

enresa

Empresa Nacional de Residuos Radiactivos, S. A.

ENRESA -- 08/93.

PROGRAMAS COMPLEMENTARIOS PARA EL ANALISIS ESTOCASTICO DEL TRANSPORTE DE RADIONUCLEIDOS

PROGRAMAS COMPLEMENTARIOS PARA EL ANALISIS ESTOCASTICO DEL TRANSPORTE DE RADIONUCLEIDOS

J. Jaime Gómez Hernández

*Departamento de Ingeniería Hidráulica y Medio Ambiente
Universidad Politécnica de Valencia*

PROGRAMAS COMPLEMENTARIOS PARA EL ANALISIS ESTOCASTICO DEL TRANSPORTE DE RADIONUCLEIDOS

Este trabajo ha sido realizado bajo contrato con ENRESA.

Las conclusiones y puntos de vista expresados en el informe corresponden a sus autores y pueden no coincidir necesariamente con los de ENRESA.

INDICE

PREFACIO	1
1. INTRODUCCION.....	5
2. GENERALIDADES	9
2.1 Distribución	11
2.2 Convenciones comunes	11
3. ESTIMACION	15
3.1 Introducción.....	17
3.2 Krigeaje indicador	17
3.2.1 Programa IK3D	19
3.2.2 Programa POSTIK.....	21
3.3 Cokrigeaje.....	24
3.3.1 Programa COKB3D	25
4. SIMULACION	29
4.1 Introducción.....	31
4.2 Simulación de campos multinormales	31
4.2.1 Programa GCOSIM3D	31
4.3 Simulación secuencial de indicadores	34
4.3.1 Programa ISIM3D	36
5. EJEMPLO DE APLICACION	41
5.1 Ejemplo de COKB3D	45
5.2 Ejemplo de IK3D y POSTIK	46
5.3 Ejemplo de GCOSIM3D	47
5.4 Ejemplo de ISIM3D	47
6. BIBLIOGRAFIA.....	77
APENDICES	81
APENDICE A. Diagrama de flujo del programa IK3D	83
APENDICE B. Diagrama de flujo del programa COKB3D	89
APENDICE C. Diagrama de flujo del programa GCOSIM3D.....	93
APENDICE D. Diagrama de flujo del programa ISIM3D	97

PREFACIO

Este informe presenta el primer grupo de programas complementarios al análisis estocástico del transporte de radionucleidos. Pertenece a la serie de informes que sobre teoría y práctica de la geoestadística están siendo publicados dentro de la colección de informes técnicos de ENRESA como producto del proyecto de investigación y desarrollo "*Uso del krigaje de variables indicadoras para la elaboración de mapas y generación de campos de permeabilidad*" (Publicaciones Técnicas 04/91 y 13/92).

La utilización de métodos estocásticos en el análisis del transporte de radionucleidos requiere de una serie de herramientas que permitan:

1. Analizar la variabilidad espacial de los datos con el fin de construir un modelo de incertidumbre de los parámetros de los modelos de flujo y transporte (por ejemplo permeabilidad) en los puntos donde se carece de los mismos.
2. Asignar valores de los parámetros en los puntos donde se carece de medidas. Dependiendo de si esta asignación se hace pretendiendo minimizar una función de pérdidas local o pretendiendo minimizar una función de pérdidas que dependa globalmente de los valores del parámetro en todos los puntos del área en estudio, esta asignación se efectúa utilizando técnicas de estimación en el primer caso y técnicas de simulación en el segundo.
3. Tener en cuenta la heterogeneidad en los datos que es consecuencia de la diferente escala en

que se toman medidas del mismo parámetro. Este problema se acusa especialmente en las permeabilidades y requiere un método riguroso que permita pasar a la misma escala todas las medidas disponibles.

En la actualidad, existen una serie de paquetes de dominio público que permiten el análisis de la variabilidad espacial de los parámetros y la estimación de los mismos por técnicas paramétricas.

Este informe presenta un conjunto de programas que complementan, así como suplementan, esos paquetes, poniendo al alcance de todos las herramientas necesarias para la aplicación de técnicas no paramétricas, tanto para la estimación como para la simulación.

En un informe adicional se acompañará un programa para el tratamiento riguroso del problema del cambio de escala en permeabilidades.

Parte de los programas aquí presentados, concretamente aquellos que se refieren a la estimación de parámetros, han sido desarrollados en la Universidad de Stanford y se encuentran en la librería conocida como GSLIB (Deutsch y Journel, 1992), las versiones de los programas de estimación que se presentan en este informe han sido corregidas por el autor. El resto han sido desarrollados por el autor, en parte durante su estancia en la Universidad de Stanford y en parte en la Universidad Politécnica de Valencia bajo contrato con ENRESA.

1

INTRODUCCION

Los programas que a continuación se presentan son el complemento informático de la metodología descrita en el informe "Geoestadística para el análisis de riesgos: Una introducción a la geoestadística no paramétrica" (Gómez Hernández, 1991) y permitirán al lector interesado el llevar a cabo un análisis de riesgo utilizando técnicas paramétricas y no paramétricas.

Los programas presentados se agrupan en dos bloques que comprenden: (1) estimación de una variable mediante krigeaje indicador y estimación de una o varias variables por cokrigeaje y (2) simulación de una o varias variables, de acuerdo a un modelo estocástico multivariado multigaussiano, y simulación de una variable de acuerdo a un modelo bivariado no gaussiano, definido por una serie de variogramas indicadores.

Estos programas permiten generar campos del parámetro en estudio, condicionales o no a los datos, asumiendo que el soporte de los datos coincide con el tamaño de la malla sobre la que se están generando los resultados. En el caso en que el parámetro en estudio sea la conductividad hidráulica (o la transmisividad) el tamaño de la malla sobre la que se resuelve el problema de flujo puede ser variable y, en general, es mayor que el soporte de los datos, por lo que antes de proceder a la solución de los problemas de flujo y transporte hay que realizar un cambio de escala en las conductividades. Este cambio de escala, en general, no se limita a tomar un valor medio de las conductividades dentro del bloque y requiere tener en cuenta el fenómeno físico del flujo en un medio poroso.

Este informe incluye una serie de programas que permitirán hacer uso de los nuevos conceptos desarrollados en geoestadística en los últimos años, particularmente en lo que se refiere a la geoestadística no paramétrica. No se incluyen en este informe programas para el análisis exploratorio de los datos (univariado o multivariado) ni para el cálculo y modelización de variogramas, ni siquiera programas para realizar el krigeaje "tradicional" (krigeaje "simple" y krigeaje "ordinario") por existir una serie de programas en el dominio público que permiten llevar a cabo esas tareas (Froidevaux, 1990; Englund and Sparks, 1988) y que pueden obtenerse simplemente escribiendo directamente a los autores a la dirección que se indica en las referencias. Por su parte, el Departamento de Ciencias de la Tierra Aplicada de la Universidad de Stanford viene desarrollando programas ilustrando la aplicación de los distintos algoritmos que allí se desarrollan. Estos programas están en el dominio público y son el resultado de muchos años de trabajo de los estudiantes que por ese departamento han pasado.

Recientemente, estos programas han sido homogeneizados, proveyéndolos con una entrada de datos y una salida de resultados coherente a todos los programas y han sido recientemente publicados junto con un manual de uso (Deutsch y Journel, 1992).

Hay disquetes con los programas a los que se refiere este trabajo. Para solicitarlos, si es de interés del lector, utilizar la hoja incluida al final de este volumen.

2

GENERALIDADES

2.1 Distribución

En el disquete adjunto a este informe se acompaña una copia del código fuente de los programas que más adelante se describen, así como de los datos de entrada de los ejemplos que se utilizan en el texto. El código fuente es una mezcla de programas escritos en Fortran 77 y en C, y en la medida de lo posible su escritura se ha adherido al estándar ANSI. Se acompañan asimismo copias ejecutables de todos los programas. Algunos de estos programas, particularmente los que llevan a cabo simulación, puede que no funcionen en un ordenador personal con menos de 2 megaoctetos de memoria, dependiendo del tamaño del área a simular.

2.2 Convenciones comunes

Entrada de datos

Todos los programas requieren un fichero de entrada con los datos que se van a utilizar como valores de partida, para realizar, bien la estimación, bien la simulación del parámetro en estudio. En todos los casos se supone que la estimación o la simulación se va a llevar a cabo sobre una malla regular cubriendo el área de interés. Para el caso de la simulación utilizando el método de las variables indicadoras, puede necesitarse un segundo archivo de entrada de datos para los datos imprecisos, como se indica más adelante.

Todos los programas requieren, asimismo, uno o más ficheros de entrada con los parámetros definiendo la malla y los variogramas a utilizar.

Salida de datos

La salida de datos consiste siempre de un fichero con los valores estimados/simulados distribuidos sobre la malla definida en el archivo de entrada de datos. Puede haber también un número de archivos intermedios de salida y siempre hay también un archivo con información complementaria sobre el desarrollo de la ejecución del programa.

Puesto que la salida de datos es siempre sobre una malla regular, se han omitido las coordenadas de cada punto. Los valores de la malla se escriben con la convención que se indica más abajo.

Definición de la malla

Aunque los algoritmos utilizados tanto para estimación como para simulación pueden utilizarse sobre mallas irregulares, el hecho de que cada punto estimado/simulado deba tener el mismo soporte que los

datos (todos los cuales deben tener el mismo soporte a su vez) hace recomendable que se utilice una malla regular para que impida la interpretación de los valores generados como teniendo un soporte irregular y con un tamaño de alguna manera inversamente proporcional a la densidad de valores generados alrededor del mismo.

Todos los programas han sido escritos para ser utilizados en tres dimensiones. Evidentemente, estos programas pueden utilizarse en una o dos dimensiones simplemente definiendo una malla en tres dimensiones en la que se restringe su tamaño a una única capa o una única fila.

La definición de la malla regular sobre la que se generan los resultados requiere indicar las coordenadas del centro de la primera celda de la malla. La distancia entre centros de celda en las direcciones x , y y z y el número de celdas en cada una de estas direcciones.

En la figura 1 se muestran los parámetros necesarios para la definición de la malla. Nótese que el nodo (1,1,1) está en la esquina inferior izquierda del nivel más bajo y no en el superior derecho del nivel más alto como es común en otros programas.

Puesto que los archivos de salida en todos los casos contienen los valores estimados o simulados sobre una malla regular, se decide escribirlos sin indicar las coordenadas de cada punto y siguiendo la convención de escribirlos con el índice iz siendo el que varía más lentamente, después el índice iy y finalmente el índice ix . Es decir, refiriéndonos a la figura 1, los resultados se escriben por capas ($iz=1, \dots, n_z$), cada capa se escribe por filas ($iy=1, \dots, n_y$) y finalmente cada fila se escribe por columnas ($ix=1, \dots, n_x$).

Definición de los variogramas

Antes de proceder a estimar/simular la variable sobre la malla hay que construir un modelo estocástico que describa la variabilidad espacial, es decir, hay que calcular y modelar una serie de variogramas. Los programas que aquí se presentan aceptan como modelos de variogramas cualquier variograma válido en tres dimensiones, compuesto por un efecto pepita y hasta un máximo de tres estructuras imbricadas de los variogramas más comunes: el esférico, el exponencial, el gaussiano o el monomial.

Todos los variogramas pueden ser anisótropos y los ejes principales de anisotropía no necesitan estar alineados con los ejes cartesianos que definen la malla. Para definir estos modelos hay que indicar la orientación del elipsoide de anisotropía y los alcances correspondientes en las direcciones principales de dicho

elipsoide. La expresión de este variograma es:

$$\gamma(h) = C_0 + \sum_{i=1}^3 C_i \gamma_i(h') \quad (1)$$

donde $\gamma(h)$ es el variograma, h es el vector para el que se calcula el variograma, C_0 es el efecto pepita, C_i es la meseta correspondiente a cada uno de los variogramas imbricados y $\gamma_i(h')$ es el valor del variograma tipo de meseta unidad y alcance unitario, siendo h' la distancia reducida asociada con el vector h y que se calcula de acuerdo a la orientación de los ejes de anisotropía y los alcances en las direcciones principales del elipsoide de anisotropía según la expresión siguiente:

$$h' = \sqrt{h_x'^2 + h_y'^2 + h_z'^2}$$

donde

$$\begin{Bmatrix} h'_x \\ h'_y \\ h'_z \end{Bmatrix} = \begin{Bmatrix} 1/a_x & 0 & 0 \\ 0 & 1/a_y & 0 \\ 0 & 0 & 1/a_z \end{Bmatrix} \cdot \begin{Bmatrix} \cos X'X \cos X'Y \cos X'Z \\ \cos Y'X \cos Y'Y \cos Y'Z \\ \cos Z'X \cos Z'Y \cos Z'Z \end{Bmatrix} \cdot \begin{Bmatrix} h_x \\ h_y \\ h_z \end{Bmatrix} \quad (2)$$

donde a_x , a_y y a_z son los alcances en las direcciones principales del elipsoide de anisotropía y los elementos de la segunda matriz son los cosenos de los ángulos formados entre los ejes principales de anisotropía $X'Y'Z'$ y los ejes cartesianos XYZ .

Los modelos que pueden utilizarse en la ecuación (1) son los que se indican a continuación con sus respectivas expresiones.

1. Esférico

$$\gamma(h) = Sph(h) = \begin{cases} \frac{1}{2} \left(3\frac{h}{a} - \frac{h^3}{a^3} \right) & \text{si } h \leq a \\ 1 & \text{si } h > a \end{cases} \quad (3)$$

el parámetro a es el alcance, es decir la distancia a la que se alcanza la meseta

2. Exponencial

$$\gamma(h) = Exp(h) = 1 - \exp\left(-\frac{h}{a}\right)$$

el modelo exponencial alcanza la meseta asintóticamente, $3a$ corresponde al alcance práctico que es aquél en el que el variograma alcanza el 95% de su meseta.

3. Gausiano

$$\gamma(h) = Gau(h) = 1 - \exp\left(-\frac{h^2}{a^2}\right)$$

similarmente al modelo exponencial, el modelo gausiano alcanza su meseta asintóticamente, $\sqrt{3}a$ corresponde al alcance práctico, es decir, aquél en el que el variograma alcanza el 95% de su meseta

4. Monomial

$$\gamma(h) = Mon(h) = h^a$$

el modelo monomial no posee meseta por lo que tiene un grado de variabilidad que aumenta con la distancia, está íntimamente relacionado con medios fractales y puede demostrarse que el exponente a es una función lineal de la dimensión fractal del fenómeno en estudio. Para que el modelo monomial sea un modelo válido de un variograma se requiere que el parámetro a esté comprendido entre 0 y 2.

La matriz de rotación en la ecuación (2) puede especificarse bien explícitamente, proporcionando los valores de los nueve coeficientes que la componen, o bien indicando los ángulos de giro que, sucesivamente aplicados, transforman los ejes cartesianos en los ejes principales del elipsoide de anisotropía.

Cualquier rotación en tres dimensiones puede expresarse como el producto de tres rotaciones consecutivas que transformarán el sistema de ejes cartesianos original XYZ en el sistema de ejes de anisotropía $X'Y'Z'$:

$$XYZ \xrightarrow{\alpha_1} X^{(1)} Y^{(1)} Z^{(1)} \xrightarrow{\alpha_2} X^{(2)} Y^{(2)} Z^{(2)} \xrightarrow{\alpha_3} X'Y'Z'$$

la primera rotación es de α_1 grados en sentido dextrógiro alrededor del eje Z hasta conseguir situar el eje y en el plano vertical de la dirección de alcance máximo, (en otras palabras, si el eje Y se supone alineado con el Norte geográfico, el ángulo α_1 es el acimut del plano vertical que contiene la dirección de máxima anisotropía), la segunda rotación es de α_2 grados en sentido levógiro alrededor del eje $X^{(1)}$, (esta rotación tiene por objeto transformar el eje $Y^{(1)}$ en el eje $Y^{(2)}$ alineado con la dirección de máxima anisotropía) y, finalmente, la tercera rotación es de α_3 grados en sentido dextrógiro alrededor del eje $Y^{(2)}$ (una vez alineado $Y^{(2)}$ con el eje de máxima anisotropía se hace un giro alrededor del mismo con el fin de situar los dos ejes restantes en la dirección de las otras dos direcciones principales de anisotropía).

Por razones históricas, en los programas de estimación, los alcances en las direcciones principales de los ejes de anisotropía no se especifican como tales, sino que los programas requieren el valor del alcance

en la dirección de alcance máximo y después la fracción que los alcances en las otras direcciones representan con respecto al valor máximo. Concretamente, después de las tres rotaciones, el eje Y' está alineado con la dirección principal de máxima anisotropía y

los ejes X' y Z' están alineados con los otros dos ejes de anisotropía. Los programas requieren el alcance en la dirección Y' y los cocientes entre los alcances en las direcciones X' y Z' y el alcance en la dirección Y' , siempre en este orden.

3

ESTIMACION

3.1 Introducción

A continuación se describen los programas que permiten interpolar, a partir de los datos, el valor de la variable de interés en los puntos de una malla regular, utilizando la técnica no paramétrica de krigeaje indicador o la técnica de cokrigeaje "tradicional".

3.2 Krigeaje indicador

El krigeaje indicador se utiliza para la estimación de funciones de distribución condicionales a los datos. Para llevar a cabo tal estimación es necesario discretizar el rango de variabilidad de la variable en estudio en K clases mediante $K-1$ umbrales. La estimación $i^*(x; z_k)$ de la variable indicadora en el punto x para el umbral z_k proporciona, no una estimación del valor que la variable indicadora puede tomar en x , es decir, 0 ó 1, sino el estimador por mínimos cuadrados de la probabilidad acumulada condicionada de que la variable en estudio esté por debajo del umbral z_k .

La estimación de los valores de la variable indicadora para todos los umbrales $z_k, k=1, \dots, K$ en el punto x proporciona una estimación de la función de probabilidad condicional acumulada. A partir de esta función de probabilidad, el usuario puede decidir entre elegir un estimador en cada punto de acuerdo a cierto criterio, por ejemplo, puede elegir el valor medio de la distribución condicional o el valor mediano, o un determinado decil de la misma. También se pueden extraer probabilidades de excedencia de un cierto valor, resultando en mapas de probabilidad de excedencia o mapas de riesgo.

El krigeaje indicador se lleva a cabo en dos fases, en la primera fase se estiman las distribuciones condicionales acumuladas y en la segunda fase se procesan éstas para proporcionar bien un estimador, bien una probabilidad asociada con un determinado umbral.

La estimación de las probabilidades condicionadas acumuladas se lleva a cabo mediante el programa **IK3D** y el proceso posterior con el programa **POSTIK**.

Una de las ventajas del krigeaje indicador con respecto a los métodos de krigeaje más tradicionales es su posibilidad de incluir información imprecisa o "blanda". Por este motivo los datos de partida pueden proporcionarse en dos formatos distintos. Por una parte se puede dar el valor numérico del dato para que el programa realice internamente la transformación en las correspondientes variables indicadoras o

bien se pueden dar directamente los valores de las variables indicadoras. Es este segundo caso el que incluye la posibilidad de incorporar información "blanda" al proceso de estimación.

Los datos "blandos" que pueden incluirse en el proceso de estimación son de dos tipos, intervalos y distribuciones de probabilidad a priori.

Sirva como ejemplo la figura 2 en la que se indica cómo hay que codificar tanto los datos "duros" como los "blandos" en la correspondiente serie de valores de variables indicadoras. Supongamos que se ha decidido utilizar como umbrales para el cálculo de las distribuciones de probabilidad condicionada los valores $\{z_i, i=1, \dots, 9\} = \{1, 10, 20, 50, 70, 80, 90, 95, 99\}$ y que los tres datos que se desea codificar directamente en las correspondientes variables indicadoras son: dato "duro" $z = 25$, dato de intervalo $z \in [30, 85]$ y dato de distribución de probabilidad condicional a priori dado por los siguientes valores $\{P(Z \leq z_i), i=1, \dots, 9\} = \{0.0, 0.0, 0.3, 0.5, 0.7, 1.0, 1.0, 1.0, 1.0\}$. Para la transformación de estos datos en series de indicadores sólo hay que aplicar la definición de variable indicadora o su interpretación probabilística.

El programa de krigeaje indicador proporciona tres opciones para el cálculo de las distribuciones condicionales de probabilidad: krigeaje indicador simple, krigeaje indicador ordinario y krigeaje indicador mediano.

Con krigeaje indicador simple, la estimación de la función de probabilidad condicionada se hace mediante krigeaje simple de las variables indicadoras. Este caso supone que las variables indicadoras son estacionarias dentro del área en estudio y su valor esperado (que es igual a la proporción de valores por debajo del umbral considerado) es conocido y se mantiene constante dentro del área de estudio. Krigeaje indicador simple requiere, como datos, los valores de la distribución de probabilidad marginal para los umbrales elegidos.

Con krigeaje indicador ordinario, la estimación de la función de probabilidad condicionada se hace usando krigeaje ordinario. Si además, se restringe el conjunto de datos a utilizar en la estimación a aquellos dentro de un determinado entorno alrededor del punto a estimar, el uso del krigeaje ordinario equivale a reestimar el valor medio de las variables indicadoras en cada entorno a partir de los datos dentro del mismo. De este modo, se pueden reflejar pequeñas desviaciones de los datos con respecto a la hipótesis de estacionariedad. Krigeaje indicador ordinario no requiere especificar los valores de la distribución de probabilidad marginal.

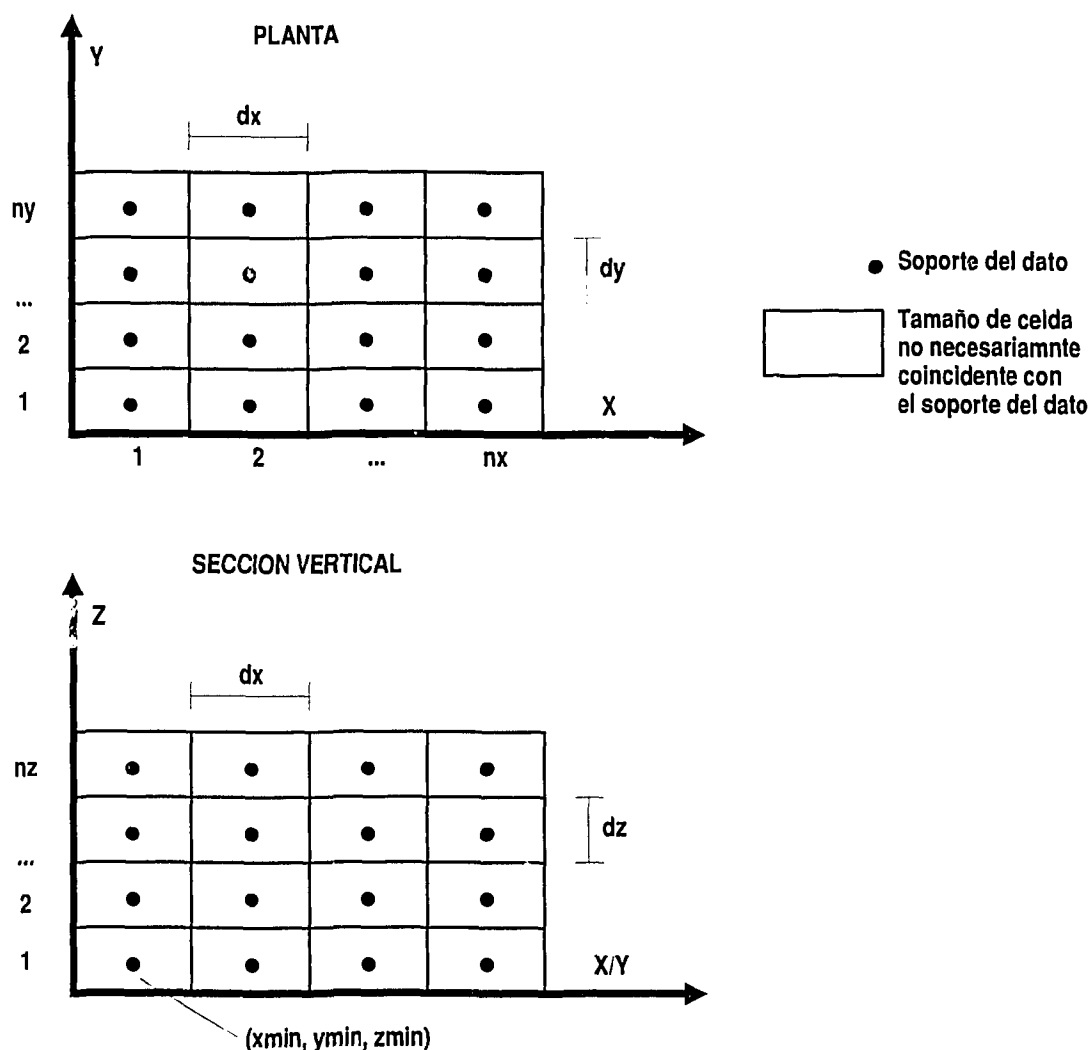


Figura 1

Definición de la malla. Los parámetros que definen la malla son las coordenadas (x_{min} , y_{min} , z_{min}) del centro de la celda con índices (1,1,1), el paso de malla (dx , dy , dz) y el número de celdas en cada dirección (nx , ny , nz)

```

25    → {0,0,0,1,1,1,1,1,1}
[30,85] → {0,0,0,?, ?, ?, 1,1,1}
a priori pdf → {0.0,0.0,0.3,0.5,0.7,1.0,1.0,1.0,1.0}

```

Figura 2

Ejemplo de codificación de información en series, posiblemente incompletas, de variables indicadoras. Los interrogantes corresponden a umbrales para los que no se dispone información y que deberán ser estimados.

Krigeaje indicador mediano corresponde a elegir un modelo de función aleatoria conocido como modelo "mosaico" (Journel, 1984). Este modelo tiene la "ventaja" de que, como el modelo multigaussiano, está completamente determinado por la función de distribución marginal y la covarianza del atributo. En el modelo mosaico los correlogramas de las variables indicadoras son todos iguales entre sí e iguales al correlograma de la variable. Los variogramas de todas las variables indicadoras tienen por tanto la misma forma aunque distinta meseta. El único correlograma que hay que estimar puede hacerse a partir de la variable original, o mejor, a partir de la variable indicadora considerando la mediana como umbral. El hecho de que todos los variogramas indicadores sean el mismo hace que el método sea muy rápido puesto que sólo es necesario resolver un sistema de krigeaje indicador.

Contrario al modelo multigaussiano, el modelo "mosaico" no implica la pérdida de correlación de los valores extremos, puesto que todas las variables indicadoras tienen la misma correlación. Cuando el número de datos es pequeño, como para no permitir inferir más allá de la covarianza, el uso de un modelo multigaussiano, o un modelo mosaico, es una decisión que el modelizador debe tomar en base al comportamiento conocido de los momentos de orden superior de ambos modelos.

3.2.1 Programa IK3D

El programa IK3D para el cálculo de las distribuciones de probabilidad condicionada tiene tres archivos de entrada y dos de salida. El diagrama de flujo del programa se muestra en el apéndice A.

Los archivos con los datos de entrada están organizados por columnas, donde cada columna corresponde a una variable distinta. El archivo de entrada con los datos "duros" contiene un primer registro con un título, el segundo registro tiene un entero indicando el número de variables, **nvar** (organizadas en columnas) en el archivo, siguen **nvar** registros con los nombres de las variables y finalmente un número de registros indeterminado con los datos, cada registro con **nvar** valores correspondientes a las **nvar** variables. Como mínimo el archivo debe contener, en cualquier orden, tantas columnas como coordenadas (2 ó 3 según sea un problema en dos o en tres dimensiones) más una columna con la variable. El ejemplo de archivo de entrada con los datos "duros" para el caso práctico de la base de datos Berea se muestra en la figura 12. Los archivos de entrada de datos pueden contener varias variables. En caso de que en cierta

posición se carezca de valor para alguna de las variables, se debe asignar un valor fuera del rango de variabilidad de las variables (positivo o negativo de gran valor absoluto). En cada registro después de los **nvar** valores puede añadirse un comentario, por ejemplo, indicando el nombre del pozo donde se ha tomado la medida.

Aun cuando el archivo de entrada puede contener sólo dos coordenadas para facilitar la entrada de datos en el caso en que éstos estén ubicados en un plano bidimensional, el programa siempre realiza una estimación en tres dimensiones, el usuario deberá indicar en el archivo de parámetros la dimensionalidad del problema dando los correspondientes valores a los parámetros **nx**, **ny**, y **nz** (véase 2.2 para la definición de estos parámetros), asimismo, si la tercera coordenada no se define, todos los datos se suponen con dicha coordenada igual a la correspondiente coordenada del origen de la malla **xmin**, **ymin** o **zmin**.

El archivo de entrada de datos "blandos" tiene la misma estructura que el de datos "duros", pero en este caso el número de variables debe ser igual a tres más el número de umbrales. Cada registro debe contener las tres coordenadas seguidas de los valores de las variables indicadoras para los umbrales seleccionados en orden ascendente. Las variables indicadoras desconocidas se indican con un valor negativo muy pequeño (menor que el parámetro **tmin** en el archivo de parámetros). La figura 25 muestra un ejemplo de entrada de datos blandos donde cada registro contiene un último campo sin uso y que no es referenciado en la lista de variables al comienzo del archivo.

