Multi-Point BME Space/Time Mapping of Environmental Variables
George Christakos
Department of Environmental Sciences & Engineering, School of Public Health, 111 Rosenau Hall, CB#7400, University of North Carolina, Chapel Hill, N.C. 27599-7400, U.S.A.
EMail: george_christakos@unc.edu

Abstract

The multi-point Bayesian Maximum Entropy (BME) approach considers physical data analysis from a modern geostatistical perspective, promoting the view that a deeper understanding of a theory of knowledge is an important prerequisite for the improved mathematical modelling and mapping of spatiotemporal environmental variables. A spatiotemporal map should depend on what we know about the phenomenon it represents, as well as how we know it (i.e., how we collect and process knowledge). BME's rich theoretical basis provides guidelines for the adequate interpretation and processing of all knowledge bases available. BME is formulated in a way that preserves earlier geostatistical results, which are its limiting cases, but it also provides novel and more general results that could not be obtained within traditional geostatistics. Numerical comparisons of BME vs. simple and indicator kriging are discussed via a porosity data set and a synthetic example.

1 Spatiotemporal Mapping

For a very large number of environmental problems, the required outcome of the analysis is one or more spatiotemporal maps. These maps may be the visual representation of information regarding the distribution of environmental variables in space/time (e.g., spring water solute contents; calcium, nitrate and chloride ions), usually obtained from data sets. Or, they may involve broader knowledge bases (general knowledge, analytic beliefs, mathematical models representing physical laws given some boundary/initial conditions, etc.). While the first viewpoint is more descriptive, the second one is more explanatory. This work favors a perspective that combines both: a spatiotemporal map representing the evolution of an environmental variable in space/time should be the outcome of an analysis that incorporates the set of observations available in space/time as well as other important knowledge bases.

Bayesian Maximum Entropy (BME) is a stochastic method for studying spatiotemporal distributions of environmental variables in terms of random fields. The main advantage of random fields is that they allow the rigorous characterization of
space/time variabilities and uncertainties which are an intrinsic part of most environmental phenomena. Let \( p = (s,t) \in \mathbb{R}^s \times T \) denote points in the Cartesian product space × time, and assume that \( X(p) \) is a spatiotemporal random field. The \( X(p) \) can be, also, viewed as a collection of correlated random variables at points \( p_{map} = [p_1 \ldots p_m p_{k_1} \ldots p_{k_p}]^T \) in space/time. A realization of the \( X(p) \) at these points is denoted by \( \chi_{\text{map}} = [\chi_1 \ldots \chi_m \chi_{k_1} \ldots \chi_{k_p}]^T \). It must be clear that we use the terms \( p_{map} \) and \( \chi_{\text{map}} \), because our goal is to obtain spatiotemporal maps displaying estimates \( \hat{\chi}_{k_\ell} \) of the unknown values \( \chi_{k_\ell} \) of \( X(p) \) at space/time points \( p_{k_\ell} \), \( (\ell = 1,\ldots,p) \), on the basis of the data vector \( \chi_{\text{data}} = [\chi_1,\ldots,\chi_m]^T \) associated with the knowledge bases available.

BME is a powerful method of scientific reasoning. Some of its appealing features are summarized below (detailed discussions may be found in Christakos, 1992, 1998; Christakos and Hristopulos, 1998; Christakos and Li, 1998):

- It satisfies sound epistemological ideals and incorporates physical knowledge bases in a rigorous and systematic manner.
- It does not require any assumptions regarding the shape of the underlying probability law (non-gaussian laws are automatically incorporated).
- It can be applied as effectively in the spatial as in the spatiotemporal domains and can model non-homogeneous/non-stationary data, in general.
- It leads to non-linear predictors, in general, and can obtain well-known kriging estimators as its limiting cases.
- It is easily extended to functional and vector environmental variables.
- It allows multi-point mapping (i.e., interdependent predictions at several space/time points simultaneously), which most traditional techniques do not offer.

