Inverse problem in hydrogeology

Jesús Carrera · Andrés Alcolea · Agustín Medina · Juan Hidalgo · Luit J. Slooten

Abstract The state of the groundwater inverse problem is synthesized. Emphasis is placed on aquifer characterization, where modelers have to deal with conceptual model uncertainty (notably spatial and temporal variability), scale dependence, many types of unknown parameters (transmissivity, recharge, boundary conditions, etc.), nonlinearity, and often low sensitivity of state variables (typically heads and concentrations) to aquifer properties. Because of these difficulties, calibration cannot be separated from the modeling process, as it is sometimes done in other fields. Instead, it should be viewed as one step in the process of understanding aquifer behavior. In fact, it is shown that actual parameter estimation methods do not differ from each other in the essence, though they may differ in the computational details. It is argued that there is ample room for improvement in groundwater inversion: development of user-friendly codes, accommodation of variability through geostatistics, incorporation of geological information and different types of data (temperature, occurrence and concentration of isotopes, age, etc.), proper accounting of uncertainty, etc. Despite this, even with existing codes, automatic calibration facilitates enormously the task of modeling. Therefore, it is contended that its use should become standard practice.

Résumé L'état du problème inverse des eaux souterraines est synthétisé. L'accent est placé sur la caractérisation de l'aquifère, où les modélisateurs doivent jouer avec l'incertitude des modèles conceptuels (notamment la variabilité spatiale et temporelle), les facteurs d'échelle, plusieurs inconnues sur différents paramètres (transmissivité, recharge, conditions aux limites, etc.), la non linéarité, et souvent la sensibilité de plusieurs variables d'état (charges hydrauliques, concentrations) des propriétés de l'aquifère. A cause de ces difficultés, le calibrage ne peut être séparé du processus de modélisation, comme c'est le cas dans d'autres cas de figure. Par ailleurs, il peut être vu comme une des étapes dans le processus de détermination du comportement de l'aquifère. Il est montré que les méthodes d'évaluation des paramètres actuels ne diffèrent pas si ce n'est dans les détails des calculs informatiques. Il est montré qu'il existe une large panoplie de techniques d ‹inversion : codes de calcul utilisables par tout-un-chacun, accommodation de la variabilité via la géostatistique, incorporation d'informations géologiques et de différents types de données (température, occurrence, concentration en isotopes, âge, etc.), détermination de l'incertitude. Vu ces développements, la calibration automatique facilite énormément la modélisation. Par ailleurs, il est souhaitable que son utilisation devienne une pratique standardisée.

Resumen Se sintetiza el estado del problema inverso en aguas subterráneas. El énfasis se ubica en la caracterización de acuíferos, donde los modeladores tienen que enfrentar la incertidumbre del modelo conceptual (principalmente variabilidad temporal y espacial), dependencia de escala, muchos tipos de parámetros desconocidos (transmisividad, recarga, condiciones limitantes, etc), no linealidad, y frecuentemente baja sensibilidad de variables de estado (típicamente presiones y concentraciones) a las propiedades del acuífero. Debido a estas dificultades, no puede separarse la calibración de los procesos de modelado, como frecuentemente se hace en otros campos. En su lugar, debe de visualizarse como un paso en el proceso de entendimiento del comportamiento del acuífero. En realidad, se muestra que los métodos reales de estimación de parámetros no difieren uno del otro en lo esencial, aunque sí pueden diferir en los detalles computacionales. Se discute que existe amplio espacio para la mejora del problema inverso en aguas subterráneas: desarrollo de códigos amigables al usuario, acomodamiento de variabilidad a través de geoestadística, incorporación de información geológica y diferentes tipos de datos (temperatura, presencia y concentración de isótopos, edad, etc), explicación apropiada de incertidumbre, etc. A pesar de esto, aún con los códigos existentes, la calibración automática facilita enormemente la tarea de mode-

Received: 10 June 2004 / Accepted: 2 November 2004
Published online: 25 February 2005

J. Carrera (✉) · A. Alcolea · A. Medina · J. Hidalgo · L. J. Slooten
School of Civil Engineering,
Technical University of Catalonia,
1-3, Jordi Girona Street, Building D2, 08034 Barcelona, Spain
e-mail: jesus.carrera@upc.edu
Tel.: +34-93-4016890
Fax: +34-93-4017251

 DOI 10.1007/s10040-004-0404-7

lado. Por lo tanto, se sostiene que su uso debería de convertirse en práctica standard.

Keywords Inverse problem · Aquifer · Groundwater · Modeling · Parameter estimation

Introduction

In broad terms, inverse modeling refers to the process of gathering information about the model from measurements of what is being modeled. This includes two related concepts: model identification and parameter estimation. The latter will be used here as being synonymous with calibration. Model identification applies to methods to find the nature (features) of the model, such as the governing equations, boundary conditions, time regime, or heterogeneity patterns. Parameter estimation, instead, is restricted to assigning values to the properties characterizing those features.

The above definitions apply to groundwater modelling without much modification. In fact, it can be argued that some sort of inversion has always been standard practice in hydrogeology. For example, hydrogeologists perform pumping tests to characterize aquifers. These tests are interpreted by, first, identifying the most appropriate model (on the basis of actual measurements and geology) and, then, estimating model parameters by curve fitting. This can be facilitated by automatic procedures, which have become rather standard in fields where analytical solutions are scarce, such as the unsaturated zone (Kool and Parker 1988; Hollenbeck and Jensen 1998). Automatic calibration has also become standard in fields where appropriate tests can be performed to characterize spatial variability: medicine, where it is termed "scanning" or "tomography" (Rudin et al. 1999; Rühli et al. 2002); geophysics (Sambridge and Mosegaard 2002); and many others. However, groundwater modeling displays several peculiarities that collectively sets it apart from modeling in other fields:

- *Cost*. Groundwater models are relatively expensive to run. They require building large systems of equations (to get an accurate and realistic picture of the system) that need to be solved for each model run. In addition, advances in computer science are used to increase the quality of the model performance rather than to reduce execution times.
- *Time dependence*. State variables such as heads and concentrations are time dependent. Many flow problems are essentially under steady state conditions. Yet, even in this case, the information might be contained in temporal fluctuations of heads, thus deserving transient modeling.
- *Heterogeneity*. Values of hydraulic conductivity, K, which is often the most dominant hydraulic property, may vary over several orders of magnitude. The same can be said about transmissivity, which is essentially equivalent to K in 2D models (hereinafter transmissivity and hydraulic conductivity will be used interchangeably).
- *Different types of parameters*. While efforts are often concentrated on transmissivity, other parameters (recharge, boundary fluxes, etc.) may be equally uncertain and relevant.
- *Scale dependence*. Parameters measured in the field often represent a small portion of the aquifer. As a result, they are qualitative and quantitatively different from what is needed in the model.
- *Model uncertainty*. Geometry of the aquifer and heterogeneity patterns are controlled by the geology, which is never known accurately.
- *Low sensitivity*. Depending on the problem, state variables may display low sensitivity to model parameters (i.e. their information content is low). In particular, heads (the most frequent and sometimes unique type of measurement) sometimes contain little information about hydraulic conductivity.

Because of the above features, the aquifer model predictions are highly uncertain. Moreover, parameter estimation cannot be formulated as clearly as in other fields. That is possible for pumping test interpretation, where one can indeed take heads straight from the test, enter them in a code and derive parameter values after a moderately qualitative model analysis. In aquifers, one is forced to cast inversion as one step of the modeling process. Many of these features are shared by groundwater's sister science, surface water hydrology, where many of these issues also have been addressed (Beven 1993; Gupta et al. 1999).

Another consequence of the singularities of groundwater inversion is its relative isolation. Many inversion methods have been developed independently from those in other fields. The earliest methods were based on simply substituting heads, assumed to be known, into the flow equation, which leads to a first order partial differential equation in transmissivity (Stallman 1956). This method, termed "direct" by Neuman (1973), is relatively simple to understand and has been widely used after Nelson (1960, 1961). In fact, it allowed deriving transmissivities from flow nets based on head measurements (Bennet and Meyer 1952). Unfortunately, this approach has several drawbacks. First, it requires knowing heads (and recharge, storage coefficient and boundary conditions) over the whole domain in space and time. This can only be achieved through interpolation, which introduces smoothing and errors, so that the estimated transmissivity values become somewhat artificial. Second, it is unstable (small errors in heads cause large errors in transmissivity). To overcome the first problem, most recent inversion methods use what Neuman (1973) termed the indirect approach, which consists of acknowledging that measurements contain errors and finding the hydraulic properties that minimize these errors. That is, parameters are found by minimizing an objective function, which may become a huge computational task. To overcome instabilities, a number of approaches can be taken: adding a

