

# El problema inverso en hidrología subterránea. Aplicaciones

A. Medina<sup>2</sup>, G. Galarza<sup>1</sup>, J. Carrera<sup>1</sup>, J. Jódar<sup>1</sup> y A. Alcolea<sup>1</sup>

1 Departamento de Ingeniería del Terreno. E.T.S.I. de Caminos, UPC, Barcelona.  
2 Departamento de Matemática Aplicada III. E.T.S.I. de Caminos, UPC, Barcelona

## RESUMEN

La complejidad de los medios geológicos, así como el que cada vez más se requieran estudios de mayor detalle, ha provocado un gran auge en el empleo de modelos numéricos en hidrogeología, fundamentalmente de flujo de agua, pero también de forma creciente de transporte de solutos. Debido a esta complejidad y a la dificultad de disponer de medidas directas de los parámetros físicos de flujo y transporte (conductividad, difusión, porosidad, etc.), resulta necesaria una calibración. Este proceso de calibración se puede llevar a cabo de forma manual o automática. Aunque actualmente se dispone de un buen número de códigos que permiten la calibración automática, todavía se sigue empleando la calibración manual. En este artículo se presenta brevemente la teoría del problema inverso y se muestra su aplicación a dos ejemplos mostrando algunas de las posibilidades de la calibración automática.

Palabras clave: calibración, función objetivo, modelos, problema inverso

## *The inverse problem in hydrogeology. Applications*

### ABSTRACT

*The complexity of geological media, together with the need of larger detail studies has led to an important increase in the use of hydrogeological models, mainly for flow, but growing in solute transport studies. Due to the complexity and difficulty of obtaining direct measurements of flow and transport physical parameters (conductivity, diffusion, porosity, etc.) calibration is required. This calibration process can be done automatically or manually. Even though a large number of codes that allow automatic calibration is available, manual calibration is still in use. In this paper, the theory of the inverse problem and its application to two examples is shown.*

*Key words: calibration, inverse problem, modeling, objective function*

## Introducción

Un modelo es una herramienta diseñada para simular algún aspecto del comportamiento de un sistema real. Los modelos hidrogeológicos tienen como finalidad simular el flujo de agua y el transporte de solutos a través de formaciones geológicas.

Contrariamente a otros tipos de modelos, ni los procesos físico-químicos que definen el flujo y el transporte, ni la estructura de los parámetros que los controlan son siempre bien conocidos.

Para situar adecuadamente el problema inverso dentro del proceso de modelación, es conveniente distinguir a grandes rasgos las siguientes fases en el desarrollo de un modelo:

**Identificación de Procesos:** Selección de los procesos físicos (y químicos), y sus ecuaciones correspondientes, que deben incluirse en el modelo y que se

consideran relevantes para el problema simulado. En muchas ocasiones, de todos los procesos que tienen lugar, sólo unos pocos son conocidos y, entre éstos, la mayoría únicamente se pueden representar mediante simplificaciones de la realidad. En definitiva, en esta fase se decide qué procesos se van a considerar y cómo se van a representar matemáticamente (flujo de Darcy, dispersión de Fick, adsorción lineal o no lineal, difusión en la matriz rocosa, etc.).

**Identificación de la Estructura del Modelo:** En la fase anterior el sistema real se ha reducido a una o más ecuaciones, que están definidas en base a los parámetros físicos (conductividad, porosidad, etc.), de los que es preciso definir su estructura, es decir, su variación espacial y temporal. Además, es necesario definir la geometría del modelo, las condiciones de contorno, etc. Normalmente los datos disponibles sobre estos parámetros físicos son escasos, por lo

que una definición exhaustiva de su variación espacial y temporal es inabordable, debiendo generalmente expresarla en función de un pequeño número de variables, que se suelen conocer como parámetros del modelo. Es importante tener presente en el proceso de modelación que los parámetros del modelo no son más que unos valores que aproximan de una cierta forma (dependiendo de cada modelo) los correspondientes parámetros físicos.

**Estimación de parámetros (problema inverso):**

Cuantificación de los parámetros del modelo. Esta fase conlleva mayores dificultades de las que en un principio cabría esperar: los datos sobre los parámetros físicos provienen de varias fuentes con distintos grados de representatividad; algunos parámetros se pueden medir en campo, pero suelen ser difíciles de obtener y propensos a errores; en ocasiones las medidas se han efectuado en escalas y bajo condiciones distintas a las del modelo, por lo que pueden ser numérica y conceptualmente distintas entre sí, etc. Por todo esto, los parámetros del modelo se buscan de modo que se obtenga la mayor similitud posible entre los valores medidos y calculados de niveles y concentraciones. Este proceso se conoce como calibración o problema inverso.

La separación de estos tres conceptos no deja de ser un tanto arbitraria, porque en muchos casos se solapan. Sin embargo, conviene mantenerla, aunque sólo sea a efectos de organizar ideas. Las dos primeras fases se suelen incluir bajo el término de conceptualización y la última en el de estimación de parámetros o problema inverso.

Debido a las dificultades ya mencionadas, en el ámbito de la hidrogeología se hace necesario la calibración de los parámetros. Ello ha provocado que se haya dedicado un gran esfuerzo de investigación relacionado con el problema inverso. Esta larga y fructífera investigación ha dado lugar también a distintos enfoques en el planteamiento del problema inverso. Los diferentes métodos se diferencian en el modo como se expresan los parámetros (en algunos casos se tratan de forma determinista y en otros son funciones aleatorias) y los algoritmos empleados en la calibración (ver, por ejemplo, los estados del arte de Yeh, 1986, Carrera, 1987, McLaughlin y Townley, 1996).

Aunque con enfoques distintos, el objetivo final de todos estos métodos consiste en buscar aquellos valores de los parámetros que proporcionan una buena similitud entre los valores medidos y calculados de niveles y/o concentraciones (u otras variables de estado, como temperatura, contenido de humedad, etc.) Para llevar a cabo este propósito, se van modificando los valores de los parámetros hasta con-

seguir dicha similitud. Esta modificación se puede llevar a cabo mediante el empleo de cualquiera de los algoritmos que se describen en las referencias anteriores (hablamos de calibración automática), o bien de forma manual, es decir, con los criterios subjetivos de cada uno se van realizando distintas simulaciones cambiando los valores de los parámetros que se considere oportuno. Con la primera metodología un programa trata de minimizar algún tipo de distancia entre los valores medidos y los calculados y ha tenido un gran auge en los últimos 25 años, fundamentalmente para la ecuación de flujo y en algo menor medida para la de transporte. El segundo método, que fue el primer método empleado, es lento y laborioso, estando muy sujeto a la experiencia y subjetividad del modelista.

*Organización y objetivos*

Se presenta una de las formulaciones de la calibración automática. Para ello se muestran primero las ecuaciones de flujo y transporte y las ecuaciones del problema inverso para dicha formulación. Se comentan los principales puntos de interés, así como algunos puntos de la calibración automática que se suelen pasar por alto o que tienen una cierta dificultad. Finalmente se muestra la aplicación de la teoría a dos ejemplos y se resumen los puntos más destacados en las conclusiones.

**Ecuaciones**

Las ecuaciones con las que se muestra la formulación son las de flujo de agua subterránea y de transporte de solutos. Por simplificar la exposición, las ecuaciones se refieren a medio bidimensional.

El flujo en acuíferos confinados está gobernado por (Bear, 1972, Custodio y Llamas, 1983):

$$\nabla(T \nabla h) + q = S \frac{\partial h}{\partial t} \text{ en } \Omega \tag{1}$$

donde  $h$  es el nivel,  $T$  es el tensor de transmisividad,  $S$  es el coeficiente de almacenamiento,  $\Omega$  es el dominio del problema,  $q$  es el término fuente/sumidero distribuido superficialmente (recarga/extracción) y  $t$  es el tiempo.

