 2004 que utiliza el CORE (Código para la simulación numérica de procesos de flujo de agua, transferencia de calor y solutos REactivos) para obtener las concentraciones y alturas piezométricas en cada uno de los nodos en los que esta discretizada el área de estudio.

El modelo consta de 12254 elementos triangulares y 6218 nodos, derivados de la subdivisión de la malla utilizada en el modelo de 1994; esto con el fin de tener un mayor número de elementos en las zonas más cercanas a la fábrica de Uranio del FUA, así como en algunos otros lugares de interés, como son los lugares en donde las concentraciones tenían gran variabilidad entre las calculadas por el modelo y las observadas durante las campañas de muestreo.

La figura 3.1 nos muestra la malla creada a partir de la información disponible en los archivos de entrada de CORE, la cual fue añadida al SIG del FUA. Posteriormente de obtener la malla se creó una cuadrícula de 1344 filas por 660 columnas, con tamaño de celdas de 5x5 metros que abarca toda el área de estudio, iniciando en la coordenada UTM 30 401920 Este, 4208450 Norte y finalizando en la en la coordenada UTM 30 408640 Este, 4211750 Norte. Esto da como resultado una cuadrícula de 887040 celdas. La Figura 3.2 nos muestra como la cuadrícula abarca toda el área de estudio.

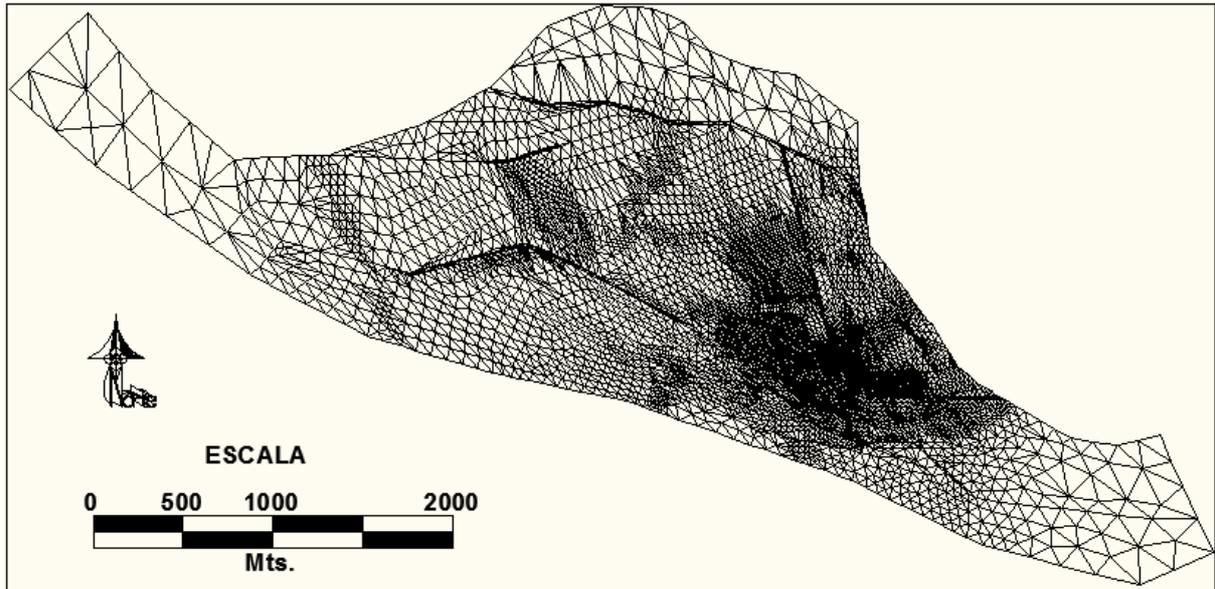


Figura 3.1 Malla de elementos triangulares creada a partir de la información de entrada utilizada por el código CORE.

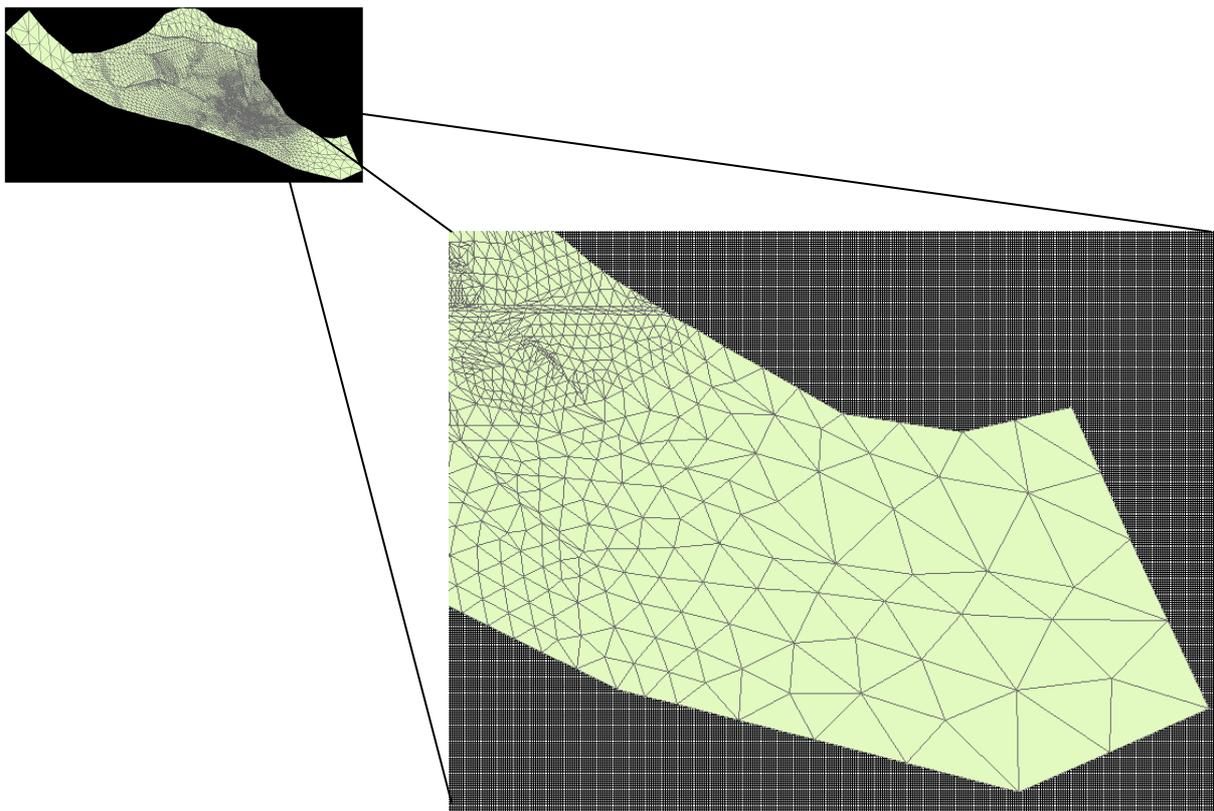


Figura 3.2 Cuadrícula de 887040 elementos que envuelve el área de estudio.

Una vez creada la cuadrícula, se obtuvieron los valores de conductividades hidráulicas de los puntos de muestreo de la red de Vigilancia. A partir de la información disponible en el SIG.

En total se tienen 154 puntos con información de conductividades que fueron extraídos del SIG para utilizarlos como información de partida para el programa SGEMS.

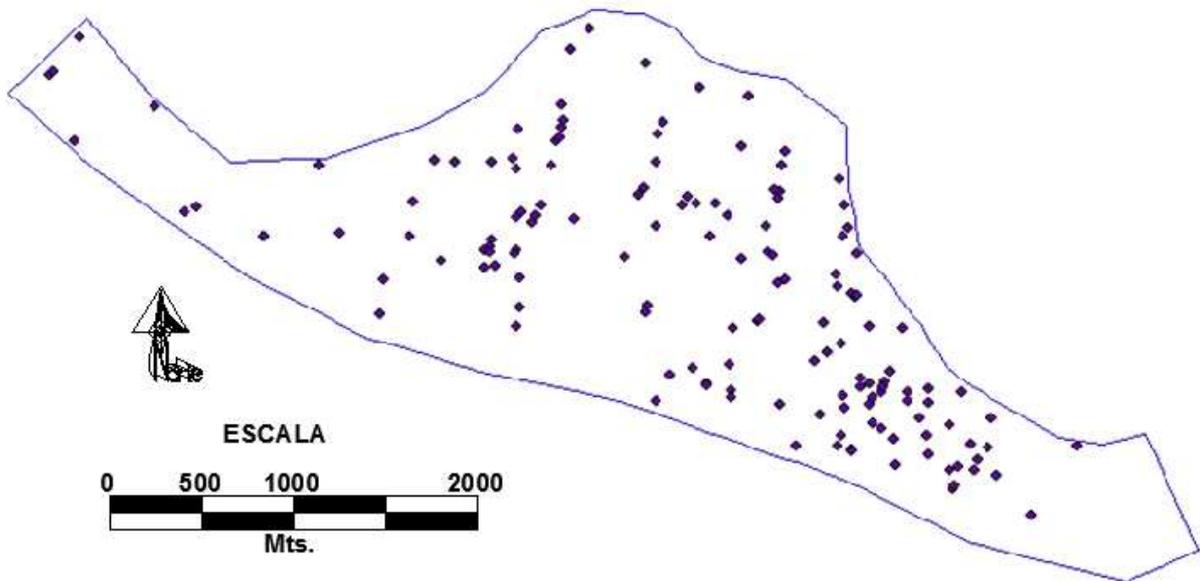


Figura 3.3 Mapa de localización de los 154 datos de Conductividades.



## 4 Análisis geoestadístico

### 4.1 Modelo de continuidad espacial

Para poder realizar un análisis Geoestadístico es necesario contar con información útil que nos ayude a comprender la variabilidad espacial de la propiedad que nos interesa conocer, ya sea obteniendo valores medidos de esta propiedad, o bien, a través de otra propiedad que esté correlacionada con la primera, en este caso se creó el campo de conductividades hidráulicas directamente del modelo FUA04 mediante el programa SGEMS introduciendo las conductividades previamente obtenidas de los puntos de observación de la red de vigilancia. La figura 4.1 nos muestra la distribución espacial de dichos puntos, así como el valor de la conductividad hidráulica para cada uno de los 154 puntos en m/d.

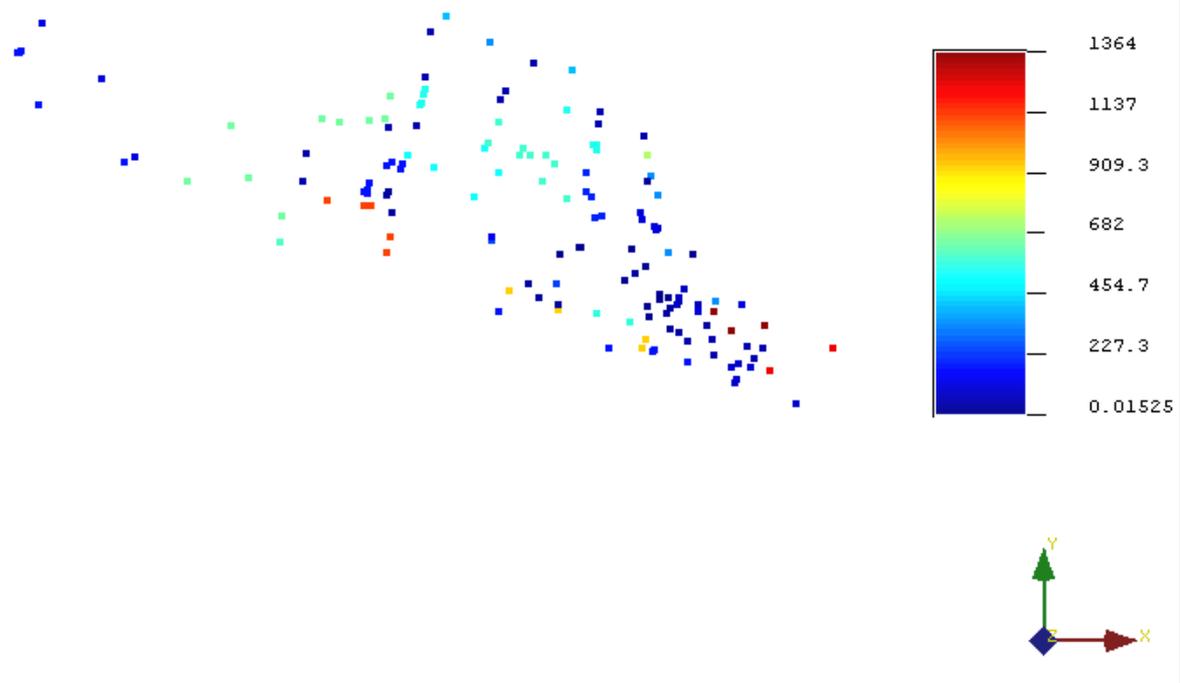


Figura 4.1 Distribución espacial de los valores de conductividad hidráulica.

Se muestra la distribución espacial de los valores de conductividad hidráulica. Cabe señalar que se tiene un aumento en la conductividad hacia el oeste de la zona de la FUA. En la figura se logra apreciar que se tiene una cierta continuidad cercana a los 120° azimuth, por lo que se generó un variograma omnidireccional para apreciar mejor el alcance que se tiene en los

datos, además se generaron diversos variogramas direccionales variando el ángulo de búsqueda en 15 grados cada vez. En cuanto a los datos estadísticos de la muestra se realizó un histograma de frecuencias para conocer los valores de la media y la desviación típica de los datos, estos se observan en la figura 4.2.

Primeramente llama la atención en el histograma el que existen una gran cantidad de valores altos de la muestra, al observar la función de distribución acumulada se aprecia que de manera general se sesga hacia valores menores de 200 m/d, se observa que los datos no están distribuidos de manera uniforme en el histograma, posiblemente debido a la gran cantidad de pozos de observación cercanos a las instalaciones de la fábrica. La desviación típica de la muestra es importante, mientras que la media llega a ser de 329.18 m/d.

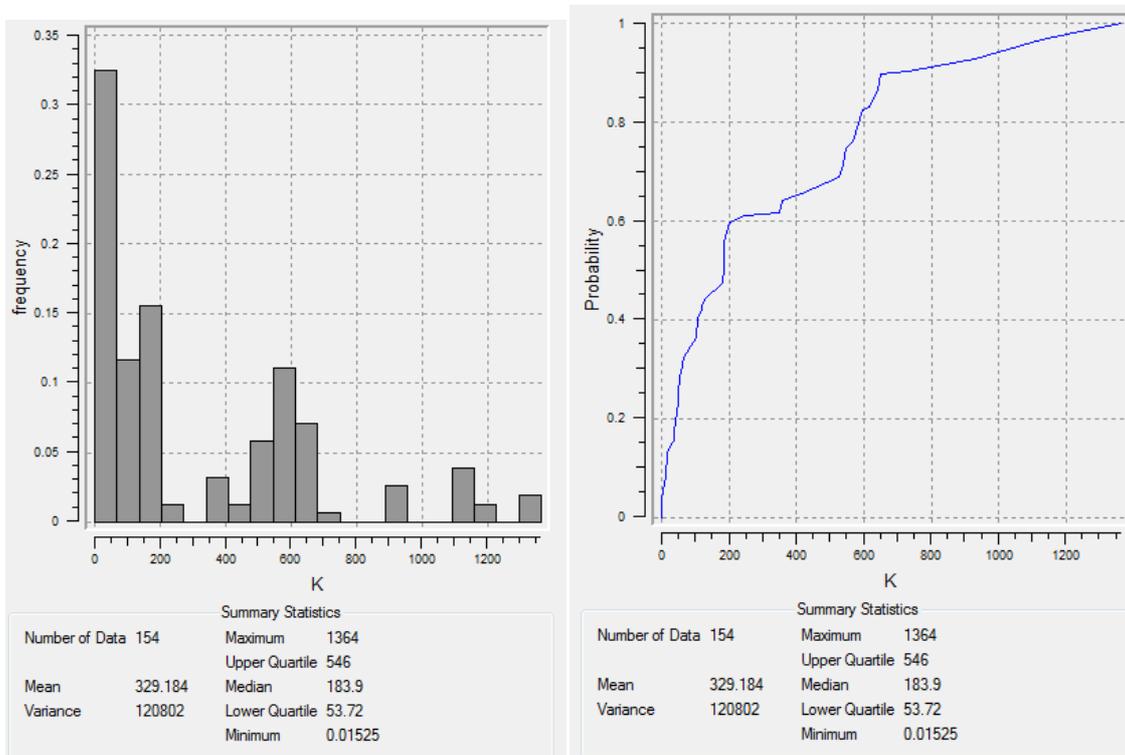


Figura 4.2 Histograma y curva de probabilidad acumulada.

Se observa una gran variabilidad en los datos, que contribuye a afirmar la hipótesis de la heterogeneidad del medio mencionada en los trabajos previos. Volviendo un poco atrás a los variogramas mencionados el variograma omnidireccional muestra que se tiene un alcance cercano a los 3000 metros de distancia, a partir del cual comenzamos a buscar la dirección de máxima continuidad. La figura 4.3 nos muestra el variograma omnidireccional.

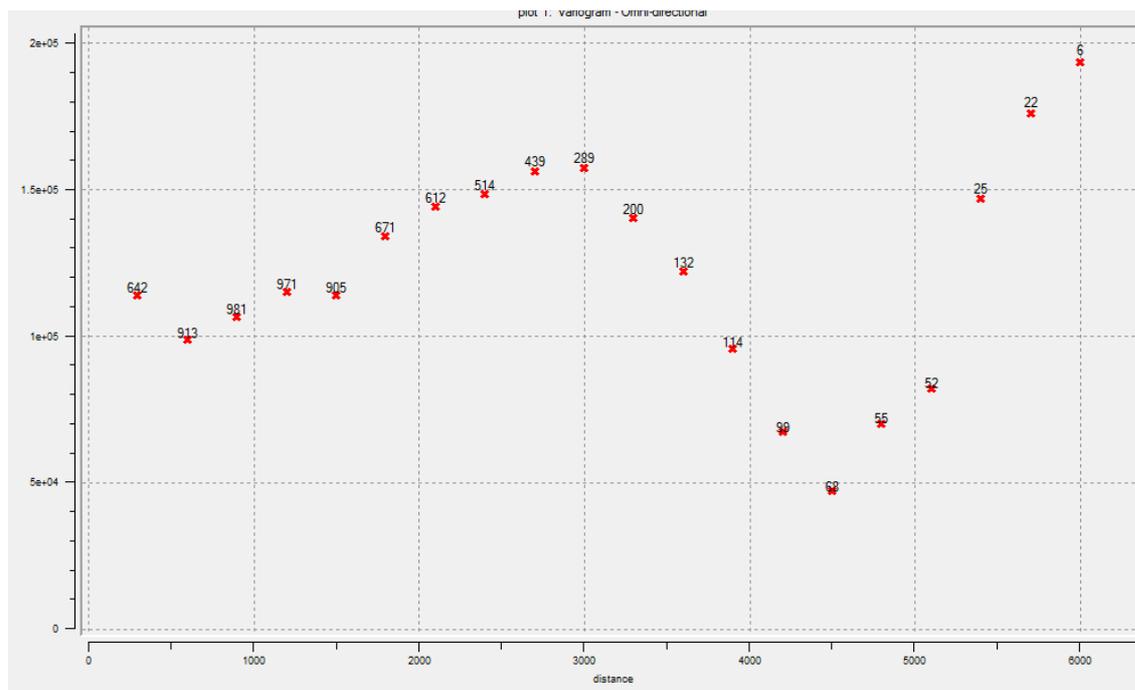


Figura 4.3 Variograma omnidireccional

Los variogramas direccionales tienen una gran variabilidad espacial, por lo que se buscó los ejes de mayor y menor continuidad, observándose que entre los 120 y los 135 grados se tenía la menor variabilidad espacial, en cuanto al eje de menor continuidad se observó que estaba entre los 30 y los 45 grados. Las figuras 4.4 y 4.5 nos muestran la gran variabilidad entre el variograma en la dirección de 30 grados y el de la dirección de 127 grados.

El variograma de 30 grados parece más un efecto pepita, quizás porque la gran mayoría de datos están distribuidos en una banda con una orientación de 120 grados azimut, sin embargo, se logra observar en la figura 4.5 que los datos tienen una mayor correlación aparente en el eje de los 120 grados.

El variograma de 120 grados además tiene una disminución de la variabilidad pasando los 3000 metros para volver a ascender nuevamente, Pyrcz y Deutsch detallan el fenómeno en su publicación haciendo hincapié a cierto ciclamiento en la variabilidad espacial de los datos que le da esa forma característica, conocido como efecto hueco (Pyrcz, Deutsch, 2003).

