



Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Journal of Hydrology 294 (2004) 57–67

Journal  
of  
**Hydrology**

[www.elsevier.com/locate/jhydrol](http://www.elsevier.com/locate/jhydrol)

## Probabilistic reconstruction of geologic facies

Laura Guadagnini<sup>a</sup>, Alberto Guadagnini<sup>a,\*</sup>, Daniel M. Tartakovsky<sup>b</sup>

<sup>a</sup>*Dipartimento di Ingegneria Idraulica, Ambientale Infrastrutture Viarie e Rilevamento (DIAR),  
politecnico di Milano, Piazza L. Da Vinci, 32, 20133 Milano, Italy*

<sup>b</sup>*Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*

### Abstract

Random domain decomposition (RDD) provides a powerful tool for quantifying uncertainty in flow simulations, when both the geologic makeup of a porous medium and its hydraulic parameters are under-specified by data. Its prior applications dealt with flows in porous media whose internal compositions are amenable to simple parameterizations. This study provides a means for probabilistic reconstruction of boundaries between geologic facies. We apply our general approach to multiple data sets to reconstruct highly permeable zones within an aquitard in the Bologna (Italy) aquifer system and demonstrate how it can be used in conjunction with RDD.

© 2004 Elsevier B.V. All rights reserved.

*Keywords:* Stochastic; Random domain decomposition; Composite media

### 1. Introduction

With rapid advances in computing power and numerical techniques, insufficient medium parameterization (site characterization) seems to be one of the remaining stumbling blocks on the road to describing efficiently and reliably flow and transport in heterogeneous subsurface environments. As the size of computational domains increases (it is not uncommon to see numerical models with millions of degrees of freedom), the need to quantify uncertainty associated with assigning values of hydraulic and transport parameters (e.g., hydraulic conductivity, porosity, and dispersivity) to the nodes of a grid where data are not available is becoming increasingly

important. Complicating the matter further is an often occurring disparity between a scale (or scales) on which data have been collected and a scale on which they are used in numerical simulations.

Stochastic methods have emerged as a powerful tool for making predictions and quantifying predictive uncertainty in subsurface modeling. However, most of these approaches are limited to mildly heterogeneous porous media. For (semi-) analytical methods, such as moment equations, this requirement is essential to guarantee the accuracy of closure approximations. For purely numerical methods, such as Monte-Carlo simulations, it is needed to keep the number of realizations manageable. Of course, the more data are available, the higher the degree of heterogeneity for which stochastic methods are reliable (Guadagnini and Neuman, 1999). Advances in data collection and data assimilation techniques make well-conditioned simulations a reality.

\* Corresponding author. Tel.: +39-023996263; fax: +39-0223996298.

*E-mail addresses:* [alberto.guadagnini@polimi.it](mailto:alberto.guadagnini@polimi.it) (A. Guadagnini); [laura.guadagnini@polimi.it](mailto:laura.guadagnini@polimi.it) (L. Guadagnini); [dmt@lanl.gov](mailto:dmt@lanl.gov) (D.M. Tartakovsky).

When data are scarce, however, the question remains how to make the best use of them within the stochastic framework. The solution we propose is to utilize available data in a way that significantly increases their information content. Specifically, we demonstrate how available data, e.g., hydro-stratigraphic measurements, can be used to estimate the statistics of a facies' geometry. These data are further combined with other types of data, e.g., hydraulic conductivity measurements, to obtain a parameter's distribution within each facies. This is in contrast with existing stochastic approaches that either ignore internal macro-structures of porous media completely (the so called homogeneous approximation) or rely on statistically homogeneous multi-modal distributions. For the detailed review of these and other methods, such as data de-trending, we refer the interested reader to Winter et al. (2003).

A probabilistic description of facies' geometry is the input required by the random domain decomposition (RDD) approach (Winter and Tartakovsky, 2000, 2002; Winter et al., 2002). The key advantage of RDD is that it provides robust closures (accurate approximations) for moment equations even when the medium is highly heterogeneous. RDD makes use of the fact that a high degree of heterogeneity usually arises from the presence of different geologic facies (understood here in a very broad sense to include fractured regions, inclusions, layering, etc.) in the subsurface environment. Specifically, RDD replaces a non-Gaussian, multi-modal hydraulic and/or transport parameter field,  $Y(\mathbf{x})$ , of large variance,  $\sigma_Y^2$ , with a two-scale random process. The large-scale randomness arises due to uncertainty in internal boundaries of geologic facies. The small-scale randomness corresponds to uncertainty in hydraulic and/or transport parameters within each facies. In other words, a non-Gaussian, multi-modal probability density function  $p(Y)$  is replaced with a joint probability density function  $p(Y, \Gamma) = p(Y|\Gamma)p(\Gamma)$ . The conditional probability density function,  $p(Y|\Gamma)$ , describes the distribution of  $Y$  within each geologic facies conditioned on the boundary location  $\Gamma$ . As such it has convenient properties, such as uni-modality, small variances, etc.

Within the RDD framework, calculating the statistics of system parameters, such as mean hydraulic head,  $\langle h \rangle$ , is carried out in two steps.

The first step consists of calculating the conditional statistics, e.g.,

$$\langle h \rangle_\Gamma = \int h p(Y|\Gamma) dY. \quad (1)$$

The second step is averaging in probability space of  $\Gamma$

$$\langle h \rangle = \int \langle h \rangle_\Gamma p(\Gamma) d\Gamma. \quad (2)$$

Evaluating Eq. (1) requires a closure approximation, e.g., a perturbation expansion in the variance of log-conductivity,  $\sigma_Y^2$ . This formally limits applicability of such approximations to mildly heterogeneous porous media, i.e., media with  $\sigma_Y^2 \ll 1$ . As demonstrated by Winter et al. (2002), the RDD approach extends the range of applicability of a perturbation closure of the moment equations to heterogeneous media with  $\sigma_Y^2$  as high as 24.