La ecuación (1) se resuelve con las condiciones iniciales:

$$h(x, y, 0) = h_0(x, y) \tag{2}$$

donde  $h_0$  puede ser cualquier función arbitraria, o la solución de la ecuación (1) en una simulación anterior, frecuentemente el régimen estacionario. También es común, sobre todo al analizar ensayos de bombeo, que  $h_0$  se tome igual a cero en todo el acuífero, en cuyo caso  $h$  representa las variaciones de nivel o descensos.

Las condiciones de contorno pueden ser de Dirichlet:

$$h = H \quad \text{en} \quad \Gamma_1$$

o mixtas:

$$(T \nabla h)n = \alpha(H - h) + Q \quad \text{en} \quad \Gamma_2 \quad (3)$$

donde  $\Gamma = \Gamma_1 \cup \Gamma_2$  es el contorno de  $\Omega$ ,  $n$  es el vector unitario normal a  $\Gamma$  y dirigido hacia el exterior,  $H$  es un nivel externo,  $Q$  es un caudal prefijado y  $\alpha$  es un coeficiente que controla el tipo de condición de contorno ( $\alpha = 0$  para caudal fijo, y  $\alpha \neq 0$  para condición mixta, en cuyo caso  $\alpha$  es un coeficiente de goteo).

La ecuación (1) y las condiciones de contorno, quedan expresadas en función de los parámetros físicos (transmisividad, recarga, coeficiente de almacenamiento, etc.). Para resolver esta ecuación, es necesario describir adecuadamente la variación espacial y temporal de los mismos (identificación de la estructura del modelo).

El transporte de solutos en un acuífero incluyendo los procesos de difusión, dispersión, convección, reacciones químicas de primer orden y adsorción está controlado por la ecuación (Bear, 1972, Custodio y Llamas, 1983):

$$b\phi R \frac{\partial c}{\partial t} = \nabla(D \nabla c) + q(c' - c) - \lambda b\phi R c \quad (4)$$

siendo  $\phi$  la porosidad del acuífero,  $b$  su espesor,  $c$  la concentración del soluto,  $c'$  es la concentración externa en fuentes distribuidas superficialmente,  $t$  el tiempo,  $R$  el coeficiente de retardo debido a fenómenos de adsorción,  $q$  es el flujo de Darcy,  $\lambda$  es el coeficiente de desintegración radioactiva,  $D$  es el tensor de dispersión:

$$D = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{pmatrix}$$

(5)

$$\begin{aligned} D_{xx} &= \phi D_d + \frac{\alpha_L q_x^2 + \alpha_T (q_y^2 + q_z^2)}{|q|} \\ D_{yy} &= \phi D_d + \frac{\alpha_T (q_x^2 + q_z^2) + \alpha_L q_y^2}{|q|} \end{aligned} \quad (5.1)$$

$$\begin{aligned} D_{zz} &= \phi D_d + \frac{\alpha_L q_z^2 + \alpha_T (q_x^2 + q_y^2)}{|q|} \\ D_{xy} &= (\alpha_L - \alpha_T) \frac{q_x q_y}{|q|} \end{aligned} \quad (5.2)$$

$$\begin{aligned} D_{xz} &= (\alpha_L - \alpha_T) \frac{q_x q_z}{|q|} \\ D_{yz} &= (\alpha_L - \alpha_T) \frac{q_y q_z}{|q|} \end{aligned} \quad (5.3)$$

siendo  $D_d$  el coeficiente de difusión molecular,  $\alpha_L$  la dispersividad longitudinal,  $\alpha_T$  la dispersividad transversal,  $q_x$ ,  $q_y$  y  $q_z$  las componentes del vector  $q$  en las direcciones de los ejes  $x$ ,  $y$ ,  $z$  respectivamente y  $|q|$  su módulo.

La ecuación (3) se resuelve con las condiciones de contorno e iniciales siguientes:

Condiciones de contorno de Dirichlet:

$$c = c_e \quad \text{en} \quad \Gamma_1 \quad (6)$$

y mixtas:

$$D \nabla c n = \beta(c_e - c) + M \quad \text{en} \quad \Gamma_2 \quad (7)$$

donde  $\Gamma = \Gamma_1 \cup \Gamma_2$  es el contorno de  $\Omega$  y siendo  $\beta = 0$  y  $M = 0$  condición de contorno de salida (flujo másico nulo),  $\beta = -qn$  condición de entrada con concentración  $c_e$  y  $\beta = 0$  y  $M \neq 0$  condición de contorno de entrada de masa.

Condiciones iniciales:

$$c(x, y, 0) = c_0(x, y) \quad \forall x, y \in \Omega \quad (8)$$

donde  $c_0(x; y)$  es una función arbitraria dada o bien la solución en régimen estacionario de la ecuación (4).

### Parametrización

Para resolver las ecuaciones de flujo y transporte de solutos (1) y (4) con sus respectivas condiciones iniciales y de contorno, es preciso conocer los valores de los parámetros de las mismas (transmisividad, dispersividad, porosidad, etc.). Además, salvo casos muy particulares y simples en los que se dispone de solución analítica, para resolver dichas ecuaciones es necesario recurrir a métodos numéricos. Para ello es preciso discretizar las ecuaciones y con ello, los valores de los parámetros. El proceso por el que se realiza esta discretización se denomina parametrización. Debido a que la información disponible sobre los parámetros físicos acostumbra a ser escasa, la variabilidad espacial se suele expresar en función de unos pocos parámetros, los parámetros del modelo. Muchas de las parametrizaciones que hay en la bibliografía se pueden representar como

$$p(x) = \sum_{j=1}^n p_j \psi_j(x) \quad (9)$$

donde  $p_j$  son los parámetros del modelo y  $\psi_j(x)$  son unas funciones de interpolación. Dependiendo de las funciones de interpolación empleadas se obtienen distintas parametrizaciones, (para ver una presentación exhaustiva de los distintos tipos de parametrizaciones consultar por ejemplo, Macaughlin y Townley, 1996). En este trabajo, se presenta la zonificación por ser una de las más sencillas y empleadas. En esta aproximación se divide el dominio en subconjuntos disjuntos denominados zonas y  $\psi_j(x) = 0$  fuera de la zona  $j$  y  $\psi_j(x) = 1$  en la zona  $j$ . Es decir, los parámetros físicos se aproximan como constantes en cada zona.

Con esta aproximación únicamente se tiene la variación espacial. Algunos parámetros presentan también variación temporal. En este caso, lo más habitual suele ser considerar de forma independiente la variación temporal y la espacial, es decir,

$$p(x,t) = p(x) f_p(t) \quad (10)$$

siendo  $f_p(t)$  una función que contiene la variación temporal del parámetro físico  $p$ . Juntando las ecuaciones (9) y (10) resulta la aproximación del parámetro físico  $p$ :

$$p(x,t) = \left[ \sum_{j=1}^n p_j \psi_j(x) \right] f_p(t) \quad (11)$$

Aquí se ha presentado un solo tipo de parametrización, no obstante, independientemente del tipo

empleado, deben recordarse los siguientes puntos:

El proceso de calibración modifica los valores  $p_j$ , en general, éste no es el valor de la aproximación del parámetro físico correspondiente, sino que ésta hay que calcularla mediante (11).

Si se desea calibrar el modelo, conviene que el número de parámetros no sea muy grande, sobre todo si se realiza de forma manual. Además, debe tenerse muy en cuenta que, en general, cuantos más parámetros se empleen, mejor ajuste (similitud entre valores calculados y medidos de niveles y/o concentraciones) cabe esperar, pero esto suele ser a costa de introducir parámetros que pueden carecer de sentido físico. Hay que tener cuidado, pues con la sobreparametrización.

El número de parámetros que se pueden calibrar es función de la cantidad de información que dispongamos.

### Problema inverso (calibración)

Como se ha mencionado en la introducción, la calibración (también conocida en la literatura como problema inverso) se puede llevar a cabo de forma manual o automática. La primera, que ha sido casi la única utilizada durante mucho tiempo, consiste en modificar sucesivamente los valores de los parámetros en base a los resultados de la simulación, utilizando criterios obtenidos con la experiencia de quien está realizando el modelo, hasta lograr un ajuste satisfactorio. Con la segunda, además de obtener (igual que la manual) unos valores estimados de los parámetros, permite también obtener informaciones adicionales que con la calibración manual no son posibles de obtener o resulta muy complicado (matriz de covarianza de los parámetros, intervalos de confianza, etc.).