2 Physical Knowledge Bases

The BME study of environmental variables is a multidisciplinary affair involving knowledge from various sciences, not the province of pure, a priori mathematics. Indeed, the physical knowledge used in the study of environmental phenomena may come from a variety of sources. We distinguish between two prime bases of knowledge: General knowledge and case-specific knowledge.

General knowledge denotes the background knowledge relative to the mapping situation overall based on general laws of science, structured patterns and assumptions, previous experience, analytic beliefs (justified on the basis of relations of concepts, theories), etc.. For the purposes of this work, the general knowledge will be expressed mathematically in terms of known statistics about the environmental variables of interest. Let \( g_\alpha(\chi_{\text{map}}) \), \( \alpha = 0,1,\ldots,N_\alpha \), be functions of \( \chi_{\text{map}} \) such that \( g_\alpha(p_{map}) \) are known \( X(p) \)-statistics (the bar denotes expectation; \( g_0 = 1 \) is a normalization constant). These statistics may come from a variety of sources, such as a theory or a physical law governing phenomena in several media.

Case-specific knowledge includes measurements and perceptual evidence, empirical propositions, expertise with the specific situation, incomplete evidence, etc.. Data sets available as case-specific knowledge are usually divided into two
groups: Hard data $\chi_{\text{hard}}$ (i.e., measurements obtained from real-time observation devices), and soft data $\chi_{\text{soft}}$ (i.e., interval values, probability statements, empirical charts, expert’s assessments, etc.). Both knowledge bases mesh in coherent interaction with the new information provided by the BME map in order to provide us with an explanatory rationale for the phenomenon.

3 The Method

Just as any other product of scientific reasoning, solutions to environmental problems should be derived by means of a sound epistemological framework. The latter can enlighten considerably our mathematical investigations for the best solution possible. The proposed epistemological framework distinguishes between three main stages of knowledge-gaining, processing and interpretation: (a) The prior stage: initial consideration of an environmental problem does not work in an intellectual vacuum, but it always starts with a basic set of assumptions and general knowledge from the relevant scientific fields. (b) The pre-posterior (meta-prior) stage, in which case-specific knowledge is collected. (c) The posterior stage, in which results from (a) and (b) are processed to produce the required space/time map.

With this epistemological framework in mind, we would want environmental solutions that are informative (i.e., they carry as much information as possible), well-evidenced, and cogent with a high probability of being correct (rather than certain correctness, which is an aspect that cannot be guaranteed before the event). BME aims at a proper balance between informativeness and cogency. Both requirements involve probabilities, but are conditional probabilities relative to different knowledge bases. At the prior stage the information we seek to maximize is conditioned to the general knowledge available. This stage assumes an inverse relation between information and probability: the more informative is an assessment about a mapping situation, the less probable it is to occur. This expresses a standard epistemic rule: the more vague and general an assumption is, the more alternatives it includes (it is, hence, more probable), but also the less informative is, and vice versa. At the pre-posterior stage we collect and organize additional case-specific knowledge that can be incorporated into the BME formulation. In real world applications, we may be dealing with a variety of possible knowledge bases, which can make the analysis at the pre-posterior stage not a trivial task (Christakos, 1998). At the posterior stage, the probability is conditioned to the case-specific knowledge and perceptual evidence. This conditioning assumes a connection between mapping predictions and the case-specific evidence available (through, e.g., physical laws or detectable space/time patterns). While we seek posterior predictions that are highly probable, we nevertheless want them to achieve this probability on the basis of case-specific knowledge and not on in terms of general knowledge alone.