 DOI 10.1007/s10040-004-0404-7

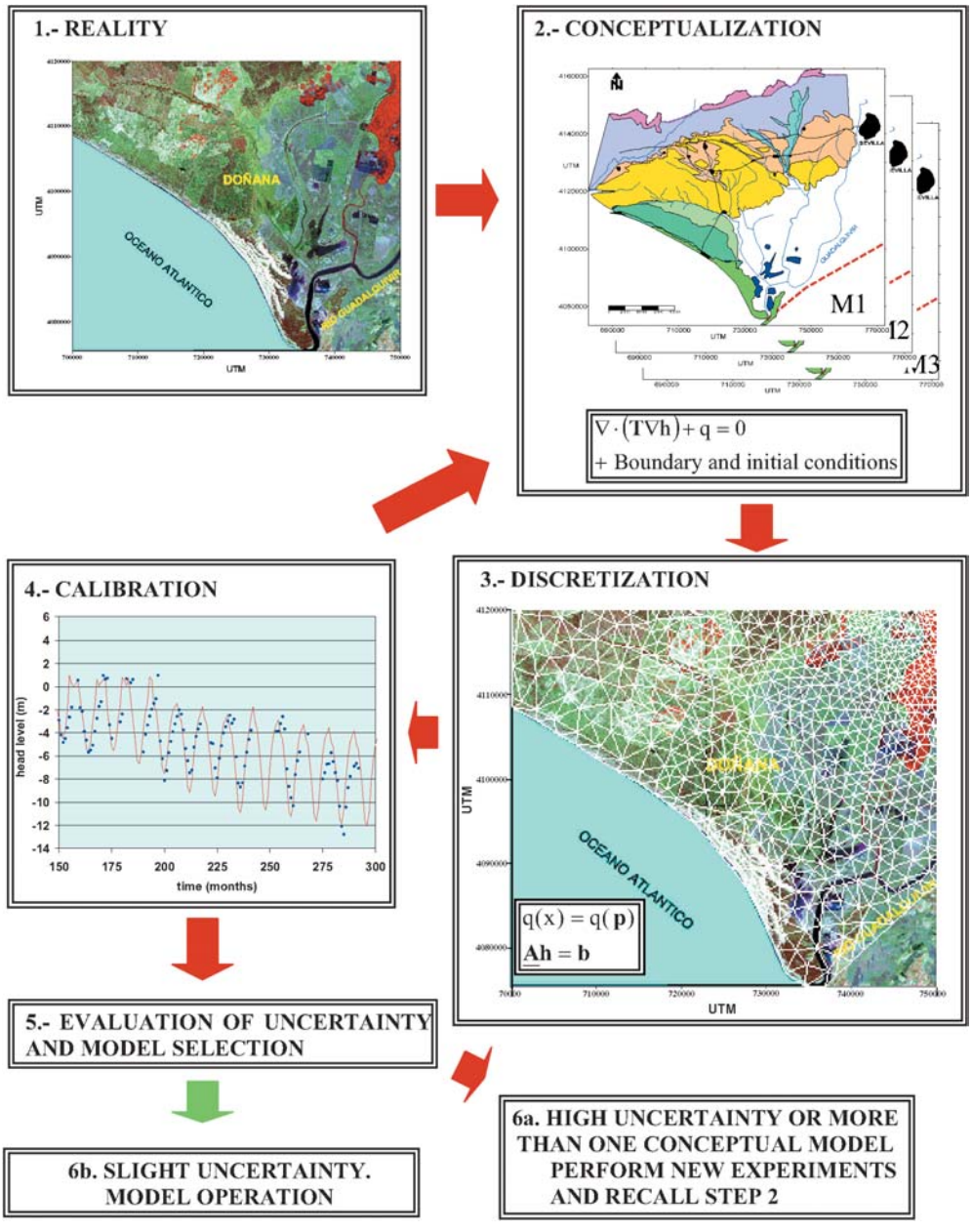

Fig. 1 Schematic representation of the modeling process (the Aznalcóllar site, Castro et al. 1999). **Step 1.** All available information about the real system is collected and used to define the plausible conceptual models (M1, M2, M3, etc.). **Step 2.** Conceptual models are expressed in terms of governing equations (in the inset, T is transmissivity, h head and q represents sink and source terms; color patterns in M1, M2, M3 differentiate different geological formations). **Step 3.** These equations are discretized and solved in a finite set of points (in the inset, **a** is the conductance matrix and **h** and **b** are vectors containing heads and sink/source terms). Also, unknown hydraulic properties (i.e. areal recharge) are discretized as a function of a set of model parameters **p**. **Step 4.** These model parameters are optimized in such a way that (1) a good fit between calculated (*red line*) and measured data (*blue dots*) is obtained and (2) derived hydraulic properties are "plausible". **Step 5.** Error analysis must be performed and, if needed, some of the initial conceptual models can be dismissed through new experiments. Otherwise, the model is operative

regularization term to the objective function to dampen unwarranted oscillations, considering additional types of data, reducing the number of parameters to be estimated without losing the ability to reproduce spatial variability, etc. These issues (computation, stabilization, different types of data, spatial variability) have remained the focus of much research. Reviews of them are presented by Yeh (1986), Carrera (1987), Kool et al. (1987), McLaughlin and Townley (1996) and de Marsily et al. (1999). Therefore, a thorough review of the state of the art in aquifer modeling inversion would be superfluous.

In this context, the objective of this paper is to describe the current state of inverse modeling for aquifer characterization. Emphasis is placed on seeking similarities of existing methods and on discussing how they are used, rather than describing them exhaustively. It is argued that calibration should be viewed as a necessary step in modeling and the paper is structured to mimic the modeling process (Fig. 1). The following section of this paper is devoted to outlining the roles of knowledge and data in the definition of the conceptual model, arguably the most frequent source of errors in groundwater modeling (Bredehoeft 2004). Methods for representing aquifer properties in terms of a hopefully small number of unknowns to be optimized (the so called model parameters) are presented in the section entitled "What is to be estimated?". The next natural step is to optimize (calibrate) the value of the model parameters. Actual calibration methods and codes are presented in the sections, "How to estimate the model parameters, p"; and "How codes work". Once the parameters have been obtained, one needs to check the quality of the resulting model, which is addressed in the

 DOI 10.1007/s10040-004-0404-7

Table 1 Classification of the types of data commonly used in hydrogeology

Type	Example	Use
Knowledge	Darcy's law validity, dimensionality, B.C.'s, etc	Governing equations
SOFT DATA		
Qualitative	Geology (heterogeneity)	Parameterization
Quantitative (indirect)	Geophysics Remote sensing Grain size analysis	
	Measurement error distribution	Calibration
HARD DATA		
Quantitative	Pumping test Tracer test	Prior Information
	Heads Concentrations of chemical constituents	Calibration, validation
	Other measurements	

section entitled, "How good is the model?". Actual application trends are discussed in the section entitled, "What is actually done?". The paper ends with a discussion on future challenges.

The conceptual model: knowledge and data

Knowledge and data are the basis of the conceptual model, which represents what is going to be actually modeled. Knowledge and data concepts have been defined in many different ways. An often encountered point of view is that "knowledge" consists in beliefs about reality. In science and philosophy it is of concern whether beliefs are justifiable and true. In this view, only if beliefs are justifiable and true are they deemed knowledge. The term "data" is used here to mean all the pieces of information about the aquifer, not only hard numbers for measurements in the field.

As shown in Fig. 1, groundwater models are based on generic scientific knowledge on the behavior of groundwater and site specific data. This leads to a site specific understanding, which can be expressed in terms of mathematical equations. These equations are often manipulated to get a set of discretized equations that can be solved numerically.

Data are used not only to define qualitatively the conceptual model, but also to define it quantitatively. Table 1 displays some of the types of data and the role they play in modeling. They can be classified according to several criteria. One distinction can be made between qualitative ("soft") and quantitative ("hard") data. Another distinction can be stated considering their temporal behavior: static data refer to quantities that remain constant in time while dynamic data relate to quantities that change in time. A third distinction can be made between data about the parameters to be estimated and measurements of state variables (heads, fluxes, etc.).

Quantitative data are acquired through measurements, which contain errors. Therefore, a probability density function (hereinafter referred to as pdf) should be used instead of single values to define measurements. A Gaussian distribution is often used, both because it has been proven to suitably describe many variables (e.g. log-transmissivity, hereinafter referred to as log-T) and because it can be fully described by the mean and the standard deviation. Given the ease of use of this distribution, data that do not follow a Gaussian pdf are often transformed so that the resulting pdf is close to Gaussian (e.g. transmissivity). This is a frequent source of errors, because the fact that point values of log-T follow a Gaussian distribution does not imply that the spatial distribution is multigaussian (Gómez-Hernández and Wen 1998). In fact, the opposite can often be argued (Meier et al. 1999). Yet, the multigaussian feature of log-T is the most frequent assumption in hydrogeology. Geostatistical software may be used to generate alternative spatial distributions, for example using Markov-chain methods (Weissmann et al. 1999).

Transformations of raw measurements of state variables may also be appropriate to increase the sensitivity of data to parameters. This topic has received a lot of attention in surface water hydrology (Meixner et al. 1999). A groundwater example of this trend is to use the total observed mass at an outflow point instead of individual concentration values when modeling a breakthrough curve. Peak concentration and peak time can also be used (e.g. Woodbury and Rubin 2000). The same can be achieved by using different types of data or by employing an adequate weighting scheme (see, e.g. Wagner and Gorelick 1987 and Anderman and Hill 1999). Concerning the use of different data types, the information content of data is problem dependent, so that it is difficult to give general rules. Including flow rate data in the calibration improves sensitivity to transmissivity (Larocque et al. 1999). Temperature data can be informative of vertical fluxes (Woodbury et al. 1987). Environmental isotopes and age data may be informative about regional flow trends (Varni and Carrera 1998). Chen et al. (2003) used subsidence rates to help in the estimation of permeability of aquitards. Streamflow gains and losses were used by Hill (1992) to help in model calibration. In short a wide range of state variables can be used.

Data can be incorporated into the inversion procedure in different ways. Most of them can be summarized by considering a vector ω containing both observations of state variables, $\mathbf{h}$, and model parameters, $\mathbf{p}$:

$$\omega = \begin{pmatrix} \mathbf{h} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{h}^* \\ \mathbf{p}^* \end{pmatrix} + \begin{pmatrix} \epsilon_{\mathbf{h}} \\ \epsilon_{\mathbf{p}} \end{pmatrix} \quad (1)$$

where $\mathbf{h}$ and $\mathbf{p}$ represent true values, $\mathbf{h}^*$and $\mathbf{p}^*$ represent measurements (or appropriate transformations; $\mathbf{p}^*$ is usually termed "prior estimate"), and $\epsilon_{\mathbf{h}}$ and $\epsilon_{\mathbf{p}}$ represents errors. It is reasonable to assume that, a priori, errors in measurements of state variables are independent of errors in those of parameters. In fact, it is often reasonable to assume that all measurement errors are independent. Dependence between data arises as a result of the

 DOI 10.1007/s10040-004-0404-7

abovementioned transformations. This is particularly true for model parameters, where measurements never coincide exactly with what is needed for the model. In order to obtain **p***, one should use all available data about the parameters, such as point measurements of the parameters, soft data (e.g. a geophysical image), or the like. All these measurements are contained in a data vector $\mathbf{z}_p$. Then, using $\mathbf{z}_p$, the prior estimation and covariance matrix of **p** can be written as a generalization of the linear conditional expectation equation:

$$\mathbf{p}^* = E\left[\mathbf{p}|\mathbf{z}_p\right] = E(\mathbf{p}) + \mathbf{Q}_{\mathrm{pze}}\mathbf{Q}_{\mathrm{zze}}^{-1}\left(\mathbf{z}_p - E\left[\mathbf{z}_p\right]\right) \quad (2)$$

$$\mathbf{C}_p = E\left[(p - p^*)(\mathbf{p} - \mathbf{p}^*)^{\mathbf{t}}|\mathbf{z}_p\right] = \mathbf{Q}_{\mathrm{pp}} - \mathbf{Q}_{\mathrm{pze}}\mathbf{Q}_{\mathrm{zze}}^{-1}\mathbf{Q}_{\mathrm{pze}}^{t} \quad (3)$$

where $\mathbf{Q}_{\mathrm{pp}}$, $\mathbf{Q}_{\mathrm{zze}}$, and $\mathbf{Q}_{\mathrm{pze}}$ are respectively the (co)variances of model parameters, of measurements in $\mathbf{z}_p$, and cross-covariances between model parameters and measurements in $\mathbf{z}_p$ and $E[\cdot]$ stands for expected value. In practice, the way in which these matrices are evaluated depends on the nature of **p** and of $\mathbf{z}_p$. If $\mathbf{z}_p$ represents point measurements, then the components of $\mathbf{Q}_{\mathrm{zze}}$ are obtained from the covariance function (i.e., $\mathbf{Q}_{\mathrm{zze}}$ would be the simple kriging matrix). In such a case, $\mathbf{Q}_{\mathrm{pze}}$ would also be given by the covariance function, possibly averaged if **p** represents block values of log-T. If the prior expectations of **p**, $E(\mathbf{p})$, and $\mathbf{z}_p$, $E\left[\mathbf{z}_p\right]$, are not known, then one may need to extend these matrices to impose unbiasedness conditions (hence the subindex e). Moreover, if $\mathbf{z}_p$ includes not only point measurements but also soft data, then the correlation among them should be included in $\mathbf{Q}_{\mathrm{zze}}$, which would then become the cokriging matrix. Estimating cross covariances is difficult, so that one may instead prefer to use universal kriging, kriging with an external drift or other forms of non-stationary kriging. It is clear that computing these matrices can be technically complex. Details can be found in geostatistics texts (e.g., Kitanidis 1997; Rubin 2003). A step-by-step description of how to handle increasingly large sets of different types of data is provided by Carrera et al. (1993a). However, regardless of the details on how these matrices are built, two issues should be clear:

- Virtually all types of data on both state variables and model parameters can be accommodated, except data that are strictly qualitative (such as geological descriptions).
- Statistical descriptions of these data are needed to properly weigh them in the inversion.