One of the main unresolved issues of the RDD approach is how to obtain  $p(\Gamma)$  from available (scarce) data. Another is evaluation of the functional integral in Eq. (2). Previous applications of RDD dealt with idealized geometries, such as square inclusion (Winter et al., 2002) or perfect layering (Guadagnini et al., 2003), which makes their statistical parameterization straightforward. This study provides a means to apply RDD to realistic settings with complicated geologic structures, which do not lend themselves to simple parameterizations.

We formulate our general approach for the probabilistic geometry reconstruction in Section 2. The strategy we pursue here is somewhat similar to that proposed by Ritzi et al. (1994). The key difference is that their method provides an estimate of the facies geometry without quantifying the uncertainty associated with such a prediction. Our method provides a probabilistic description of the internal boundary  $\Gamma$  in the form of a probability density function  $p(\Gamma)$ . We demonstrate applicability of the approach to real world simulations, by using data sets collected at an aquifer system of Bologna, Italy (Section 3). The robustness and accuracy of the approach are analyzed in detail in Section 4 for a synthetic layered media. RDD is then used in Section 5 to calculate, without resorting to expensive Monte-Carlo simulations, mean hydraulic head and head variance. Section 6 compares our approach with alternative methods.

## 2. General approach

A typical site characterization yields multiple data sets that describe different, but often interconnected, features of a porous medium, such as its hydraulic conductivity, stratigraphy, etc. Often, these data sets correspond to different support volumes and possess variable degrees of reliability. We leave these important factors for future studies, and assume that all data are error free and are collected on the same support scale.

Let  $N_p$  be a number of parameters  $\mathcal{A}^{(p)}$  ( $p = 1, \dots, N_p$ ). Each parameter  $\mathcal{A}^{(p)}$  is sampled at locations  $\mathbf{x}_i^{(p)}$  ( $i = 1, \dots, N_s^{(p)}$ ), with  $\mathcal{A}_i^{(p)} = \mathcal{A}^{(p)}(\mathbf{x}_i)$  indicating the corresponding measurements. Reconstructing boundaries of geologic facies from small-scale data and estimating the corresponding uncertainties are our main goals. For the sake of simplicity, we assume that a medium consists of two geologic facies,  $F_1$  and  $F_2$ . A typical pdf for a parameter  $\mathcal{A}^{(p)}$  becomes bi-modal (see, for example, Fig. 2 that shows a frequency distribution for a sedimentologic data set used below in our analysis of the Bologna site).

The procedure we propose consists of the following steps.

### Step 1

*Constructing the indicator function.* Several parameters are often sampled at the same locations. Let  $N_0$  be the number of locations where measurements of more than one parameter are available. Then, the total of  $N = \sum_{j=1}^{N_p} N_s^{(j)} - \sum_{j=1}^{N_0} N_s^{(j)} + N_0$  locations have at least one measurement. To each of these locations,  $\mathbf{x}_i$  ( $i = 1, \dots, N$ ), we assign a (random) indicator function

$$I(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in F_1 \\ 0 & \mathbf{x} \in F_2 \end{cases}. \quad (3)$$

The following rule is used to infer values of the indicator function from the data sets  $\mathcal{A}^{(p)}$  ( $p = 1, \dots, N_p$ ). If a measurement falls within the interval  $A_1^- \leq \mathcal{A}_i^{(p)} \leq A_1^+$ , then the measurement point  $\mathbf{x}_i \in F_1$ , i.e.,  $I(\mathbf{x}_i) = 1$ . Otherwise, the measurement location  $\mathbf{x}_i \in F_2$ , i.e.,  $I(\mathbf{x}_i) = 0$ . The bounds  $A_1^-$  and  $A_1^+$  are inferred from analyzing corresponding bi-modal distributions.

### Step 2

*Estimating the relative volumes occupied by each facies.* Volumetric fractions  $V_1$  and  $V_2$  occupied by the facies  $F_1$  and  $F_2$ , respectively, are estimated by computing the global de-clustered mean of the indicator function  $I(\mathbf{x}_i)$ , where  $i = 1, \dots, N$ . De-clustering is required to avoid systematic bias introduced by uneven distributions of measurement points (Issaks and Srivastava, 1989).

### Step 3

*Structural analysis and spatial statistics of  $I(\mathbf{x})$ .* A correlation structure of the indicator function  $I(\mathbf{x})$  is computed via a sample variogram. The ensemble mean,  $\langle I(\mathbf{x}) \rangle$ , and variance,  $\sigma_I^2(\mathbf{x})$ , of the indicator function are computed by ordinary Kriging. This yields the probability,  $P[\mathbf{x} \in F_1]$ , of encountering the facies  $F_1$  at a point  $\mathbf{x}$ , since  $P[\mathbf{x} \in F_1] = \langle I(\mathbf{x}) \rangle$ . Equally important to note is that  $\langle I(\mathbf{x}) \rangle$  represents an estimate of the local volumetric fraction of the facies  $F_1$ .

### Step 4

*Calculating the probability distribution of  $I$  at each estimation point.* This step requires an assumption that  $I(\mathbf{x})$  is a truncated Gaussian field, so that the statistics computed by the point Kriging (mean and variance) uniquely specifies  $p(I; \mathbf{x})$ , a single-point probability density function of  $I(\mathbf{x})$ .

