

## Nearest-neighbor classification for facies delineation

Daniel M. Tartakovsky,<sup>1,2</sup> Brendt Wohlberg,<sup>2</sup> and Alberto Guadagnini<sup>3</sup>

Received 13 February 2007; revised 11 May 2007; accepted 4 June 2007; published 20 July 2007.

[1] Geostatistics has become the dominant tool for probabilistic estimation of properties of heterogeneous formations at points where data are not available. Ordinary kriging, the starting point in the development of other geostatistical techniques, has a number of serious limitations, chief among which is the intrinsic hypothesis of the (second-order) stationarity of the underlying random field. Attempts to overcome this limitation have led to the development of ever more complex flavors of kriging. We pursue an opposite strategy that consists of finding the simplest possible technique that is adequate for the task of facies delineation. Guided by the principle of parsimony, we identify nearest-neighbor classification (NNC) as a viable alternative to geostatistics among deterministic techniques. We demonstrate that when used for the purpose of facies delineation, NNC, which has no fitting parameters and operational assumptions, outperforms indicator kriging, which has several parameters.

**Citation:** Tartakovsky, D. M., B. Wohlberg, and A. Guadagnini (2007), Nearest-neighbor classification for facies delineation, *Water Resour. Res.*, 43, W07201, doi:10.1029/2007WR005968.

### 1. Introduction

[2] Our knowledge of the spatial distribution of the physical properties of geologic formations is uncertain because of ubiquitous heterogeneity and the sparsity of data. Consequently, one is often faced with a problem of using measurements  $K_i = K(\mathbf{x}_i)$  of a parameter  $K(\mathbf{x})$  sampled at locations  $\mathbf{x}_i$  ( $i \in \{1, \dots, N\}$ ) to estimate its values at locations  $\mathbf{x}$ , where data are not available. This task is routinely accomplished by means of kriging, which is defined by *Olea* [1991, p. 14] as “a collection of generalized linear regression techniques for minimizing an estimation variance defined from a prior model for a covariance” [*Deutsch and Journel*, 1992].

[3] In other words, kriging computes an estimate  $\mathcal{K}$  of a true value of  $K$  as a weighted average of the data,

$$\mathcal{K} = \sum_{i=1}^N w_i K_i, \quad \sum_{i=1}^N w_i = 1, \quad (1)$$

where the weights  $\{w_i\}_{i=1}^N$  are determined by minimizing the error variance

$$\sigma_{\mathcal{K}}^2 = \frac{1}{N} \sum_{i=1}^N (\mathcal{K} - K_i)^2. \quad (2)$$

Different ways to approach this minimization problem give rise to different flavors of kriging (e.g., simple kriging,

ordinary kriging, etc), but they all share the common philosophy: a parameter  $K$  is treated as a random field. This step implicitly invokes the ergodicity hypothesis to substitute the sample spatial statistics, which can be calculated, for the ensemble statistics, which are actually required. Ergodicity cannot be proved and requires a number of modeling assumptions [*Rubin*, 2003, section 2.7, and references therein].

[4] On technical level, kriging allows one to estimate only the mean and variance of a random field  $K$ , while a variety of applications [e.g., *Isaaks and Srivastava* 1990, chapter 8] demand the knowledge of its full distribution. An assumption of subsurface parameters, e.g., hydraulic conductivity, having multi-Gaussian distributions or being nonlinear maps thereof would resolve this limitation, since such distributions are completely specified by their first two moments, i.e., by their means and variances. Unfortunately, the experimental evidence leaves little doubt that this assumption is often inadequate [e.g., *Gómez-Hernández and Wen*, 1998; *Winter et al.*, 2003, and references therein].