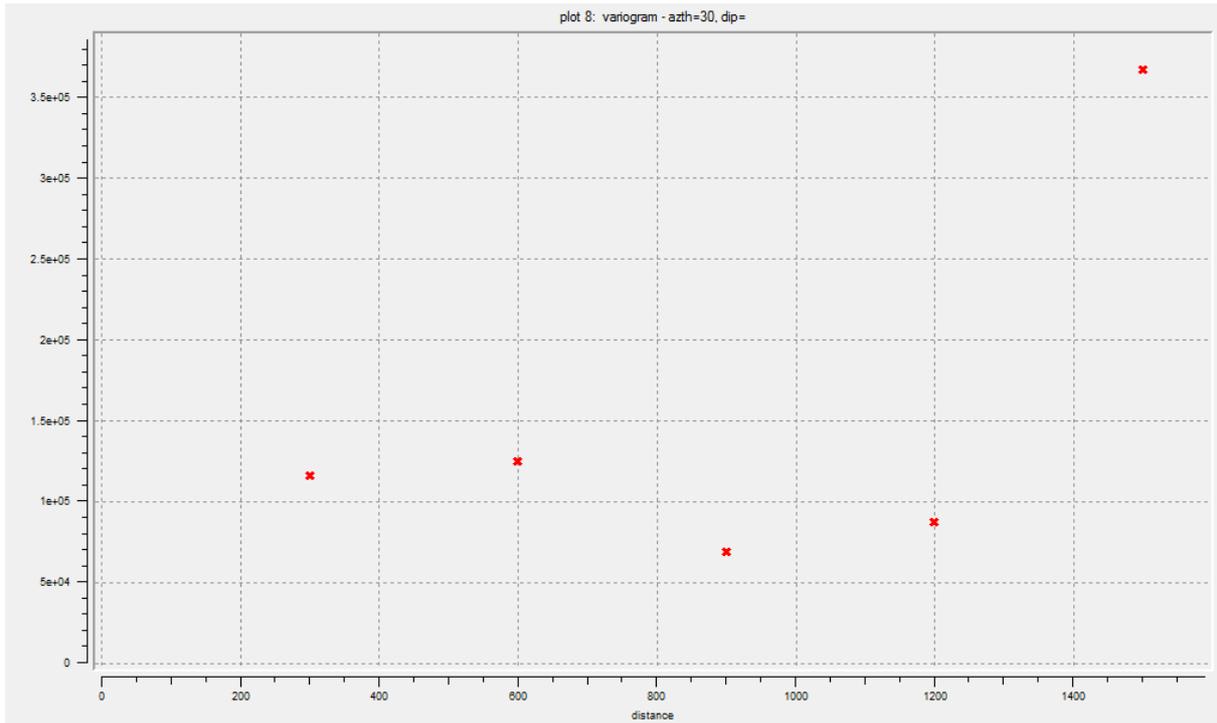


Figura 4.4 Variograma unidireccional en la dirección de los 30 grados azimut.

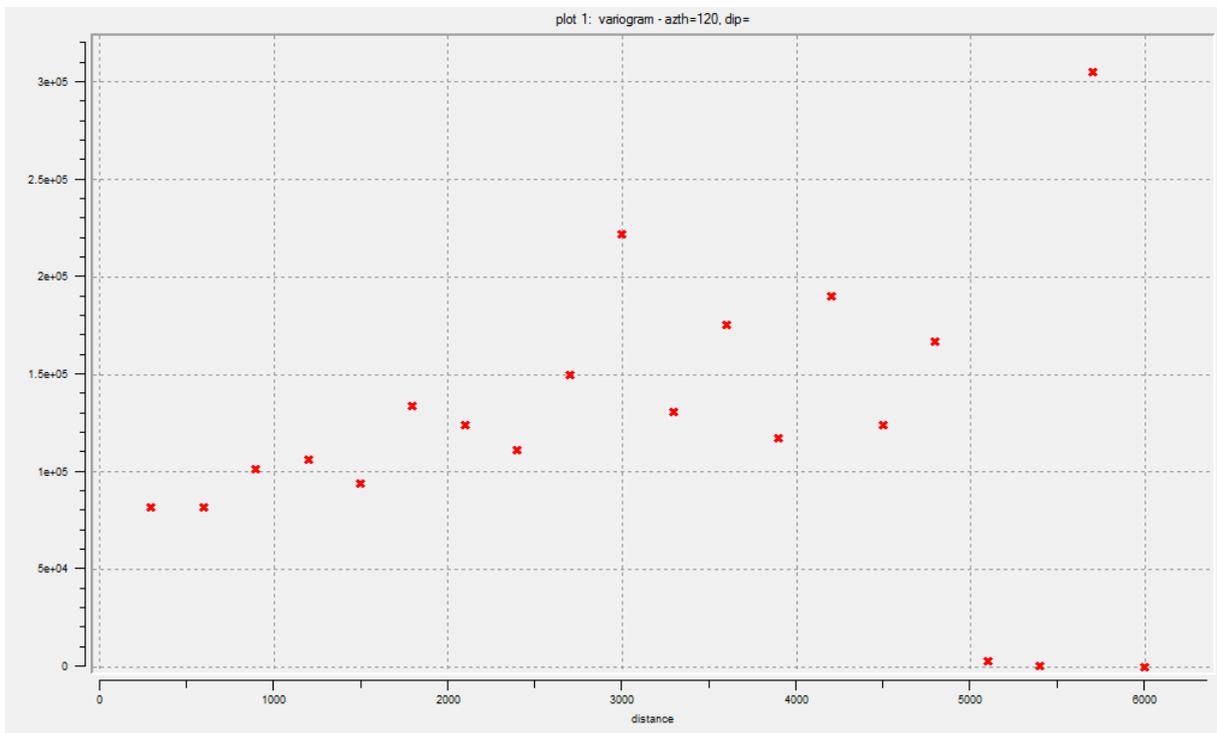


Figura 4.5 Variograma unidireccional en la dirección de los 120 grados azimut.

Tras ajustar un poco en busca de los ejes de máxima y mínima continuidad se determinó que el eje de máxima continuidad estaba en los 127° azimut, mientras que el de mínima continuidad era perpendicular a este último, tal y como se muestra en la figura 4.6.

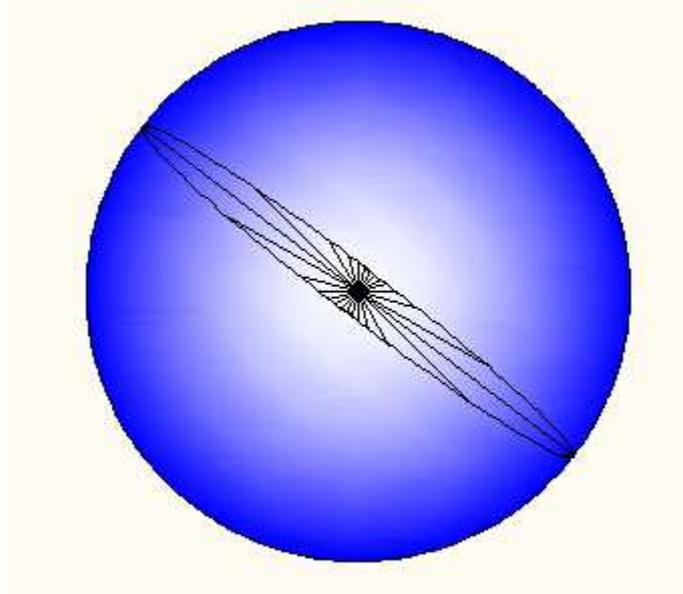


Figura 4.6 Ejes de máxima y mínima continuidad de los datos de conductividad hidráulica.

Ya conocidos los ejes principales se definió un modelo de continuidad para los datos, el cual está definido por un efecto pepita de 10000 unidades y una contribución de 140000 unidades añadidas con un modelo de variograma esférico:

$$\gamma(h) = \left[ \frac{3 \|h\|}{2 a} - \frac{1}{2} \left( \frac{\|h\|}{a} \right)^3 \right] \quad (4.1)$$

El modelo resultante es el siguiente:

$$\gamma(h) = 10000 + 140000\gamma^{(1)}(h) \quad (4.2)$$

Donde el primer termino corresponde al efecto pepita puro y el segundo termino es el aporte proveniente del variograma esférico con un alcance de 100000m y un rango de 5670m en el eje de máxima continuidad  $\alpha = 125^\circ$  y un rango de 2940m en el eje de mínima continuidad  $\beta = 35^\circ$ .



## 5 Estimación

### 5.1 Krigeado ordinario

La estimación por medio de krigeado ordinario se elabora siguiendo las bases del algoritmo básico generalizado para la regresión lineal utilizada en el krigeado simple (C. V. Deutsch, A. G. Journel, 1998 ):

$$[Z_{SK}^*(u) - m(u)] = \sum_{\alpha=1}^n \lambda_{\alpha}(u)[Z(u_{\alpha}) - m(u_{\alpha})] \quad (5.1)$$

Donde  $Z(u)$  es el modelo de la variable aleatoria en la locación  $u$ ,  $u_{\alpha}$  son las  $n$  locaciones de cada uno de los datos,  $m(u) = E\{Z(u)\}$  es el valor esperado dependiente de la locación de la variable aleatoria  $Z(u)$  y por último  $Z_{SK}^*(u)$  es el estimador llamado comúnmente estimador del krigeado simple (SK).

El conjunto de pesos  $\lambda_{\alpha}(u)$  es obtenido del sistema de ecuaciones que tienen la siguiente forma:

$$\sum_{\beta=1}^n \lambda_{\beta}(u)C(u_{\beta}, u_{\alpha}) = C(u, u_{\alpha}), \alpha = 1, \dots, n \quad (5.2)$$

El krigeado simple requiere conocer a priori el número de  $n + 1$  medias  $m(u)$ ,  $m(u_{\alpha})$ ,  $\alpha = 1, \dots, n$ , y la matriz de covarianzas de  $(n + 1)$  por  $(n + 1)$  elementos  $[C(u_{\alpha}, u_{\beta}), \alpha, \beta = 1, \dots, n]$  con  $u = u_0$ , lo que requiere asumir que la función aleatoria  $Z(u)$  es estacionaria.

Si esto sucede entonces la función  $Z(u)$  tiene media  $m$  constante y la covarianza puede ser calculada de la siguiente manera:

$$C(h) = C(u, u + h), \forall u \quad (5.3)$$

Reduciendo así el estimador del krigeado simple a su versión estacionaria:

$$[Z_{SK}^*(u) - m(u)] = \sum_{\alpha=1}^n \lambda_{\alpha}(u)[Z(u_{\alpha}) - m(u_{\alpha})] \quad (5.4)$$

Donde el sistema de krigeado simple está representado por:

$$\sum_{\beta=1}^n \lambda_{\beta}(u)C(u_{\beta} - u_{\alpha}) = C(u - u_{\alpha}), \alpha = 1, \dots, n \quad (5.5)$$

El krigeado ordinario por su parte filtra la media del estimador del krigeado simple forzando que los pesos del krigeado sumen la unidad, es decir:

$$Z_{OK}^*(u) = \sum_{\alpha=1}^n \lambda_{\alpha}^{(OK)}(u)Z(u_{\alpha}) \quad (5.6)$$

Donde:

$\lambda_{\alpha}^{(OK)}(u)$  es la serie de pesos para cada locación ( $u$ )

Convirtiéndose el sistema estacionario del krigeado simple en:

$$\sum_{\beta=1}^n \lambda_{\beta}^{(OK)}(u)C(u_{\beta} - u_{\alpha}) + \mu(u) = C(u - u_{\alpha}), \quad \alpha = 1, \dots, n \quad (5.7)$$

Sujeta a la restricción:

$$\sum_{\beta=1}^n \lambda_{\beta}^{(OK)}(u) = 1 \quad (5.8)$$

Donde:

$\mu(u)$  es el parámetro de Lagrange asociado con la restricción de que los pesos sumen la unidad.

Se utilizó el Programa SGEMS (Stanford Geostatistical Modeling Software), un programa de tipo GUI (interfaz gráfica de usuario) creado en C++ y el cual tiene como base la programación genérica, esta forma de programación que a partir de una serie de algoritmos generales, realiza una tarea específica, fue adoptada para la creación de la G<sub>3</sub>TL (Geostatistical Template Library), de la cual se deriva posteriormente SGEMS al introducir los diversos algoritmos de estimación simulación de la librería a un entorno amigable para el usuario.

Se realizó un krigado ordinario en toda la zona del modelo del FUA04 el cual arroja valores muy parecidos a los datos, suavizando claro está debido al método mismo del Krigado ordinario. Para obtener el mapa se utilizó un vecindario de búsqueda mediante una elipse de 3000m en su semieje mayor y 3000m en el semieje menor. Se condicionó a que existieran como mínimo 1 dato y como máximo 40 datos. La figura 5.1 muestra el área de estudio obtenida luego de aplicar el método.

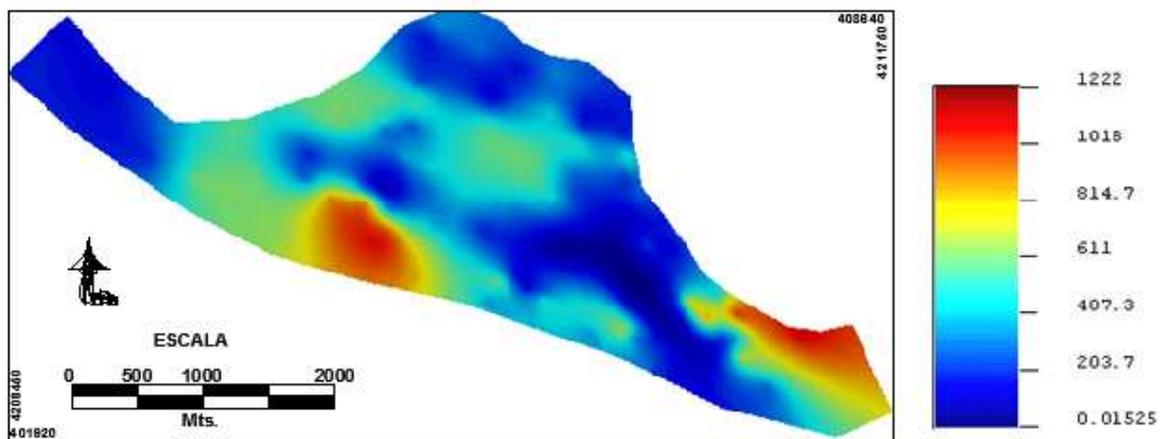


Figura 5.1 Mapa de conductividades obtenido mediante el Krigado Ordinario.

La figura mostrada anteriormente revela que las zonas cercanas a la fábrica cuentan con una muy baja conductividad hidráulica, no obstante el método de Krigado nos ofrece un panorama poco realista ya que no toma en cuenta la gran heterogeneidad que puede llegar a presentarse a menor escala, por lo que se procedió a tomar en cuenta la aleatoriedad de la zona por medio de métodos estocásticos.

Por otra parte la varianza del campo Krigado está entre 0 y 94500, aunque el valor esperado de la varianza es de 25135. Esta se observa suavizada en la figura 5.2. Se observan claramente los puntos difuminados característicos del Krigado, por otra parte, la zona sur este del mapa tiene un incremento en la varianza debido a que existen datos con conductividades muy altas a poca distancia de datos con variabilidades extremadamente bajas.

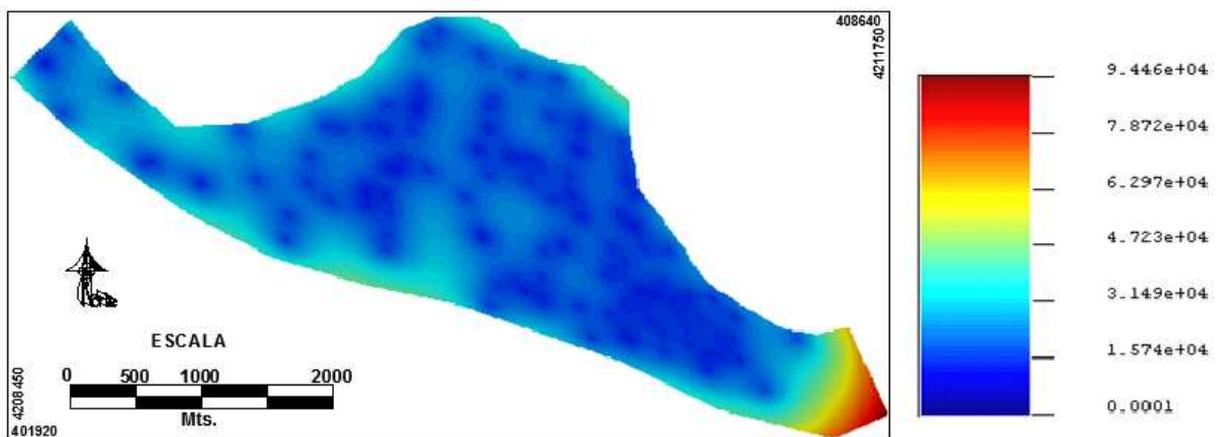


Figura 5.2 Mapa de la varianza en la conductividad hidráulica de la zona de estudio

## 6 Simulación

### 6.1 Simulación gaussiana secuencial

Una simulación estocástica es el proceso de construir modelos alternativos de distribución espacial igualmente probables de una variable con una resolución que logre captar la heterogeneidad del medio en la escala que pretendamos trabajar.

Para obtener un campo más heterogéneo es necesario considerar el gran número de distribuciones posibles de las variables a estudiar en el medio, para ello hay que conocer los principios de las simulaciones estocásticas.

Para empezar consideramos una distribución sobre el campo de estudio  $A$  de un atributo (1 en nuestro caso, aunque podrían ser mas) el cual llamaremos  $z(\mathbf{u})$ ,  $z$  tiene un cierto valor para cada ubicación  $\mathbf{u}$  perteneciente al campo de estudio es decir  $\mathbf{u} \in A$ .

Cada una de las distribuciones será denotada con el superíndice  $l$  por lo que se tendrá:

$$\{z^{(l)}(\mathbf{u}), \mathbf{u} \in A\} \tag{6.1}$$

Si el resultado de las simulaciones tiene los mismos valores que los datos en las coordenadas de los mismos, entonces se dice que es una simulación condicional y se denota como:

$$Z^{(l)}(\mathbf{u}_\alpha) = z(\mathbf{u}_\alpha), \quad \forall l \tag{6.2}$$

Dentro de la gran cantidad de algoritmos usados para reproducir la heterogeneidad, los algoritmos gaussianos son usados comúnmente debido a la manejabilidad y facilidad de los mismos, ya que tienen la propiedad de que al sumar varias distribuciones espaciales de probabilidad similares suelen asemejarse a un modelo gaussiano. Esto se conoce como el teorema del límite central. Este teorema es llamado “central” debido a Polyá (1920) quien describió el rol que cumple este teorema en la teoría de probabilidades.

La importancia del teorema radica en que desvela las razones por las cuales, en muchos campos de aplicación, se encuentran en todo momento distribuciones normales o casi normales. (Blaiotta, Delieutraz. 2004).

Para poder utilizar el teorema se debe tener en cuenta que:

- El teorema del límite central garantiza una distribución normal cuando  $n$  es suficientemente grande.
- Existen diferentes versiones del teorema, en función de las condiciones utilizadas para asegurar la convergencia.
- La aproximación entre las distribuciones es, en general, mayor en el centro de las mismas que en sus extremos o colas, motivo por el cual es llamado teorema del “límite Central”.

De manera general puede ser escrito como:

$$Z(\mathbf{u}) = \sum_{k=1}^K Y_k(\mathbf{u}) \approx \text{Gaussiana}$$

(6.3)

Estando sujeta mayormente al hecho de que las variables  $Y_k(\mathbf{u})$  que se suman sean independientes, en primer término e idénticamente distribuidas, en segundo término; sin embargo el hecho de no congeniar al 100% con las características rígidamente dichas de los modelos gaussianos no se opone al hecho de que los modelos gaussianos multivariados son en sí mismos extremadamente congeniantes. Si se aplican las condiciones del teorema del límite central estrictamente entonces se asume que los modelos gaussianos multivariados no son aplicables en las ciencias de la tierra donde los diversos procesos geológicos y biológicos tienen gran correlación entre ellos; en lugar de apegarnos a este criterio es mejor definir la función gaussiana multivariada a través de sus propiedades características, es decir que una función aleatoria  $Y(\mathbf{u}) = \{ Y(\mathbf{u}), \mathbf{u} \in A \}$  es multivariada normal si y solo si:

- Todos los subconjuntos de la función aleatoria, por ejemplo  $\{Y(\mathbf{u}), \mathbf{u} \in B \subset A\}$  son también normales multivariados.
- Tener covarianza cero (o correlación) hace que las variables sean independientes:

Si  $Cov \{Y(\mathbf{u}), Y(\mathbf{u}')\} = 0$  las dos variables aleatorias  $Y(\mathbf{u})$  y  $Y(\mathbf{u}')$  no son solo no correlacionadas sino que también son independientes.

- Todas las combinaciones lineales de las variables aleatorias que componen  $Y(\mathbf{u})$  (en modelos univariados) son normalmente distribuidas:

$$X = \sum_{\alpha=1}^n \omega_{\alpha} Y(\mathbf{u}_{\alpha}) \text{ Tiene distribución normal,}$$

$$\forall n, \forall \text{ los pesos } \omega_{\alpha} \text{ siempre y cuando } \mathbf{u}_{\alpha} \in A$$

(6.4)

- Todas las distribuciones condicionales de todos los subconjuntos de la función aleatoria  $Y(\mathbf{u})$ , teniéndose la realización de cualquier otro subconjunto, son distribuciones normales multivariadas. Por ejemplo la distribución condicional de las  $K$  variables aleatorias  $\{Y_k(\mathbf{u}'_k) \ k = 1, \dots, K\}$  siempre que  $\{\mathbf{u}'_k \in A\}$ , dada la realización  $y(\mathbf{u}_{\alpha}) = y_{\alpha}, \alpha = 1, \dots, n$  es variada normal de orden  $K, \forall K, \forall \mathbf{u}'_k, \forall n, \forall \mathbf{u}_{\alpha}, \forall y_{\alpha}$ .

En el caso en que  $K = 1, \mathbf{u}'_1 = \mathbf{u}_o$ , donde la variable aleatoria  $Y(\mathbf{u}_o)$  modela la incertidumbre de un valor no muestreado  $y(\mathbf{u}_o)$ , nos es de gran interés debido a que: la función condicionada de densidad acumulada  $Y(\mathbf{u}_o)$ , dados los  $n$  datos  $y_{\alpha}$ , es normal y completamente caracterizada por:

- La media o valor esperado, el cual es obtenido mediante los pesos del krigeado simple.
- La varianza condicional del krigeado simple.

Siendo solamente necesaria la resolución del sistema del sistema matricial del Krigeado Simple para obtener una función de densidad de probabilidad que nos permita obtener el valor:

$$Z(\mathbf{u}) = E \{Y_k(\mathbf{u})\} \tag{6.5}$$

Definido esto, en nuestro caso de estudio se eligió la simulación Gaussiana secuencial, donde cada variable es simulada secuencialmente de acuerdo con la función condicionada de densidad acumulada; siendo la condicionante todos los valores de los datos originales y todos los valores previamente simulados dentro de un vecindario de búsqueda determinado.

