Available online at www.sciencedirect.com
SCIENCE DIRECT®
Advances in Water Resources 29 (2006) 605–623
Advances in Water Resources
www.elsevier.com/locate/advwatres

The cost of uniqueness in groundwater model calibration

Catherine Moore *, John Doherty

School of Engineering, University of Queensland, St. Lucia, Queensland 4072, Australia

Received 25 February 2004; received in revised form 13 June 2005; accepted 7 July 2005
Available online 29 August 2005

Abstract

Calibration of a groundwater model requires that hydraulic properties be estimated throughout a model domain. This generally constitutes an underdetermined inverse problem, for which a solution can only be found when some kind of regularization device is included in the inversion process. Inclusion of regularization in the calibration process can be implicit, for example through the use of zones of constant parameter value, or explicit, for example through solution of a constrained minimization problem in which parameters are made to respect preferred values, or preferred relationships, to the degree necessary for a unique solution to be obtained.

The "cost of uniqueness" is this: no matter which regularization methodology is employed, the inevitable consequence of its use is a loss of detail in the calibrated field. This, in turn, can lead to erroneous predictions made by a model that is ostensibly "well calibrated". Information made available as a by-product of the regularized inversion process allows the reasons for this loss of detail to be better understood. In particular, it is easily demonstrated that the estimated value for an hydraulic property at any point within a model domain is, in fact, a weighted average of the true hydraulic property over a much larger area. This averaging process causes loss of resolution in the estimated field. Where hydraulic conductivity is the hydraulic property being estimated, high averaging weights exist in areas that are strategically disposed with respect to measurement wells, while other areas may contribute very little to the estimated hydraulic conductivity at any point within the model domain, this possibly making the detection of hydraulic conductivity anomalies in these latter areas almost impossible. A study of the post-calibration parameter field covariance matrix allows further insights into the loss of system detail incurred through the calibration process to be gained. A comparison of pre- and post-calibration parameter covariance matrices shows that the latter often possess a much smaller spectral bandwidth than the former. It is also demonstrated that, as an inevitable consequence of the fact that a calibrated model cannot replicate every detail of the true system, model-to-measurement residuals can show a high degree of spatial correlation, a fact which must be taken into account when assessing these residuals either qualitatively, or quantitatively in the exploration of model predictive uncertainty.

These principles are demonstrated using a synthetic case in which spatial parameter definition is based on pilot points, and calibration is implemented using both zones of piecewise constancy and constrained minimization regularization.

Keywords: Regularization; Model calibration; Pilot points; Resolution matrix; PEST

1. Introduction

Model calibration is the process whereby model parameters are adjusted until a satisfactory fit between model outputs and field measurements is achieved. In most cases the calibration process seeks to infer a unique parameter set, which can then be used by the model in making predictions of future system behavior. However in all but the most homogenous of geological environments, solution of the groundwater inverse problem is inherently nonunique, as model parameters representing the hydraulic properties of a heterogeneous and complex "real world" must be inferred from only a handful of

* Corresponding author.
E-mail addresses: s4024009@student.uq.edu.au (C. Moore), jdoherty@gil.com.au (J. Doherty).

0309-1708/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.advwatres.2005.07.003

measurements made at discrete points, contaminated with measurement "noise".

Notwithstanding the inherent complexity of groundwater systems, a unique solution to the inverse problem can generally be obtained by introducing additional "simplifying" information to the calibration process. Traditionally this additional information takes the form of a user-supplied zonation pattern, in which the model domain is subdivided into a limited number of regions of constant parameter value. The spatial zonation pattern may, or may not, be inferable from geological mapping, and is normally used to represent medium to large-scale variability in hydraulic properties. Alternatively, additional information can be introduced through the inclusion of mathematical regularization in the parameter estimation process. If regularized inversion is implemented using a so-called "Tikhonov scheme" this extra information can take the form of some "preferred parameter condition" (for example a smoothness or homogeneity condition), from which departures will be tolerated only to the extent that they are demanded by the calibration dataset. The end product of regularized inversion can be considered to be a parameter field that reflects, to the greatest extent possible, any information pertaining to spatial hydraulic property detail that is available in the calibration dataset. However where information is lacking, the preferred parameter condition prevails, this being a necessity if a unique solution to the inverse problem is to be attained. Alternatively "subspace methods" such as "truncated singular value decomposition" can be employed for regularization purposes. Using these methods a stable solution to the parameter estimation problem is sought within an "estimable subspace" of parameter space normally associated with parameter variations of low spatial frequency. Estimation of the high spatial frequency parameter components is often compromised by system noise, or even rendered numerically intractable due to singularity of the inverse problem when attempts are made to estimate parameters to this level of detail. Though the discussion in the present paper is limited to Tikhonov regularization, results documented herein are just as applicable to model calibration contexts in which regularized inversion is implemented using subspace methods.

While calibration based on parameter zonation or regularization allows the estimation of a unique parameter set, both of these methods yield a vastly simplified picture of the real world. The extent to which such simplification is necessary to achieve a unique solution to the inverse problems depends on the information content (or lack of information content) of the calibration dataset. Whether this simplification matters or not depends on the types of predictions the model will be required to make. Where a model prediction is sensitive to spatial hydraulic property detail, this prediction may be seriously flawed. Contaminant transport is an obvious example of a process that is highly sensitive to hydraulic property detail.

The smoothing or averaging that is an inherent part of model calibration has been the subject of a number of theoretical and practical studies. For example Backus and Gilbert [5–7], Jackson [34], Menke [43], Lines and Treitel [38], Vasco et al. [59] (to mention just a few) in the geophysical context, and Kitanidis [35], Guadagnini and Neuman [28], Yeh and Liu [64], Clemo et al. [19], Gorokhovski [27], Vasco et al. [58] and Butler et al. [11] in the groundwater (particularly tomography) context, all discuss the ability (or lack thereof) of the calibration process to infer spatial detail of earth properties from a set of continuous or point-based measurements. The literature on this subject is not confined to earth sciences, for it is a commonly recurring theme wherever imaging (for example medical imaging) is undertaken; see for example Lutkenhoner and Menendez [39], Menendez et al. [42] and Babiloni et al. [4]. All of these studies demonstrate that the interpretation of measurements made externally to a system, or at a number of discrete points within a system, leads to spatially integrated system properties, with the integration kernel being dependent on the disposition and spatial density of field measurements.

Regularized inversion presents a powerful means of rapidly fitting model outputs to field data. There is also a wealth of information available as a by-product of the regularized inversion process which can be used to gain important insights into the limitations of the calibration process in inferring details of a particular groundwater system. These limitations exist whether or not formal regularization is employed in the calibration process. In fact, because estimation of zone-based parameters can be easily recast as a regularized inversion problem (as is demonstrated in this paper), the same methods of analysis can be used in conjunction with traditional calibration methodologies. Hence the conclusions drawn in this study on the limitations of the calibration process are as applicable to traditional zone-based model calibration as they are to formal regularized inversion.

The purpose of this paper is to examine the use of regularized inversion in the context of groundwater model calibration. Regularized inversion has been discussed in the groundwater literature prior to this paper (see for example Doherty [22] and Mantoglou [41]). The use of regularized inversion in the interpretation of densely-sampled groundwater data at the site scale has also been documented; see for example Vasco and coworkers [57,58,60] and Clemo et al. [19]. However literature reports of its application, of the benefits to be gained through its use, and of the information available as a by-product of its deployment, in cases that can be considered typical of those encountered in everyday groundwater modeling work are, to our knowledge,

rare. Given the increasing use of regularized inversion in groundwater model calibration as a result of its ready availability in public domain and commercial groundwater modeling packages, further discussion of these matters is, in our opinion, urgently needed. Furthermore, whether calibration is achieved through regularized inversion, or using tradition "manual" regularization devices such as zones of piecewise constancy, it is our impression that there is insufficient appreciation in the groundwater modeling community of the relationship between true hydraulic properties and those estimated through the calibration process. This paper attempts to address these issues by presenting some of the theory of regularized inversion, and some extensions of this theory as it applies to groundwater model calibration. Through application of this theory to a synthetic case study, some important aspects of the calibration process that are often unappreciated in everyday modeling work are made readily apparent.

We first present a brief overview of the theory of regularized inversion and demonstrate the means by which this theory can be applied to the task of groundwater model calibration. Next we describe the "resolution matrix" as a device for exploring the relationship between a calibrated hydraulic property field on the one hand, and "hydraulic reality" on the other. The concept of parameter resolution then forms the basis of the discussions and illustrations that follow. In particular:

(1) Resolution analysis is employed to compare the properties of parameters estimated on the basis of traditional zone based calibration methods with those estimated through regularized inversion.
(2) The concept of parameter resolution is extended to that of the "calibrated field covariance matrix", this being an explicit indicator of the ability of the calibration process to capture (or not) the structural fine detail existing within an aquifer whose stochastic properties are known.
(3) The nature of "structural noise" as a component of "measurement error", and the dependence of this noise on the hydraulic property heterogeneity of an aquifer is described.
(4) A methodology through which inferences of "true hydraulic property heterogeneity" can be drawn on the basis of a calibrated property field is presented.

The concepts introduced above are illustrated using a synthetic model. This model is also used to demonstrate the loss of parameter detail incurred by the calibration process, and the degree of predictive error that this loss of detail can potentially incur, even where a perfect fit is obtained between model outputs and field measurements. It is hoped that the model is simple enough for the matters raised herein to be demonstrated in a manner that is unclouded by issues unrelated to those we endeavor to discuss, yet is complex enough to be representative of many real-world groundwater modeling applications.

2. Regularized parameter estimation

2.1. Background

Classical parameter estimation theory as presented in works such as Bard [8], Mikhail [44], Draper and Smith [24], and many others demonstrates that the number of parameters estimated through the calibration process cannot exceed the number of observations used for their estimation if these parameters are to be estimated uniquely. In fact in most practical cases, the number of parameters that can be uniquely estimated is much smaller than the number of observations comprising the calibration dataset due to the limited information content of most such datasets. Modelers are advised to "live within their means" when formulating an inverse problem, basing that problem on a parameterization scheme that conforms to the dictates of the "principle of parsimony"; see for example Hill [31], Carerra and Neuman [13], Akaike [1]. Where zones of parameter constancy are used as a method of spatial parameterization for groundwater models, a limited number of these zones is thus employed.

Limitations in the information content of a calibration dataset with respect to the parameters that require estimation result in parameter insensitivity and correlation; see for example Hill and Osterby [32]. At best this leads to a high degree of uncertainty in estimated parameters; at worst it leads to numerical instability in the parameter estimation package used for numerical solution of the inverse problem, often making the solution of that problem impossible. Fortunately, statistical information forthcoming from the parameter estimation process allows the modeler to readily identify insensitive and correlated parameters; the parameter estimation problem can then be reformulated to mitigate this effect, normally through reducing the number of parameters that require estimation.

Mathematical regularization as a device for model calibration provides an important and useful alternative to spatial parameterization based on a necessarily small number of zones of constant parameter value. In its broadest terms, mathematical regularization can be considered as a means of obtaining a unique solution to an inherently nonunique inverse problem, with the necessary level of simplification required to achieve this solution being determined as part of the inversion process itself. When applied to the estimation of spatially distributed parameters, it often involves the imposition of smoothness or equality constraints on those parameters;

see for example Tikhonov [54], Haber [29] and Engl et al. [25]. While routinely used in geophysical data interpretation (e.g. Mackie and Madden [40], Yao and Roberts [63], Nolet et al. [48], Bertete-Aguirre et al. [9] and many others including references already cited), regularized inversion is rarely employed in groundwater model calibration, though, as stated above, it is presently gaining increasing acceptance.

