Advances in Water Resources 63 (2014) 22–37

Contents lists available at ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

Review

Inverse methods in hydrogeology: Evolution and recent trends

Haiyan Zhou [a,b,*], J. Jaime Gómez-Hernández [a], Liangping Li [a,b]

[a] Research Institute of Water and Environmental Engineering, Universitat Politècnica de València, 46022 Valencia, Spain
[b] Center for Petroleum and Geosystems Engineering Research, University of Texas at Austin, Austin, TX 78712, USA

ARTICLE INFO

Article history:
Received 4 July 2013
Received in revised form 24 October 2013
Accepted 28 October 2013
Available online 7 November 2013

Keywords:
Heterogeneity
Parameter identification
Data assimilation
Uncertainty
Groundwater modeling

ABSTRACT

Parameter identification is an essential step in constructing a groundwater model. The process of recognizing model parameter values by conditioning on observed data of the state variable is referred to as the inverse problem. A series of inverse methods has been proposed to solve the inverse problem, ranging from trial-and-error manual calibration to the current complex automatic data assimilation algorithms. This paper does not attempt to be another overview paper on inverse models, but rather to analyze and track the evolution of the inverse methods over the last decades, mostly within the realm of hydrogeology, revealing their transformation, motivation and recent trends. Issues confronted by the inverse problem, such as dealing with multiGaussianity and whether or not to preserve the prior statistics are discussed.

Contents

1. Introduction

Mathematical modeling of subsurface flow and mass transport is needed, for instance, for groundwater resources management

* Corresponding author at: Center for Petroleum and Geosystems Engineering Research, University of Texas at Austin, Austin, TX 78712, USA.

E-mail addresses: haiyanzhou@utexas.edu (H. Zhou), jaime@dihma.upv.es (J.J. Gómez-Hernández), liangpingli@utexas.edu (L. Li).

0309-1708/$ - see front matter
http://dx.doi.org/10.1016/j.advwatres.2013.10.014

or for contaminant remediation. The forward model requires specification of a variety of parameters, such as, hydraulic conductivity, storativity and sources or sinks together with initial and boundary conditions. However, in practice, it is impossible to characterize the model exhaustively from sparse data because of the complex hydrogeological environment; for this reason, inverse modeling is a valuable tool to improve characterization. Inverse models are used to identify input parameters at unsampled locations by incorporating observed model responses, e.g., hydraulic conductivities are derived based on hydraulic head and/or solute concentration data. Deriving model parameters from model state observations is common in many other disciplines, such as petroleum engineering, meteorology and oceanography. This work mostly focuses on inverse methods used in hydrogeology.

1.1. The forward problem and the inverse problem

The forward problem involves predicting model states, e.g., hydraulic head, drawdown and solute concentration, based on a prior model parameterization. Combining mass conservation and Darcy's laws, the forward groundwater flow model in an incompressible or slightly compressible saturated aquifer can be written as [1]

$$\nabla \cdot (K\nabla h) = S_s \frac{\partial h}{\partial t} + Q \tag{1}$$

subject to initial and boundary conditions, where $\nabla\cdot$ is the divergence operator $\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial z}\right)$, ∇ is the gradient operator $\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)^T$, K is hydraulic conductivity ($\mathrm{L\,T^{-1}}$), h is hydraulic head (L), S_s is specific storage ($\mathrm{L^{-1}}$), t is time (T), and Q is source or sink ($\mathrm{T^{-1}}$). The differential equation governing non-reactive transport in the subsurface is:

$$\phi \frac{\partial C}{\partial t} = -\nabla \cdot (qC) + \nabla \cdot (\phi \mathrm{D} \nabla C) \tag{2}$$

subject to initial and boundary conditions, where C is the concentration of solute in the liquid phase ($\mathrm{M\,L^{-3}}$), ϕ is porosity (–), D is the local hydrodynamic dispersion tensor ($\mathrm{L^2\,T^{-1}}$) usually defined as $D_i = \alpha_i|q| + D_m$ where α_i refers to the longitudinal and transverse dispersivities (L) and D_m is the molecular diffusion coefficient ($\mathrm{L^2\,T^{-1}}$), and q is the Darcy velocity ($\mathrm{L\,T^{-1}}$) given by Darcy's law as $q = -K\nabla h$.