El Software *SGEMS* fue utilizado debido a la manejabilidad que tiene y a que cuenta con un algoritmo para realizar la simulación gaussiana secuencial bastante estable (el cual fue desarrollado previamente para la librería *G<sub>s</sub>TL* y posteriormente, anexo al software), el programa realiza los siguientes pasos para cada realización:

Primero obtiene la función de densidad de probabilidad acumulada,  $F_Z(z)$  del área en estudio.

Se define un recorrido aleatorio con el cual se visita cada nudo de la malla (la cual en principio no importa si es regular o no, aunque se observó en la realización de este trabajo que no produce los mismos resultados) una vez, esto mediante un valor “semilla” que hace que se visiten los nodos como mínimo una vez, ya que si el nodo visitado no tiene datos cercanos se pasa a un segundo nodo y el primero será visitado posteriormente en espera de tener más datos disponibles. En cada nodo  $\mathbf{u}$ , se retiene un cierto número de valores (especificados previamente) condicionantes incluyendo tanto datos iniciales como valores previamente simulados.

En el algoritmo de la simulación gaussiana secuencial se utiliza comúnmente un krigeado simple con el modelo de un variograma normal a fin de determinar la media y la varianza de la función condicionada de densidad acumulada de la función aleatoria  $F(\mathbf{u})$  en el sitio  $\mathbf{u}$ . En nuestro caso utilizamos un krigeado ordinario a fin de poder asimilar el valor esperado

$E\{Y(\mathbf{u})\}$  no estacional para cada locación  $\mathbf{u}$ , existiendo así mayor variabilidad espacial entre datos cercanos.

Después de determinar la media y la varianza en el sitio  $\mathbf{u}$ , se toma un valor simulado  $y^{(l)}(\mathbf{u})$  de la función condicionada de densidad acumulada.

Se añade el valor de  $y^{(l)}(\mathbf{u})$  a los datos iniciales.

Se procede con la siguiente locación y se comienza de nuevo el proceso hasta que todos los nodos están simulados.

Por último se convierten los valores normalizados  $\{y^{(l)}(\mathbf{u}), \mathbf{u} \in A\}$  de vuelta a valores congruentes a los que tiene la variable original  $\{z^{(l)}(\mathbf{u}) = \varphi^{-1}(y^{(l)}(\mathbf{u})) \mid \mathbf{u} \in A\}$ . Para ello se interpola o extrapola dependiendo de si el valor a convertir está en la cola inferior o superior de la función, o bien en el centro de la distribución mediante distintos modelos posibles:

Cola inferior	Centro	Cola Superior
Modelo lineal	Modelo lineal	Modelo lineal
Modelo exponencial	Modelo exponencial	Modelo exponencial
A partir de valores tabulados	A partir de valores tabulados	A partir de valores tabulados
		Modelo Hiperbólico

Tabla 6.1 Modelos de interpolación o extrapolación utilizados al transformar los valores.

Para realizar las simulaciones para el área de estudio se condicionó un vecindario de búsqueda de forma elíptica con un semieje mayor de 3000m y un semieje menor perpendicular de 3000m, además se condicionó a tener un máximo de 40 valores con los cuales realizar la estimación del valor de cada celda.

Se realizaron un total de 500 simulaciones las cuales arrojan valores esperados entre los 0.00001 y 1365 m/d, con una media de 444.5 m/d y una varianza de 44495 (m/d)<sup>2</sup>.

Las figuras 6.1, 6.2 y 6.3 nos muestran los resultados de 3 simulaciones distintas la simulación 22, la simulación 35 y la 43 respectivamente, además la figura 6.4 nos muestra el mapa con los valores esperados.

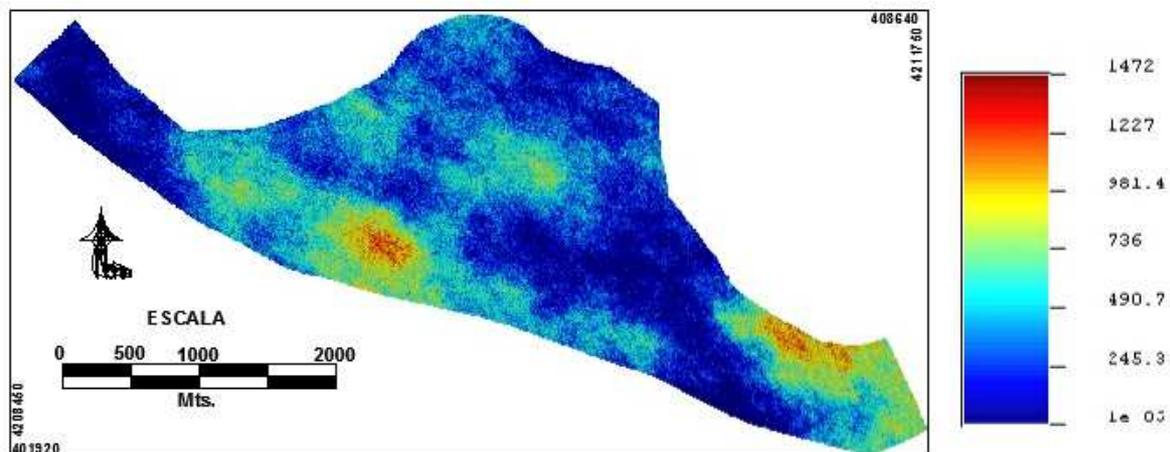


Figura 6.1 Mapa de conductividades hidráulicas (m/d) resultado de la simulación 22.

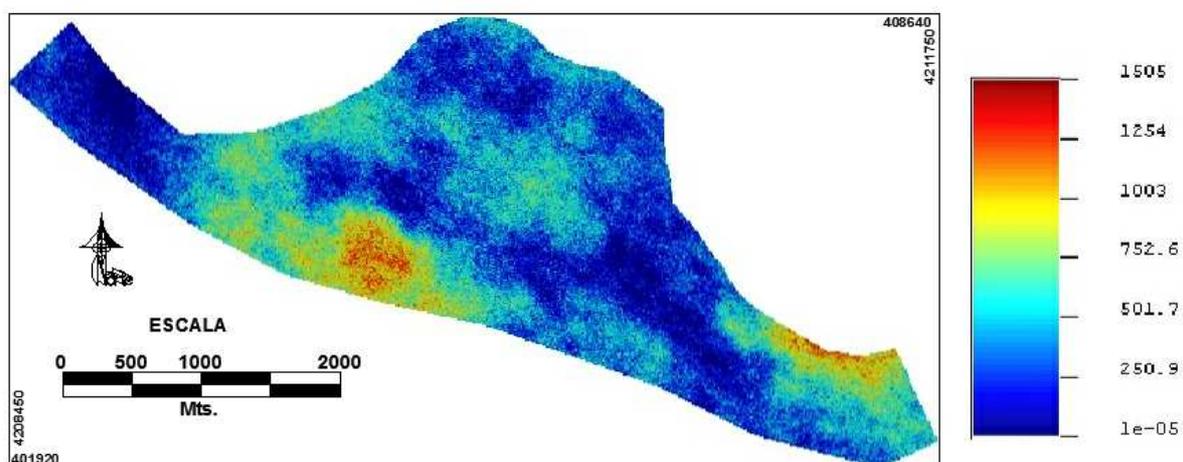


Figura 6.2 Mapa de conductividades hidráulicas (m/d) resultado de la simulación 35.

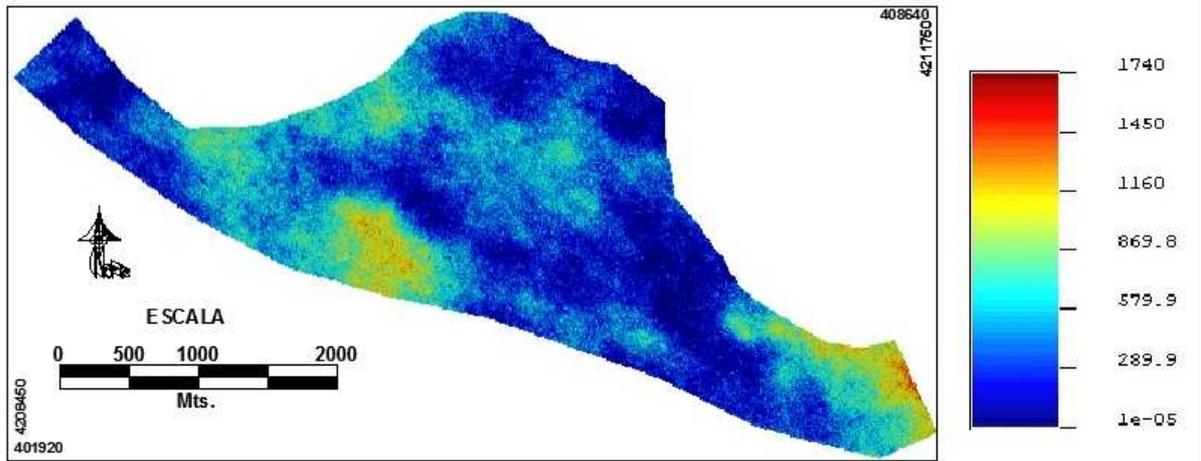


Figura 6.3 Mapa de conductividades hidráulicas (m/d) resultado de la simulación 43.

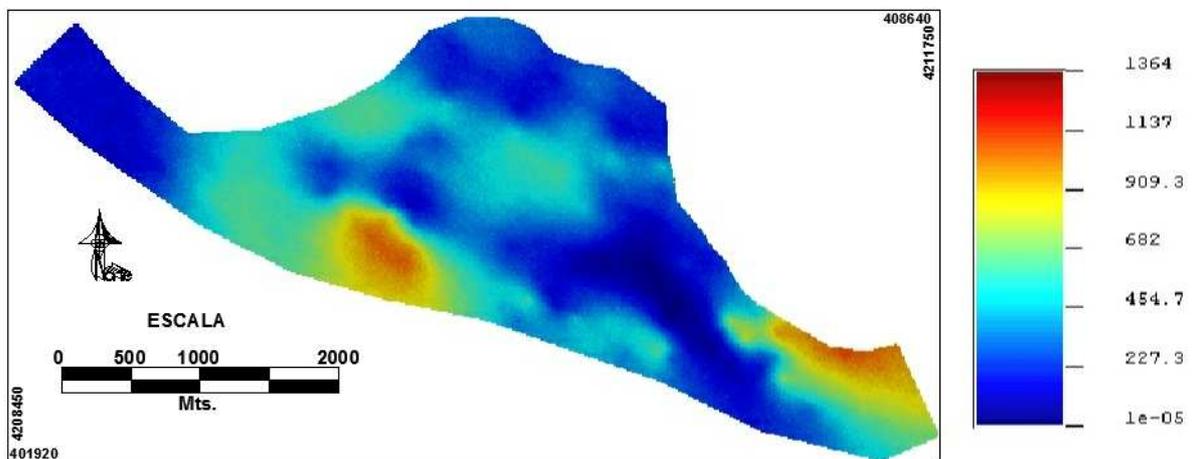
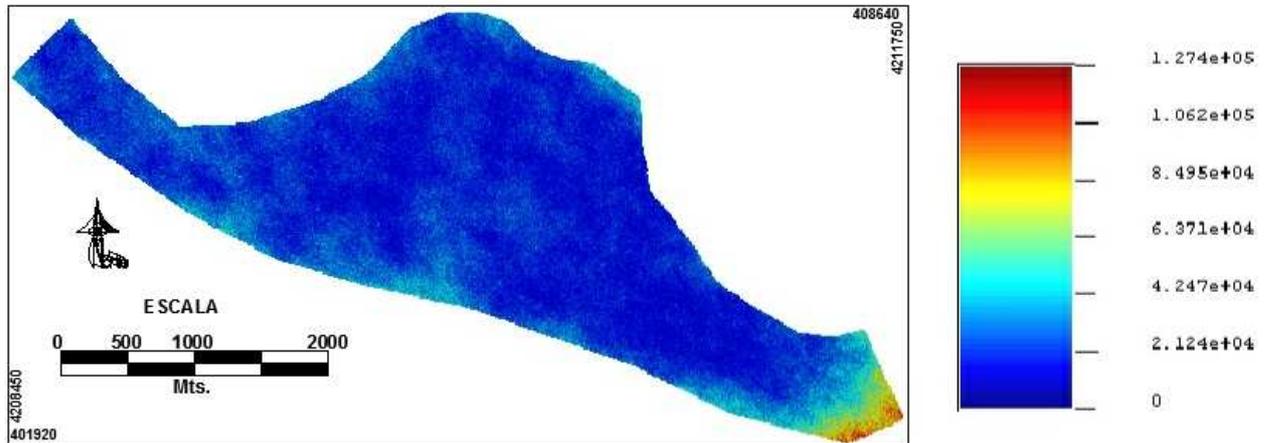


Figura 6.4 Mapa de los valores esperados de las conductividades hidráulicas (m/d).

Cabe mencionar que la figura 6.4 es en sí muy parecida a la figura obtenida con el krigado ordinario (figura 5.1), no obstante, en las otras 3 figuras se ve claramente una heterogeneidad en el terreno introducida en cada simulación lo que le da ese coloreado punteado con un aspecto similar al que crea un crayón de cera y que es más parecido a la realidad.

La figura 6.5 nos muestra el campo de Varianza entre las 500 simulaciones para cada uno de los elementos dentro de la zona de estudio. Cabe señalar que la mayor varianza está en la parte sudoriental del mapa, pues coexisten datos con conductividades muy altas a poca distancia de los valores más bajos de toda el área de estudio.



Figuras 6.5 Varianza Condicional en del Área de estudio.

## 7 CORE

### 7.1 Aspectos teóricos del código

El código utilizado (CORE2D V2.0) es una versión mejorada del código CORE-LE- 2D VO (Samper et al., 1998), el cual es un código de elementos finitos que, como bien se dijo anteriormente resuelve las ecuaciones de flujo, transporte de solutos y transferencia de calor en medios con contornos irregulares y propiedades físicas y geoquímicas no uniformes.

#### 7.1.1 CORE acuífero confinado

El movimiento del agua en un medio poroso es gobernado por la Ley de Darcy, la cual relaciona el flujo de agua  $q$  con el gradiente de presión  $p$  de la siguiente manera:

$$q = \frac{k}{\mu} (\nabla p + \rho g \nabla z) \tag{7.1.1.1}$$

Donde:

$\rho$  es la densidad del agua ( $ML^{-3}$ )

$\mu$  es la viscosidad dinámica ( $MT^{-1}L$ )

$k$  es el tensor de la permeabilidad intrínseca ( $L^2$ )

$g$  es la aceleración de la gravedad ( $LT^{-2}$ )

Si la densidad del agua sufre cambios despreciables, entonces la ley de Darcy puede escribirse en términos de altura piezométrica como sigue:

$$q = -K \nabla h \tag{7.1.1.2}$$

Donde  $K$  se obtiene de la siguiente manera:

$$K = \frac{k \rho g}{\mu} \quad (7.1.1.3)$$

Siendo  $K$  el tensor de conductividades hidráulicas ( $LT^{-1}$ )

Y la altura piezométrica se define como:

$$h = \frac{p}{\rho g} + z \quad (7.1.1.4)$$

$z$  es la altura medida a partir de un nivel de referencia (L)

Combinando la ley de Darcy con la ecuación de balance de masas tenemos que:

$$\nabla \cdot (K \nabla h) + w = S_s \frac{\partial h}{\partial t} \quad (7.1.1.5)$$

Luego:

$w$  representa el termino fuente/sumidero [ $L^3/L^3/T$ ]

$S_s$  es coeficiente de almacenamiento específico [ $L^{-1}$ ]

Como nos encontramos en un acuífero confinado podemos integrar el flujo dentro de todo el espesor del acuífero  $b$ ,  $b$  está definido como la diferencia entre la elevación del fondo y el techo del acuífero llamadas  $z_l$  y  $z_b$  respectivamente. La ecuación resultante es:

$$\nabla \cdot (T \nabla h) + r = S \frac{\partial h}{\partial t} \quad (7.1.1.6)$$

Donde:

$r$  es el termino fuente/sumidero por unidad de superficie [ $L^3/L^2/T$ ]

$T$  es el tensor de transmisividades [ $L^2T^{-1}$ ] y  $S$  es el coeficiente de almacenamiento [-] que se define

como:

$$T(x, y) = \int_{z_b}^{z_t} K(x, y, z) dz = \bar{K}b \quad (7.1.1.7)$$

$$S(x, y) = \int_{z_b}^{z_t} S_s(x, y, z) dz = \bar{S}_s b \quad (7.1.1.8)$$

$\bar{S}_s$  y  $\bar{K}$  son los promedios de las conductividades hidráulicas y los coeficientes de almacenamiento en el sentido vertical.

### 7.1.2 CORE acuífero libre

En un acuífero libre el límite superior coincide con la lamina de agua, en estos casos la trasmisividad  $T$  está dada por  $\bar{K}(h - z_b)$  y depende de la piezometría del acuífero. Además, el coeficiente de almacenamiento se calcula de la siguiente manera:

$$S = \bar{S}_s(h - z_b) + S_y \quad (7.1.2.1)$$

Donde  $s_y$  es el rendimiento específico [-].

Se debe además complementar la ecuación de flujo con condiciones iniciales apropiadas, como podría ser el caso de satisfacer las condiciones del estado estacionario:

$$\nabla \cdot (T \nabla h_o) + r_o = 0 \quad (7.1.2.2)$$

o bien conociendo a priori el estado del sistema, tomando además en cuenta las diferentes condiciones de contorno posibles:

- Dirichlet conociendo la altura piezométrica en el contorno  $\Gamma_1$
- Neuman teniendo un flujo prescrito en el contorno  $\Gamma_2$
- Cauchy con un flujo dependiente de la altura piezométrica en  $\Gamma_3$

### 7.1.3 CORE zona no saturada

Si se tiene un flujo en un medio con saturación variable, la altura piezométrica se escribe como la suma de la altura de presión  $\Psi$  y la elevación  $z$ . Siendo  $\Psi$  positiva en medios saturados y negativa en regiones parcialmente saturadas. Por tanto se puede calcular el flujo por medio de la ley de Darcy y la ley de conservación de la masa, tomando en cuenta la altura de presión  $\Psi$ :

$$\nabla \cdot [K_r K \nabla (\Psi + z)] + w = \left( \phi \frac{\partial S_w}{\partial \Psi} + S_w S_s \right) \frac{\partial \Psi}{\partial t} \quad (7.1.3.1)$$

De esta ecuación el término de la conductividad hidráulica es el producto de la conductividad relativa  $K_r$  [-] y la conductividad saturada  $K$ .

El término que multiplica a la derivada parcial de la altura de presión con respecto al tiempo es la capacidad de almacenamiento, el primer término es importante en zonas no saturadas, mientras que el segundo término es relevante en zonas saturadas.

Ambos términos anteriormente señalados, así como la conductividad relativa  $K_r$  dependen de la altura de presión, lo que hace a esta ecuación extremadamente no lineal, tomando en cuenta que además se debe calcular el contenido de humedad de la siguiente manera:

$$S_w = \frac{\theta}{\phi} \quad (7.1.3.2)$$

Donde:

$\emptyset$  es la porosidad del medio

$\theta$  es el contenido de humedad

Al no poder calcular de manera directa  $S_w$ , este se calcula con la función de van Genuchten(1980):

$$S_w(\Psi) = S_r + (1 - S_r)[1 + (-\alpha\Psi)^n]^{-m} \quad (7.1.3.3)$$

$$K_r(\Psi) = \sqrt{S_e}[1 - (1 - S_e^{1/m})^m]^2 \quad (7.1.3.4)$$

De donde:

$$m = 1 - \frac{1}{n} \quad (7.1.3.5)$$

$$S_e = \frac{S_w - S_r}{1 - S_r} \quad (7.1.3.6)$$

Siendo:

$S_r$  la saturación residual

$m, n$  y  $\alpha$  ( $1/L^{-1}$ ) son constantes que se obtienen al ajustar esta función a datos experimentales.

$\alpha$  puede ser interpretado como el valor recíproco de la succión crítica  $\Psi_c$ :

$$\alpha = \frac{1}{\Psi_c} \quad (7.1.3.7)$$

Obteniéndose a través de la porosidad y del tamaño efectivo de partícula  $d_{10}$  (Bear 1972).

Finalmente al igual que en el flujo en el acuífero, el flujo en la zona de transición se calcula a partir de condiciones iniciales y de contorno. Las condiciones iniciales deben ser conocidas en  $t = 0$ :

$$\Psi(x, z, t = 0) = \Psi_0(x, z)$$

(7.1.3.8)

En cuanto a las condiciones de contorno deben ser de cualquiera de las siguientes formas:

- Dirichlet conociendo la altura piezométrica en el contorno  $\Gamma_1$
- Neuman teniendo un flujo prescrito en el contorno  $\Gamma_2$
- Cauchy con un flujo dependiente de la altura piezométrica en  $\Gamma_3$
- Condición de drenaje Libre, en la cual se tiene un gradiente hidráulico vertical unitario cuando la zona de interés es la no saturada.

## 7.2 Solución numérica del flujo en CORE

Como se comentó al inicio del capítulo CORE utiliza el método de elementos finitos para aproximar la solución numérica del problema de flujo.

A continuación se plantean los algoritmos de resolución del problema de flujo para acuíferos confinados, acuíferos libres y la resolución de las ecuaciones que gobiernan el movimiento del agua en medios porosos no saturados según lo explica el manual de usuario del CORE 2.0 (ENRESA, 2000).

### 7.2.1 Solución del flujo en el acuífero

El método de elementos finitos aproxima la solución exacta  $h(x, y, t)$  mediante una solución numérica  $h^N(x, y, t)$  de la siguiente manera:

$$h(x, y, t) \cong h^N(x, y, t) = \sum_{m=1}^N h_m(t) \xi_m(x, y) \quad (7.2.1.1)$$

Donde  $h_m(t)$  es la altura piezométrica en el nodo  $m$  en un tiempo  $t$  y  $\xi_m(x, y)$  son las  $N$  funciones bases que son linealmente independientes en el dominio del flujo, que por conveniencia se definen como 1 en el nodo y cero en todos los otros nodos, está definida así:

$$\xi_m(x, y) = \delta_{nm} \quad (7.2.1.2)$$

Donde:

$\delta_{nm}$  se conoce como la función delta de Kronecker.