Los parámetros definiendo la geometría del área a estimar, tipo de krigeaje indicador a utilizar, definición de variogramas indicadores, etc., están contenidos en un archivo de parámetros como se explica a continuación.

Los parámetros dentro del archivo de parámetros no comienzan hasta después de la línea que dice, literalmente y en mayúsculas, COMIENZO DE PARÁMETROS. A partir de esta línea hay que indicar:

1. Nombre del fichero con el archivo de entrada de datos "duros" (menos de 40 caracteres). Nota. Los 40 primeros caracteres de este registro son interpretados como el nombre del archivo, por lo que no debe ponerse ningún comentario en esa línea antes de la columna 41, puesto que pasaría a formar parte del nombre del archivo.
2. **icolx**, **icoly**, **icolz**, **icolvar**, valores enteros que indican las columnas dentro del archivo de datos "duros" que contienen las coordenadas *x*, *y* y *z* de los datos y la columna conteniendo la va-

riable. En caso de que los datos se encuentren en un plano y el archivo sólo contenga dos de las tres coordenadas, la coordenada que falta puede indicarse introduciendo en este registro el valor 0 para la coordenada sin especificar.

3. Nombre del fichero de datos "blandos" (menos de 40 caracteres). Este fichero debe contener las tres coordenadas y los valores de las variables indicadoras. Por una ligera inconsistencia, aun cuando los datos sean bidimensionales, la tercera coordenada debe especificarse igualmente. Aun cuando no se utilicen datos "blandos" este fichero debe contener un valor.
4. **tmin**, **tmax**, sólo los datos con valores contenidos entre **tmin** y **tmax** son considerados en el análisis. En los datos de intervalo en el fichero de datos "blandos" los indicadores a estimar se especifican con un valor por debajo de **tmin**.
5. Nombre del fichero de salida de resultados.
6. Nivel de depuración. Un entero que puede tomar los valores 0, 1, 2 ó 3 dependiendo de la cantidad de información intermedia que se desea imprimir. Lo normal es situar el nivel de depuración a 0, a no ser que se suponga que se ha detectado algún error en el programa y se desea conocer los pasos intermedios.
7. Nombre del fichero de salida con la información del análisis de depuración.
8. **nx**, **xmin**, **deltax**, véase sección 2 y figura 1.
9. **ny**, **ymin**, **deltay**
10. **nz**, **zmin**, **deltaz**, en el caso que el archivo de datos duros sólo contuviera dos coordenadas y se desee hacer una estimación en el plano de los datos (estimación bidimensional), el valor de **nx**, **ny** o **nz** correspondiente a la coordenada sin especificar debe valer 1 y, automáticamente, todos los datos toman como valor de la coordenada sin especificar el correspondiente a **xmin**, **ymin** o **zmin**. Si también se incluye un fichero de datos blandos y éstos están colocados en el mismo plano que los datos duros, el valor que debe darse a la tercera coordenada que no se especificó en el fichero de datos duros debe ser el correspondiente **xmin**, **ymin** o **zmin**.
11. **mindat**, **maxdat**, sólo si en el entorno de búsqueda se encuentran al menos **mindat** datos se llevará a cabo la estimación de la probabilidad condicional; si se encuentran más de **maxdat**, sólo los **maxdat** valores más próximos al punto

a estimar son utilizados. En caso que se encuentren menos de **mindat** datos, se escribe un mensaje de aviso en el archivo de depuración y el valor estimado que se escribe en el archivo de salida es un valor negativo muy alto.

12. **radio**, máxima distancia en cualquier dirección que se usa para localizar los datos a utilizar en el krigeaje indicador
13. **ang1**, **ang2**, **ang3**, **anis1**, **anis2**, la esfera (o círculo) definido en el registro anterior como entorno de búsqueda se gira y distorsiona para constituir un elipsoide de búsqueda. Los ángulos **ang1**, **ang2** y **ang3** tienen el mismo significado que los ángulos α_1 , α_2 , α_3 , definidos en la sección 2.2. Tras la rotación, el semieje en la dirección Y' mantiene el valor **radio** del registro anterior, el semieje en la dirección X' se reduce en el factor **anis1** y el semieje en la dirección Z' se reduce en el factor **anis2**.
14. **noct**, máximo número de datos que se retienen dentro del entorno de búsqueda por octante. Si **noct** = 0 entonces no se utiliza la búsqueda por octante y se retienen los **maxdat** más próximos independientemente del octante al que pertenecen. Si **noct** \neq 0 se retienen hasta **maxdat** dentro del entorno pero restringiendo el máximo número de datos por octante igual a **noct**.
15. **mik**, **mikcut**, si **mik** es igual a 0 se realiza krigeaje indicador completo con un variograma distinto para cada variable indicadora. Si **mik** es igual a 1, entonces se realiza krigeaje indicador mediano, usando para todas las variables indicadoras, el variograma de la variable indicadora correspondiente al umbral más próximo a **mikcut**.
16. **ktipo**, si **ktipo**=1, utilizar krigeaje ordinario, si **ktipo**=0, utilizar krigeaje simple.
17. **numb**, número de umbrales.

Los registros siguientes hay que introducirlos para todas las variables indicadoras, incluso cuando se utiliza krigeaje indicador mediano. En el caso de krigeaje indicador mediano, sólo uno de los variogramas es utilizado, aunque es necesario completar, con algún valor, los registros de los umbrales que no se usan. El orden debe ser ascendente en el valor del umbral.

1. **umbral**, **cdf**, **nest**, **c0**, umbral que define la variable indicadora, valor de la distribución marginal acumulada para este umbral, número de estructuras imbricadas que definen el variograma

(no contando el efecto pepita), valor del efecto pepita.

Los siguientes dos registros se repiten para cada una de las *nest* estructuras imbricadas indicadas en el registro anterior.

2. **vtipo**, **alcance**, **meseta**, tipo de variograma, 1 para esférico, 2 para exponencial, los modelos gaussiano y monomial (**vtipo** = 3 ó 4), no son permitidos como modelos válidos para variables indicadoras; **alcance** en la dirección de máxima continuidad (la que tras la rotación especificada en el registro siguiente queda alineada con el eje Y' , véase sección 2.2 y valor de la meseta de esta estructura.
3. **ang1**, **ang2**, **ang3**, **anis1**, **anis2**, ángulos que definen la rotación de los ejes cartesianos XYZ para transformarlos en ejes paralelos a los ejes de anisotropía, según se explica en 2.2, y factores de anisotropía de los alcances en la dirección de los ejes X' y Z' , en ese orden, con respecto del alcance en la dirección Y' .

En la figura 24 se muestra un ejemplo de archivo de parámetros para **IK3D**.

De los dos archivos de salida, uno contiene información sobre la ejecución del programa y el otro la salida de datos propiamente dicha.

La cantidad de información sobre la ejecución del programa que se imprime en el archivo de depuración depende del nivel de depuración, indicado en el registro 6 del archivo de parámetros. Para el nivel cero, sólo contiene una copia de los parámetros de entrada y un informe sobre el número y magnitud de las relaciones de orden violadas, y para el nivel tres, un informe detallado de los sistemas de ecuaciones resueltos, las soluciones de dichos sistemas y las funciones de probabilidad condicional obtenidas.

El archivo de salida tiene el mismo formato que los ficheros de entrada de datos descritos en 3.2.1: un primer registro con un título, el siguiente registro es el número de umbrales en que se discretiza la función de probabilidad condicionada, siguen tantos registros como umbrales con un nombre describiendo cada umbral y, a continuación, tantos registros como puntos donde se ha estimado la función de distribución condicional, cada registro con tantos campos como umbrales con la probabilidad que la variable esté por debajo del umbral correspondiente. El orden en que estos registros están escritos fue descrito en 2.2

3.2.2 Programa POSTIK

El archivo de salida de **IK3D** contiene las distribuciones de probabilidad condicional en cada nudo de la malla dados los datos disponibles (tanto "duros" como "blandos"). A partir de estas funciones de probabilidad el usuario debe decidir cómo procesar esta información para obtener un estimador que minimice una cierta función de pérdida o una probabilidad para un cierto umbral para la elaboración de un mapa de riesgo.

Este procesamiento de los resultados de salida de **IK3D** se lleva a cabo mediante el programa **POSTIK**. El programa **POSTIK** es relativamente limitado, puesto que sólo permite la obtención de dos tipos de estimadores: el valor esperado condicional y el cuantil q correspondiente a un determinado nivel de probabilidad p , además de la probabilidad condicional p correspondiente a un determinado umbral.

El valor esperado condicional es el "mejor" estimador para el caso en que la función de pérdida es directamente proporcional al cuadrado del error. El cuantil correspondiente a $p = 0.5$, es decir la mediana, es el "mejor" estimador para una función de pérdida que es directamente proporcional al valor absoluto del error, la elección de otros cuantiles es óptima para ciertas funciones de pérdida no simétricas (véase Journel, 1983 y Isaaks, 1990, para una discusión detallada sobre distintas funciones de pérdida).

El archivo de salida de **IK3D** sólo proporciona los valores de la función de probabilidad condicionada correspondientes a los umbrales seleccionados. Es por ello que para la determinación de ciertos cuantiles sea necesario interpolar la función de distribución entre los valores calculados y, más importante aún, extrapolar estas funciones de distribución por debajo del umbral más pequeño y por encima del más grande. Estas extrapolaciones pueden tener un impacto muy alto en el cálculo de valores esperados o de cuantiles (y probabilidades) en las colas de la distribución.

El programa **POSTIK** proporciona un número de opciones para llevar a cabo dichas interpolaciones/extrapolaciones.

Interpolación

Dados dos umbrales $z_{k-1} < z_k$ a los que les corresponde sendas estimaciones de la función de probabilidad condicionada $F(z_{k-1}) < F(z_k)$, la interpolación de la función de probabilidad para los valores de z en el intervalo $[z_{k-1}, z_k]$ puede hacerse mediante cual-

quier función monótonicamente creciente. En concreto, puede hacerse mediante una función lineal que equivale a suponer que la distribución de z dentro de la clase tiene una distribución uniforme. Como alternativa se ofrece la posibilidad de utilizar el modelo polinómico:

$$F_{\omega, [z_{k-1}, z_k]}(z) = F(z_{k-1}) + (F(z_k) - F(z_{k-1})) \left(\frac{z - z_{k-1}}{z_k - z_{k-1}} \right)^\omega, \\ \forall z \in [z_{k-1}, z_k], \omega > 0 \quad (4)$$

Este modelo varía en función del parámetro ω como se indica en la figura 3. Para $\omega = 1$ la interpolación es lineal, si $\omega > 1$ la interpolación tiene un sesgo negativo y si $\omega < 1$ la interpolación tiene un sesgo positivo.

Extrapolación

La extrapolación de los valores de $F(z)$ por encima del valor correspondiente al umbral máximo $F(z_k)$ requiere primero establecer el valor para el cual $F(z)=1$, es decir, el valor máximo del rango de variabilidad de z . Si éste es un valor finito z_{\max} , entonces se puede elegir el modelo polinómico anterior (3.2.2) para interpolar entre $F(z_k)$ y $F(z_{\max}) = 1$ con un cierto parámetro ω . Si el límite superior del rango de variabilidad es infinito, entonces se puede utilizar un modelo hiperbólico:

$$F_{\omega, [z_k, \infty]}(z) = 1 - (1 - F(z_k)) \left(\frac{z_k}{z} \right)^\omega, \quad \forall z \in [z_k, \infty]$$

el parámetro ω en este caso controla la rapidez de la convergencia de la función de extrapolación a su asíntota $F(\infty) = 1$ (figura 4). Cuanto más pequeño es ω mayor es la cola de la distribución, por tanto, mayor es la probabilidad de encontrar valores mucho mayores que el umbral máximo.

Para extrapolación por debajo del mínimo umbral sólo se da la opción de utilizar un modelo de tipo monómico entre $F(z_{\min}) = 0$ y $F(z_1)$.

Como archivos de entrada se utilizan el archivo resultante de **IK3D** y un archivo de parámetros con los siguientes valores a partir de la línea, después de aquella que comienza literalmente por **COMIENZO DE PARAMETROS**.

1. Fichero de entrada (que es un fichero de salida de **IK3D** con las probabilidades condicionadas).
2. Fichero de salida con los resultados (éstos dependen del tipo de salida que se desee).

3. **ipar, par**. Si **ipar=1**, se calcula el valor esperado de la distribución de probabilidad condicionada; si **ipar=2**, se calcula la probabilidad de estar por debajo de **par** y los valores medios por encima y por debajo de **par**; si **ipar=3**, se calcula el valor de z que corresponde a una probabilidad condicional acumulada = **a par**.

4. **numb**, número de umbrales utilizados para discretizar la distribución.

5. **umbr(i), i=1,numb**, valores de los umbrales. Aunque tanto el número de umbrales como los valores de los mismos podrían obtenerse del fichero de salida de **IK3D**, deben de repetirse aquí con el cuidado de que coincidan con los valores utilizados para generar dicho fichero.

6. **zmin, zmax**. Valores mínimo y máximo que limitan el rango de variabilidad de la variable. Incluso en el caso que se desee utilizar la aproximación hiperbólica para extrapolar por encima del valor máximo de los umbrales, se supone que la cola no se extiende más allá de dicho máximo.

7. **colainf, param**, opción para extrapolar por debajo del umbral más pequeño hasta el mínimo **zmin**, **colainf=1** extrapolación lineal, **colainf=2** extrapolación monómica según la figura 3 con parámetro $\omega=\text{param}$.

8. **inter, param**, opción de interpolación, **inter=1** interpolación lineal, **colinf=2** interpolación monómica según la figura 3 con parámetro $\omega=\text{param}$.

9. **colasup, param**, opción para extrapolar por encima del umbral más grande, hasta el máximo **zmax**, **colasup=1** extrapolación lineal, **colasup=2** extrapolación monómica según la figura 3 con parámetro $\omega=\text{param}$, **colasup=3** extrapolación hiperbólica según la figura 4 con parámetro $\omega=\text{param}$, cuando se usa esta opción, la función de probabilidad condicionada se trunca en el valor **zmax**.

10. **maxdis**. Número máximo de subintervalos en que se subdivide el intervalo $[0,1]$, para la aproximación mediante una suma finita de las integrales que representan los valores medios que se utilizan en algunas de las opciones de salida. **maxdis=50** es un valor que da buenos resultados.

En la figura 27 se muestra un ejemplo de archivo de parámetros para **POSTIK**.

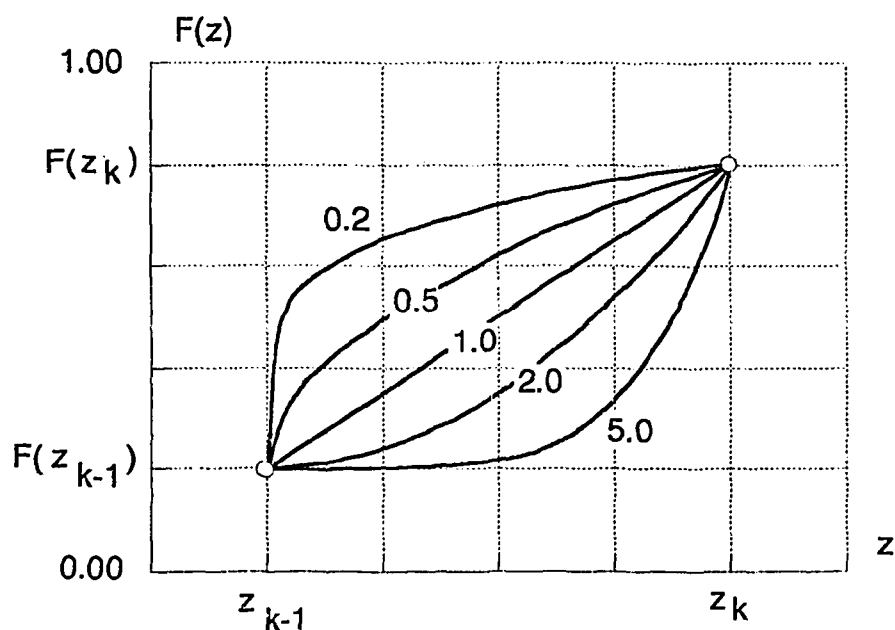


Figura 3

Modelos de interpolación de la función de distribución condicional. Familia de interpoladores válidos entre $F(z_{k-1})$ y $F(z_k)$ dados por la ecuación (4) en función del parámetro ω .

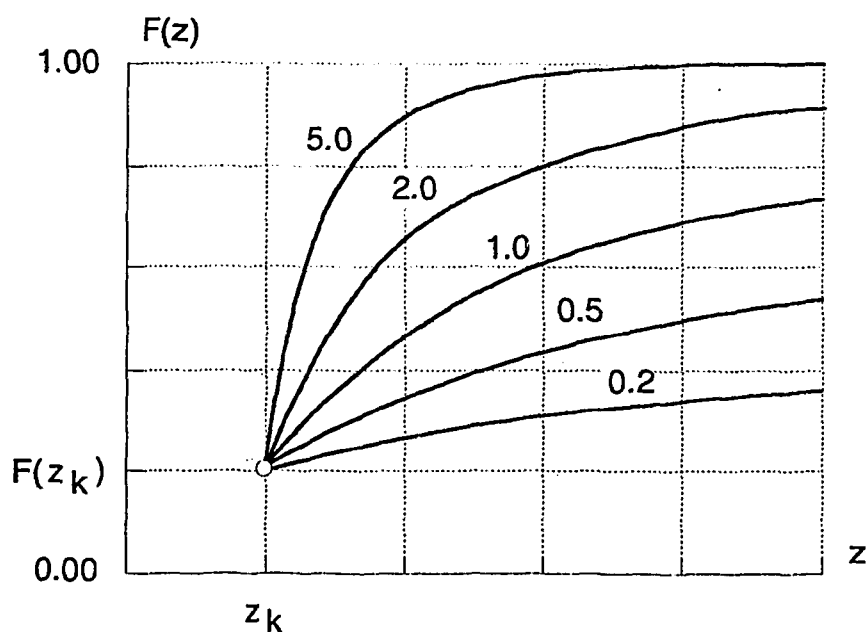


Figura 4

Modelos de extrapolación de la función de distribución condicional. Familia de extrapoladores válidos entre $F(z_k)$ y 1 dados por la ecuación (3.2.2) en función del parámetro ω .

El único archivo de salida contiene los resultados de acuerdo con la opción especificada en el tercer registro del archivo de parámetros.

3.3 Cokrigeaje

Además del programa genérico **IK3D** de krigeaje indicador múltiple, se incluye el programa **COKB3D** que permite llevar a cabo el krigeaje tradicional (ordinario o simple) de una variable o el cokrigeaje de una variable en función, no sólo de los valores de dicha variable en los alrededores, sino también los valores de variables secundarias.

El programa **COKB3D** permite el krigeaje de bloques. Aprovechando la linealidad del algoritmo de krigeaje se pueden realizar estimaciones de combinaciones lineales de la variable en estudio directamente sin necesidad de estimar cada uno de los términos de la combinación lineal independientemente. En particular se pueden estimar valores de bloque cuando éstos son iguales a la media aritmética de los puntos dentro del bloque. El programa **COKB3D** permite estimar:

$$\frac{1}{N} \sum_{i=1}^N z(x'_i)$$

que es una aproximación del valor de bloque

$$z_V(x) = \frac{1}{V} \int_{V(x)} z(x') dx'$$

donde $V(x)$ es un bloque de volumen $|V|$ centrado en x y x'_j , son N puntos discretizando el bloque.

El cokrigeaje requiere la entrada, tanto de los variogramas de cada una de las variables, como de los variogramas cruzados entre cada par de variables. Para asegurar que el modelo de variabilidad espacial conjunta de las diferentes variables es válido, el conjunto de variogramas y variogramas cruzados debe satisfacer una serie de condiciones que son bastante restrictivas. Si el modelo es incorrecto, el sistema de ecuaciones resultante es indeterminado, puesto que la matriz de coeficientes tiene un rango inferior al número de ecuaciones. El programa **COKB3D** no comprueba que el modelo es correcto excepto cuando detecta que se trata de un modelo lineal de correogionalización (Journel y Huijbregts, 1987; p. 259).

Las diversas opciones de krigeaje y cokrigeaje disponibles en el programa se explican a continuación.

Krigeaje simple

La estimación $z_{ok}^*(x)$ por krigeaje simple de la variable z en el punto x en función de los valores de la misma en el contorno $\{z(x_i), i = 1, \dots, n_i\}$ se hace mediante la siguiente combinación lineal:

$$z_{ok}^*(x) = m_z + \sum_{i=1}^{n_i} \lambda_i(x) (z(x_i) - m_z)$$

donde m_z es el valor medio y los parámetros de la combinación lineal $\lambda_i(x)$ se obtienen tras la solución de un sistema de ecuaciones normales, a las que generalmente se refiere como sistema de ecuaciones de krigeaje simple (Journel y Huijbregts, 1978). Los coeficientes de este sistema dependen del variograma de la función aleatoria Z . Esta formulación asegura la condición de estimador insesgado independientemente de los valores de los coeficientes de la combinación lineal, por lo que éstos no están sujetos a ninguna restricción.

Krigeaje ordinario

La estimación por krigeaje ordinario no requiere el conocimiento del valor medio. El estimador $z_{ok}^*(x)$ viene dado por la expresión:

$$z_{ok}^*(x) = \sum_{i=1}^{n_i} \lambda_i(x) z(x_i)$$

en este caso, para asegurar la condición de estimador insesgado es necesario requerir que la suma de los parámetros sea igual a la unidad

$$\sum_{i=1}^{n_i} \lambda_i(x) = 1$$

Los parámetros se obtienen como solución de un sistema de ecuaciones lineales al que se refiere como sistema de ecuaciones de krigeaje ordinario.

El uso del krigeaje ordinario, junto con un entorno de búsqueda reducido, puede compensar las variaciones locales del valor medio que no pueden ser tenidas en cuenta en un modelo estacionario.

Cokrigeaje simple

En el cokrigeaje (tanto simple como ordinario) además de los valores de la propia variable que está sien-

do estimada, se tiene en cuenta los valores de una o más variables secundarias que estén correlacionadas con la variable a estimar. Para el caso que se considere una variable secundaria y de la que se disponen datos en un conjunto de puntos, no necesariamente coincidiendo con los mismos puntos en los que se dispone de valores de la variable principal, $\{y(x_j), j=1, \dots, n_2\}$, el estimador $z_{csk}^*(x)$ viene dado por:

$$z_{csk}^*(x) = m_Z + \sum_{i=1}^{n_1} \lambda_i(x) (z(x_i) - m_Z) + \sum_{j=1}^{n_2} \lambda_j(x) (y(x_j) - m_Y)$$

donde m_Y es el valor medio de la segunda variable. Los coeficientes $\lambda_i(x)$ y $\lambda_j(x)$ se obtienen como solución de un sistema de ecuaciones lineales cuyos coeficientes son función de los variogramas de Z y de Y , así como del variograma cruzado entre Z e Y . Evidentemente, si el variograma cruzado es un efecto pepita puro, la falta de correlación entre ambas variables hace que el estimador z_{csk}^* sea igual al estimador resultante del kriging simple, z_{sk}^* .

La formulación utilizada, similarmente a lo que ocurre con el estimador de kriging simple, asegura la condición de estimador insesgado, independientemente de los valores que tomen los coeficientes de la combinación lineal.

Cokriging ordinario

El cokriging ordinario no requiere el conocimiento de los valores medios m_Z y m_Y . El estimador en este caso se expresa como

$$z_{csk}^*(x) = \sum_{i=1}^{n_1} \lambda_i(x) z(x_i) + \sum_{j=1}^{n_2} \lambda_j(x) y(x_j)$$

La condición de estimador insesgado se satisface imponiendo que la suma de los coeficientes que multiplican a $z(x_i)$ sumen la unidad y que la suma de los coeficientes que multiplican a $y(x_j)$ sumen cero.

$$\sum_{i=1}^{n_1} \lambda_i(x) = 1$$

$$\sum_{j=1}^{n_2} \lambda_j(x) = 0$$

Esta condición, que es suficiente pero no necesaria, resulta generalmente en valores de los coeficientes $\lambda_j(x)$ muy pequeños; por lo tanto, dando poco peso a la variable secundaria, aun cuando su correlación con la variable principal sea buena.

Cokriging ordinario estandarizado

Para intentar suprimir el problema mencionado en el párrafo anterior, producido por la condición de que los pesos aplicados a la variable secundaria sumen cero, se ha propuesto transformar la variable secundaria en otra que tenga la misma media que la variable principal y forzar que la suma de todos los pesos sea la unidad:

$$z_{csk}^*(x) = \sum_{i=1}^{n_1} \lambda_i(x) z(x_i) + \sum_{j=1}^{n_2} \lambda_j(x) (y(x_j) - m_Y + m_Z)$$

$$\sum_{i=1}^{n_1} \lambda_i(x) + \sum_{j=1}^{n_2} \lambda_j(x) = 1$$

el tipo de transformación propuesta para la segunda variable no cambia el variograma cruzado entre la variable principal y la secundaria.

3.3.1 Programa COKB3D

El programa **COKB3D** para la estimación de la variable z en función de los valores de dicha variable en el contorno y posiblemente en función de los valores de una o varias variables secundarias, tiene dos archivos de entrada y dos de salida. El diagrama de flujo del programa se muestra en el apéndice B.

Uno de los ficheros de entrada contiene los valores de la variable a estimar y las posibles variables secundarias si es que hay alguna. El formato de este fichero es el mismo que el descrito en 3.2.1.

El otro fichero de entrada es el fichero de parámetros. Al igual que el fichero de parámetros para el programa **IK3D**, las primeras líneas del fichero pueden utilizarse para incluir una explicación de los contenidos del mismo, ya que los parámetros de entrada no empiezan a leerse hasta después de la línea que comienza textualmente por **COMIENZO DE PARAMETROS**. A partir de esta línea se requieren los siguientes registros:

1. Nombre del fichero con los datos de entrada (40 caracteres).
2. Número de variables. Este valor será 1 para kriging ordinario o simple. En caso de cokriging

el valor será 1 más el número de variables secundarias.

3. Columnas para las coordenadas x , y , z , variable principal y variables secundarias. En caso de que alguna de las coordenadas no se haya especificado, introducir un cero.
4. **tmin**, **tmax**, sólo los datos (de cualquier variable) comprendidos entre **tmin** y **tmax** son considerados en el análisis.
5. **ktipo**, tipo de krigeaje, 0 = krigeaje o cokrigeaje simple, 2 = krigeaje o cokrigeaje ordinario, 1 = cokrigeaje ordinario estandarizado
6. **medias**, valor medio de la variable principal y todas las secundarias. Estos valores sólo se usan en el caso de krigeaje o cokrigeaje simples, aunque un valor debe ser introducido en todos los casos.
7. Nombre del fichero de salida de resultados.
8. Nivel de depuración. Un entero que puede tomar los valores 0, 1, 2 ó 3 dependiendo de la cantidad de información intermedia que se desea imprimir. Lo normal es situar el nivel de depuración a 0, a no ser que se suponga que se ha detectado algún error en el programa y se desea conocer los pasos intermedios.
9. Nombre del fichero de salida con la información del análisis de depuración.
10. **nx**, **xmin**, **deltax**, véase sección 2.2 y figura 1.
11. **ny**, **ymin**, **deltay**, ídem.
12. **nz**, **zmin**, **deltaz**, en el caso que el archivo de datos duros sólo contuviera dos coordenadas y se desee hacer una estimación en el plano de los datos (estimación bidimensional), el valor de **nx**, **ny** o **nz** correspondiente a la coordenada sin especificar debe valer 1 y, automáticamente, todos los datos toman como valor de la coordenada sin especificar el correspondiente a **xmin**, **ymin** o **zmin**.
13. **nxdisc**, **nydisc**, **nzdisc**. En el caso de krigeaje por bloques, estos tres enteros indican la densidad de discretización del bloque en las direcciones x , y y z . En el caso de krigeaje puntual, sólo se desea estimar un punto por bloque, por tanto, estas tres variables deben valer 1.
14. **mindatprim**, **maxdatprim**, **maxdatsec**, sólo si en el entorno de búsqueda se encuentran al menos **mindatprim** datos de la variable principal

se llevará a cabo la estimación; como máximo se usarán **maxdatprim** datos de la variable principal y un total de **maxdatsec** datos en total entre todas las variables secundarias.

15. **radioprim**, **radiosec**, se usan dos entornos de búsqueda, uno de radio **radioprim** para la variable principal y otro de radio **radiosec** para las secundarias.