The inverse problem aims at determining the unknown model parameters by making use of the observed state data. In the early days of groundwater modeling, it was common to start with a prior guess of the model parameters, run the forward model to obtain the simulated states, and then enter in a manual loop iteratively modifying the parameters, and then running the forward model, until observed and simulated values were close enough so as to accept the model parameter distribution as a good representation of the aquifer. This "trial and error" method falls into the scope of "indirect methods" as opposed to the "direct methods" which do not require multiple runs of the forward model to derive the model parameters [2] as will be discussed below.

1.2. Why is the inverse problem necessary?

Sagar et al. [3] classified the inverse problem into five types according to the unknowns, i.e., model parameters, initial conditions, boundary conditions, sources or sinks and a mixture of the above. Most documented inverse methods fall into the first type, that is, they try to identify model parameters, which contribute largely to the model uncertainty due to the inherent heterogeneity of aquifer properties. Parameter identification is of importance considering the fact that no reliable predictions can be acquired without a good characterization of model parameters. Parameter identification is a broad concept here including not only the property values within facies but the facies distribution, or in other words, geologic features. The effect of geologic uncertainty in groundwater modeling is examined, for instance, by He et al. [4] in a real case study. Furthermore, data scarcity deteriorates the characterization of the model parameters and raises the uncertainty. Besides estimating aquifer parameters, the inverse methods also play a critical role in assessment of uncertainty for the predictions. Furthermore, the inverse problem might be used as a guide for data collection and the design of an observation network. The reader is referred to Poeter and Hill [5] who discussed the benefits of inverse modeling in depth. In this work we are mainly concerned with the uncertainty introduced by the unknown model parameters and thus the inverse methods that are used to characterize these parameters.

1.3. Why is the inverse problem difficult?

A problem is properly posed if the solution exists uniquely and varies continuously as the input data changes smoothly. However, most of the inverse problems in hydrogeology are ill-posed and they cannot be solved unless certain assumptions and constraints are specified. Ill-posedness may give rise to three problems: non-uniqueness, non-existence and non-steadiness of the solutions, among which non-uniqueness is the most common. Non-uniqueness primarily stems from the fact that the number of parameters to be estimated exceeds that of the available observation data. Another reason is that the observations are sometimes not sensitive to the parameters to be identified; in other words, the information content of the observations is very limited. For instance, hydraulic heads close to the prescribed head boundaries are more influenced by the boundaries than by the nearby hydraulic conductivities (i.e., the hydraulic heads are not so sensitive to the conductivities), and on the contrary, the hydraulic heads close to the prescribed flux boundaries are determined to a large extent by the hydraulic conductivities nearby [6].

A series of suggestions have been proposed to alleviate the ill-posedness:

1. Reduce the number of unknown parameters, e.g., using zonation, or collect more observation data so that the numbers of data and unknowns are balanced.
2. Consider the prior information or some other type of constraint to restrict the space within which parameters may vary.
3. Impose regularization terms to reduce fluctuations during the optimization iterations [2].
4. Maximize the sensitivity of observations to model parameters, for instance, by designing properly the observation network.
5. Minimize the nonlinearity in the model equation. Carrera and Neuman [6] argued that working with the logarithm of hydraulic conductivity reduces the degree of non-convexity during optimization. An alternative is to infer hydraulic conductivity using fluxes rather than heads as done by Ferraresi et al. [7], since the relationship between hydraulic conductivity and flux is linear (Darcy's law) while the relationship between hydraulic conductivity and head is nonlinear.

Detailed discussions on this subject can be found in [2,6,8,9] among others.

Besides the ill-posedness problem, computational burden is the second main hurdle for inverse problems [10]. There are several reasons for the high CPU time requirement. Since many inverse models are iterative, the forward model has to be run many times until an acceptable parameter distribution is obtained. The time