Si bien se comentó anteriormente que el área de estudio está dividida en una malla triangular, cabe decir que la solución de los elementos finitos es definida por los valores de la altura piezométrica en los nodos de la malla  $h_m(t)$ , por lo que la interpolación dentro del

área de los elementos triangulares define un plano con las aristas en los tres vértices del triángulo.

La solución numérica aproxima la solución exacta, la cual utiliza la siguiente ecuación:

$$L(h) = \nabla \cdot (T\nabla h) + r - S \frac{\partial h}{\partial t} = 0 \quad (7.2.1.3)$$

Como la solución numérica no logra obtener una solución completamente idéntica a la exacta, se tiene un error en el balance de masas  $\varepsilon(x, y, t)$ :

$$L(h^N) = \nabla \cdot (T\nabla h^N) + r - S \frac{\partial h^N}{\partial t} = \varepsilon \neq 0 \quad (7.2.1.4)$$

El cual es impuesto a que sea ortogonal a alguna función de pesos  $w_n(x, y)$ :

$$\int_{\Omega} L(\varepsilon)w_n dx = 0 \quad (7.2.1.5)$$

Que se extiende por todo el dominio  $\Omega$ .

El método de elementos finitos de Galerkin está basado en adoptar como pesos el mismo conjunto de funciones base  $\xi_m(x, y)$ , por lo que la función puede escribirse de la siguiente manera:

$$\int_{\Omega} \left[ \nabla \cdot (T\nabla h^N) + r - S \frac{\partial h^N}{\partial t} \right] \xi_n dx = 0 \quad (7.2.1.6)$$

Puesto que en CORE los elementos y las funciones base son lineales, la derivada de  $\xi_n$  de primer orden da como resultado una constante y por lo tanto se pierden términos como la divergencia en las alturas piezométricas. Tomando en cuenta esto, el código reescribe la integral del primer término mediante el término identidad de Green, que para cualquier vector se escribe:

$$\int_{\Omega} \nabla \cdot f \xi_n(x) dx = \int_{\Omega} \nabla \cdot (f \xi_n) dx - \int_{\Omega} f \cdot \nabla \xi_n dx \quad (7.2.1.7)$$

El cual al aplicar el teorema de divergencia al primer término del lado derecho de la ecuación se convierte en:

$$\int_{\Omega} \nabla \cdot f \xi_n dx = \int_{\Gamma} f \cdot n \xi_n d\Gamma - \int_{\Omega} f \cdot \nabla \xi_n dx \quad (7.2.1.8)$$

Para  $f = T\nabla h$ , la primera identidad de Green puede ser aplicada a la integración del primer término, modificando la ecuación 7.2.1.6:

$$\int_{\Gamma} (T\nabla h)^N \cdot n \xi_n d\Gamma - \int_{\Omega} (T\nabla h)^N \cdot \nabla \xi_n dx + \int_{\Omega} r \xi_n dx = 0 \quad (7.2.1.9)$$

Tomando en cuenta el planteamiento inicial de la ecuación 7.2.1.1 y usando la condición de contorno general:

$$(T\nabla h) \cdot n = -q \cdot n = \alpha(H - h) + Q \quad (7.2.1.10)$$

La ecuación de elementos finitos se convierte en:

$$\sum_{m=1}^N h_m \int_{\Omega} (T \nabla \xi_m) \cdot \nabla \xi_n dx - \int_{\Gamma} \left[ \alpha(H - \sum_{m=1}^N h_m \xi_m) + Q \right] \xi_n + \sum_{m=1}^N \frac{dh_m}{dt} \int_{\Omega} S \xi_m \xi_n dx = \int_{\Omega} r \xi_n dx \quad (7.2.1.11)$$

Según se explica en el manual, todas las integrales dentro del dominio  $\Omega$  deben ser calculadas sobre una base en los elementos ya que las funciones de forma son dependientes linealmente de la variable en este caso los elementos. Por tanto las ecuaciones pueden ser escritas de la siguiente manera:

$$\sum_{m=1}^N h_m A_{mn} + \sum_{m=1}^N \frac{dh_m}{dt} D_{mn} = d_n \quad (7.2.1.12)$$

Donde:

$$A_{mn} = \sum_e A_{mn}^e + \sum_s A_{mn}^e \quad (7.2.1.13)$$

$$D_{mn} = \sum_e D_{mn}^e \quad (7.2.1.14)$$

$$d_n = \sum_e d_n^e + \sum_e d_n^s \quad (7.2.1.15)$$

El superíndice  $e$  denota el número de elemento que contiene los nodos  $m$  y  $n$ ,  $s$  son segmentos adyacentes al nodo  $n$  que juegan un papel de frontera. La expresión de  $A_{mn}^e$ ,  $D_{mn}^e$ ,  $d_n^e$ ,  $A_{mn}^s$  y  $d_n^s$  son:

$$A_{mn}^e = \frac{1}{4\Delta^e} [T_{xx}^e b_m^e b_n^e + T_{xy}^e (b_n^e c_m^e + b_m^e c_n^e) + T_{yy}^e c_n^e c_m^e] \quad (7.2.1.16)$$

$$D_{mn}^e = \frac{S^e \Delta^e}{12} \eta_{mn}$$

$$\text{Donde: } \eta_{mn} = \begin{cases} 2 & \text{si } m = n \\ 1 & \text{si } m \neq n \end{cases} \quad (7.2.1.17)$$

$$d_n^e = \frac{r^e \Delta^e}{3} \quad (7.2.1.18)$$

$$A_{mn}^s = \frac{\alpha^s L^s}{6} \eta_{mn} \quad (7.2.1.19)$$

$$d_n^s = \frac{(\alpha^s H^s + Q^s) L^s}{2} \quad (7.2.1.20)$$

Donde  $\Delta^e$  es el área del elemento,  $b_m^e$  y  $c_m^e$  son coeficientes de los elementos que dependen de la geometría del elemento.

La ecuación 2.3.11 pueden ser escritas en forma de matriz como:

$$Ah + D \frac{dh}{dt} = d \quad (7.2.1.21)$$

Donde  $A$  es una matriz de conductancias con  $N^2$  valores,  $D$  es una matriz de capacitancias con  $N^2$  valores,  $d$  es un vector de términos independientes en forma de columna, y  $h$  es un vector en forma de columna de alturas piezométricas. Este sistema puede resolverse por medio del método de diferencias finitas resultando en:

$$A[\varepsilon h^{k+1} + (1-\varepsilon)h^k] + D \frac{h^{k+1} - h^k}{\Delta t} = d^{k+1} \quad (7.2.1.22)$$

Donde:

$\Delta t = t_{k+1} - t_k$ , es decir, es el resultado de aplicar el método a la derivada respecto a  $t$   $dt$   
 $0 \leq \varepsilon \leq 1$ ,  $h^k$  y  $h^{k+1}$  son los vectores de los niveles piezométricos en los tiempos  $t_k$  y  $t_{k+1}$  siendo el esquema de integración explícito  $\varepsilon = 0$  implícito  $\varepsilon = 1$  o con un esquema Crank-Nicholson  $\varepsilon = 0.5$

Así el sistema se resuelve secuencialmente una vez conocido  $h^k$  se procede a resolver para el tiempo  $h^{k+1}$  mediante la expresión:

$$\left(\varepsilon A + \frac{D}{\Delta t}\right) h^{k+1} = d^{k+1} + \left[(\varepsilon - 1)A + \frac{D}{\Delta t}\right] h^k \quad (7.2.1.23)$$

La transmisividad  $T$  es el producto del espesor del acuífero  $b$  y la conductividad hidráulica  $K$ . Si el acuífero es confinado, el espesor del acuífero es constante con respecto al tiempo, variando solamente de entre elementos, CORE asume un valor de  $b=1$  para el flujo vertical en acuíferos confinados.

Para el flujo en 3 dimensiones con simetría axial en acuíferos confinados CORE calcula  $b^e$  como la longitud promedio de la coordenada radial del elemento:

$$b^e = 2\pi \frac{x_1^e + x_2^e + x_3^e}{3} \quad (7.2.1.24)$$

Siendo  $x_1^e, x_2^e, x_3^e$  las coordenadas radiales de cada elemento.

Si el acuífero es libre, entonces el espesor saturado varía tanto con la distancia como con el tiempo, por lo que es necesaria la utilización de métodos iterativos para resolverlo. El espesor saturado promedio es actualizado a partir de la iteración anterior de acuerdo a la siguiente expresión:

$$b^e = \frac{h_1^e + h_2^e + h_3^e}{3} - z_b^e \quad (7.2.1.25)$$

Donde:

$h_1^e$  son las alturas piezométricas en los nodos y  $z_b^e$  es la base del elemento  $e$ .

Este cálculo iterativo se repite hasta que se cumpla la siguiente condición:

$$\max_i \left[ \frac{h_i^{k+1,s+1} - h_i^{k+1,s}}{h_i^{k+1,s+1} + h_i^{k+1,s}} \right] \leq \omega \quad i = 1, 2, \dots, N \quad (7.2.1.26)$$

Donde  $k$  y  $s$  indican el intervalo de tiempo y el nivel de iteración respectivamente; y  $\omega$ , es un parámetro específico que determina tolerancia del error en la altura piezométrica. En el primer intervalo y la primera iteración  $b$  es calculado a partir de la distribución inicial de  $h$ , resolviendo primero el sistema en estado estacionario.

Una vez definido  $h$  se resuelven los componentes  $x$  e  $y$  del vector de velocidades de acuerdo a la ley de Darcy:

$$q_x^e = - \left( K_{xx}^e \sum_{i=1}^3 h_i^e b_i^e + K_{xy}^e \sum_{i=1}^3 h_i^e c_i^e \right) \quad (7.2.1.27)$$

$$q_y^e = - \left( K_{yy}^e \sum_{i=1}^3 h_i^e c_i^e + K_{xy}^e \sum_{i=1}^3 h_i^e b_i^e \right) \quad (7.2.1.28)$$



## 7.2.2 Solución del flujo en la zona de transición

Los métodos numéricos para la solución del flujo en zonas con saturación variable son similares a las del flujo en acuíferos, solo que dependen de la altura de presión. La ecuación del flujo (7.2.1.12), puede ser discretizada por medio del método de elementos finitos para obtener:

$$\sum_{m=1}^N \Psi_m A_{mn} + \sum_{m=1}^N \frac{d\Psi_m}{dt} D_{mn} = d_n + B_n \quad (7.2.2.1)$$

Donde:

$A_{mn}$ ,  $D_{mn}$  y  $d_n$  son las ecuaciones 7.2.1.13, 7.2.1.14, y 7.2.1.15, en donde  $A_{mn}^e$ ,  $D_{mn}^e$  puede ser expresado como:

$$A_{mn}^e = K_F^e C_{mn}^e \quad (7.2.2.2)$$

Donde  $C_{mn}^e$  tienen la siguiente expresión:

$$C_{mn}^e = \frac{1}{4\Delta^e} [K_{xx}^e b_m^e b_n^e + K_{xy}^e (b_n^e c_m^e + b_m^e c_n^e) + K_{yy}^e c_n^e c_m^e] \quad (7.2.2.3)$$

Y  $D_{mn}^e$  es calculado de la siguiente manera:

$$C_{mn}^e = \frac{S^e \Delta^e}{12} \eta_{mn} = \begin{cases} 2 & \text{si } m = n \\ 1 & \text{si } m \neq n \end{cases} \quad (7.2.2.4)$$

De Acuerdo con la ecuación 7.2.1.12, para flujo no saturado, la capacidad de almacenamiento  $S^e$  esta dada por:

$$S^e = \phi^e \left. \frac{\partial S_w}{\partial \psi} \right|^e + S_w^e S_s^e \quad (7.2.2.5)$$

El término  $B_n$  de la ecuación 7.2.2.1 se deriva del término gravitatorio, y tiene la siguiente expresión:

$$B_n = - \frac{K_\Gamma^e}{2} (K_{xx}^e c_n + K_{xy}^e b_n^e) \quad (7.2.2.6)$$

Donde  $x$  e  $y$  son las coordenadas vertical y horizontal respectivamente.

El valor promedio en cada elemento de la conductividad hidráulica relativa, la saturación de agua y los demás términos derivados son evaluados con la altura de presión promedio de los tres nudos, o bien:

$$K_\Gamma^e = K_\Gamma(\psi^e) \quad (7.2.2.7)$$

$$S_w^e = S_w(\psi^e) \quad (7.2.2.8)$$

$$\left. \frac{\partial S_w}{\partial \psi} \right|^e = \frac{\partial S_w}{\partial \psi}(\psi^e) \quad (7.2.2.9)$$

Donde:

$$\psi^e = \left( \frac{\psi_1 + \psi_2 + \psi_3}{3} \right)^e \quad (7.2.2.10)$$

Los términos de la ecuación 7.2.2.1 correspondientes a la conductancia  $A_{mn}$ , capacidad de almacenamiento  $D_{mn}$  y el término independiente  $B_n$  dependen de la altura de presión, lo que ocasiona que sea una ecuación no lineal en gran medida. Para resolverla se utiliza el método iterativo de Newton-Raphson modificado por Galarza (1993) mediante la formulación 3 de este último. A partir de esto se modifica la ecuación 7.2.2.1 de la siguiente manera:

$$A^{k+\varepsilon}\psi^{k+\varepsilon} + D^{k+\varepsilon}\frac{\psi^{k+1} - \psi^k}{\Delta t} = d^{k+1} + B^{k+\varepsilon} \quad (7.2.2.11)$$

Donde:

$$A^{k+\varepsilon} = A(\psi^{k+\varepsilon}) \quad (7.2.2.12)$$

$$D^{k+\varepsilon} = D(\psi^{k+\varepsilon}) \quad (7.2.2.13)$$

$$\psi^{k+\varepsilon} = \varepsilon\psi^{k+1} + (1 - \varepsilon)\psi^k \quad (7.2.2.14)$$

Y el valor de  $\varepsilon$  debe ser  $0 \leq \varepsilon \leq 1$  la ecuación 7.2.2.11 puede ser reescrita de la siguiente manera:

$$f = A^{k+\varepsilon}\psi^{k+\varepsilon} + D^{k+\varepsilon}\frac{\psi^{k+1} - \psi^k}{\Delta t} - d^{k+1} - B^{k+\varepsilon} = 0 \quad (7.2.2.15)$$

Donde  $\psi^k$  es conocida,  $\psi^{k+1}$  es desconocida, y  $\psi^{k+\varepsilon}$  es una combinación de ambos. El método iterativo de Newton-Raphson está basado en las series de expansión de Taylor. Por lo que al reacomodar estas expansiones el sistema de ecuaciones queda de la siguiente manera:

$$\sum_{i=1}^N \frac{\partial f_n}{\partial \psi_i^{k+1}} \Delta \psi_i^{k+1} = -f_n \quad n = 1, 2, \dots, N$$

(7.2.2.16)

Donde  $N$  es el número de nodos. En forma matricial la ecuación se reduce a la siguiente forma:

$$J \Delta \psi^{k+1} = -f$$

(7.2.2.17)

Donde  $J$  es la matriz Jacobiana que contiene los términos diferenciales de la ecuación 7.2.2.16 y  $f$  es el vector de residuos. Una vez se han calculado los valores de  $\Delta \psi^{k+1}$  estos son utilizados para actualizar  $\psi^{k+1}$  de acuerdo a:

$$\psi^{k+1,s+1} = \psi^{k+1,s} + \Delta \psi^{k+1}$$

(7.2.2.18)

$$\psi^{k+1,0} = \psi^k$$

(7.2.2.19)

Donde el subíndice  $s$  es el número de iteración. El proceso iterativo es realizado hasta que el valor absoluto del incremento relativo de la altura de presión para todos los nodos es menor que el especificado con un valor de tolerancia en la convergencia (en general un número muy pequeño):

$$\max_i \left( \left| \frac{\Delta \psi^{k+1}}{\psi^{k+1,s}} \right| \right) \leq \omega \quad i = 1, 2, \dots, N$$

(7.2.2.20)

La matriz Jacobiana  $J$  se obtiene tomando derivadas de la ecuación 7.2.2.15 con respecto a la incógnita  $\Delta \psi^{k+1}$ , y tomando en cuenta las ecuaciones 7.2.2.12 a 7.2.2.14, o bien:

$$J = \varepsilon \left( \frac{\partial A^{k+\varepsilon}}{\partial \psi^{k+\varepsilon}} \psi^{k+\varepsilon} \right) + \frac{\varepsilon}{\Delta t} \left( \frac{\partial D^{k+\varepsilon}}{\partial \psi^{k+\varepsilon}} (\psi^{k+1} - \psi^k) \right) + \varepsilon A^{k+\varepsilon} + \frac{D^{k+\varepsilon}}{\Delta t} - \varepsilon \frac{\partial B^{k+\varepsilon}}{\partial \psi^{k+\varepsilon}} \quad (7.2.2.21)$$

La matriz Jacobiana  $J$  es una matriz en banda no simétrica con  $N \times N$  que puede resolverse al multiplicar elemento por elemento (es decir, teniendo 2 matrices  $(X; Y)$  el primer elemento de la matriz  $X$  por el primer elemento de la matriz  $Y$ ) y luego reasignando las contribuciones de cada elemento de una manera similar a como estaba la matriz de conductancia  $A$  original.

### 7.3 Datos de entrada al modelo CORE

El modelo CORE consta de 3 archivos de entrada, escritos en archivos de texto que, actuando en conjunto, son leídos por el código para resolver problemas de flujo, transporte de solutos reactivos y transporte de calor; combinando el flujo con el transporte de solutos reactivos o con el transporte de calor o bien todos en conjunto.

El archivo ROOT.inp (siendo ROOT el nombre que se utilizara para esa simulación en específico) tiene los parámetros de control generales del problema, como es el número de parámetros, número de iteraciones, los pesos iniciales con los que se comienza a ajustar la función de flujo, el valor mínimo y máximo de los parámetros, las coordenadas y nombres de los puntos de muestreo y observación, el número de especies químicas con que lidiara el código, etc.

El archivo ROOT\_che.inp es un archivo en donde se detallan valores de distintos complejos acuosos, gases, cationes de intercambio y complejos superficiales, así como las distintas especies químicas que tratara el problema. Este archivo debe contener la temperatura inicial del problema (por defecto 25 grados centígrados) y en general los parámetros que controlan el transporte de reactivos. También contiene los tipos de agua de contorno y de recarga con que tratara el sistema, así como la función de tiempo que corresponde a cada tipo de agua.

El archivo ROOT\_tra.inp es el archivo en donde se recoge toda la información espacial del modelo, coordenadas de nodos, número de nodos, número de elementos, número de zonas de material que se encuentra en la zona, tipo de acuífero, condiciones iniciales y de contorno.

También en este archivo se detalla la información de la discretización temporal del problema, las funciones de flujo y de transporte que utilizara el modelo, el número y nombre de cada uno de los archivos de salida que generará el modelo (máximo 21), los datos que tomara como criterio de convergencia el modelo (el cual se detalla en la sección anterior), las variables que controlan la creación de la matriz Jacobiana, los nodos de los que se van a imprimir resultados en los archivos de salida, etc.

Además el código lee datos de valores observados, los espesores del acuífero, los datos de concentraciones de entrada, etc. Especificados en archivos de texto con la extensión .dat y deben ser creados para cada problema en específico. No obstante dentro de los archivos .dat se encuentran 2, el archivo Masterte.dat y el archivo master25.dat que son los que controlan todas las reacciones e iteraciones de compuestos químicos con los que puede trabajar CORE, y que, no se contemplan en el alcance que tiene este estudio. Para mayor detalle véase el manual de usuario de CORE V 2.0 (ENRESA 200).

El modelo CORE ha sido reestructurado desde el código fuente para poder utilizar una cantidad mayor de elementos que los que se tenían contemplados en el modelo de 1994, así como también se ha utilizado una mayor cantidad de zonas de material que las que tenía el modelo de 2004.

## 7.4 Simulación de alturas piezométricas modelo CORE

Se realizó una simulación con los datos y parámetros de las simulaciones realizadas con el modelo del 2004 que abarcan desde octubre de 2003 hasta octubre de 2020, sirviendo de caso de partida. Se discretizó temporalmente con un incremento de 15 días para que el código resolviera las ecuaciones de flujo.

Luego se modificó el caso base a fin de tomar en cuenta la heterogeneidad a menor escala, incrementando el número de zonas de material, hasta obtener 12254 zonas distintas que corresponden exactamente con los 12254 elementos triangulares con los que cuenta la malla del modelo del año 2004. Se realizó de esta manera debido a que así es posible introducir mayor heterogeneidad al medio sin ser necesarios grandes cambios en la estructura del código fuente de CORE, que de haberlos realizado, hubiese sido necesario el reestructurar casi en su mayoría el código, además de posiblemente ralentizar los cálculos, pues CORE realiza acciones tanto en los nodos como en los elementos. No obstante, la metodología elegida es apropiada pues evita grandes cambios estructurales y toma en cuenta todos los elementos dentro del área de estudio, obteniendo resultados similares a los que se obtendrían al utilizar los elementos de una malla fina.

Se calculó la media geométrica del conjunto de elementos de la cuadrícula generada para el cálculo de los campos de conductividades que estuviese dentro del elemento triangular donde ocupase mayor área cada elemento de la cuadrícula en particular. Para lograr lo anterior primeramente se utilizó el proyecto creado en ArcGis, el cual cuenta con las capas de los elementos triangulares, así como la cuadrícula de 5x5m para obtener las coordenadas de los vértices de cada elemento de cada una de las capas, a continuación se creó un programa de comparación de coordenadas, que, mediante una serie de bucles, asignó para cada uno de los elementos triangulares los elementos de la cuadrícula que tuviesen la mayoría de su área dentro del triángulo. Esto dio como resultado una lista de 887040 elementos que tenían asignados el elemento triangular apropiado, existiendo claro está, elementos de la cuadrícula que no tenían ningún elemento asignado, pues se encuentran fuera del área de estudio.