A continuación hay que indicar los variogramas de cada variable y los variogramas cruzados. Primero hay que indicar mediante dos índices a qué variables se refiere el variograma o variograma cruzado. Índice "1" se refiere a la variable principal e índices "2", "3", etc., se refieren a las variables secundarias, según se indique en el registro 3. Así, para introducir el variograma de la variable principal, estos dos índices deben valer 1 y 1; para el variograma cruzado entre la variable principal y la primera de las secundarias, los índices deben valer 1 y 2; para el variograma cruzado entre la primera de las variables secundarias y la segunda de las variables secundarias, estos índices deben ser 2 y 3 y así sucesivamente. Para cada uno de los variogramas y variogramas cruzados hay que repetir los registros siguientes:

1. **i**, **j**, índices indicando las variables a las que corresponde este variograma (si es el variograma directo, ambos índices deben tener el mismo valor).
2. **nest**, **c0**, número de estructuras imbricadas que componen el variograma, sin contar el efecto pepita, y valor de este último.

Los siguientes registros hay que repetirlos para cada una de las estructuras imbricadas:

3. **vtipo**, **alcance**, **meseta**, tipo de variograma, 1 para esférico, 2 para exponencial, 3 para gaussiano y 4 para monomial; **alcance** en la dirección de máxima continuidad (la que tras la rotación especificada en el registro siguiente queda alineada con el eje Y' , véase sección 2.2 y valor de la meseta de esta estructura.
4. **ang1**, **ang2**, **ang3**, **anis1**, **anis2**, ángulos que definen la rotación de los ejes cartesianos XYZ para transformarlos en ejes paralelos a los ejes de anisotropía, según se explica en 2.2, y factores de anisotropía de los alcances en la dirección de los ejes X' y Z' , en ese orden, con respecto del alcance en la dirección Y' .

En la figura 20 se muestra un ejemplo de archivo de parámetros para **COKB3D**.

De los dos archivos de salida, uno es el que contiene información sobre la ejecución del programa en función del valor del nivel de depuración. El otro es un archivo con el formato descrito en la sección 3.2.1, que contiene en dos columnas el estimador por krigeaje y la varianza de la estimación. Por ser una

salida de datos sobre malla regular, las coordenadas no se dan explícitamente sino que, como también se explica en 3.2.1, están escritos comenzando por la celda correspondiente a las coordenadas x_{\min} , y_{\min} , z_{\min} , capa por capa, dentro de cada capa fila por fila y, en cada fila, columna por columna.

4

SIMULACION

4.1 Introducción

En esta sección se describen los programas de simulación. Estos programas, en lugar de generar un único campo, como es el caso de los programas de estimación, generan una serie de campos equiprobables, todos los cuales reproducen los datos medidos y la variabilidad inferida de los mismos. Sin tener en cuenta ninguna otra información, cualquiera de estos campos podría ser el real, por ello se llaman equiprobables.

En contraposición a los campos obtenidos por estimación, que no pretenden reflejar la variabilidad espacial observada en los datos, (más al contrario, su variabilidad espacial es siempre muy inferior a la real), los campos obtenidos por simulación pretenden obtener valores que localmente son óptimos bajo un determinado criterio de optimalidad.

Los programas que se describen permiten la generación de campos tanto multinormales como no y ambos están basados en la técnica de simulación conocida como de simulación secuencial (Alabert, 1987; Journel, 1989; Isaaks, 1990; Gómez-Hernández y Srivastava, 1990; Gómez-Hernández, 1991; Gómez-Hernández y Journel, 1993).

4.2 Simulación de campos multinormales

Las técnicas más sencillas de simulación de campos son aquéllas que generan campos correspondientes a una función aleatoria con distribución multinormal. El uso de estas técnicas requiere que del análisis de los datos podamos concluir que la variable en estudio tiene unas características que la hacen buena candidata para una modelización mediante este tipo de función aleatoria.

La condición primera que deben satisfacer los datos para poder modelizarse con una función multinormal, es que su distribución univariada de frecuencia acumulada (posiblemente después de aplicar alguna técnica de desaglomeración de los datos) debe ser normal. Si esto no ocurre, siempre se puede recurrir a alguna transformación monotónica no lineal que convierta los datos, inicialmente con una distribución lejos de la normal, en otro conjunto de datos que sí tienen esa distribución. Esta transformación puede tener una expresión analítica o no paramétrica. El ejemplo más corriente en hidrogeología es la transformación de la permeabilidad en log-permeabilidad.

No basta con esta transformación univariada para asegurar multinormalidad de los datos, puesto que, si

bien su distribución univariada se ha convertido en normal, la transformación efectuada no asegura que los momentos de orden superior sean los correspondientes a una distribución multinormal.

Este hecho es en la mayoría de los casos olvidado. La razón del olvido es que, en general, no se dispone de suficiente cantidad de datos para poder calcular con cierta fiabilidad momentos de orden superior. Sin embargo, el aceptar los momentos de orden superior que vienen implícitamente impuestos por la elección del modelo multinormal como válidos, puede conducir a resultados que lleven a hacer conclusiones demasiado optimistas o demasiado pesimistas.

La gran ventaja del modelo multinormal, una vez se ha decidido utilizarlo, es que queda totalmente definido en función de la media y la covarianza (o la media, la varianza y el variograma), valores éstos que, generalmente, pueden inferirse a partir de los datos. De todos modos, queremos insistir que, el que no se puedan inferir momentos de orden superior a partir de los datos no quiere decir que hay que aceptar aquéllos que corresponden a una distribución multinormal. El usuario debe, tras transformar los datos de manera que la transformada sea univariada normal, decidir si acepta el modelo multinormal para los momentos de orden superior o prefiere otro modelo, en cuyo caso necesitará recurrir a técnicas no gaussianas de simulación.

4.2.1 Programa GCOSIM3D

El programa **GCOSIM3D** (Gómez-Hernández y Journel, 1993) permite la generación de campos de varias variables mutuamente relacionados a través de una distribución multinormal multivariada. Es el contrapunto del programa de estimación **COKB3D**. El diagrama de flujo del programa se muestra en el apéndice C.

La estimación de las funciones de distribución condicionada se hacen por cokrigeaje simple, no permitiéndose ninguna de las otras opciones disponibles en el programa de estimación. En el caso en que la distribución conjunta de las variables sea multinormal, la función de distribución condicional es normal, con media igual al valor estimado por cokrigeaje simple y con varianza dada por el valor de la varianza de estimación del cokrigeaje.

El programa necesita tres archivos de entrada y produce dos archivos de salida. Los archivos de entrada deben tener los nombres **gcosim3d.geo**, **gcosim3d.var** y **gcosim3d.dat**. Los archivos de salida tienen los nombres **gcosim3d.dbg** y **gcosim3d.out**.

Sus formatos y contenidos se describen a continuación.

Fichero gcosim3d.geo

Fichero que contiene, entre otros parámetros, la descripción de la geometría del sistema. El archivo de parámetros comienza en la primera línea (a diferencia de los archivos de parámetros de los programas de estimación, cuya primera línea válida es la siguiente a la que textualmente dice **COMIENZO DE PARAMETROS**), cada línea debe contener tantos parámetros como se indica a continuación y no se permiten caracteres adicionales al final de la línea. El formato de los ficheros de entrada es bastante estricto. En cualquier caso, todos los datos de entrada se imprimen junto a una explicación de los mismos en el fichero de salida **gcosim3d.dbg**, que conviene comprobar para identificar errores en la confección de los archivos de entrada.

1. **seed, nsim**, semilla que se utilizará para inicializar el generador de números aleatorios y número de simulaciones deseadas. Valores enteros, sin unidades.
2. **dx, dy, dz**, tamaño de la malla en las tres direcciones coordenadas. Valores reales, en unidades de longitud.
3. **xmin, ymin, zmin**, coordenadas del centro de la celda (1,1,1) (véase figura 1). Valores reales, en unidades de longitud.
4. **nx, ny, nz**, número de celdas en las direcciones *x*, *y* y *z*. Valores enteros, sin unidades.

En algunos casos y en especial cuando el área a ser simulada es muy grande, conviene hacer una primera prueba simulando una pequeña subárea de la totalidad; esta subárea se especifica indicando los índices de los dos vértices opuestos de la misma. Los dos registros siguientes indican que el subárea a simular va desde la celda con índices **desde.x, desde.y, desde.z** hasta la celda con índices **hasta.x, hasta.y, hasta.z**. En caso que se desee simular la totalidad del área se deberá especificar **desde.x=1, desde.y=1, desde.z=1** y **hasta.x=nx, hasta.y=ny, hasta.z=nz**. Esta opción es también interesante cuando se desea hacer una simulación bidimensional pero utilizando datos que están fuera del plano de la simulación.

5. **desde.x, desde.y, desde.z**, índices que referencian la primera celda de la subárea que se va a simular. Valores enteros, sin unidades.

6. **hasta.x, hasta.y, hasta.z**, índices que referencian la celda diagonalmente opuesta a la anterior. Valores enteros, sin unidades.

7. **semix, semiy, semiz**, longitud de los semiejes principales del entorno elipsoidal de búsqueda. La eficacia del programa depende de una buena elección de estos parámetros. Si el entorno de búsqueda es demasiado grande, el número de celdas que contiene es muy alto y el tiempo invertido en la localización de los puntos dentro del entorno es grande. Si el entorno de búsqueda es pequeño, la búsqueda es rápida pero la reproducción de los variogramas de entrada se deteriora. El entorno de búsqueda debe ser algo más grande que los alcances de los distintos variogramas en cada dirección. Si el alcance es muy grande y no se dispone de muchos datos iniciales que puedan controlar la reproducción del variograma para distancias largas, hay que recurrir a una simulación en etapas, en la primera etapa se utiliza una discretización gruesa que permite usar un entorno de búsqueda grande sin que tenga que contener muchas celdas. Cuando este primer mallado grueso queda completado se utilizan los datos generados sobre esta malla como datos condicionantes para la segunda etapa, en la que el entorno de búsqueda no necesita ser tan grande porque la reproducción de los variogramas a largas distancias está asegurada por los datos de la primera etapa. Generalmente, la simulación en dos etapas es suficiente aunque el concepto se puede extender a múltiples etapas. Valores reales, en unidades de longitud.

8. **cosx'x, cosx'y, cosx'z**, cosenos de los ángulos que forma el semieje **semix** con los ejes cartesianos XYZ. Valores reales, sin unidades.

9. **cosy'x, cosy'y, cosy'z**, cosenos de los ángulos que forma el semieje **semiy** con los ejes cartesianos XYZ. Valores reales, sin unidades.

10. **cosz'x, cosz'y, cosz'z**, cosenos de los ángulos que forma el semieje **semiz** con los ejes cartesianos XYZ. Valores reales, sin unidades.

11. **maxoct, maxdata**, sólo los puntos dentro del entorno de búsqueda hasta un máximo de **maxoct** por octante serán considerados para la estimación de la función de probabilidad condicionada. La velocidad de ejecución del programa es inversamente proporcional a este parámetro, mientras que la reproducción de los variogramas de entrada es directamente proporcional al mismo. En cualquier caso, se recomienda un va-

lor igual o superior a 2 sin sobrepasar 8. Valores enteros adimensionales.

12.dbg, un entero indicando el nivel de información intermedia deseada. Normalmente debe valer cero pero si se le da un valor de 3 o más, la cantidad de información enviada al fichero **gcosim3d.dbg** es enorme. Tal opción sólo debe utilizarse cuando el subárea indicada en los registros 5 y 6 sea pequeña. Valor entero adimensional.

En la figura 35 se muestra un ejemplo de archivo **gcosim3d.geo** para el ejemplo práctico de la base de datos Berea.

Fichero **gcosim3d.var**

Este fichero contiene la descripción de los variogramas directos y cruzados que determinan el modelo multinormal multivariado a utilizar. Al igual que para el fichero anterior, el formato de entrada es muy estricto por lo que no se deben incluir campos extra ni se debe alterar el orden de los registros, tal y como se indica a continuación.

1. **nvar**, número de variables que se van a utilizar. Valor entero adimensional.
2. **nvar_sim**, número de variables que se van a simular. No todas las variables de las que se disponen datos deben ser simuladas, en algunos casos sólo se desea simular una variable condicionada a los valores de la misma y a los de otra u otras variables con las que guarda correlación toman en el entorno. Valor entero adimensional.
3. Vector con **nvar_sim** números enteros con un índice indicando qué variables son las que deben simularse. Este índice se utiliza para identificar los datos en el fichero de entrada y para identificar los variogramas que se especifican más abajo. Enteros adimensionales.
4. Vector con los valores medios de las **nvar** distribuciones marginales de cada una de las variables. Valor real con unidades de la variable a la que represente.

Los registros que siguen definen todos los variogramas y variogramas cruzados. Puesto que los variogramas cruzados son simétricos en las variables para las que se calculan, es decir, el variograma entre

las variables z_1 y z_2 es el mismo que el variograma entre z_2 y z_1 ⁽¹⁾, se necesita especificar **nvar(nvar-1)/2** variogramas. Estos variogramas se especifican en el siguiente orden: primero el variograma de la variable 1, segundo el variograma cruzado entre las variables 1 y 2, tercero el variograma de la variable 2, cuarto el variograma cruzado entre las variables 1 y 3, quinto el variograma cruzado entre las variables 2 y 3, sexto el variograma de la variable 3 y así sucesivamente.

Cada uno de los variogramas necesita los siguientes registros:

1. **pepita**, efecto pepita. Valor real con unidades en función de la variable a que represente.
2. **cmax**, un valor real superior al máximo que puede tomar el variograma dentro del entorno de búsqueda especificado en **gcosim3d.geo**.
3. **num_estruc**, número de estructuras imbricadas que definen el variograma. Entero adimensional.
4. **tipos**, vector indicando el tipo de las **num_estruc** estructuras siendo los tipos 1 para esférico, 2 para exponencial, 3 para gaussiano y 4 para monomial (véase sección 2.2). Enteros adimensionales.
5. **mesetas**, vector con las mesetas de las **num_estruc** estructuras imbricadas. Valores reales cuyas unidades dependen de la variable.
6. **alcancex'**, vector con los alcances de las **num_estruc** estructuras en la dirección principal de anisotropía X' , no necesariamente la de máxima continuidad. Valores reales con unidades de longitud.
7. **alcancey'**, vector con los alcances de las **num_estruc** estructuras en la dirección principal de anisotropía Y' .
8. **alcancez'**, vector con los alcances de las **num_estruc** estructuras en la dirección principal de anisotropía Z' .
9. **cosx'x**, **cosx'y**, **cosx'z**, cosenos de los ángulos formados entre el semieje de anisotropía X' y los ejes cartesianos XYZ. El elipsoide de anisotropía puede ser diferente del entorno elipsoidal de búsqueda. De hecho, este último debe de tener en cuenta todos los elipsoides de anisotropía

⁽¹⁾ Nótese que este hecho no es cierto para las covarianzas, es decir, la covarianza entre z_1 y z_2 no es necesariamente igual a la covarianza entre z_2 y z_1

de todas las estructuras puesto que, en principio, éstos pueden ser diferentes para diferentes variogramas. El elipsoide de anisotropía debe ser el mismo para todas las estructuras imbricadas de un mismo variograma. Valores reales adimensionales.

10. **cosy*x, cosy*y, cosy*z**, cosenos de los ángulos formados por el semieje de anisotropía Y' y los ejes cartesianos XYZ . Valores reales adimensionales.

11. **cosz*x, cosz*y, cosz*z**, cosenos de los ángulos formados por el semieje de anisotropía Z' y los ejes cartesianos XYZ . Valores reales adimensionales.

En la figura 36 se muestra un ejemplo de archivo **gcosim3d.var** para el ejemplo práctico de la base de datos Berea.

Fichero gcosim3d.dat

Fichero conteniendo los datos que sirven para condicionar las simulaciones. En la versión actual, sólo los datos que se encuentran dentro de la malla, se utilizan. Si se desea utilizar datos que están fuera de la malla hay que definir una malla más amplia que los contenga y después restringir la generación a la subárea de interés utilizando los registros 5 y 6 del fichero **gcosim3d.geo**.

1. **ndata**, número de datos a leer (puede ser menor que el número total de datos contenidos en el fichero).

A continuación tantos registros como **ndata** conteniendo los datos. Cada dato requiere:

2. **x, y, z, var, value**, coordenadas x , y y z , un entero indicando a qué variable corresponde el dato y el valor del mismo. Tres reales con unidades de longitud, un entero adimensional y un real cuyas unidades son las de la variable que representa.

En la figura 37 se muestra un ejemplo de archivo **gcosim3d.dat** para el ejemplo práctico de la base de datos Berea.

Fichero gcosim3d.dbg

Fichero de salida con información sobre la ejecución del programa. La cantidad de información vertida a este fichero depende del nivel de información intermedia fijado en el registro 12 del archivo **gcosim3d.geo**. Para el nivel de depuración mínimo sólo se reproducen los valores de los archivos de entrada.

Fichero gcosim3d.out

El fichero de salida es similar a los ficheros de salida de los anteriores programas, comienza con una línea de texto, el siguiente registro contiene el número de variables que han sido simuladas (**nvar_sim**), los siguientes **nvar_sim** registros son etiquetas con los índices de orden de las variables. A continuación siguen los registros conteniendo las **nsim** simulaciones de las **nvar_sim** variables. Cada registro contiene una columna por cada variable simulada. El orden de impresión de los registros es una simulación tras otra, dentro de cada simulación, una capa tras otra (donde capa va asociado con la coordenada z), dentro de cada capa una fila tras otra (donde fila va asociado con la coordenada y) y dentro de cada fila una columna tras otra (donde columna se asocia a la coordenada x). Todos los bucles empiezan siempre por el valor que tiene las coordenadas más bajas.

4.3 Simulación secuencial de indicadores

La simulación secuencial de indicadores puede utilizarse para generar campos que no corresponden a una función aleatoria multinormal. Más concretamente, la simulación secuencial de indicadores pretende producir realizaciones de una función aleatoria con una función de distribución bivariada cualquiera.

El programa **ISIM3D** es el contrapunto del programa **IK3D**. Mientras que éste se utiliza para la generación de un único campo que minimiza localmente una cierta función de pérdidas sin pretender reproducir la variabilidad espacial de los datos, aquél genera múltiples realizaciones equiprobables, todas ellas reproduciendo los datos y con la variabilidad espacial que se observa en los mismos.

Antes de pasar a describir el manejo del programa es conveniente aclarar cuáles son las ventajas que tiene la simulación secuencial de indicadores frente a técnicas de simulación multigaussianas y comentar cuándo debe utilizarse una u otra.

La mayor ventaja, y a la vez el mayor inconveniente, de las técnicas de simulación multigaussianas es que el modelo de función aleatoria subyacente queda completamente definido en función de un valor medio y una covarianza. La ventaja es que el número de parámetros necesarios para caracterizar la función aleatoria es muy pequeño y generalmente se puede inferir a partir de los datos. El inconveniente es que, tomada la decisión de utilizar un modelo multigaussiano, no sólo se ha decidido cuál debe ser la media y la covarianza de los campos generados, sino todos sus

parámetros estadísticos. En concreto, las covarianzas de las variables indicadoras correspondientes a una función aleatoria multinormal están perfectamente caracterizadas y muestran las siguientes peculiaridades:

1. Las covarianzas de variables indicadoras correspondientes a umbrales simétricos respecto a la media son iguales, lo que implica que la correlación espacial de los valores altos debe ser la misma que la de los valores bajos. No sólo la magnitud de los alcances sino también la orientación de los ejes de anisotropía.
2. A medida que el umbral que define la variable indicadora se aparta del valor medio, el alcance del variograma se reduce y el efecto pepita relativo aumenta, lo que implica una pérdida de continuidad espacial en los valores extremos.

Estas dos propiedades son características de todos los campos generados con técnicas multigaussianas. La decisión de utilizar una técnica multigaussiana o no deberá estar basada en la conveniencia o no que los campos generados, además de reproducir la media y la varianza, presenten este tipo de comportamiento de las covarianzas indicadoras y lo que ello implica en cuanto a continuidad de los valores extremos.

Si la cantidad de datos es suficiente para calcular variogramas indicadores para un cierto número de umbrales, entonces estos variogramas indicadores deben compararse con los que corresponden al modelo multigaussiano con la media y covarianza de los datos. Si la comparación es favorable, no hay necesidad de utilizar simulación de indicadores, puesto que en principio deberían proporcionar los mismos resultados. Si la comparación no lo es, entonces debe ponderarse la desviación existente y las implicaciones que tendría la generación de campos multigaussianos con las características antes mencionadas.

Cuando no hay suficientes datos para poder calcular los variogramas indicadores con un mínimo de confianza, entonces el uso de un generador multigaussiano o uno basado en simulación secuencial de indicadores es una decisión que debe tomar el modelizador. Hay una combinación infinita de posibles variogramas indicadores para una determinada covarianza de la variable original, del mismo modo que hay un número infinito de posibles funciones de probabilidad acumulada con una media dada.

Una de dichas combinaciones es la que corresponde a la distribución multigaussiana con las características de variabilidad espacial para los valores extremos. Otras combinaciones de variogramas indica-

dores pueden tener, por ejemplo, la misma estructura espacial para todos los umbrales, es decir, la correlación espacial de los valores extremos es la misma que la de los valores medianos, es el que se conoce como modelo "mosaico".

Hay otros modelos que pueden tener mayor correlación espacial para los valores extremos muy altos y poca (o con direcciones de anisotropía diferentes) para los valores extremos muy bajos. Evidentemente, en el caso de decidirse por una función no multigaussiana, el modelizador debe especificar los variogramas indicadores para un número de umbrales, con la precaución de que la covarianza de la variable original sea reproducida.

El problema de la dificultad de la inferencia de los variogramas indicadores que en algunos casos se achaca a las técnicas no paramétricas no existe. Si los variogramas indicadores no se pueden inferir por falta de datos no implica que los variogramas indicadores, implícitamente considerados cuando se elige un modelo multigaussiano, sean los correctos. Por ello, en tanto en cuanto se asegure la reproducción del variograma de la variable original, cualquier selección de variogramas indicadores es válida hasta que se demuestre lo contrario.

La mayor desventaja de las técnicas de simulación de indicadores es la necesidad de discretizar el rango de variación de la variable en estudio en un número finito de clases, con lo que las simulaciones basadas en indicadores proporcionan un resultado con una resolución que es función del número y amplitud de las clases seleccionadas.

La solución a este problema consiste en asignar un valor real continuo a cada celda después que mediante la simulación secuencial de indicadores se ha determinado la clase a la que pertenece. Esta asignación se puede hacer manteniendo una cierta correlación espacial dentro de cada clase pero independientemente una de otra, o, más simple es asignar un valor seleccionado aleatoriamente dentro de cada clase de acuerdo a una función de probabilidad marginal predeterminada. En ambos casos, más notablemente en el segundo, se introduce un cierto ruido en el variograma de la variable continua así generada que se traduce en un efecto pepita no deseado.

El programa **ISIM3D** internamente realiza la asignación de un valor real a cada una de las celdas, de acuerdo con un par de modelos que se explican más adelante.

4.3.1 Programa ISIM3D

El programa **ISIM3D** (Gómez-Hernández y Srivastava, 1990) permite la generación de campos en los que la variable ha sido previamente discretizada en un número finito de clases y cada variable indicadora, correspondiente a los límites que definen dichas clases, tiene un variograma no necesariamente correspondiendo a una función aleatoria multigaussiana. El diagrama de flujo del programa se muestra en el apéndice C.

La estimación de las funciones de distribución condicionada se pueden hacer por cokrigeaje indicador simple u ordinario. La función de distribución condicionada se estima como en el programa **IK3D**, con la salvedad que, tras estimar dicha función de distribución, un valor es extraído aleatoriamente de la misma y asignado al nodo correspondiente, valor éste que pasa directamente a engrosar el conjunto de datos en los que debe condicionarse la función de distribución condicional de los nodos restantes.

Debido al algoritmo de simulación secuencial (Gómez-Hernández y Srivastava, 1990), la reproducción de los variogramas de entrada sólo está asegurada hasta una distancia de separación igual a los semiejes del entorno elipsoidal de búsqueda que se define en el fichero **isim3d.geo**.

Con el objetivo de asegurar la reproducción de los variogramas de entrada, para distancias de correlación grandes, con respecto al tamaño de discretización de la malla, sin necesidad de recurrir a un entorno de búsqueda desproporcionado (véase comentario al registro número 7 del fichero de entrada **geo-sim3d.geo**), se permite la posibilidad de realizar la simulación en dos etapas; en la primera etapa se utiliza una malla gruesa (cuyo espaciado debe ser un múltiplo entero del espacio en la malla final) sobre la que se realiza una primera simulación y que servirá para imponer las distancias de correlación grandes en la malla final. En esta primera etapa, el paso de malla es grande y el entorno de búsqueda puede tener unas dimensiones físicas grandes sin que necesariamente contenga un elevado número de celdas. Cuando se ha completado esta etapa, la segunda etapa consiste en el relleno de la malla fina en aquellas celdas que no corresponden a una posición previamente simulada en la malla gruesa. En este caso, y puesto que la reproducción de la correlación espacial a grandes distancias está asegurada por la simulación sobre la malla gruesa, el entorno de búsqueda no debe tener unas dimensiones físicas demasiado grandes.

El programa necesita tres archivos de entrada, (cuatro si se utiliza información blanda), que deben tener

los nombres **isim3d.geo**, **isim3d.var**, **isim3d.dat**, (y **isim3d.ind**). Los archivos de salida tienen los nombres **isim3d.out**, **isim3d.cou**, **isim3d.rst** y **isim3d.dbg**. Sus formatos y contenidos se describen a continuación.

Fichero **isim3d.geo**

Fichero con la descripción geométrica del campo a generar y algunos otros parámetros de control. Cada registro debe contener tantos parámetros como se indica a continuación, permitiéndose un comentario opcional al final del registro después del último parámetro. Es importante comprobar que el archivo ha sido escrito correctamente, por lo que todos los parámetros de entrada se vuelven a escribir en el archivo de depuración **isim3d.dbg**.

1. **seed, nsim**, semilla utilizada para inicializar el generador de números aleatorios y número de simulaciones a generar. Valores enteros sin unidades.
2. **dx, dy, dz**, tamaño de la malla final en la que se desea obtener las simulaciones en las tres direcciones coordenadas. Valores reales en unidades de longitud.
3. **dxg, dyg, dzg**, tamaño de la malla gruesa intermedia sobre la que se realiza una primera simulación que permita preservar las distancias de correlación a largas distancias. Los cocientes **dxg/dx**, **dyg/dy** y **dzg/dz** deben ser números enteros aunque no necesariamente iguales, de esta manera se asegura que los valores simulados sobre la malla gruesa están localizados exactamente sobre un nodo de la malla final. Valores reales de longitud.
4. **xmin, ymin, zmin**, coordenadas del centro de la celda (1,1,1) (véase la figura 1). Valores reales de longitud.
5. **nx, ny, nz**, número de celdas en las direcciones *x*, *y* y *z* en la malla final. El número de celdas de la malla gruesa es calculado internamente en función de las relaciones **dxg/dx**, **dyg/dy** y **dzg/dz**. Valores enteros, sin unidades.

En algunos casos, y en especial cuando el área a ser simulada es muy grande, conviene hacer una primera prueba simulando una pequeña subárea de la totalidad, esta subárea se especifica indicando los índices de los dos vértices opuestos de la misma. Los dos registros siguientes indican que el subárea a simular va desde la celda con índices **desde.x**, **desde.y**, **desde.z** hasta la celda

con índices **hasta.x**, **hasta.y**, **hasta.z**. En caso que se desee simular la totalidad del área se deberá especificar **desde.x=1**, **desde.y=1**, **desde.z=1** y, **hasta.x=nx**, **hasta.y=ny**, **hasta.z=nz**. Esta opción es también interesante cuando se desea hacer una simulación bidimensional pero utilizando datos que están fuera del plano de la simulación.

6. **desde.x**, **desde.y**, **desde.z**, índices que referencian la primera celda de la subárea que se va a simular. Valores enteros, sin unidades.
7. **hasta.x**, **hasta.y**, **hasta.z**, índices que referencian la celda diagonalmente opuesta a la anterior. Ambos índices referidos a la malla final. Valores enteros, sin unidades.
8. **semix**, **semy**, **semiz**, longitud de los semiejes principales del entorno elipsoidal de búsqueda en la malla final. (Ver comentario al registro 7 en el fichero **gcosim3d.geo**). Valores reales con dimensiones de longitud.
9. **semixg**, **semyg**, **semizg**, longitud de los semiejes principales del entorno elipsoidal de búsqueda en la malla gruesa. Los dos entornos de búsqueda tienen la misma orientación. Valores reales con dimensiones de longitud.
10. **cosx'x**, **cosx'y**, **cosx'z**, cosenos de los ángulos que forma el semieje **semix** con los ejes cartesianos XYZ. Valores reales, sin unidades.
11. **cosy'x**, **cosy'y**, **cosy'z**, cosenos de los ángulos que forma el semieje **semy** con los ejes cartesianos XYZ. Valores reales, sin unidades.
12. **cosz'x**, **cosz'y**, **cosz'z**, cosenos de los ángulos que forma el semieje **semiz** con los ejes cartesianos XYZ. Valores reales, sin unidades.
13. **maxoct**, **okmin**, **maxdata**, sólo los puntos dentro del entorno de búsqueda hasta un máximo de **maxoct** por octante serán considerados para la estimación de la función de probabilidad condicionada. En cualquier caso y como máximo la estimación se hace con los **maxdata** más próximos. La velocidad de ejecución del programa es inversamente proporcional a estos dos parámetros, mientras que la reproducción de los variogramas de entrada es directamente proporcional a los mismos. En cualquier caso, se recomienda un valor para **maxoct** entre 2 y 8. Si el número de puntos dentro del entorno de búsqueda es menor que **okmin**, se calcula la función de probabilidad acumulada utilizando krigeaje indica-

dor simple, si es mayor o igual se utiliza krigeaje indicador ordinario. Si el número de datos iniciales es grande, es preferible utilizar krigeaje ordinario desde el principio, puesto que este tipo de krigeaje permite corregir desviaciones de la estacionaridad que tiene el modelo de función aleatoria. Si el número de datos es bajo (o nulo), entonces debe forzarse el uso de krigeaje simple utilizando un valor alto de **okmin**. Valores enteros sin unidades.