The principal advantage of using regularized inversion as a means of groundwater model calibration is that a large number of parameters can be distributed throughout the model domain to define the spatial distribution of the hydraulic property requiring estimation. To prevent the onset of numerical instability that would inevitably result from the estimation of so many parameters, so-called "Tikhonov schemes" are often employed as a regularization device. Using such a scheme, each parameter is accompanied by at least one "observation" pertaining directly to that parameter (and maybe to others as well), thereby defining a "preferred condition" for that parameter while at the same time ensuring that there are more "observations" than parameters present in the inversion process. Parameter uniqueness, and hence numerical stability of the inverse problem, is guaranteed through the assignment of appropriate weights to such "regularization observations", often through the application of a "weight multiplier" to all of them. This weight multiplier is calculated as part of the regularized inversion process such that it is high enough to allow parameter uniqueness to be attained, yet not so high that it compromises the goodness of fit between model outputs and field data beyond that which is required to prevent "over fitting" to measurement noise. As already mentioned, other regularization methodologies are also available, notably those based on parameter subspaces. For simplicity, these will not be discussed herein. However conclusions drawn below pertaining to loss of parameterization detail, and its effect on model predictive error, are just as pertinent to these schemes as they are to Tikhonov schemes.

In the present study, regularized model calibration is undertaken in a similar manner to that described by Doherty [22], and is numerically implemented through the use of the PEST [23] parameter estimation package. The theory is now briefly outlined.

2.2. Basic theory

In unregularized, or "overdetermined" linear parameter estimation, parameter values are calculated using the equation:

$$\underline{\mathbf{p}} = (\mathbf{X}^t\mathbf{Q}_c\mathbf{X})^{-1}\mathbf{X}^t\mathbf{Q}_c\mathbf{c} \tag{1}$$

where $\underline{\mathbf{p}}$ is the vector of estimated parameters; $\mathbf{c}$ is the vector of measurements; $\mathbf{X}$ is the model Jacobian matrix (i.e. the matrix of derivatives of model outputs to which there are observed equivalents with respect to adjustable model parameters, often referred to also as the "sensitivity matrix"); and $\mathbf{Q}_c$ is the measurement weight matrix, often chosen as the inverse of the measurement covariance matrix which we denote by $C(\boldsymbol{\varepsilon})$, $\boldsymbol{\varepsilon}$ specifying the noise associated with the measurements $\mathbf{c}$.

Eq. (1) is derived through minimization of an objective function (see Eq. (4)) comprised of weighted squared differences between model outputs and field measurements; see for example Bard [8], Mikhail [44], Draper and Smith [24] and many others. For nonlinear models $\mathbf{X}$ is a function of parameter values. Eq. (1) is then solved by an iterative procedure in which $\mathbf{X}$ is recalculated during every iteration, and in which parameter improvements are calculated on the basis of current model-to-measurement residuals, rather than directly from the measurements themselves.

If we supplement $\mathbf{c}$ with a set of "regularization observations" $\mathbf{z}$, and if we supplement the model with a set of "regularization conditions" on estimated parameters encapsulated in a matrix $\mathbf{Z}$, the following equation describing regularized parameter estimation is obtained through a simple extension of Eq. (1):

$$\underline{\mathbf{p}} = (\mathbf{X}^t\mathbf{Q}_c\mathbf{X} + \mu\mathbf{Z}^t\mathbf{Q}_z\mathbf{Z})^{-1}(\mathbf{X}^t\mathbf{Q}_c\mathbf{c} + \mu\mathbf{Z}^t\mathbf{Q}_z\mathbf{z}) \tag{2a}$$

where $\mathbf{Z}$ is a matrix which describes the relationships between parameters that collectively define a "preferred system state". These relationships will prevail in the absence of any evidence to the contrary available through observations encapsulated in $\mathbf{c}$; $\mathbf{z}$ is the vector of values defining the "preferred system state", referred to herein as the set of "regularization observations"; $\mathbf{Q}_z$ is a "regularization weight matrix"; and μ is a "weight multiplier" applied to all regularization observations.

For the remainder of this paper we consider only cases for which regularization observations are all zero, and therefore for which the $\mathbf{Z}^t\mathbf{Q}_z\mathbf{z}$ term of Eq. (2a) is zero. This involves no loss of generality, as it can easily be achieved in most practical cases through appropriate parameter transformation and/or offsetting. With $\mathbf{z}$ equated to zero, Eq. (2a) can be re-written:

$$\underline{\mathbf{p}} = (\mathbf{X}^t\mathbf{Q}_c\mathbf{X} + \mu\mathbf{Z}^t\mathbf{Q}_z\mathbf{Z})^{-1}\mathbf{X}^t\mathbf{Q}_c\mathbf{c} \tag{2b}$$

The regularization relationships encapsulated in $\mathbf{Z}$ may be linear or nonlinear; in the latter case $\mathbf{Z}$ (which then represents the derivatives of those relationships with respect to parameter values), must be recalculated during every optimization iteration. Often-used regularization relationships include discrete parameter equality (in which case $\mathbf{Z}$ is simply a diagonal matrix), or equality of parameters to each other (in which case each row of $\mathbf{Z}$ is comprised of a "+1" and a "−1" at elements pertaining to the respective parameters).

As well as being considered a "weight multiplier", it can be shown that μ is also equivalent to the reciprocal

of a Lagrange multiplier used in the solution of the constrained minimization problem posed as follows:

$$\text{Minimize } \Phi_r = (\mathbf{Z}\underline{\mathbf{p}} - \mathbf{z})^t \mathbf{Q}_z (\mathbf{Z}\underline{\mathbf{p}} - \mathbf{z}) \tag{3a}$$

or as $\mathbf{z} = \mathbf{0}$ in this case,

$$\text{Minimize } \Phi_r = \underline{\mathbf{p}}^t \mathbf{Z}^t \mathbf{Q}_z \mathbf{Z}\underline{\mathbf{p}} \tag{3b}$$

under the constraint that

$$\Phi_m = (\mathbf{X}\underline{\mathbf{p}} - \mathbf{c})^t \mathbf{Q}_c (\mathbf{X}\underline{\mathbf{p}} - \mathbf{c}) \leqslant \Phi_m^l \tag{4}$$

where Φ_r is referred to herein as the "regularization objective function", Φ_m is referred to as the "measurement objective function", and Φ_m^l is referred to as the "limiting measurement objective function".

As formulated through Eqs. (3) and (4), Tikhonov-based regularized inversion seeks to minimize the sum of squared outcomes of the regularization relationships encapsulated in $\mathbf{Z}$ (and hence maximize the extent to which estimated parameters respect the preferred system condition encapsulated in those relationships), while insisting that the departure of model outputs from field observations, as measured in the weighted least squares sense, is no greater than a user-defined threshold, Φ_m^l. This threshold should be set such that the sum of squared model-to-measurement residuals is commensurate with expected field measurement uncertainty as described by the measurement covariance matrix $C(\boldsymbol{\varepsilon})$; thus "over fitting" or "under fitting" is prevented. In implementing the constrained minimization problem described by Eqs. (3) and (4), a new weight factor μ must be computed during each optimization iteration. In PEST, this is done using an iterative procedure based on local linearization of the constrained optimization problem; see Doherty [23] for details.

2.3. The resolution matrix

With appropriate choice of μ, Eq. (2) provides a unique solution to the constrained minimization problem described by Eqs. (3) and (4). However it does not provide a unique solution for the estimated parameter set $\underline{\mathbf{p}}$ itself, for different solutions can be obtained using different regularization operators $\mathbf{Z}$. Some regularization operators are more appropriate for use in some circumstances than others (see, for example, Portniaguine and Zhdanov [50] and Vogel and Oman [62]), their appropriateness being judged partly on the basis of the characteristics of the estimated property field $\underline{\mathbf{p}}$ produced as an outcome of their deployment. One way in which the characteristics of the estimated property field can be explored, and in which different regularization methodologies can be compared, is through the so-called "resolution matrix", which will now be defined.

The relationship between the set of field measurements $\mathbf{c}$ and the true hydraulic property field $\mathbf{p}$ is described by the equation:

$$\mathbf{c} = \mathbf{X}\mathbf{p} + \boldsymbol{\varepsilon} \tag{5}$$

where $\boldsymbol{\varepsilon}$ is a "noise vector", as described above, normally assumed to be random with a mean of zero and a covariance matrix $C(\boldsymbol{\varepsilon})$. If Eq. (5) is substituted into Eq. (3) the following equation is obtained:

$$\underline{\mathbf{p}} = (\mathbf{X}^t \mathbf{Q}_c \mathbf{X} + \mu \mathbf{Z}^t \mathbf{Q}_z \mathbf{Z})^{-1} \mathbf{X}^t \mathbf{Q}_c (\mathbf{X}\mathbf{p} + \boldsymbol{\varepsilon}) \tag{6a}$$

or

$$\underline{\mathbf{p}} = \mathbf{R}\mathbf{p} + \mathbf{G}\boldsymbol{\varepsilon} \tag{6b}$$

where the "resolution matrix" $\mathbf{R}$ is given by the equation:

$$\mathbf{R} = (\mathbf{X}^t \mathbf{Q}_c \mathbf{X} + \mu \mathbf{Z}^t \mathbf{Q}_z \mathbf{Z})^{-1} \mathbf{X} \mathbf{Q}_c \mathbf{X} \tag{7a}$$

and $\mathbf{G}$ is defined by:

$$\mathbf{G} = (\mathbf{X}^t \mathbf{Q}_c \mathbf{X} + \mu \mathbf{Z}^t \mathbf{Q}_z \mathbf{Z})^{-1} \mathbf{X}^t \mathbf{Q}_c \tag{7b}$$

If measurement noise is small enough to be ignored Eq. (6) becomes simply:

$$\underline{\mathbf{p}} = \mathbf{R}\mathbf{p} \tag{8}$$

Thus the matrix $\mathbf{R}$ describes the relationship between the true property field $\mathbf{p}$ and its estimated counterpart $\underline{\mathbf{p}}$. Eq. (8) tells us that each element of $\underline{\mathbf{p}}$ is the scalar product of the pertinent row of $\mathbf{R}$ with the true property vector $\mathbf{p}$. If the matrix $\mathbf{R}$ is the identity matrix (implying perfect resolution of the estimated field), $\underline{\mathbf{p}}$ and $\mathbf{p}$ would be equal. In practice the best that can be hoped for is that $\mathbf{R}$ is diagonally dominant, and that the components of $\mathbf{R}$ decay rapidly to the left and right of its diagonal. To the extent that off-diagonal elements of $\mathbf{R}$ are significant, the estimated hydraulic property at any point within the model domain (for example the i'th component of $\underline{\mathbf{p}}$) is a weighted average of true hydraulic property values. In many circumstances, this averaging can occur over a large part of the model domain. Thus the "image" of $\mathbf{p}$ (i.e. $\underline{\mathbf{p}}$) produced as an outcome of the calibration process is "blurred", and the estimated field at any point is unlikely to be equal to the true field at that point. The closer that $\mathbf{R}$ approaches the identity matrix, the more representative is the calibrated field of the true field. Conversely, the greater the bandwidth of $\mathbf{R}$, the larger is the spatial extent of the "integration kernel" through which the estimated field is derived by spatial averaging of the true field. Note that for over-determined inversion based on Eq. (1), $\mathbf{R}$ is equal to $\mathbf{I}$. This means that the model calibration process yields perfect resolution of system complexity under the assumption that the complexity of the real world is no greater than that of the model which represents it.

2.4. The regularized parameter covariance matrix

Prior to model calibration much is often known about hydraulic properties prevailing within a study

area as a result of geological investigations and other site characterization activities. This knowledge usually takes the form of most likely values for hydraulic properties within different parts of the model domain based on lithologic inference or pump test interpretation, together with the degree of variability to be expected of these properties. Conceptually, pre-calibration parameter uncertainty can be summarized through the use of a parameter covariance matrix $C(\mathbf{p})$. If enough information is available within a study area to construct a hydraulic property variogram, $C(\mathbf{p})$ can be calculated from that.