The meaning of **p** in Eqs. (1), (2) and (3) has gone purposefully undefined. This is the subject of the next section.

What is to be estimated?

The coefficients in the governing equations (see Fig. 1) are the hydraulic properties, whose value is to be estimated. Regardless of their scalar or tensorial character, all of them vary in space and some of them may vary in time as well. To obtain the value of the hydraulic properties at every point of a continuous model domain is impossible. Therefore, a discrete representation is required. The process of expressing hydraulic properties in terms of a hopefully small number of model parameters (unknowns to be found during the inversion process) is termed parameterization.

Choosing the set of model parameters is not easy. Modelers may tend to use many degrees of freedom to get an accurate description of variability. However, if the number of model parameters is large, inverse problems may become ill-posed (Hadamard 1902). Therefore, a compromise is needed, which is the motivation of parameterization methods. Most of them express aquifer properties as linear combinations of the unknown model parameters (those to be estimated). That is, any hydraulic property q (e.g., areal recharge) can be written as:

$$q(\mathbf{x}, t) = q_0(\mathbf{x}, t) + \Sigma_{i=1}^{N} p_i \cdot \alpha_i^q(\mathbf{x}, t) \quad (4)$$

Here, N is the number of model parameters p_i, $q_0(\mathbf{x})$ is an additive factor of the hydraulic property (e.g. the initial value if initial model parameters are zero) and $\alpha_i^q(\mathbf{x}, \mathbf{t})$ are interpolation functions used to parameterize areal recharge in space and time. Discretization allows one to write Eq. (4) in matrix form

$$\mathbf{q} = \mathbf{q}_0 + \mathbf{A}^q\mathbf{p} \quad (5)$$

where $\mathbf{A}^{\mathrm{q}}$ is a matrix containing the areal recharge interpolation functions and **p** a vector containing the model parameters. Thus, parameterization methods are defined in terms of: (1) an estimation of the model parameters **p**, (2) an interpolation method that describes the way matrix **A** is calculated, and (3) an initial distribution of the hydraulic property $\mathbf{q}_0$. Parameterization methods differ in the way in which these vectors and matrices are defined. The most widely used are outlined below.

Zonation

Parameterization is accomplished by partitioning the domain in a set of subdomains (zones). Typically, each component of the vector of model parameters, p_i, is associated to one subdomain. Within each of them, properties $\mathbf{q}(\mathbf{x})$ are assumed constant or prescribed to vary in a predefined manner and the value of the interpolation function in Eq. (4) is zero if point **x** falls outside the zone being considered.

Often, time and space variability are decoupled (Carrera and Neuman 1986a), so that

$$q = (\mathbf{x}, \mathbf{t}) = p_i f_x(\mathbf{x}) f_t(\mathrm{t}) \quad (6)$$

where f_x is a spatial function (e.g., f_x may be aquifer thickness if p_i represents hydraulic conductivity of the zone and q is transmissivity) and f_t is a time function (e.g.

 DOI 10.1007/s10040-004-0404-7

f_t may represent time varying recharge obtained by mass balance in the soil and p_i would be an unknown factor if q is areal recharge).

The main advantage of zonation is its generality and flexibility to accommodate the geological information (e.g., zones may represent geological units or portions of them). It should be stressed, however, that zonation does not preclude the use of geostatistics. In fact, Clifton and Neuman (1982) used zonation coupled to kriging.

While the zones need not be large, the original spirit of zonation is to reduce the dimension of **p**, while ensuring geological consistency (Stallman 1956). In fact, Carrera et al. (1993b) argue that, when available, geological information about parameter variability is so compelling (in the sense that it can be included deterministically) that it overcomes the advantages of conventional geostatistics. Zonation is sometimes criticized as rigid. Hence, it is not surprising that efforts have been made to optimize the geometry of zones. A particularly appealing one is "geomorphing" where the geometry of zones is derived during the calibration process (Roggero and Hu 1998).

Point estimation

It can be viewed as the limiting case of zonation, as the size of zones tends to zero (actually, to the element or cell size). The formalism of Eq. (4) can still be used (e.g. Meier et al. 2001). However, the dimension of the parameter space becomes so large that it may be more appropriate to seek alternative formulations (Kitanidis and Vomvoris 1983; Dagan 1985; McLaughlin and Townley 1996). These will be outlined in the next section.

Heuristic interpolation functions

The interpolation functions α_i in Eq. (4) can be chosen arbitrarily. Different types have been chosen, including finite elements (Yeh and Yoon 1981), Ridge functions (Mantoglou 2003), or others. These approaches offer significant flexibility, but it is not clear how to define prior information on model parameters.

Pilot points

In this case, p represents the unknown values of the property q at a set of (pilot) points. The method, originally devised by de Marsily et al. (1984), has become very popular (e.g. Ramarao et al. 1995; Vesselinov et al. 2001; Hernandez et al. 2003), to the point of becoming the standard for non-linear geostatistical inversion. It can be considered as a generalization of Eq. (4) for the case in which measurements at the pilot points are uncertain. In this case, Eq. (4) can be written as:

$$q(\mathbf{x}) = \Sigma_{i=1}^{\dim(zp)} \lambda_i^{\mathrm{zp}}(\mathbf{x}) z_{\mathrm{pi}} + \Sigma_{j=1}^{N} \lambda_j^{\mathrm{PP}}(\mathbf{x}) p_j \qquad (7)$$

where λ_i^{zp} and λ_i^{PP} are the (co-) kriging weights for measurements and pilot points, respectively. Comparing Eqs. (7) and (4), it is clear that q_0 is the first term on the right hand side and $\alpha_i^q(\mathbf{x}, \mathbf{t})$ are equal to $\lambda_i^{\mathrm{PP}}(\mathbf{x})$. The main advantage of the method is its relative ease and flexibility. Surprisingly, it has not been until 2004 that the need to account for prior information on the parameters has been recognized. Traditionally, direct measurements were used for the first term in Eq. (7), but were disregarded during the inversion process. Doherty (2003) includes a regularization criterion penalizing non-homogeneity of the model parameters and he does not use prior information. Recently, Kowalsky et al. (2004) used a plausibility term for the first time in the context of the pilot points method, but the role of this term is not explored.

Conditional simulation

Methods described so far are implicitly based on seeking some sort of optimal estimation. As will be seen in subsequent sections, it is sometimes preferable to seek equally likely simulations of $q(\mathbf{x}, t)$. Neuman and Wierenga (2003) present a comprehensive strategy in the context of simulation. Alternatives that have been used include the self calibration approach of Sahuquillo et al. (1992), Gómez-Hernández et al. (1997) and Capilla et al. (1998), in which $q_0(x)$ represents a simulation of the log-T field conditioned by all (soft and hard) available information and the terms $\alpha_i(\mathbf{x}, t)p_i$ represent perturbations imposed by a set of master points. Another possibility is to express $q(x)$ as a linear combination of random functions (Roggero and Hu 1998; Hu 2002), where α_i's represent conditional simulations and the model parameters are simply weights of those simulations.

In summary, first, a representation of the variability of the hydraulic properties is necessary and, second, the most common parameterization schemes can be written using Eq. (4) or, in spite of their quite different appearance. The question now is how to find **p**.

How to estimate the model parameters, p

The estimation problem deals with the concept of "best" set of model parameters. A good question is what does "best" exactly mean? There is no perfect answer to this question. In fact, there may not be a single set of model parameters leading to a "good" representation of reality, which motivates conditional simulation methods Gómez-Hernández et al. (1997) that aim at finding a collection of equally probable parameters sets. Following is a description of the most widely used methods.

Optimization methods

In these methods, the parameter set is defined as the minimum of an objective function. Originally, the objective function was motivated toward ensuring a good match between computed and measured data (model fit). This agreement (performance criterion in this case) can be measured as the squared difference between computed and measured heads. This squared difference is, in mathematical terms, a norm (called L_2 norm). However, other norms are available and some have been used in groundwater. For example, errors can be quantified as the absolute value of the difference between measured and computed values (L_1 norm). Each norm has its own ad-

 DOI 10.1007/s10040-004-0404-7

vantages: the L_2 norm is more sensitive to outliers than L_1, which is more difficult to work with than the L_2 norm (see Woodbury et al. (1987) for details and Xiang et al. (1992) for an application to modeling). Defining a performance criterion leads naturally to a minimization problem, i.e., to seek the model parameters that make the chosen norm minimum. Using the L_2 norm, the objective function to be minimized can be written as

$$F_h = (\mathbf{h} - \mathbf{h}^*)^t \mathbf{C}_h^{-1} (\mathbf{h} - \mathbf{h}^*) \quad (8)$$

where $\mathbf{C}_h^{-1}$ is a matrix of weights. When F_h is used as objective function, the problem often becomes ill-posed. The solution becomes unstable, i.e. different parameter sets lead to very similar values of F_h. Sometimes, this is shown as extreme sensitivity to initial parameters, which prompts modelers to argue that the solution is non-unique. To overcome these problems, some researchers tend to set upper and lower bounds on estimated parameters. In itself, this does not solve the problem. The solution simply fluctuates between those (arbitrary) bounds, but its reliability is not improved. This prompted Neuman (1973) to add a plausibility term, F_p, to the objective function:

$$F = F_h + \lambda F_p \quad (9)$$

$$F_p = (\mathbf{p} - \mathbf{p}^*)^t \mathbf{C}_p^{-1} (\mathbf{p} - \mathbf{p}^*) \quad (10)$$

This equation was originally based on matching and stability arguments. The term F_p can be viewed as a regularization term in the sense of Tihonov (1963). In hydrology, Emselem and de Marsily (1971) used it to dampen oscillations. Also, Eq. (9) can be derived by statistical means in which case $\mathbf{C}_h$ is viewed as the covariance matrix of measurement errors. Gavalas et al. (1976) derived it by maximizing the posterior pdf of the model parameters (maximum a posteriori, MAP). The obtained objective function is equal to Eq. (9) with λ equal to 1. Carrera and Neuman (1986a) derived Eq. (9) by maximizing the likelihood of the parameters given the data (maximum likelihood estimation, MLE). The advantages of formulating Eq. (9) in a statistical framework lie in the fact that it yields ways to estimate not only the parameters controlling aquifer properties but also those controlling their uncertainty (variances, variogram and the like). As it turns out, the latter are no less important than the former (Zimmerman et al. 1998).

The minimization of the objective function is an arduous task because the relation between state variables and parameters is usually non-linear. This is why formulations of the inverse problem can be classified as linear and non-linear.