Independientemente del tipo de calibración que se vaya a efectuar, hay que definir con precisión criterios objetivos para su realización (sin duda, en la manual no es tan necesario, puesto que es posible introducir con relativa facilidad criterios subjetivos basados en la experiencia personal).

El proceso de calibración tiene como objetivo obtener unos parámetros con los que se reproduzca de forma adecuada la evolución del sistema, evolución que viene en este caso definida por los valores de niveles y concentraciones. Se trata, pues, de hallar los parámetros que den como resultado unos valores de niveles y de concentraciones lo más parecidos a la realidad observada. Para ello es necesario expresar matemáticamente el problema anterior, es decir, es preciso definir algún tipo de distancia entre los valo-

res calculados y los medidos que controle la semejanza entre unos y otros. De inmediato surge un problema: No es posible conocer realmente la diferencia entre las soluciones (exactas) de las ecuaciones  $h_a, c_a$  y los valores (reales) de niveles y concentraciones, porque ninguno de éstos es conocido, el primero porque la resolución de las ecuaciones suele ser numérica (lo que conlleva de forma implícita errores de discretización, de redondeo, del esquema numérico, etc.), y el segundo, porque todo proceso de medida conlleva un cierto error.

Si se denota por  $z$  el vector formado por los niveles y concentraciones calculados,  $z_a$  el vector formado por las soluciones exactas de las ecuaciones (1) y (4),  $\hat{z}$  el vector formado por los niveles y concentraciones reales y por  $z^*$  el vector formado por los niveles y concentraciones medidos, se pueden distinguir los tres tipos de errores siguientes:

1.-  $z - z_a$  (error numérico, diferencia entre el valor calculado numéricamente y la solución analítica)

2.-  $z_a - \hat{z}$  (error debido a la imprecisión en los valores de los parámetros y la estructura del modelo, diferencia entre la solución analítica y los valores reales de niveles y concentraciones)

3.-  $\hat{z} - z^*$  (error de medida, diferencia entre los valores reales de niveles y concentraciones y los valores medidos)

Como se observa fácilmente, no es posible conocer el valor de ninguno de estos errores por separado, aunque sí la suma de todos,  $z - z^*$  (diferencia entre el valor calculado y el medido), motivo por el cual, en lugar de buscar los valores de los parámetros con los que se consigue una mejor aproximación entre los valores exactos de la ecuación y los "reales" (que será lo mejor), hay que conformarse con buscar aquellos valores de los parámetros que conlleven a una mejor aproximación entre los valores calculados y los medidos. El planteamiento resulta, pues, muy intuitivo: se buscan los parámetros del modelo que dan lugar a una distancia mínima entre  $z$  y  $z^*$ . Aunque hay varias formas de plantear esta distancia, en este artículo hablaremos exclusivamente del método de máxima verosimilitud.

Se denomina verosimilitud (y se denota por  $L(H=R)$ ) de la hipótesis  $H$  dadas las observaciones  $R$  a  $kf(R=H)$ , siendo  $k$  una constante arbitraria y  $f(R=H)$  la función de densidad de probabilidad de obtener los resultados  $R$  dada la hipótesis  $H$ , de acuerdo a un cierto modelo estadístico. En nuestro caso,  $R$  es el conjunto de observaciones (los valores medidos) y la hipótesis  $H$  está formada por los valores de los parámetros y por las ecuaciones de flujo y transporte. El método de máxima verosimilitud consiste en obtener los parámetros que maximizan el valor de  $L(H=R)$ . El

fundamento de este método radica en el axioma de verosimilitud (también denominado principio) (Edwards, 1972):

"Dentro del marco de trabajo de un modelo estadístico, toda la información que los datos proporcionan respecto al mérito de dos hipótesis, está contenida en la razón de verosimilitud de ambas respecto de los datos y la razón de verosimilitud se interpreta como el grado respecto del cual los datos soportan mejor una hipótesis que la otra."

### La función objetivo

Bajo las hipótesis de distribución normal de los residuos y de la información previa, así como de independencia entre los errores de medida de niveles, concentraciones e información previa, la función de verosimilitud viene dada por la expresión:

$$L = (2\pi)^{-\frac{n}{2}} \left( |C_h| \cdot |C_c| \cdot |C_p| \right)^{-\frac{1}{2}} \exp \left[ -\frac{1}{2} \left( (h - h^*)' C_h^{-1} (h - h^*) + (c - c^*)' C_c^{-1} (c - c^*) + (p - p^*)' C_p^{-1} (p - p^*) \right) \right] \quad (12)$$

donde  $h$  y  $c$  son vectores con los valores calculados de niveles y concentraciones respectivamente,  $h^*$  y  $c^*$  son vectores con los valores medidos,  $p$  es un vector formado por los parámetros calculados,  $p^*$  es la información previa de los mismos,  $C_h$ ,  $C_c$  y  $C_p$  son respectivamente las matrices de covarianza de los errores de niveles, concentraciones y parámetros respectivamente (si la información previa de los distintos tipos de parámetros no están correlacionadas entre sí,  $C_p$  tiene una estructura diagonal por bloques -un bloque para cada uno de los tipos de parámetros-) y  $n$  es el número total de datos (niveles medidos + concentraciones medidas + información previa de parámetros).

Puesto que generalmente es más sencillo establecer las correlaciones de forma relativa que de forma absoluta, siguiendo a Neuman y Yakowitz (1979), suponemos que las matrices de covarianza se pueden expresar como:

$$C_h = \tau_h V_h, \quad C_c = \tau_c V_c, \quad C_i = \tau_i V_i \quad (13)$$

siendo  $\tau_h$ ,  $\tau_c$  y  $\tau_i$  ( $i=1, \dots$ , número de tipos de parámetros) escalares desconocidos y  $V_h$ ,  $V_c$  y  $V_i$  matrices simétricas definidas positivas con la estructura de correlación de las observaciones. En la ecuación (13) se está suponiendo que los errores en la información

previa de los distintos tipos de parámetros no están correlacionados, es decir, la matriz de covarianza de la información previa tiene una estructura diagonal por bloques, que son las matrices  $V_i$  ( $i=1, \dots$ , número de tipos de parámetros).

El método de máxima verosimilitud consiste pues, en maximizar la expresión definida en (12). Por simplicidad, en lugar de maximizar esta expresión, se minimiza  $S = -2\ln(L)$ . Si además se suprimen los términos que contienen parámetros estadísticos, queda:

$$F = (h - h^*)^t V_h^{-1} (h - h^*) + \lambda_c (c - c^*)^t V_c^{-1} (c - c^*) + \sum_i \lambda_i (p_i - p_i^*)^t V_i^{-1} (p_i - p_i^*) \quad (14)$$

dónde

$$\lambda_c = \frac{\tau_h}{\tau_c}, \quad \lambda_i = \frac{\tau_h}{\tau_i}$$

La ecuación (14) contiene tres sumandos, cada uno de ellos relacionado con un tipo de información independiente. El primer término hace referencia al ajuste entre niveles medidos y calculados, cuanto más parecidos sean éstos, menor será el primer sumando de (14). Lo mismo sucede con los otros términos de concentración y de información previa. Los parámetros representan la ponderación relativa entre cada uno de estos tipos de información (por ejemplo, si  $\lambda_c = 0$  y  $\lambda_i = 0$ , la única información que se tendría en cuenta en la calibración sería la de niveles). Si la información estadística contenida en las matrices  $V$  es correcta, entonces los pesos  $\lambda$  deberían ser unitarios. Aunque en esta exposición se han incluido únicamente datos de nivel y de concentraciones, es posible incluir todos los tipos de datos que se deseen en la formulación (caudales, contenido de humedad, presiones, y en general, cualquier variable que se pueda obtener en función de los niveles y/o las concentraciones).