Once $C(\mathbf{p})$ is known (or estimated) simple matrix relationships for propagation of covariance allow us to calculate a covariance matrix of the estimated field, i.e. $C(\underline{\mathbf{p}})$, from Eqs. (6) as

$$\begin{aligned} C(\underline{\mathbf{p}}) = {} & (\mathbf{X}^t\mathbf{Q}_c\mathbf{X} + \mu\mathbf{Z}^t\mathbf{Q}_z\mathbf{Z})^{-1}\mathbf{X}^t\mathbf{Q}_c\mathbf{X}C(\mathbf{p}) \\ & \times \mathbf{X}^t\mathbf{Q}_c\mathbf{X}(\mathbf{X}^t\mathbf{Q}_c\mathbf{X} + \mu\mathbf{Z}^t\mathbf{Q}_z\mathbf{Z})^{-1} \\ & + (\mathbf{X}^t\mathbf{Q}_c\mathbf{X} + \mu\mathbf{Z}^t\mathbf{Q}_z\mathbf{Z})^{-1}\mathbf{X}^tC(\boldsymbol{\varepsilon}) \\ & \times \mathbf{X}(\mathbf{X}^t\mathbf{Q}_c\mathbf{X} + \mu\mathbf{Z}^t\mathbf{Q}_z\mathbf{Z})^{-1} \end{aligned} \tag{9a}$$

where, for simplicity, $C(\boldsymbol{\varepsilon})$, the covariance matrix of measurement noise, is assumed to be the inverse of $\mathbf{Q}_c$. Note that derivation of Eq. (9a) from Eq. (6) assumes that no correlation exists between measurement noise and hydraulic property values in different parts of the model domain.

If measurement noise is neglected Eq. (9a) becomes:

$$\begin{aligned} C(\underline{\mathbf{p}}) = {} & (\mathbf{X}^t\mathbf{Q}_c\mathbf{X} + \mu\mathbf{Z}^t\mathbf{Q}_z\mathbf{Z})^{-1}\mathbf{X}^t\mathbf{Q}_c\mathbf{X}C(\mathbf{p}) \\ & \times \mathbf{X}^t\mathbf{Q}_c\mathbf{X}(\mathbf{X}^t\mathbf{Q}_c\mathbf{X} + \mu\mathbf{Z}^t\mathbf{Q}_z\mathbf{Z})^{-1} \end{aligned} \tag{9b}$$

A short example may improve understanding of the $C(\underline{\mathbf{p}})$ matrix. Suppose that a thousand different realisations of a hydraulic conductivity field are generated within a model domain based on the same $C(\mathbf{p})$. Suppose that each one of these hydraulic conductivity fields is then used to parameterise a model, which is then run in order to generate a set of outputs. Suppose now that each of these outputs is sampled at identical locations, and that regularized inversion (using the same inversion methodology in each case) is then undertaken based on each synthetic measurement set. The result would be one thousand estimated hydraulic property fields, the statistical properties of which would be approximated by $C(\underline{\mathbf{p}})$ (''approximated'' because the model is not linear, and because even if it were, identical values of the weight factor μ would not be calculated in each case). The diagonal elements of $C(\underline{\mathbf{p}})$ describe variability of individual parameters between these fields, that is, parameter variability at any one point in the model domain over the notional one thousand calibration exercises. The off-diagonal elements represent the degree to which field estimates at different points within the model domain are statistically related.

2.5. Spectral decomposition of the true and estimated fields

In many instances of groundwater model calibration, the matrix $\mathbf{X}$ does not have full column rank because either the number of parameters exceeds the number of observations, and/or because the information content of the calibration dataset is insufficient to provide unique estimation of parameters pertaining to all of the detail represented in the model. The matrix $\mathbf{X}^t\mathbf{Q}_c\mathbf{X}$ of Eqs. (9a) and (9b) is thus rank-deficient (Koch [36] p. 28), having a maximum rank equal to the number of observations (i.e. the dimension of $\mathbf{c}$), though often considerably less than this. The matrix $C(\underline{\mathbf{p}})$ calculated by those equations is thus rank-deficient, and is therefore singular.

Notwithstanding its singularity, $C(\underline{\mathbf{p}})$ is a symmetric, positive semidefinite matrix. As such it has real eigenvalues, with the number of non-zero eigenvalues being equal to its rank. These non-zero eigenvalues, and corresponding eigenvectors, can easily be determined through singular value decomposition. As a result of such an analysis, the stochastic character of the estimated parameter field $\underline{\mathbf{p}}$ can be represented by independent random variations in directions which are mutually orthogonal in parameter space. These directions are those of the eigenvectors of $C(\underline{\mathbf{p}})$ that correspond to its non-zero eigenvalues; the variance, or stochastic variability, in each of these directions is given by the corresponding eigenvalue.

Similarly, eigenvalue/eigenvector decomposition of $C(\mathbf{p})$, the spatial covariance matrix of the true field, results in a set of independent, uncorrelated random variables pertaining to orthogonal directions in parameter space (i.e. the directions of the eigenvectors of $C(\mathbf{p})$). The variance in each such direction is equal to the corresponding eigenvalue. A comparison of the eigenvalues/eigenvectors of $C(\mathbf{p})$ with those of $C(\underline{\mathbf{p}})$ can reveal much about the loss of detail incurred through the calibration process as will be demonstrated for a synthetic example later in this paper.

Note that discussions of covariance matrix decomposition in relation to inversion problems in the stochastic framework can be found in Braud and Obled [10] and Delay et al. [21].

2.6. Inference of true hydraulic property variability

Eqs. (9) relate the spatial variability of an estimated hydraulic property field to that of a true hydraulic property field. Consequently they provide a mechanism for testing plausibility of a posited true field covariance matrix $C(\mathbf{p})$, once an estimated hydraulic property field has been obtained through regularized inversion.

To test the validity of a hypothesized $C(\mathbf{p})$, that $C(\mathbf{p})$ can be inserted into the right side of either of Eqs. (9).

The hypothesis that $C(\mathbf{p})$ represents a true descriptor of hydraulic property variability within the model domain can then be accepted or rejected at a certain level of confidence by testing the compatibility of the estimated field $\underline{\mathbf{p}}$ derived using Eqs. (2), with $C(\underline{\mathbf{p}})$ calculated through Eqs. (9). A testing procedure is now described.

For the ith non-zero eigenvalue v_i of $C(\underline{\mathbf{p}})$, define η_i as

$$\eta_i = \underline{\mathbf{p}}^t \mathbf{e}_i / v_i^{1/2} \tag{10}$$

where $\mathbf{e}_i$ is the eigenvector of unit length corresponding to v_i. If $C(\underline{\mathbf{p}})$ as calculated by Eqs. (9) properly characterizes the spatial covariance structure of $\underline{\mathbf{p}}$, then each η_i should have a variance of 1, and will be statistically independent from all other η_i. Furthermore, if $C(\mathbf{p})$ is multi-Gaussian and if model outputs are reasonably linear with respect to parameters, then the η_i should be normally distributed, and possess a mean of zero. Thus the sum of all squared η_i values corresponding to non-zero eigenvalues of $C(\underline{\mathbf{p}})$ should belong to a chi-square distribution with order equal to the number of elements summed. In practice, not all of these eigenvalues can be used for testing chi-square behaviour due to amplification of numerical noise in directions of low spatial variability as a result of near-zero denominators in Eq. (10) corresponding to low eigenvalues. Hence the summation must be restricted to a sufficient number of terms to account for most of the variability of $C(\underline{\mathbf{p}})$ as calculated through summing eigenvalues, without contaminating this sum through the inclusion of noise.

(It is worth noting in passing that rarely, if ever, is real-world heterogeneity characterized by a multi-Gaussian stochastic structure. The analysis above should, however, in many circumstances provide some insights into the level of "uncaptured heterogeneity" prevailing in the real hydraulic property field; in most circumstances such insights are likely to be qualitative rather than quantitative.)

2.7. Zones as a regularization device

As has already been discussed, in many groundwater modeling applications, zones of constant parameter value are used to represent geological heterogeneity. A fundamental distinction between parameters based on a limited number of zones, and the use of regularized inversion based on a large number of parameters is the number of degrees of freedom that exist in the space of estimated parameters under these two different calibration methodologies. As was discussed above, where parameters outnumber observations in regularized inversion, the number of degrees of freedom in parameter space from which the calibrated field can be constructed is potentially as high as the number of observations (though it may be considerably less than this if the information provided by different observations is not distinct or is shrouded in measurement noise). Furthermore regularized inversion allows the parameter estimation process to automatically place heterogeneity where it is most needed. In contrast, where regularization is achieved by restricting the calibration process to the estimation of parameters pertaining to a limited number of zones of piecewise constancy, the number of degrees of freedom that is available in parameter space is equal to the number of zones. This places restrictions on the fit that can be achieved between model outcomes and field data, and potentially suppresses some of the information residing in field measurements.

Through a slight expansion of the concepts presented so far, the relationship between estimated and true hydraulic properties when zones of piecewise constancy are employed as a pre-calibration regularization device can be explored using a "zonal resolution matrix". Suppose that zonal parameters are represented by the vector $\mathbf{s}$ whose dimensionality is low enough to ensure that a unique solution to the inverse problem is obtained using Eq. (1); that is

$$\underline{\mathbf{s}} = (\mathbf{Y}^t \mathbf{Q}_c \mathbf{Y})^{-1} \mathbf{Y}^t \mathbf{Q}_c \mathbf{c} \tag{11}$$

which is identical to Eq. (1) except for the fact that $\mathbf{Y}$ is used instead of $\mathbf{X}$, where $\mathbf{Y}$ is the Jacobian matrix as it pertains to the limited number of lumped parameters encapsulated in $\mathbf{s}$, and thus has far fewer columns than $\mathbf{X}$, but has the same number of rows. If Eq. (5) is now substituted for $\mathbf{c}$ we obtain:

$$\underline{\mathbf{s}} = (\mathbf{Y}^t \mathbf{Q}_c \mathbf{Y})^{-1} \mathbf{Y}^t \mathbf{Q}_c \mathbf{X} \mathbf{p} + (\mathbf{Y}^t \mathbf{Q} \mathbf{Y})^{-1} \mathbf{Q}_c \boldsymbol{\varepsilon} \tag{12a}$$

That is

$$\underline{\mathbf{s}} = \mathbf{R}' \mathbf{p} + \mathbf{G}' \boldsymbol{\varepsilon} \tag{12b}$$

where

$$\mathbf{R}' = (\mathbf{Y}^t \mathbf{Q}_c \mathbf{Y})^{-1} \mathbf{Y}^t \mathbf{Q}_c \mathbf{X} \tag{13}$$

and

$$\mathbf{G}' = (\mathbf{Y}^t \mathbf{Q}_c \mathbf{Y})^{-1} \mathbf{Y}^t \mathbf{Q}_c \tag{14}$$

It is apparent that each row of the matrix $\mathbf{R}'$ represents the "integration kernel" or "averaging weights" by which the elements of the lumped parameters $\underline{\mathbf{s}}$ are obtained from the "true" hydraulic properties $\mathbf{p}$.