14. **depura**, **depura** es un entero indicando la cantidad de información intermedia que debe vertirse al fichero **isim3d.dbg**. Normalmente debe valer cero, un valor de 3 o superior imprime todos los detalles de la simulación. Valor entero adimensional.

15. **frec_reini**, **indi_reini**, estos dos valores permiten recuperar parte del trabajo perdido si por una interrupción brusca del programa éste es abortado. El uso de esta opción tiene mayor sentido cuando el tamaño de la simulación supera los cientos de miles de celdas, con lo que el tiempo invertido en la generación de una sola imagen es alto. **frec_reini** es la frecuencia, contada en número de celdas simuladas, a la que quiere hacerse una copia del estado de la simulación en el disco. Si no se quiere utilizar esta opción sólo hay que asignar a esta variable un valor superior al número total de celdas. **indi_reini** es un indicador que puede tomar los valores 0 ó 1. Si es 0, la simulación comienza desde el principio, si es 1, es porque el programa ha sido abortado en medio de la simulación y se desea comenzar de nuevo desde el punto en que se quedó la última vez que se grabó la información en disco (con arreglo a la frecuencia indicada por **frec_ini**). Las nuevas simulaciones se añaden al final del fichero **isim3d.out**.

En la figura 41 se muestra un ejemplo de archivo **isim3d.geo** para el ejemplo práctico de la base de datos Berea.

Fichero **isim3d.var**

Este fichero contiene la descripción de los variogramas indicadores para los distintos umbrales seleccionados. Al igual que el fichero anterior, cada línea de este fichero debe contener los campos que se indican, pudiendo, a continuación, incluirse un comentario al final de cada línea.

1. **nind**, número de variables indicadoras. Valor entero adimensional.

2. **minimo, maximo**, límites absolutos que delimitan el rango de variabilidad de la variable. Si algún dato de entrada está fuera de estos límites se descarta. Estos valores también controlan los valores que puede tomar la variable simulada tras la selección de un valor real dentro de cada clase. Valores reales con las unidades de la variable en estudio.

Los siguientes registros deben repetirse para cada una de las variables indicadoras. El orden de entrada es de menor a mayor umbral.

1. **umbral, pdf**, umbral que define la variable indicadora y valor de la distribución de probabilidad marginal para dicho umbral (este valor sólo se usa para la estimación de la función de probabilidad condicionada por krigeaje indicador simple). Valores reales, el primero con unidades de la variable en estudio y el segundo adimensional comprendido entre 0 y 1.
2. **pepita**, efecto pepita. Valor real adimensional entre 0 y $\text{pdf}(1-\text{pdf})$.
3. **num_estruc**, número de estructuras imbricadas que definen el variograma. Entero adimensional.
4. **tipos**, vector indicando el tipo de las **num_estruc** estructuras. Sólo se permiten los tipos 1 para esférico y 2 para exponencial. Valores enteros adimensionales.
5. **mesetas**, vector con las mesetas de las **num_estruc** estructuras imbricadas. Valores reales adimensionales.
6. **alcancex'**, vector con los alcances de las **num_estruc** estructuras en la dirección principal de anisotropía X' , no necesariamente la de máxima continuidad. Valores reales con unidades de longitud.
7. **alcancey'**, vector con los alcances de las **num_estruc** estructuras en la dirección principal de anisotropía Y' .
8. **alcancez'**, vector con los alcances de las **num_estruc** estructuras en la dirección principal de anisotropía Z' .
9. **cosx'x, cosx'y, cosx'z**, cosenos de los ángulos formados entre el semieje de anisotropía X' y los ejes cartesianos XYZ. El elipsoide de anisotropía puede ser diferente del entorno elipsoidal de búsqueda. De hecho, este último debe tener en cuenta todos los elipsoides de anisotropía de todas las estructuras, puesto que, en principio, éstos pueden ser diferentes para diferentes variogramas. El elipsoide de anisotropía debe ser el mismo para todas las estructuras imbricadas de un mismo variograma. Valores reales adimensionales.
10. **cosy'x, cosy'y, cosy'z**, cosenos de los ángulos formados por el semieje de anisotropía Y' y los ejes cartesianos XYZ. Valores reales adimensionales.
11. **cosz'x, cosz'y, cosz'z**, cosenos de los ángulos formados por el semieje de anisotropía Z' y los ejes cartesianos XYZ. Valores reales adimensionales.

En la figura 42 se muestra un ejemplo de archivo **isim3d.var** para el ejemplo práctico de la base de datos Berea.

Fichero isim3d.dat

Fichero que contiene los datos "duros" que condicionan la simulación. Sólo se utilizan los datos que están dentro de la malla. Si se desea utilizar datos que están fuera de la malla para condicionar la simulación de la misma, hay que definir en **isim3d.geo** una malla suficientemente grande que los contenga y después reducir la simulación a la zona de interés, utilizando los registros de definición de subárea en dicho fichero.

El fichero contiene tantos registros como datos existan, cada registro con las coordenadas x , y y z del dato y el valor del mismo. Valores reales, los tres primeros con unidades de longitud y el último con las unidades de la variable.

En la figura 43 se muestra un ejemplo de archivo **isim3d.dat** para el ejemplo práctico de la base de datos Berea.

Fichero isim3d.ind

Fichero que contiene los datos "blandos" que condicionan la simulación. Dicho fichero puede no existir si no se utilizan datos "blandos".

En caso de que exista dicho fichero contendrá las coordenadas x , y y z del dato y a continuación los valores de las variables indicadoras para los umbrales definidos en el archivo **isim3d.var**. Si el dato es de intervalo, los indicadores que falta por completar se indican con un valor negativo. Los datos que corresponden a una función de probabilidad condicionada a priori se indican dando los valores de dicha distribución. Los valores de las variables indicadoras se dan en orden ascendente de los umbrales, por lo que ne-

cesariamente dichos valores deben formar una secuencia monótonica no descendente (excepto por el valor negativo indicativo de un valor desconocido). En caso de que así no lo sea el dato es descartado. Valores reales, los tres primeros en unidades de longitud, los restantes adimensionales.

En la figura 44 se muestra un ejemplo de archivo **isim3d.ind** para el ejemplo práctico de la base de datos Berea.

Fichero **isim3d.dbg**

Fichero de salida con información sobre la ejecución del programa. El nivel de información vertida a este fichero depende del nivel de información intermedia fijado en el archivo **isim3d.geo**.

Fichero **isim3d.out**

Fichero de salida con las simulaciones. Este fichero contiene un registro por cada celda con tres valores. El primer valor es un entero entre 1 y el número de clases en que se ha discretizado el rango de variabilidad de la variable. Este entero indica a qué clase corresponde el valor simulado. La resolución de la simulación secuencial de indicadores se limita a determinar la clase a la que pertenece el valor simulado. Para asignar un valor real a cada celda, en función de la clase a la que pertenece, hay que postular un modelo para dicha transformación. Los dos valores reales que acompañan el valor entero en cada registro de salida corresponden a dos de estos modelos. Conocida la clase a la que el valor simulado pertenece, se conocen los valores de las probabilidades marginales de los límites inferior y superior de dicha clase. Se elige, aleatoriamente, un valor de la probabilidad entre dichos límites. Esta probabilidad se utiliza para determinar, a partir del histograma (sin desaglomerar) de los datos, el valor asignado a la celda. Esta probabilidad también se utiliza para determinar un valor real por simple interpolación lineal entre los límites de la clase si la clase es intermedia, o por interpolación monómica si es una clase extrema. En resumen, el primer valor real tiene un histograma igual que el de los datos duros y el segundo valor real tiene un histograma que viene definido por los umbrales que definen las variables indicadoras y sus probabilidades marginales del fichero **isim3d.var**. Si

las frecuencias relativas acumuladas calculadas a partir de los datos duros sin desaglomerar que corresponden a los umbrales que definen las variables indicadoras son muy distintas de las probabilidades marginales asignadas a dichos umbrales en el archivo **isim3d.var**, los dos modelos pueden resultar en realizaciones con histogramas muy distintos. Excepto cuando se dispone de muchos datos duros (más de cincuenta) sin grandes aglomeraciones espaciales, es preferible utilizar los resultados del modelo basado en las probabilidades marginales de los umbrales indicadas en **isim3d.var**. Los valores **mínimo** y **máximo** especificados en el fichero **isim3d.var** proporcionan los límites inferior y superior de la primera y última clase respectivamente. A aquellas celdas que corresponden a un dato duro se les asigna el valor de dicho dato.

Puesto que la selección de los valores reales dentro de cada clase se hace aleatoriamente sin considerar ningún tipo de correlación espacial con los valores asignados a celdas contiguas, el variograma de la variable continua siempre muestra un efecto pepita superior al observado en los datos, aún cuando los variogramas indicadores son fielmente reproducidos por las realizaciones. En general, este efecto pepita es pequeño, excepto si hay muy pocas clases discretizando un gran rango de variabilidad, y su impacto es mínimo. Para evitar este efecto pepita se puede recurrir a métodos más elaborados para hacer la clasificación de los datos dentro de cada clase, posiblemente entremezclando las realizaciones así obtenidas con realizaciones gaussianas.

Fichero **isim3d.cou**

Fichero de salida con los resultados de la simulación sobre la malla gruesa. El interés fundamental de este archivo es mínimo. El formato de salida es muy compacto, está escrito en el orden usual de escritura de los archivos con valores sobre una malla regular, es decir, por capas y cada capa por filas, comenzando por la celda con coordenadas más pequeñas.

Fichero **isim3d.rst**

Fichero intermedio binario necesario para el reinicio de una simulación después de una terminación anormal.

5

EJEMPLO DE APLICACION

Para demostrar el uso de los programas descritos se ha utilizado el conjunto de datos Berea descrito por Giordano *et al.* (1985), y consistente en 1600 medidas de permeabilidad con minipermeámetro distribuidas regularmente en una malla de 40 por 40 sobre una sección de arenisca de Berea de 61 cm por 61 cm. Este conjunto de datos ha sido aumentado con un conjunto de 1600 medidas de porosidad generadas sintéticamente, asegurando la existencia de cierta correlación cruzada entre ambas variables.

Por conveniencia en la descripción de los conjuntos de datos y del análisis de los mismos se utiliza el milidarcy (md) como medida de permeabilidad $1 \text{ md} \approx 10^{-5} \text{ m/s}$ cuando el fluido es agua y el tanto por ciento como medida de la porosidad. Igualmente por conveniencia se utilizará como paso de malla 1 unidad sin especificar dimensiones.

Las figuras 5 y 6 muestran, en un mapa de grises, la distribución espacial de los datos de ambas muestras. En ambas figuras se observa la alta continuidad, tanto de los valores de permeabilidad como de los de porosidad en la dirección N57W; a destacar la barrera de baja permeabilidad y baja porosidad en la esquina inferior izquierda y la zona de alta permeabilidad y alta porosidad en la esquina superior derecha.

Los histogramas y algunos estadísticos de permeabilidad y porosidad se muestran en las figuras 7 y 8 respectivamente. Ambos histogramas pueden considerarse como representativos de una muestra procedente de una distribución normal. El coeficiente de variación de ambas muestras no es muy alto. La figura 9 muestra un gráfico con los valores de permeabilidad y porosidad correspondientes a las 1600 mediciones. La recta de regresión de permeabilidad en porosidad también se muestra en dicha figura. El coeficiente de correlación entre ambas variables es de 0.64.

De este conjunto exhaustivo de datos se ha seleccionado un conjunto de 16 datos duros (al que nos referiremos como **b16**) y un conjunto de 500 datos blandos (al que nos referiremos como **b500**). El conjunto **b16** contiene datos de permeabilidad y porosidad en 16 puntos elegidos por muestreo aleatorio estratificado. Tras dividir el área en estudio en 16 paneles de igual tamaño, se ha elegido una muestra al azar en cada uno de los paneles. En cada uno de los 16 puntos se ha muestreado permeabilidad y porosidad. Las figuras 10 y 11 muestran la posición y valores de permeabilidad y porosidad del conjunto **b16**. Los valores numéricos correspondientes a estas figuras se muestran en la figura 12.

El conjunto **b500** pretende asemejarse al tipo de información que podría obtenerse de una campaña de geofísica sísmica en 3-D. Se considera que se dispone de información blanda a lo largo de cada una de las 14 líneas (7 horizontales y 7 verticales) sísmicas a lo largo de las filas 5, 10, 15, 20, 25, 30 y 35 y a lo largo de similar número de columnas. En total son $40 \times 14 = 560$ medidas de las que se han eliminado aquellas que coinciden con otro dato blando o con alguno de los datos duros. Se supone que se dispone de un conjunto de datos de calibración que nos permite acotar la variabilidad de la permeabilidad en cada uno de los 500 puntos en los que se dispone de medidas geofísicas. Las medidas geofísicas permiten acotar el valor de la permeabilidad en un punto dado a pertenecer a uno de los siguientes intervalos: por debajo de 55 md, entre 45 y 65 md, por encima de 55 md, por debajo de 65 md y por encima de 45 md. Puesto que los límites elegidos (45, 55 y 65 md) coinciden con los cuartiles y la mediana del conjunto exhaustivo (ver figura 7), los datos blandos son capaces de reducir en un 50% en los tres primeros intervalos, y en un 25% en los dos restantes, la incertidumbre acerca del valor real. En cada uno de los 500 puntos con datos blandos se ha elegido un intervalo de entre los 5 descritos que contenga el valor real. La figura 13 muestra la posición de los datos blandos y los intervalos que representan. Considerando la información que aporta cada dato blando se puede considerar que el conjunto **b500** contiene la misma cantidad de información que un conjunto de unos 50 datos.

Para la demostración de los programas de estimación y simulación no paramétricas se han seleccionado tres variables indicadoras, correspondiendo con el primer cuartil, la mediana y el tercer cuartil de la distribución exhaustiva de datos de permeabilidad. Estos valores son, respectivamente, 45, 55 y 65 md.

Puesto que el objetivo de esta demostración de los programas no es ni la inferencia de los variogramas ni un análisis de la robustez de la estima de los variogramas poblacionales a partir de los variogramas muestrales, se ha decidido utilizar como variogramas de entrada a los distintos programas aquellos que se obtienen del análisis del conjunto exhaustivo de datos.

Se han calculado y modelizado los variogramas (y variogramas cruzados) exhaustivos necesarios para los distintos programas. Aprovechando el que se dispone de un conjunto exhaustivo de datos distribuido sobre una malla regular, se ha calculado el valor de los variogramas para todos los vectores (h_x, h_y) , donde h_x y h_y son valores enteros entre 0 y 20. Las figu-

ras 14 a 19 muestran, en función de h_x y h_y , el variograma de la permeabilidad, el variograma cruzado permeabilidad-porosidad, el variograma de la porosidad, el variograma indicador correspondiente al primer cuartil de la permeabilidad, el variograma indicador correspondiente a la mediana y el variograma indicador correspondiente al tercer cuartil.

A efectos de una mejor comparación, todos los variogramas han sido normalizados dividiéndolos por la varianza de la variable o la covarianza entre variables y se ha usado en todas las figuras la misma escala de grises.

Los variogramas de porosidad, permeabilidad y el variograma cruzado entre estas dos variables muestran todos ellos una mayor continuidad de las respectivas variables en la dirección N57W, con un alcance muy superior a 30 en dicha dirección y de aproximadamente de 10 en la dirección ortogonal. El efecto pepita es difícil de apreciar en estos gráficos. Un análisis más detallado de esas dos direcciones nos lleva a postular los siguientes modelos para estos variogramas:

Variograma de la permeabilidad γ_k

$$\gamma_k(h_x, h_y) = 37.4 + 211.7 \cdot \text{Sph}_{a_x=70, a_y=12, \alpha=-57}(h_x, h_y) \quad (5)$$

donde

$$\text{Sph}_{a_x, a_y, \alpha}(h_x, h_y) = \text{Sph} \left(\sqrt{\left(\frac{h_x \cos \beta - h_y \sin \beta}{a_x} \right)^2 + \left(\frac{h_x \sin \beta + h_y \cos \beta}{a_y} \right)^2} \right)$$

donde, $\beta = 90 - \alpha$ y $\text{Sph}(h)$ viene dado por la ecuación 3.

Variograma cruzado de permeabilidad-porosidad $\gamma_{k,\phi}$

$$\gamma_k(h_x, h_y) = 7.6 + 17.8 \cdot \text{Sph}_{a_x=70, a_y=12, \alpha=-57}(h_x, h_y)$$

Variograma de la porosidad γ_ϕ

$$\gamma_k(h_x, h_y) = 2.3 + 3.5 \cdot \text{Sph}_{a_x=70, a_y=12, \alpha=-57}(h_x, h_y)$$

Este modelo para la variabilidad conjunta de permeabilidad y porosidad representa un modelo de coregionalización lineal. La característica principal de este tipo de modelos es que todos los variogramas

contienen las mismas estructuras pero con distintos coeficientes; en este caso, todos los variogramas constan de un efecto pepita y de una estructura esférica con elipse de anisotropía con semieje mayor de 70 en un ángulo de 143° del eje x (acimut de -57° medido en el sentido de las agujas del reloj con respecto al eje y) y el semieje menor de 12. Finalmente, para asegurar la validez del modelo es necesario que las matrices que definen los coeficientes de cada estructura sean definidas positivas, lo que en este caso se cumple puesto que:

$$\begin{vmatrix} 37.4 & 7.6 \\ 7.6 & 2.3 \end{vmatrix} = 28.3 \geq 0$$

$$\begin{vmatrix} 211.7 & 17.8 \\ 17.8 & 3.5 \end{vmatrix} = 424.1 \geq 0$$

Los variogramas de las variables indicadoras muestran un comportamiento claramente no multigaussiano del conjunto de datos Berea. Si el conjunto de datos Berea fuera multigaussiano, los variogramas de las variables indicadoras serían idénticos para umbrales situados simétricamente respecto a la mediana; en este caso, los variogramas de las variables indicadoras correspondientes al primer y tercer cuartiles deberían ser iguales y no lo son (véase las figuras 17 y 19).

Asimismo, el alcance máximo en una variable multigaussiana se da para la variable indicadora correspondiente a la mediana y en este caso se da para el primer cuartil.

Las direcciones de máxima y mínima continuidad de todos los variogramas son aproximadamente las mismas que las del variograma de la permeabilidad, es decir, N57W y N33E. Tras un análisis más detallado de los variogramas en estas direcciones se postulan los siguientes modelos:

Variograma indicador $k < 45$

$$\gamma_I(h_x, h_y; 45) = 0.077 + 0.115 \cdot \text{Sph}_{a_x=90, a_y=7, \alpha=-57}(h_x, h_y)$$

Variograma indicador $k < 55$

$$\gamma_I(h_x, h_y; 55) = 0.010 + 0.149 \cdot \text{Sph}_{a_x=50, a_y=10, \alpha=-57}(h_x, h_y)$$

Variograma indicador $k < 65$

$$\gamma_I(h_x, h_y; 65) = 0.055 + 0.128 \cdot \text{Sph}_{a_x=30, a_y=10, \alpha=-57}(h_x, h_y)$$

5.1 Ejemplo de COKB3D

En este ejemplo se pretende estimar tanto las permeabilidades como las porosidades sobre la malla de 1600 puntos a partir del conjunto **b16**. Aunque, en este caso, en los 16 puntos se dispone de valores tanto de permeabilidad como de porosidad, no es necesario que esto sea así para la aplicación del método; podría haber puntos en los que no se dispusiera de valores de todas las variables.

La figura 20 muestra el archivo de parámetros de entrada utilizado para la estimación de la permeabilidad. El cuarto campo del tercer registro indica cuál es la variable principal que se va a estimar; en este caso, es la variable en la tercera columna del archivo de entrada de datos que corresponde a la permeabilidad. El campo siguiente (en este caso sólo uno) indica que la variable en la cuarta columna de **b16.dat** es la que se utiliza como variable secundaria. El tercer campo de este registro indica que el archivo de datos sólo contiene las coordenadas *x* e *y* de los datos. El programa, puesto que es para problemas tridimensionales, internamente asigna el valor **zmin** (leído más abajo) como tercera coordenada. Se ha elegido como tipo de cokrigeaje aquél que convierte los datos de las variables secundarias en datos con la misma media que la variable principal. Los valores medios que se utilizan para hacer esta transformación son los dados en el registro con las **medias**, no los valores medios de los datos. En el registro de **medias** se da primero la media de la variable principal, seguida de las medias de las variables secundarias en el mismo orden en el que se especificaron en el tercer registro. Puesto que el problema es bidimensional **nz** vale 1.

La búsqueda de datos para el cokrigeaje se extiende hasta una distancia de 45 para la variable principal y sólo de 12 para la variable secundaria. En cualquier caso, el número máximo de datos de la variable principal se limita a 8 y el número máximo de datos de la variable secundaria se limita también a 8.

Se supone que el soporte de los datos es igual al tamaño de las celdas sobre las que se va a hacer la discretización, por lo que las celdas no necesitan discretización en celdillas más pequeñas. Si el soporte de los datos hubiera sido, por ejemplo, 1/9 del tamaño de la discretización de la malla, podría haberse estimado el valor de la celda asumiendo que éste sería igual a la media aritmética de los 9 valores que, con el soporte de los datos, se hubieran podido estimar dentro de cada celda. Para ello habría que haber especificado una discretización de bloque de 3, 3, 1 en las direcciones *x*, *y* y *z*.

A continuación se especifican los tres variogramas que definen la variabilidad espacial conjunta de las variables principal y secundaria. El orden es indisinto puesto que las variables a las que corresponden estos variogramas se especifican con los índices *i* y *j* que aparecen antes de cada variograma. En este caso se han indicado: primero, el variograma de la variable principal (*i*=1, *j*=1), a continuación el variograma cruzado entre la variable principal y la primera variable secundaria (*i*=1, *j*=2) y, finalmente, el variograma de la variable secundaria (*i*=2, *j*=2).

Los parámetros que definen los variogramas provienen de la expresión de los modelos dada antes. Como ejemplo, el variograma de la permeabilidad, que viene dado por la ecuación (5), tiene 1 estructura, un efecto pepita de 37.4, esta única estructura es del tipo 1 (esférica) con una elipse de anisotropía cuyo semieje mayor tiene una longitud de 70 y una meseta de 211.7. La orientación de la elipse de anisotropía viene dada por un ángulo de -57° que forman el semieje mayor con el eje *y* medido en sentido dextrógiro. El factor de anisotropía **anis1** es el cociente entre los alcances menor y mayor: 12/70. El factor de anisotropía **anis2** no influye en los cálculos. Los ángulos de giro **ang2** y **ang3** valen 0.0 puesto que no se requieren más giros para alinear los ejes cartesianos con los ejes del elipsoide de anisotropía.

El archivo de entrada de datos con los 16 valores de permeabilidad y porosidad se muestra en la figura 12.

Parte del archivo de salida de resultados correspondiente a estos dos archivos de entrada se muestra en la figura 21. Un mapa de grises de los resultados se muestra en la figura 22. Una comparación de esta figura con la que muestra el conjunto exhaustivo de datos (figura 5) proporciona una clara idea del efecto suavizador del krigeaje. La variabilidad del campo obtenido por krigeaje es mucho menor que la del campo real. La razón fundamental de esta diferencia es que la estimación por krigeaje no pretende obtener un mapa con unas características multivariadas similares a las de los datos originales, sino que pretende obtener un mapa que localmente sea "el mejor" en el sentido de mínimos cuadrados. Por ello, cada punto del mapa krigado es "el mejor" estimador dados los datos, considerado independientemente del resto de valores estimados. Cuando todos los valores estimados se consideran en conjunto, el mapa así representado será un buen estimador del mapa real, dependiendo de la aplicación que se vaya a dar al mismo: para la estimación de las reservas globales de una mina sí será un buen estimador pero no para el análisis de transporte de radionucleidos, puesto que la es-

timación siempre infraestimarán los valores altos, sobreestimarán los valores bajos y enmascarará la variabilidad real de la variable.

Similarmente a como se ha estimado la permeabilidad se puede estimar la porosidad. El resultado de esta estimación se muestra en la figura 23.

5.2 Ejemplo de IK3D y POSTIK

En este ejemplo se pretende estimar los valores de la permeabilidad utilizando la técnica de estimación por krigeaje indicador. Se muestran dos ejemplos: el primero sólo utiliza los 16 datos duros de permeabilidad contenidos en el archivo **b16**, el segundo utiliza también los datos blandos del conjunto **b500**.

El uso de **IK3D** frente a **COKB3D** puede justificarse por las características no gaussianas mostradas por los variogramas indicadores o por la capacidad de incorporar información imprecisa.

Como ya se indicó, se han elegido tres umbrales para discretizar el rango de variabilidad de la permeabilidad coincidiendo con el primer cuartil (45 md), la mediana (55 md) y el tercer cuartil (65 md) del conjunto exhaustivo de datos.

El archivo con los parámetros de entrada se muestra en la figura 24. Este fichero es el mismo tanto si se utilizan como si no los datos blandos. En el caso en que no se utilicen los datos blandos, el fichero de datos blandos indicado en el tercer registro no debe existir, aunque este registro debe incluirse siempre en el archivo de parámetros.

Se utiliza un entorno de búsqueda de forma elipsoidal definida por su semieje mayor y las rotaciones y anisotropías definidas en 2.2.

Se ha elegido realizar krigeaje indicador completo en lugar de krigeaje mediano. Asimismo, se ha elegido el krigeaje indicador simple frente al krigeaje indicador ordinario.

Finalmente, se indican los variogramas indicadores. En este caso, el orden sí es importante y deben indicarse los variogramas en orden creciente de los umbrales. Los parámetros definiendo los variogramas indicadores son similares a los descritos en el ejemplo anterior, con la salvedad de que aquí hay que introducir, para cada umbral, el valor de probabilidad marginal acumulada.

El archivo de datos duros de entrada se muestra en la figura 12, parte del archivo de datos blandos se muestra en la figura 25. Este archivo contiene una séptima columna con el valor real de la variable para comprobación de que el dato blando es coherente con el correspondiente dato duro. Nótese que en este ar-

chivo de entrada no se permite la omisión de la tercera coordenada que ha tenido que ser proporcionada explícitamente. Para el caso de simulación bidimensional, en el que la tercera coordenada no ha sido especificada en el archivo de entrada de datos duros, este valor debe ser igual al valor **zmin** del archivo de parámetros. A las variables indicadoras indefinidas se les ha asignado un valor fuera de los límites indicados por el cuarto registro del fichero de parámetros.

El krigeaje indicador proporciona estimadores de las probabilidades condicionales en cada punto. Estas probabilidades condicionales deben ser procesadas por el programa **POSTIK** para producir un estimador de la variable o de la probabilidad de superar un cierto umbral. La figura 26 muestra parte del archivo de salida para el caso que sólo se utilizan datos duros.

La figura 27 muestra el archivo de parámetros de entrada a **POSTIK** para el cálculo del valor esperado de la función de distribución condicional. El significado de los parámetros utilizados en este archivo se explicó en la descripción de este programa.

Los valores mínimo y máximo sirven para acotar los límites extremos de la primera y última clases.

Parte del archivo de salida correspondiente a este archivo de parámetros de entrada se muestra en la figura 28. Las figuras 29 y 30 muestran sendos mapas de grises con los valores medios de las funciones de probabilidad condicional resultantes del krigeaje indicador, sin incluir datos blandos e incluyéndolos, respectivamente.

Estas dos figuras muestran el problema inherente a la inclusión de datos blandos de tipo intervalo en la estimación y es que los resultados correspondientes a la primera y última clases siempre son desvirtuados. La inclusión de información blanda abundante hace aumentar la variabilidad espacial del correspondiente mapa estimado pero, a la vez, infraestima la proporción de datos en las clases extremas.

Similarmente, se han obtenido mapas de probabilidad de que la permeabilidad esté por debajo de 45 md, utilizando las distribuciones de probabilidad condicionales obtenidas sin datos blandos y con ellos. Estos se muestran en las figuras 31 y 32. En este caso y puesto que se ha elegido la probabilidad de pertenecer a la clase extrema más baja, la inclusión de los datos blandos desvirtúa los resultados, siendo más plausibles, a la vista del conjunto de datos exhaustivo, los resultados obtenidos utilizando sólo los 16 datos duros.