Resumiendo, el problema inverso (hallar los valores de los parámetros del modelo) se ha reducido a la minimización de la función (14). Es decir, buscamos los parámetros que den lugar al menor valor de (14). Por este motivo, a esta función se suele denominar función objetivo. Llegados a este punto es conveniente hacer algunas reflexiones.

- Los algoritmos de minimización (que buscan los parámetros que dan lugar al menor valor de (14)) son iterativos y como todo proceso iterativo, para iniciarlo es necesario seleccionar unos valores iniciales de los parámetros. Estos valores se suelen tomar iguales a la información previa.
- La minimización de una expresión compleja como

(14) no es tarea sencilla (Fletcher, 1987) y en algunos casos pueden presentarse problemas de convergencia (no alcanzar el valor mínimo). Cuando se tienen dudas sobre si se ha alcanzado, puede ser conveniente comenzar desde otros valores de los parámetros. Los parámetros estimados pueden depender del valor inicial que se elija. En estos casos hay que analizar en profundidad el problema.

- El valor de la función objetivo (14) para un cierto conjunto de parámetros no es por sí solo representativo. Sabiendo únicamente que la función objetivo es de 1000 por ejemplo, no se puede deducir nada sobre la calidad del ajuste del modelo. Únicamente puede proporcionar una cierta información relativa, es decir, si para un conjunto de valores la función (14) vale 1000 y para otro 100, el último valor es el que proporciona un mejor ajuste.
- El valor asignado a  $\lambda$  puede ser importante. Una de las dificultades del problema inverso es que puede estar mal condicionado. En este sentido, el término de la información previa de los parámetros en (14) proporciona una regularización al problema (Tikhonov y Arsenin, 1997). Cuanto mayor sea el valor de  $\lambda_i$  mejor condicionado está el problema. No obstante, imponer un valor excesivamente grande no es solución, porque entonces el valor mínimo de (14) se obtiene para unos parámetros muy parecidos a la información previa, reduciendo mucho la importancia de los otros términos y teniendo por tanto un ajuste pobre en niveles y/o concentraciones.
- Si por el contrario, se reduce el valor de  $\lambda_i$  y se hace tender a cero, el resultado es que se consigue el máximo ajuste de niveles (por simplificar el razonamiento, suponemos que no hay datos de concentración). Esto, no obstante conlleva dos problemas:

*Problema numérico.* Cuanto menor es  $\lambda_i$ , debido a la relación no lineal entre  $h$  y los parámetros de flujo, la minimización de (14) puede ser compleja y fallar el algoritmo de minimización empleado (Fletcher, 1987). En ocasiones se puede evitar esta dificultad resolviendo varios problemas inversos consecutivos empezando por un valor de  $\lambda_i$  grande e ir reduciéndolo. Los valores iniciales de cada problema inverso se toman como los finales del problema anterior (con el  $\lambda_i$  mayor).

*Conceptual.* Si las medidas de nivel,  $h^*$  no tuviesen error, interesaría que el ajuste de niveles fuera lo mejor posible, es decir, que el primer sumando de (14) alcanzara el menor valor posible. No obstante, como las medidas tienen error, no resulta

conveniente ajustar en exceso los valores calculados y medidos, porque se podría estar ajustando el error. Conviene que la diferencia media entre los valores medidos y calculados (en valor absoluto) no sea mucho menor que el error de las medidas.

- Es importante remarcar que siempre es básica la definición del modelo conceptual adecuado. Si el modelo conceptual no responde a la realidad que se desea representar, de nada sirve intentar buscar el mínimo de (14), porque éste puede hallarse con unos valores de los parámetros imposibles en el sistema que se está modelando. En este caso, la minimización de la función objetivo nos puede indicar que hay que buscar un nuevo modelo conceptual.

Como se deduce de las reflexiones anteriores, la elección del valor de  $\lambda$  puede ser crítica en el proceso de calibración, por tanto sería conveniente disponer de métodos para hallar el mejor valor de estos pesos. No obstante, aún no se dispone de un algoritmo que permita calcularlos de forma precisa y directa. En la mayoría de casos, lo que se hace es partir de un valor inicial e ir modificando éstos sucesivamente.

Debido a todas las incertidumbres, no es suficiente con tener un valor estimado de los parámetros, sino que resulta conveniente tener también información sobre su grado de incertidumbre. En la calibración manual se obtiene a partir de simulaciones realizadas variando los valores de los parámetros, en la automática se puede obtener a partir de la matriz de covarianza de los parámetros estimados. Con el método de máxima verosimilitud, es posible obtener una cota inferior de la matriz de covarianza como

$$\Sigma = \left[ \tau_h^{-1} J_h^t V_h^{-1} J_h + \tau_c^{-1} J_c^t V_c^{-1} J_c + C_p^{-1} \right]^{-1} \quad (15)$$

Al utilizar esta expresión, debe tenerse presente que:

- 1.-  $\Sigma$  es en realidad una aproximación de la verdadera cota de varianza mínima.
- 2.- El estimador de Máxima Verosimilitud alcanza la varianza mínima de forma asintótica, es decir, cuanto mayor sea la muestra, más cercana a la cota estaría la varianza real de los parámetros. Como es obvio, todas las muestras son finitas.

#### *Criterios de selección de modelos*

Como se ha comentado en la introducción, para la creación de un modelo, debe realizarse en primer lugar la fase de conceptualización, en la que se seleccionan los procesos físicos que se consideran relevantes, su tratamiento matemático, etc. De esta etapa

no siempre se obtiene un único modelo conceptual, sino que en función de las simplificaciones que se consideren, y de los fenómenos físicos que se tengan en cuenta, pueden resultar varios modelos, a priori todos igualmente plausibles. La "realidad" nunca se hallaría entre éstos, siendo meras aproximaciones de la misma. Nuestro objetivo será pues, tratar de elegir aquel que mejor aproxime la "realidad". Como ésta no se conoce, es necesario encontrar herramientas que permitan elegir de la forma más objetiva posible, entre varios modelos, aquel que represente de modo más fidedigno el sistema físico que se desea modelar.

El axioma de máxima verosimilitud permite comparar diferentes hipótesis en el marco de una estructura específica. Por ello, no sirve para comparar diferentes modelos basados en estructuras distintas.

Posiblemente el método más empleado (y sin duda el más antiguo) es el análisis de residuos (entendiendo por residuo la diferencia entre los valores medidos y los calculados). Utilizando los datos históricos de niveles y de concentraciones, y comparando con los obtenidos por el modelo, es posible sugerir algunas modificaciones del modelo conceptual, para obtener una mejor aproximación de éste a la realidad (entre las modificaciones más utilizadas están el añadir nuevas zonas de algún parámetro, variar su geometría, cambiar alguna condición de contorno, etc.). Estas modificaciones deben estar respaldadas por algún tipo de información relativa al sistema físico (aunque sea únicamente de forma cualitativa), como litología, estructura geológica, hidroquímica, etc., sino se corre el riesgo que el modelo se vaya alejando de la realidad.

En ocasiones un residuo elevado puede estar motivado por más de una causa (p. ej., un nivel alto puede significar que el modelo está considerando una recarga excesiva o una transmisividad muy baja), haciendo difícil su interpretación para su mejora, pudiendo incluso no tener nada que ver el error con los parámetros del modelo, sino con la misma discretización empleada. Por este motivo, es importante que antes de la calibración se haya verificado que la discretización es correcta.

Una forma relativamente simple y que por tanto resulta tentadora, consiste en aumentar el número de parámetros a estimar, lo cual puede mejorar el resultado (tener un mejor ajuste entre valores medidos y calculados), pero puede llevar en muchos casos a una sobreparametrización, es decir, que los parámetros tengan una gran incertidumbre, pudiendo perder su sentido físico. Por ello son necesarios criterios menos subjetivos para analizar el comportamiento de un modelo conceptual respecto de otro.