A general BME formulation is summarized in Table 1. The coefficients $\mu_a (p_{\text{map}})$, $\alpha = 0, 1, \ldots, N_\alpha$, in Eqs. (1) depend on the general knowledge available; and the operator $\Psi[\cdot]$ in Eqs. (2) depends on the case-specific knowledge. The $A$ is a normalization constant and $f_{\chi}(\chi_k|\chi_{\text{data}})$ is the multi-point posterior probability density function (pdf), in which $\chi_{\text{data}} = \chi_{\text{hard}} \cup \chi_{\text{soft}}$ serves as a pointer to a context of knowledge and $\chi_k|\chi_{\text{data}}$ stands for the possible values $\chi_k = [X_{k1}, \ldots, X_{kN}]^T$ of the map for the knowledge bases specified by $\chi_{\text{data}}$. The BME steps are as follows:
Table 1: Mathematical formulation of the BME space/time mapping approach

\[
\begin{align*}
\mathbf{g}_\alpha(p_{\text{map}}) &= \left\{ d\mathbf{X}_{\text{map}} g_{\alpha}(\mathbf{X}_{\text{map}}) \exp \hat{\mathbb{S}}[\mathbf{X}_{\text{map}}; p_{\text{map}}] \right\}, \quad \alpha = 0,1,\ldots, N_i \\
\hat{\mathbb{S}}[\mathbf{X}_{\text{map}}; p_{\text{map}}] &= \sum_{\alpha=0}^{N_i} \mu_{\alpha}(p_{\text{map}}) g_{\alpha}(\mathbf{X}_{\text{map}}) \\
\frac{\partial}{\partial \mathbf{x}_i} \psi[\mathbf{X}_{\text{soft}}, \exp \hat{\mathbb{S}}[\mathbf{X}_{\text{map}}; p_{\text{map}}]]_{\mathbf{x}_i = \hat{\mathbf{x}}_i} &= 0, \quad i = 1,\ldots,P
\end{align*}
\]  

(i) Eqs. (1) are solved for \( \mu_{\alpha}(p_{\text{map}}) \). (ii) The \( \mu_{\alpha}(p_{\text{map}}) \) are substituted into Eqs. (2), which are then solved for the unknown estimates \( \mathbf{X}_k = \hat{\mathbf{x}}_k \), with \( \hat{\mathbf{x}}_k = [\hat{x}_{k_1} \ldots \hat{x}_{k_P}]^T \) (the number \( P \) of points to be estimated simultaneously may depend on theoretical, computational and physical considerations). (iii) The posterior pdf is obtained from Eq. (3). To assure maximum pdf estimates, the Hessian of the pdf at \( \mathbf{x}_k = \hat{\mathbf{x}}_k \) must be negative-definite. Clearly, the single-point estimation is a special case of Table 1 for \( P = 1 \). While multi-point mapping offers several interrelated estimates at a time, single-point mapping gives one estimate at a time independently of its neighboring estimates; the former is, hence, more informative than the latter.

4 Two Applications

In the first application, BME is compared to simple kriging (SK, Olea, 1997). Porosity data were collected at the West Lyon field in West-Central Kansas. The geomorphology includes Mississippian (Lower Carboniferous) sediments deposited in the shallow epicontinental seas that covered much of the North America in the Late Paleozoic. The porosity data were collected over an area of approximately 2.5 x 4.5 miles\(^2\). A total of 76 data were available (Fig. 1). Our analysis of the porosity field is two-fold. In the first part we will allow BME to use all 76 hard data. The general knowledge consists of the mean and the covariance (Fig. 1) of the porosity data set. Based on these data, BME produces the porosity map of Fig. 2. Note that this BME map is the same as the SK map obtained using the same data, mean and covariance. In the second part of the analysis we used only 56 hard data, while at the remaining 20 points we assumed soft interval data with a width of 1 unit and centered at the measured porosity value (assumed unknown). BME analysis accounts for both hard and soft data, and produces the map of Fig. 3a; note that despite the uncertainty introduced by the soft data, it closely resembles the spatial structure of the previous map (Fig. 2). SK can use only hard data; hence, the map it produces (Fig. 3b) using the 56 hard data is considerably less accurate than the BME map.

In the second application, BME is compared to indicator kriging (IK: Journel, 1989). IK can incorporate interval soft data, though not in a systematic and rigorous way as BME does. For a fixed set of 7 spatial points (Fig. 4), 500 realizations of a
Figure 1: Complete porosity data configuration (measurements are available at points indicated by x); and the covariance of the porosity data.

