Journal of Applied Geophysics 68 (2009) 60–70

Contents lists available at ScienceDirect

Journal of Applied Geophysics

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

Simulated-annealing-based conditional simulation for the local-scale characterization of heterogeneous aquifers

B. Dafflon, J. Irving, K. Holliger *

Institute of Geophysics, University of Lausanne, CH-1015 Lausanne, Switzerland

ARTICLE INFO

Article history:
Received 31 October 2007
Accepted 19 September 2008

Keywords:
Data integration
Georadar
Simulated annealing
Stochastic methods
Porosity
Conditional simulation

ABSTRACT

Simulated-annealing-based conditional simulations provide a flexible means of quantitatively integrating diverse types of subsurface data. Although such techniques are being increasingly used in hydrocarbon reservoir characterization studies, their potential in environmental, engineering and hydrological investigations is still largely unexploited. Here, we introduce a novel simulated annealing (SA) algorithm geared towards the integration of high-resolution geophysical and hydrological data which, compared to more conventional approaches, provides significant advancements in the way that large-scale structural information in the geophysical data is accounted for. Model perturbations in the annealing procedure are made by drawing from a probability distribution for the target parameter conditioned to the geophysical data. This is the only place where geophysical information is utilized in our algorithm, which is in marked contrast to other approaches where model perturbations are made through the swapping of values in the simulation grid and agreement with soft data is enforced through a correlation coefficient constraint. Another major feature of our algorithm is the way in which available geostatistical information is utilized. Instead of constraining realizations to match a parametric target covariance model over a wide range of spatial lags, we constrain the realizations only at smaller lags where the available geophysical data cannot provide enough information. Thus we allow the larger-scale subsurface features resolved by the geophysical data to have much more due control on the output realizations. Further, since the only component of the SA objective function required in our approach is a covariance constraint at small lags, our method has improved convergence and computational efficiency over more traditional methods. Here, we present the results of applying our algorithm to the integration of porosity log and tomographic crosshole georadar data to generate stochastic realizations of the local-scale porosity structure. Our procedure is first tested on a synthetic data set, and then applied to data collected at the Boise Hydrogeophysical Research Site.

1. Introduction

A key control on groundwater flow and contaminant transport in the subsurface is the spatial distribution of hydrological properties. Accurate characterization of these properties is crucial for developing reliable numerical models of flow and transport, which are required to design effective and cost-efficient aquifer remediation and groundwater management strategies. It is well understood that spatial variability needs to be defined at a wide range of scales for effective modeling of hydrological phenomena (e.g., Sudicky and Huyakorn, 1991; Gelhar, 1993; Zheng and Gorelick, 2003; Hubbard and Rubin, 2005). However, conventional hydrological measurement techniques tend to lie at two ends of a spectrum in terms of resolution and sampling volume, leaving a significant gap in a range that is expected to contain particularly critical hydrological information. Whereas pumping and tracer tests tend to yield only gross average properties over a relatively large region, core samples and borehole logs yield high-resolution estimates of aquifer properties, but only along sparse 1-D profiles. Consequently, over the past two decades, much work has been done on the use of geophysical methods for aquifer characterization. Such methods can bridge the gap between the analysis of cores or logs and well tests, and have proven to be extremely useful not only for aquifer zonation but also for estimating the spatial distribution of hydrological parameters (e.g., McKenna and Poeter, 1995; Hyndman et al., 2000; Chen et al., 2001; Tronicke et al., 2002; Hubbard and Rubin, 2005; Kowalsky et al., 2005; Paasche et al., 2006). An important and still largely unresolved issue with the use of geophysical data in hydrological studies, however, is that of data integration. That is, how do we quantitatively integrate geophysical data with an existing database of other measurements to best constrain our knowledge of the spatial distribution of one or several target parameters?

The integration of different types of data for subsurface characterization has been a subject of much investigation in the petroleum industry, and has received increased attention in groundwater studies

* Corresponding author. Tel.: +41 21 692 44 03; fax: +41 21 692 44 05.
E-mail address: klaus.holliger@unil.ch (K. Holliger).

0926-9851/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jappgeo.2008.09.010

in recent years. Whereas much data integration and joint inversion work in the past has involved the determination of a single model of the subsurface parameters of interest, a number of recent efforts have focused on the creation of sets of multiple realizations that are consistent with all of the available data, and represent the uncertainty in our knowledge of the spatial distribution of subsurface properties (e.g., McKenna and Poeter, 1995; Bosch, 1999; Avseth et al., 2001; Caers et al., 2001; Mukerji et al., 2001; Ramirez et al., 2005; Hansen et al., 2006). The idea behind such conditional simulation approaches is that they can be used in combination with complex hydrological models to make predictions regarding groundwater flow and contaminant transport within a framework of uncertainty. Work with these methodologies has increased over recent years as a product of continually growing computer capacity, and also the ever increasing realization that "mean models" of subsurface properties do not adequately represent subsurface heterogeneity for reliable flow and transport predictions (e.g., Goovaerts, 1997).