[5] This inadequacy is commonly addressed by means of indicator kriging (IK). The approach is based on a mapping of parameter  $K$  onto an indicator function [e.g., *Isaaks and Srivastava* 1990, p. 423]

$$I(k, \mathbf{x}) = \begin{cases} 1 & K(\mathbf{x}) \leq k \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

so that  $F_K(k)$ , the cumulative distribution function of  $K$ , estimated from data  $\{K_i\}_{i=1}^N$  mapped onto  $\{I_i = I(\mathbf{x}_i)\}_{i=1}^N$ , can be obtained as

$$F_K(k) = \frac{1}{N} \sum_{i=1}^N I_i(k). \quad (4)$$

An estimate  $\mathcal{I}$  of a true value of  $I$  at locations where data are not available is obtained by applying the kriging procedure

<sup>1</sup>Department of Mechanical and Aerospace Engineering, University of California, San Diego, La Jolla, California, USA.

<sup>2</sup>Group T-7, Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA.

<sup>3</sup>Dipartimento di Ingegneria Idraulica, Ambientale, Infrastrutture Viarie e Rilevamento, Politecnico di Milano, Milan, Italy.

(1)–(2) to  $\{I_i = I(\mathbf{x}_i)\}_{i=1}^N$ . If one can identify the threshold value  $k = k^*$  of  $K$  that distinguishes one geologic facies from another, then IK can be used as a tool for facies delineation [Ritzi et al., 1994; Guadagnini et al., 2004].

[6] IK shares some of the key limitations of other kriging techniques, chief among which is the intrinsic hypothesis of kriging, i.e., an assumption of the (second-order) stationarity of the underlying random field  $K$ . It requires, among other things, that its mean  $\langle K \rangle$  be constant in space and its two-point covariance function  $C_K$  depend on the separation distance between two points rather than on their locations. Some of the limitations that are peculiar to IK, including the so-called order relation problems, are discussed by Glacken and Blackney [1998], Goovaerts [1999], and Emery and Ortiz [2004].

[7] When faced with these challenges, the standard modus operandi has been to increase the complexity of geostatistical tools. For example, the assumption of stationarity has been relaxed by replacing it with the notion of local stationarity, and by replacing ordinary kriging with universal kriging. The latter, in turn, can yield biased estimates and can have the serious operational difficulties identified by Matheron [Cressie, 1986]. To overcome these, Cressie [1986] proposed median polish kriging. Likewise, potential pitfalls of IK, including a theoretical possibility of obtaining negative probability estimates and probability distributions that do not sum up to 1 (the order relation problems), have led to recent calls to either substantially modify IK [Emery and Ortiz, 2004] or abolish it altogether in favor of more conceptually and computationally complex techniques, such as Bayesian maximum entropy [Christakos and Li, 1998; Bogaert, 2002].

[8] Are there simpler alternatives to IK? Tartakovsky and Wohlberg [2004] and Wohlberg et al. [2006] provide one such alternative, support vector machines, that slightly outperforms IK, while requiring fewer modeling assumptions and data points. In this study, guided by the principle of parsimony, we search for a viable alternative to geostatistics among deterministic techniques. In section 2 we briefly review several deterministic methods of data analysis, the simplest of which, nearest-neighbor classification (NNC), is described in detail in section 3. To test its performance, we use NNC to reconstruct boundaries between two materials in synthetically generated porous media (section 4). This section also contains a comparison of the NNC performance with that of IK. Section 5 provides some concluding remarks.

## 2. Deterministic Analyses of Data

[9] Deterministic approaches to the analysis of spatial data include polygonal methods, various flavors of triangulation, inverse distance methods, splines, nonparametric regressions, etc. Such methods were widely used for subsurface characterization before the advent of geostatistics in the early 60s, which was spearheaded by G. Matheron [e.g., Matheron, 1963]. In fact, today it might come as a surprise that the appearance of geostatistics “has been countered by widespread negative reactions ranging from mild to total skepticism” (Laslett [1994, p. 391]; see also references therein). Since then, a plethora of studies have aimed to reestablish deterministic methods as viable alternatives to geostatistics.