2.8. Spatial correlation of structural noise

Suppose that the presence of measurement noise dictates that a less than perfect fit be sought between model outputs and field measurements through the regularized inversion process. The target measurement objective function Φ_m^l of Eq. (4) should thus be set at a value that reflects the level of expected measurement noise. From Eqs. (5) and (6):

$$\begin{aligned} \mathbf{c} &= \mathbf{X}\mathbf{p} + \boldsymbol{\varepsilon} = \mathbf{X}\mathbf{p} - \mathbf{X}\underline{\mathbf{p}} + \mathbf{X}\underline{\mathbf{p}} + \boldsymbol{\varepsilon} \\ &= \mathbf{X}\underline{\mathbf{p}} + \mathbf{X}(\mathbf{I} - \mathbf{R})\mathbf{p} + (\mathbf{I} - \mathbf{X}\mathbf{G})\boldsymbol{\varepsilon} \end{aligned} \tag{15}$$

The last two terms of Eq. (15) define model-to-measurement misfit of the calibrated model (i.e. residuals). The last of these is present even where calibration is effected using over-determined inversion based on Eq. (1). The middle term, however, appears because of the fact that model calibration is undertaken using regularized inversion; its disappearance in the overdetermined case (in which the system is assumed to be completely characterized by the limited number of parameters employed by the model) is readily demonstrated by setting $\mathbf{R}$ in this equation to $\mathbf{I}$. Thus this term is a "structural noise" term that describes the action of the model operator $\mathbf{X}$ on parameterization detail unrepresented in the calibrated model; that is, it represents the action of $\mathbf{X}$ on $\mathbf{p} - \underline{\mathbf{p}}$.

Where regularization is based on pre-calibration parameter lumping a similar equation can be derived. In that case we again start with the equation:

$$\mathbf{c} = \mathbf{X}\mathbf{p} + \boldsymbol{\varepsilon}$$

but through substitution of Eq. (12) we obtain:

$$\mathbf{c} = \mathbf{Y}\mathbf{s} + (\mathbf{X} - \mathbf{Y}\mathbf{R}')\mathbf{p} + (\mathbf{I} - \mathbf{Y}\mathbf{G}')\boldsymbol{\varepsilon} \tag{16}$$

Once again, structural noise is represented by the second term on the right side of this equation. It expresses the difference in system states $\mathbf{c}$ calculated using the "true model" $\mathbf{X}$ on the one hand and its lumped equivalent $\mathbf{Y}$ on the other hand, acting on the "true parameters" $\mathbf{p}$ in the first case and their lumped equivalents $\mathbf{R}'\mathbf{p}$ in the second case.

Measurement noise $\boldsymbol{\varepsilon}$ is often assumed to be uncorrelated. However structural noise does indeed show spatial correlation; therefore so too will calibration residuals. To demonstrate this, we first represent the structural noise term by $\boldsymbol{\delta}$ using the equation:

$$\boldsymbol{\delta} = \mathbf{X}(\mathbf{I} - \mathbf{R})\mathbf{p} \tag{17a}$$

or

$$\boldsymbol{\delta} = (\mathbf{X} - \mathbf{Y}\mathbf{R}')\mathbf{p} \tag{17b}$$

depending on whether regularization is undertaken through the regularized inversion process itself, or through pre-emptive parameter zonation. The covariance matrix of structural noise is then easily calculated as

$$\mathbf{C}(\boldsymbol{\delta}) = \mathbf{X}(\mathbf{I} - \mathbf{R})\mathbf{C}(\mathbf{p})(\mathbf{I} - \mathbf{R})^t\mathbf{X}^t \tag{18a}$$

or

$$\mathbf{C}(\boldsymbol{\delta}) = (\mathbf{X} - \mathbf{Z}\mathbf{R}')\mathbf{C}(\mathbf{p})(\mathbf{X} - \mathbf{Z}\mathbf{R}')^t \tag{18b}$$

It is apparent from Eqs. (18) that $\mathbf{C}(\boldsymbol{\delta})$ is a function of $\mathbf{C}(\mathbf{p})$ and thus depends on the level of heterogeneity present (though not necessarily capturable) in the real-world system that is simulated by the model. Furthermore, $\mathbf{C}(\boldsymbol{\delta})$ will not, in general, be a diagonal matrix; thus structural noise is likely to be spatially correlated.

This short analysis demonstrates an important, though seldom appreciated, property of model-to-measurement residuals. Conventional wisdom as applied to groundwater model calibration leads us to seek, at the end of a calibration exercise, model-to-measurement residuals that are uncorrelated in space, see Anderson and Woessner [3] and modelling guidelines such as STOWA [53] and Murray–Darling Basin Commission [47]. However Eqs. (18) inform us that model-to-measurement residuals will, in fact, have a spatial correlation structure that depends on the level and type of heterogeneity prevailing in the true property field, and on the regularization process employed to achieve a unique solution to the inverse problem. An example of spatial correlation of model-to-measurement residuals in a zone-based model will be presented later in this paper.

The fact that model-to-measurement residuals are likely to show spatial correlation can have important repercussions when attempts are made to quantify the uncertainty of parameters, and/or of model predictions, using methods such as that proposed by Poeter and Hill [49] and Hill [31] (based on an assumed linear relationship between model parameters and model outputs) or by Vecchia and Cooley [61], Christensen and Cooley [15,16] in which no such linearity assumption is invoked. In all of these cases, the calculated confidence intervals of model parameters and/or predictions are based on the assumed probability structure of model-to-measurement noise; in most cases statistical independence of this noise is assumed. However where model-to-measurement misfit is principally an outcome of "model structural error" (as it usually is), neglecting the spatially correlated nature of this term invalidates calculation of parameter and prediction uncertainty. This has been recognised by Cooley [20], Christensen and Cooley [17], and Christensen et al. [18] who have explicitly included $C(\boldsymbol{\delta})$ in their analysis of parameter and model predictive error variance.

3. Synthetic case

3.1. Description

Fig. 1a shows the domain of a groundwater model. It is rectangular, with dimensions 500 m by 800 m and thickness 10 m. Along the top of the model is an inflow boundary, with water entering the system at a rate of 0.1 m^3 per day per metre of boundary. Along the bottom of the domain is a fixed head boundary, with the head set at 0.0 m uniformly along this boundary. There is no recharge within the model domain itself. The simulated aquifer is confined and conditions are steady state.

An hydraulic conductivity field was synthesised using a multi-Gaussian field generator based on a log-exponential variogram. The equation for this type of variogram is given by:

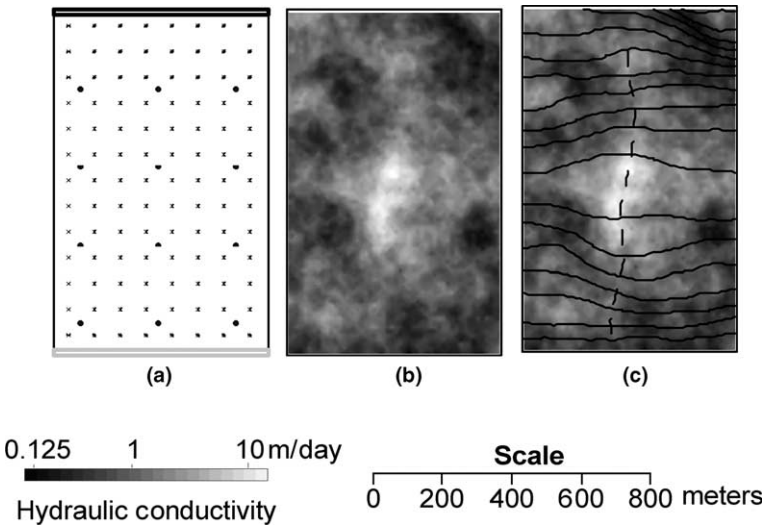

Fig. 1. (a) Model domain and grid, 104 pilot point locations (crosses), and 12 observation well locations (circles), (b) "true" hydraulic conductivity field, (c) piezometric contours (solid lines) and particle path (dotted line).

$$\gamma(h) = b(1 - e^{-ah}) \quad (19)$$

where b is the sill, and the coefficient a is often considered to be about one third of the effective range of the variogram. In the present case b was assigned a value of 0.2 while a was assigned a value of 200 m; an average log conductivity value of 0.0 (corresponding to an hydraulic conductivity of 1.0 m/day) was employed. This conductivity distribution was used as the basis for generation of "field data" employed in the calibration exercises documented below. In the discussion that follows, this is referred to as the "true" hydraulic conductivity distribution for the study area; it is depicted in Fig. 1b.

Groundwater flow was simulated using MODFLOW 2000 [30] using a finite difference grid comprised of 4000 cells, each 10 m square. Particle movement was simulated using the MODFLOW 2000 ADV package [2].

Fig. 1a shows the locations of 12 observation wells (as circles), evenly distributed throughout the model domain. For many of the calibration exercises documented in this paper, the heads at these 12 wells comprised the calibration dataset. In other exercises the heads in a greater number of wells were used for model calibration.

Fig. 1c shows piezometric contours throughout the model domain calculated by MODFLOW on the basis of the true hydraulic conductivity field shown in Fig. 1b. Also shown is the MODFLOW-calculated track of a particle released into the model domain 100 m south of the upper inflow boundary. Time of travel to the lower fixed-head boundary is calculated as 3256 days; this is referred to as the "true" particle travel time.

3.2. Parameter estimation

For the purpose of model calibration, pilot points were employed to represent spatial hydraulic conductivity variation within the groundwater model domain. Using this scheme, each element of $\underline{\mathbf{p}}$ represents the log of hydraulic conductivity at the location of a pilot point. During the inversion process, hydraulic conductivity is estimated at each of these locations; hydraulic conductivities for all of the cells of the model grid are then calculated through spatial interpolation (using kriging in the present case) from this set of pilot points. The use of pilot points as a device for groundwater model parameterization dates back to Certes and de Marsily [14]. However their use in conjunction with regularized inversion (which allows deployment of a far greater number of pilot points than would otherwise be possible) was first demonstrated by Doherty [22]. In the present study the estimated hydraulic conductivity distribution was characterized using 104 pilot points; these are shown as crosses in Fig. 1a. It is readily admitted that the use of pilot points, like the use of zones, introduces a certain degree of "pre-emptive averaging" into the estimation of hydraulic properties undertaken through the calibration process; the smaller the number of pilot points employed, the greater the degree of this averaging. However this effect is ignored in the present study as the use of 104 pilot points is sufficient to ensure that the conclusions drawn herein are minimally affected by pilot-point induced parameter averaging [46].

Though many options exist for the $\mathbf{Z}$ matrix of Eqs. (2), only two are employed in this study. The first option (implementing what is referred to herein as "smoothing regularization") is the same as that described by Doherty [22]. The basis of this scheme is that the difference between the logs of hydraulic conductivity values at pairs of pilot points is "observed" to be zero. Because each row of the $\mathbf{Z}$ matrix of Eq. (2) thus constitutes a relationship in which the difference between two

parameter values is taken, each such row contains only two non-zero elements, these being "+1" and "−1". A diagonal $\mathbf{Q}_z$ matrix was employed, the magnitude of each diagonal element being greater for pilot points which are close together than for pilot points which are further apart; this weight was calculated as the square of the inverse of the hydraulic conductivity variogram value for the respective pilot point separation distance. This regularization scheme thus attempts to force pilot points which are closer together to have similar values (unless this is contradicted by the calibration dataset), with specified weights taken from the variogram which is assumed to describe spatial variability of that hydraulic property.

For the second $\mathbf{Z}$ matrix option (implementing what is referred to herein as "preferred value regularization"), the log of the hydraulic conductivity at each pilot point was individually assigned a preferred value of zero (corresponding to an actual hydraulic conductivity value of 1.0 m/day). While this preferred value is obviously suitable for the synthetic model used in this study given the fact that generation of the synthetic hydraulic conductivity field used to represent "reality" was based on a mean log-value of zero, it is important to note that this "preferred value" regularization scheme is readily extended to more complex calibration contexts where the mean hydraulic property value differs from unity. For example, parameter deviations from another uniform value, or from a "trend value" - possibly determined as part of the calibration process itself - can easily be accommodated. In the present work, the regularization covariance matrix $\mathbf{Q}_z$ of Eq. (2) was selected as the inverse of the spatial covariance matrix describing the hydraulic conductivity distribution, i.e. $C^{-1}(\mathbf{p})$; however this is not essential for implementation of "preferred value" regularization.