### Step 5

*Assigning probabilistic weights to the boundaries.* In the spirit of Ritzi et al. (1994), we assume that the mean boundary between the materials  $F_1$  and  $F_2$  is defined by points  $\mathbf{x}$ , where  $P[\mathbf{x} \in F_1] = V_1$ . This preserves the relative volumetric fractions occupied by each facies, as inferred from available data in Step 2. Then for any suitable spatial discretization  $\Delta$  of the pdf  $p(I; \mathbf{x})$ , we compute

$$\mathcal{W}(\mathbf{x}) = \int_{V_1 - \Delta/2}^{V_1 + \Delta/2} p(I; \mathbf{x}) dI. \quad (4)$$

The isolines  $\mathcal{W}(\mathbf{x}) = \mathcal{W}_i$  define boundaries  $\Gamma_i$  between the facies  $F_1$  and  $F_2$  corresponding to probabilistic weights  $\mathcal{W}_i$ .

Once the weights  $\mathcal{W}_i$  are calculated, statistics of the system states, such as the mean hydraulic head in Eq. (2), is readily approximated by

$$\langle h \rangle = \sum_i \mathcal{W}_i \langle h \rangle_{\Gamma_i}. \quad (5)$$

### 3. Application to the Bologna aquifer system (Italy)

We use our probabilistic facies reconstruction to analyze composition of an aquitard that separates two aquifers in the alluvial aquifer system of the city of Bologna, Italy (Fig. 1). The aquitard serves as a natural barrier that separates the polluted upper aquifer from the lower aquifer that is used for municipal water supplies. The available 39 geognostical boreholes logs and 183 well-logs reveal that the aquitard's thickness is highly variable, changing from 1 to 3 m in the vicinity of the peak of an alluvial fan to 8–12 m near the well fields, to even larger values in the northern part (Guadagnini et al., 2002). The deposits are mainly silty-clayey, with local interbedding of coarser material. A quantity  $[gr + sa]/th$ , representing the cumulative thickness of gravel (gr) and sand (sa) relative to the total thickness (th), is generally less than 0.2. However, it displays local peaks larger than 0.8, indicating possible discontinuities within the aquitard itself. We use hydrostratigraphic data to categorize materials within the aquitard Alpha into two classes (i.e., low- and high-permeability facies) according to the frequency distribution of  $[gr + sa]/th$  shown in Fig. 2. Presence of the highly conductive regions indicates

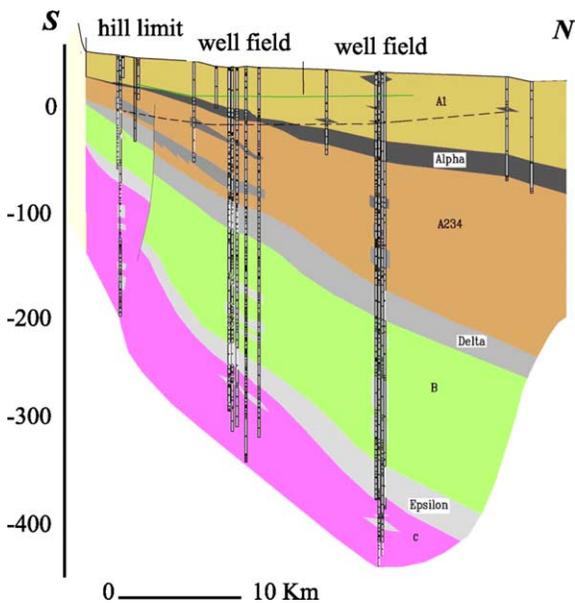


Fig. 1. Aquitard Alpha separating two aquifers in the Bologna aquifer system.

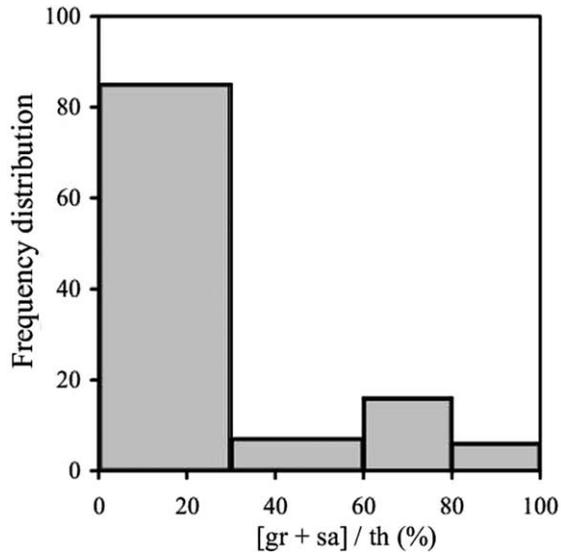


Fig. 2. Frequency distribution of  $[gr + sa]/th$ , the cumulative thickness of gravel (gr) and sand (sa) relative to the total thickness (th) in the aquitard Alpha. Such a distribution is typical for parameters that characterize porous media composed of two distinct facies.

possible connections between the upper and lower aquifers.

Following the steps outlined in Section 2, we use sedimentological and stratigraphic data sets to assign the indicator function  $I(x)$  to the low-conductivity facies,  $F_1$ . Specifically, a point  $x$  is assigned to either low- or high-conductivity facies according to values of both (a) local thickness of the aquitard (as estimated by stratigraphic analysis) and (b) percentage of the coarse-grain materials integrated along a stratigraphic column within the identified thickness. A global de-clustered average indicates that  $V_1 = 0.81$ , i.e., that the low-conductivity facies occupies about 80% of the sampled aquitard.

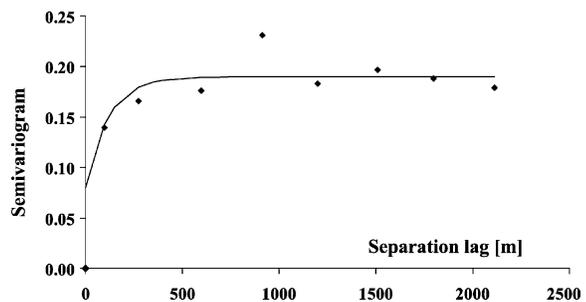


Fig. 3. A sample variogram for the indicator function  $I(x)$ .