Los criterios de selección de modelos se deducen

de forma teórica basándose en minimizar una distancia entre el sistema real y el modelo usando teoría de la información, utilizar medidas de probabilidad que aseguren que la probabilidad de elegir el modelo correcto tienda a 1, etc. (Carrera, 1984). Aquí presentamos el criterio de Kashyap (1982):

$$d = S + M \ln(N/2\pi) + \ln|F| \quad (16)$$

donde  $S = -2\ln(L)$ ,  $N$  es el número total de datos,  $M$  es el número de parámetros estimados  $|F|$  es el determinante de la matriz de información de Fisher.

Dados dos modelos conceptuales distintos, aquel cuyo valor de (16) sea menor, debería ser el seleccionado según la teoría. Obsérvese que aparte de las consideraciones teóricas, se está ponderando por un lado el valor de  $S$  (a menor valor, mejor ajuste) y por otro lado se está penalizando el número de parámetros (para evitar la sobreparametrización) y la incertidumbre de éstos que viene dada por el determinante de la matriz de información de Fisher.

## Aplicaciones

Se presentan dos aplicaciones distintas, una perteneciente a un caso real y la otra a un ejemplo sintético, para mostrar algunas de las posibilidades de aplicación del problema inverso. En la primera se presenta una aplicación a un caso real del problema inverso geoestadístico. En la segunda se muestra una aplicación de los criterios de selección de modelos.

### Problema inverso geoestadístico

El planteamiento en este caso es muy similar al presentado en la ecuación (1), con la diferencia que los parámetros pueden ser funciones aleatorias. En el caso del ejemplo que se presenta, concretamente se supone que la transmisividad es un campo aleatorio (se supone que el campo de transmisividades sigue una distribución log-normal). La función objetivo a la que se llega es la misma que en (14). El problema inverso geoestadístico se aplica cuando se sabe que alguno de los parámetros físicos presenta una gran heterogeneidad. En la literatura hay diferentes aproximaciones a este problema, Gómez-Hernández *et al.* (1997), Carrera *et al.* (1992), Kitanidis y Vomvoris (1983). En este ejemplo se muestra su aplicación a la calibración del campo de transmisividades de una fractura del proyecto GAM (GAs Migration in Shear zones) cuyo objetivo es el de obtener un mejor cono-

cimiento de los procesos involucrados en el transporte de gas en fracturas.

La fractura que se estudia en el proyecto GAM se encuentra ubicada en la parte sur del Grimsel Test Site (GTS). En esta zona existe una familia de planos de fractura que buzan en la dirección NE-SO y muestran estructuras internas similares. Macroscópicamente las zonas de cizalla se caracterizan por la existencia de zonas de deformación dúctil que forman bandas de milonita ricas en mica. Existen zonas donde la deformación es muy elevada dando lugar a la aparición de horizontes de material muy frágil con espesores que van desde unos pocos milímetros hasta centímetros. En estas zonas de mayor deformación hay un material muy fino y no cohesivo llamado harina de falla (fault gouge) que podría formar una red de canales preferentes para el flujo de agua en la fractura. Bozzart y Mazurek (1991) estimaron la porosidad de esta harina de falla que oscila entre 0.1 y 0.3. Por otro lado, estudios recientes realizados por Marshall y Croisé (1999) muestran que la distribución de aperturas en la zona de mayor deformación se puede describir mediante una distribución log-normal. La fractura del proyecto GAM ha sido perforada con 18 sondeos equipados con obturadores que aíslan el tramo del sondeo de interés (la intersección entre el sondeo excavado y la fractura). La intersección de la galería con la fractura ha sido sellada con resina para impedir el flujo de agua desde la fractura hacia la intersección de esta con la galería (Fig. 1).

Para caracterizar hidráulicamente la fractura se han realizado siete ensayos de inyección de pulso en otros tantos sondeos para estimar la transmisividad en un entorno pequeño alrededor del punto de inyección, y cinco ensayos de bombeo para obtener información sobre la heterogeneidad de la fractura en una zona más amplia. Aquí se presentan los resultados de la calibración conjunta de los cinco ensayos de bombeo y niveles en régimen estacionario, es decir, como información para la estimación se emplearon los datos de descenso en los puntos de observación en cada ensayo y los datos de nivel estacionario. La fractura se ha modelado como un dominio rectangular bidimensional de 184\*184 m<sup>2</sup>, centrada alrededor de la galería. La transmisividad se trata como una variable regionalizada y se ha dividido el dominio en 483 zonas de T, siendo el tamaño de éstas menor donde se requiere un mayor grado de refinamiento. El coeficiente de almacenamiento se toma homogéneo para todo el dominio. El modelo geoestadístico adoptado tiene un variograma exponencial con una meseta (varianza de  $\ln(T)$ ) de 7.4, una estructura de correlación anisótropa con rangos de 0.5 m en la dirección horizontal y 0.3 m en la dirección vertical. La hipóte-



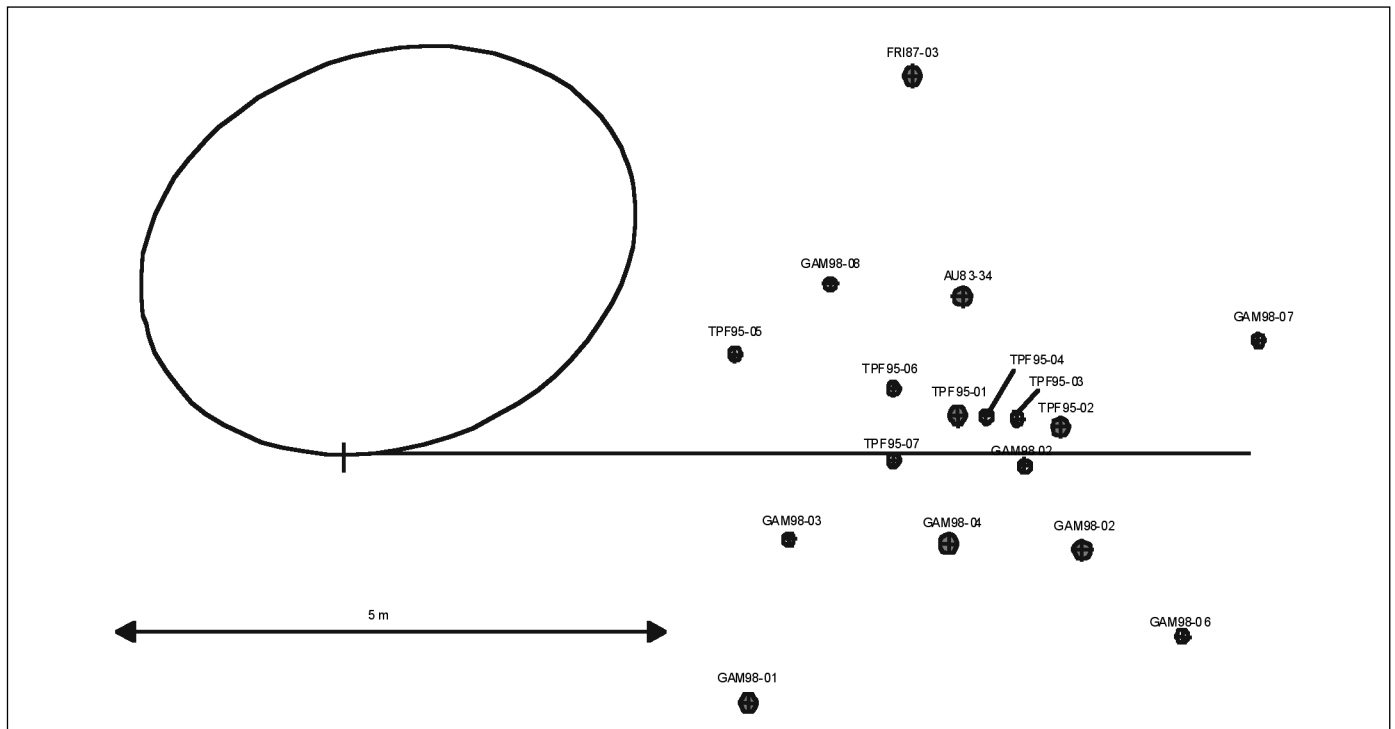


Fig. 1: Plano de la fractura del proyecto GAM. Los puntos indican las intersecciones de los sondeos con el plano de fractura. La elipse es la intersección del túnel desde el que se han excavado los sondeos con el plano de la fractura

Fig. 1: Fracture plane within the GAM project. Points correspond to the intersections of the bore holes with the fracture plane. The ellipse is the intersection of the fracture with the tunnel from which the bore holes have been excavated

sis de anisotropía en la estructura de correlación está basada en la información cualitativa que se tiene sobre canales de alta conectividad hidráulica existentes en la fractura. Para la calibración automática se ha empleado el programa TRANSIN-II (Medina *et al.*, 1996). El campo de transmisividades estimado se puede ver en la Fig. 2, donde se aprecian zonas de alta y baja transmisividad correspondiendo a posibles canales y barreras de flujo respectivamente. Se incluyen también algunos de los ajustes obtenidos.