Finalmente, se han obtenido también los mapas correspondientes a la permeabilidad que tiene una pro-

babilidad de no ser superada del 20%. Estos mapas corresponderían a un estimador que en término medio infraestima la variable pero que puede corresponder a un estimador óptimo en el caso en que la pérdida asociada con una infraestimación sea menor que la asociada con una sobreestimación. Las figuras 33 y 34 muestran sendos mapas de grises correspondientes a la estimación sin datos blandos y con ellos.

5.3 Ejemplo de GCOSIM3D

En este ejemplo se pretende generar realizaciones de permeabilidad y porosidad que sean fieles a los datos medidos y que muestren la variabilidad que se ha inferido del conjunto de datos exhaustivo, medida por los variogramas y variogramas cruzados de permeabilidad y porosidad.

La figura 35 muestra el archivo de entrada con los parámetros que definen la geometría del sistema. El número de realizaciones a generar es de 5. Aunque la simulación es de una malla de dos dimensiones, como se indica en el primer registro, el programa es tridimensional y por ello necesita que explícitemente se le proporcionen los valores de los parámetros para las tres dimensiones. Por esto, el número de niveles (n_z) es de 1, no 0. Puesto que se desea generar toda la malla, los valores de los registros con los parámetros **desde** y **hasta** son 1,1,1 y 40,40,1, respectivamente. El entorno de búsqueda se especifica proporcionando los valores de los semiejes del elipsoide alineados con un sistema $X'Y'Z'$, en este caso 50, 14, 0.1, y con la matriz de rotación que transforma los ejes XYZ en los ejes $X'Y'Z'$.

La figura 36 muestra el fichero en el que se definen los variogramas de las variables que se van a utilizar en la simulación. El primer registro indica que se dispone de datos de 2 variables, el segundo registro que se van a simular 2 variables y el tercer registro que las variables a simular son las que tienen índices 1 y 2. A continuación siguen las medias de las variables para las que se dispone de datos (independientemente de si todas se van a simular o no) y los variogramas y variogramas cruzados entre todas las variables de las que se dispone de datos, empezando con el variograma de la variable con índice 1, seguido con el variograma cruzado entre las variables 1 y 2 y acabando con el variograma de la variable 2. El elipsoide de anisotropía de estos variogramas se expresa de la misma manera que el elipsoide de búsqueda, indicando los alcances en las tres direcciones de un sistema de ejes cartesianos $X'Y'Z'$ y la matriz de rotación que transforma el sistema XYZ en el $X'Y'Z'$.

La figura 37 muestra el fichero de datos con los mismos 16 valores de permeabilidad y porosidad utilizados en el ejemplo de estimación. En este caso, cada dato, ya sea de porosidad o de permeabilidad, ocupa un registro, por lo que el fichero contiene 32 registros. También es necesario dar la tercera coordenada que se podía omitir en los ejemplos de estimación. El cuarto campo de cada registro es un entero indicando la variable a la que corresponde el dato; en este caso, la variable con índice 1 es la permeabilidad y la 2 la porosidad.

Un fragmento del archivo de resultados se muestra en la figura 38, que contiene dos columnas con los valores de permeabilidad y porosidad de las 5 simulaciones de 1600 valores. Las simulaciones de permeabilidad se muestran en la figura 39. Las correspondientes simulaciones de porosidad se muestran en la figura 40. Cada una de las imágenes en esta última figura va emparejada con la imagen en la figura anterior que ocupa la misma posición.

5.4 Ejemplo de ISIM3D

Finalmente se incluye un ejemplo de simulación secuencial de indicadores con y sin información blanda.

La figura 41 muestra el archivo con los parámetros que describen la geometría del problema. Al igual que en el caso anterior, aunque el problema es bidimensional, hay que describirlo como tridimensional con un único nivel en la dirección z . La única peculiaridad de este fichero con el equivalente para simulación gaussiana, es la posibilidad de definir una malla gruesa sobre la que realizar una primera simulación, con el fin de reducir el tamaño del entorno de búsqueda en la simulación final sobre la malla fina. En este caso, el uso de esta posibilidad es puramente simbólico, porque se puede definir un entorno de búsqueda en la malla fina que comprenda la totalidad del área a simular sin incurrir en un entorno de tamaño desproporcionado. Por esto, el paso de malla gruesa reduce ésta a un único punto, ($\text{gruesa_delta.x} = \text{gruesa_delta.y} = 40$). El tamaño del entorno de búsqueda es el mismo tanto para la malla gruesa como para la fina.

La figura 42 muestra el fichero con la definición de los variogramas de las variables indicadoras. Después del número de umbrales se indica el mínimo y el máximo que acotan la primera y la última clases. A estos siguen la definición de los variogramas en orden creciente del umbral.

La figura 43 muestra el fichero con los datos duros. Nótese la diferencia con el fichero de entrada de datos al programa de simulación gaussiana, en que el número de datos no debe especificarse en el primer registro y que el cuarto campo en el que se indicaba el índice de la variable no existe.

La figura 44 muestra parte del fichero de entrada con los datos blandos del conjunto **b500**. Este fichero sólo existe cuando se desean utilizar los datos blandos.

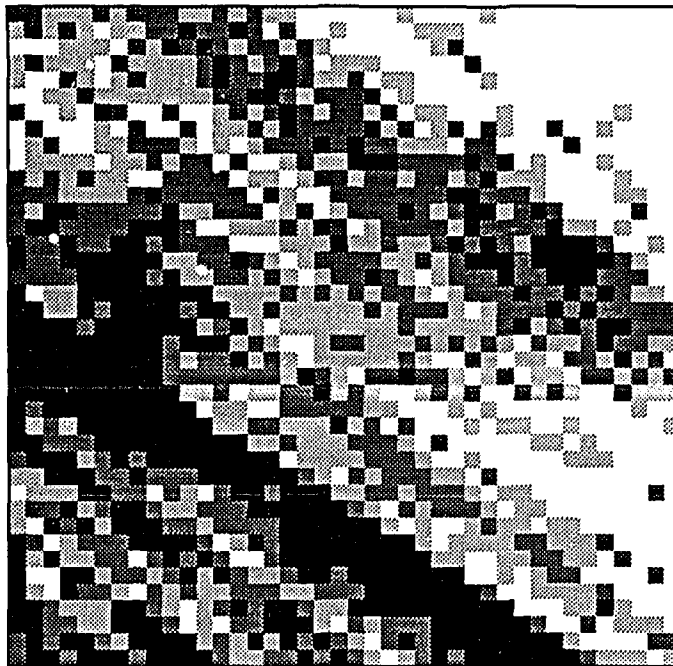
La figura 45 muestra parte del fichero de salida del programa **ISIM3D**. En este caso, y dado que con 16 datos duros no se puede construir una función de frecuencias acumuladas significativa, la segunda columna que está basada en el histograma de los datos del

conjunto **b16** proporciona resultados menos fiables que la tercera columna que, al menos, respeta las probabilidades marginales especificadas en el fichero **isim3d.var**.

La figura 46 muestra cinco realizaciones obtenidas con **ISIM3D** utilizando sólo los datos duros, mientras que la figura 47 muestra otras cinco realizaciones generadas incluyendo los datos blandos.

La ventaja de utilizar la información blanda es evidente en este caso. La variabilidad que se observa entre las distintas realizaciones es mucho más pequeña que la que se observa entre las realizaciones generadas sin considerar esta información blanda, tanto para el caso gaussiano como para el de indicadores.

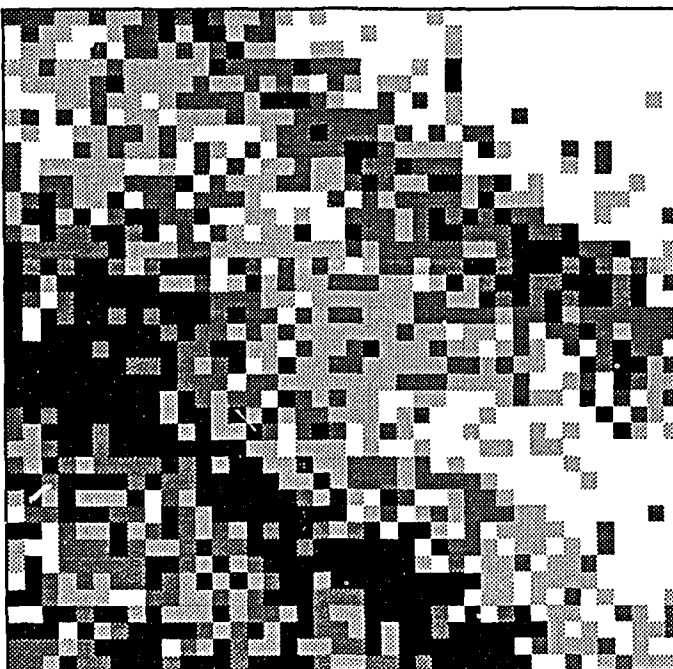
Base de datos
Berea
Permeabilidad



milidarcys
 65.0 - 111.5
 55.0 - 65.0
 45.0 - 55.0
 19.5 - 45.0

Figura 5
Mapa de grises de permeabilidades de la base de datos Berea.

Base de datos
Berea
Porosidad



Porcentaje
 11.4 - 18.5
 9.9 - 11.4
 8.4 - 9.9
 4.7 - 8.4

Figura 6
Mapa de grises de porosidad de la base de datos Berea.

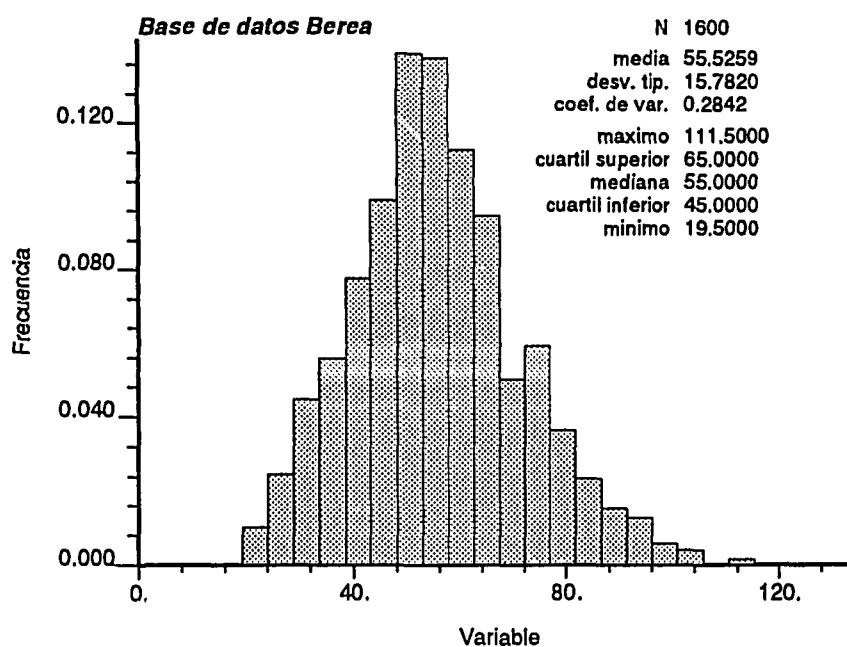


Figura 7
 Histograma de permeabilidades de la base de datos Berea.

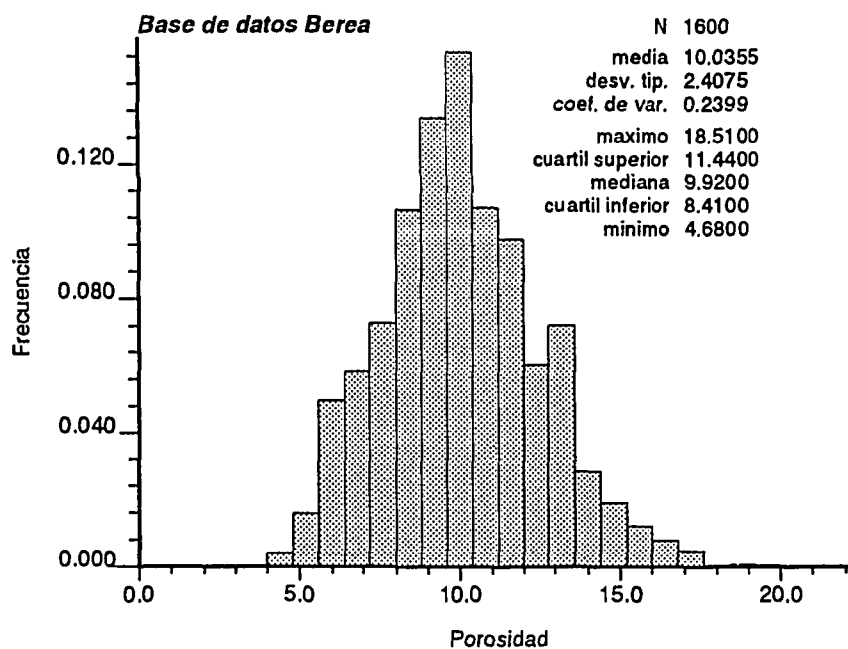


Figura 8
 Histograma de porosidades de la base de datos Berea.

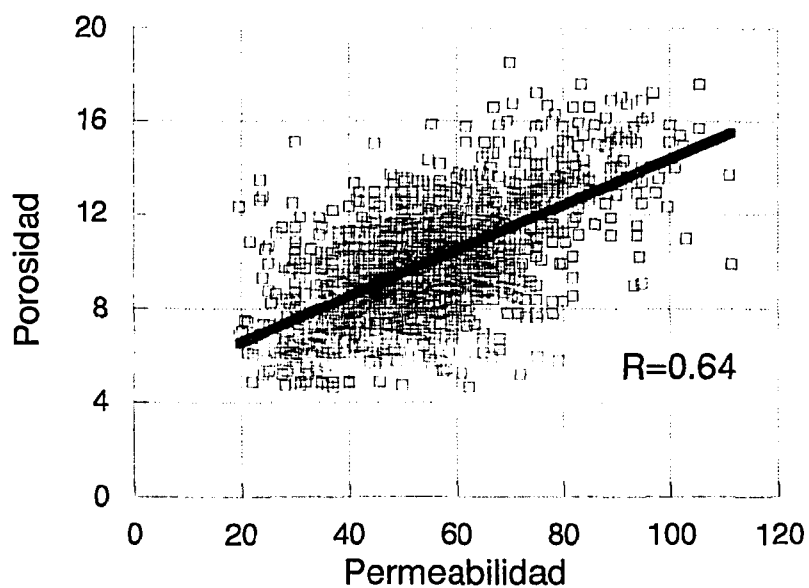


Figura 9
Correlación entre K y ϕ en la base de datos Berea.

Conjunto
de Datos B16
Permeabilidad

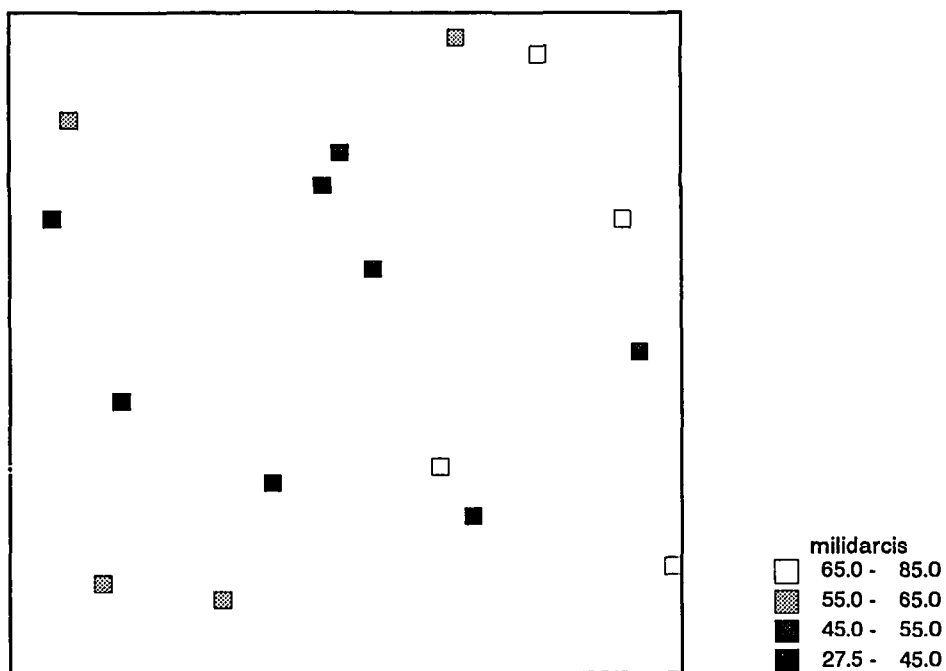


Figura 10
Permeabilidades del conjunto b16.

Conjunto
de Datos B16
Porosidad



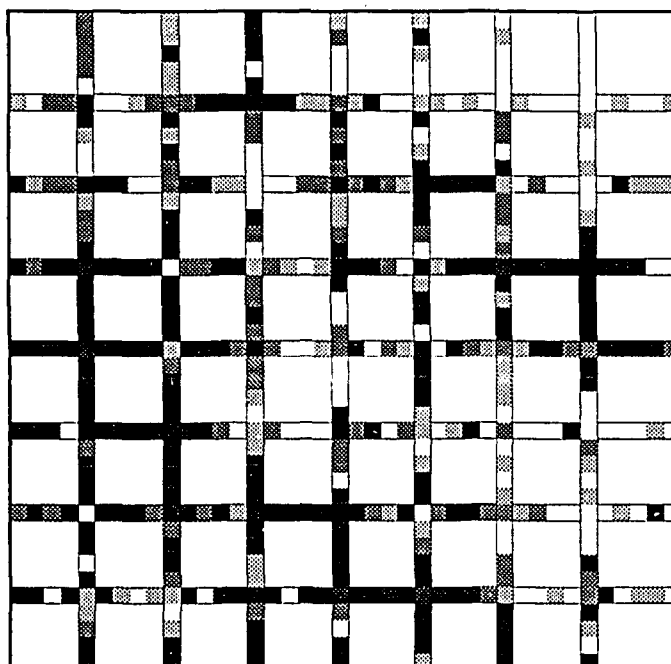
Figura 11
Porosidades del conjunto b16.

Sample of 16 data of k and phi

4			
Easting			
Northing			
Perm			
Synth. Por.			
6	6	58.50	10.53
7	17	28.50	5.37
3	28	40.00	8.18
4	34	56.00	9.62
13	5	55.00	9.09
16	12	27.50	5.67
19	30	50.00	10.38
20	32	47.00	8.71
28	10	47.50	9.70
26	13	67.50	11.82
22	25	50.00	10.38
27	39	64.00	10.99
40	7	85.00	13.42
38	20	50.00	9.16
37	28	66.00	14.63
32	38	82.00	16.61

Figura 12
Datos duros. Valores numéricos de los datos mostrados en las figuras 10 y 11. Fichero b16.dat en el disco de distribución

Conjunto
de Datos B500
Intervalos de
permeabilidad



intervalos

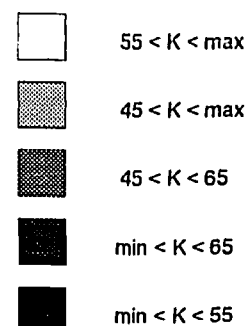


Figura 13
Conjunto de datos blandos.

Permeabilidad

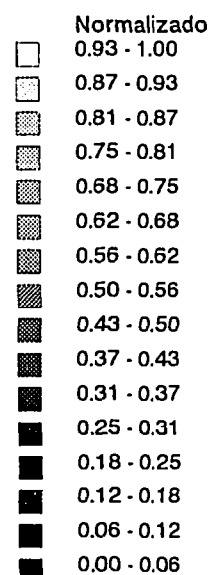
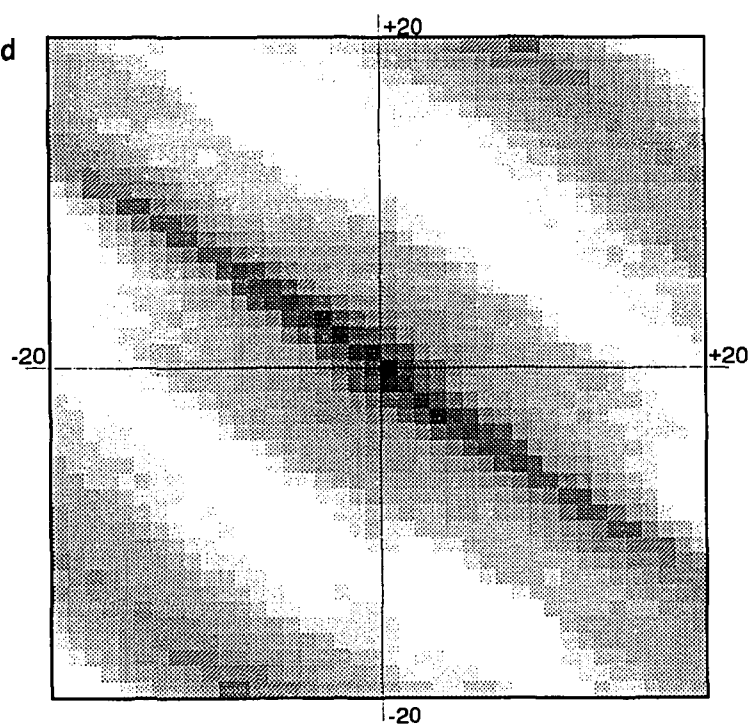


Figura 14
Variograma de permeabilidad de la base de datos Berea.

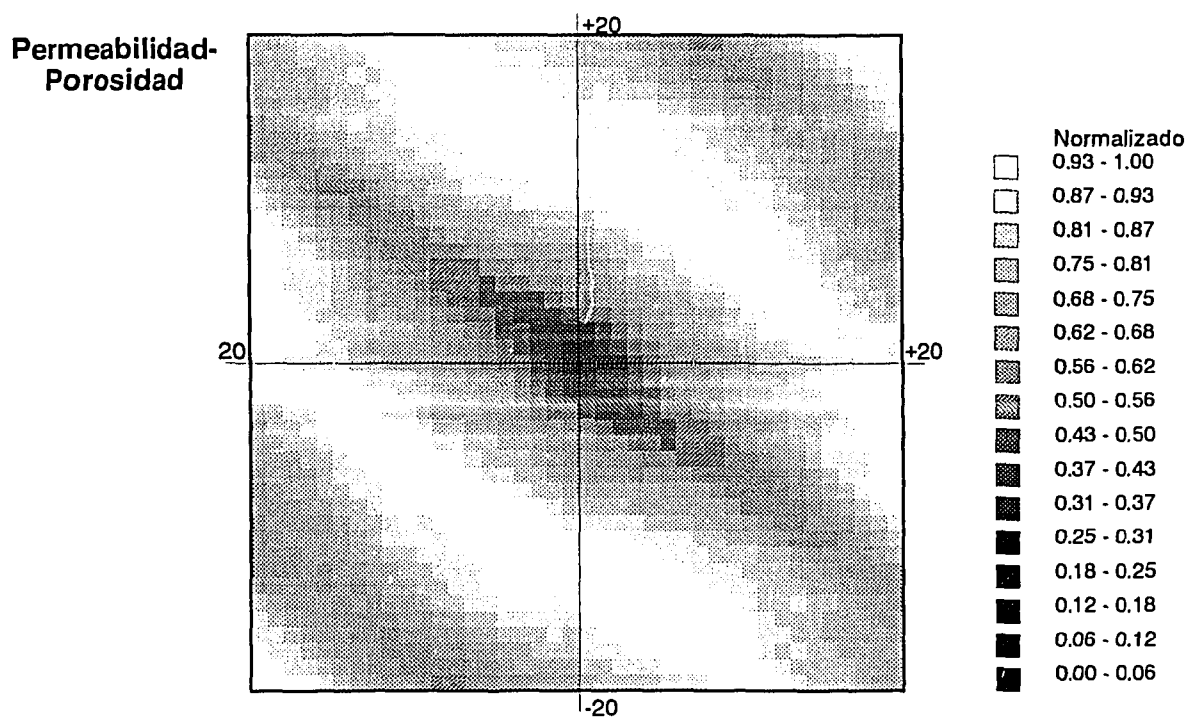


Figura 15
Variograma cruzado de permeabilidad-porosidad de la base de datos Berea.

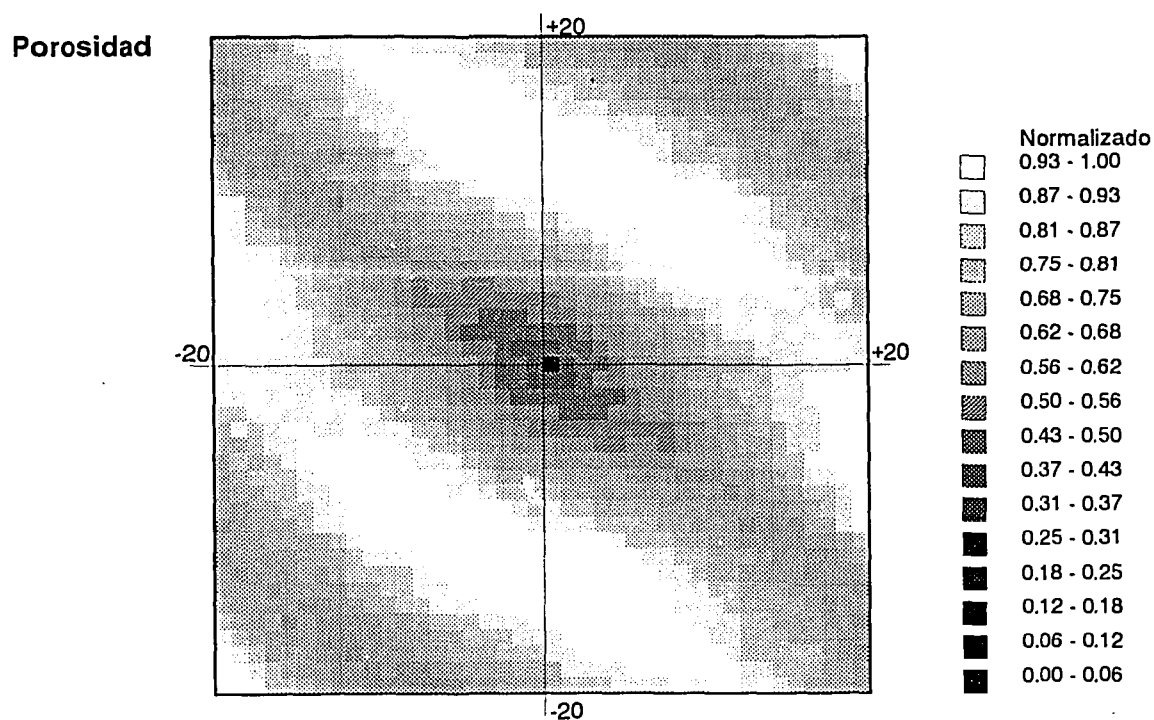


Figura 16
Variograma de porosidad de la base de datos Berea.

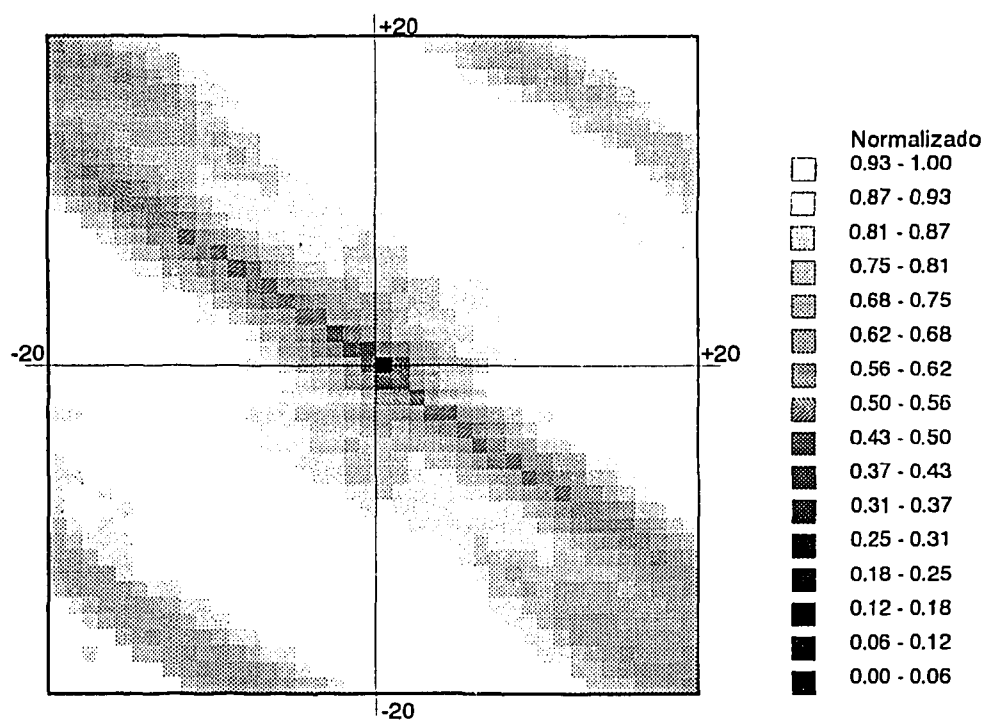
Cuartil
Inferior

Figura 17

Variograma indicador de la base de datos Berea. Umbral en el cuartil inferior.