A partir de esta lista se realizó el cálculo de las medias geométricas. Fue necesaria la creación de un programa para calcular de manera eficiente cada una de las medias geométricas, El cálculo de las medias geométricas se realizó mediante la siguiente ecuación:

$$x = \sqrt[n]{\prod_{i=1}^n x_i}$$

(7.4.1)

Básicamente, el programa busca los valores de las conductividades hidráulicas que tiene cada uno de los elementos de las simulaciones provenientes de SGems que cumplan con la restricción de estar localizados dentro de un mismo elemento triangular y posteriormente realiza el cálculo de la media geométrica. Se eligió el cálculo de la media geométrica ya que se ha comprobado que para casos bidimensionales isotrópicos esta logra obtener los valores más precisos debido a que considera todos los valores de la distribución y es menos sensible que la media aritmética a los valores extremos.

Esto dio como resultado un total de 12254 zonas con conductividades distintas las cuales se introdujeron en los archivos de entrada del CORE. Las figuras 7.4.1.1, 7.4.1.2, 7.4.1.3 y 7.4.1.4 nos muestran 2 simulaciones, así como los valores esperados de las 12254 zonas y la varianza condicional de las mismas.

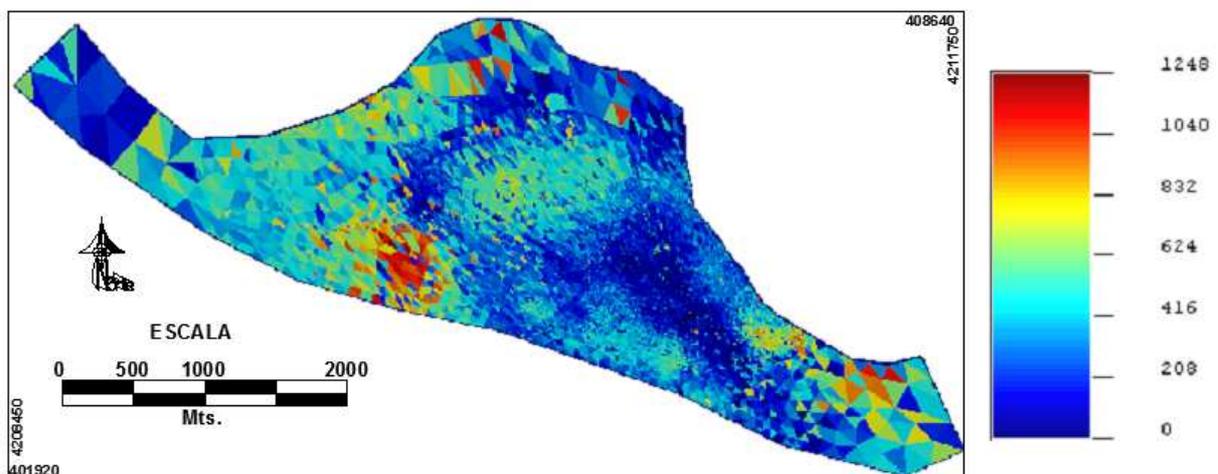


Figura 7.41.1 Mapa con 12254 zonas de conductividades a partir de la simulación 11.

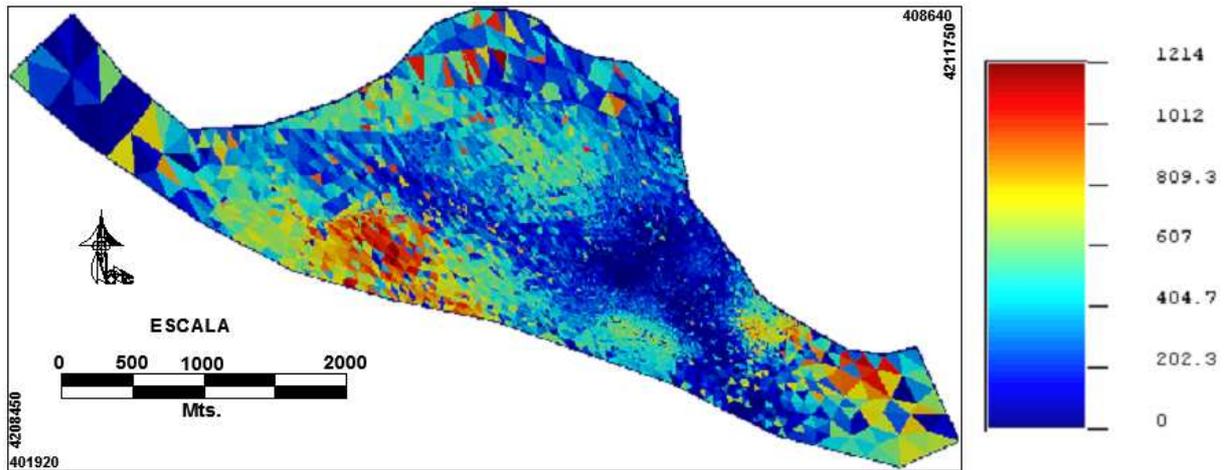


Figura 7.41.2 Mapa con 12254 zonas de conductividades a partir de la simulación 61.

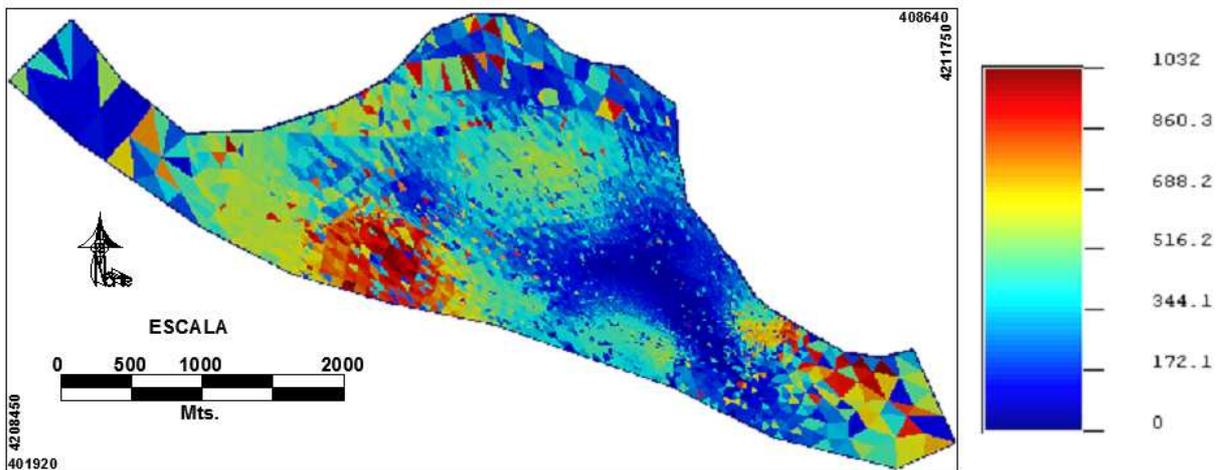


Figura 7.41.3 Mapa de valores esperados de las 12254 zonas de conductividades a partir del conjunto de simulaciones.

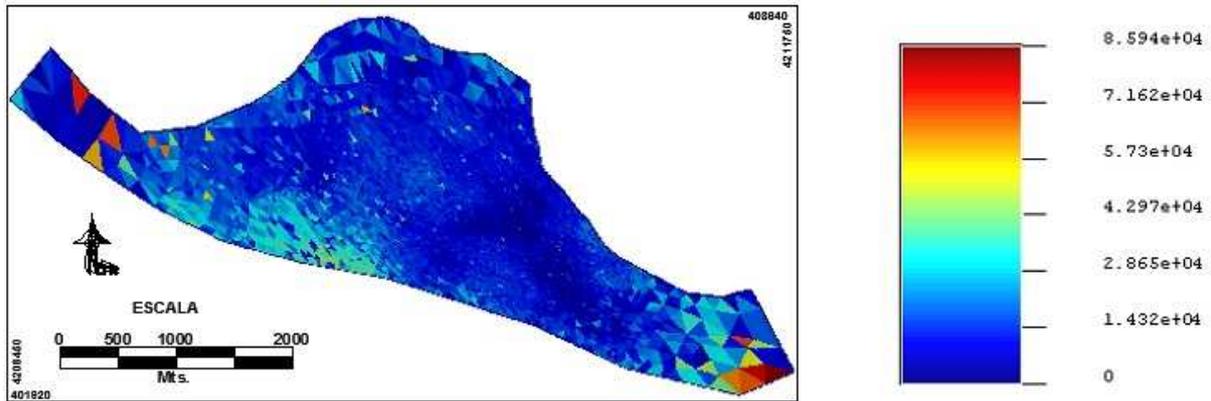


Figura 7.41.4 Mapa de varianza condicional de las 12254 zonas de conductividades a partir del conjunto de simulaciones.

Finalmente se compararon con los valores observados, según se describen en las figuras de la 7.4.1.5 a la 7.4.1.9. En ellas se observa la evolución de los niveles piezométricos para cuatro pozos, el pozo de control 1 PC1, el pozo de control 5 PC5 y dos pozos más el pozo 611 y el pozo 472.

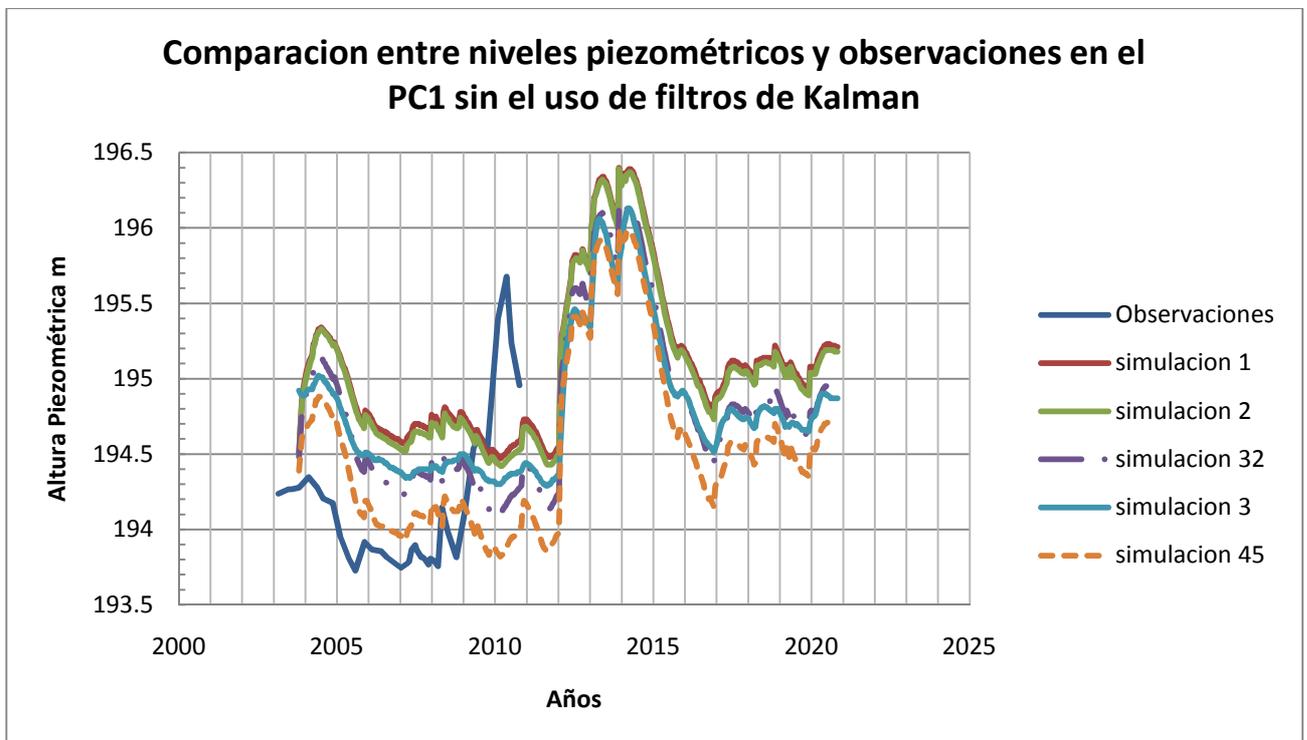


Figura 7.4.1.5 Alturas Piezométricas observadas en el PC1 comparadas con los valores que reproduce el CORE para el periodo 2003-2020.

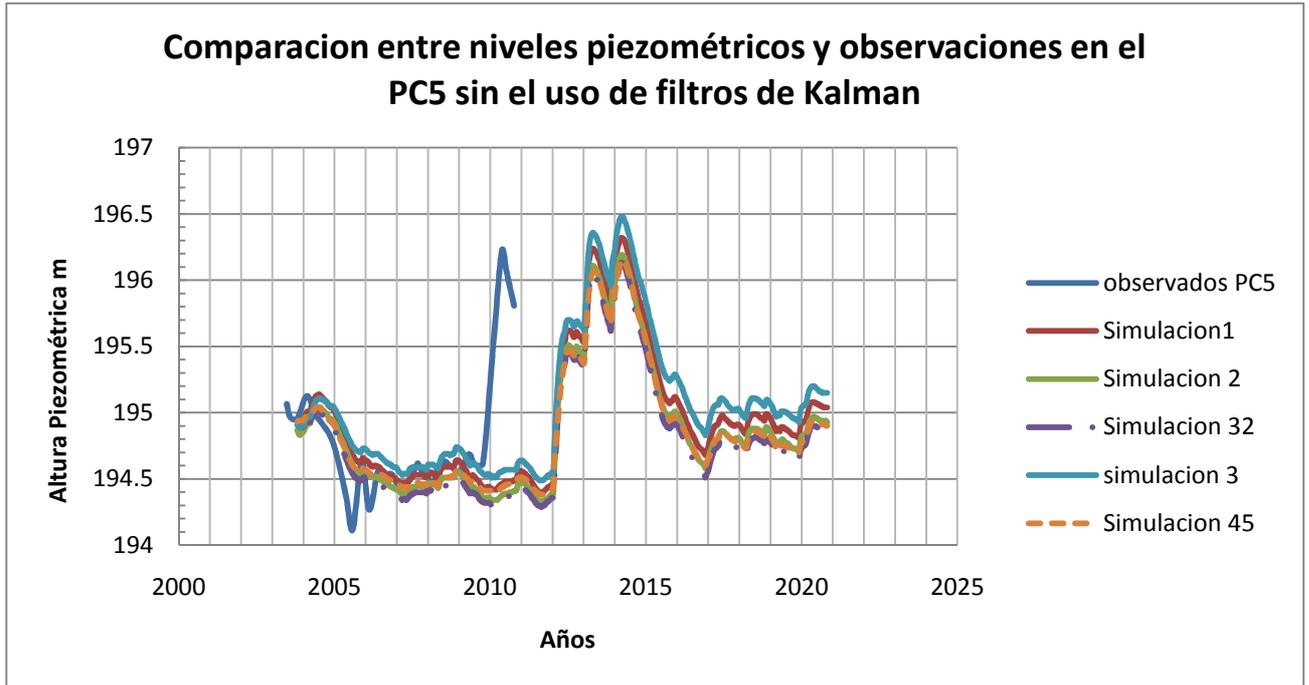


Figura 7.4.1.6 Alturas Piezométricas observadas en el PC5 comparadas con los valores que reproduce el CORE para el periodo 2003-2020.

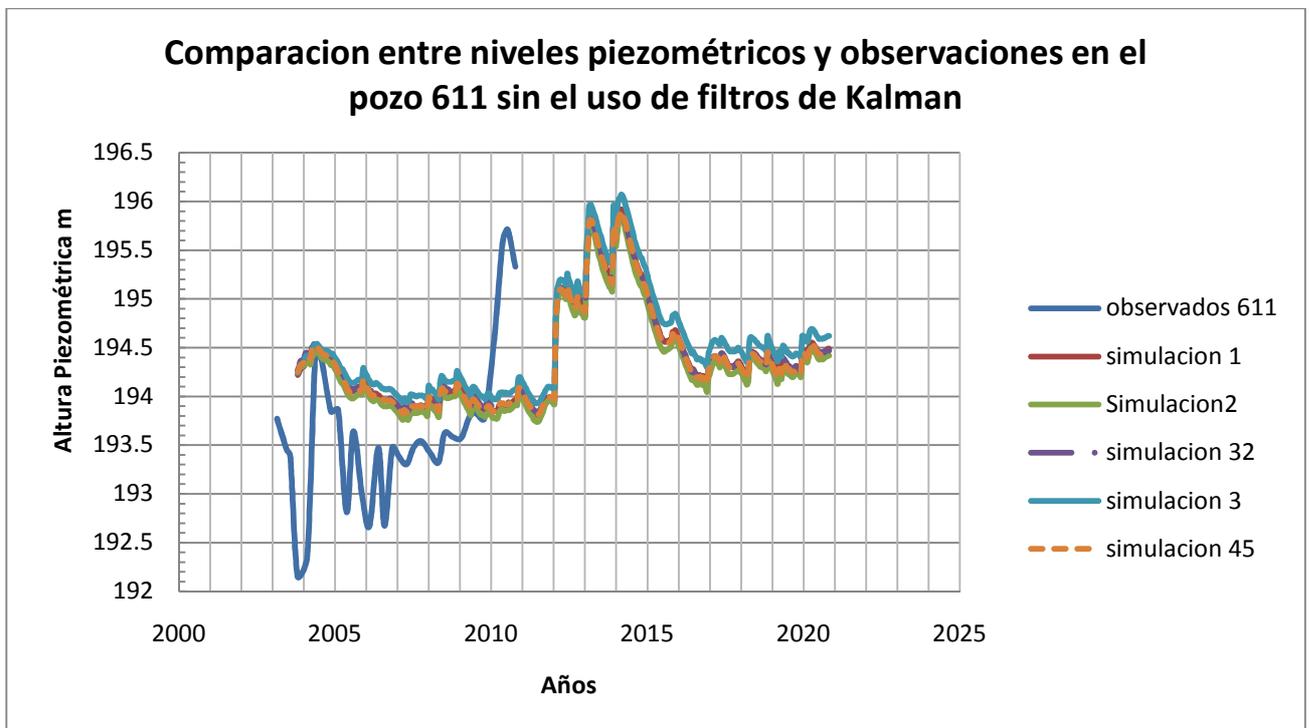


Figura 7.4.1.7 Alturas Piezométricas observadas en el pozo 611 comparadas con los valores que reproduce el CORE para el periodo 2003-2020.

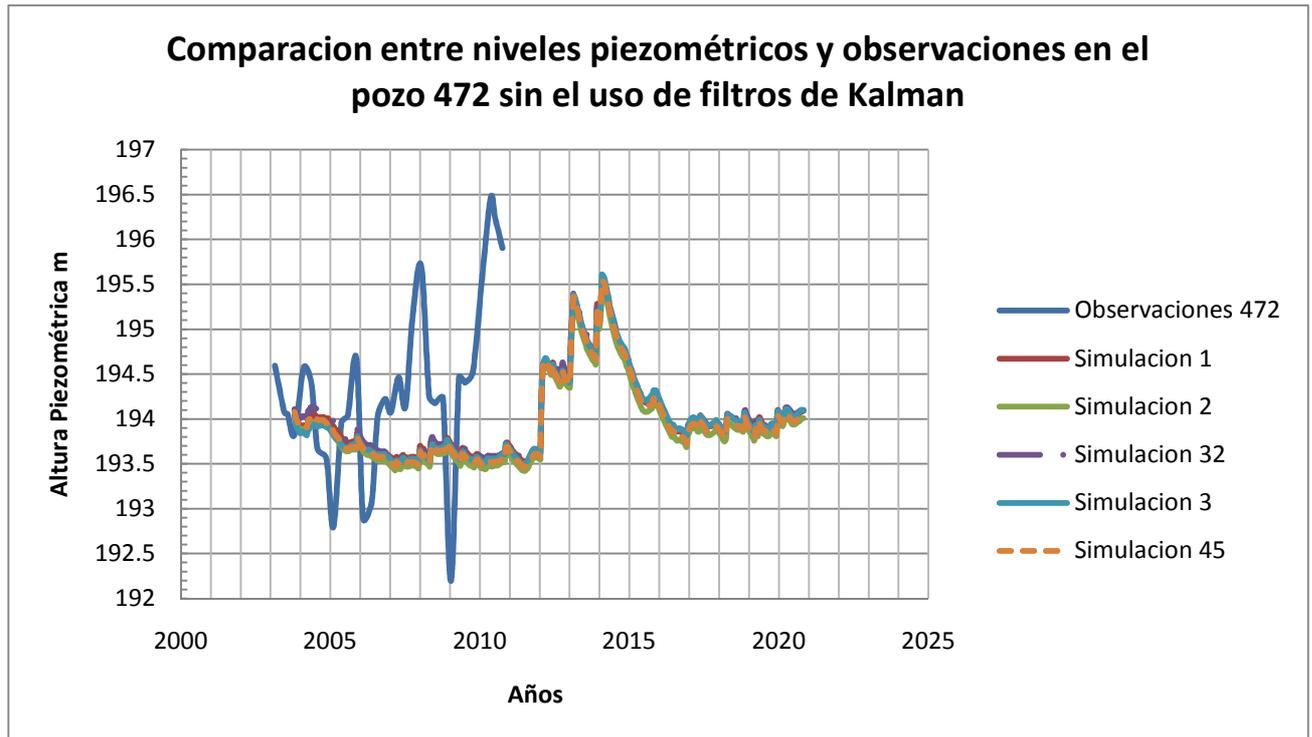


Figura 7.4.1.8 Alturas Piezométricas observadas en el pozo 472 comparadas con los valores que reproduce el CORE para el periodo 2003-2020.

## 8 Filtro de Kalman de conjuntos

### 8.1 Filtro de Kalman de conjuntos aplicado al FUA

Para medir la incertidumbre del modelo en la estimación de los niveles piezométricos se hizo uso de los filtros de Kalman. El programa creado se puede subdividir en tres pasos:

PASO 1. Modelo de predicción. La ecuación de flujo es resuelta en por medio de CORE, y que es representada por la ecuación:

$$Y_k = f(X_{k-1}, Y_{k-1}) \quad (8.1.1)$$

Donde  $Y_k$  es el estado del sistema (nivel piezométrico) en el tiempo  $t_k$ ,  $f$  representa el modelo de flujo (incluidas las condiciones de contorno, recargas y parámetros), y  $X_{k-1}$  denota los parámetros del modelo (conductividades hidráulicas) después de la última actualización del modelo. Por ultimo  $Y_{k-1}$  es el estado previo del sistema.