3.3. Smoothing of hydraulic conductivity fields and predictive error

In order to explore the ability of the calibration process to reproduce spatial variability of hydraulic properties, the model of Fig. 1 was calibrated using "observed" heads in different numbers of wells spread throughout the model domain. The "preferred value" regularization scheme described above was employed in this exercise.

Fig. 2a–f shows the estimated hydraulic conductivity fields resulting from 6 different calibration exercises; the locations of wells whose heads were used in each of these calibration exercises are superimposed on these estimated fields. In all cases model-to-measurement head misfit at all of these locations was negligible (as they should be, as no noise was added to the MODFLOW-calculated heads before using these heads for model calibration). The grey scale used to represent the log of hydraulic conductivity in each of these plots is the same as that used to represent the true field in Fig. 1. Collectively, these figures provide a graphic illustration of the relationship between true and estimated hydraulic property fields. At no point in the model domain is the calibrated field equal to the true field, in spite of the near-perfect fit between model outputs and field measurements. More importantly, as less data is made available for calibration, there is a serious reduction in the ability of the calibrated field to reproduce spatial detail contained in the true field.

This loss of ability to infer hydraulic property detail may or may not matter, depending on how the model is to be used. However, if the model is used to make a prediction that is dependent on system detail beyond that which can be captured by the calibration process, then the consequences of using the calibrated field to make that prediction can be substantial. For each of Fig. 2a–f, the particle travel time predicted using the respective calibrated field is recorded below the figure; recall that the "true" travel time (i.e. the travel time calculated on the basis of the true hydraulic conductivity distribution) is 3256 days. It is immediately obvious that significant errors are incurred when any of the "perfectly calibrated" models illustrated in Fig. 2 are used to predict this travel time, with this error generally increasing as data support for the calibration process is reduced.

Fig. 3a–d shows another series of estimated hydraulic conductivity fields. For each of these cases heads from 84 wells were used in the calibration process (see Fig. 2a for the locations of these wells). However for these experiments progressively larger amounts of Gaussian noise were added to the head measurements in order to explore the effect of measurement inaccuracy on the calibration process. With the addition of Gaussian noise, the user-specified value for the limiting measurement objective function (Φ_m^l of Eq. (4)) was increased to a level commensurate with this noise. In this way the regularized inversion process was prevented from "over fitting" the measurement data beyond the level of noise that was present within it; nor was it allowed to "under fit" the measurement data through acceptance of an inappropriately large level of model-to-measurement misfit.

It is apparent from Fig. 3 that the addition of noise to the calibration dataset has a similar effect to the progressive removal of data. That is, the ability of the calibration process to infer hydraulic property detail decreases with increasing measurement noise. An inspection of predicted travel times recorded under these figures reveals that this loss of detail has, once again, serious consequences for the accuracy of these model predictions.

In order to compare predictive error introduced through zone-based parameterization with that introduced through regularized inversion, calibration of the synthetic model was repeated using zones. Fig. 4a shows

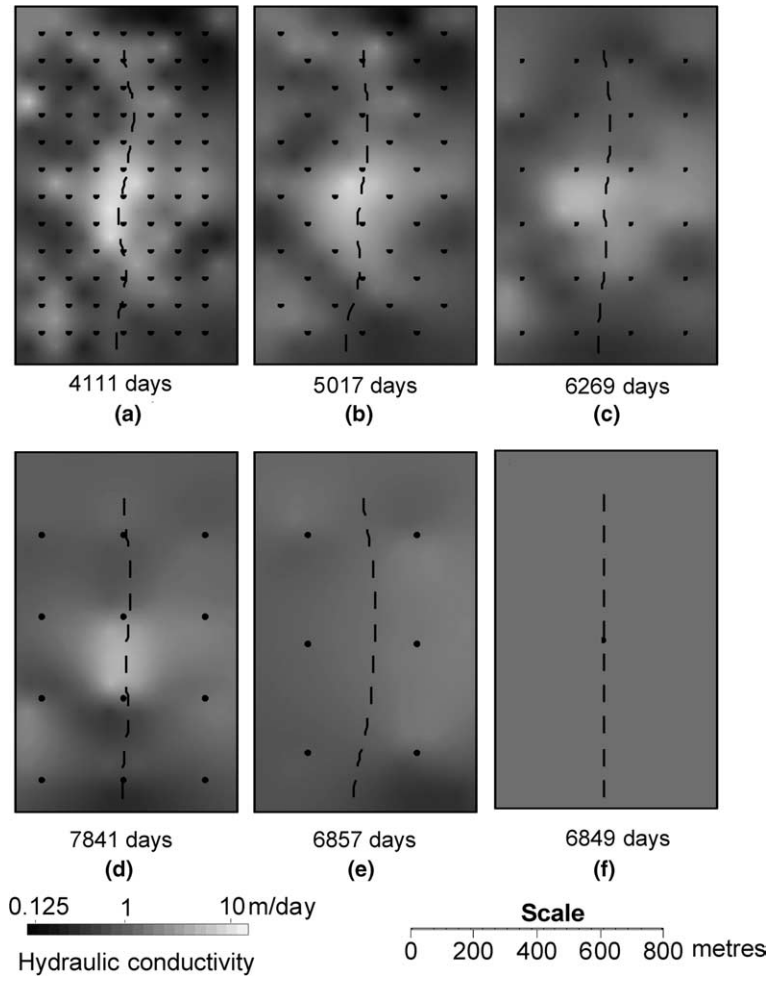

Fig. 2. Hydraulic conductivity fields and particle exit times calculated by the model calibrated against heads measured in (a) 84 wells, (b) 42 wells, (c) 24 wells, (d) 12 wells, (e) 6 wells, (f) 1 well.

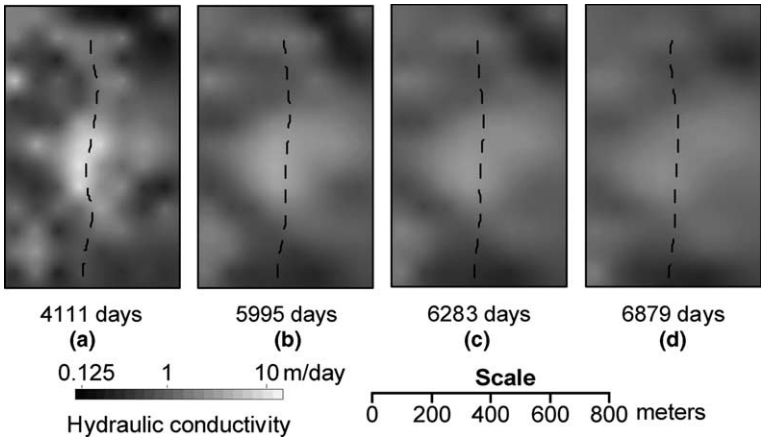

Fig. 3. Hydraulic conductivity fields and particle exit times calculated by the model calibrated against head measurements in 84 wells with (a) no noise; (b) noise with a standard deviation of 0.05 m; (c) noise with a standard deviation of 0.1 m; (d) noise with a standard deviation of 0.2 m. The locations of the 84 wells are shown in Fig. 2a.

the zonation pattern chosen for this exercise; boundaries for three (non-contiguous) zones were drawn against the background of the true field in an attempt to ensure that zone boundaries coincided, as much as possible, with areas of distinctly different conductivities in the true field (a luxury not available in normal modeling practice).

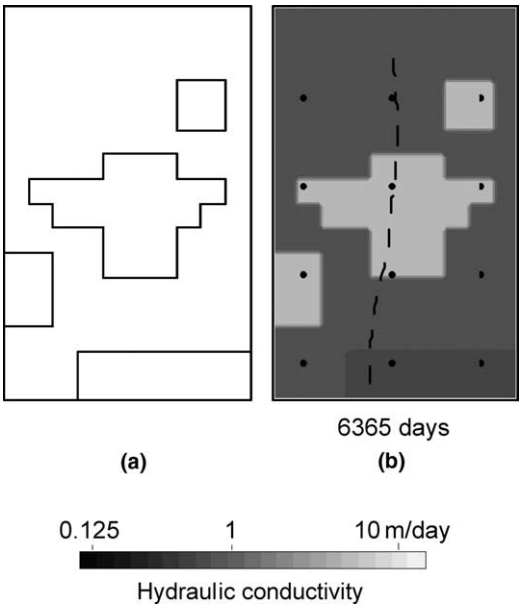

Fig. 4. Model domain showing (a) disposition of zones of parameter constancy; (b) calibrated hydraulic conductivity field based on this zonation; particle path and measurement well locations are also shown.

Estimation of the three zone-based hydraulic conductivities was undertaken using Eq. (1) on the basis of a calibration dataset comprised of 12 measured heads with no noise added to these heads. The estimated hydraulic conductivity field is shown in Fig. 4b; the root mean square misfit between model outputs and field measurements was 0.065 m, this being the minimum that could be achieved using this parameterization scheme. Particle travel time to the lower boundary was calculated as 6365 days, demonstrating, once again, the potential error associated with model predictions as an outcome of the loss of detail incurred through the calibration process.

This exercise was repeated with 5 and then 8 non contiguous zones. While the root mean square fit improved slightly to 0.063 m and 0.039 m respectively, the particle travel time to the lower boundary remained similar at 6466 days and 6249 days respectively.

We conclude this sub-section by pointing out that it would be wrong to draw the conclusion that travel times predicted by calibrated models are always larger than those which occur in reality. Indeed, other numerical experiments that we have undertaken using the same model framework, but with different "true" hydraulic conductivity fields, have demonstrated that the opposite can occur just as easily. What is constant in all cases however is that the calibration process fails to reproduce the detail of the true field. When an estimated field is then used to make a prediction that is as sensitive to the existence and disposition of hydraulic property heterogeneity as travel time predictions are, that prediction has the potential to be considerably in error as a direct result of that loss of detail.

3.4. The resolving power of calibration

The resolution matrix was discussed in Section 2.3 of this paper. There it was pointed out that each row of this matrix contains the weights by which true hydraulic properties at different locations within the model domain are multiplied in order to calculate the hydraulic property value assigned to a particular point of that domain. In the present context, where spatial parameterization is achieved through the use of pilot points, each row of the resolution matrix represents the weights by which true hydraulic properties at pilot point locations are multiplied in order to obtain the estimated hydraulic conductivity at the pilot point pertaining to that row. (Actually, this is not strictly correct as our "true field" was not generated on the basis of pilot points, but was actually generated on a cell-by-cell basis by our multi-Gaussian simulator; nevertheless, because of the large number of pilot points used in the calibration process, this is a close enough approximation for the purposes of this discussion. Numerical experiments have demonstrated that errors incurred through the making of this approximation are slight; see Moore [46].)

The spatial averaging that is inherent in model calibration is best demonstrated if elements of the resolution matrix are plotted in a spatial setting. This is done in Fig. 5, which depicts elements of the resolution matrix calculated on the basis of a calibration dataset comprised of 12 wells (shown as circles in that figure), for a number of specific pilot points. Model calibration was achieved using "smoothing regularization". Each of the four plots comprising this figure depicts components of one row of the resolution matrix; the pilot point corresponding to each such row is marked in the figure as a cross. Each element of the respective row of the resolution matrix is plotted at the location of its pertinent pilot point; components pertaining to these elements are then contoured. Thus each of these contour maps represents the "weighting function" through which true hydraulic properties over the model domain are averaged to obtain the estimated hydraulic property at the pilot point to which the map pertains.