[10] For example, Yakowitz and Szidarovsky [1985] found nonparametric regression to be more robust and accurate than kriging, when used on nonstationary fields with trends. Laslett [1994] provided a detailed review of the use of splines as a deterministic alternative to kriging and conducted a systematic comparison between the two. Gotway et al. [1996] concluded that the relative performances of an inverse distance method and ordinary kriging depend on the coefficient of variation of a field being reconstructed. Reed et al. [2004] compared the performance of inverse distance and nonlinear least squares inverse distance methods with several kriging techniques. They found that both deterministic approaches are highly competitive vis à vis ordinary kriging, but underperform relative to more complicated kriging approaches, such as quantile kriging.

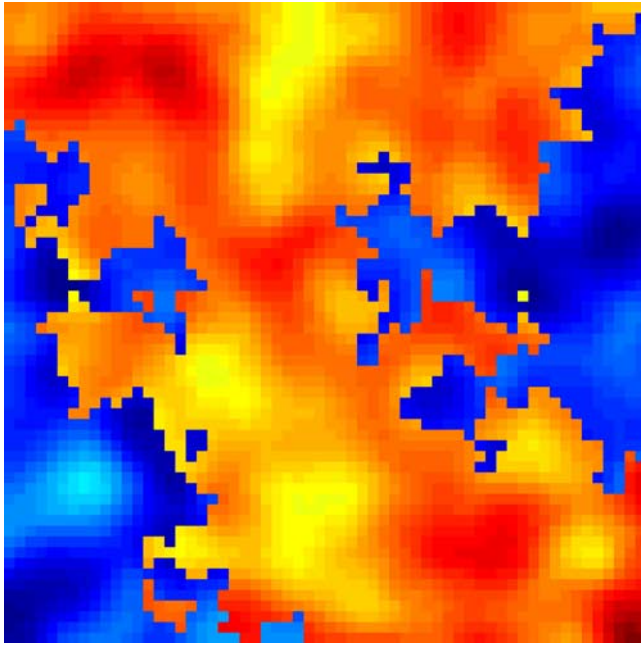
[11] These and other similar studies point to an emerging consensus: deterministic interpolation schemes provide viable alternatives to geostatistics because of their simplicity and the ease of implementation. The main, and often crucial, drawback of the deterministic approaches is their inability to quantify predictive uncertainty. In this study, we circumvent this limitation by applying a deterministic approach to data for the indicator function (3) obtained from an analysis of the corresponding parameter data. We rely on the principle of parsimony to select the deterministic approach. This leads us to nearest-neighbor classification (NNC), a class of classification methods that requires the smallest number of assumptions (none), is simplest both conceptually and in terms of implementation, and provides a performance that is highly competitive with that of IK. As we discuss in the following section, the only requirement for NNC to be workable is that the field to be reconstructed possesses some spatial correlation or smoothness, so that the class of a data point is relevant in determining the class of a neighboring point. This requirement is, of course, essential for any reconstruction effort, since no amount of data would be sufficient to estimate parameters of an uncorrelated (white noise) field.

## 3. Nearest-Neighbor Classification

[12] Geostatistics revolves around the four key questions “Are there enough nearby samples? Are there too many samples? Are there nearby samples that are redundant? Are the nearby samples relevant?” [Isaaks and Srivastava, 1990, p. 338]. The nearest-neighbor classification (NNC) we propose provides a radical answer to these questions: Only one sample located closest to the point of interest is enough for classification purposes!

[13] Nearest-neighbor methods form a large class of algorithms, many of which are described by Dixon [2002] and a few have been used in subsurface applications. This can cause a considerable confusion, since the term is often applied to vastly different computational schemes. For example, Wilkinson et al. [1983] proposed a nearest-neighbor approach that requires data trends to be locally quadratic. This and other limitations were deemed by Cressie [1986] to be overly restrictive for subsurface applications.