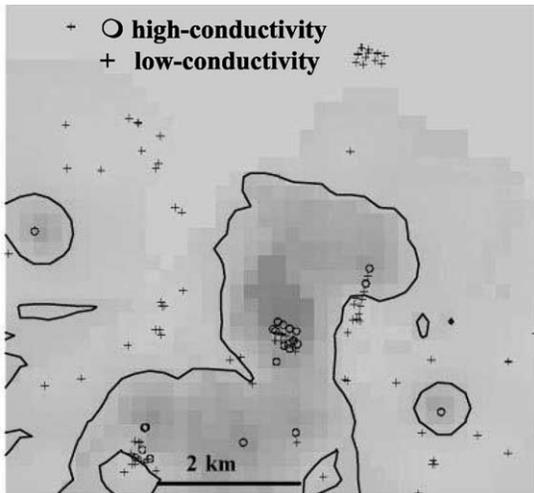


Fig. 4. Estimate of the facies geometry.

The next step is to use a sample variogram to estimate a spatial correlation of  $I(\mathbf{x})$ . We computed several directional variograms using an angular tolerance of  $30^\circ$  along the directions oriented at azimuths of 0, 45, 90 and  $135^\circ$  from the North. The sample variograms exhibit no clear evidence of anisotropy. Fitting an isotropic exponential model with a nugget to the sample variograms results in nugget = 0.08, sill = 0.11 and correlation scale = 350 m. Fig. 3 shows the corresponding variogram.

Using point Kriging, and selecting probability cutoff  $P[\mathbf{x} \in F_1] = V_1 = 0.81$ , yields an estimate

(ensemble mean) of the facies geometry, which is shown in Fig. 4. Fig. 5 depicts boundaries between the two facies that correspond to the probabilistic weights 0.069 and 0.044. The contour levels of cumulative frequency corresponding to these boundary configurations are 74 and 87%, respectively.

Reliability of our facies reconstruction approach is corroborated by the analysis of available stratigraphic data and geologic cross-sections, which suggest connections between upper and lower aquifers within the identified areas. To further investigate accuracy and robustness of the proposed approach, we consider a synthetic example. Of particular interest in this example is the influence of measurement locations on one's ability to delineate facies.

#### 4. Synthetic example

We start by generating a heterogeneous  $Y = \ln K$  field, whose properties, including facies' geometry, are known. Next we select 100 points, where conductivity is known (measured), and use these data to reconstruct the boundary. To investigate the impact of measurement locations we consider two scenarios. In the first, the data points are uniformly distributed throughout the domain, without any regard for a boundary location. In the second, the measurements are denser around a supposed boundary between the two materials. The latter scenario occurs

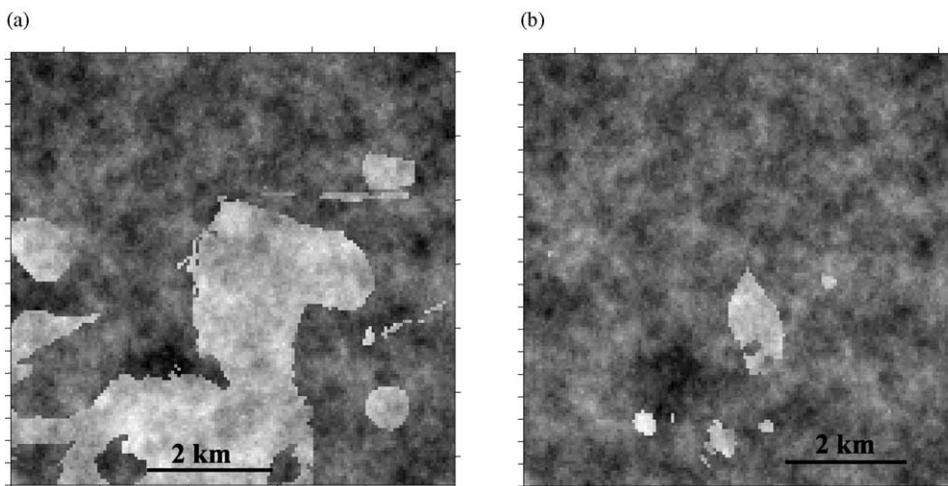


Fig. 5. Facies geometry corresponding to contour levels of 74% (a) and 87% (b) of cumulative frequency.

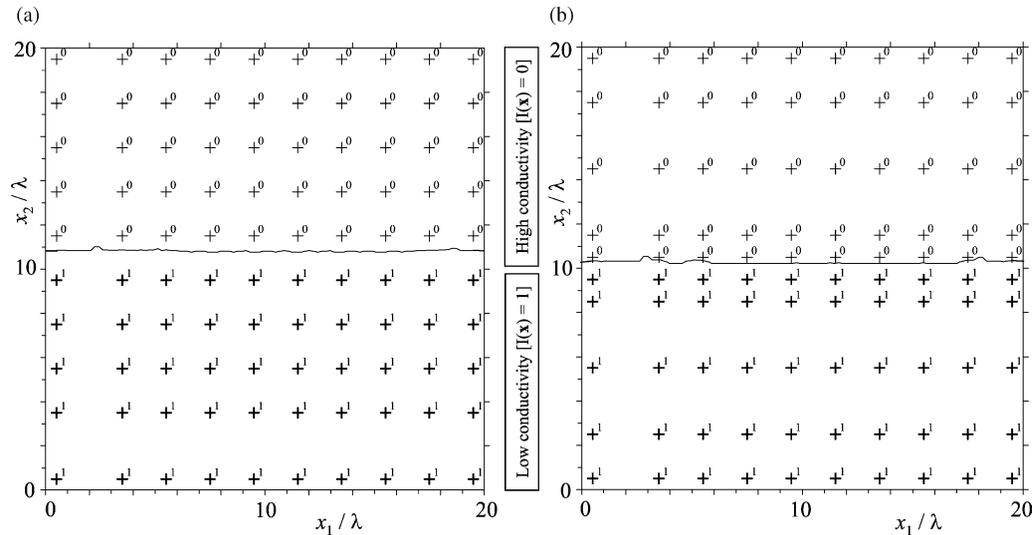


Fig. 6. Layered medium with (a) uniformly and (b) non-uniformly distributed measurement points. The points are assigned values of the indicator function, such that  $I(\mathbf{x}) = 1$  if  $\mathbf{x}$  is in the low-conductivity layer and 0 otherwise.

in practice, when an expert knowledge is incorporated into modeling process.