The maps of Fig. 5 possess important features that are relevant not just to the present calibration exercise, but to groundwater model calibration in general. Firstly, the peak value of the weighting factor is nowhere greater than 0.14 for these particular pilot points, providing a graphic illustration of the fact that hydraulic conductivities estimated through the calibration process are representative of no one point in space, but are rather a spatial average of true hydraulic conductivities occurring throughout the model domain; furthermore, some of the averaging weights involved in this process are

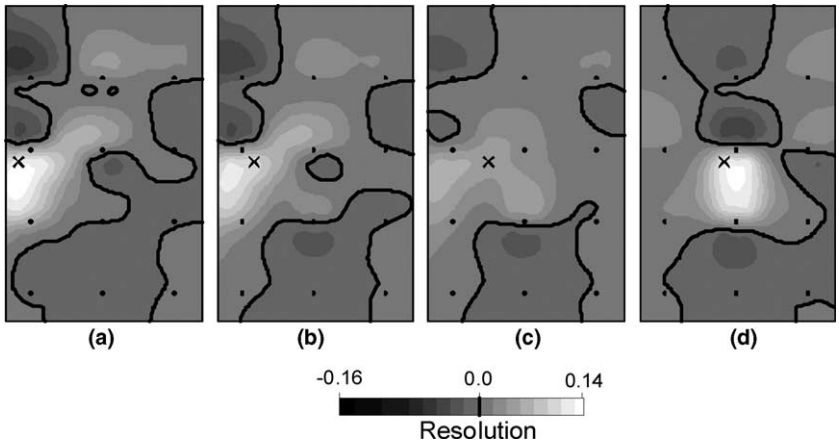

Fig. 5. Each of these pictures depicts the elements of one row of the resolution matrix; the pilot point to which the row pertains is shown as a cross. The zero contour is enhanced.

actually negative. Secondly (and perhaps surprisingly), the largest positive contribution to the estimated hydraulic conductivity at any particular point within the model domain does not necessarily come from the immediate vicinity of that point. Rather, in the present case at least, it comes from the area between the nearest pair of measurement wells (in the direction of groundwater flow). Thirdly, there are many locations within the model domain that do not contribute significantly to the hydraulic conductivity estimated for any pilot point (predominantly places that do not lie between pairs of measurement wells along lines of head gradient). Thus the hydraulic conductivities in these areas are simply unsampled through the calibration process. Other numerical studies that we have undertaken indicate that large, and important, hydraulic property anomalies could exist in such areas, and be virtually undetected by the calibration process.

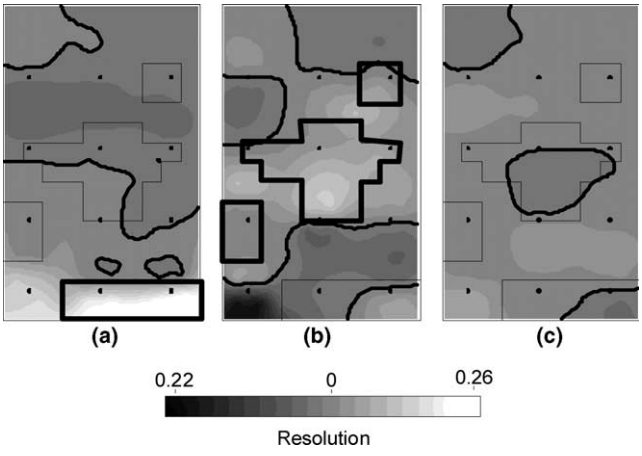

Fig. 6. Each of these pictures depicts the elements of one row of the zonal resolution matrix; the zone to which the row pertains is highlighted in figures (a) and (b), and the remainder of the model domain which comprises the third zone is shown in figure (c). The zero contour is enhanced. (Note that for the central figure the parameter zone is non-contiguous, being comprised of three spatially separate subzones.)

Fig. 6 shows similar plots to those provided in Fig. 5. However these pertain to each of the three rows of the "zonal resolution matrix" defined in Eq. (13); recall that each row of this matrix depicts the weights by which "true parameter values" are multiplied in formulation of zonal parameter values. It is obvious from this figure that estimated zonal parameter values do not represent average hydraulic properties within each respective zone. While, for each of the zones illustrated in Fig. 6, the bulk of the contribution to the estimated zonal property originates from within the zone, the averaging kernel is far from uniform within the zone. Furthermore, substantial contributions are made by parts of the model domain existing outside the pertinent zone.

3.5. Eigen-analysis of pre- and post-calibration parameter covariance matrices

Section 2.5 of this paper discussed eigenvalue/eigenvector analysis as a mechanism for comparing $C(\mathbf{p})$ with $C(\underline{\mathbf{p}})$.

Fig. 7 shows eigenvalues, arranged in order of decreasing value, calculated for $C(\mathbf{p})$ and $C(\underline{\mathbf{p}})$ where calculation of $\underline{\mathbf{p}}$ was achieved through regularized inversion ("preferred value" option) based on a measurement dataset comprised of heads (with no noise added) in 12 observation wells. As has been discussed, the spatial variation of estimated hydraulic conductivity $\underline{\mathbf{p}}$ is characterised by 104 pilot points, the locations of which are shown in Fig. 1a. These same points were used to sample the true hydraulic conductivity field $\mathbf{p}$ for the purpose of calculating the eigenvectors and eigenvalues of $C(\mathbf{p})$. Because this sampled $\mathbf{p}$ has 104 components, $C(\mathbf{p})$ has 104 eigenvalues. Thus spatial variation of $\mathbf{p}$ can be characterised by independent variation of 104 random variables in parameter space; these individual

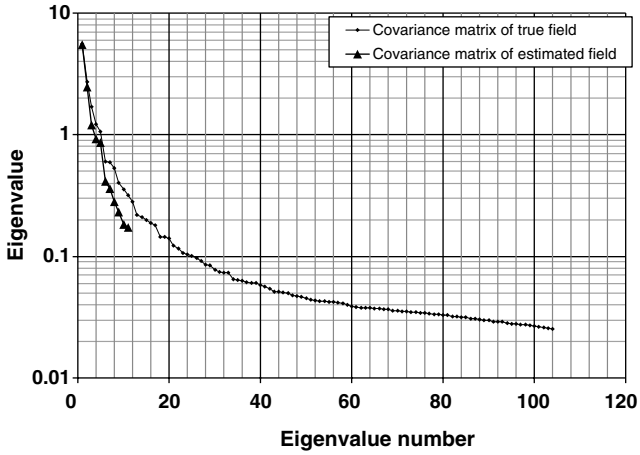

Fig. 7. Eigenvalues of true and estimated covariance matrices for the case where 12 observations are used for model calibration.

random values pertain to linear combinations of the 104 elements of $\mathbf{p}$ corresponding to the respective eigenvectors of $C(\mathbf{p})$. The components of the normalized eigenvectors corresponding to the highest three eigenvalues and the lowest three eigenvalues of $C(\mathbf{p})$ are shown in Fig. 8. It is readily apparent that the directions in parameter space of highest variability for the true parameter field correspond to broad scale variations (i.e. variations of low spatial frequency), while directions of smallest variability correspond to small scale variations (i.e. variations of high spatial frequency).

In stark contrast to $C(\mathbf{p})$, the spectrum of $C(\underline{\mathbf{p}})$ only has 12 non-zero eigenvalues, this being a direct outcome of the fact that only 12 observations are used for model calibration and that $C(\mathbf{p})$ is thus singular with a rank of 12. Furthermore, as shown in Fig. 7, variability in $C(\underline{\mathbf{p}})$ decreases more rapidly with eigenvalue number than for $C(\mathbf{p})$. Components of the normalized eigenvectors corresponding to the highest and lowest three eigenvalues of $C(\underline{\mathbf{p}})$ are shown in Fig. 9.

It is once again apparent from this example that most of the "fine detail" present in the true parameter field is lost in the calibrated field. This results from the fact that variability in the estimated field is restricted to only as many independent directions in parameter space as there are observations; furthermore, variability in those directions corresponding to non-zero eigenvalues is attenuated, with increasing attenuation occurring in

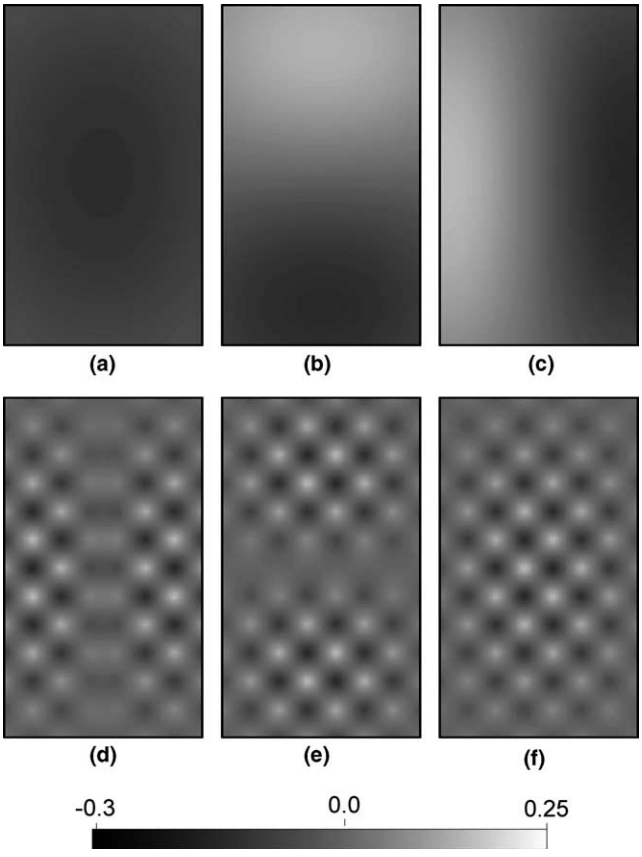

Fig. 8. Components of normalised eigenvectors (plotted at locations of corresponding parameters) of the true covariance matrix corresponding to (a–c) the largest three eigenvalues, and (d–f) the smallest three eigenvalues.

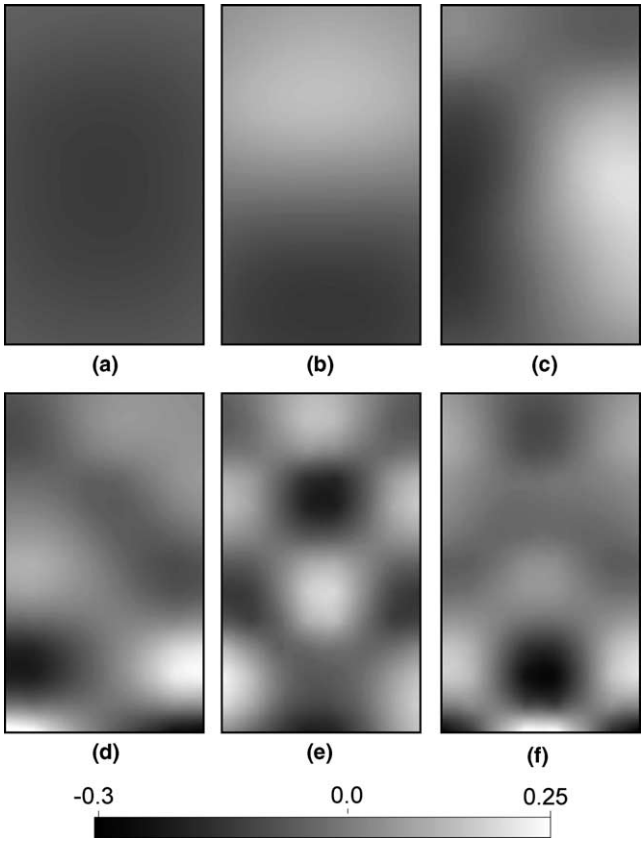

Fig. 9. Components of normalised eigenvectors (plotted at locations of corresponding parameters) of the covariance matrix of the estimated hydraulic conductivity field based on 12 observations, corresponding to (a–c) the largest three eigenvalues and (d–f) the smallest three eigenvalues.

directions that reflect finer system detail. Thus, conceptually at least, calibration can be viewed as a kind of low pass spatial filter, in which low spatial frequencies represented in the true hydraulic conductivity field are allowed to pass through to the estimated field, whereas higher spatial frequencies are either attenuated, or lost completely.