Linear methods

Regardless of the formulation, the problem can be linearized as:

$$\mathbf{h}(\mathbf{p}) = \mathbf{h}_0 + \mathbf{J}_{hp}(\mathbf{p} - \mathbf{p}^*) \quad (11)$$

where $\mathbf{J}_{hp}$ is the Jacobian (sensitivity matrix), containing the derivatives of $\mathbf{h}$ with respect to $\mathbf{p}$. Its computation is discussed later (sensitivity equations). Assuming this equation to hold, Eq. (2) can be generalized to obtain parameters as the conditional expectation of $\mathbf{p}$ given $\mathbf{h}$ (Dagan 1985; Rubin and Dagan 1987). For this purpose, one needs to define the cross-covariances between data and parameters $\mathbf{Q}_{ph}$ and the covariance matrix of data $\mathbf{Q}_{hh}$, given by:

$$\mathbf{Q}_{ph} = \mathbf{C}_{pp}\mathbf{J}_{hp}^t \quad (12)$$

$$\mathbf{Q}_{hh} = \mathbf{J}_{hp}\mathbf{C}_{pp}\mathbf{J}_{hp}^t + \mathbf{C}_h \quad (13)$$

Then

$$\hat{p} = E(p|\mathbf{p}^*, \mathbf{h}^*) = \mathbf{p}^* + \mathbf{Q}_{ph}\mathbf{Q}_{hh}^{-1}(\mathbf{h}^* - \mathbf{h}(\mathbf{p}^*)) \quad (14)$$

By extending matrices $\mathbf{Q}_{ph}$ and $\mathbf{Q}_{hh}^{-1}$ in a way similar to what was done in Eqs. (2), (14) can also be viewed as cokriging (Kitanidis and Vomvoris 1983; Hoeksema and Kitanidis 1984), which results from minimizing the variance of estimated parameters. Possibly, the most important thing is that Eq. (14) and its kriging variants do not rely explicitly on any geometrical parameterization scheme. Once $\mathbf{Q}_{hh}$ has been found, one can theoretically estimate $\mathbf{p}$ at any point. In fact, for Kitanidis and Vomvoris (1983), the only parameters to be estimated are the ones characterizing the statistical properties of the log-T field. In general, the covariance of measurement errors is also needed.

Non linear methods

One question that arises from Eq. (11) is why should the linearization be performed at the prior estimates of the parameters (linearization is made around the constant mean in the original papers). Actually, Carrera and Glorioso (1991) prove that it is much better to linearize around the estimated parameters themselves, which should be close to the true ones and may be far from the prior estimates. In practice, this is achieved by iteratively linearizing at the parameters obtained in the previous iteration. Carrera and Glorioso (1991) also showed that the result is identical to the one obtained by minimizing Eq. (9), and that Eq. (14) yields the same results as the first iteration of a Gauss-Newton method to minimize Eq. (9). This closes the loop. As it turns out, all the methods discussed previously and summarized in Table 2, are identical from the estimation viewpoint.

Differences among methods are restricted to the characterization of statistical parameters and to the evaluation of uncertainty. Without discussing how different methods deal with these issues, the next section is devoted to arguing that indeed they are important.

How codes work

In essence, most codes work by following an iterative process such as:

 DOI 10.1007/s10040-004-0404-7

Table 2 Summary of estimation methods. Notice that they are all quite similar despite their apparent differences. Linear methods look different, but Carrera and Glorioso (1991) showed that they can be viewed as the first iteration of non-linear methods. All methods are explicitly stated in the section, "How to estimate the model parameters, p"

Method	Type	Estimator	Algorithm (Eq. no.)	Reference
Least squares	Non linear	F_h	(8)	Cooley (1977) Hill et al. (1998)
Maximum Likelihood Estimation (MLE)	Non linear	$-2\ln P(\mathbf{p}\|\mathbf{h})$	(9)	Carrera and Neuman (1986a)
Conditional Expectation	Linear or Non linear	$E(\mathbf{p}\|\mathbf{h})$	(14)	Dagan (1985) Carrera and Glorioso (1991)
Cokriging	Linear or Non linear	Estimation variance	(14)	Kitanidis and Vomvoris (1983), Carrera et al. (1993a, b)
Maximum a posteriori (MAP)	Non linear	$P(\mathbf{p}\|\mathbf{h})$	(9) with λ=1[1]	Gavalas et al. (1976)

[1] Weighting factor of the regularization/plausibility term

Table 3 Comparison of minimization methods used in groundwater inverse modeling

Method	Comput. needs	Order of convergence	Convergence[1]	References
Gauss-Newton, Marquardt	F[2], $\mathbf{g}$[3], $\mathbf{J}$[4]	2	Local	Cooley (1977)
Conjugate gradients	F, $\mathbf{g}$	1	Local	Carrera and Neuman (1986b)
Simulated annealing, Genetic algorithms	F	0	Global	Rao et al. (2003) Tsai et al. (2003)

[1] Global means that the method may theoretically escape from local minima. Local means that it is difficult for the method to escape from a local minimum
[2] Objective function
[3] Gradient vector
[4] Jacobian matrix

Table 4 Synthetic comparison of methods to compute derivatives. Cost per Gauss-Newton iteration is estimated in terms of simulation runs for a hypothetical problem of 100 parameters and 20 observation points. In the case of direct nonlinear problems (i.e., nonlinear governing equations, e.g. unsaturated flow problem), it is assumed that 10 iterations are needed for the direct problem

Method	Advantages	Disadvantages	Cost per iteration	
			Linear problem	Nonlinear problem
Direct derivation	Exact	Hard to program	101[1]	110[1]
Adjoint state	Exact	Hard to program	21[1]	n/a
Finite differences	Easy programming	Not exact	101	1010

[1] The computing cost of exact derivation and adjoint methods can be reduced if one takes advantage of the fact that the system matrices are identical for all the parameters

1. Initialization: Read input data, set iteration counter i=0, initialise parameters, $\mathbf{p}^0$
2. Solve the simulation problem, $\mathbf{h}(\mathbf{p}^i)$, compute the objective function F^i, and possibly its gradient (assuming that it is continuously differentiable), $\mathbf{g}^i$ and the Jacobian matrix, $\mathbf{J}_{hp}$.
3. Compute an updating vector, $\mathbf{d}$, possibly using information on previous iterations, as well as $\mathbf{g}^i$ and $\mathbf{J}_{hp}$.
4. Update parameters, $\mathbf{p}^{i+1} = \mathbf{p}^i + \mathbf{d}$
5. If convergence has been reached, then stop. Otherwise, set i=i+1 and return to 2.

These steps are rather straightforward, except for the definition of the updating direction, $\mathbf{d}$, and the computation of $\mathbf{g}$ and $\mathbf{J}_{hp}$. The most frequent methods for these are outlined below.

Computing the updating step, d

Computation of $\mathbf{d}$ falls within the realm of optimization. Literature on numerical optimization methods is extensive, but only some of these have been applied to the solution of the groundwater inverse problem. They are summarized in Table 3. Cooley (1977) proposed using Marquardt's method (Marquardt 1963), which is based on minimizing a quadratic approximation of the objective function while limiting the size of the step to be taken at each iteration. This method is quite powerful, in the sense that it often converges in a small number of iterations, but expensive because it requires computing the Jacobian matrix. Carrera and Neuman (1986b) propose using a combination of quasi-Newton and conjugate gradient methods. These are not as robust as Marquardt's method but their cost per iteration is much smaller, as they only need to compute the gradient of the objective function. The choice of one or the other is problem dependent.

 DOI 10.1007/s10040-004-0404-7

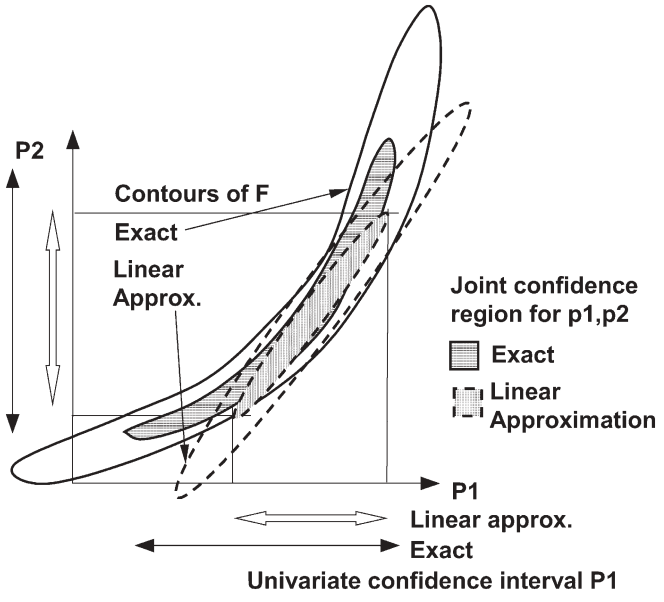

Fig. 2 Contour lines of the objective function in the parameters P1 and P2. The ellipses (dashed) represent the linear approximation implicit in Eq (3) and similar ones. Notice that the actual uncertainty is larger

Cooley (1985) compared these two families of methods and concluded that Marquardt method is superior (see also Hill 1990, for additional comparisons).

Most codes use these methods, which are generically termed descent methods, because they seek an improvement of the objective function at each iteration. As a result they tend to get stuck at local minima. Overcoming local minima is the motivation of many methods based on different variations of random searches such as simulated annealing (Rao et al. 2003) or genetic algorithms (Tsai et al. 2003). One method that has proven highly effective in surface water hydrology is the "Shuffled Complex Evolution" (Duan et al. 1992), which has not yet been tested in groundwater. These methods have not been widely used in practice both because their performance decreases when the number of parameters is large (say, larger than 20), which is frequent in aquifer modeling, and because they demand many evaluations of the objective function, which is rather expensive in groundwater models.

The complexity of the forward problem is relevant when comparing different optimization methods. For example, if the direct problem is non-linear, the cost of evaluating the objective function increases much more than the cost of computing the gradient or the Jacobian matrix, because the computation of these is non iterative. This tends to favor order of convergence 2 methods over either order 0 or 1 methods for non-linear problems.

Sensitivity equations

Sensitivities are the derivatives of state variables with respect to the model parameters. They are useful for two reasons. On the one hand, they can be employed in some of the above optimization methods. On the other hand, they yield useful information about the reliability of both the model and its estimated parameters. In addition, as they identify which data are most sensitive to which parameters, they can be used in the design of optimal networks (see section entitled, "How good is the model?"). In essence, sensitivities can be obtained in three ways: direct derivation, adjoint state and finite differences. A comparison of these methods can be found in Carrera et al. (1990a) and Carrera and Medina (1994).

Direct derivation is based simply on taking derivatives of the simulation equations with respect to **p**. In the steady state case of Fig. 1, this leads to

$$\mathrm{A}\frac{\partial \mathbf{h}}{\partial \mathbf{p}} = \frac{\partial \mathbf{b}}{\partial \mathbf{p}} - \frac{\partial \mathbf{A}}{\partial \mathbf{p}}\mathbf{h} \tag{15}$$

It should be noticed that Eq. (15) allows the computation of the derivatives of nodal heads with respect to the parameters, but they can be computed at every point in space using the finite elements method (FEM) interpolation functions. The derivatives of any variable that depends on heads can be also computed as e.g. flows.