En este ejemplo se observa como la calibración automática permite obtener campos heterogéneos que con la calibración manual no resulta posible. A pesar de la gran cantidad de datos para ajustar, la semejanza entre los descensos medidos y calculados es bastante buena como se ve en la Fig. 2.

### Criterios de selección de modelos

En este ejemplo se muestra la aplicación de los criterios de selección. Para ello, se define un modelo sintético que es "la realidad". Sobre esta "realidad"

se realizan tres conceptualizaciones, de las que una coincide con "la realidad".

- *Modelo B (real)*. El pozo de bombeo está abierto en una zona de gravas y arena, por encima de la cual hay una capa de limos formando un acuitardo que durante el proceso de bombeo recargaría (gotearía) a la parte inferior. De forma simplificada, puede suponerse como primera aproximación que el flujo en el acuífero es prácticamente horizontal y en los limos que recargan a éste, el flujo es prácticamente vertical.
- *Modelo A*. La situación es muy similar a la del modelo real, con la única diferencia de suponer que el agua adicional de recarga de los limos no viene de éstos, sino que por encima de ellos hay otro acuífero que no se explota directamente y cuyos niveles no se ven afectados por el bombeo (de donde proviene el agua de recarga del acuífero inferior). Desde el punto de vista de modelación, ello implica imponer una condición de contorno de nivel prescrito en los nudos superiores de la malla vertical. Físicamente, la diferencia con el modelo "real" está en que con este modelo los

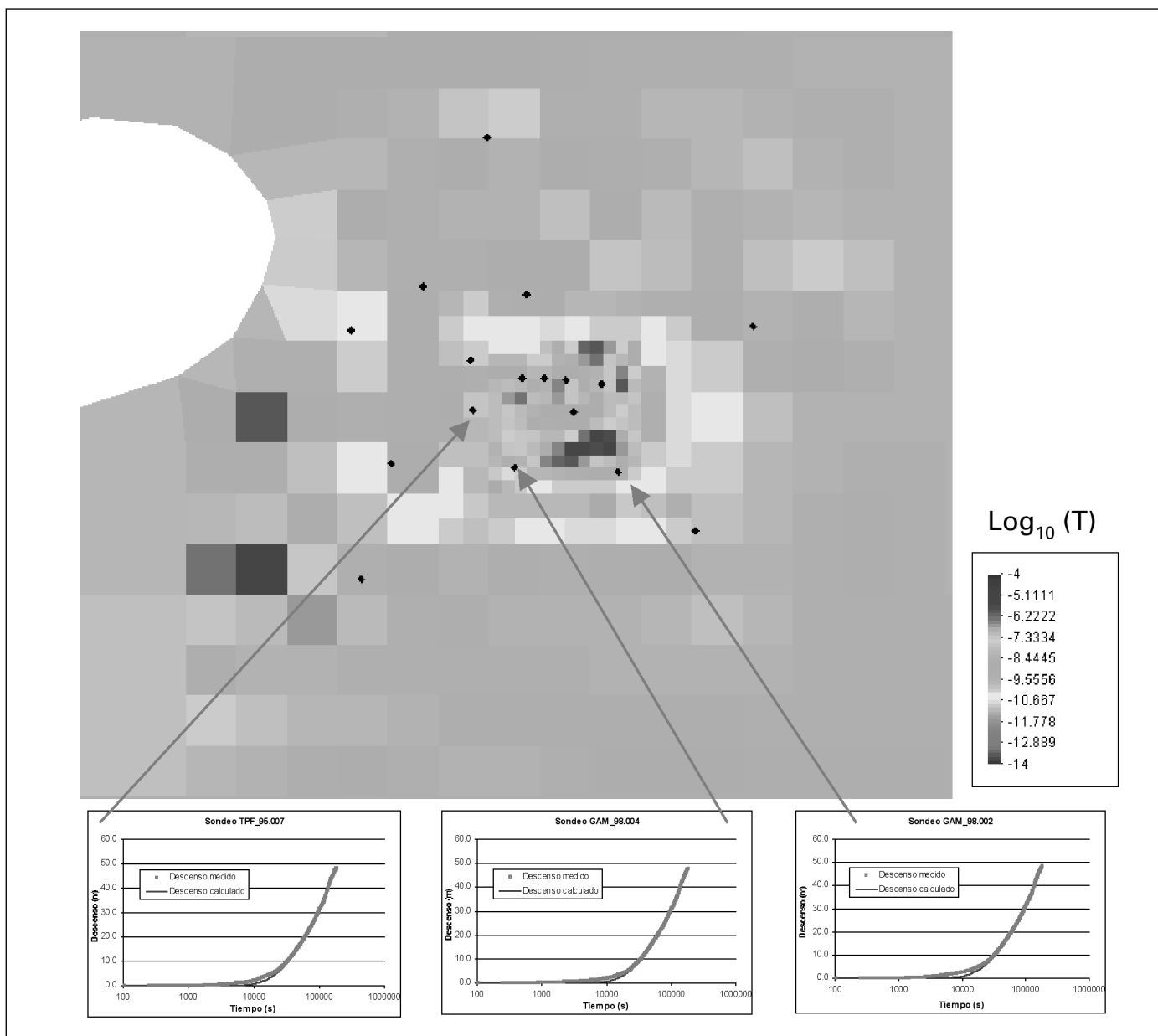


Fig. 2. Campo de Log-transmisividades obtenido tras la calibración conjunta de los cinco ensayos de bombeo y los niveles estacionarios mediante el problema inverso geostatístico. Se incluyen los descensos medidos y calculados en algunos de los puntos de observación

Fig. 2: Log-transmissivity field obtained after joint calibration of five pumping tests and steady-state heads using geostatistical inverse problem. Some examples of measured and computed drawdowns in selected observation points are shown

niveles se estabilizan llegando a un régimen estacionario, mientras que en el "real" los niveles se estabilizan durante un cierto período de tiempo (fase de drenaje diferido) para seguir descendiendo posteriormente.

- *Modelo C.* Este modelo más simple supone que la recarga procedente de los limos es lo suficiente-

mente pequeña como para considerarla despreciable. Es decir el acuífero se trata como confinado, reduciéndose el problema a uno unidimensional. Este modelo conceptual no podría explicar que los niveles tiendan a estabilizarse (fase de transición del modelo de drenaje diferido). Una posible forma de explicarlo con este modelo sería que el

acuífero fuese confinado en un entorno del pozo de bombeo y libre fuera de éste. Para ello, se ha supuesto que existen dos zonas de almacenamiento: una próxima al pozo de bombeo, en la que el acuífero funciona como confinado, y otra más alejada, en la que funciona como libre.

La discretización elegida para el modelo real consiste en una malla unidimensional, dado que el problema de flujo tiene simetría radial. Para poder tener en cuenta la aportación del acuitardo en forma de goteo hacia el acuífero, se considera una malla vertical compuesta por dos nudos sobre cada uno de los de la malla unidimensional del acuífero (en el modelo C que no hay goteo, se utiliza únicamente la parte horizontal de la malla). Como condiciones de contorno se toma de caudal constante en el pozo de bombeo y de nivel fijo a 500 m del mismo, para que no influya esta condición en el desarrollo del problema. Como condición inicial se toma un nivel constante e igual a 10 m en todos los nudos. Las condiciones iniciales de concentración son de concentración nula en todos los puntos, salvo en el de inyección, donde se pone una masa de 2400 gramos, lo cual equivale a la

inyección de un pulso de trazador en el momento de inicio del bombeo.