PASO 2. Análisis. Utilizando los datos de alturas piezométricas observadas se actualizan los parámetros del modelo.

1 Se construye el conjunto de vectores  $\Psi_k$ , que incluye los parámetros  $X$  y la predicción del estado del sistema  $Y$ . Este vector se puede dividirse en tantos miembros como realizaciones se tengan en el conjunto mediante:

$$\Psi_{k,j} = \begin{bmatrix} X \\ Y \end{bmatrix}_{k,j} \quad (8.1.2)$$

Siendo  $j$  el numero de miembros del conjunto en el tiempo  $t_k$ .

2 El conjunto de vectores es actualizado, realización por realización, asimilando las observaciones ( $Y_k^{obs}$ ):

$$\Psi_{k,j}^a = \Psi_{k,j}^e + G_k(\Psi_{k,j}^{obs} + \epsilon - H\Psi_{k,j}^e) \quad (8.1.3)$$

Donde el superíndice  $a$  y  $e$  denota análisis y estimación, respectivamente;  $\epsilon$  es un vector de error de observación aleatorio;  $H$  es un interpolador lineal que interpola la altura piezométrica estimada a la ubicación donde se encuentra la observación más cercana.

En el caso de  $H$  se utiliza una función que sigue la siguiente forma:

La ganancia del filtro de Kalman  $G_k$  esta dada por:

$$G_k = P_k^e H^T (H P_k^e H^T + R_k)^{-1} \quad (8.1.4)$$

Donde

$R_k$  es la matriz de covarianza de errores de medida.

$P_k^e$  contiene las covarianzas entre los diferentes componentes del vector de estado y puede ser estimado a partir de los resultados del conjunto estimado:

$$P_k^e \approx E \left[ (\Psi_{k,j}^e - \bar{\Psi}_{k,j}^e)(\Psi_{k,j}^e - \bar{\Psi}_{k,j}^e)^T \right] \\ \approx \sum_{j=1}^{N_r} \frac{(\Psi_{k,j}^e - \bar{\Psi}_{k,j}^e)(\Psi_{k,j}^e - \bar{\Psi}_{k,j}^e)^T}{N_r} \quad (8.1.5)$$

Donde  $N_r$  es el número de realizaciones en el conjunto, y la barra denota el promedio del conjunto.

PASO 3. El estado actualizado del sistema se convierte en el estado actual y el proceso comienza nuevamente.

## 8.2 Descripción del programa

El programa toma como input los archivos de salida de niveles piezométricos ROOT\_HH1.out y ROOT\_HH2.out provenientes de CORE, así como el conjunto de archivos ROOT\_NRE.inp donde ROOT es el nombre del problema que se está tratando, y NRE es el numero de realización del campo de conductividades que utiliza el archivo de entrada en CORE para resolver las ecuaciones de flujo.

Primeramente toma un archivo de texto con cada una de las realizaciones de conductividades hidráulicas del conjunto para modificar un archivo base, y crear un conjunto de archivos de entrada para CORE. Luego busca el primer valor de observaciones que se tenga en un archivo de texto que contiene los resultados de las campañas de medición de niveles piezométricos, enseguida compara la fecha de la observación con el intervalo de tiempo en CORE para modificar el conjunto de archivos de entrada de CORE y así, calcular hasta la fecha correcta. Se realiza una serie de cálculos en CORE para cada uno de los archivos de entrada. Esto genera un conjunto de conductividades hidráulicas, así como una serie de niveles piezométricos medidos en el intervalo temporal que se asignó a CORE para cada uno de los pozos de la red de vigilancia.

Luego se introduce el conjunto completo en la subrutina que crea los distintos vectores que utilizan los filtros de Kalman de conjuntos, se genera la ganancia del Filtro de Kalman mediante las matrices de covarianzas que utiliza la metodología de los filtros de Kalman de Conjuntos, se actualiza el sistema y posteriormente se lee nuevamente el archivo de observaciones para saber si existen más observaciones disponibles, de ser así entonces se actualizan los archivos de entrada a CORE con los nuevos valores de conductividades y se añade el intervalo de tiempo que se va a calcular para volver a resolver el sistema con CORE, esto se repite hasta que todas las observaciones sean tomadas en cuenta.

El diagrama de flujo que representa el proceso anterior se observa en la figura 8.2.1.

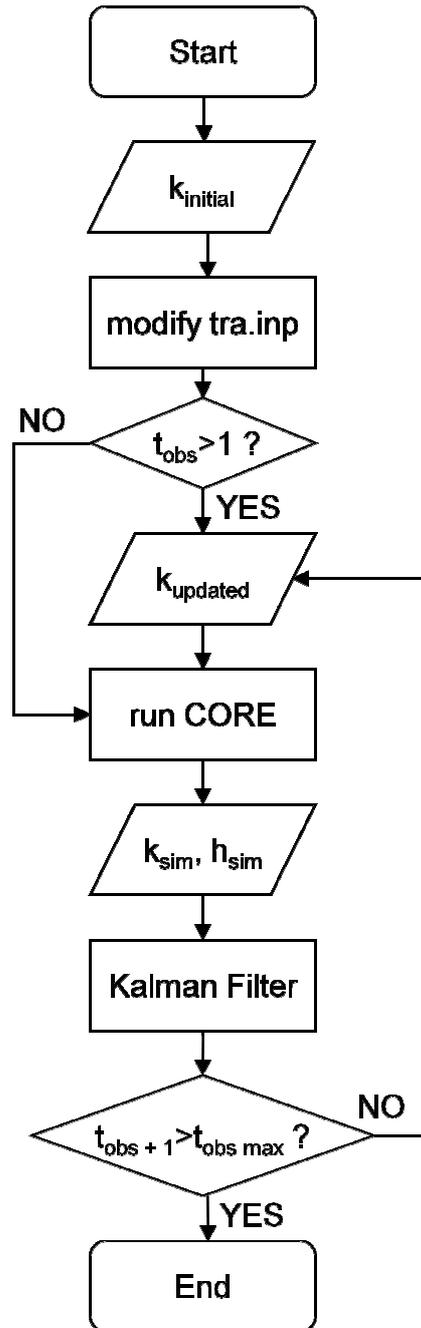


Figura 8.2.1 Diagrama de flujo del programa de Filtros de Kalman de Conjuntos añadido al código CORE.

## 9 Resultados y conclusiones

### 9.1 Resultados

Dentro de los resultados obtenidos se aprecia que el programa va ajustando las conductividades para obtener un mejor ajuste de manera global en el modelo. Las figuras 9.1.1 a 9.1.7 muestran el ajuste gradual para los pozos de control PC1, PC2, PC3, PC5, PC6, así como los pozos 505 y 611, para la realización 2. Para cada una de las graficas se muestran los resultados que va obteniendo el filtro para cada intervalo de tiempo desde el intervalo 7 hasta el 96 con un incremento de 6 pasos de tiempo. Hay que recordar que cada intervalo de tiempo es de 15 días, por lo que la simulación 7 es el día 105 y la 13 el 195 y así sucesivamente.

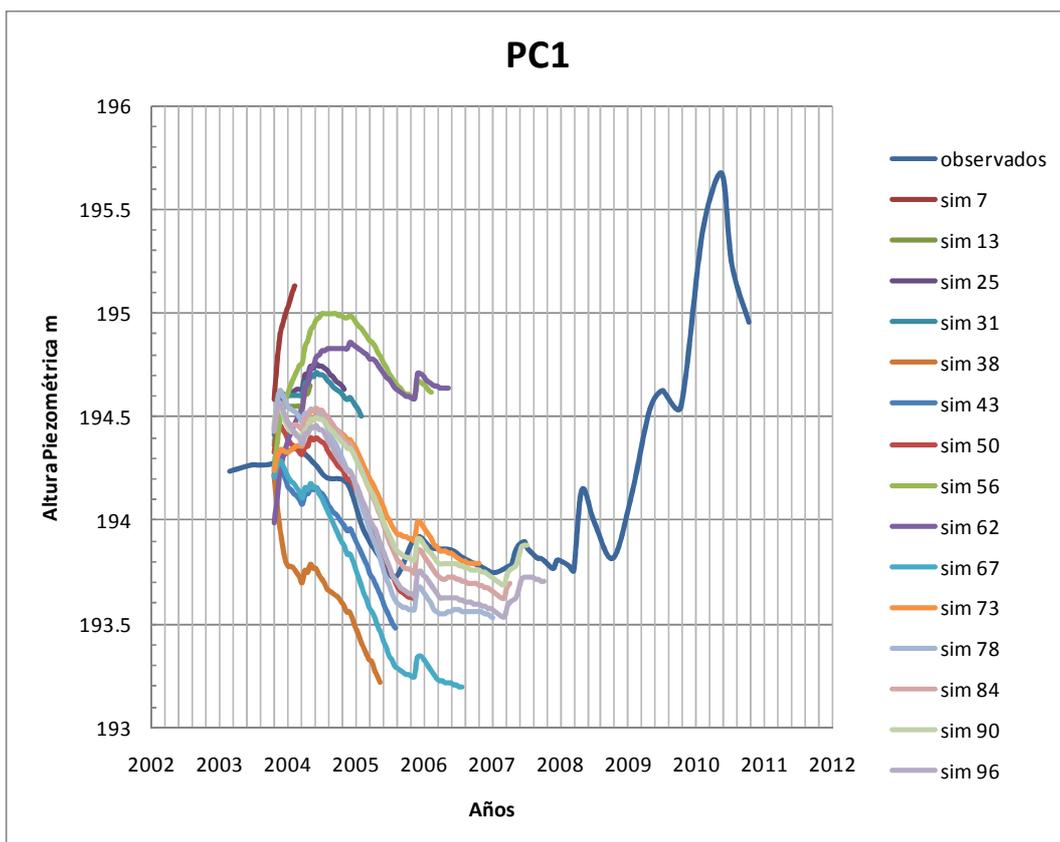


Figura 9.1.1 Ajuste de conductividades realizado mediante Filtros de Kalman de conjuntos para el PC1 realización 2.

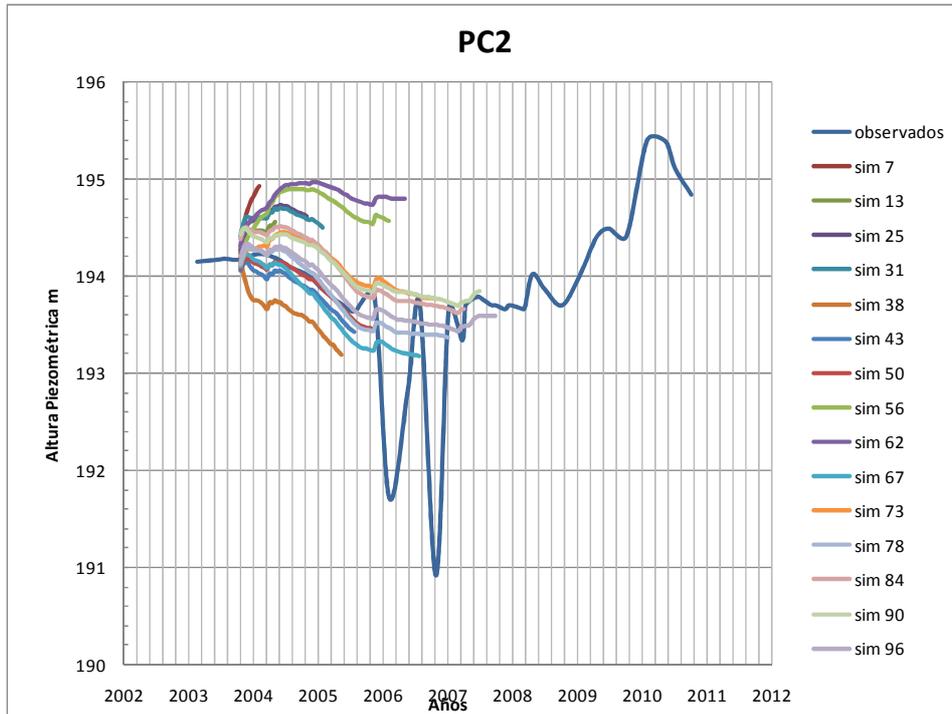


Figura 9.1.2 Ajuste de conductividades realizado mediante Filtros de Kalman de conjuntos para el PC2 realización 2.

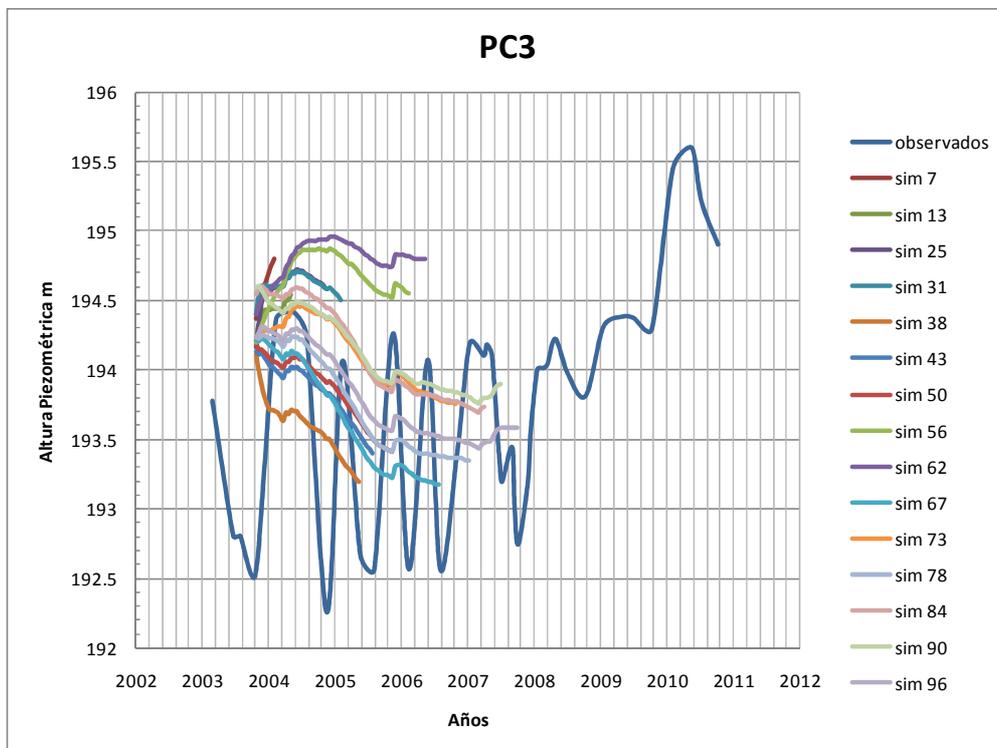


Figura 9.1.3 Ajuste de conductividades realizado mediante Filtros de Kalman de conjuntos para el PC3 realización 2.

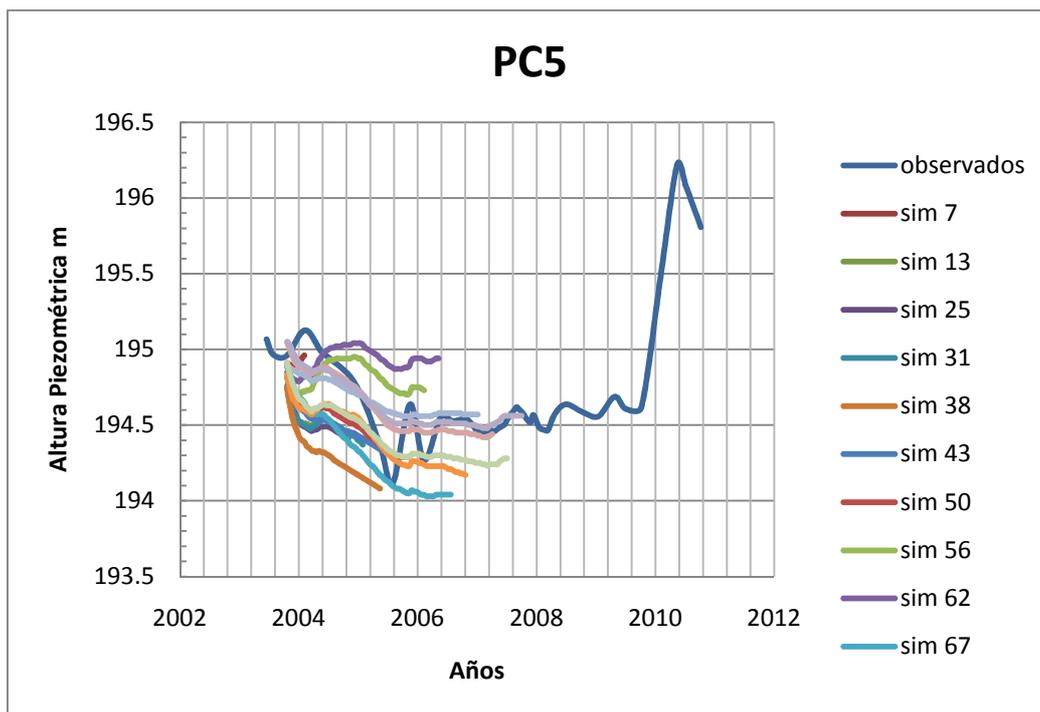


Figura 9.1.4 Ajuste de conductividades realizado mediante Filtros de Kalman de conjuntos para el PC5 realización 2.

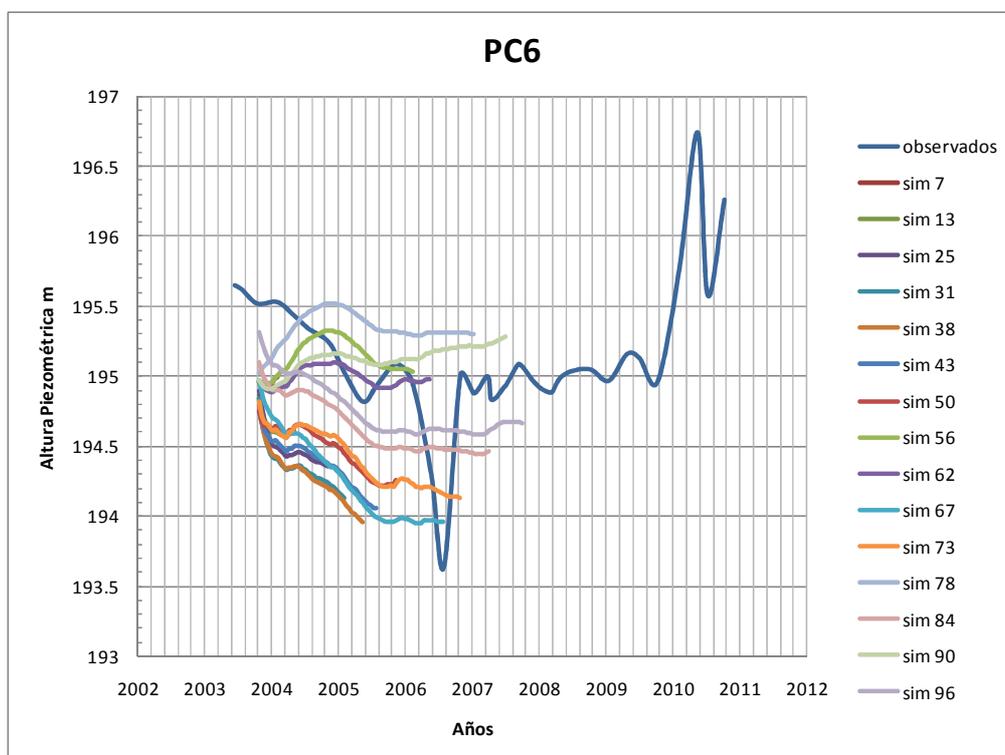


Figura 9.1.5 Ajuste de conductividades realizado mediante Filtros de Kalman de conjuntos para el PC6 realización 2.

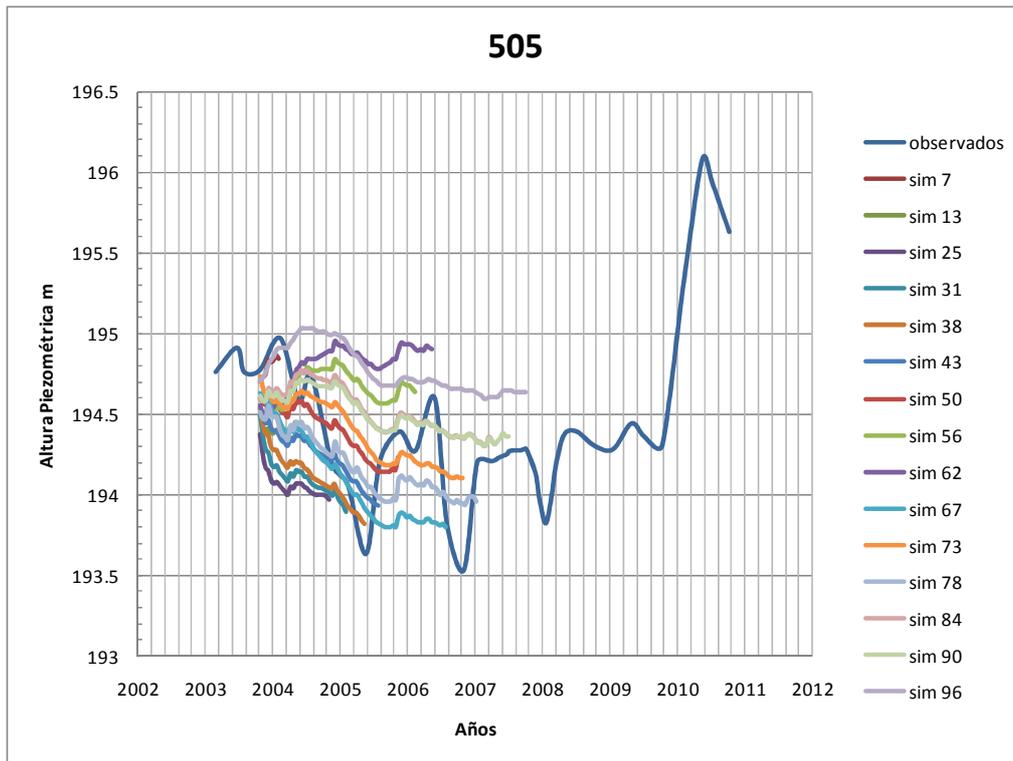


Figura 9.1.6 Ajuste de conductividades realizado mediante Filtros de Kalman de conjuntos para el pozo 505 realización 2.