Adjoint state method equations. This method is based on viewing the minimization of F as an optimization problem with respect to both **h** and **p**, while keeping the state equations (Fig. 1) as equality constraints. The adjoint state vector λ is the set of Lagrange multipliers of the joint optimization problem, which can be obtained from:

$$\lambda^t \mathbf{A} = -\mathbf{2}(\mathbf{h} - \mathbf{h}^*)^t \mathbf{C}_h^{-1} \tag{16}$$

The gradient of the objective function then becomes:

$$\frac{\partial F}{\partial \mathbf{p}} = \lambda^t \left(\frac{\partial \mathbf{A}}{\partial \mathbf{p}}\mathbf{h} - \frac{\partial \mathbf{b}}{\partial \mathbf{p}} \right) + \mathbf{2}\mathbf{C}_p^{-1}(\mathbf{p} - \mathbf{p}^*) \tag{17}$$

Adjoint state equations can also be used for computing the Jacobian matrix, being advantageous when the number of observation points is smaller than the number of parameters, or for the exact computation of the Hessian matrix (second order derivatives of the objective function with respect to model parameters; (see computation details in Carrera and Medina 1994; Medina and Carrera 2003).

Finite differences are based on approximating the derivatives by the incremental ratio.

$$\frac{\partial F}{\partial p} = \frac{F(p + \Delta p) - F(p - \Delta p)}{2\Delta p} + O(\Delta p^2)$$
$$= \frac{F(p + \Delta p)}{\Delta p} + O(\Delta p) \tag{18}$$

Finite differences can be used without added cost to compute the Jacobian matrix. Mehl and Hill (2003) compare the ways to perform this calculation.

The nature of the direct problem also affects the choice of the method to compute derivatives. If it is non linear (e.g. unsaturated flow), the computational effort of the direct problem is very large, making finite differences more expensive than exact methods (see Table 4). However, the ease of the finite difference method and its ability to work from outside a simulation code has caused

 DOI 10.1007/s10040-004-0404-7

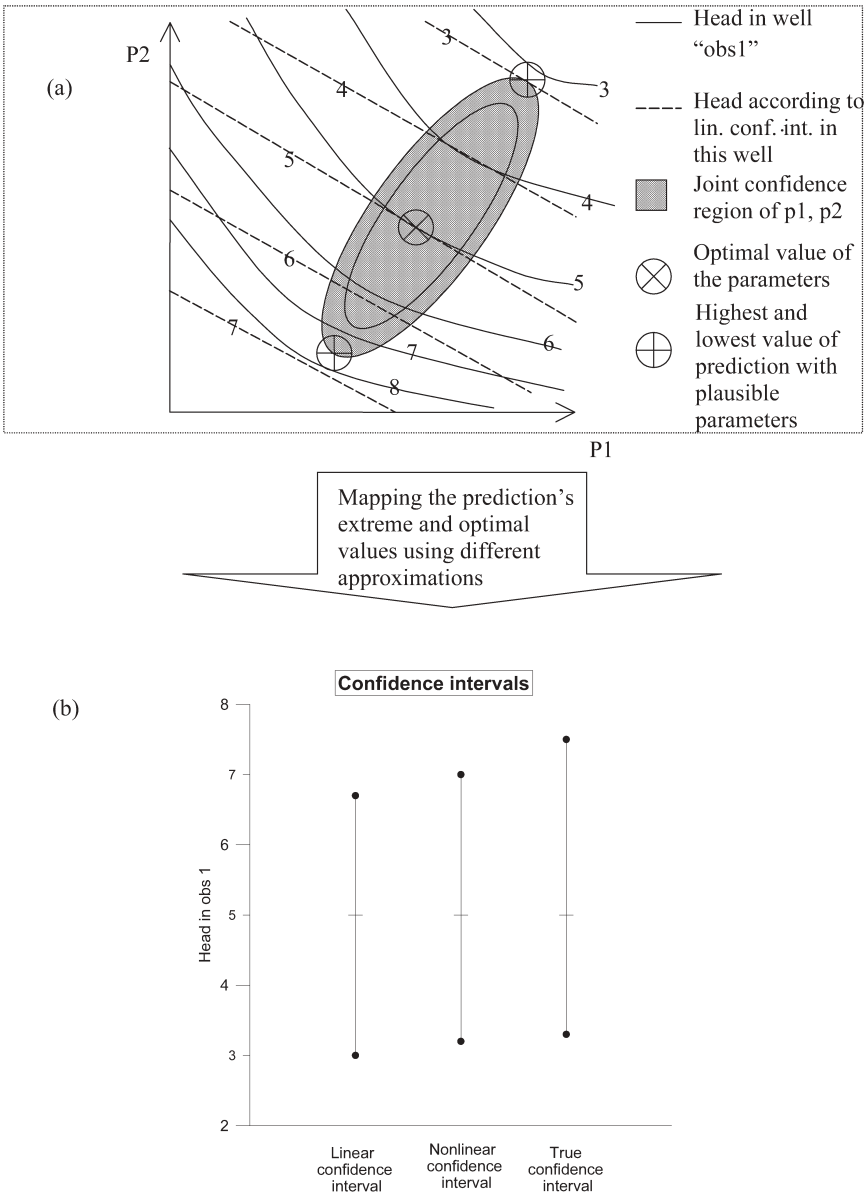

Fig. 3 **a** Contours of predicted value in the parameter space, according to the model and according to a linearization of the model. **b** Confidence interval (conf. int.) of prediction according to the model and according to a linearization of the model. A third type of confidence interval is shown, based on a nonlinear approximation of the model. The contours of this nonlinear approximation are not shown in (**a**) for clarity

an increase in the use of codes based on it, such as UCODE and PEST (Poeter and Hill 1998; Doherty et al. 2002).

How good is the model?

Following the modeling procedure outlined in Fig. 1, once the parameters have been estimated, it is necessary to assess how good are both the model concept and estimated parameters. Model quality is affected by three factors (Beck 1987): (1) uncertainty about the model structure, (2) uncertainty about the values of the parameters appearing in the model structure and (3) uncertainty associated with predictions of the future behavior of the system. The term uncertainty is used here to mean not only random fluctuations in errors, but also biases (see, e.g., Barth et al. 2001).

Uncertainty about the model structure

Conceptual models have many uncertain features because data are never exhaustive and contain inconsistencies. Modelers are forced to make simplifying assumptions. Errors are introduced in the parameterization, in the discretization, in the selection of boundary conditions, etc. Addressing this uncertainty often requires posing several conceptual models. Different models can be compared in terms such as model fit Eq. (8), residual distribution, parameter correlation, and confidence intervals for parameters and predictions. Ideally, a good model should lead to a good match with observations, uncorrelated residuals, and reasonable parameter values. Still, several models may fit all these criteria and one may need to select one among them. Several model selection criteria have been defined from the field of time series analysis and applied to groundwater (Akaike 1974, 1977; Rissanen

 DOI 10.1007/s10040-004-0404-7

1978; Schwarz 1978; Hannan 1980; Kashyap 1982). Carrera and Neuman (1986c) applied four of these criteria to a synthetic test case and concluded that the Kashyap criterion was the best. Similar results were obtained by Carrera et al. (1990b) and by Medina and Carrera (1996).

Still, one might question the wisdom of selecting only one model. This implies rejecting the others, which may not be logical if they are consistent with available knowledge and data. This line of argument leads to accepting a large set of models and using them all to characterize uncertainty in model predictions (Beven and Binley 1992; Beven and Freer 2001)

Difficulties associated with the calibration process

The most important difficulties associated with the optimization are the problems of non-uniqueness, non-identifiability and instability. Non identifiability occurs when more than one set of parameters leads to a given solution of the forward problem. Non-uniqueness is present when more than one set of parameters leads to minima of the objective function. Instability is present when small changes in the observations lead to large changes in the estimated parameters, but is usually identified by a dependence of the solution on the initial parameters. In addition, optimization algorithms such as Marquardt and conjugate gradients may get stuck into a local minimum, a set of model parameters whose performance cannot be improved by small changes in data, but which does not represent the overall optimum. Carrera and Neuman (1986b) discuss extensively these concepts and show that they are closely related. In fact, they can all be characterized by the region of uncertainty, even if it is linearized (Fig. 2).

Instability and large uncertainty are different concepts, but both are often associated to elongated confidence regions, which can be characterized by the eigenvalues and eigenvectors of the posterior covariance matrix. The eigenvectors of this matrix form a set of orthogonal vectors, each of which is associated to an eigenvalue. The vector associated to the largest eigenvalue represents the linear combination of parameters that has the largest uncertainty, whereas the vector with the smallest eigenvalue defines the direction with least uncertainty. If eigenvalues are dramatically different, then one should expect instabilities to occur. One of the effects is parameter correlation. The parameters in Fig. 2 suffer from correlation: a shift in the estimation of one parameter (say, P1) causes a shift in the optimal value of the other (P2). In other words, the estimated parameters are not independent. This dependence causes the confidence intervals of the parameters to be larger than they would have been if they were independent.

To summarize the above, instabilities and large uncertainties, though frequent, can be easily identified and characterized. Tactics to combat these problems include:

- *Regularization*: This is what motivated adding F_p to the objective function in (9). Weiss and Smith (1998) comment on which prior information will most effectively reduce uncertainty.
- *Reducing the number of parameters*: This is what motivates the parameterization schemes discussed in the section entitled, "What is to be estimated?".
- *Increasing the number and types of data*, which was discussed in the section, "The conceptual model: Knowledge and data".
- *Optimizing the observation scheme:* Observation networks and experiments can be designed to minimize model uncertainty and/or to increase the ability of data to discriminate among alternative models (Knopman and Voss 1989; Usunoff et al. 1992)

Despite this, one may have to acknowledge that it is not possible to find a unique solution to the problem. This motivates some researchers to use stochastic simulation of parameter fields conditional to data, rather than estimation (e.g. Gómez-Hernández et al. 1997). These techniques generate large numbers of e.g. transmissivity fields that satisfy the available head and transmissivity data. In this way, one ends up with a number of models, rather than one. Uncertainty is associated with the ensemble set of all simulations, rather than to statistical measures of uncertainty. As another alternative, Yapo et al. (1998) study instability using a concept known as the pareto optimum, which denotes the set of parameter vectors for which improving one component of the objective function causes a deterioration in another component.

Difficulties associated with predictions of the future behavior of the system

All the above difficulties cause the model parameters to have some uncertainty, which is inherited by the model predictions. The possibility of multiple conceptual models causes further uncertainty in the model predictions. Therefore, evaluating uncertainty in prediction requires analyzing the effect of both parameter and model uncertainty. The latter is usually analyzed by simulating with the models available and evaluating the range of predictions (e.g., Medina and Carrera 1996). While systematization is needed, the fact that conceptual model building is not systematic makes this objective hard to meet. Hence, what follows concentrates on evaluating the effect of parameter uncertainty.

There exist various methods for quantifying the prediction uncertainty: Linear approximation, non-linear approximation and Monte Carlo methods. They are outlined in Fig. 3. The linear method is relatively simple. It is based on the a posteriori covariance matrix:

$$\Sigma_p = \mathbf{J}_{hp}^t \mathbf{C}_h^{-1} \mathbf{J}_{hp} + \mathbf{C}_p^{-1} \tag{19}$$

where $\mathbf{J}_{\mathrm{hp}}$ is the Jacobian matrix, $\mathbf{C}_h$ the measurements covariance matrix and $\mathbf{C}_p$ the parameters covariance matrix.