[14] There are practically no similarities, except for the name, between the Wilkinson et al. [1983] method and the NNC we propose. Our approach consists of the following two steps. First, we assign the indicator to each data point.



**Figure 1.** Synthetic example 1: values of  $\ln K$  ranging between  $-2.04$  and  $9.89$  on a  $60 \times 60$  grid.

Then we employ the  $k$ -nearest-neighbor classification [Hastie et al., 2001, section 13.3], in which the classification of a test point is determined by majority vote amongst the  $k$  nearest-neighbor points in the training set. Here, we consider the simplest of these classifiers, with  $k = 1$ , for which the classification of each point in the domain is determined by finding the nearest training point, and assigning the known class of that point. Given a set of training data points  $\{\mathbf{x}_i\}_{i=1}^N$  with corresponding known indicator function values  $I(\mathbf{x}_i)$ , the NNC classification for an arbitrary point  $\mathbf{x}$  in the domain is computed as follows:

[15] 1. Define  $j$  as the index of the training data point, from the set  $\{\mathbf{x}_i\}_{i=1}^N$ , which is closest to query point  $\mathbf{x}$ ; that is,  $j = \operatorname{argmin}_i \|\mathbf{x} - \mathbf{x}_i\|_2$ .

[16] 2. Assign the indicator function value of training data point  $\mathbf{x}_j$  (i.e.,  $I(\mathbf{x}_j)$ ) as the indicator function value of query point  $\mathbf{x}$ . This classification is simple to compute, and has no free parameters to estimate.

#### 4. Computational Example

[17] To test the NNC classification described above, we reconstruct, from a few data points randomly selected according to a uniform distribution, the boundaries between heterogeneous geologic facies in two synthetic porous media shown in Figures 1 and 2. These examples idealize two commonly found geological settings. Figure 1 mimics a planar view of a subsurface environment in which a highly conductive material has been deposited on top of a less conductive basement and is laterally constrained by a less permeable rock. Figure 2 represents a vertical cross section of a high-conductivity aquifer with a narrow and elongated inclusion of a low-conductivity material. Both the hydraulic conductivities of each material and the internal boundaries between materials are generated randomly (see Wohlberg et al. [2006] for a detailed description of the generation

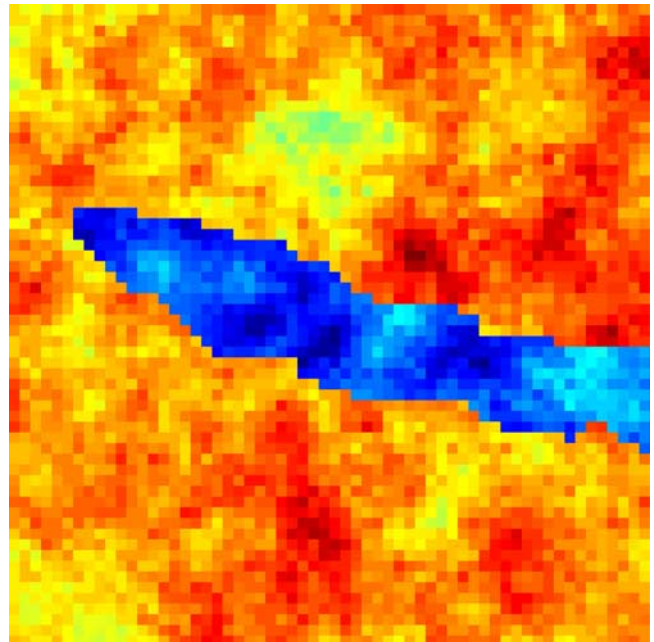
procedure) to avoid any bias in favor of a particular reconstruction method. Indicator fields  $I$  corresponding to data in Figures 1 and 2 were obtained by setting a threshold value of 4.0 for example 1, and 3.0 for example 2, as described by Wohlberg et al. [2006].