Como nos encontramos en un ejemplo sintético, hay que generar "las medidas". El proceso de generación consta de tres pasos:

1. En primer lugar se eligen unos valores de parámetros que se van a considerar como valores "verdaderos" del sistema físico y que se muestran en la Tabla 1.
2. Con estos valores "reales" y utilizando el modelo "real" se efectúa una simulación obteniendo los valores "verdaderos" de niveles y concentraciones en los puntos de observación (pozo de bombeo y de inyección para niveles y pozo de bombeo para concentraciones).
3. Como todo proceso de medida está sujeto a errores, a los valores "verdaderos" se les añade un ruido con distribución normal y media nula, con tres valores de desviación:  $\sigma=0.02$ ,  $\sigma=0.1$  y  $\sigma=0.4$  para niveles y  $\sigma=0.02$ ,  $\sigma=0.1$  y  $\sigma=2$  para concentraciones. Estos tres valores representan tres grados de precisión en el proceso de medición. De este modo se obtienen tres conjuntos de medidas

	Par. Verd.	$\sigma=0,02$			$\sigma=0,1$			$\sigma=0,04$		
		A	B	C	A	B	C	A	B	C
F.obj.	-	3800	93.62	31878	249.5	93.63	1343	101.7	93.30	94.76
$d_k$	-	-47.32	-350.9	-40.11	-4.42	-80.48	33.20	242.2	239.3	233.9
PARÁMETROS										
T	5.0	5.002	4.999	5.070	5.012	4.997	5.056	5.075	4.969	5.017
$K_v$	$10^{-1}$	0.1102	$9.93 \cdot 10^{-2}$	-	0.1057	$9.65 \cdot 10^{-2}$	-	$8.99 \cdot 10^{-2}$	$8.72 \cdot 10^{-2}$	-
$S_s$	$10^{-4}$	$1.29 \cdot 10^{-4}$	$9.9310^{-5}$	$1.2710^{-4}$	$1.22 \cdot 10^{-4}$	$9.65 \cdot 10^{-5}$	$1.25 \cdot 10^{-4}$	$1.00 \cdot 10^{-4}$	$8.94 \cdot 10^{-5}$	$1.04 \cdot 10^{-4}$
$S_1$	-	-	-	0.1456	-	-	0.1345	-	-	$9.70 \cdot 10^{-2}$
$S_y$	0.3	-	0.2948	-	-	0.2751	-	-	0.2288	-
$\alpha_L$	1.0	1.030	1.002	1.429	1.041	1.011	1.427	1.241	1.219	1.645
$\theta$	0.1	0.1068	0.1002	0.1219	0.1076	0.1010	0.1222	0.1197	0.1130	0.1352

Tabla 1. Resultados obtenidos mediante la estimación conjunta de parámetros de flujo y transporte

$d_k$ : Criterio de Kashyap (ec. 16). T: Transmisividad.  $K_v$ : Conductividad vertical de los limos.  $S_s$ : Coeficiente de almacenamiento específico.  $S_1$  porosidad drenable del acuífero en el modelo C.  $S_y$ : Porosidad acuitardo.  $\alpha_L$ : Dispersividad longitudinal del acuífero.  $\theta$ : Porosidad del acuífero

Table 1: Results from the joint estimation of flow and transport parameters.

$d_k$ : Kashyap's criterion (eq. 16). T: Transmissivity.  $K_v$ : Vertical conductivity of the silts.  $S_s$ : Specific storage coefficient.  $S_1$ : Aquifer drainable porosity in model C.  $S_y$ : Porosity of the aquitard.  $\alpha_L$ : Aquifer longitudinal dispersivity.  $\theta$ : Aquifer porosity.

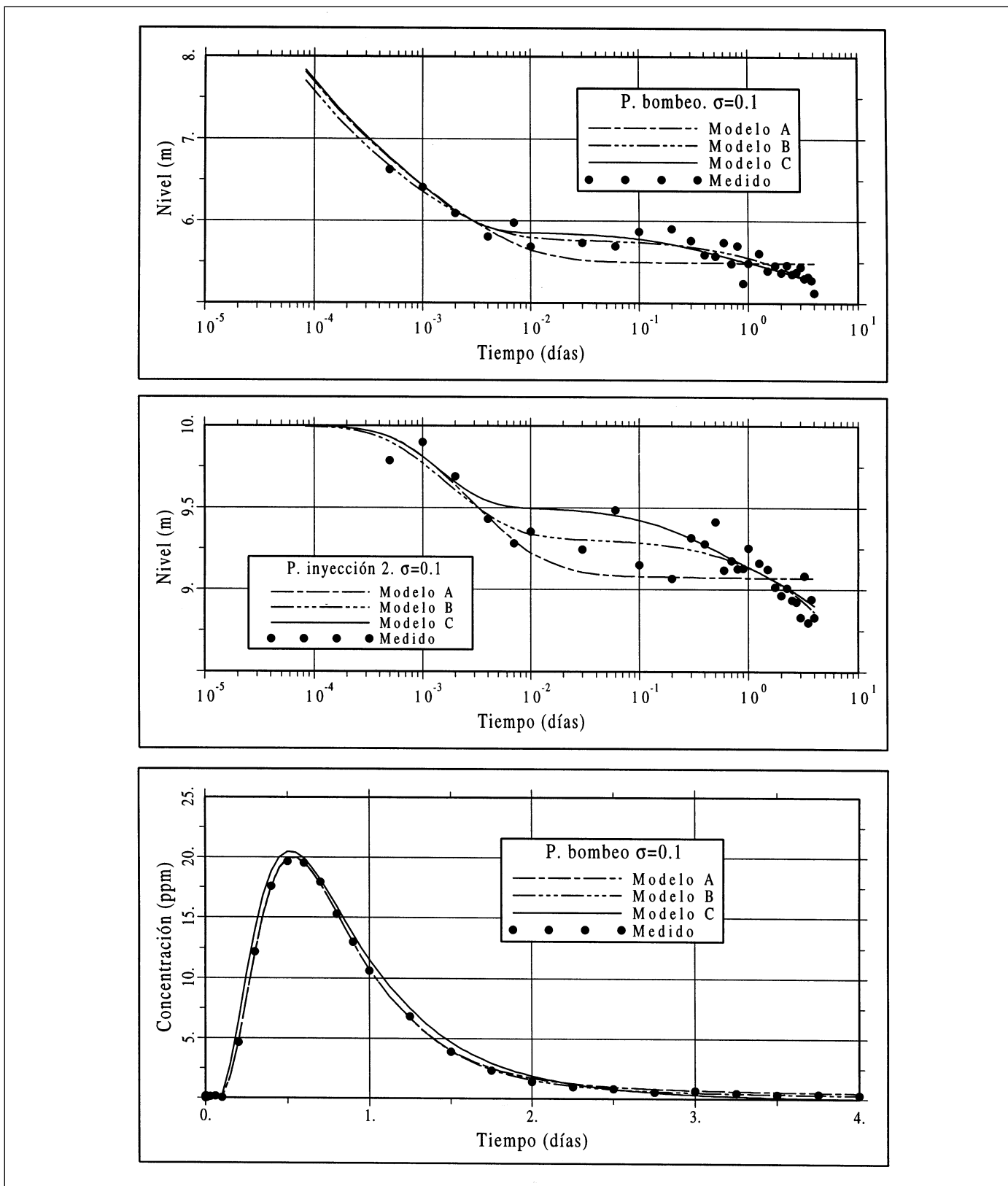


Fig. 3. Niveles y concentraciones medidos (con ruido  $\sigma = 0.1$ ) y calculados

Fig. 3: Measured and computed heads and concentrations (noise  $s=0.1$ )

sobre los que poder analizar los diferentes criterios.