Consider a square domain within a layered medium composed of two contrasting materials. The layers are separated by a straight line that passes through the center of the square.

The domain is discretized by a grid of  $10^4$  square elements (100 rows and 100 columns) of uniform size,  $\Delta = 0.2$ , with five points per correlation length of  $Y$ . The log-conductivity is correlated within each layer but there is no correlation between the conductivities of different layers. Values of  $Y$  are generated at the center of elements using GSLIB (Deutsch and Journel, 1992). The two sampling strategies are shown in Fig. 6.

Following the procedure outlined in Section 2 leads to

- (1) the indicator function  $I(\mathbf{x}) = 1$  when  $\mathbf{x} \in F_1$  (a low-conductivity facies) and 0 when  $\mathbf{x} \in F_2$  (a high-conductivity facies);
- (2) the volumetric fractions  $V_1 = V_2 = 0.50$ ;
- (3) the sample variograms displaying a zonal anisotropy. Regardless of the sampling strategy, fitting an isotropic Gaussian model with a nugget to the sample variograms gives nugget = 0.02 and range = 20. For the uniform sampling, sill = 0.5 in the N/S direction and 0.09 in

the E/W direction. The alternative sampling results in sill = 0.5 in the N/S direction and 0.07 in the E/W direction;

- (4) the estimate (ensemble mean) of the boundary, which is defined by  $P[\mathbf{x} \in F_1] = V_1 = 0.50$ ; and the probabilistic weights associated with each boundary configuration (Figs. 7 and 8).

Not surprisingly, the non-uniform sampling increases accuracy of the estimates (ensemble means) of the boundary. At the same time, Fig. 7 shows that the two sampling strategies play virtually no role in assigning probabilistic weights to the boundaries. This is also confirmed by Fig. 8 that compares the boundary's probabilistic reconstructions corresponding to the two sampling strategies. This somewhat unexpected result can be explained by the simple geometry (a straight line) combined with the particular arrangements of data points used in our example. We expect that for more complicated spatial configurations, such as the one considered in Section 3, the placement of data points will play a crucial role in reducing uncertainty in the boundary reconstruction.

It is interesting to note that the maximum weight does not correspond to the horizontal line located in the middle of the square domain. This is so, because some of the measurement points with  $I = 1$  are closer

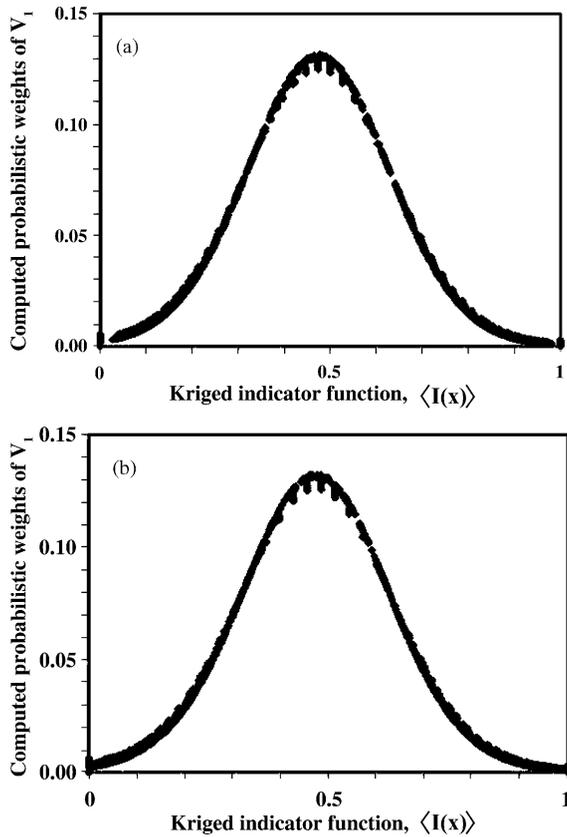


Fig. 7. Distributions of weights for the layered medium with (a) uniformly and (b) non-uniformly distributed measurement points.

to the ‘true’ separation line than the points with  $I = 0$  (Fig. 6). This results in a mean (estimated) boundary that somewhat differs from the ‘true’ location of the layering.

### 5. Quantifying uncertainty in flow simulations

To demonstrate how the probabilistic facies delineation can be used in conjunction with RDD, we consider flow through a medium with the internal boundaries analyzed in Section 1. Flow simulations in this section differ from reality in two important aspects. First, we impose artificial boundary conditions, which translates into an artificial flow regime. Second, due to lack of conductivity data, we arbitrarily select heterogeneous distributions of the hydraulic conductivity of each facies.