[18] We used NNC to reconstruct the boundary between the geologic facies in Figures 1 and 2 from data sets corresponding to various sampling densities. The number of data points ranged from 10 (a sampling density of 0.28%) to 110 (a sampling density of 3.06%). For each sampling density, an ensemble of 20 randomly generated realizations of the sample locations was constructed to reduce the influence of “lucky” and “unlucky” measurement locations. Classification errors for each sampling density were computed as the fraction of misclassified grid points averaged over the classification results for each of the 20 realizations at that sampling density.

[19] The results of these simulations are presented in Figures 3 and 4, in comparison with the geostatistical approach of Ritzi et al. [1994] and the support vector machine (SVM) approach of Wohlberg et al. [2006]. These latter methods are described briefly in Appendix A.

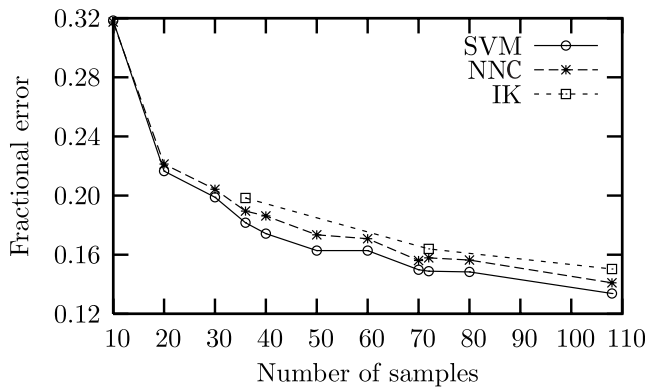
[20] One can see that NNC outperforms IK, slightly in the first example and significantly in the second. The SVM slightly outperforms NNC in the first example but NNC outperforms it in the second. As the number of samples decreases, the use of IK becomes questionable, while both the SVM and NNC remain functional, albeit at the cost of the exponentially increasing fractional error.

[21] Moreover, the accuracy of NNC increases exponentially with the sampling density (the number of measurements), whereas the accuracy of IK might either increase (Figure 3) or decrease (Figure 4) depending on the medium being reconstructed. This is because the blind selection of measurement locations fails to provide an accurate description of the localized and highly anisotropic low-conductiv-



**Figure 2.** Synthetic example 2: values of  $\ln K$  ranging between  $-3.00$  and  $10.44$  on a  $60 \times 60$  grid.





**Figure 3.** Example 1: boundary estimation errors with the indicator kriging (IK) approach, the nearest-neighbor classification (NNC), and support vector machines (SVM).

ity geological structure in Figure 2. The advantage of NNC becomes more striking if one recalls that, unlike IK and SVM, it has no fitting parameters.

## 5. Conclusions

[22] We explored the potential of a simple deterministic approach, nearest-neighbor classification (NNC), to reconstruct geologic facies from a sparse data set containing measurements of material properties. The comparison of its performance with that of indicator kriging (IK) leads to the following major conclusions.

[23] 1. NNC outperforms IK in terms of the reconstruction error.

[24] 2. NNC has lower data requirements than IK, for which too few data points render the construction of a variogram meaningless.

[25] 3. Unlike IK, NNC has no fitting parameters.

[26] These findings, combined with the principle of parsimony, strongly suggest that nearest-neighbor classification is a viable alternative to geostatistics.

[27] It is important to recognize that our use of NNC is limited to the problem of facies delineation. There are other applications, e.g., spatial interpolation of parameter values, where NNC might or might not outperform geostatistics. Their analysis lies outside the scope of the present study and is a subject of our ongoing investigation.

## Appendix A: Alternative Approaches

[28] The NNC classification algorithm is compared with two alternatives, the geostatistical approach of *Ritzi et al.* [1994] and the support vector machine approach of *Wohlberg et al.* [2006], each of which is discussed briefly below.