Due to their inherent flexibility with regard to imposing constraints and their conceptual simplicity, simulated-annealing-type conditional stochastic simulations seem to be particularly promising for subsurface data integration (e.g., Deutsch and Wen, 1998, 2000; Kelkar and Perez, 2002). The simulated annealing (SA) approach is not limited to simple Gaussian statistics, and is able to incorporate any constraint on the output realizations that can be expressed in the form of an objective function, with the caveat that efficiency in terms of computation time and convergence decreases with constraint complexity. With SA, parameter fields satisfying all of the available data are obtained through minimization of a global, generally multi-component objective function, and multiple realizations can be generated by running the algorithm with different initial conditions. It should be emphasized that the variability seen in such multiple realizations depends on the applied constraints and stopping criteria, and as a result the realizations should not be confused with samples drawn from a posterior probability density function. Nevertheless, the SA method still allows evaluation of the variability in flow and transport behavior associated with uncertain data and constraints.

Recently, Tronicke and Holliger (2005) explored the use of SA for hydrogeophysical data integration through a synthetic model study. Starting with a simulated porosity model of a heterogeneous alluvial aquifer, they generated synthetic porosity logs and crosshole georadar traveltime data. These data, along with geostatistical constraints, were then used as conditioning information in a SA-based optimization procedure to generate porosity models that were consistent with all of the available information. In their work, Tronicke and Holliger (2005) pursued the classical SA approach of gradually "organizing" an uncorrelated random initial field through repeated swapping of values in the simulation grid, while adherence to the geophysical and geostatistical data was accomplished through matching the correlation coefficient between the realization and geophysical data to a prescribed value, and matching a prior parametric covariance model, respectively. Although the results obtained using this methodology clearly demonstrated that SA has much potential for hydrogeophysical data integration, we have found that the lateral continuity of the resulting porosity models is in general inadequate, which in turn significantly reduces the predictive value of such models in subsequent flow and transport simulations. Closer inspection indicates that this problem likely arises from the fact that it is inherently difficult with purely stochastic simulations to effectively impose constraints with regard to the underlying deterministic structure of the target parameter, as provided, for example, by high-resolution geophysical data.

In this paper, we present a novel SA-type conditional simulation procedure that aims to address and resolve this issue, as well as to improve the convergence and computational efficiency of the traditional SA method, which are known to be suboptimal as a result of having a relatively complex objective function. To begin, we review the overall methodology and describe an approach to more effectively account for the larger-scale deterministic information contained in geophysical data. Next, we test our conditional stochastic simulation algorithm on a synthetic data set consisting of crosshole georadar data and porosity logs from a highly heterogeneous, realistic aquifer model. Finally, we use our method to integrate field crosshole georadar and neutron porosity log data collected at the Boise Hydrogeophysical Research Site (BHRS) near Boise, Idaho, USA.

2. Conditional stochastic simulation using simulated annealing

2.1. Background

Simulated annealing is a directional Monte-Carlo-type optimization procedure, whose central idea is based upon the thermodynamics of a cooling melt. Atoms can move freely throughout a melt at high temperatures, but as the temperature is lowered, their mobility progressively decreases. Eventually, the system reaches its thermodynamic minimum-energy state and the atoms assume fixed positions within a crystal lattice. In SA, there are a large number of possible initial states, but during the cooling or annealing process all possible states converge to a final acceptable one. A flowchart describing the general methodology of SA for conditional simulation is shown in Fig. 1 (e.g., Deutsch, 2002; Kelkar and Perez, 2002; Tronicke and Holliger, 2005). The classical approach begins with an uncorrelated random field generated from an inferred/assumed probability distribution for the target parameter. The optimization process that follows consists of repeatedly perturbing individual values of this random field in order to satisfy a global objective function, O, which generally consists of the weighted sum of several component objective functions O_i that represent constraints on fitting the output realization to the available data or information:

$$O = \sum_{i=1}^{n} \omega_i O_i, \tag{1}$$

where n is the number of component objective functions and ω_i are the weights. All perturbations that lower the global objective function are accepted in the algorithm, whereas those that do not are accepted according to a Boltzmann-type exponential probability distribution controlled by a temperature parameter T. This "decision rule" is

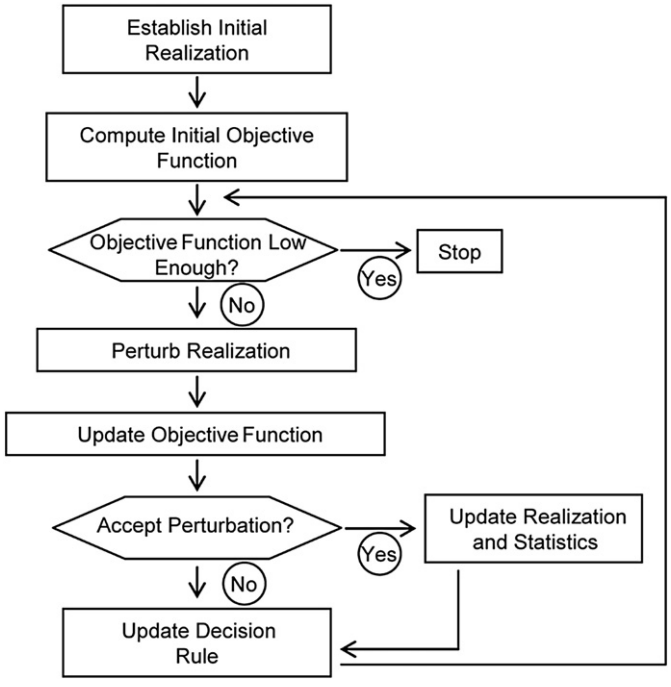

Fig. 1. Flowchart of SA approach for conditional stochastic simulation.