Mediana

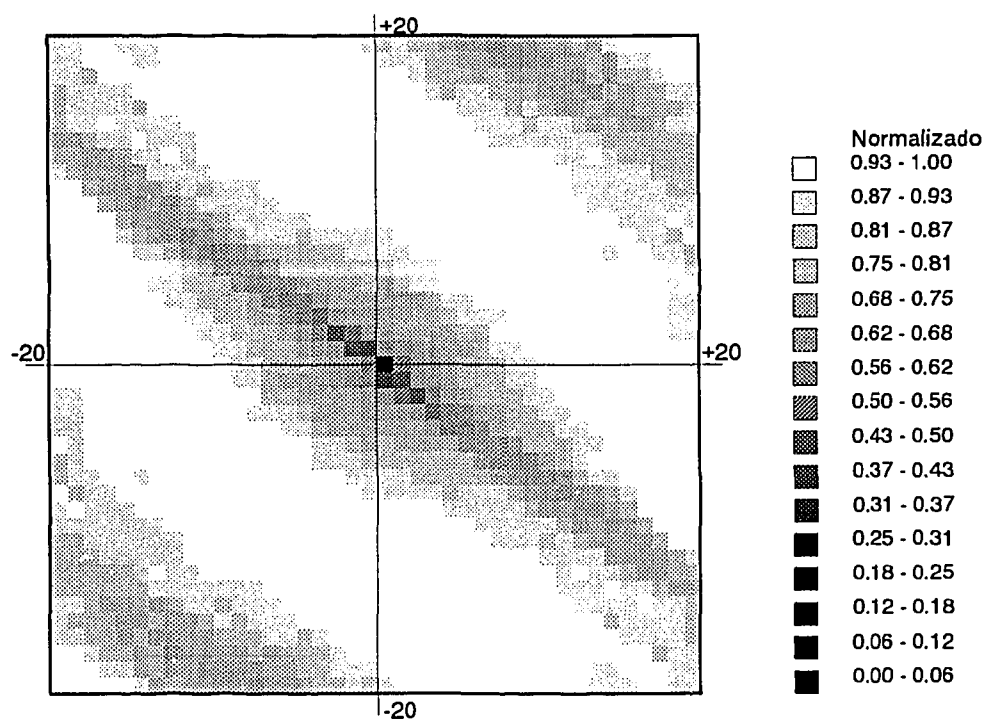


Figura 18

Variograma indicador de la base de datos Berea. Umbral en el cuartil mediana.

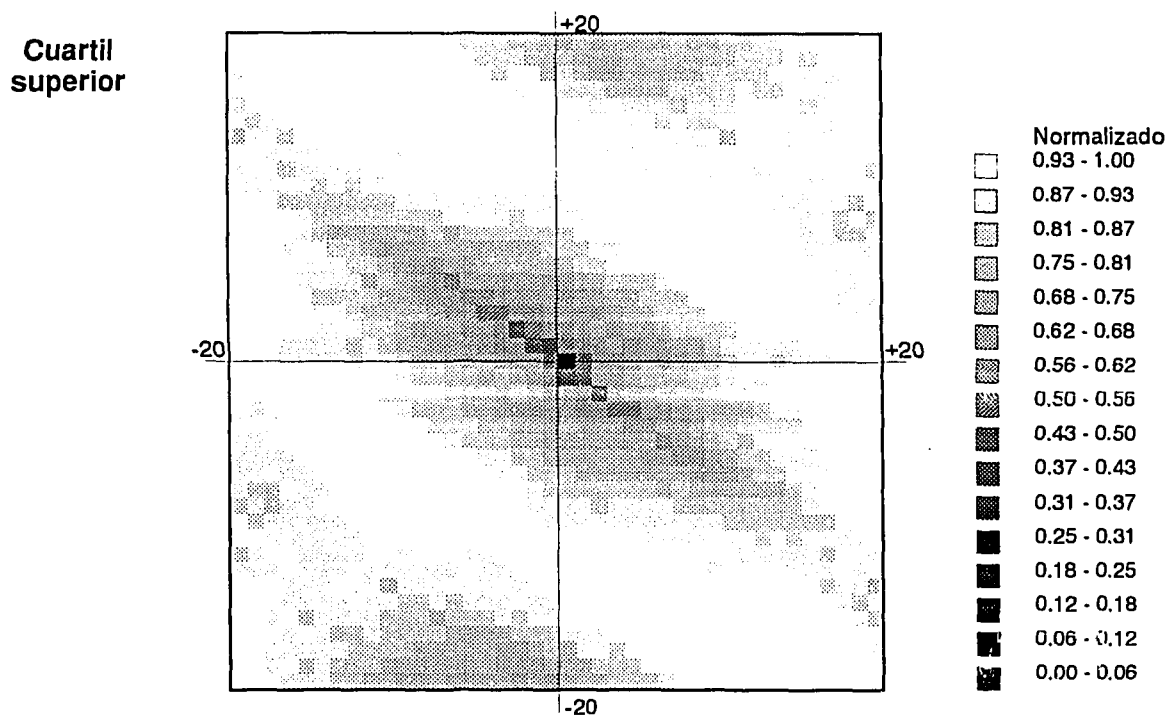


Figura 19
Variograma indicador de la base de datos Berea. Umbral en el cuartil superior.

Fichero con los parametros de entrada para el programa COKB3D

COMIENZO DE PARAMETROS:

```

b16.dat          \ fichero de datos
2               \ numero de vars. principal y otras
1   2   0   3   4   \ col de x,y,z,primaria,secundarias
-1.0e21   1.0e21   \ limites para las variables
1             \ tipo de krig (0=KS,1=K0,2=K0-trad)
55.5   9.95     \ medias(i),i=1,nvar
cokb3dk.out      \ fichero de salida de resultados
0              \ 0,1,2,3
cokb3d.dbg       \ fichero de info. complementaria
40   1.0   1.0   \ nx,xmin,dx
40   1.0   1.0   \ ny,ymin,dy
1    1.0   1.0   \ nz,zmin,dz
1    1      1   \ x, y, and z discretizacion bloque
1    8      8   \ min, max princ., max secund.
45.0  12.0     \ maximo radio busqueda: princ,sec
1    1         \ variograma de "i" y "j"
1    37.4      \ num estruc, efecto pepita
1    70.0  211.7 \ tipo,alcance,meseta: estr 1
-57.  0.0   0.0  .17 1.0 \ angl,ang2,ang3,anis1,anis2:
1    2         \ variograma de "i" y "j"
1    7.6       \ num estruc, efecto pepita
1    70.0  17.8 \ tipo,alcance,meseta: estr 1
-57.  0.0   0.0  .17 1.0 \ angl,ang2,ang3,anis1,anis2:
2    2         \ variograma de "i" y "j"
1    2.3       \ num estruc, efecto pepita
1    70.0  3.5  \ tipo,alcance,meseta: estr 1
-57.  0.0   0.0  .17 1.0 \ angl,ang2,ang3,anis1,anis2:

```

Figura 20

Fichero de parámetros para COKB3D. Fichero cokb3d.par en el disco de distribución.

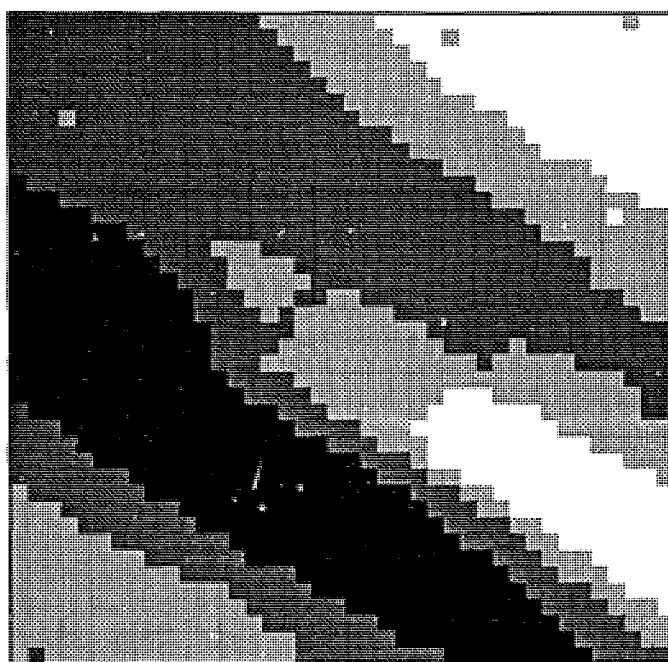
```

COKB3D with: Sample of 16 data of k and phi
2
estimate
estimation variance
55.6616      242.7905
54.9986      235.6221
55.1344      230.0646
55.2976      223.3306
55.4864      215.3044
55.6993      205.9016
.....
.....
64.4551      218.6974
66.6318      229.4262
65.7818      235.9678

```

Figura 21
Fichero de salida de COKB3D. Interpolación de permeabilidades.

**Cokrigeaje
gausiano
Permeabilidad**



Permeabilidad

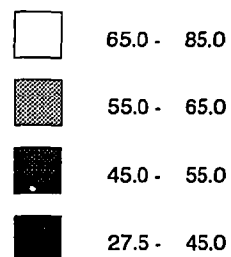
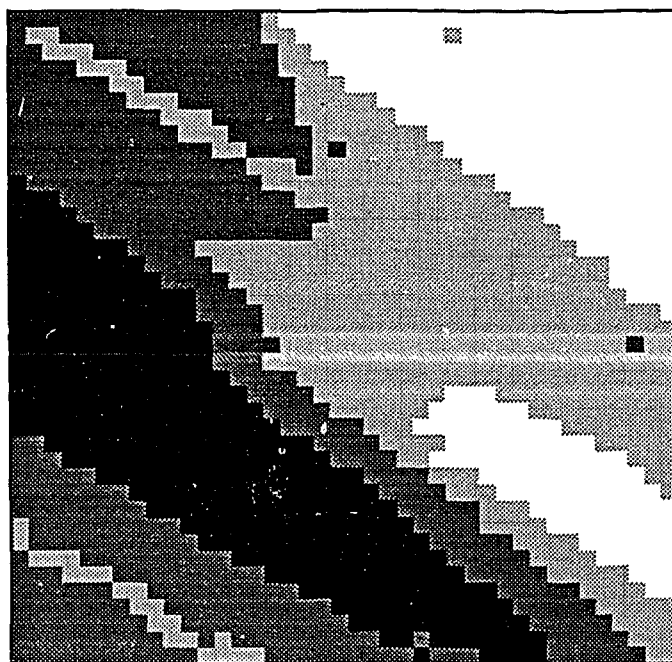


Figura 22
Resultados de COKB3D. Interpolación de permeabilidades.

Cokrigearje
gausiano
Porosidad



porcentaje

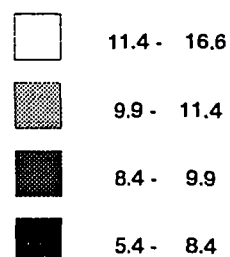


Figura 23
Resultados de COKB3D. Interpolación de porosidades.

Fichero con los parametros de entrada para el programa IK3D

COMIENZO DE PARAMETROS:

b16.dat	\fichero de datos
1 2 0 3	\col de x,y,z y variable
b500.dat	\fichero de datos blandos
-1 1.0e21	\limites para las variables
ik3dsi.out	\fichero de salida de resultados
2	\nivel de depuracion: 0.1,2,3
ik3d.dbg	\fichero de info. complementaria
40 1.0 1.0	\nx,xmin,dx
40 1.0 1.0	\ny,ymin,dy
1 1.0 1.0	\nz,zmin,dz
0 16	\min, max data para krigeaje
45.0	\maximo radio de busqueda
123 0.0 0.0 1.0 0.2	\busqueda: angl,2,3,anis1,2
0	\max por octante (0-> sin uso)
0 55. <- valor necesario	
aunque sin uso	
0	\0=KI completo, 1=KI med (umbral)
3	\0=KS, 1=K0
	\numero umbrales
45 0.26 1 0.077	\umbral, marg prob, nst, pepita
1 90.0 0.115	\ tipo, alcance, meseta
-57. 0.0 0.0 0.0777 1.0	\ angl,ang2,ang3,anis1,2
55 0.53 1 0.099	\umbral, marg prob, nst, pepita
1 50.0 0.149	\ tipo, alcance, meseta
-57. 0.0 0.0 0.2 1.0	\ angl,ang2,ang3,anis1,2
65 0.76 1 0.055	\umbral, marg prob, nst, pepita
1 30.0 0.127	\ tipo, alcance, meseta
-57. 0.0 0.0 0.333 1.0	\ angl,ang2,ang3,anis1,2

Figura 24

Fichero de parámetros para IK3D. Fichero ik3d.par en el disco de distribución.

```

Berea Direct IK
6
Easting
Northing
Dummy
K<=45
K<=55
K<=65
1      5      1      -99      1      1      28.000
2      5      1      -99     -99      1      56.000
3      5      1       0       0     -99      72.000
4      5      1     -99     -99      1      44.000
6      5      1     -99      1      1      45.000
7      5      1       0     -99     -99      56.000
8      5      1       0       0     -99      65.000
.....
.....
35     35      1       0       0     -99      90.000
35     36      1       0       0     -99      77.000
35     37      1       0       0     -99      82.000
35     38      1       0       0     -99      89.000
35     39      1       0       0     -99      85.000
35     40      1       0       0     -99      92.000

```

Figura 25

Fichero con datos blandos. A las variables indicadoras indeterminadas se les ha asignado un valor negativo fuera de los límites indicados en el archivo de parámetros.

```

IK3D Estimates with: Sample of 16 data of k and
phi
3
Cutoff: 1 at 45.00000
Cutoff: 2 at 55.00000
Cutoff: 3 at 65.00000
0.2158 0.5271 0.7791
0.2146 0.5069 0.7853
0.1955 0.4501 0.8198
.....
.....
0.1911 0.4877 0.5699
0.1959 0.5093 0.5934
0.1484 0.5466 0.5743

```

Figura 26

Fichero de salida de IK3D. Sólo se han utilizado los datos duros.

Archivo con los parametros de entrada a POSTIK

COMIENZO DE PARAMETROS:

ik3dno.out	\entrada (salida de IK3D)
postne.out	\fichero de salida
1 55.	\opcion salida, parametro
3	\numero de umbrales
45. 55. 65.	\umbrales
19 112	\valores minimo y maximo
2 2.5	\cola inf : opcion, parametro
1 1.0	\intermedio: opcion, parametro
2 .25	\cola sup : opcion, parametro
100	\maxima discretizacion

Figura 27

Fichero de parámetros para POSTIK. Opción de estimación de los valores esperados de las distribuciones condicionales de probabilidad.

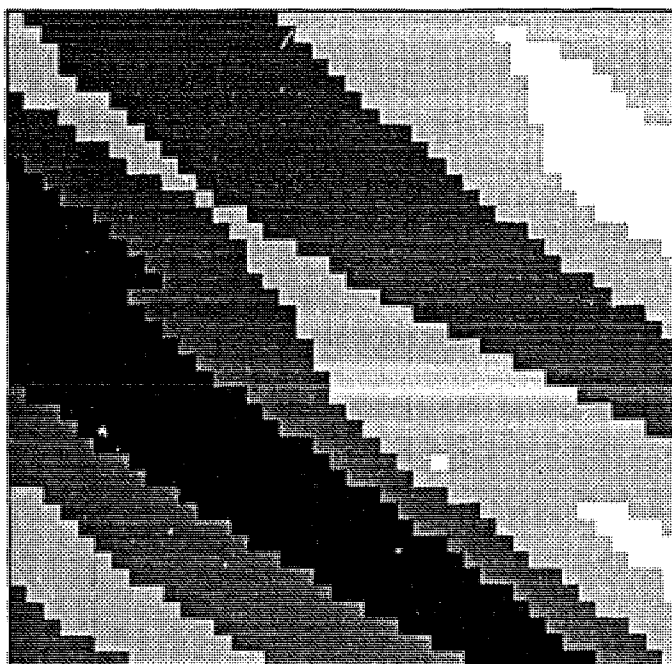
E-type mean values

1
mean
52.8208
52.9625
53.4829
53.6280
.....
.....
58.3232
57.5351
56.8086
56.1442
57.1643

Figura 28

Fichero de salida de POSTIK. Opción de cálculo del valor medio de la función de probabilidad de condicional.

Estimación
por indicadores
Permeabilidad



milidarcis

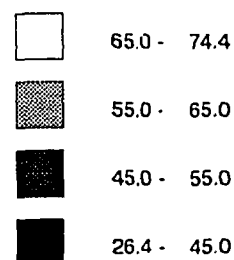
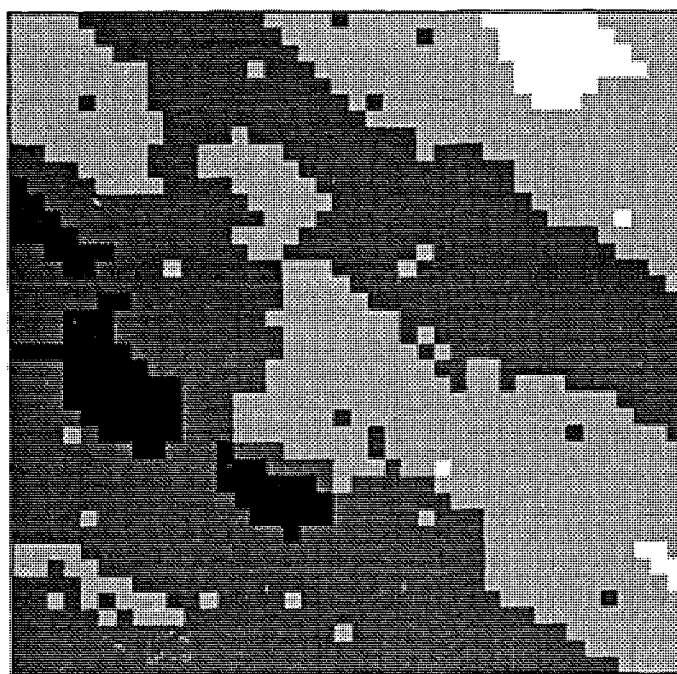


Figura 29

Resultado de IK3D tras proceso por POSTIK. Valores medios de las probabilidades condicionadas resultantes de IK3D sin considerar la información blanda.

Krigeaje
indicador blando
Permeabilidad



milidarcis

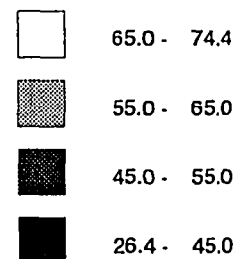


Figura 30

Resultado de IK3D tras proceso por POSTIK. Valores medios de las probabilidades condicionadas resultantes de IK3D considerando la información blanda.

Estimación
por indicadores
Probabilidad $K < 45$

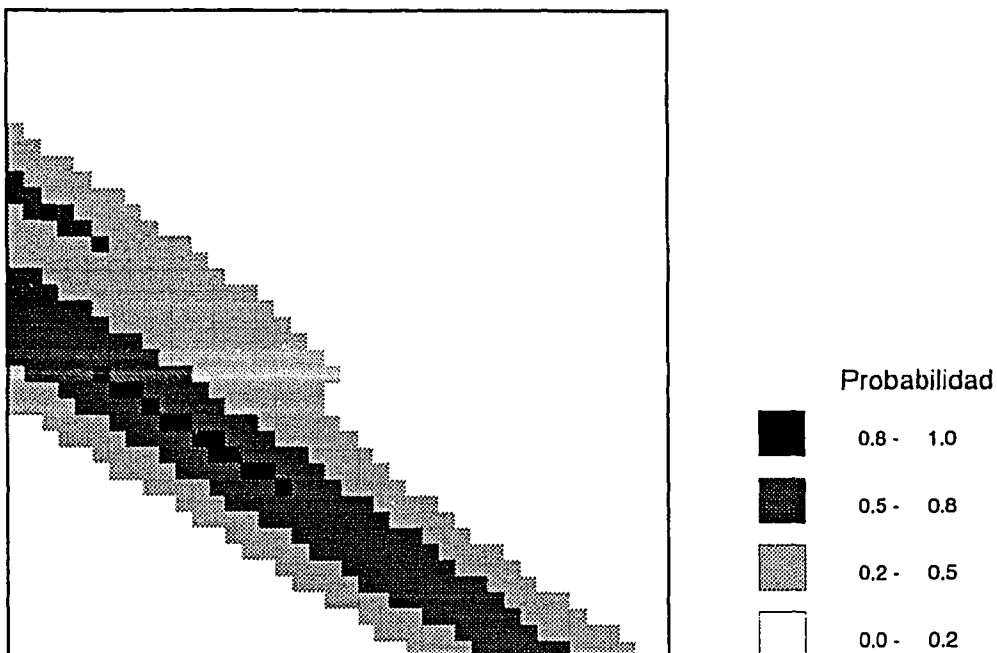


Figura 31

Resultado de IK3D tras proceso por POSTIK. Probabilidades de que la permeabilidad esté por debajo de 45 md.

Krigeaje
indicador blando
Probabilidad $K < 45$

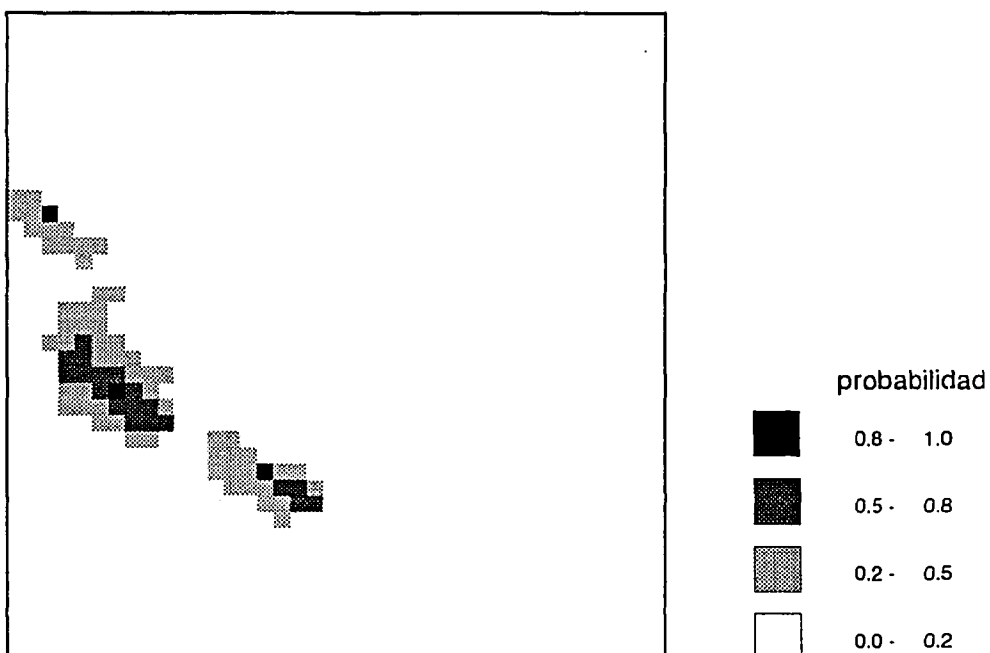
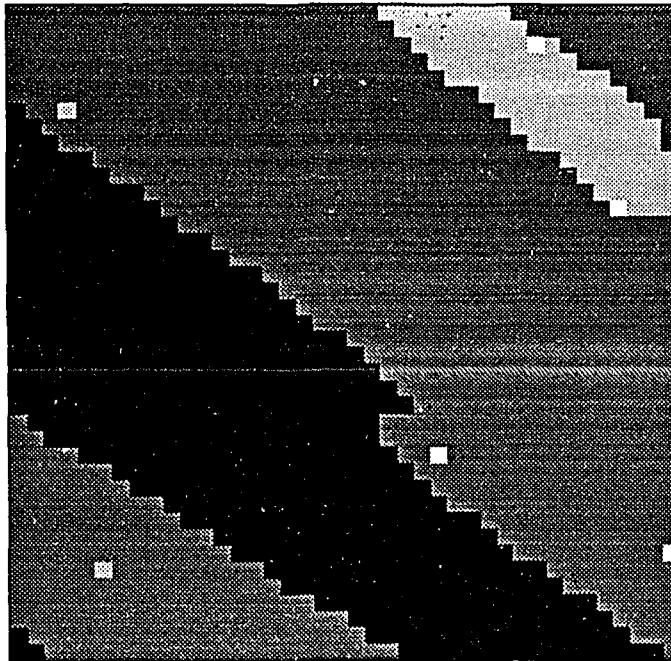


Figura 32

Resultado de IK3D tras proceso por POSTIK. Probabilidades de que la permeabilidad esté por debajo de 45 md. Considerando información blanda.

Estimación
por indicadores
Permeabilidad con
prob. <0.2



Permeabilidad

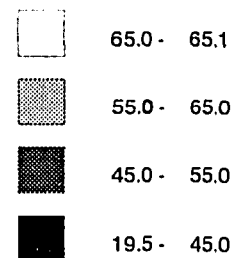
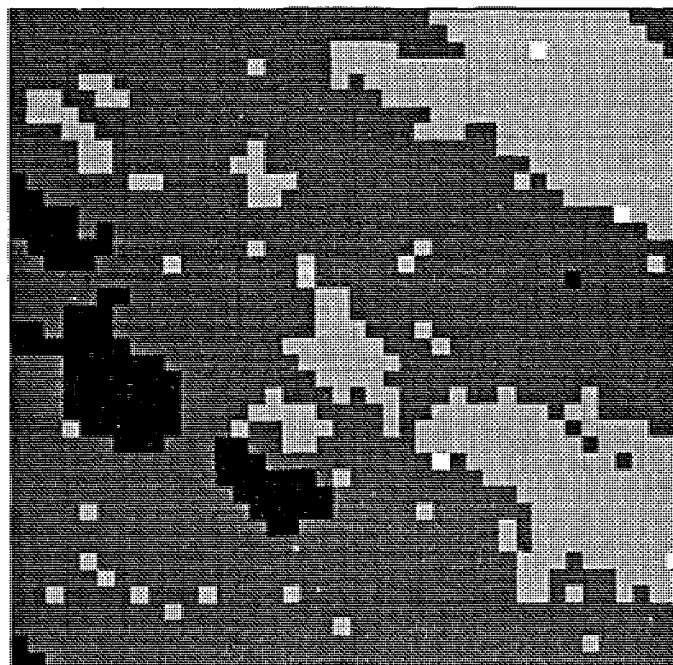


Figura 33

Resultado de IK3D tras proceso por POSTIK. Permeabilidades con probabilidad de 0.2 de no ser superadas.

Krigeaje
indicador blando
Permeabilidad con
prob. <0.2



milidarcis

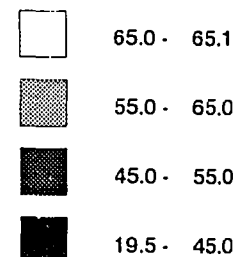


Figura 34

Resultado de IK3D tras proceso por POSTIK. Permeabilidades con probabilidad de 0.2 de no ser superadas. Considerando información blanda.

5. Ejemplo de aplicación

```

30 5 2
1 1 1
1.0 1.0 1.0
40 40 1
1 1 1
40 40 1

```

```

50 14 0.1

```

```

0.84 -0.54 0
0.54 0.84 0
0 0 1

```

```

3
0

```

Figura 35

Fichero de geometría de entrada a GCOSIM3D. Fichero gcosim3d.geo en el disco de distribución.

```

2
2
1 2
55.5 9.95

```

```

37.4
500.
1
1
211.7
70.0
12.0
1.

```

```

0.84 -0.54 0
0.54 0.84 0
0 0 1

```

```

7.6
500.
1
1
17.8
70.0
12.0
1.

```

```

0.84 -0.54 0
0.54 0.84 0
0 0 1

```

```

2.3
500.
1
1
3.5
70.0
12.0
1.

```

```

0.84 -0.54 0
0.54 0.84 0
0 0 1

```

Figura 36

Fichero de variogramas de entrada a GCOSIM3D. Fichero gcosim3d.var en el disco de distribución.

```

32
6 6 1 1 58.50
7 17 1 1 28.50
3 28 1 1 40.00
4 34 1 1 56.00
.....
.....
27 39 1 2 10.99
40 7 1 2 13.42
38 20 1 2 9.16
37 28 1 2 14.63
32 38 1 2 16.61

```

Figura 37

Fichero de datos duros de entrada a GCOSIM3D. Fichero gcossim3d.dat en el disco de distribución.

```

Output from GCOSIM3D,
5 simulations of
the subarea
2
Variable number 1
Variable number 2
42.382    7.456
53.633   10.498
52.705   11.062
55.697    8.810
-----
5 simulaciones
de 1600 datos
-----
42.757    8.729
49.023   10.973
52.910   12.117
47.653   10.325

```

Figura 38

Fichero de salida de GCOSIM3D. Co-simulación gaussiana de permeabilidad y porosidad.