Specifically, we assume that within each facies the log hydraulic conductivity,  $Y = \ln K$ , is a statistically homogeneous Gaussian field with an exponential correlation function. Mean log-conductivities of the low- and high-permeability zones are set to  $\langle Y_{\text{low}} \rangle = 3.5$  and  $\langle Y_{\text{high}} \rangle = 7.0$ , respectively, when hydraulic conductivities are expressed in cm/day. We further assume that log-conductivity within each facies has the same variance  $\sigma_Y^2 = 1$ , correlation scale  $\lambda$ , and that the conductivities of the two facies are uncorrelated. All these assumptions are made for the sake of

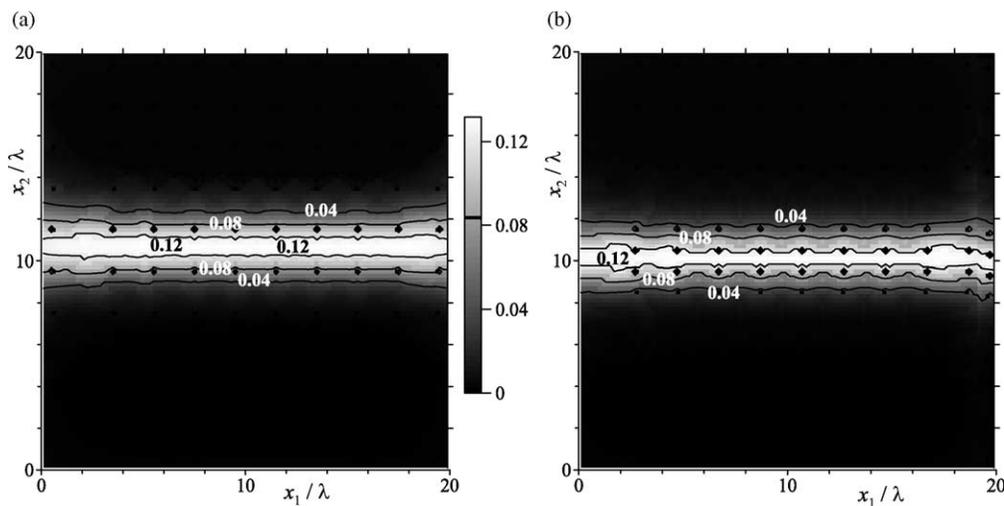


Fig. 8. Probabilistic reconstructions of the boundary between the two layers obtained from (a) uniformly and (b) non-uniformly distributed measurement points. Each isoline represents a boundary configuration between the two layers, which corresponds to a given weight.

convenience only, since the general theory of RDD allows for different covariance structures within facies and for cross-correlations between facies.

A correlation function for the hydraulic conductivity of the composite medium is obtained by averaging the conditional correlation functions over all possible realizations of the materials distribution. Even though the two materials are assumed to be uncorrelated, there exists a transitional zone, where the points from the two materials are correlated. Within this zone, membership of a given point in a particular material is uncertain. Averaging over the boundary distribution smooths the conditional correlation function of conductivity.

Consider steady-state flow through the rectangular domain shown in Figs. 4 and 5. Flow is driven by (mean) head gradient of about 0.2% due to the constant heads  $H_A = 21.0$  m and  $H_B = 1.0$  m that are imposed on the left and right hand boundaries of the domain, respectively. The remaining two boundaries are impermeable. Both the size of the flow domain ( $7.2 \times 6.8$  km<sup>2</sup>) and the background hydraulic gradient are representative of the actual field conditions. A pumping well is located in the middle of the field and operates at the constant flow rate of 100 m<sup>3</sup>/day.

The flow domain is discretized by a grid of 19,484 square elements (144 rows and 136 columns) of uniform size,  $\Delta = 50$  m, with five points per correlation length of  $Y$ . We obtain the conditional hydraulic head statistics (conditional mean and variance) by solving the RDD moment equations

(Winter and Tartakovsky, 2000, 2002) with the finite element code of Guadagnini and Neuman (1999). The accuracy of approximations that is required to derive these moment equations is assessed by comparison with Monte-Carlo simulations. Since a complete stabilization of the Monte-Carlo statistics is not necessary for such a comparison to be meaningful (Guadagnini and Neuman, 1999), 3000 conditional Monte-Carlo simulations for each of the log-conductivity fields were used. Since 21 realizations of the material distributions were considered in this study, we performed a total of  $21 \times 3000 = 63,000$  Monte-Carlo simulations. Such a huge computational burden makes the use of Monte-Carlo simulations to compute statistics of transient flow and/or transport problematic. RDD makes the stochastic analysis of such systems feasible.

We find that an overall agreement between the two solutions is excellent, except in the vicinity of the pumping well. This is in line with previous results of Guadagnini and Neuman (1999). The mean and variance of the hydraulic head computed with RDD are shown in Fig. 9.

To ascertain the relative importance of the two sources of uncertainty (facies geometry versus facies conductivity), we show in Fig. 10 the mean and variance of the hydraulic head corresponding to the random facies geometry but the deterministic (corresponding to their respective geometric means) hydraulic conductivities. Comparing Figs. 9 and 10 reveal that this simplification leads to similar

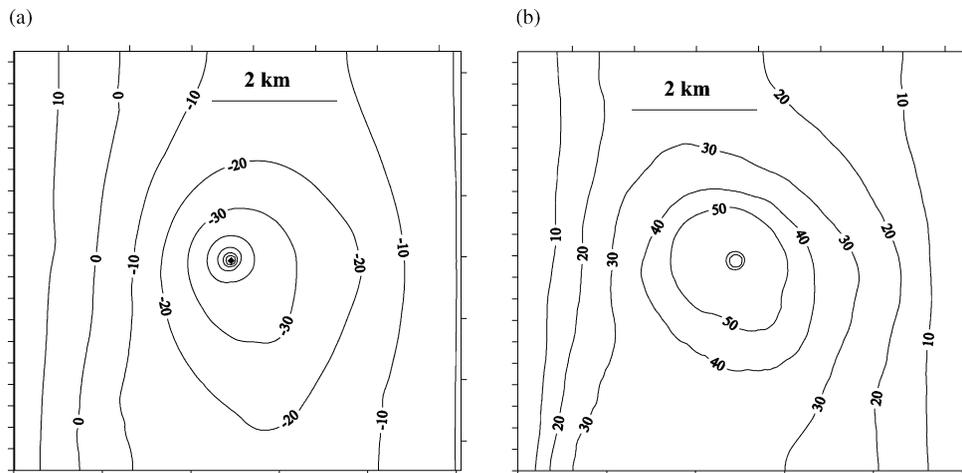


Fig. 9. Mean (a) and variance (b) of hydraulic head resulting from uncertain facies geometry and hydraulic conductivity.