If f is a prediction to be made with the model (f is a function of parameters, p), a lower bound of its variance is given by:

 DOI 10.1007/s10040-004-0404-7

$$\mathrm{Var}(f) = \left(\frac{\partial f}{\partial \mathbf{p}}\right)^t \Sigma_p \left(\frac{\partial f}{\partial \mathbf{p}}\right) + \sigma_{p0}^2 \quad (20)$$

where σ_{p0}^2 represents model errors independent of parameter uncertainties. This equation represents a linear approximation, so that actual uncertainty may be larger than that represented. This is why more sophisticated approaches may be needed. Still, this equation allows one to spell out quantitatively how the different factors affect prediction uncertainty. In essence, prediction uncertainty grows with:

1) Sensitivity of predictions to model parameters and initial conditions. Obviously, a parameter is a source of concern only if predictions are sensitive to it, as measured by $\partial f/\partial \mathbf{p}$. This is why sensitivity analyses are sometimes performed instead of a formal error analysis.
2) Uncertainty in model parameters (and/or initial state). This is measured by the covariance matrix Σ_p.

Non linear methods are more difficult to apply. Vecchia and Cooley (1987) present a way to compute nonlinear confidence intervals. They conclude that (i) corresponding linear and nonlinear confidence intervals are often offset or shifted towards each other, and the nonlinear ones are often larger, (ii) the variability in sizes of nonlinear confidence intervals is usually larger than the corresponding variability in linear confidence interval sizes, (iii) the difference between the sizes of linear and nonlinear confidence intervals increases as the sizes of the intervals increase, and (iv) prior information can alter the size of the confidence intervals. Christensen and Cooley (1999) present a measure to quantify model non-linearity. The prediction analyzer of PEST currently includes nonlinear confidence intervals.

The Monte Carlo method is the most computationally intensive method. It is based on many forward problem evaluations with different sets of parameters. Its main advantages are that it is easy to understand, it yields a probability density function and does not require difficult assumptions. Its main problem is that it is hard to ascertain the required number of simulations. An example of this is the GLUE methodology (Generalized Likelihood Uncertainty Estimation; Beven and Freer 2001).

What is actually done?

Application trends

The literature in scientific journals tends to concentrate on the development of new methods and new interpretations. Therefore it is not appropriate for identifying application trends. Applications are most often found in internal reports or special sessions of congresses. None of these are easy to track exhaustively. Therefore, this section is biased by what can effectively be found and by the authors' personal views. Despite the above, a search on the "web of science" (http://go5.isiknowledge.com/portal.cgi), which keeps track of all papers published in major journals, was performed. Results suggest that the number of papers about inverse modeling remains more or less steady (around 5% of those about groundwater modeling), while the number of papers using inverse modeling has slowly but steadily increased in the last 13 years. These papers cover a broad range of topics. Many have been cited in previous sections and some will be cited in the remaining sections. It is difficult, however, to identify trends. The only ones that emerge is that geostatistical inversion tends to be used in relatively small scale problems, such as the interpretation of hydraulic tests, while large scale models tend to be based on zonation. Since most groundwater applications can be classified in one of these two categories, they deserve further attention and are discussed below.

Geostatistical inversion

The use of geostatistics is motivated by the need to address the variability of hydraulic properties (specifically transmissivity) when modeling aquifers. In practice, however, this need can be understood in two different ways: (1) to constrain model parameters and (2) because it is deemed necessary for model predictions. While these two views are not exclusive, they rarely go together.

Methodologically, geostatistical inversion follows the steps originally proposed by Clifton and Neuman (1982). That is, one starts by proposing a stochastic model (i.e., whether the log-T is stationary, what is its variogram and mean, etc). Second, available data is used to produce a prior, best estimate of log-T and its covariance, using Eqs. (2) and (3). Finally, these are used to obtain an estimate of model parameters either by minimizing Eq. (9) or using Eq. (14). Kitanidis and Vomvoris (1983) and Carrera and Neuman (1986a) modified this concept by emphasizing the need for optimal estimation of statistical parameters (variances of errors, correlation distance and the like). The importance of these parameters has been recognized by many. In fact, this is one of the conclusions of Zimmerman et al. (1998), after comparing different geostatistical inversion techniques.

Successful applications to real field data are restricted to relatively small-scale problems. These include hydraulic test interpretation (Yeh and Liu 2000; Meier et al. 2001; Vesselinov et al. 2001), well capture zone delineation (Vassolo et al. 1998; Kunstmann et al. 2002; Harrar et al. 2003), and others (Barlebo et al. 2004). The fact that applications to large-scale aquifers are scarce reflects the difficulty in modeling them, but also points out the two limitations of geostatistics as it is most frequently used: it fails to include geological information and it fails to reproduce actual variability.

Geological information is normally expressed in terms that are difficult to account for during inversion (depositional patterns, orientation of conductive features, and the like). This type of information can become precise at the large scale but rarely at the test scale. Here, the only thing that can be said is that permeability is variable,

 DOI 10.1007/s10040-004-0404-7

which is properly acknowledged by geostatistics. Still, when discrete features are identified, inversion gains significantly by explicitly including them as deterministic features. This was acknowledged by Meier et al. (2001), who used information about the stress state of a shear zone to define the anisotropy direction of fracture transmissivity.

Ironically, the main problem with geostatistical estimation is the fact that it fails to reproduce actual variability. As discussed in the section, "How to estimate the model parameters, p", Eqs. (9) and (14) yield the conditional expectation of log-T given the available data. As such, the estimation says little about actual deviations because the expected value filters them out. When data are abundant, which is frequent in detailed hydraulic tests, variability patterns can be delineated with some certainty (e.g. Meier et al. 2001). Otherwise, only the estimation covariance provides some information about variability and about the continuity of high conductivity zones (paths for solutes migration) or low conductivity barriers. Another source of hope for large-scale geostatistical inversion is the use of geophysical data (Hubbard and Rubin 2000).

Gómez-Hernández and coworkers (Gómez-Hernández et al. 1997; Capilla et al. 1998) address the above issue by rejecting optimal estimation altogether and, instead, performing conditional simulations. Ideally, the average of all these simulations should be equal to the conditional estimation, but each of them reproduces the assumed variability. As such, when used for predicting processes that are sensitive to variability (e.g., contaminant transport), simulations are much more appropriate.

Geologically based inversion

The main difference between aquifer scale models and test scale models is the degree of reliance on geology when defining variability and the multiplicity of parameter types. Regarding variability, geological data are rarely precise but cannot be ignored. Different formations, or different units within a formation, may have different properties. Thus, when these units can be outlined, the worth of this information cannot be ignored. Unfortunately, boundaries between units are rarely known accurately. Hence, a lot of work may be needed to test different geometries whenever the effect of geometry is found to be important. The process involves defining geometry and testing it against available data, which is repeated until a satisfactory fit is found. The procedure is tedious, it involves interactions between modelers and geologists, and it is not systematic. In fact, it does not get properly documented, so that if the model is revised years later, one is not sure about the reason behind the selected model structure. Things are made worse by the fact that not only transmissivity but also other types of parameters need to be specified. In the experience of the authors of this paper, ambiguities frequently exist about recharge (both average amount and time variability), boundary fluxes, pumping rates (the main finding of Castro et al. 1999, was that official pumping rates were badly underestimated) and the nature of river-aquifer interaction.

The somewhat chaotic nature of the process makes it hard to describe in detail. Attempts have been made by Cooley et al. (1986) and Hill (1998). Some trends can be identified:

- Point values of hydraulic conductivity and transmissivity are prone to error. Moreover they may be of little use when modeling at scales much larger than the pumping test in which they are based. These measurements need to be put into the related geological context.
- Dominant features (i.e. conductive fractured zones, paleochannels, or the like) must be included in the model even if they are not known accurately.
- Much information about aquifer behavior is contained in discrete events (floods, big rainfalls). Taking full advantage of these requires transient simulations.
- Model calibration is rarely unique (i.e., different model structures may fit hard data satisfactorily). This uncertainty ought to be acknowledged when performing model predictions. Reducing it often requires the use of complementary data as discussed in the section, "The conceptual model: Knowledge and data".

This kind of approach displays several drawbacks. On one the hand, it does not account properly for uncertainty. The resulting covariance matrix of model parameters is conditioned not only on hard data, but also on the many subjective decisions the modeler has made. While these can be taken into account by making predictions with several conceptual models, it is rarely done. On the other hand, one is never sure about the validity of the model beyond calibration conditions. In summary, the procedure needs to be systematized. This is best done in a geostatistical framework, hence the need to seek geostatistical descriptions that take advantage of qualitative geology data.

What comes next?

The discussion in the previous section makes it clear that it is believed that the time is ripe for standard use of inverse modeling in groundwater studies aimed at aquifer characterization and management, a view which we share with Poeter and Hill (1997).

In the 1970s, the U.S. Geological Survey promoted the use of numerical modeling for aquifer studies. This led to a significant rise in the quality of understanding of groundwater flow by many hydrologists. Numerical modeling forced them to be quantitatively consistent when integrating different data types. Since this is never straightforward, hydrologists were forced to test different parameter values, to perform sensitivity analyses, and to guess at what could be the cause behind observed data. In the end, they might not be fully successful, but they

 DOI 10.1007/s10040-004-0404-7

gained understanding, which is what counts for proper decision making.

The situation is now changing on several accounts. For one, hydrologists are increasingly expected to make hard decisions for which qualitative understanding is not sufficient. Instead, accurate quantitative models are needed. Second, the volume of data is also increasing. Long data sets of heads, pumping history, and hydrogeochemical measurements are becoming available. While they contain valuable qualitative information, it is clear that much more can be gained by using them quantitatively, i.e., by building models that can match those data. Third, modeling exercises such as INTRAVAL (Larsson 1992) have made it clear that the most important issue when modeling an aquifer is the conceptual model. Since manual calibration is very tedious, modellers have not been able to concentrate sufficiently on conceptual issues. Automatic calibration should change that.