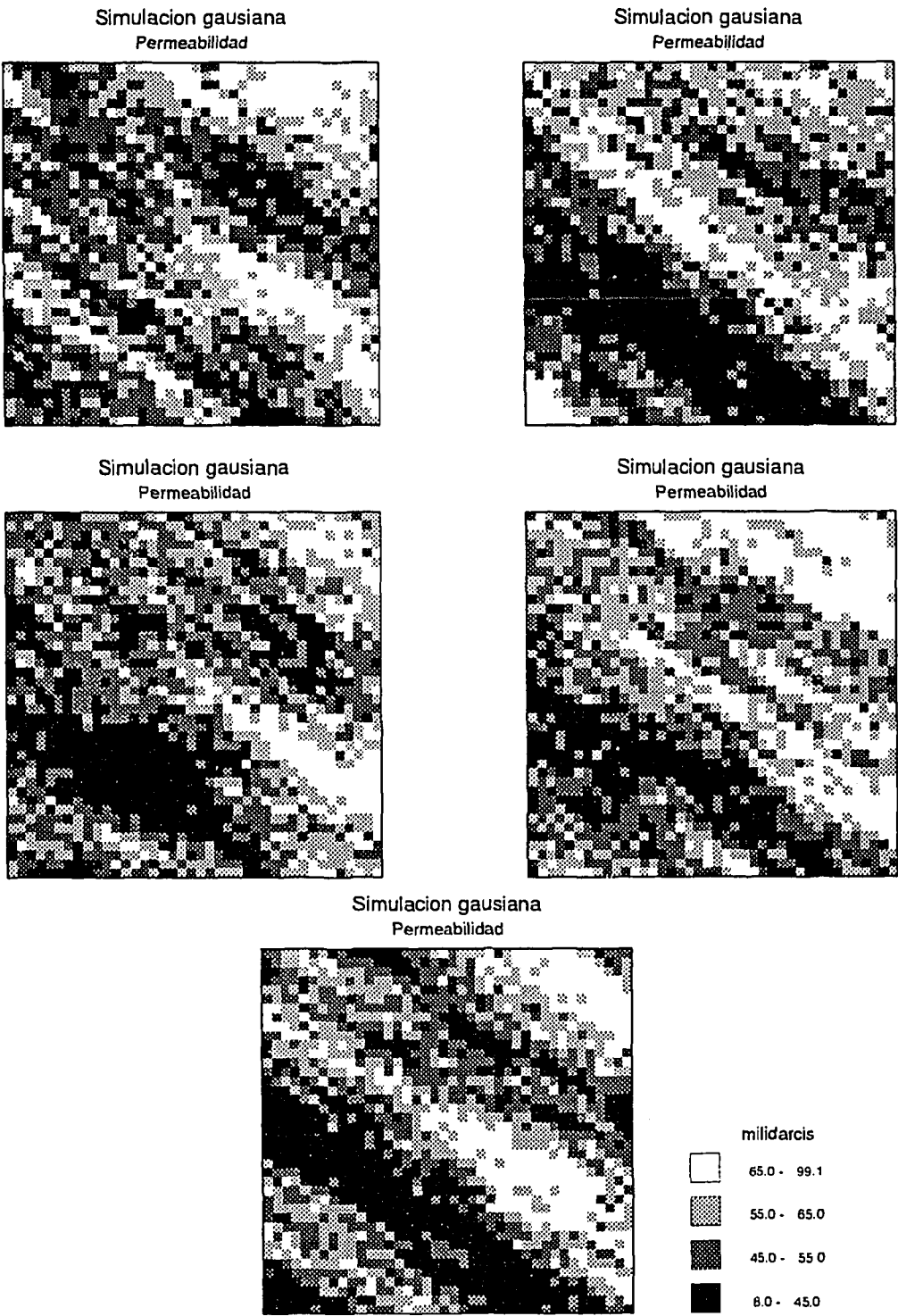


Figura 39
Simulaciones gaussianas de permeabilidad. Cinco realizaciones condicionadas a los datos duros del conjunto b16.

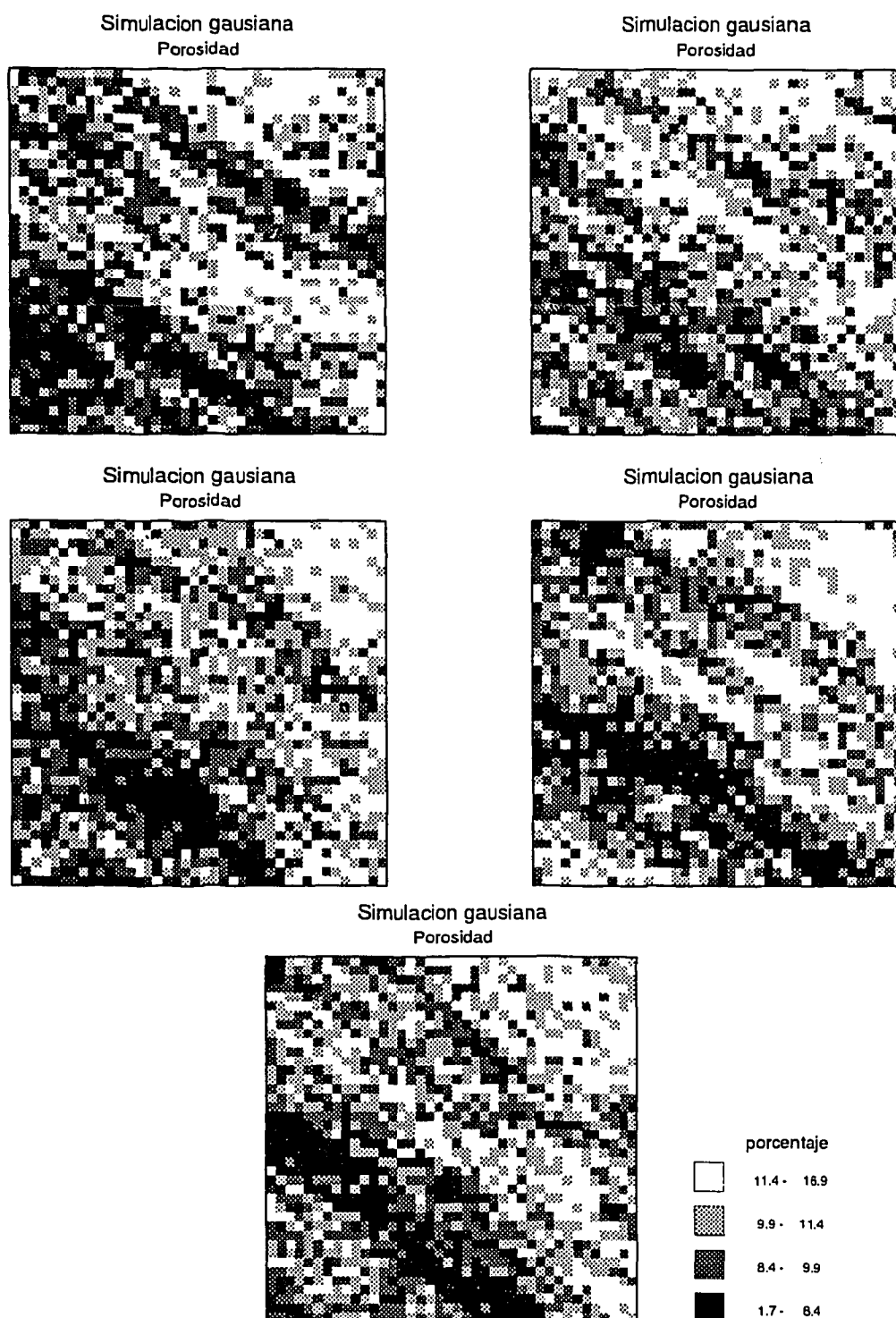


Figura 40

Simulaciones gaussianas de porosidad. Cinco realizaciones condicionadas a los datos duros del conjunto b16.

5. Ejemplo de aplicación

```
2      5      'semilla entera, numero de simulaciones
1      1      1      'delta.x delta.y delta.z
40     40     1      'gruesa_delta.x gruesa_delta.y gruesa_delta.z
1.     1.     1.     'origen.x, .y, .z
40     40     1      'n.x, n.y, n.z
1      1      1      'desde
40     40     1      'hasta
50     13     0.2     'radio de busqueda.x, .y, .z
50     13     0.2     'radio de busqueda.x, .y, .z malla gruesa

0.84   -0.54  0      'cosenos
0.54   0.84   0
0      0      1      'matriz de rotacion del entorno de busqueda

3      1000  12      'max_por_octante, contador para usar K0, max_data
0      'nivel depuracion

20000  0      'frecuencia reinicio, indicador reinicio
```

Figura 41
Fichero de geometría de entrada a ISIM3D. Fichero isim3d.geo en el disco de distribución.


```

3          'nind
19.   111. 'minimo, maximo

45.   0.26 'umbral, probabilidad acumulaa marginal

0.077     'efecto pepita
1          'num_estructuras
1          'tipo 1=esferico
0.115     'meseta
90.0      'alcance.x
7.0       'alcance.y
1.         'alcance.z
0.84      -0.54 0 'cosenos que definen la
0.54      0.84 0 '  matriz de rotacion
0         0     1 '  del elipse de anisotropia

55.   0.53 'umbral, probabilidad acumulaa marginal

0.1        'efecto pepita
1          'num_estructuras
1          'tipo 1=esferico
0.149     'meseta
50.0      'alcance.x
10.0      'alcance.y
1.         'alcance.z
0.84      -0.54 0 'cosenos que definen la
0.54      0.84 0 '  matriz de rotacion
0         0     1 '  del elipse de anisotropia

65.   0.76 'umbral, probabilidad acumulaa marginal

0.054     'efecto pepita
1          'num_estructuras
1          'tipo 1=esferico
0.129     'meseta
30.0      'alcance.x
10.0      'alcance.y
1.         'alcance.z
0.84      -0.54 0 'cosenos que definen la
0.54      0.84 0 '  matriz de rotacion
0         0     1 '  del elipse de anisotropia

```

Figura 42

Fichero de variogramas de entrada a ISIM3D. Fichero isim3d.var en el disco de distribución.

5. Ejemplo de aplicación

```

6 6 1 58.50
7 17 1 28.50
3 28 1 40.00
4 34 1 56.00
13 5 1 55.00
16 12 1 27.50
19 30 1 50.00
20 32 1 47.00
28 10 1 47.50
26 13 1 67.50
22 25 1 50.00
27 39 1 64.00
40 7 1 85.00
38 20 1 50.00
37 28 1 66.00
32 38 1 82.00

```

Figura 43

Fichero de datos duros de entrada a ISIM3D. Fichero isim3d.dat en el disco de

```

1 5 1 -99 1 1 28.000
2 5 1 -99 -99 1 56.000
3 5 1 0 0 -99 72.000
4 5 1 -99 -99 1 44.000
.....
.....
35 36 1 0 0 -99 77.000
35 37 1 0 0 -99 82.000
35 38 1 0 0 -99 89.000
35 39 1 0 0 -99 85.000
35 40 1 0 0 -99 92.000

```

Figura 44

Fichero de datos blandos de entrada a ISIM3D. Fichero isim3d.ind en el disco de distribución.

```

Simulations from ISIM3D 2.31
3
Class
No parametric
Linear model between thresholds
1 31.655 37.650
1 20.713 21.492
1 27.884 33.550
1 38.044 39.804
3 58.316 59.902
2 52.347 53.822
.....
.....
2 47.368 45.689
2 50.000 51.739
2 50.000 51.104
2 49.468 47.979
4 80.061 75.468

```

Figura 45

Fichero de salida de ISIM3D.

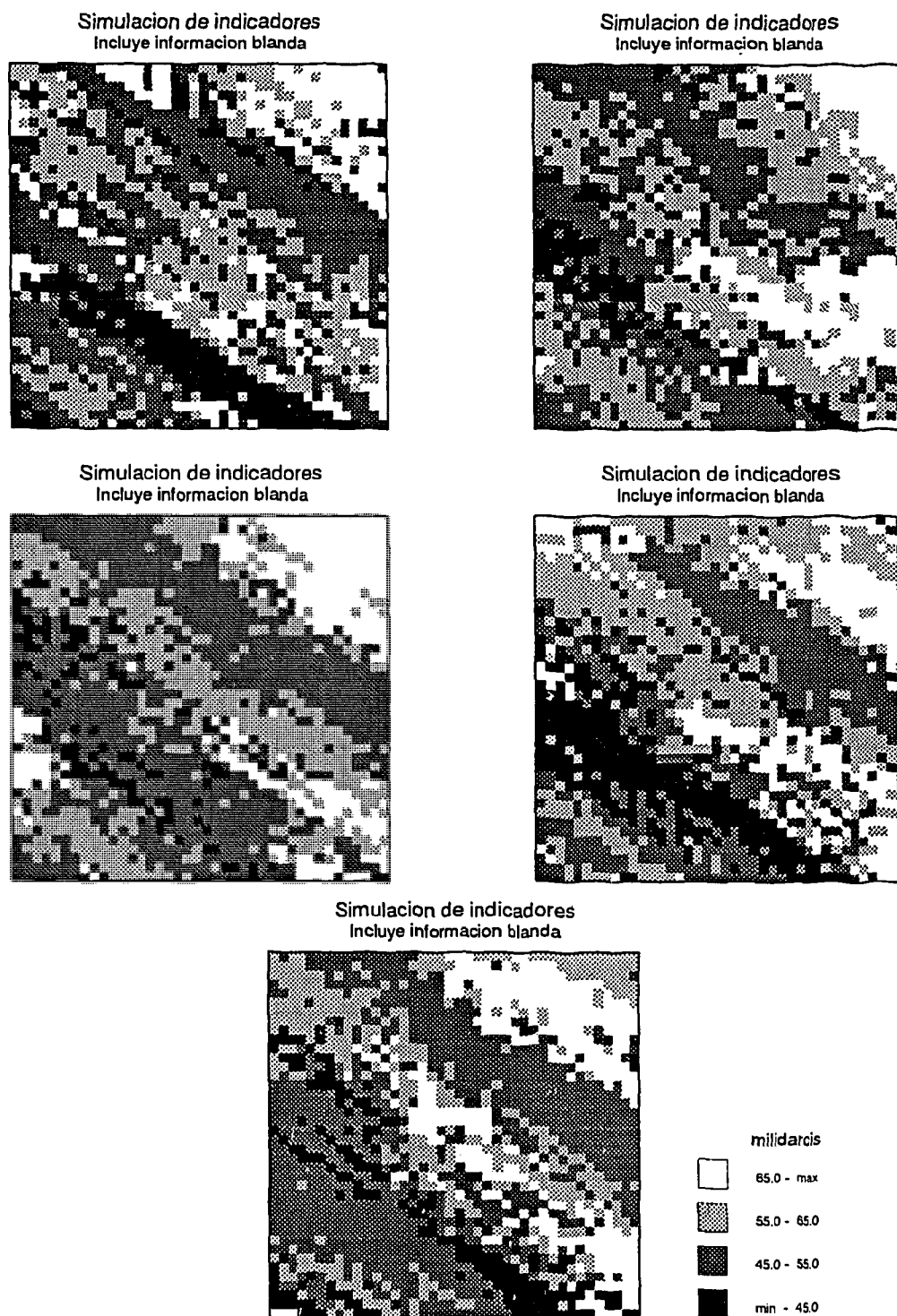


Figura 46

Simulaciones de indicadores de permeabilidad. Cinco realizaciones condicionadas sólo a los datos duros del conjunto b16.

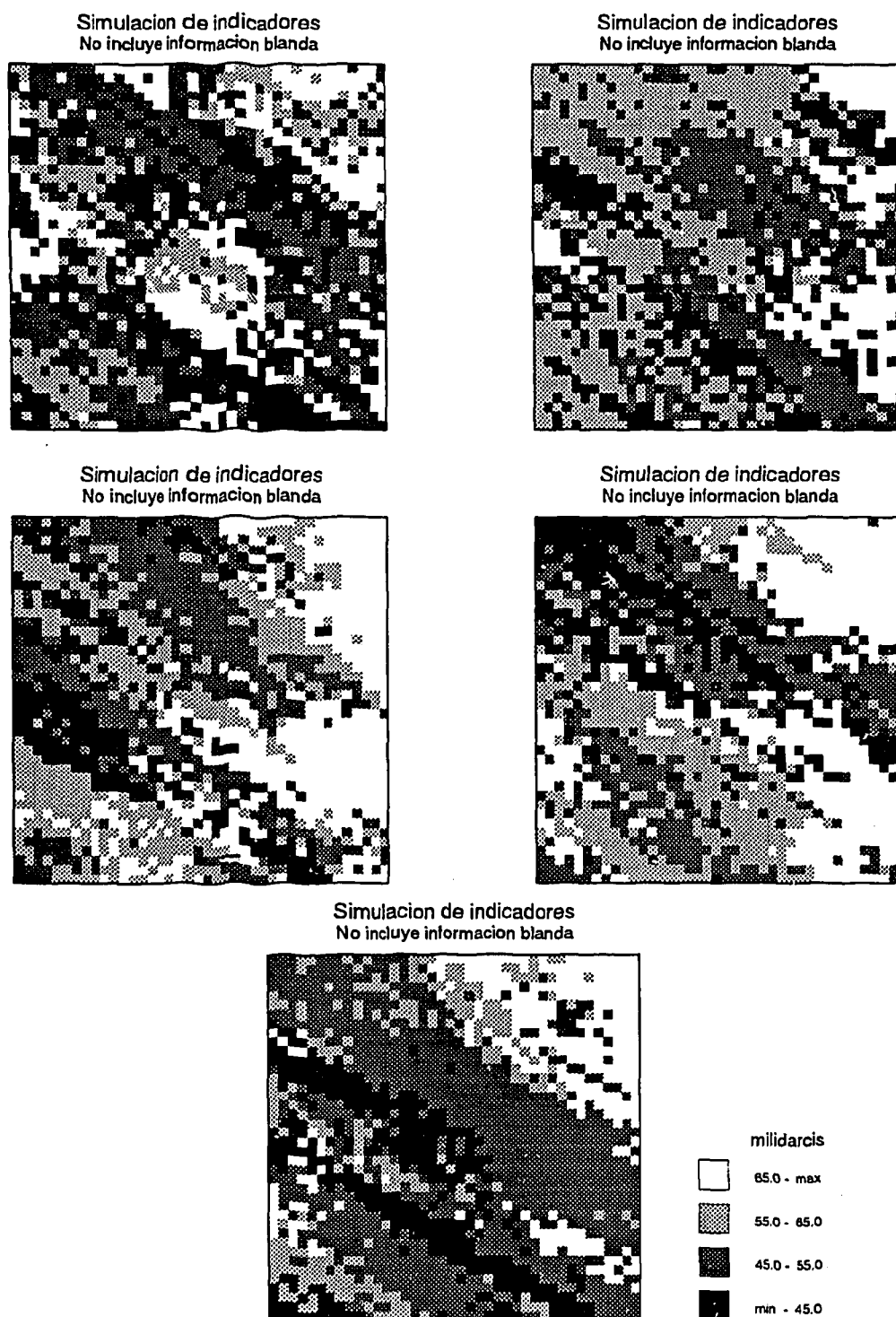


Figura 47
Simulaciones de indicadores de permeabilidad. Cinco realizaciones condicionadas a los datos duros del conjunto b16 y a los datos blandos del conjunto b500.

BIBLIOGRAFIA

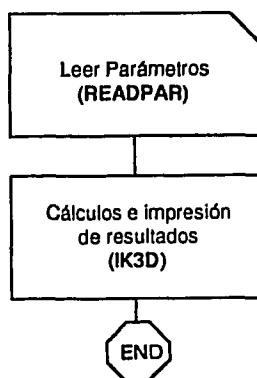
- Alabert, F.G. (1987). Stochastic imaging of spatial distributions using hard and soft data. Master's thesis, Stanford University, CA, Branner Earth Sciences Library.
- Englund, E. y Sparks, A. (1988). Geo-eas 1.2.1 user's guide. Newblock Technical Report 60018-91/008, EPA-EMSL, Las Vegas, NV 89193--3478.
- Froidevaux, R. (1990). Geostatistical toolbox primer, version 1.30. Technical report, FSS International, 10 Chemin de Drize, Troinex, Switzerland.
- Giordano, R.M., Salter, S.J., y Mohanty, K.K. (1985). The effect of permeability variations on flow in porous media. *SPE* 14365.
- Gómez-Hernández, J.J. (1991). *A Stochastic Approach to the Simulation of Block Conductivity Values Conditioned Upon Data Measured at a Smaller Scale*. PhD thesis, Stanford University.
- Gómez Hernández, J.J. (1991). Geoestadística para el análisis de riesgos: Una introducción a la geoestadística no paramétrica. Publicación Técnica 05/91 ENRESA.
- Deutsch, C.V. y Journel, A.G. (1992). *GSLIB, Geostatistical Software Library and User's Guide*. Oxford University Press, New York.
- Gómez-Hernández, J.J. y Journel, A.G. (1993). Joint simulation of multiGaussian random variables. In Soares, A., editor, *Geostatistics Tróia '92, volume 1*, páginas 85--94. Kluwer.
- Gómez-Hernández, J.J. and Srivastava, R.M. (1990). ISIM3D: An ANSI-C three dimensional multiple indicator conditional simulation program. *Computer and Geosciences*, 16(4):395--440.
- Isaaks, E.H. (1990). *The Application of Monte-Carlo Methods to the Analysis of Spatially Correlated Data*. PhD thesis, Stanford University, Branner Earth Sciences Library.
- Journel, A.G. (1983). Mad and the conditional quantile estimators. In Verly, G., David, M., Journel, A.G., and Marechal, A., editors, *Geostatistics for Natural Resources Characterization*, pages 915--934. NATO Advanced Study Institute, South Lake Tahoe, California, September 6-17, D. Reidel, Dordrecht, Holland.
- Journel, A.G. (1984). The place of non-parametric geostatistics. In Verly, G., David, M., Journel, A.G., and Marechal, A., editors, *Geostatistics for Natural Resources Characterization*. Proceedings of the NATO Advanced Study Institute, South Lake Tahoe, California, September 6-17, D. Reidel, Dordrecht, Holland.
- Journel, A.G. (1989). *Fundamentals of Geostatistics in Five Lessons*, volume 8 of *Short Courses in Geology*. AGU, Washington D.C.
- Journel, A.G. y Huijbregts, C.J. (1978). *Mining Geostatistics*. Academic Press, London.