Consideration of the calibration process from a spectral point of view allows a sharp philosophical divide to be drawn between calibration based on a set of user-supplied zones on the one hand, and calibration based on a large number of parameters estimated through regularized inversion on the other. Earlier in this paper, model calibration based on 12 observations was undertaken using three zones of parameter constancy. In this zone-based calibration the fit between model outputs and field data was not as good as that achieved through regularized inversion. Depending on the modeling circumstances, this may not necessarily be a bad thing. However of larger concern when using zone-based calibration is the fact that the estimated parameter field can have only as many directions of variability in parameter space as there are zones, three in that case. In contrast, when calibration is undertaken using regularized inversion, the estimated field is characterized by as many degrees of freedom as the data allows (12 in this case). It is thus apparent that regularized inversion allows a greater amount of heterogeneity information to be extracted from a given calibration dataset than zone based parameterization, as use of the latter device can preclude some of this measurement information being reflected in the estimated field by placing inappropriate restrictions on the number of degrees of freedom through which this information can express itself.

3.6. Inferring the geostatistical structure of the true hydraulic properties

For the synthetic calibration problem of Fig. 1, we tested the compatibility of a suite of possible true field variograms with the estimated field depicted in Fig. 2d; recall that this field was calculated using "preferred value" regularization on the basis of 12 observation wells (with no noise added to the heads in these wells). Fig. 10 shows the results of this analysis for hypothesized variograms with sills ranging from 0.01 to 1, and variogram a values ranging from 25 m to 500 m, where a is the coefficient of distance in the exponent of Eq. (19). Our 'true' field variogram (which has a sill of 0.2 and an a value of 200 m) falls well within the 80% confidence interval depicted in this figure.

Of note in Fig. 10 is the fact that the variability of variogram sills that are admissible at the 80% confidence level decreases to a minimum at a variogram range of about 100 m, and increases for variogram ranges above and below this value. For variogram ranges that are very small, spatial variability of the true hydraulic property field is virtually undetectable through the calibration process as Fig. 7 shows; hence the amplitude of spatial variation becomes less and less consequential as variogram range is reduced. For high variogram ranges, higher sills are also possible because the variogram range is comparable with (or larger than) the size of the model domain itself. The calibration process is thus able to "see" only that variability of the true field which occurs within the domain of the model, with the "effective range" of the variogram thus reduced to the size of the model domain, the "effective sill" of the variogram is limited to that variability which can occur within that effective range; thus higher "true sills" can be compatible with the calibrated model.

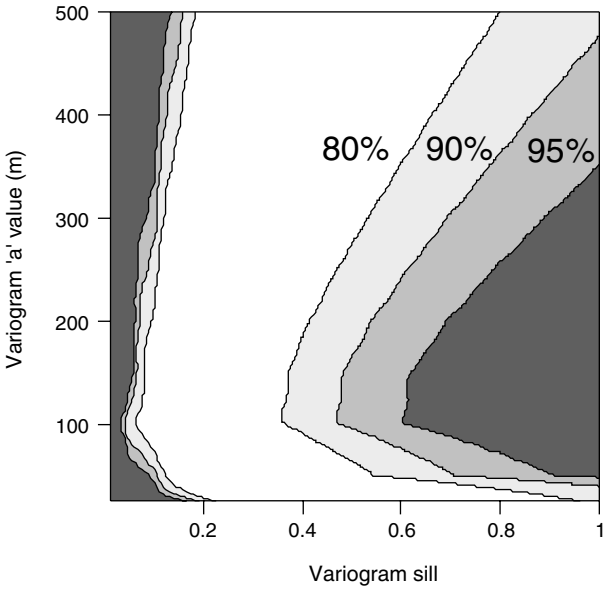

Fig. 10. Chi-square confidence regions for true field variogram sill and a values consistent with calibrated hydraulic conductivity field of Fig. 2d.

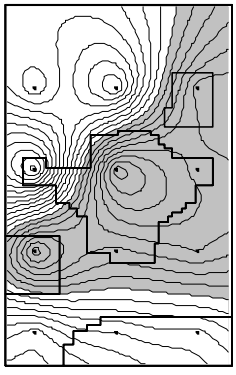

Fig. 11. Contoured residuals for zone-based calibration; the shaded area depicts that part of the model domain where residuals are negative. Observation wells are shown as filled circles.

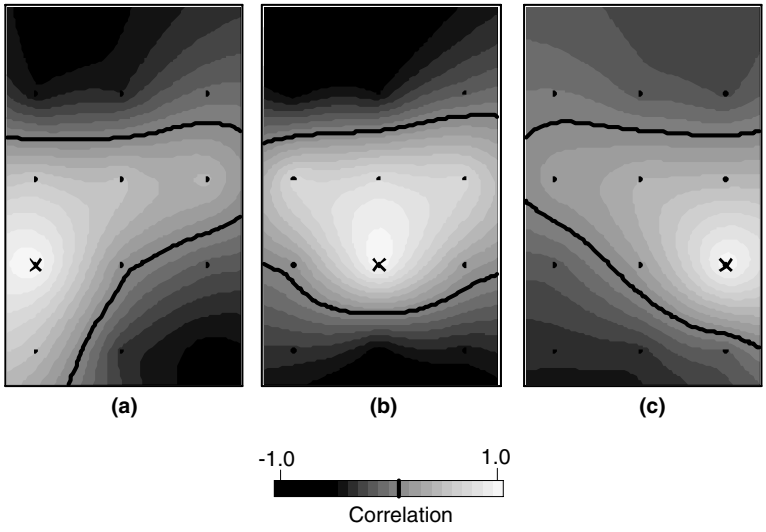

Fig. 12. Residual correlation coefficients for three head measurement points, each such head measurement location being depicted with a cross.

3.7. Spatial correlation of residuals

Fig. 11 shows contoured residuals for the zone-based calibration of Fig. 4. Recall that the calibration dataset in this case consisted of 12 measurement heads. No noise was added to the data; however structural noise with an rms value of 0.065 m was incurred through the use of zones as a parameterization device. It is hard to reconcile this picture (which typifies the spatial distribution of residuals encountered in most cases of model calibration) with the notion of zero spatial correlation of residuals as suggested by some authors as being indicative of a good calibration, as discussed in a previous section. To further illustrate the point that spatial correlation of residuals is in fact to be expected, Eq. (18b) can be used to calculate the spatial correlation structure of model-to-measurement residuals for our synthetic case. Fig. 12a–c depicts contoured correlation coefficients between residuals at each of three measurement wells (these wells are marked by a cross in each case) and all of the other wells used in the calibration process. As expected, a correlation coefficient of 1 exists at the measurement well to which each plot pertains; however the fact that correlation is significantly non-zero at all other measurement well locations (being positive at some locations and negative at others) is obvious from these figures.

4. Discussion and conclusions

Rightly or wrongly, most modelling exercises (often pursued at great expense) seek to produce a ''calibrated model''. Implicit in this endeavour is the pursuit of a unique solution to the inverse problem of model calibration, and thus the use of some kind of regularization device.

It is apparent from the examples and theory presented in this paper that, whether regularization is undertaken manually through use of zones of piecewise constancy, or as an inherent part of the inversion methodology itself, an estimated property value assigned to any point within a model domain through the calibration process is unlikely to be equal to the true hydraulic property value at that point. In fact, calibration-determined hydraulic property values can be locally quite erroneous, especially where data density is low. Furthermore, no matter how regularization is achieved, detail in the estimated parameter field cannot exist beyond a level that is sustainable by the calibration dataset. Fortunately, information made available as a by-product of the regularized inversion process allows an assessment to be made of the degree to which such detail is potentially lost. Sadly, however, this loss of detail may seriously degrade a model's ability to make accurate predictions of those aspects of system behaviour that are sensitive to this lost detail. In fact, such predictions may be seriously in error. This is the cost of obtaining a unique solution to the inverse problem!

Examination of the spatial covariance structure of the calibrated field has allowed insights to be gained into the nature of the calibration process. In this paper an analogy has been drawn between model calibration and low pass spatial filtering. The increasing attenuation of spatial detail with decreasing spatial wavelength (i.e. increasing spatial frequency) is an outcome of the fact that the calibration process is actually a process of spatial integration as evinced by the role of the resolution matrix in the calibration process. As discussed by many

authors (e.g. Stearns [52]) integration is actually a form of low-pass filtering. Hence it is to be expected that an estimated parameter field will possess a much lower spatial frequency content than the true field from which it was derived.

An analogy can be drawn between model calibration and sampling theory. Sampling theory tells us that sample points need to be spaced at an interval that is less than half the smallest wavelength contained in the sampled field if a valid reconstruction of that field from its samples is to take place. As the number of observations comprising any calibration dataset is finite, there is thus a lower bound on spatial wavelengths that can be represented in an estimated field. Because observation points are normally separated by a greater distance than half the wavelength of the maximum spatial frequency represented in the true hydraulic property field, shorter spatial wavelengths existing within this sampled field are "aliased" to appear as larger spatial wavelengths in the calibrated field.

The use of many parameters in conjunction with regularized inversion as a methodology for groundwater model calibration allows the calibration process to capture more information from a calibration dataset than that which can be captured using a limited number of zones of piecewise constancy. Looked at in the "space domain", this is an outcome of the fact that use of a multiplicity of parameters increases the chance that parameters exist where they need to exist, allowing the calibration process to introduce heterogeneity into the model domain where this is required in order to reproduce field measurements. The more parameters that are used (i.e. the more pilot points in our case), the more likely this is to be true. However the sampling theory analogy outlined above may allow us to place limits on the number of pilot points required for a particular calibration exercise if something is known about the degree of heterogeneity that is likely to exist within the true hydraulic property field, and hence the estimated field through application of Eqs. (9). Alternatively because the shortest spatial wavelength that can be represented in an estimated field is ultimately determined by the density of sample points (for example, observation wells), sampling theory allows us to determine how many pilot points will be required for adequate representation of spatial heterogeneity in fields that are estimated on the basis of a limited calibration dataset. Basically, a pilot point density that is commensurate with observation point density should be employed, with the spatial density of the former increasing where the spatial density of the latter increases (thereby allowing inference of a greater range of heterogeneity potentially existing in the real property field). However where observation point density is very large, there may not be a need for observation points to be matched "one for one" by pilot points, for if observation point spacing is less than half the shortest wavelength represented in the hydraulic property field, then its sample density is more than adequate; thus there is no requirement that pilot point density should be commensurate with it.

Knowing that the calibration process can incur a considerable loss of detail, it follows that model predictions must be viewed with great caution in situations where these predictions depend on system detail beyond that which can be inferred through the calibration process. Contaminant movement is an obvious example of a process that is highly dependent on system detail. Other examples abound, for example the transient response of water levels to climatic or pumping extremes, groundwater-surface water interaction, etc. Predictions of these types (and others), made with parameter fields estimated through calibration against an always incomplete dataset may be seriously in error. An assessment of the possible magnitude of this error should be an essential component of model deployment. Means by which this can be achieved are presently available through methods such as calibration-constrained Monte Carlo analysis (see for example Gomez Hernandez [26], Ramarao et al. [51] and LaVenue et al.'s [37], Capilla et al. [12], Hu [33]). Unfortunately the computational cost of these methods is often very high. An alternative, and much less computationally expensive, method, is presented in [45].