Presentamos los resultados obtenidos calibrando los datos de niveles y concentraciones con el código TRANSIN-II. En la tabla 1 se presentan los valores estimados de los parámetros y el criterio de Kashyap. Lo primero que se observa en la Tabla 1 es que los valores estimados de los parámetros están bastante ajustados a la realidad para los tres niveles de ruido, aunque evidentemente, la semejanza con el valor real disminuye a medida que los datos contienen un error mayor. Los ajustes entre los niveles y concentraciones medidos y calculados para un nivel de ruido  $\sigma = 0:1$  se pueden ver en la Fig. 3. También hay que destacar en la Tabla 1 el valor del criterio de Kashyap. Para los dos primeros niveles de ruido selecciona correctamente el modelo "verdadero" (recuérdese que el modelo seleccionado es el que tiene menor valor numérico de este criterio), pero falla en el nivel de ruido mayor. Se ve también que el modelo C es el peor considerado excepto en el caso de mayor error.

Se deduce, pues, que los criterios de selección de modelos pueden ser de utilidad, aunque no están muy extendidos y verificado su rango de aplicación. Aunque la presentación se ha realizado en un ejemplo sintético, es posible su empleo en casos reales (Medina y Carrera, 1996).

## Conclusiones

Se ha presentado una breve descripción de la teoría del problema inverso para una de las posibles formulaciones. Asimismo se ha mostrado su aplicabilidad en dos ejemplos. A lo largo del desarrollo se han discutido algunos puntos que conviene destacar:

En primer lugar, independientemente del método que se vaya a emplear para calibrar, resulta fundamental disponer de un modelo conceptual que represente de la forma más precisa posible el sistema que se desea modelar. Hay que tener presente además, que el proceso de calibración nos puede conducir en cualquier momento a replantearse el modelo conceptual, debido a resultados anómalos: valores absurdos de los parámetros, imposibilidad de explicar los niveles y/o las concentraciones con el modelo, etc.

En segundo lugar, hay que tener muy presente que todos los métodos de calibración automática llevan a la minimización de una función análoga a (14) que es la que controla la diferencia entre los valores medidos y calculados de niveles y/o concentraciones. Como ya se ha mencionado antes, la minimización de una función de este tipo no es tarea simple, pudiendo tener problemas de convergencia que lleven a pará-

metros absurdos o que el algoritmo de optimización sea incapaz de alcanzar el mínimo. En ocasiones, empezar desde unos valores iniciales distintos puede arrojar alguna luz.

Es importante cuantificar bien los errores de cada uno de los tipos de medidas, para ponderar adecuadamente las diferentes fuentes de información (niveles, concentraciones, etc.), pues como ya se ha dicho los pesos asignados a las diferentes piezas de información juega un papel importante en todo el proceso de calibración. En cualquier caso, no es conveniente ajustar en exceso los niveles y/o concentraciones medidos y calculados, porque si los errores de las medidas son importantes, se puede estar forzando al modelo a buscar valores de los parámetros que no tienen sentido físico. Concretamente, no resulta conveniente ajustar dichos valores por debajo de su error de medida.

En ocasiones puede suceder que el algoritmo de minimización no converja para pesos de la información previa ( $\lambda_i$ ) bajos. En estos casos puede convenir empezar con un valor de  $\lambda_i$  alto e ir reduciéndolo aplicando el problema inverso de forma sucesiva. Este proceso se ha realizado en el primero de los ejemplos aquí presentados.

Aunque no se ha presentado en ninguno de los ejemplos la aplicación de la matriz de covarianza ( $\Sigma$ , ecuación 15), ésta puede ser de utilidad en la selección de los parámetros a calibrar. Por ejemplo, si un parámetro tiene una varianza de estimación alta (valores de la diagonal de  $\Sigma$ ), significa que los datos son poco sensibles al mismo, por lo que posiblemente sea mejor fijar su valor obtenido mediante alguna fuente de información (incluso cualitativa) ya que el valor que obtengamos sería poco representativo dada la poca sensibilidad de los datos a éste.

En resumen, la calibración automática evita al modelista el trabajo tedioso de modificar los valores de los parámetros y realizar un sinnúmero de simulaciones, pero hay que emplearla con un conocimiento hidrogeológico profundo sobre el sistema que se desea modelar y con una gran capacidad crítica para analizar los resultados obtenidos en dicha calibración.

## Agradecimientos

Este trabajo se ha realizado en el marco de proyectos de investigación financiados por ENRESA a través del convenio 770067. También se ha contado con financiación del proyecto de investigación de la CICYT HID99-0838.

## Referencias

- Bear, J. 1972. "Dynamics of fluids in porous media". Dover publications.
- Bossart, P. y Mazurek, M. 1991. "Structural geology and water flow paths in the migration shear zone". NAGRA Internal Report 91-12, Wettingen, Switzerland.
- Carrera, J., 1984. "Estimation of aquifer parameters under transient and steady state conditions." *Ph. D. dissertation*, Dep. of hydrol. and Water Resour., Univ. of Arizona, Tucson.
- Carrera, J. y Neuman, S. P. 1986. "Estimation of aquifer parameters under steady-state and transient conditions: I. Background and Statistical framework", *Wat. Resour. Res.*, 22, 2, 199-210.
- Carrera, J. "State of the art of the inverse problem applied to the flow and solute transport equations". In *Analytical and Numerical Groundwater Flow and Quality Modelling* (E. Custodio, A. Gurgui and J. P. Lobo Ferreira, eds.) NATO-ARW Series Reidel, 1987.
- Carrera J., Medina, A. y Sánchez-Vila, X. 1992. "Geostatistical Formulation of Groundwater Coupled Inverse Problems". TROIA'92. Fourth Int. Geostatistics Congress. Vol. 2, pp. 779-793.
- Custodio, E. y Llamas, M. R. 1983. "Hidrología subterránea". Segunda edición. Editorial Omega. Barcelona. 2350 pp.
- Edwards, A. W. F. 1972. "Likelihood". University Press, Cambridge.
- Fletcher, R. 1987. "Practical methods of optimization". John Wiley & Sons.
- Gómez-Hernández, J. J., Sahuquillo, A. y Capilla J. E. 1997. "Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data. 1. Theory". *Journal of Hydrology*, 1-4(203), 162-174.
- Kashyap, R. L. 1982. "Optimal choice of AR and MA parts in autoregressive moving average models". *IEEE Trans Pat. Anal. Mach. Intell.*, PAMI-4(2):99-104.
- Kitanidis and Vomvoris (1983). "A geostatistical approach to the inverse problem in groundwater modeling (steady-state) and one-dimensional simulations". *Water Res. Res.*, 19 (3), pp 677-690.
- Marschall, P. y Croisé, J., 1999. Grimsel Test Site. Determination of two-phase flow properties in shear zones. NAGRA Internal Report NTB 97-06, Wettingen, Switzerland.
- McLaughlin, D y Townley, L. R. 1996. "A reassessment of the groundwater inverse problem". *Wat. Resour. Res.* 32 (5), pp. 1131-1161.
- Medina, A., Galarza, G. y Carrera, J., 1996. Transin-II. Fortran code for solving the coupled flow and transport inverse problem in saturated conditions. In *El Berrocal Project*. Topical Report, 4(16).
- Medina, A. y Carrera, J. 1996. "Coupled estimation of flow and solute transport parameters". *Wat. Resour. Res.* 32 (10), pp. 3063-3076.
- Neuman, S. P. y Yakowitz, S. 1979. "A statistical approach to the inverse problems of aquifer hydrology, 1. Theory". *Wat. Resour. Res.* 15 (4), 845-860.
- Tikhonov, A. N. y Arsenin, V. Y. 1977. "Solutions of ill-posed problems". Halsted press/Wiley, New York.
- Yeh, W. W.-G. 1986. "Review of parameter identification procedures in groundwater hydrology: The inverse problem". *Water Resour. Res.* 22 (2), 95-108.