Figure 2: The BME porosity map using 76 hard porosity data.
random field \( X(p) \) were generated which follow a \((0, 1)\) multivariate gaussian law; the covariance is \( c_x(h) = \exp[-3(h^2/a^2)] \), \( a = 2.7639 \). While the 5 points with known values are the hard data points, at 2 points (the same for all simulations) only soft data \( \chi(p_j) \in [\alpha_j, \alpha_{j+1}] \) or \( \chi(p_j) \notin [\alpha_j, \alpha_{j+1}] \) are considered \( (j = 1, \ldots, 13) \). The intervals are chosen so that \( \alpha_1 = -\infty \), \( \alpha_{13} = +\infty \), \( \alpha_2 = F^{-1}_x(0.01) \), \( \alpha_{12} = F^{-1}_x(0.99) \), \( \alpha_{i+2} = F^{-1}_x(0.1i) \), \( i = 1, \ldots, 9 \); \( F^{-1}_x(p) \) is the \( p \)-quantile of the zero mean/unit variance gaussian law. Estimation is sought at point \( p_k \) (the same for all simulations; Fig. 4). To avoid methodological issues, all parameters of the multivariate gaussian law are assumed known for both BME and IK. Also, the same intervals were chosen for all points, where the limits of the intervals are precisely the threshold values for IK. This last condition is, however, not required for BME which can deal with any combination of length intervals. Since complete information about the distribution is available, we use simple IK where the probabilities and indicator covariances are found directly from the bivariate gaussian law. BME provides directly an estimate (e.g., the mode) at point \( p_k \) as the solution of Eq. (2). Since neither the mode nor the mean can be reliably determined for the IK posterior distributions (e.g., non-monotonicity and extreme discretization of the cumulative distribution), we used the median of the distributions. The results are shown in Fig. 5. All estimated values

Figure 3: (a) BME porosity map using 56 hard and 20 soft data; (b) SK porosity map using 56 hard data.
Figure 4: Data configuration. Hard data are available at points indicated by ●; soft data are available at points ○; the estimation point is indicated by x.

Figure 5: Estimation error distributions: (a) BME; (b) IK.
have been centered with respect to the known values at \( p_k \) and, thus, each plot gives the corresponding estimation error distribution. Clearly, BME performs considerably better than IK, while it does not suffer any of the computational shortcomings of IK.

## 5 Conclusion - Discussion

The process that leads to maps of environmental phenomena is a combination of knowledge bases (theoretical concepts, physical knowledge, assumptions, models, etc.) that goes far beyond the accumulation and massaging of observational data. On the basis of a meaningful map, theoretical interpretation can lead to a useful picture of reality. The BME approach may be schematically represented as follows:

\[
\begin{array}{cccc}
\text{Physical knowledge} & \rightarrow & \text{Theory} & \rightarrow & \text{Spatiotemporal map} & \rightarrow & \text{Interpretation} & \rightarrow & \text{Picture of reality} \\
\hline
\end{array}
\]

Given the important connections between scientific explanation (interpretation) and environmental mapping (prediction), an ideal situation should consist of knowledge base-driven improvements in environmental mapping performance that can be explained within the context of our theoretical understanding. Ignoring the theoretical rationale underlying the mapping process can only damage our scientific interpretation of what the map represents. The lack of sound theoretical underpinnings and unifying principles is at the heart of the shortcomings of the cook-book approaches to geostatistical mapping. A mapping algorithm should be the end-result of an analysis that goes deeper into the fundamentals of the problem rather than a collection of techniques and procedures without any clear underlying rationale.

## Acknowledgment

The computational help offered by Dr. P. Bogaert and Mr. M. Serre is greatly appreciated. Thanks are also owed to Dr. R.A. Olea for providing the Lyon field data set. The work was supported by a grant from the Department of Energy (DE-FC09-93SR18262).

## References