It is anticipated that the use of regularized inversion as a methodology for groundwater model calibration will increase dramatically in the future, for the calibrated parameter field produced as an outcome of its use is more responsive to the information content of an often costly environmental dataset, thus "doing justice" to the cost of acquiring that data in the first place. Furthermore, information made available to the modeller as a by-product of regularized inversion, some of which has been discussed in this paper, allows the modeller to assess a calibrated model not just in terms of the detail which it has captured, but also in terms of the detail that it has failed to capture. In addition to this, the computational cost of undertaking regularized inversion, and of including many parameters in the formulation of the inverse problem to take maximum advantage of the benefits that regularized inversion has to offer, need not be great. The use of adjoint state methods for calculation of derivatives where parameters outnumber observations can reduce this burden considerably (see, for example, Townley and Wilson [56] and Clemo et al. [19]). So too can the use of "SVD-assisted" parameter estimation [55], which combines the benefits of subspace regularization with that of Tikhonov regularization in a single, computationally efficient, hybrid scheme that dispenses with the need to calculate derivatives of all model outputs with respect to all parameters during all iterations of a nonlinear inversion process.

References

[1] Akaike H. A new look at the statistical model identification. IEEE Trans Automat Contr 1974;AC19(6):716–23.

[2] Anderman ER, Hill MC. MODFLOW-2000, the US Geological Survey modular ground-water model—documentation of the ADVective-Transport observations (ADV2) Package: US Geological Survey Open-File Report 01-54, 2001, 69p.

[3] Anderson MP, Woessner WW. Applied groundwater modeling—simulation of flow and advective transport. Academic Press; 1992.

[4] Babiloni F, Carducci F, Cincotti F, Del Gratta C, Pizzella V, Romani GL, et al. Linear inverse source estimate of combined EEG and MEG data related to voluntary movements. Hum Brain Map 2001;14(4):197–209.

[5] Backus G, Gilbert F. Numerical application of a formalism for geophysical inverse problems. Geophys J Roy Astr S 1967;13:247–76.

[6] Backus G, Gilbert F. The resolving power of gross earth data. Geophys J Roy Astr S 1969;16:169–205.

[7] Backus GE, Gilbert F. Uniqueness in the inversion of inaccurate gross earth data. Philos Tr R Soc S A 1970;266(1173):123–92.

[8] Bard Y. Nonlinear parameter estimation. New York: Academic Press; 1974.

[9] Bertete-Aguirre H, Cherkaev E, Oristaglio M. Non-smooth gravity problem with total variation penalization functional. Geophys J Int 2002;149(2):499–507.

[10] Braud I, Obled C. On the use of empirical orthogonal functions analysis in the simulation of random fields: stochastic hydrol. Hydraul 1991;5(2):125–34.

[11] Butler Jr JJ, McElwee CD, Bohling GC. Pumping tests in networks of multilevel sampling wells: Motivation and methodology. Water Resour Res 1999;35(11):3553–60.

[12] Capilla JE, Rodrigo J, Gómez-Hernández JJ. Simulation of non-Gaussian hydraulic conductivity fields honoring piezometric data and integrating soft and secondary information. Math Geol 1999;31(7):907–27.

[13] Carrera J, Neuman SP. Estimation of aquifer parameters under transient and steady-state conditions. 2. Uniqueness, stability, and solution algorithms. Water Resour Res 1986;22(2):211–27.

[14] Certes C, de Marsily G. Application of the pilot points method to the identification of aquifer transmissivities. Adv Water Resour 1991;14(5):284–300.

[15] Christensen S, Cooley RL. Evaluation of prediction intervals for expressing uncertainties in groundwater flow model predictions. Water Resour Res 1996;35(9):2627–39.

[16] Christensen S, Cooley RL. Evaluation of confidence intervals for a steady-state leaky aquifer model. Adv Water Resour 1999; 22(8):807–17.

[17] Christensen S, Cooley RL. Experiences from the testing of a theory for modeling groundwater flow in heterogeneous media. Proceedings of ModelCARE'2002, Prague, Czech Republic, 17–20 June 2002. IAHS publ. no. 2002; p. 22–7.

[18] Christensen S, Moore C, Doherty J. Comparison of stochastic and regression based methods for quantification of predictive uncertainty of model-simulated wellhead protection zones in heterogeneous aquifers. ModelCARE'2005 conference paper, The Hague, Netherlands, 6–9 June 2005.

[19] Clemo T, Michaels P, Lehman RM. Transmissivity Resolution Obtained from the Inversion of Transient and Pseudo-Steady Drawdown Measurements. MODFLOW and More 2003: Understanding through Modeling—Conference proceedings 2003; Poeter,Zheng, Hill & Doherty. Downloadable from: www.mines.edu/igwmc.

[20] Cooley RL. A theory for Modeling Ground-Water Flow in Heterogeneous Media. Reston, VA, US Geological Survey, Professional Paper 1679, 2004.

[21] Delay F, Buoro A, Marsily Gde. Empirical orthogonal function analysis applied to the inverse problem in hydrogeology: Evaluation of uncertainty and simulation of new solutions. Math Geology 2001;33(8):927–49.

[22] Doherty J. Groundwater model calibration using pilot points and regularization. Ground Water 2003;41(2):170–7.

[23] Doherty J. Version 5 of PEST Manual. Watermark Numerical Computing, Brisbane Australia; 2003. Downloadable from: http://www.sspa.com/pest.

[24] Draper NR, Smith H. Applied regression analysis. Third ed. New York: John Wiley & Sons, Inc.; 1998.

[25] Engl HW, Hanke M, Neubauer A. Regularization of inverse problems. Kluwer Academic Publishers; 1996.

[26] Gómez-Hernández JJ, Sahuquillo A, Capilla JE. Stochastic simulation of hydraulic conductivity fields conditional to both hydraulic conductivity and piezometric data. 1. Theory J Hydrol 1997;20:162–74.

[27] Gorokhovski VM. Problem-dependence of ground-water model identification: significance, extent, and treatment. Ground Water 1996;34(3):461–9.

[28] Guadagnini A, Neuman SP. Nonlocal and localized analyses of conditional mean steady state flow in bounded, randomly uniform domains, 1, theory and computational approach. Water Resour Res 1999;35(10):2999–3018.

[29] Haber E. Numerical strategies for the solution of inverse problems. PhD thesis, University of British Columbia, Canada, 1997.

[30] Harbaugh AW, Banta ER, Hill MC, McDonald MG. The US Geological Survey Modular Ground-Water Model—User Guide to Modularization Concepts and the Ground-Water Flow process. US Geological Survey Open File Report 00-92. Reston, Virginia, 2000.

[31] Hill MC. Methods and guidelines for effective model calibration. US Geological Survey, Water-Resources Investigations Report 98-4005, 1998.

[32] Hill MC, Osterby O. Determining extreme parameter correlation in Ground Water Models. Ground Water 2003;41(4):420–30.

[33] Hu LY. Gradual deformation and iterative calibration of gaussian-related stochastic models. Math Geol 2000;32(1):87–108.

[34] Jackson DD. Interpretation of inaccurate, insufficient and inconsistent data. Geophys J Roy Astr S 1972;28:97–109.

[35] Kitanidis PK. The minimum structure solution to the inverse problem. Water Resour Res 1997;33(10):2263–72.

[36] Koch KR. Parameter estimation and hypothesis testing in linear models. Berlin: Springer-Verlag; 1987.

[37] LaVenue AM, RamaRao BS, de Marsily G, Marietta MG. Pilot point methodology for automated calibration of an ensemble of conditionally simulated hydraulic conductivity fields. 2. Appl Water Resour Res 1995;31(3):495–516.

[38] Lines LR, Treitel ST. A review of least-squares inversion and its application to geophysical problems. Geophys Prospect 1984; 32(2):159–86.

[39] Lutkenhoner B, Menendez RGDP. The resolution-field concept. Electroen Clin Neuro 1997;102(4):326–34.

[40] Mackie RL, Madden TR. Three-dimensional magnetotelluric inversion using conjugate gradients. Geophys J Int 1993; 115:215–29.

[41] Mantoglou A. Estimation of heterogeneous aquifer parameters from piezometric data using ridge functions and neural networks. Stoch Env Res Risk A 2003;17(5):339–52.

[42] Menendez RGDP, Andino SLG, Lutkenhoner B. Figures of merit to compare distributed linear inverse solutions. Brain Topogr 1996;9(2):117–24.

[43] Menke W. Geophysical data analysis: discrete inverse theory. Academic Press Inc; 1984.

[44] Mikhail EM with contributions by Ackerman F. Observations and Least Squares. New York: IEP—A Dun-Donnelley publisher; 1976.
[45] Moore C, Doherty J. The role of the calibration process in reducing model predictive error. Water Resour Res 2005.
[46] Moore C. The use of regularized inversion in groundwater model calibration and prediction uncertainty analysis. PhD Thesis. University of Queensland, Australia, 2005.
[47] Murray Darling Basin Commission. Murray–Darling Basin Commission Groundwater Flow Modeling Guideline. Aquaterra Consulting Pty Ltd. Project 125. 2000. Downloadable from: http://www.mdbc.gov.au/publications/pdf/model_guide.pdf.
[48] Nolet G, Montelli R, Virieuz J. Explicit, approximate expressions for the resolution and a posteriori covariance of massive tomographic systemes. Geophys J Int 1999;138:36–44.
[49] Poeter EP, Hill MC. Documentation of UCODE, a computer code for universal inverse modeling. US Geological Survey Water-Resources Investigations 1998; Report 98-4080, 122p.
[50] Portniaguine O, Zhdanov MS. Focusing geophysical inversion images. Geophysics 1999;64(3):874–87.
[51] RamaRao BS, LaVenue AM, de Marsily G, Marietta MG. Pilot point methodology for automated calibration of an ensemble of conditionally simulated hydraulic conductivity fields, 1, theory and computational experiments. Water Resour Res 1995; 31(3):475–93.
[52] Stearns SD. Digital signal analysis. NJ: Hayden Book Company, Inc; 1975.
[53] STOWA. Good Modeling Practice Handbook. STOWA 1999; Report 99-05 file gmpuk.pdf. Downloable from: http://water-land.net/riza/aquest.
[54] Tikhonov AN, Arsenin VY. Solution of ill-posed problems. Washington DC: VH Winston and Sons; 1977.
[55] Tonkin M, Doherty J. A hybrid regularised inversion methodology for highly parameterised environmental models. Water Resour Res 2005.
[56] Townley LR, Wilson JL. Computationally efficient algorithms for parameter estimation and uncertainty propagation in numerical models of groundwater flow. Water Resour Res 1985;21(12): 1851–60.
[57] Vasco DW, Akhil Datta-Gupta. Asymptotic solutions of solute transport: A formalism for tracer tomography. Water Resour Res 1990;35(1):1–16.
[58] Vasco DW, Datta-Gupta A, Long JCS. Resolution and uncertainty in hydrologic characterization. Water Resour Res 1997;33(3):379–97.
[59] Vasco DW, Petersen JE, Majer EL. Resolving seismic anisotropy: sparse matrix methods for geophysical inverse problems. Geophysics 1998;63(3):970–83.
[60] Vasco DW. Estimation of flow properties using surface deformation and head data: A trajectory-based approach. WRR 2004;40:W10104. doi:10.1029/2004WR003272.
[61] Vecchia AV, Cooley RL. Simultaneous confidence and prediction intervals for nonlinear regression models with application to a groundwater model. Water Resour Res 1987;23(7):1237–50.
[62] Vogel CR, Oman ME. Fast total variation based reconstruction of noisy, blurred images. IEEE T Image Process 1998;7(6):813–24.
[63] Yao ZS, Roberts RG. A practical regularization for seismic tomography. Geophys J Int 1999;138:293–9.
[64] Yeh TCJ, Liu SY. Hydraulic tomography: Development of a new aquifer test method. Water Resour Res 2000;36(8):2095–105.