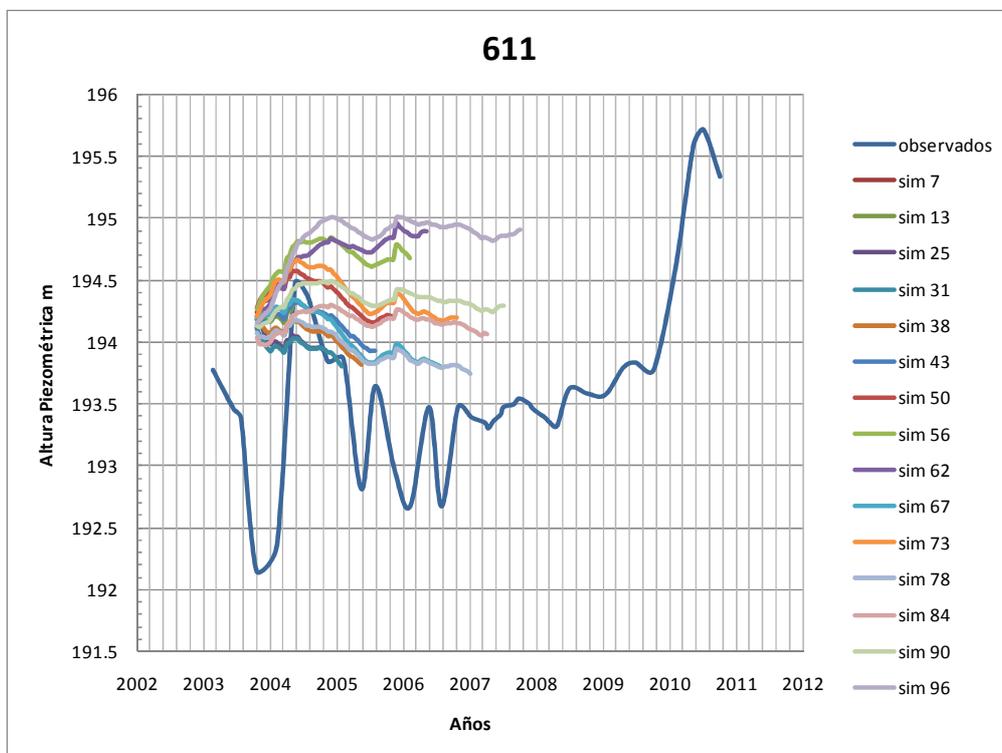
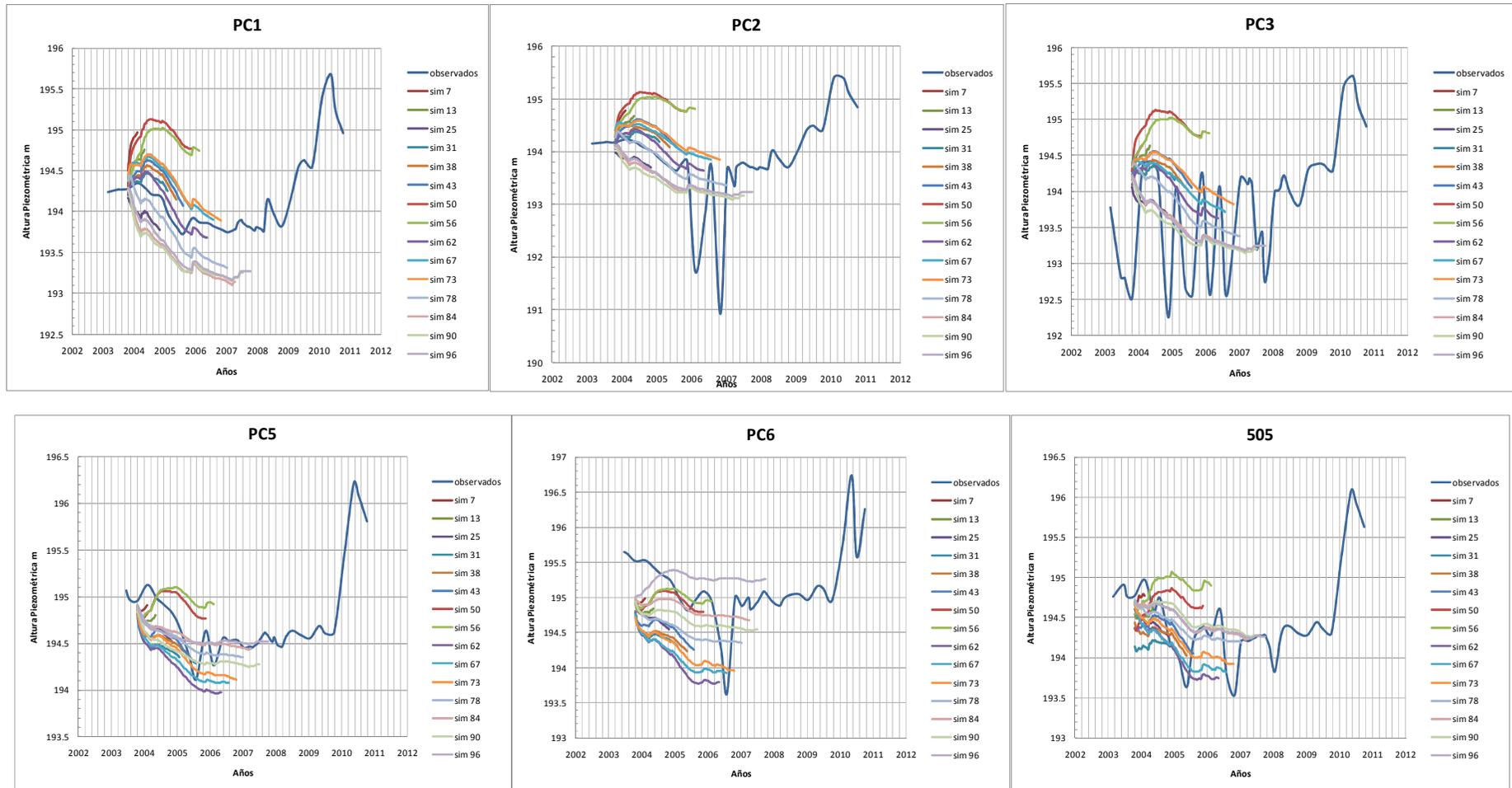
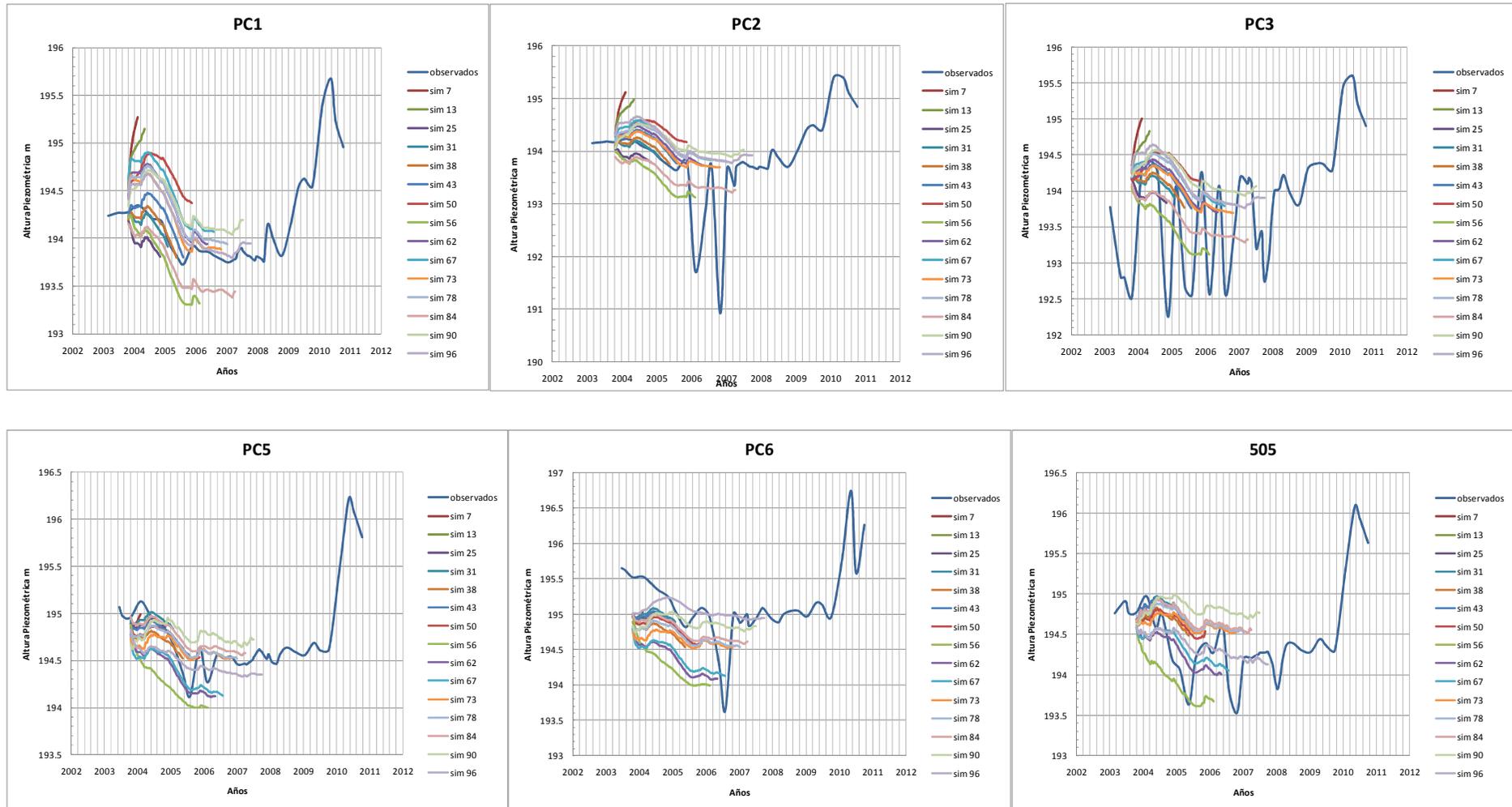


Figura 9.1.7 Ajuste de conductividades realizado mediante Filtros de Kalman de conjuntos para el pozo 611 realización 2.

En las siguientes figuras se detallan los mismos pozos de las figuras 9.1.1 a 9.1.6 pero para las realizaciones 3 y 9. Estas se muestran en las figuras 9.1.8 a 9.1.19



Figuras 9.1.8 a 9.1.13 Ajuste de conductividades realizado mediante Filtros de Kalman de conjuntos con la realización 2.



Figuras 9.1.14 a 9.1.19 Ajuste de conductividades realizado mediante Filtros de Kalman de conjuntos con la realización 9.

Se ha observado la gran variabilidad de los valores de conductividades en la zona central, esto se muestra en las figuras 9.1.20, 9.1.21 y 9.1.22.

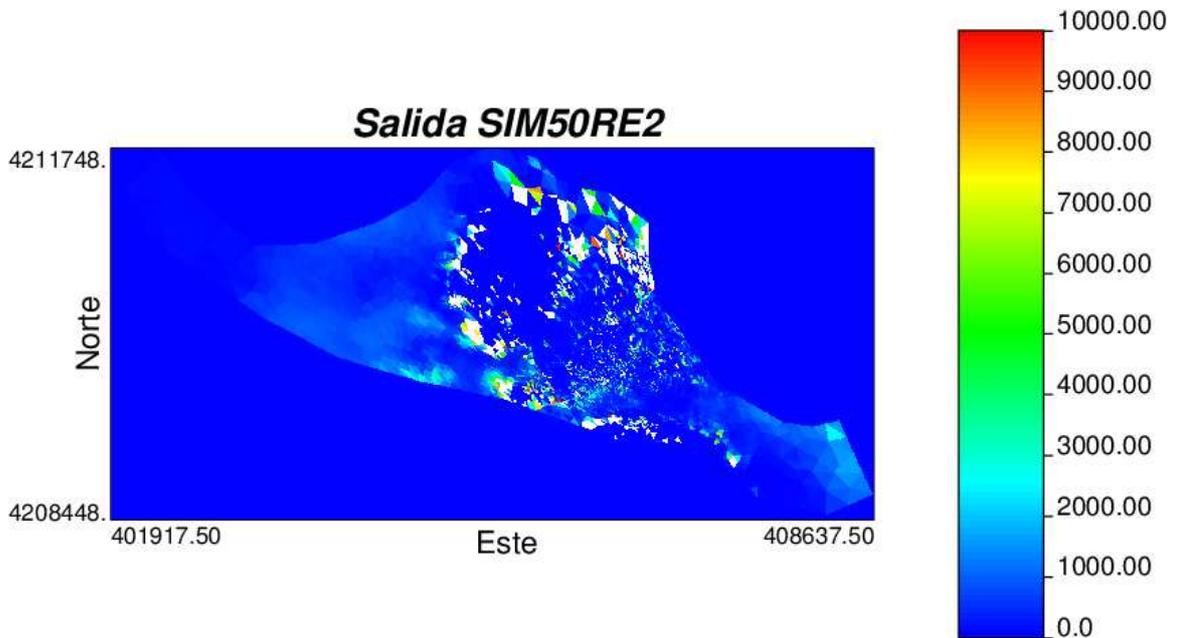


Figura 9.1.20 Campo de conductividades obtenido a partir de los filtros de Kalman de Conjuntos para la Realización 2.

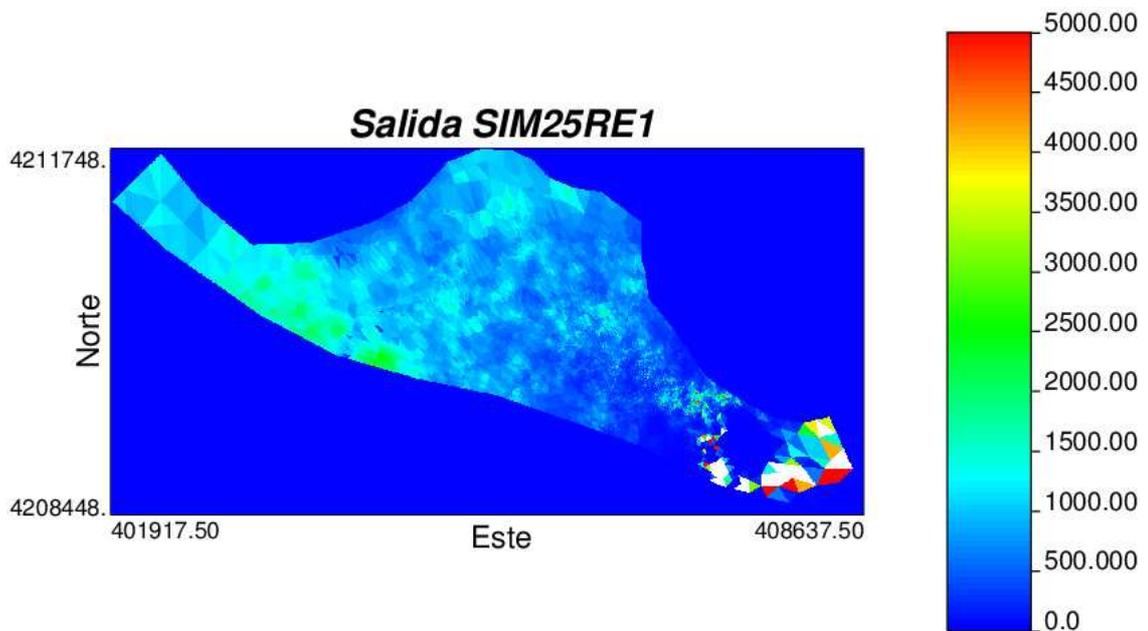


Figura 9.1.21 Campo de conductividades obtenido a partir de los filtros de Kalman de Conjuntos para la Realización 1.

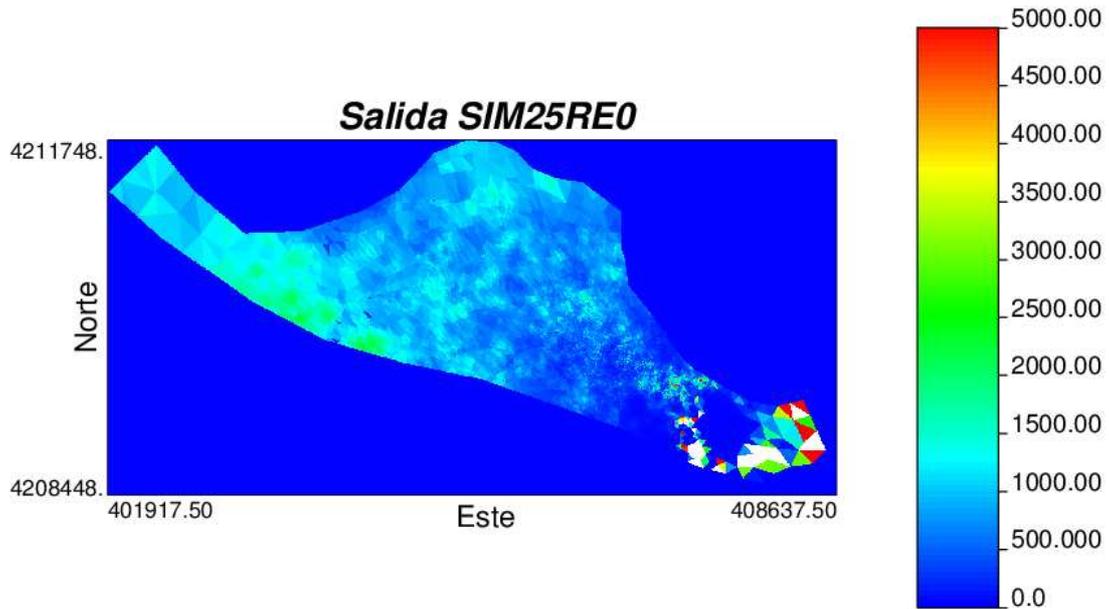


Figura 9.1.22 Campo de conductividades obtenido a partir de los filtros de Kalman de Conjuntos para la Realización 0.

Nótese que los campos de conductividades obtenidos al finalizar el Filtrado están numerados de 0 a 49. Los campos de color blanco son valores altos de conductividad que se van ajustando de manera iterativa.

Se observó también que en años posteriores a 2005 el uso de Filtros de Kalman no logra ajustar de manera adecuada las conductividades a los niveles observados. Esto puede deberse a que CORE utiliza 21 funciones de flujo discretizadas en intervalos de 15 días, siendo alguna de las 21 la utilizada en cada una de las zonas de material, por lo que parece ser que a partir de 2005 los niveles piezométricos no logran converger debido a la cantidad de flujo que se tiene, por lo que el filtrado tiende a desajustar los valores previamente ajustados para poder ajustarse a los valores piezométricos más recientes. No obstante se ha observado que los filtros de Kalman de Conjuntos logran asimilar el patrón de observaciones a medida que se avanza en la discretización temporal.

## 9.2 Análisis de sensibilidad

Para comprobar la sensibilidad que tiene el área de estudio se ha realizado una comparación entre las diferentes realizaciones tomando solamente las 10 primeras del conjunto. Se puede apreciar en las graficas la gran variabilidad que tienen los niveles piezométricos debido a los cambios en las conductividades. Las figuras 9.2.1 a 9.2.4 nos muestran estos cambios bruscos. Se observo que las simulaciones realizadas con los campos de conductividades de las realizaciones 6 y 7 no obtenían buenos resultados de ajuste, esto debido a que son campos de conductividades con cambios más bruscos que los otros campos introducidos al inicio del filtrado, y al no poder proyectar mas a futuro las simulaciones, no logran ajustarse bien a los valores de las alturas.

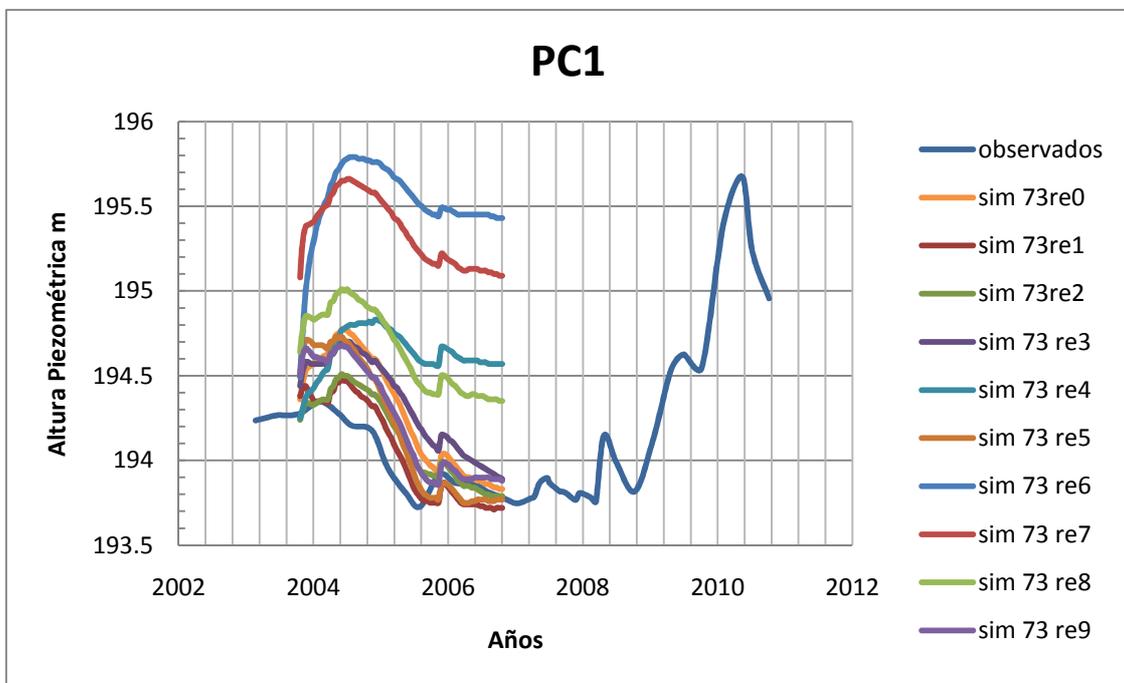


Figura 9.2.1 Comparación de alturas piezométricas obtenidas a partir de 10 campos de conductividades distintos para el PC1.