In short, it is contended that conceptually sound and quantitatively consistent modeling requires automatic inversion. This allows the modeler to concentrate on the qualitatively important issues, such as the definition of the conceptual model, representation of spatial variability, discussion of alternative management strategies, etc. Also, this allows improved model building and use. Therefore, it should be one of the challenges in the future of hydrogeology. As it is, however, inverse modeling has not reached a desired level of maturity. A number of difficulties must be overcome. Some are listed below:

- *Difficulty in running inversion codes*. Since one has to introduce data on parameterization, observations, their reliability, etc., general purpose inversion codes are cumbersome and hard to run. The emergence of codes such as PEST or UCODE has changed that trend significantly, but places an additional burden on the modeler by requiring him or her to know programming and the inside of codes. This could be alleviated by developing 'utility programs' to enable the interfacing of the inversion engine with the forward model.
- *Incorporation of geological information*. As discussed earlier, zonation is the most convenient and widely used approach for realistic incorporation of geological data available in the form of maps. This is too rigid and incomplete. Geological information is usually "soft", in the sense that boundaries between formations underground are rarely known accurately. Moreover, important features (paleochannels, water conducting faults, etc.) may have gone unmapped. Furthermore, even within a formation, there may be a lot of qualitative information (depositional patterns, continuity of water conducting features, gradation of materials and the like) that is difficult to incorporate in a zonation framework. For solute transport problems, these sources of small scale variability may be just as important as large scale trends. Presumably, they can be best handled in a geostatistical framework. While some approaches are available for incorporating them (i.e., treat each zone geostatistically, use geological maps as soft information), existing codes lack sufficient robustness and flexibility to be of general use.
- *Incorporation of age, environmental isotopic data, temperature and other sources of information*. As discussed, using different types of data improves dramatically the stability of inversion and the robustness of the model. In fact, a number of authors have shown that incorporating these data does improve the reliability of the model. The problem, again, lies in the availability of easy-to-run, flexible codes.
- *Representation of uncertainties*. It has also been discussed that dealing with uncertainty is an integral part of modeling. This is true both at the characterization stage, where data contain errors and model structure is never accurately known, and at the prediction stage. Well informed decisionmaking can not be based on single model predictions. Acknowledging uncertainties in both model concept and parameters is required. As discussed, linear estimates of uncertainty are very poor (they may underestimate actual errors in orders of magnitude; Carrera and Glorioso 1991). Alternative, non-linear estimates of error are difficult to use for hydrologists who are not familiar with statistics, even though tools are becoming increasingly user-friendly. Therefore, Monte Carlo simulation remains the most appealing. An advantage is the relative ease to accommodate conceptual model uncertainty. Still, at present, Monte Carlo methods are extremely expensive and again, not easy to use.
- *Coupling to GIS*. One of the drawbacks of inverse modeling is the need to incorporate all causes of variability; geology, soil properties and use, pumping, etc. If any of them is missed (e.g., if the effects of a pumping well are not incorporated), the algorithm will react by modifying the parameters to circumvent the effect of the error so as to best fit the measurements (e.g., reduce recharge, increase transmissivity, etc.). The resulting model is thus erroneous not only because of the missing factor, but also because of those modifications. Conventional modeling is less sensitive to this kind of error (one may be aware that model calculations may be erroneous in areas affected by unknown factors, but that does not affect the rest of the model). Therefore, inverse modeling requires careful accounting. As the history of aquifers becomes increasingly complex (growing number of pumping wells, evolving soil uses, improved knowledge of aquifer geometry), so does the difficulty to incorporate all factors. The tendency to incorporate all territorial data in GIS sheds some hope on the possibility of incorporating all these data in a model. In fact, there have been a number of efforts to link GIS to traditional models (Gogu et al. 2001; Chen et al. 2002). It is clear, however, that this need is more pressing for inverse models. Efforts along the line of Graphical User Interfaces are starting.

All the above issues might suggest that inverse modeling is not yet mature. It should be stressed, however, that all

 DOI 10.1007/s10040-004-0404-7

the above-mentioned difficulties are just that, difficulties. They can be, and indeed are, overcome in practice through laborious work. Moreover, these difficulties are not specific of inverse modeling; they are mostly shared by conventional modeling (although they are often ignored). Finally, while it is true that further improvement of programs is needed, it is not clear that their applicability is more difficult than it was for the use of early codes when the U.S. Geological Survey made modeling a routine for aquifer studies. Routine application of inverse modeling is the future. The sooner it starts, the better prepared will hydrologists be to face the challenges of the soon-to-come future.

Acknowledgments The final manuscript benefitted from comments by Mary Hill, Matt Tonkin and Johan Valstar

References

Akaike H (1974) A new look at statistical model identification. IEEE Trans Automat Contr AC-19:716–722

Akaike H (1977) On entropy maximization principle. In: Krishnaiah PR (ed) Applications of statistics. North Holland, Ámsterdam, pp 27–41

Anderman ER, Hill MC (1999) A new multi-stage ground-water transport inverse method, Presentation, evaluation, and implications. Water Resour Res 35(4):1053–1063

Barlebo HC, Hill MC, Rosbjerg D (2004) Identification of groundwater parameters at Columbus, Mississippi, using three-dimensional inverse flow and transport model. Water Resour Res 40(4):W0421110

Barth GR, Hill MC, Illangasekare TH, Rajaram H (2001) Predictive modeling of flow and transport in a two-dimensional intermediate-scale, heterogeneous porous media. Water Resour Res 37(10):2503–2512

Beck MB (1987) Water quality modelling: a review of the analysis of uncertainty. Water Resour Res 23(8):1393–1442

Bennet RR, Meyer RR (1952) Geology and groundwater resources of the Baltimore area. Mines and Water Resour Bull 4, Maryland Dept of Geology

Beven K (1993) Prophecy, Reality and uncertainty in distributed hydrological modeling. Adv Water Resour 16(1):41–51

Beven KJ, Binley AM (1992) The future of distributed models: model calibration and uncertainity prediction. Hydrol Process 6(3):279–298

Beven KJ, Freer J (2001) Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. J Hydrol 249:11–29

Bredehoeft J (2004) Modeling: the conceptualization problem-surprise. Hydrogeol J (this issue)

Capilla JE, Gómez-Hernández JJ, Sahuquillo A (1998) Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric head data—3. Application to the Culebra formation at the waste isolation pilot plan (WIPP), New Mexico, USA. J Hydrol 207(3–4):254–269

Carrera J (1987) State of the art of the inverse problem applied to the flow and solute transport problems. In: Groundwater flow and quality modeling, NATO ASI Ser: 549–585

Carrera J, Neuman SP (1986a) Estimation of aquifer parameters under transient and steady-state conditions, 1. Maximum likelihood method incorporating prior information. Water Resour Res 22(2):199–210

Carrera J, Neuman SP (1986b) Estimation of aquifer parameters under transient and steady-state conditions, 2. Uniqueness, stability and solution algorithms. Water Resour Res 22(2):211–227

Carrera J, Neuman SP (1986c) Estimation of aquifer parameters under transient and steady-state conditions, 3. Application to synthetic and field data. Water Resour Res 22(2):228–242

Carrera J, Navarrina F, Vives L, Heredia J, Medina A (1990a) Computational aspects of the inverse problem. In Proc. of VIII international conference on computational methods in water resources. CMP, pp 513–523

Carrera J, Heredia J, Vomvoris S, Hufschmied P (1990b) Fracture Flow Modelling: Application of automatic calibration techniques to a small fractured Monzonitic Gneiss Block. In: Neuman N (ed) Proc hydrogeology of low permeability environments, IAHPV, Hydrogeology, selected papers, vol 2, pp 115–167

Carrera J, Glorioso L (1991) On Geostatistical Formulations of the Groundwater Flow Inverse Problem. Adv Water Resour 14(5):273–283

Carrera J, Medina A, Galarza G (1993a) Groundwater inverse problem. Discussion on geostatistical formulations and validation. Hydrogéologie (4):313–324

Carrera J, Mousavi SF, Usunoff E, Sanchez-Vila X, Galarza G (1993b) A discussion on validation of hydrogeological models. Reliability Eng Syst Saf 42:201–216

Carrera J, Medina A (1994) An improved form of adjoint-state equations for transient problems. In: Peters, Wittum, Herrling, Meissner, Brebbia, Grau, Pinder (eds) Proc X international conference on methods in water resources, pp 199–206

Castro A, Vazquez-Suñe E, Carrera J, Jaen M, Salvany JM (1999) Calibración del modelo regional de flujo subterráneo en la zona de Aznalcóllar, España: ajuste de las extracciones [Calibration of the groundwater flow regional model in the Aznalcollar site, Spain: extractions fit]. In Tineo A (ed) Hidrología Subterránea. II,13. Congreso Argentino de Hidrogeología y IV Seminario Hispano Argentino sobre temas actuales de la hidrogeologia

Chen CX, Pei SP, Jiao JJ (2003) Land subsidence caused by groundwater exploitation in Suzhou City, China. Hydrogeol J 11(2):275–287

Chen Z, Huang GH, Chakma A, Li J (2002) Application of a GIS-based modeling system for effective management of petroleum-contaminated sites. Env Eng Sci 9(5):291–303

Christensen S, Cooley RL (1999) Evaluation of confidence intervals for a steady state leaky aquifer model. Adv Water Resour 22(8):807–817

Clifton PM, Neuman SP (1982) Effects of kriging and inverse modeling on conditional simulation of the Avra valley aquifer in southern Arizona. Wat Resour Res 18(4):1215–1234

Cooley RL (1977) A method of estimating parameters and assessing reliability for models of steady state groundwater flow, 1, Theory and numerical properties. Water Resour Res 13(2):318–324

Cooley RL (1985) A comparison of several methods of solving nonlinear-regression groundwater-flow problems. Water Resour Res 21(10):1525–1538

Cooley RL, Konikow LF, Naff RL (1986) Nonlinear regression groundwater-flow modeling of a deep regional aquifer system. Water Resour Res 22(13):1759–1778

Dagan G (1985) Stochastic modeling of groundwater flow by unconditional and conditional probabilities: the inverse problem. Water Resour Res 21(1):65–72

de Marsily GH, Lavedan G, Boucher M, Fasanino G (1984) Interpretation of interference tests in a well field using geostatistical techniques to fit the permeability distribution in a reservoir model. In: Verly et al (ed) Proc Geostatistics for natural resources characterization. Part 2. D. Reidel Pub. Co. : pp 831–849

de Marsily G, Delhomme JP, Delay F, Buoro A (1999) 40 years of inverse problems in hydrogeology. Comptes Rendus de l'Academie des Sciences Series IIA. Earth and Planet Sci 329(2):73–87. Elsevier Science

Doherty J, Brebber L, Whyte P (2002) PEST-Modelling dependent parameter estimation. Water Mark Computing. Corinda (Australia)