### A1. Geostatistical Approach

[29] The *Ritzi et al.* [1994] algorithm for facies reconstruction from indicator data is as follows. First, ordinary kriging [*Deutsch and Journel*, 1992] is used to construct a map of the ensemble average of the indicator function  $\langle I(\mathbf{x}) \rangle$  from the data  $\{I(\mathbf{x}_i)\}_{i=1}^N$ . The ensemble mean  $I(\mathbf{x})$  is the probability that a point  $\mathbf{x}$  lies in Material 1,  $\langle I(\mathbf{x}) \rangle = P[\mathbf{x} \in M_1]$ . Then a boundary between the two materials is defined as an isoline  $P[\mathbf{x} \in M_1] = c$ , where  $c$  is a number of

data points in Material 1 (or 2) relative to the total number of data points, after accounting for data clustering.

[30] In some cases, this value of  $c$  does not guarantee that the kriging estimate of the fraction of the total area covered by the low-conductivity material equals the declustered global mean of the original indicator data, resulting from the raw data. In such cases,  $c$  is set to a value of the kriged indicator field which allows one to recover a reconstruction that honors the empirical relative volumetric fractions of the two materials.

### A2. Support Vector Machines

[31] The Support Vector Machine (SVM) is a tool from Statistical Learning Theory that has been recently used by *Tartakovsky and Wohlberg* [2004] and *Wohlberg et al.* [2006] to solve the problem of facies delineation posed in section 4. (See also *Kanevski and Maignan* [2004] for other uses of SVMs in geosciences.) The theory and the implementation of SVMs is described in great detail by *Wohlberg et al.* [2006].

[32] In essence, the SVM solves the problem of facies delineation by minimizing the quadratic functional

$$\max_{\gamma} \left\{ \sum_{i=1}^N \gamma_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \gamma_i \gamma_j J_i J_j \mathcal{K}(\mathbf{x}_i, \mathbf{x}_j) \right\}, \quad (\text{A1})$$

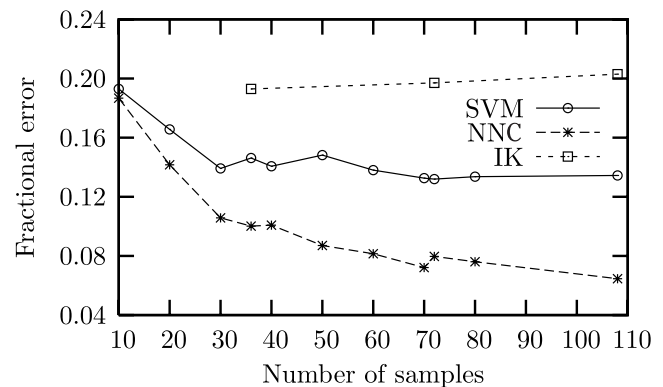
where  $J(\mathbf{x})$  is a map of the indicator function  $I(\mathbf{x})$  defined as  $J(\mathbf{x}) = 1$  for  $I(\mathbf{x}) = 1$  and  $J(\mathbf{x}) = -1$  for  $I(\mathbf{x}) = 0$ , and  $\mathcal{K}(\mathbf{x}_i, \mathbf{x}_j)$  is a given Mercer kernel, subject to the constraints

$$0 \leq \gamma_i \leq C \quad \text{and} \quad \sum_{i=1}^N \gamma_i J_i = 0. \quad (\text{A2})$$

This optimization problem has a well defined global minimum that is influenced by the choice of the fitting parameter  $C$ . The resulting facies boundary is given by

$$J(\mathbf{x}) = \text{sgn} \left( \sum_{i=1}^N \gamma_i^* J_i \mathcal{K}(\mathbf{x}, \mathbf{x}_i) + b^* \right), \quad (\text{A3})$$

where  $\max_{\gamma}^*$  is the solution of the optimization (A1), and  $b^*$  is determined as described by *Wohlberg et al.* [2006].