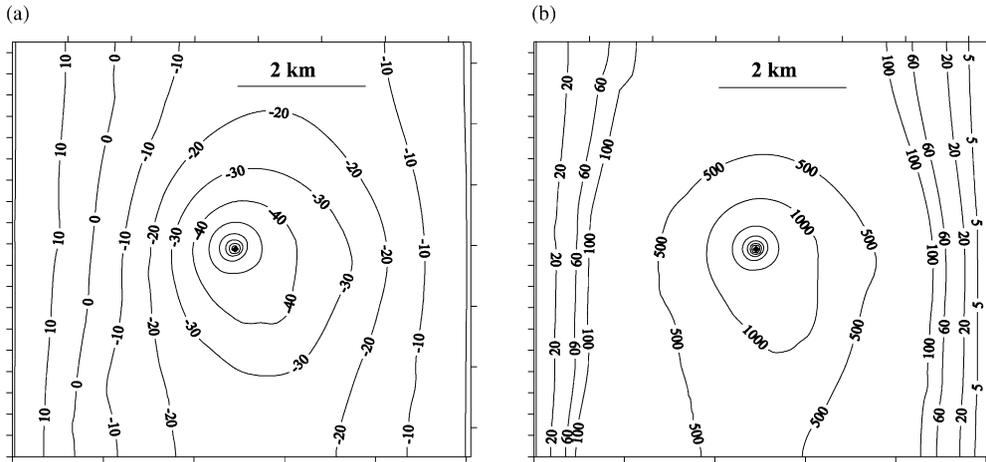


Fig. 10. Mean (a) and variance (b) of hydraulic head resulting from uncertain facies geometry and known (deterministic) hydraulic conductivity.

qualitative spatial patterns of the mean head and the head variance. However, quantitatively results differ considerably between the two models, especially with regard to hydraulic head variance. This is in accordance with earlier results of Guadagnini et al. (2003) obtained for stratified media and of Winter et al. (2003) obtained for a low-permeability inclusion in a high-permeability matrix.

### 6. Comparison with alternative models

There exist several approaches that implicitly account for the presence of different geologic facies,

without explicitly preserving the facies topology. These include a homogeneous approximation, deterministic trend models, and models resulting in multi-modal distributions of hydraulic conductivity. (We refer the interested reader to Winter et al., 2003 for a review of these and other similar approaches.). Such approaches result in statistically homogeneous conductivity distributions. In contrast, RDD gives rise to conductivity fields that are essentially inhomogeneous, in that their (ensemble) means, variances, and correlation functions are all space dependent. This distinction is important, since it is reasonable to expect that points within a geologic facies are better correlated than points in different

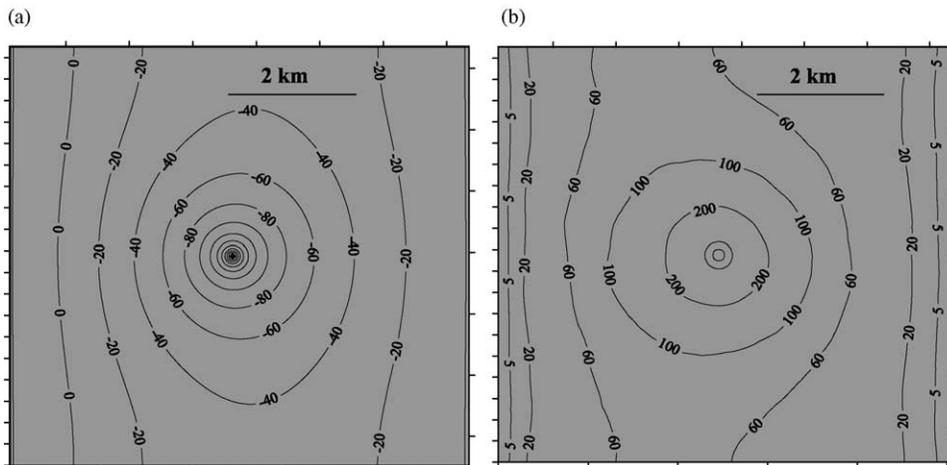


Fig. 11. Mean (a) and variance (b) of hydraulic head computed with the homogeneous approximation.

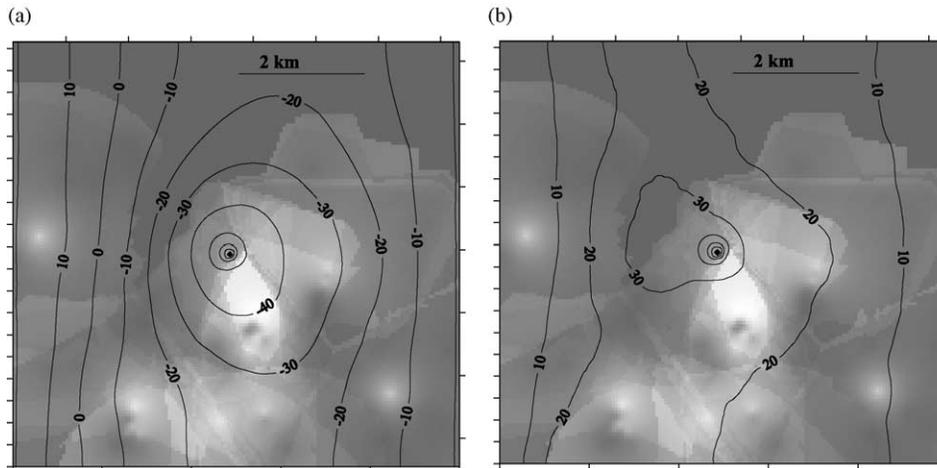


Fig. 12. Mean (a) and variance (b) of hydraulic head resulting from the bi-modal conductivity distribution. The latter is computed by Eq. (6) and is shown on the gray scale that varies from dark (low-conductivity) to white (high-conductivity).

facies. RDD accounts for this fact, while the other models do not.