Doherty J (2003) Groundwater model calibration using pilot points and regularization. Ground Water 41(2):170–177
Duan QY, Sorooshian S, Gupta V (1992) Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour Res 28(4):1015–1031
Emselem Y, de Marsily G (1971) An automatic solution for the inverse problem. Wat Resour Res 7(5):1264–1283
Gavalas GR, Shaw PC, Seinfeld JH, (1976) Reservoir history matching by Bayesian estimation. Soc Petrol Eng J 261:337–350
Gogu RC, Carabin G, Hallet V, Peters V, Dassargues A (2001) GIS-based hydrogeological databases and groundwater modeling. Hydrogeol J 9(6):555–569
Gómez-Hernández JJ, Sahuquillo A, Capilla JE (1997) Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data. 1. Theory. J Hydrol 204(1–4):162–174
Gómez-Hernández JJ, Wen XH (1998) To be or not to be multi-Gaussian? A reflection on stochastic hydrogeology. Adv Water Resour 21(1):47–61
Gupta HV, Bastidas LA, Sorooshian S, Shuttleworth WJ, Yang ZL (1999) Parameter estimation of a land surface scheme using multicriteria methods. J Geophys Res-Atmos 104(D16):19491–19503
Hadamard J (1902) Sur les problemes aux derivees partielles et leur signification physique. [On the problems about partial derivatives and their physical significance]. Bull Univ Princeton 13:49–52
Hannan ES (1980) The estimation of the order of an ARMA process. Ann Stat (8):1071–1081
Harrar WG, Sonnenborg TO, Henriksen HJ (2003) Capture zone, travel time, and solute-transport predictions using inverse modeling and different geological models. Hydrogeol J 11(5):536–548
Hernandez AF, Neuman SP, Guadagnini A, Carrera J, (2003) Conditioning mean steady state flow on hydraulic head and conductivity through geostatistical inversion. Stochas Env Res Risk Assess 17(5):329–338
Hill MC (1990) Relative efficiency of four parameter-estimation methods in steady-state and transient ground-water flow models. In: Gambolati, Rinaldo, Brebbia, Gray, Pinder (eds) Proc Computational Methods in Subsurface Hydrology, International Conference on Computational Methods in Water Resources, pp 103–108
Hill MC (1992) A computer program (MODFLOWP) for estimating parameters of a transient, three-dimensional, ground-water flow model using nonlinear regression. U.S. Geological Survey
Hill MC (1998) Methods and guidelines for effective model calibration. US geological survey. Water-Resour Investigat Rep 98–4005, 91 pp, Colorado
Hill MC, Cooley RL, Pollock DW (1998) A controlled experiment in ground water flow model calibration. Ground Water 36(3):520–535
Hoeksema RJ, Kitanidis PK (1984) Comparison of Gaussian conditional mean and kriging estimation in the geostatistical solution to the inverse problem. Water Resour Res 21(6):337–350
Hollenbeck KJ, Jensen KH. (1998) Maximum-likelihood estimation of unsaturated hydraulic parameters. J Hydrol 210(1–4):192–205
Hu LY (2002) Combination of Dependent Realizations within the gradual deformation method. Math Geol 34(8):953–963
Hubbard S, Rubin Y (2000) A review of selected estimation techniques using geophysical data. J Contamin Hydrol 45(2000):3–34
Kashyap RL (1982) Optimal choice of AR and MA parts in autoregressive moving average models. IEEE Trans Pattern Anal Mach Intel PAMI 4(2):99–104
Kitanidis PK, Vomvoris EG (1983) A geostatistical approach to the inverse problem in groundwater modelling (steady state) and one dimensional simulations. Water Resour Res 19(3):677–690
Kitanidis PK (1997) Introduction to geostatistics: applications to hydrogeology. Cambridge University Press, Cambridge, NY
Knopman DS, Voss CI (1989) Multiobjective sampling design for parameter-estimation and model discrimination in groundwater solute transport. Water Resour Res 25(10):2245–2258
Kool JB, Parker JC, Van Genuchten MT (1987) Parameter estimation for unsaturated flow and transport models. A Review J Hydrol 91:255–293
Kool JB, Parker JC (1988) Analysis of the inverse problem for transient unsaturated flow. Water Resour Res 24(6):817–830
Kowalsky MB, Finsterle S, Rubin Y (2004) Estimating flow parameter distributions using ground-penetrating radar and hydrological measurements during transient flow in the vadose zone. Adv Water Resour 27:583–599
Kunstmann H, Kinzelbach W, Siegfried T (2002) Conditional first-order second-moment method and its application to the quantification of uncertainty in groundwater modeling. Water Resour Res 38(4):Art. No. 1035
Larocque M, Banton O, Ackerer P, Razack M (1999) Determining karst transmissivities with inverse modeling and an equivalent porous media. Ground Water 37(6):897–903
Larsson A (1992) The International Projects INTRACOIN, HYDROCOIN and INTRAVAL. Adv Water Resour 15(1):85–87
Mantoglou A (2003) Estimation of Heterogeneous Aquifer Parameters from Piezometric Head Data using Ridge Functions and Neural Networks. Stochas Environmen Risk Assess 17:339–352
Marquardt DW (1963) An algorithm for least-squares estimation of non-linear parameters. J Soc Indust Appl Math 11(2)
McLaughlin D, Townley LLR (1996) A reassessment of the groundwater inverse problem. Water Resour Res 32(5):1131–1161
Medina A, Carrera J (1996) Coupled estimation of flow and solute transport parameters. Water Resour Res 32(10):3063–3076
Medina A, Carrera J (2003) Geostatistical inversion of coupled problems: dealing with computational burden and different types of data. J Hydrol 281:251–264
Mehl SW, Hill MC (2003) Locally refined block-centered finite-difference groundwater models. In: Kovar K, Zbynek H (eds) Evaluation of parameter sensitivity and the consequences for inverse modelling and predictions. IAHS Publication 277, p. 227–232
Meier P, Carrera J, Sanchez-Vila X (1999) A numerical study on the relation between transmissivity and specific capacity in heterogeneous aquifers. Ground Water 37(4):611–617
Meier P, Medina A, Carrera J (2001) Geoestatistical inversion of cross-hole pumping tests for identifyingpreferential flow channels within a shear zone. Ground Water 39(1):10–17
Meixner T, Gupta HV, Bastidas LA, Bales RC (1999) Sensitivity analysis using mass flux and concentration. Hydrol Proc 13(14–15):2233–2244
Nelson RW (1960) In place measurement of permeability in heterogeneous media, 1. Theory of a proposed method. J Geophys Res 65(6):1753–1760
Nelson RW (1961) In place measurement of permeability in heterogeneous media, 2. Experimental and computational considerations. J Geophys Res 66:2469–2478
Neuman SP (1973) Calibration of distributed parameter groundwater flow models viewed as a multiple-objective decision process under uncertainty. Water Resour Res 9(4):1006–1021
Neuman SP, Wierenga PJ (2003) A comprehensive strategy of hydrogeologic modeling and uncertainty analysis for nuclear facilities and sites. NU-REG/CR-6805, US Nuclear Regulatory Commision, Washington, DC
Poeter EP, Hill MC (1997) Inverse models: A necessary next step in groundwater modeling. Ground Water 35(2):250–260
Poeter EP, Hill MC (1998) Documentation of UCODE: a computer code for universal inverse modeling. U.S. Geological Survey Water-Resources Investigations Report 98–4080: 116 pp
Ramarao BS, Lavenue AM, de Marsily GH, Marietta MG (1995) Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields, 1. Theory and computational experiments. Water Resour Res 31(3):475–493

 DOI 10.1007/s10040-004-0404-7

Rao SVN, Thandaveswara BS, Bhallamudi SM (2003) Optimal groundwater management in deltaic regions using simulated annealing and neural networks. Water Resour Manag 17(6): 409–428

Rissanen J (1978) Modeling by shortest data description. Automatica 14:465–471

Roggero F, Hu LY (1998) Gradual deformation of continuous geostatistical models for history matching. In annual technical conference, SPE 49004

Rubin Y (2003) Applied stochastic hydrogeology. Oxford University Press, New York 391 pp

Rubin Y, Dagan G (1987) Stochastic Identification of Transmissivity and Effective Recharge in Steady Groundwater-Flow, 1 Theory. Water Resour Res 23(7):1185–1192

Rudin M, Beckmann N, Prosas R., Reese T, Bochelen D, Sauter A. (1999) In vivo magnetic resonance methods in pharmaceutical research: current status and perspectives. NMR Biomed 12(2):69–97

Rühli FJ, Lanz C, Ulrich-Bochsler S, Alt KW (2002) State-of-the-art imaging in palaeopathology: the value of multislice computed tomography in visualizing doubtful cranial lesions. Int J Osteoarchaeol 12(5):372–379

Sahuquillo A, Capilla J, Gómez-Hernández JJ, Andreu J (1992) Conditional simulation of transmissivity fields honouring piezomètrica data. In: Blain WR, Cabrera E Fluid Flow Modeling, Comput. Mech., Billerica, Mass, pp 201–212

Sambridge M, Mosegaard K (2002) Monte Carlo methods in geophysical inverse problems. Rev Geophys 40(3):Art. No. 1009

Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6(2):461–464

Stallman RW (1956) Numerical analysis of regional water levels to define aquifer hydrology. Am Geophys Union Trans 37(4):451–460

Tihonov AN (1963) Regularization of incorrectly posed problems, Sov. Math Dokl 4:1624–1627

Tsai FTC, Sun NZ, Yeh WG (2003) Global-local optimization methods for the identification of threedimensional parameter structure in groundwater modeling. Water Resour Res 39(2) Art:1043

Usunoff E, Carrera J, Mousavi SF (1992) An approach to the design of experiments for discriminating among alternative conceptual models. Adv Water Resour 15(3):199–214

Varni M, Carrera J (1998) Simulation of groundwater age distribution. Wat Resour Res 34(12):3271–3281

Vassolo S, Kinzelbach W, Schafer W (1998) Determination of a well head protection zone by stochastic inverse modelling. J Hydrol 206(3–4):268–280

Vecchia AV, Cooley RL (1987) Simultaneous confidence and prediction intervals for nonlinear regression models with application to a groundwater flow model. Water Resour Res 23(7):1237–1250

Vesselinov VV, Neuman SP, Illman WA (2001) Three-dimensional numerical inversion of pneumatic cross-hole tests in unsaturated fractured tuff 2. Equivalent parameters, high-resolution stochastic imaging and scale effects. Water Resour Res 37(12):3019–3041

Wagner BJ, Gorelick SM (1987) Optimal groundwater quality management under parameter uncertainty. Water Resour Res 23(7):1162–1174

Weiss R, Smith L (1998) Efficient and responsible use of prior information in inverse methods. Ground Water 36(1):151–163

Weissmann GS, Carle SA, Fogg GE (1999) Three-dimensional hydrofacies modeling based on soil survey analysis and transition probability geostatistics. Water Resour Res 35(6):1761–1770

Woodbury AD, Rubin Y (2000) A full-Bayesian approach to parameter inference from tracer travel time moments and investigation of scale effects at the Cape Cod experimental site. Water Resour Res 36(1):159–171

Woodbury AD, Smith JL, Dunbar WS (1987) Simultaneous inversion of temperature and hydraulic data, 1. Theory and application using hydraulic head data. Water Resour Res 23(8):1586–1606

Xiang Y, Sykes JF, Thomson NR (1992) A composite L_1 parameter estimator for model fitting in groundwater flow and solute transport simulation. Water Resour Res 29(6):1661–1673

Yapo PO, Gupta HV, Sorooshian S (1998) Multi-objective global optimization method for hydrological models. J Hydrol 204:83–87

Yeh TCJ, Liu SY (2000) Hydraulic tomography: Development of a new aquifer test method. Water Resour Res 36(8):2095–2105

Yeh WWG, Yoon YS (1981) Aquifer parameter identification with optimum dimension in parameterization. Wat Resour Res 17(3):664–672

Yeh WWG (1986) Review of parameter estimation procedures in groundwater hydrology: The inverse problem. Water Resour Res 22:95–108

Zimmerman DA, de Marsily G, Gotway CA, Marietta MG, Axness CL, Beauheim RL, Bras RL, Carrera J, Dagan G, Davies PB, Gallegos DP, Galli A, Gómez-Hernández JJ, Grindrod P, Gutjahr AL, Kitanidis PK, Lavenue AM, McLaughlin D, Neuman SP, RamaRao BS, Ravenne C, Rubin Y (1998) A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow. Water Resour Res 34(6):1373–1413

 DOI 10.1007/s10040-004-0404-7