**Figure 4.** Example 2: boundary estimation errors with the indicator kriging (IK) approach, the nearest-neighbor classification (NNC), and support vector machines (SVM).

[33] In the simulations reported here, we have used the exponential kernel  $\mathcal{K}(\mathbf{x}_i, \mathbf{x}_j) = \exp[-\|\mathbf{x}_i - \mathbf{x}_j\|/(2\sigma^2)]$ . To select the optimal values of parameters  $C$  and  $\sigma$ , we have used the leave-one-out approach [Wohlberg et al., 2006] on the grid  $(C, \sigma) \in \{10^{k/2} | k \in \{-2, -1, 0, 1, 2\}\} \times \{10^{k/4} | k \in \{-2, -1, 0, 1, 2, 3, 4\}\}$ .

[34] It is worth emphasizing a technical detail which was not mentioned in [Wohlberg et al., 2006]; we only became aware of this detail after implementing our own SVM software for the experiments described here, since it was previously obscured by our use of a publicly available SVM library. While we described the standard SVM formulation with bias term, those results, as well as the results presented here, are computed using the zero-bias SVM formulation [Vogt, 2002; Huang and Kecman, 2004], in which optimization (A1) has only the single constraint  $0 \leq \gamma_i \leq C$ , and the facies boundary is given by

$$J(\mathbf{x}) = \operatorname{sgn}\left(\sum_{i=1}^N \gamma_i^* J_i \mathcal{K}(\mathbf{x}, \mathbf{x}_i)\right). \quad (\text{A4})$$

For reasons which we are unable to explain, we have found the SVM with bias term to have a performance which is inferior in this application, to the extent that it is significantly outperformed by NNC and slightly outperformed by IK.

[35] **Acknowledgments.** We would like to thank H.-J. Hendricks Franssen and other two reviewers for their insightful comments. This research was supported by the DOE Office of Science Advanced Scientific Computing Research (ASCR) program in Applied Mathematical Sciences and by the Italian Ministry of Education, Universities and Research (MIUR) under grant PRIN2006.