APÉNDICES

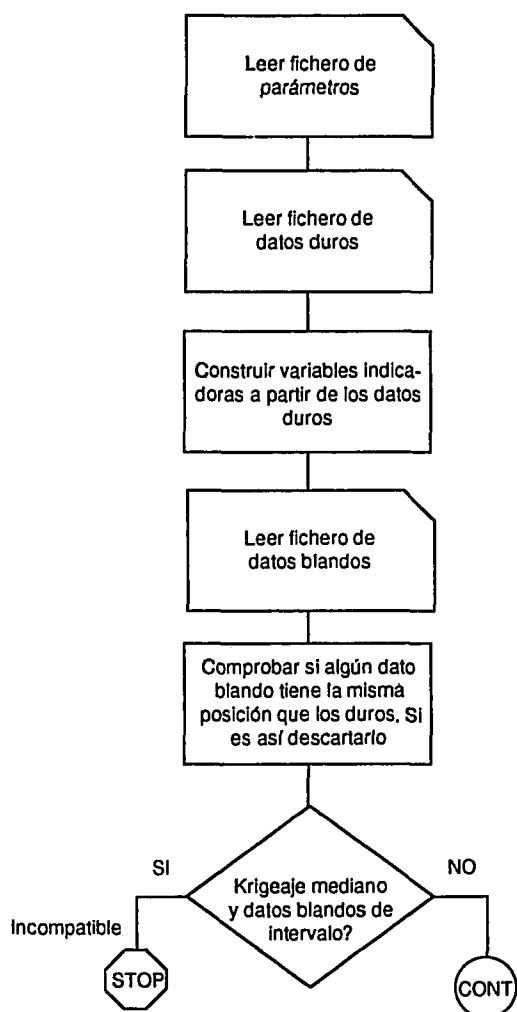
APENDICE A

DIAGRAMA DE FLUJO DEL PROGRAMA IK3D

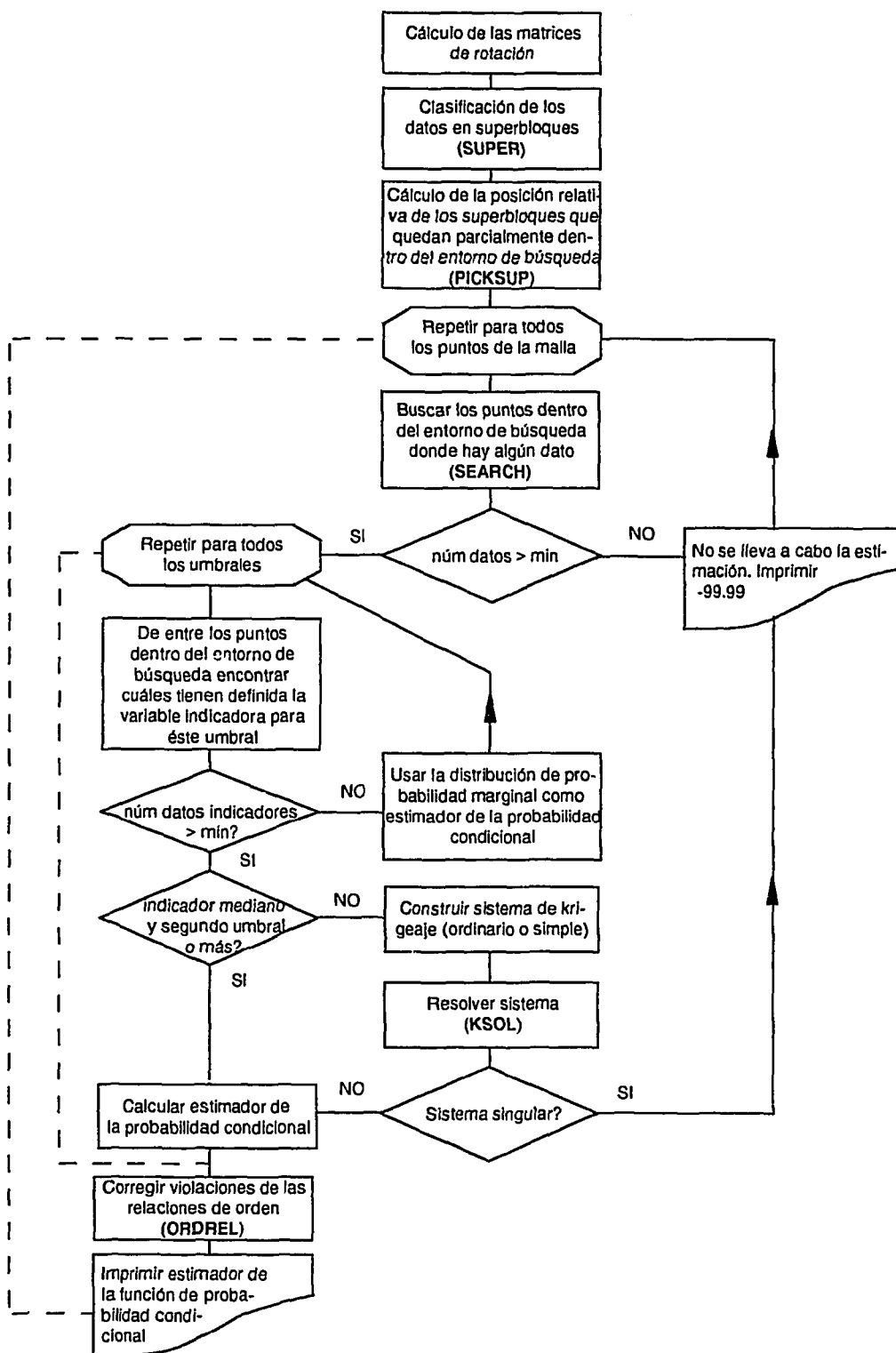
PROGRAMA PRINCIPAL



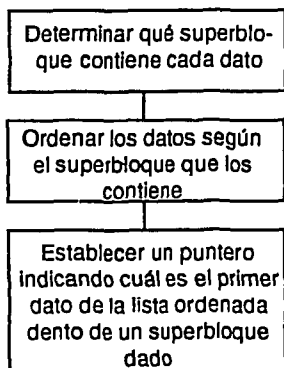
SUBROUTINA READPAR



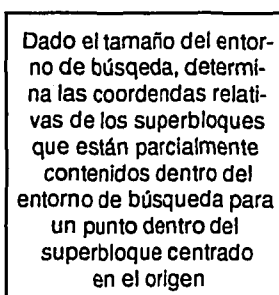
SUBROUTINA IK3D



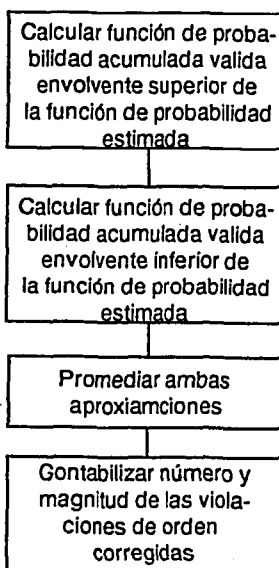
SUBROUTINA SUPER



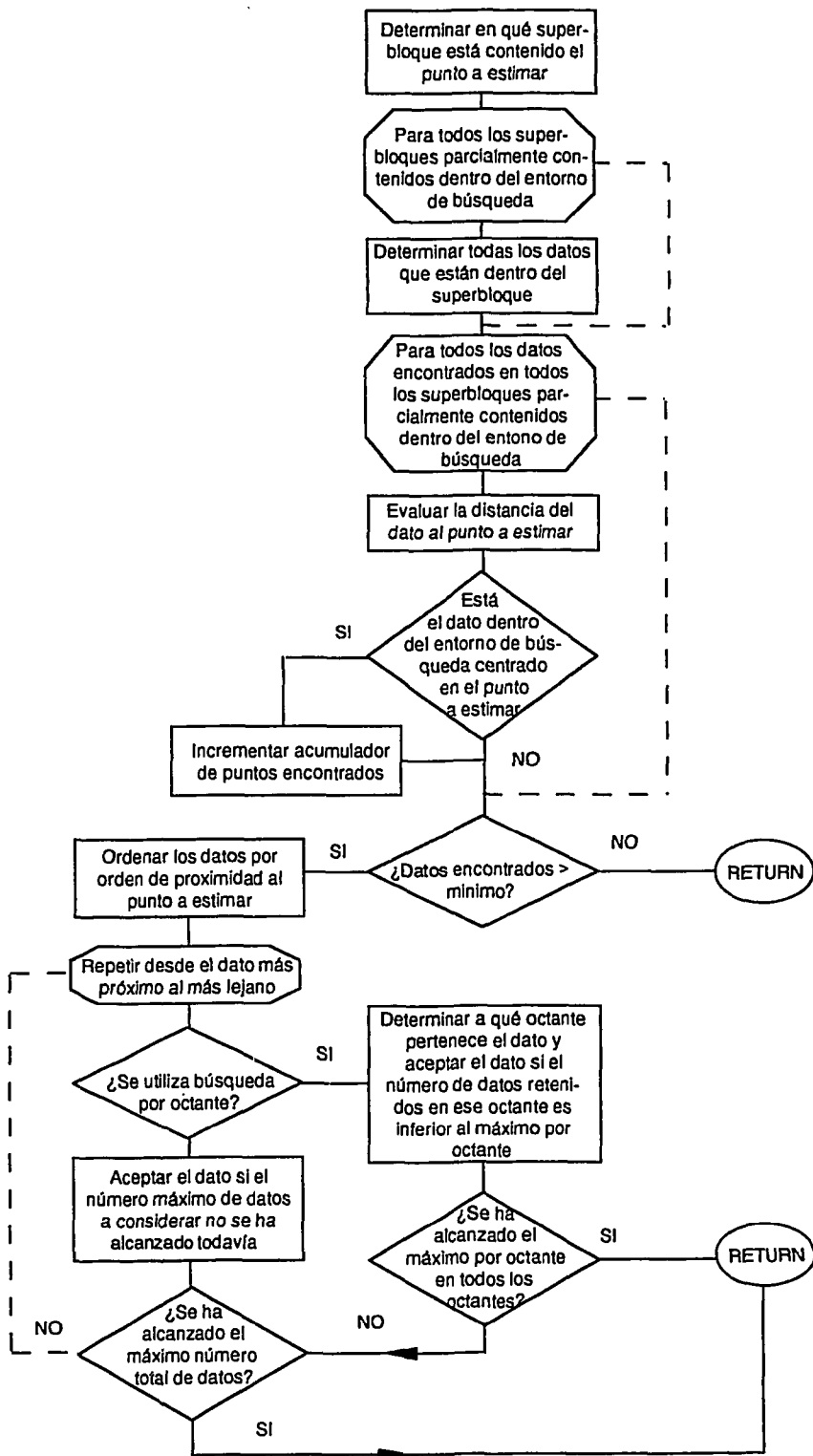
SUBROUTINA PICKSUP



ORDREL



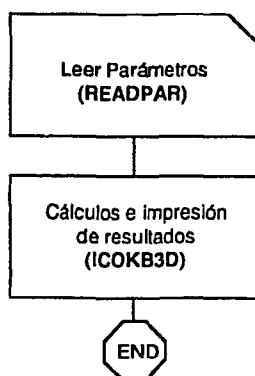
SUBROUTINA SEARCH



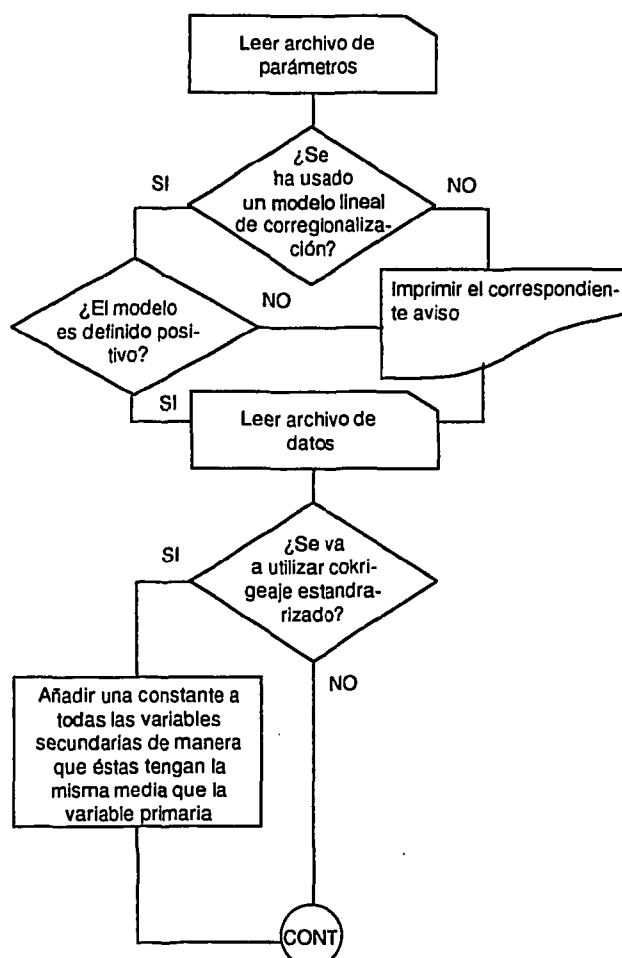
APENDICE B

DIAGRAMA DE FLUJO DEL PROGRAMA COKB3D

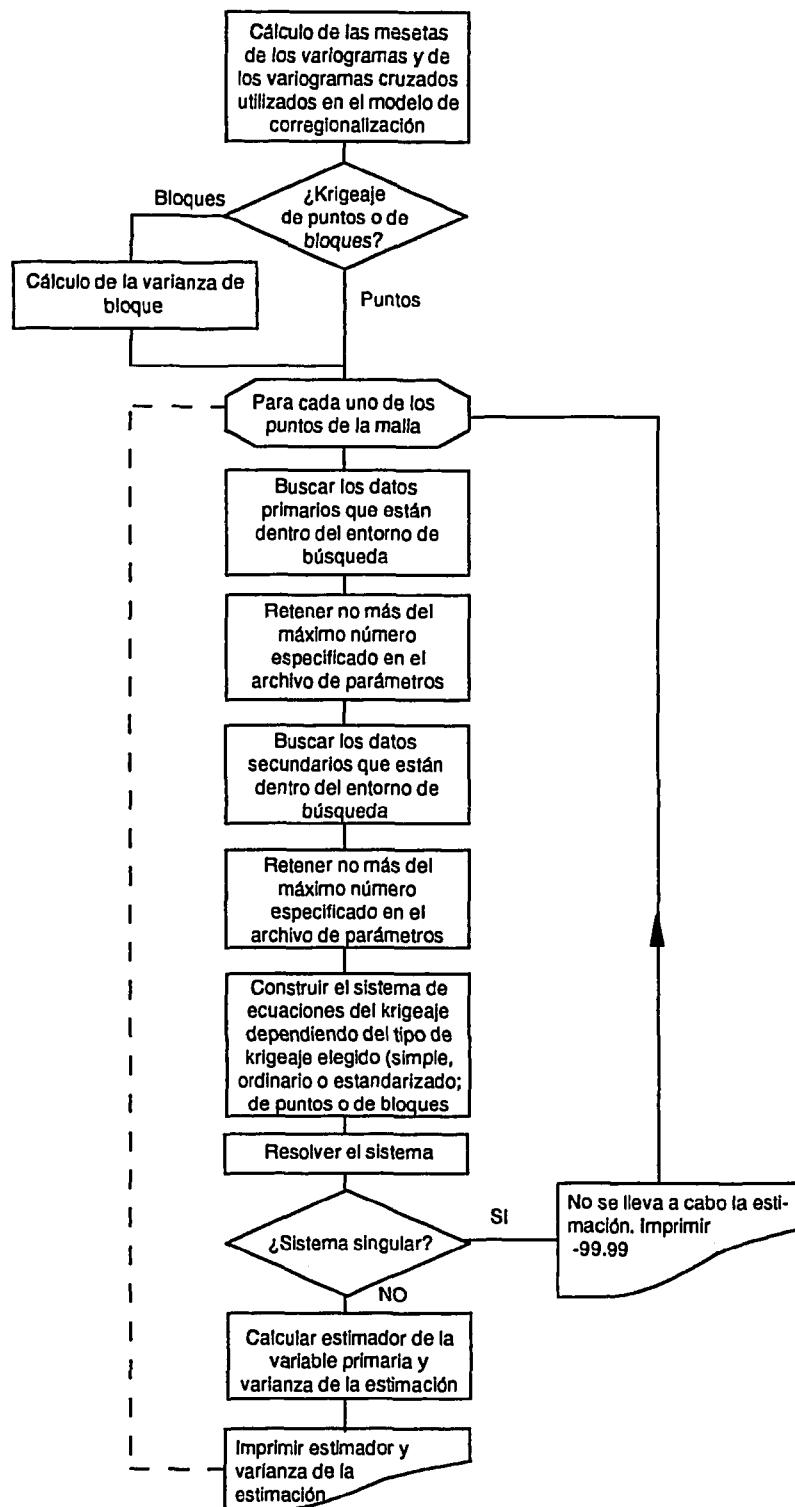
PROGRAMA PRINCIPAL



SUBROUTINA READPAR



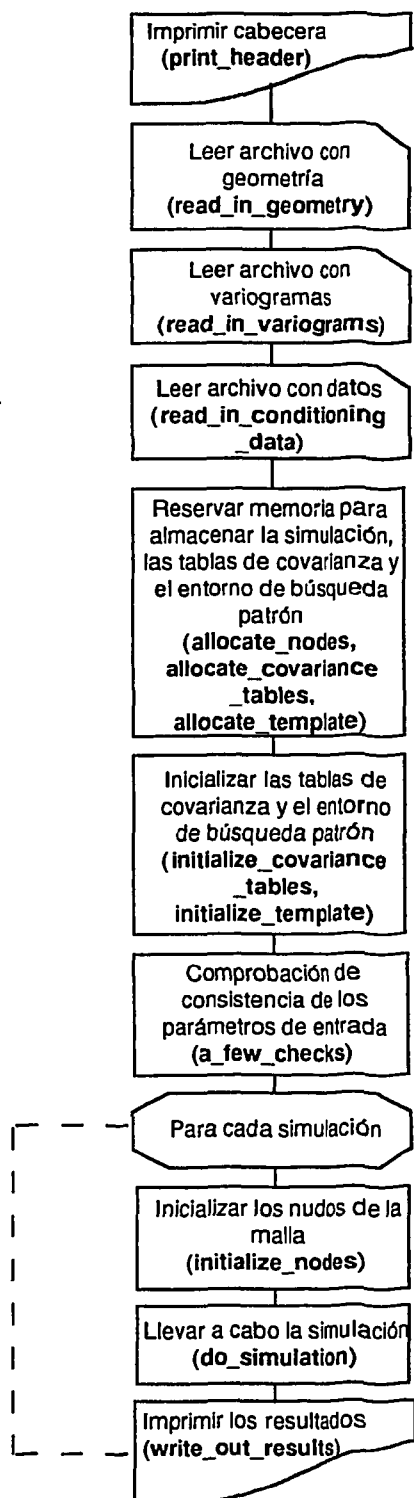
SUBROUTINA COKB3D



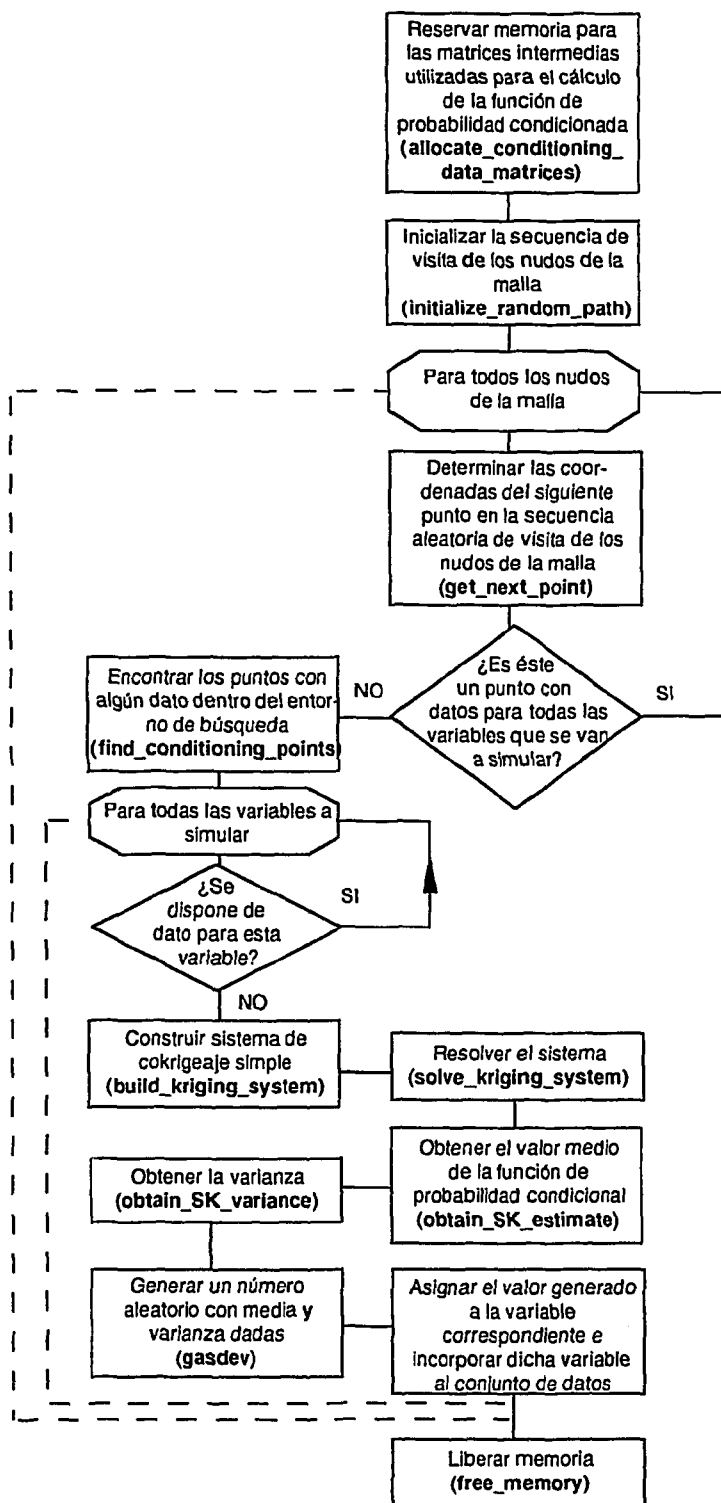
APENDICE C

DIAGRAMA DE FLUJO DEL PROGRAMA GCOSIM3D

PROGRAMA PRINCIPAL



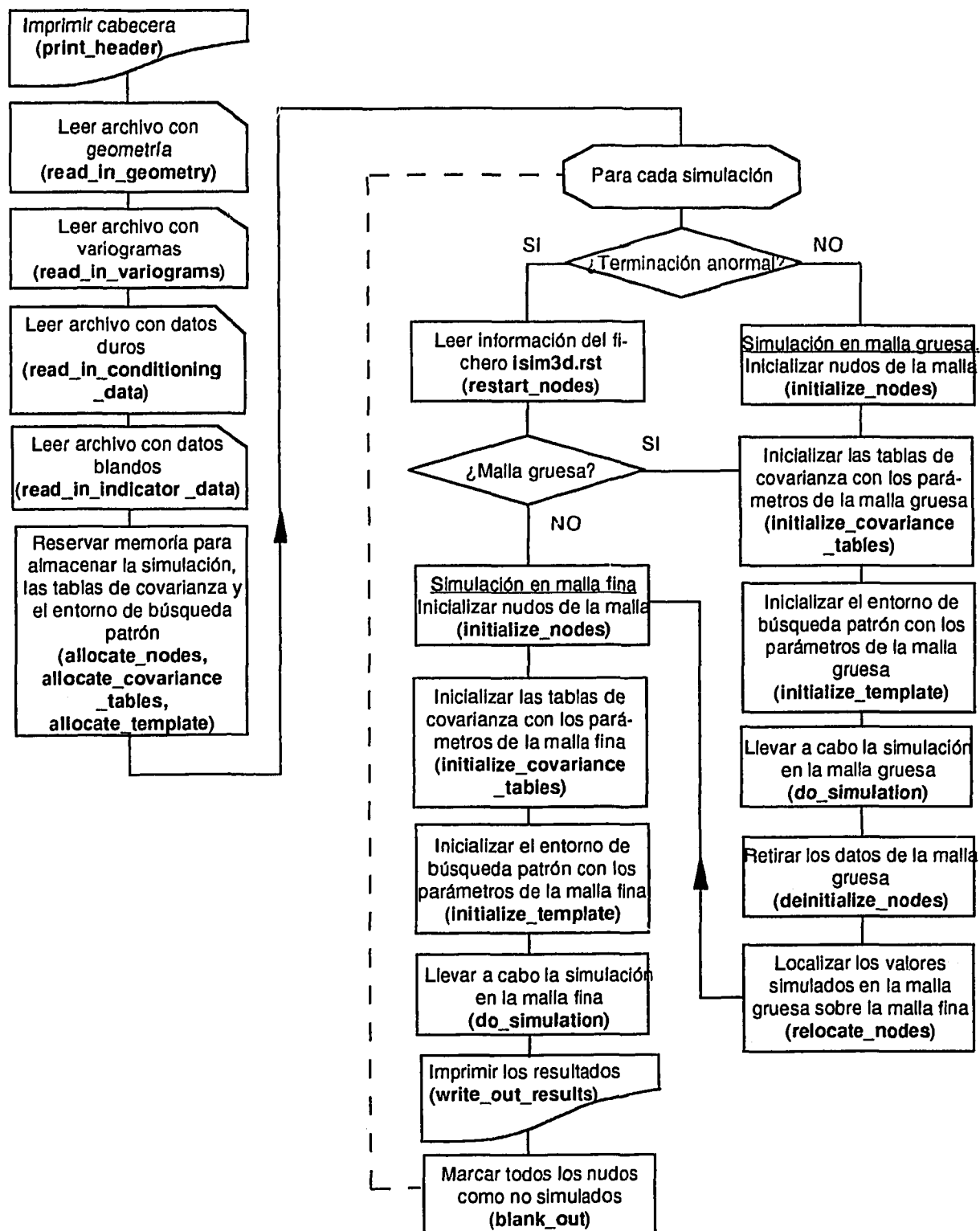
FUNCION DE SIMULACION



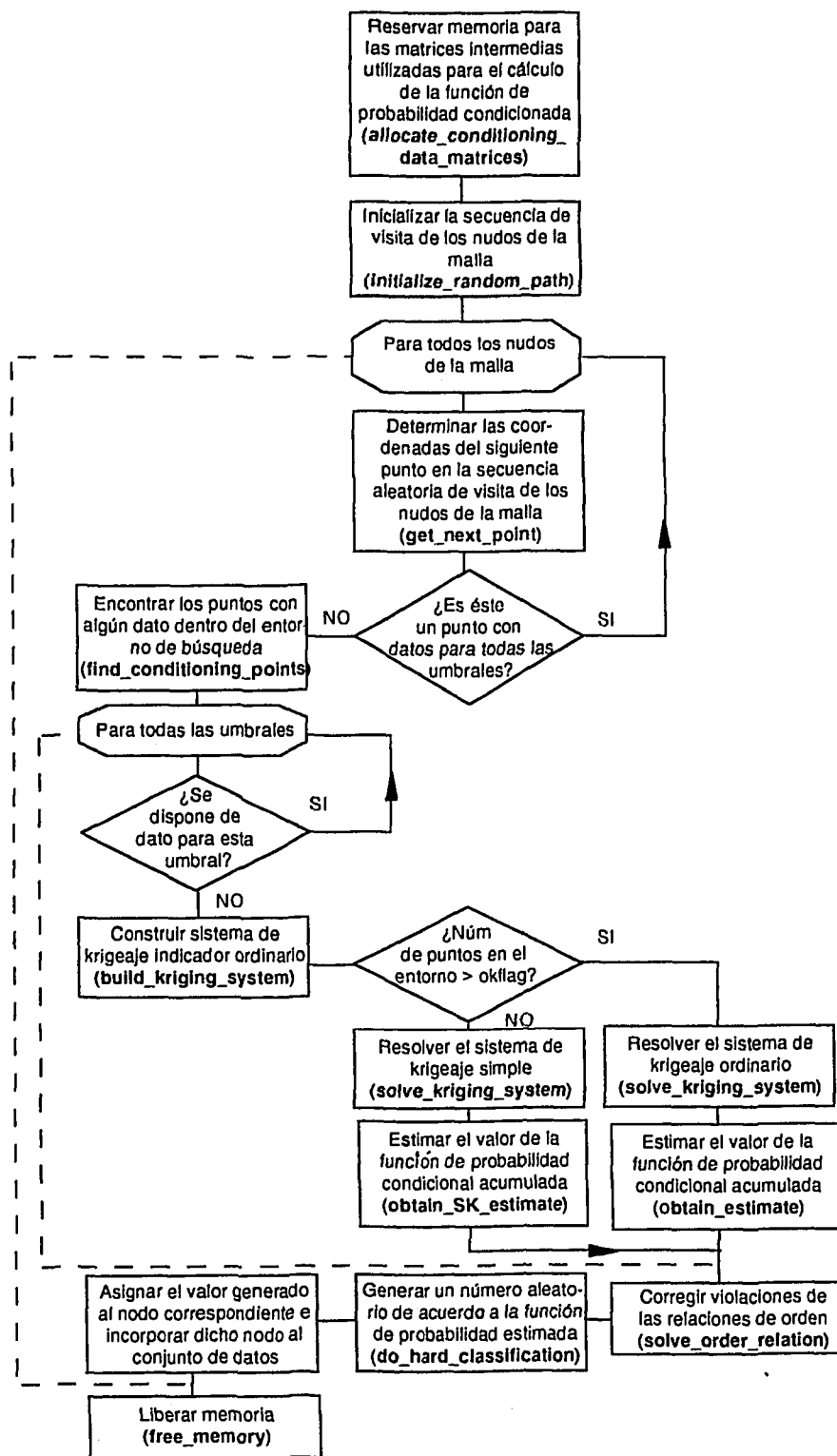
APENDICE D

DIAGRAMA DE FLUJO DEL PROGRAMA ISIM3D

PROGRAMA PRINCIPAL



FUNCION DE SIMULACION



Títulos publicados:

1991

- 01/ REVISION SOBRE LOS MODELOS NUMERICOS RELACIONADOS CON EL ALMACENAMIENTO DE RESIDUOS RADIATIVOS.
- 02/ REVISION SOBRE LOS MODELOS NUMERICOS RELACIONADOS CON EL ALMACENAMIENTO DE RESIDUOS RADIATIVOS.
ANEXO 1. Guía de códigos aplicables.
- 03/ PRELIMINARY SOLUBILITY STUDIES OF URANIUM DIOXIDE UNDER THE CONDITIONS EXPECTED IN A SALINE REPOSITORY.
- 04/ GEOESTADISTICA PARA EL ANALISIS DE RIESGOS.
Una introducción a la Geoestadística no paramétrica.
- 05/ SITUACIONES SINOPTICAS Y CAMPOS DE VIENTOS ASOCIADOS EN "EL CABRIL".
- 06/ PARAMETERS, METHODOLOGIES AND PRIORITIES OF SITE SELECTION FOR RADIOACTIVE WASTE DISPOSAL IN ROCK SALT FORMATIONS.

1992

- 01/ STATE OF THE ART REPORT: DISPOSAL OF RADIOACTIVE WASTE IN DEEP ARGILLACEOUS FORMATIONS.
- 02/ ESTUDIO DE LA INFILTRACION A TRAVES DE LA COBERTERA DE LA FUA.
- 03/ SPANISH PARTICIPATION IN THE INTERNATIONAL INTRAVAL PROJECT.
- 04/ CARACTERIZACION DE ESMECTITAS MAGNESICAS DE LA CUENCA DE MADRID COMO MATERIALES DE SELLADO.
Ensayos de alteración hidrotermal.
- 05/ SOLUBILITY STUDIES OF URANIUM DIOXIDE UNDER THE CONDITIONS EXPECTED IN A SALINE REPOSITORY.
Phase II
- 06/ REVISION DE METODOS GEOFISICOS APLICABLES AL ESTUDIO Y CARACTERIZACION DE EMPLAZAMIENTOS PARA ALMACENAMIENTO DE RESIDUOS RADIATIVOS DE ALTA ACTIVIDAD EN GRANITOS, SALES Y ARCILLAS.

- 07/ COEFICIENTES DE DISTRIBUCION ENTRE RADIONUCLEIDOS.
- 08/ CONTRIBUTION BY CTN-UPM TO THE PSACON LEVEL-S EXERCISE.
- 09/ DESARROLLO DE UN MODELO DE RESUSPENSION DE SUELOS CONTAMINADOS. APLICACION AL AREA DE PALOMARES.
- 10/ ESTUDIO DEL CODIGO FFSM PARA CAMPO LEJANO. IMPLANTACION EN VAX.
- 11/ LA EVALUACION DE LA SEGURIDAD DE LOS SISTEMAS DE ALMACENAMIENTO DE RESIDUOS RADIATIVOS. UTILIZACION DE METODOS PROBABILISTAS.
- 12/ METODOLOGIA CANADIENSE DE EVALUACION DE LA SEGURIDAD DE LOS ALMACENAMIENTOS DE RESIDUOS RADIATIVOS.
- 13/ DESCRIPCION DE LA BASE DE DATOS WALKER

1993

- 01/ INVESTIGACION DE BENTONITAS COMO MATERIALES DE SELLADO PARA ALMACENAMIENTO DE RESIDUOS RADIATIVOS DE ALTA ACTIVIDAD. ZONA DE CABO DE GATA, ALMERIA.
- 02/ TEMPERATURA DISTRIBUTION IN A HYPOTHETICAL SPENT NUCLEAR FUEL REPOSITORY IN A SALT DOME.
- 03/ ANALISIS DEL CONTENIDO EN AGUA EN FORMACIONES SALINAS. Su aplicación al almacenamiento de residuos radiactivos
- 04/ SPANISH PARTICIPATION IN THE HAW PROJECT. Laboratory Investigations on Gamma Irradiation Effects in Rock Salt.
- 05/ CARACTERIZACION Y VALIDACION INDUSTRIAL DE MATERIALES ARCILLOSOS COMO BARRERA DE INGENIERIA.
- 06/ CHEMISTRY OF URANIUM IN BRINES RELATED TO THE SPENT FUEL DISPOSAL IN A SALT REPOSITORY (I).
- 07/ SIMULACION TERMICA DEL ALMACENAMIENTO EN GALERIA-TSS.

Edita:

enresa

empresa nacional de residuos radiactivos, s. a.

Diseño y coordinación editorial: TransEdit

Fotomecánica: Sistemas Gráficos, S.A.

Imprime: Artes Gráficas JESMA, S.A.

D.L.: M-26382-1991

Por favor, péguelo

Dada la naturaleza del presente trabajo, existe a disposición de las personas interesadas un disquete con los programas explicados en el texto.

Todo aquél que desee recibirlo, puede utilizar este impreso una vez cumplimentado. Sepárelo y pliegue por la línea de puntos.

.....

Deseo recibir el disquete con los programas que se describen en la Publicación Técnica 08/93 de ENRESA.

APELLIDOS: _____

NOMBRE: _____

INSTITUCION: _____

DIRECCION: _____

C.P. _____ POBLACION: _____

PAIS _____

Envíe esta tarjeta de pedido a ENRESA

**Empresa Nacional
de Residuos Radiactivos, S. A.**

Emilio Vargas, 7
Tels. 519 52 55 - 519 52 63
Fax 519 52 68
28043 MADRID

AGOSTO 93