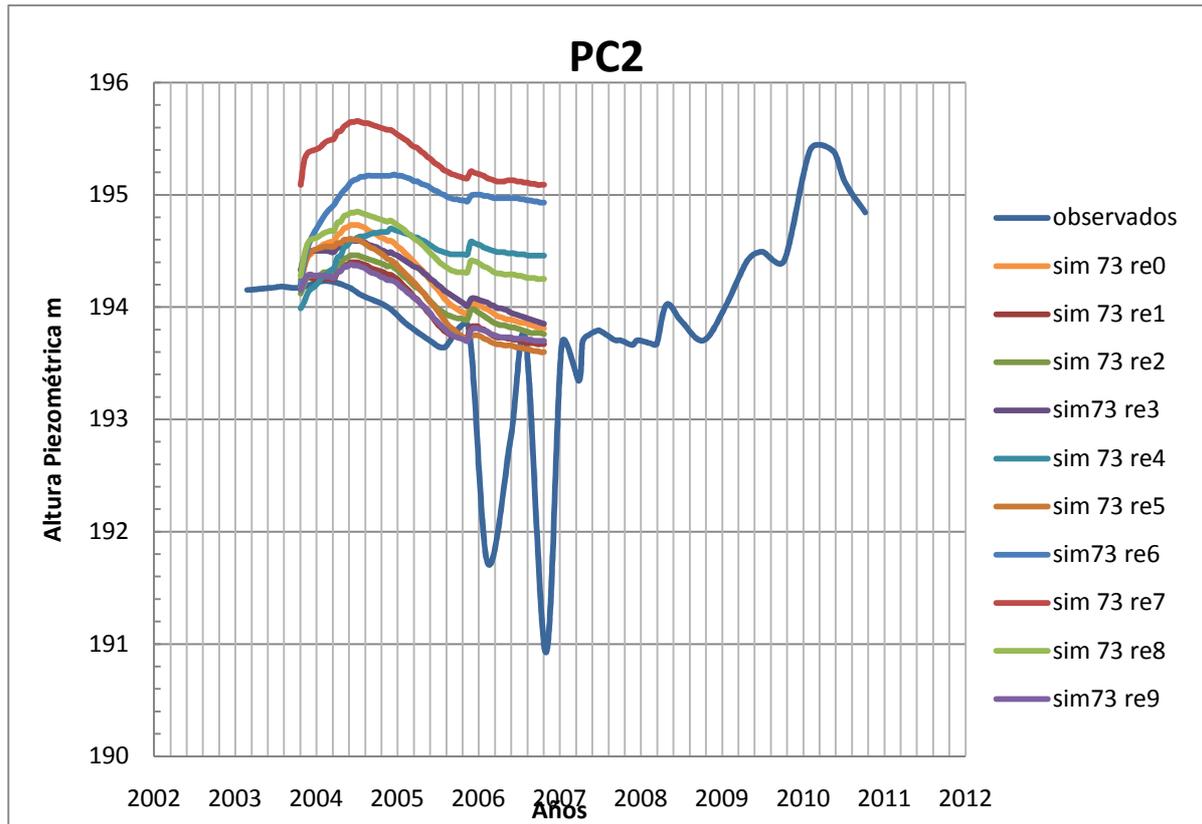


Figura 9.2.2 Comparación de alturas piezométricas obtenidas a partir de 10 campos de conductividades distintos para el PC2.

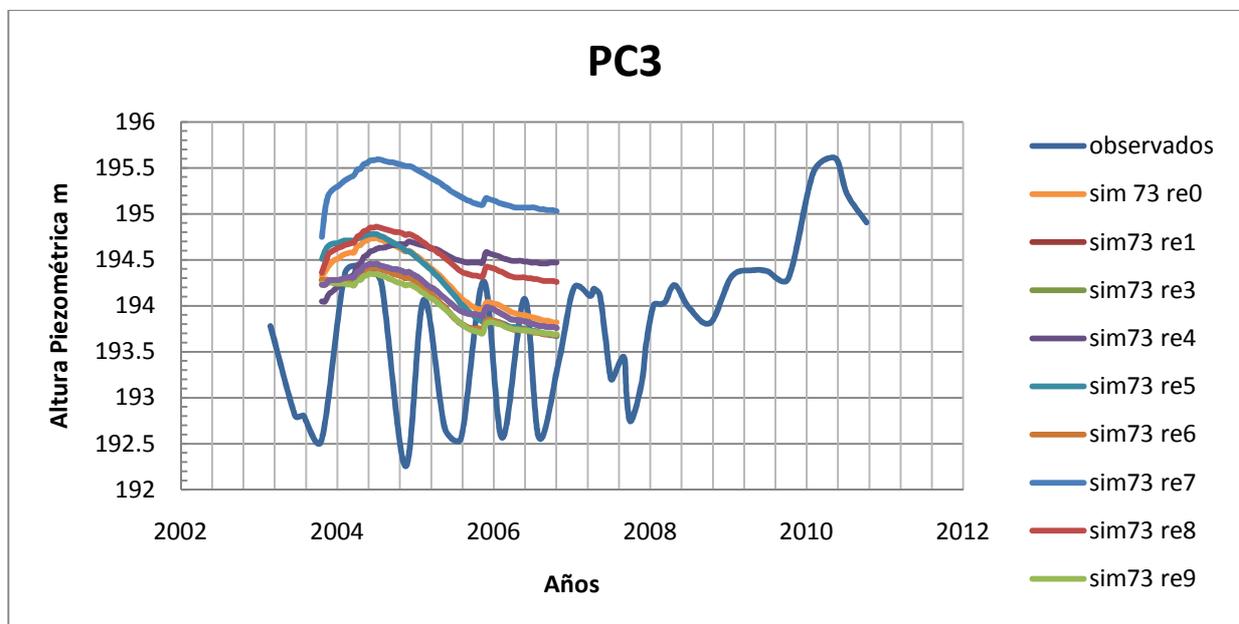


Figura 9.2.3 Comparación de alturas piezométricas obtenidas a partir de 10 campos de conductividades distintos para el PC3.

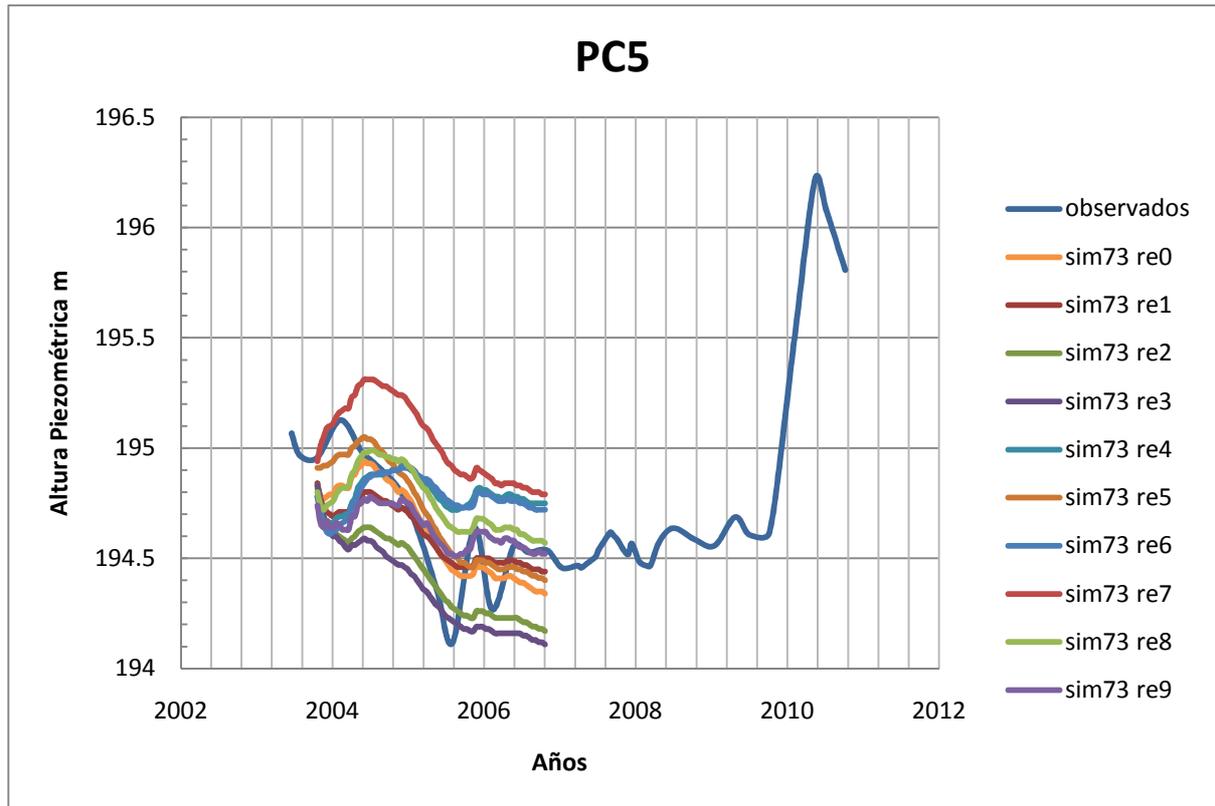


Figura 9.2.4 Comparación de alturas piezométricas obtenidas a partir de 10 campos de conductividades distintos para el PC5.



## 9.2 Análisis de incertidumbre

Se realizó un análisis de incertidumbre a partir de los 8 campos de conductividades que ajustan mejor los valores observados de la simulación 73, las conductividades hidráulicas para los pozos de control PC1 al PC6 se observan en las siguientes figuras, tomando como caso base el valor de conductividad que tenía la zona donde se encuentran cada uno de estos pozos en el modelo del FUA de 2004. La tablas 9.2.1 a nos muestran a que zona corresponde cada uno de los pozos de control, así como las conductividades correspondientes.

Pozo	Zona	Conductividad
PC1	2515	4.55E+01
PC2	2544	3.59E+02
PC3	3492	5.64E+01
PC4	2455	4.11E+03
PC5	2302	1.13E+02
PC6	6472	4.60E+06

Tabla 9.2.1 Pozos de control y zonas correspondientes para la realización 0.

Pozo	Zona	Conductividad
PC1	2515	4.70E+01
PC2	2544	3.70E+02
PC3	3492	6.88E+01
PC4	2455	4.08E+03
PC5	2302	2.90E+01
PC6	6472	5.00E+00

Tabla 9.2.2 Pozos de control y zonas correspondientes para la realización 1.

Pozo	Zona	Conductividad
PC1	2515	7.34E+01
PC2	2544	1.99E+02
PC3	3492	2.13E+01
PC4	2455	3.12E+03
PC5	2302	9.23E+01
PC6	6472	0.00E+00

Tabla 9.2.3 Pozos de control y zonas correspondientes para la realización 2.

Pozo	Zona	Conductividad
PC1	2515	2.12E+01
PC2	2544	3.74E+01
PC3	3492	6.31E+01
PC4	2455	5.56E+01
PC5	2302	6.48E+01
PC6	6472	2.11E+02

Tabla 9.2.4 Pozos de control y zonas correspondientes para la realización 3.

Pozo	Zona	Conductividad
PC1	2515	2.32E+01
PC2	2544	3.45E+04
PC3	3492	5.90E+01
PC4	2455	2.35E+04
PC5	2302	1.12E+02
PC6	6472	2.01E+05

Tabla 9.2.5 Pozos de control y zonas correspondientes para la realización 4.

Pozo	Zona	Conductividad
PC1	2515	1.73E+01
PC2	2544	1.43E+01
PC3	3492	1.07E+03
PC4	2455	5.00E+00
PC5	2302	1.86E+01
PC6	6472	3.20E+04

Tabla 9.2.6 Pozos de control y zonas correspondientes para la realización 5.

Pozo	Zona	Conductividad
PC1	2515	5.00E+00
PC2	2544	5.00E+00
PC3	3492	5.00E+00
PC4	2455	5.00E+00
PC5	2302	1.33E+01
PC6	6472	2.59E+03

Tabla 9.2.7 Pozos de control y zonas correspondientes para la realización 6.

Pozo	Zona	Conductividad
PC1	2515	2.43E+04
PC2	2544	8.76E+07
PC3	3492	1.22E+03
PC4	2455	5.00E+00
PC5	2302	5.10E+02
PC6	6472	7.91E+02

Tabla 9.2.8 Pozos de control y zonas correspondientes para la realización 7.

Pozo	Zona	Conductividad
PC1	2515	7.25E+00
PC2	2544	8.20E+01
PC3	3492	1.11E+02
PC4	2455	4.31E+05
PC5	2302	1.02E+02
PC6	6472	4.99E+06

Tabla 9.2.9 Pozos de control y zonas correspondientes para la realización 8.

Pozo	Zona	Conductividad
PC1	2515	4.21E+01
PC2	2544	4.07E+02
PC3	3492	3.16E+01
PC4	2455	1.93E+03
PC5	2302	5.52E+03
PC6	6472	0.00E+00

Tabla 9.2.10 Pozos de control y zonas correspondientes para la realización 9.

Se concluye que los valores observados y calculados en algunas zonas tienen una gran variabilidad, por lo que aunque el uso de Filtros de Kalman de Conjunto ajusta el modelo para que reproduzca los valores observados lo mejor posible, no logra ajustarlos de manera congruente, creando así una varianza significativa en las conductividades.

### 9.3 Conclusiones

El presente trabajo se ha realizado para implementar el filtro de Kalman de conjuntos a un caso real y medir la incertidumbre que presentan los niveles piezométricos al introducir variabilidad en la heterogeneidad del terreno, esto mediante una serie de realizaciones de distintos campos de conductividades.

El análisis geoestadístico corroboró las hipótesis acerca de la dirección y variabilidad espacial que se hicieron al crear los distintos modelos de la FUA. Se observó que, tal y como detalla Samper et. al (2004) se tienen bien definidas una serie de zonas con conductividades hidráulicas semejantes, y que, en su mayoría tienen un eje de máxima continuidad cercano a los 125°.

Se generaron una serie de 500 simulaciones las cuales siguen la misma tendencia que los datos originales.

Se observó que el solo hecho de crear campos de conductividades hidráulicas con variabilidad espacial a menor escala no es suficiente como para representar adecuadamente los niveles piezométricos observados utilizando el modelo de la FUA del año 2004. También se observó que, se tiene un cierto grado de incertidumbre al introducir distintos valores de conductividades que, si bien es cierto modifican las alturas piezométricas, aumentando o disminuyendo el nivel en ciertos puntos, no logran adecuarse a los valores observados pues la tendencia global no se ve alterada.

Se realizó un programa que mediante el uso de Filtros de Kalman de Conjuntos, asimila información en cada paso iterativo modificando las conductividades para reproducir mejor los valores observados. De esto cabe señalar dos aspectos:

Primero que los filtros de Kalman de conjuntos son una potente herramienta que ha demostrado su eficacia en distintos campos de aplicación, y que, como en este estudio logra reproducir de manera más eficiente el patrón observado al asimilar información reciente. En definitiva los niveles piezométricos son más parecidos con el uso de filtros de Kalman de conjuntos que sin ellos.

El segundo aspecto a señalar es que, aunque los filtros de Kalman ajustan las conductividades de manera iterativa para que reproduzcan mejor los valores observados, se tienen conductividades hidráulicas muy elevadas en algunos puntos. Esto debido a que los valores de las distintas funciones de flujo del modelo del FUA del año 2004 no son suficientes para representar los valores observados hasta el año 2010, siendo medianamente ajustable la piezometría hasta el año 2005. La calibración que se hizo en el año 2004 a partir de la información disponible desde 1977 hasta 2003 y con la que se realizó la predicción para el periodo 2004-2020 no es suficiente para poder ajustar de manera satisfactoria el nuevo conjunto de valores observados, por lo que será necesario una nueva calibración del modelo para este periodo.

Cabe señalar que, si bien no se logro ajustar las alturas piezométricas para el periodo completo de observaciones, si se observo la gran diferencia que supone el uso de los filtros de Kalman de conjuntos y la ventaja que se tiene al agregar información reciente.

Este trabajo también sirve de antesala para en un futuro próximo verificar las herramientas creadas para el cálculo de filtros de Kalman de conjuntos en un periodo anterior con funciones de flujo ajustadas. Posteriormente se incluirán las funciones de flujo a los parámetros de ajuste de los filtros de Kalman de conjuntos. Se extenderá el uso de las herramientas creadas, para poder ajustar las concentraciones en el modelo. Finalmente mediante las herramientas creadas es posible observar la inflación de la covarianza en el área de estudio, así como la localización de áreas con una variabilidad muy grande.

## 10 Bibliografía

Almendral-Vazquez A. Randi Syversveen A. (2006). The ensemble Kalman filter- theory and applications in oil industry. Norsk Regnesentral. SAND/05/06.

Chen Y., Zhang D. (2006). Data assimilation for transient flow in geologic formations via ensemble Kalman filter. *Advances in Water Resources* 29 (8), 1107-1122.

Cox H. (1964). On the estimation of state variables and parameters for noisy dynamic systems. *IEEE Transactions on Automatic Control* 9 (1), 5-12.

Deutsch C.V. , Journel A.G. (1998). *GSLIB Geostatistical software Library and Users's Guide*. S.E. Oxford University Press. New York.

ENRESA. (2004). Actualización del modelo de flujo y transporte de Andújar. Tomos 1 a 4.

Evensen G. (1994). Sequential data assimilation with a nonlinear quasigeostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* 99 (C5), 10, 143-10, 162.

Evensen G. (2003). The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics* 53, 343-367.

Gomez M. (2011). Uncertainty Analisis of groundwater flow and solute transport model FUA04. MSc Proyect, UPV.

Haugen et al. (2008). History matching using the ensemble Kalman filter on a North Sea field Case. *SPE Journal* 13 (4), 382-391.

Hendricks Franssen et al. (2011). Operational real-time modeling with ensemble Kalman filter of variably saturated subsurface flow including stream-aquifer interaction and parameter updating. *Water Resources Reserch*, Vol. 47 W02532, 20 PP., 2011

Hendricks Franssen, H., Kinselbach, W. (2008). Real-time groundwater flow modeling with the Ensemble Kalman Filter: Joint estimation of states and parameters and the filter inbreeding problem. *Water Resources Reserch* 44 (9), W09408.

Isaaks E., Srivastava R. (1989). An Introduction to applied Geostatistics. Oxford University Press. New York.

Kalman R. (1960). A new Approach to Linear Filtering and Prediction Problems, Transactions of the ASME-Journal of Basic Engineering, Vol. 82 , Series D, 35-45.

Keper J. , Sun X. , Steinle P. (2003). Atmospheric data assimilation using the Ensemble Kalman Filter at BMRC. Bureau of Meteorology Research Centre.

Kopp R., Orford R. (1963). Linear regression applied to system identification for adaptive control systems. AIAA Journal 1 (10), 2300-2306.

Li et al. (2011). Jointly Mapping Hydraulic Conductivity and Porosity by Assimilating Concentration Data via Ensemble Kalman Filter.

Naevdal et al. (2005). Reservoir monitoring and continuous model updating using ensemble Kalman filter. SPE Journal 10 (1), 66-74.

Pyrz M.J. Deutsch C.V. The whole story on the Hole Effect. Centre for computational Geostatistics, University of Alberta.

Remi N. (2001). GsTL: The Geostatistical Template Library in C++. A report submitted to the department of petroleum engineering of Stanford University for the Degree of Master of Science.

Remi N. (2004). Geostatistical Earth Modeling Software: User's Manual.

Samper et al. (1993). Calibración de los modelos de flujo y transporte de contaminantes en el agua subterránea.(Condición A.9.1, orden 1-2-1991). Informe elaborado para ENRESA.

Samper et al. (1988). Application of an automatic calibration technique to modelling an alluvial aquifer. IAHS Publ. 195. 1990.

Samper et al. (1991). Revision del modelo de Flujo y Transporte de contaminates en las aguas subterraneas del acuífero de la FUA. Informe CIMNE, IT-45

Samper et al. (2000). CORE 2D. A code for non-isothermal water flow and reactive solute transport. Users manual version 2. ENRESA.

Samper J. Pisani B. (1998). CORE\_LE\_2D V0 Users Manual. ENRESA. Barcelona

Schöninger A. Nowak W. Hendricks Franssen H.J. (2012). Parameter estimation by ensemble Kalman filters with transformed data: Approach and application to hydraulic tomography. *Water Resources Research*, Vol. 48, W04502, 18PP.

Skjervheim et al. (2007). Incorporating 4d seismic data in reservoir simulation models using ensemble Kalman filter. *SPE Journal* 12 (3), 282-292, 95789-PA.

Sun A., Morris A., Mohanty S. (2009). Sequential updating of multimodal hydrogeologic parameter fields using localization and clustering techniques. *Water Resources Research* 45, 15 PP.

Thulin K., Skaug H., Aanonsen S., Naevdal G. (2012). Dual Ensemble Kalman Filters. Submitted to: *Computers y Geoscience*. January 2012.

Zhou et al. (2011). Handling nongaussian distributions with Ensemble Kalman Filter. *Advances in Water Resources*. In press, doi:10.1016/j.advwatres.2011.04.014.

Zhou et al. (2011). Pattern Recognition in a bimodal Aquifer using the Normal-Score Ensemble kalman Filter. *Mathematical Geosciences*. Under review.