## References

- Bogaert, P. (2002), Spatial prediction of categorical variables: The Bayesian maximum entropy approach, *Stochastic Environ. Res. Risk Assess.*, 16(6), 425–448.
- Christakos, G., and X. Li (1998), Bayesian maximum entropy analysis and mapping: A farewell to kriging estimators?, *Math. Geol.*, 30(4), 435–462.
- Cressie, N. (1986), Kriging nonstationary data, *J. Am. Stat. Assoc.*, 81(395), 625–634.
- Deutsch, C. V., and A. G. Journel (1992), *Geostatistical Software Library and User's Guide*, Oxford Univ. Press, New York.
- Dixon, P. M. (2002), Nearest neighbour methods, in *Encyclopedia of Environmetrics*, vol. 3, edited by A. H. El-Shaarawi and W. W. Piegorsch, pp. 1370–1383, John Wiley, New York.
- Emery, X., and J. M. Ortiz (2004), Shortcomings of multiple indicator kriging for assessing local distributions, *Appl. Earth Sci.*, 113(4), 249–259, doi:10.1179/174327504X27242.
- Glacken, I., and P. Blackney (1998), A practitioners implementation of indicator kriging, paper presented at Symposium on Beyond Ordinary Kriging, Geostat. Assoc. of Australas., Perth, W. Aust., Australia.
- Gómez-Hernández, J. J., and X.-H. Wen (1998), To be or not to be multi-Gaussian? A reflection on stochastic hydrogeology, *Adv. Water Resour.*, 21(1), 47–61.
- Goovaerts, P. (1999), Geostatistics in soil science: State-of-the-art and perspectives, *Geoderma*, 89(1–2), 1–45.
- Gotway, C. A., R. B. Ferguson, G. W. Hergert, and T. Peterson (1996), Comparison of kriging and inverse-distance methods for mapping soil parameters, *Soil Sci. Soc. Am. J.*, 60(4), 1237–1247.
- Guadagnini, L., A. Guadagnini, and D. M. Tartakovsky (2004), Probabilistic reconstruction of geologic facies, *J. Hydrol.*, 294, 57–67.
- Hastie, T., R. Tibshirani, and J. Friedman (2001), *The Elements of Statistical Learning*, Springer, New York.
- Huang, T. M., and V. Kecman (2004), Bias term  $b$  in SVMs again, paper presented at European Symposium on Artificial Neural Networks, Bruges, Belgium, Inst. of Electr. and Electron. Eng., 28–30 April.
- Isaaks, E. H., and R. M. Srivastava (1990), *An Introduction to Applied Geostatistics*, Oxford Univ. Press, New York.
- Kanevski, M., and M. Maignan (2004), *Analysis and Modelling of Spatial Environment Data*, Marcel Dekker, New York.
- Laslett, G. M. (1994), Kriging and splines: An empirical comparison of their predictive performance in some applications, *J. Am. Stat. Assoc.*, 89(426), 391–400.
- Matheron, G. (1963), Principles of geostatistics, *Econ. Geol.*, 58(8), 1246–1266.
- Olea, R. (Ed.) (1991), *Geostatistical Glossary and Multilingual Dictionary*, Oxford Univ. Press, New York.
- Reed, P. M., T. R. Ellsworth, and B. S. Minsker (2004), Spatial interpolation methods for nonstationary plume data, *Ground Water*, 42(2), 190–202, doi:10.1111/j.1745-6584.2004.tb02667.x.
- Ritzi, R. W., D. F. Jayne, A. J. Z. Jr., A. A. Field, and G. E. Fogg (1994), Geostatistical modeling of heterogeneity in glaciofluvial, buried-valley aquifer, *Groundwater*, 32(4), 666–674.
- Rubin, Y. (2003), *Applied Stochastic Hydrogeology*, Oxford Univ. Press, New York.
- Tartakovsky, D. M., and B. E. Wohlberg (2004), Delineation of geologic facies with statistical learning theory, *Geophys. Res. Lett.*, 31, L18502, doi:10.1029/2004GL020864.
- Vogt, M. (2002), SMO algorithms for support vector machines without bias term, technical report, Lab. for Control Eng. and Process Autom., Inst. of Autom. Control, Tech. Univ. Darmstadt, Darmstadt, Germany.
- Wilkinson, G. N., S. R. Eckert, T. W. Hancock, and O. Mayo (1983), Nearest neighbour (NN) analysis of field experiments, *J. R. Stat. Soc., Ser. B*, 45(2), 151–211.
- Winter, C. L., D. M. Tartakovsky, and A. Guadagnini (2003), Moment equations for flow in highly heterogeneous porous media, *Surv. Geophys.*, 24(1), 81–106.
- Wohlberg, B., D. M. Tartakovsky, and A. Guadagnini (2006), Subsurface characterization with support vector machines, *IEEE Trans. Geosci. Remote Sens.*, 44(1), 47–57.
- Yakovitz, S. J., and F. Szidarovsky (1985), A comparison of geostatistical and nonparametric regression methods, *J. Multivariate Anal.*, 16(1), 21–53.

A. Guadagnini, DIAR, Politecnico di Milano, Piazza L. Da Vinci, 32, I-20133 Milano, Italy. (alberto.guadagnini@polimi.it)

D. M. Tartakovsky, Department of Mechanical and Aerospace Engineering, University of California, San Diego, 9500 Gilman Dr., Mail Code 0411, La Jolla, CA 92093, USA. (dmt@ucsd.edu)

B. Wohlberg, Group T-7, Theoretical Division, Los Alamos National Laboratory, MS B284, Los Alamos, NM 87545, USA. (brendt@t7.lanl.gov)