We have analyzed errors introduced by deterministic trend models and have contrasted them with RDD elsewhere (Winter et al., 2002). Here, we provide a similar comparison for the homogeneous model and a model with a bi-modal conductivity distribution. The former constructs a statistically homogeneous field, whose statistics is determined as the mixture. The latter expresses the local conductivity,  $K_{eq}$ , as a weighted sum of the conductivities within each facies, i.e.

$$K_{eq} = P[I(\mathbf{x}) = 1]K_{low} + P[I(\mathbf{x}) = 0]K_{high}. \quad (6)$$

The probabilities  $P[I(\mathbf{x}) = 1]$  and  $P[I(\mathbf{x}) = 0] = 1 - P[I(\mathbf{x}) = 1]$  are determined by the Kriging estimate of  $I(\mathbf{x})$ . For each facies we generate 3000 log-conductivity fields with the same statistics as used before and then construct realizations of  $K(\mathbf{x})$  according to Eq. (6). Standard approaches for deriving bi-modal (dual continuum) distributions of hydraulic conductivity (Shvidler, 1988; Rubin, 1995) assume that the volumetric fractions of the materials are constant over an entire flow domain. The approach we use results in a statistically inhomogeneous conductivity field.

Figs. 11 and 12 show the hydraulic head statistics computed with the homogeneous approximation and the bi-modal distribution model, respectively.

The homogeneous approximation significantly overestimates the head and uncertainty (as quantified by the head variance) at the well. These results are consistent with those reported by Guadagnini et al. (2003). The bi-modal distribution model leads to the mean hydraulic head that is qualitatively and quantitatively similar to that obtained by considering only randomness in boundaries between materials. However, it significantly underestimates the hydraulic head variance.

## 7. Conclusions

We presented an approach for the probabilistic reconstruction of boundaries between geologic facies comprising natural porous media. Advantages of our approach are:

- It can assimilate in a seamless manner different sources of information, such as pumping and tracer tests, well-logs, and geophysics.
- It provides a required input for the RDD approach (Winter and Tartakovsky, 2000, 2002), which allows one to quantify uncertainties in both geologic makeup of porous media and hydraulic (and transport) parameters within each geologic facies.
- It alleviates the need for a closure approximation or Monte-Carlo simulations required by RDD to

compute ensemble averages in the probability space of geologic facies' geometries.

The proposed approach represents a first attempt at probabilistic reconstruction of geologic facies and relies on an assumption that the indicator function is a truncated Gaussian field. When this assumption is not valid, alternative approaches must be explored.

### Acknowledgements

This work was performed under the auspices of the US Department of Energy (DOE): DOE/BES (Bureau of Energy Sciences) Program in the Applied Mathematical Sciences contract KC-07-01-01. This work made use of shared facilities supported by SAHRA (Sustainability of semi-Arid Hydrology and Riparian Areas) under the STC Program of the National Science Foundation under agreement EAR-9876800. The authors acknowledge support from the European Commission under contract EVK1-CT-1999-00041 (W-SAHARA).

### References

- Deutsch, C.V., Journel, A.G., 1992. Geostatistical Software Library and User's Guide, Oxford University Press, New York.
- Guadagnini, A., Neuman, S.P., 1999. Nonlocal and localized analyses of conditional mean steady state flow in bounded, randomly nonuniform domains. 1. Computational examples. *Water Resources Research* 35, 3019–3040.
- Guadagnini, L., Farina, M., Frontini, S., Simoni, M., 2002. Geostatistical modeling of a heterogeneous alluvial aquifer. in: *Proceedings of the International Conference on Calibration and Reliability in Groundwater Modeling, Model-CARE2002*, Prague, 167–171.
- Guadagnini, A., Guadagnini, L., Tartakovsky, D.M., Winter, C.L., 2003. Random domain decomposition for flow in heterogeneous stratified aquifers. *Stochastic Environmental Research and Risk Assessment* 17, 394–407, doi 10.1007/s00477-003-0157-1, 2003.
- Issaks, E.H., Srivastava, R.M., 1989. *Applied Geostatistics*, Oxford University Press, New York.
- Ritzi, R.W., Jayne, D.F., Zahradnik, A.J., Field, A.A., Fogg, G.E., 1994. Geostatistical modeling of heterogeneity in glaciofluvial, buried-valley aquifers. *Ground Water* 32, 666–674.
- Rubin, Y., 1995. Flow and transport in bimodal heterogeneous formations. *Water Resources Research* 31, 2461–2468.
- Shvidler, M.I., 1988. Multicontinuum description of percolating flow through periodic inhomogeneous porous media. *Fluid Dynamics* 23, 894–901.
- Winter, C.L., Tartakovsky, D.M., 2000. Mean flow in composite porous media. *Geophysical Research Letters* 27, 1759–1762.
- Winter, C.L., Tartakovsky, D.M., 2002. Groundwater flow in heterogeneous composite aquifers. *Water Resources Research* 38(8) (doi:10.1029/2001WR000450).
- Winter, C.L., Tartakovsky, D.M., Guadagnini, A., 2002. Numerical solutions of moment equations for flow in heterogeneous composite aquifers. *Water Resources Research* 38(5) (doi: 10.1029/2001WR000222).
- Winter, C.L., Tartakovsky, D.M., Guadagnini, A., 2003. Moment equations for flow in highly heterogeneous porous media. *Surveys in Geophysics* 24, 81–106.