Spatial Random Field Models Inspired from Statistical Physics with Applications in the Geosciences

Dionissios T. Hristopulos

Department of Mineral Resources Engineering
Technical University of Crete
Chania 73100, Greece

Abstract

The spatial structure of fluctuations in spatially inhomogeneous processes can be modeled in terms of Gibbs random fields. A local low energy estimator (LLEE) is proposed for the interpolation (prediction) of such processes at points where observations are not available. The LLEE approximates the spatial dependence of the data and the unknown values at the estimation points by low-lying excitations of a suitable energy functional. It is shown that the LLEE is a linear, unbiased, non-exact estimator. In addition, an expression for the uncertainty (standard deviation) of the estimate is derived.

PACS numbers: 02.50.Ga,02.50.Fz,05.40.-a,05.10.Ln,89.60.-k

Keywords: spatial correlation, excitations, hamiltonian, stochastic estimation

Electronic address: dionisi@mred.tuc.gr; URL: \texttt{http://www.mred.tuc.gr/dionisi.htm}

\footnote{This work is partially supported by the EPEAEK Program: Environment Pithagoras II.}
I. INTRODUCTION

Spatial random fields (SRF’s) have applications in hydrology [9, 11], oil reservoir engineering [4], environmental pollutant mapping and risk assessment [1], mining exploration and reserves estimation [3], as well as environmental health studies [2]. SRF’s model spatial correlations in variables such as mineral concentrations, dispersion of environmental pollutants, soil and rock permeability, and flow fields in oil reservoirs. Knowledge of spatial correlations enables (i) generating predictive iso-level contour maps (ii) estimating the uncertainty of predictions and (iii) developing simulations that partially reconstruct the process of interest. Geostatistics provides mathematical tools for these tasks. The classical approach is based on Gaussian SRF’s (GSRF’s) and various generalizations for non-Gaussian distributions [10, 13]. For GSRF’s the spatial structure is determined from the covariance matrix, which is estimated from the distribution of the data in space.

An SRF state (realization) can be decomposed into a deterministic trend \( m_x(s) \), a correlated fluctuation \( X_\lambda(s) \), and an independent random noise term, \( \epsilon(s) \) so that \( X(s) = m_x(s) + X_\lambda(s) + \epsilon(s) \). The trend represents large-scale variations of the field, which can be obtained in principle by ensemble averaging, i.e. \( m_x(s) = E[X(s)] \). In practice, the trend is often determined from a single available realization. The fluctuation term corresponds to ‘fast variations’ that reveal structure at small scales, which nonetheless exceed a cut-off \( \lambda \). The random noise represents non-resolved inherent variability due to resolution limits, purely random additive noise, or non-systematic measurement errors. It is typically assumed that the fluctuation is a second-order stationary SRF, or an intrinsic SRF with second-order stationary increments [15]. The observed SRF after detrending is a zero-mean fluctuation: \( X^*(s) = X_\lambda(s) + \epsilon(s) \).

In statistical physics the probability density function (pdf) of a fluctuation field \( x(s) \) governed by an energy functional \( H[x(s)] \) is expressed as \( f_x[x(s)] = Z^{-1} \exp \{-H[x(s)]\} \), where \( Z \) is the partition function. Using this representation, the Gaussian joint pdf in classical geostatistics is expressed in terms of the functional:

\[
H[x(s)] = \frac{1}{2} \int ds \int ds' x(s) \left[ G_x \right]^{-1}(s - s') x(s'). \tag{1}
\]

In Eq. (1), \( \left[ G_x \right]^{-1}(s - s') \) is the inverse of the covariance function \( G_x(s - s') \), which determines the spatial disorder. While statistical physics plays an increasingly important role in under-
standing the behavior of complex geophysical systems [12], its applications in geostatistical analysis have not yet been explored.

Spartan Spatial Random Fields (SSRF’s) model spatial correlations in terms of ‘inter-
actions’, in the spirit of Markov SRF’s [14]. In general properties and permissibility condi-
tions were derived for the fluctuation-gradient-curvature (FGC) SSRF model, with the following energy functional:

$$H_{fgc}[X_\lambda] = \frac{1}{2\eta_0 \xi} \int ds \left\{ [X_\lambda(s)]^2 + \eta_1 \xi^2 [\nabla X_\lambda(s)]^2 + \xi^4 [\nabla^2 X_\lambda(s)]^2 \right\}.$$ (2)

For this model, a moment-based method for parameter estimation was proposed and tested with simulated data; methods for SSRF non-constrained simulation were presented in [7]; systematic reduction of anisotropic disorder, based on the covariance tensor identity, was investigated in [5,8]. The FGC model [6] has three main parameters: the scale coefficient $\eta_0$, the covariance-shape coefficient $\eta_1$, and the correlation length $\xi$. Bochner’s theorem [1] for the covariance function requires $\eta_1 > -2$. A coarse-graining kernel is used to cut off the fluctuations at $k_c \propto \lambda^{-1}$ [6,7], leading to band-limited covariance spectral density and differentiable field configurations (in the mean square sense) [7].

II. OPERATOR NOTATION

Let $\Omega \in \mathbb{R}^d$ denote the area of interest and $A(\Omega)$ its boundary. Consider an SSRF defined over this area with parameters $\eta_0, \eta_1, \xi$, with a finite variance $\sigma_\xi^2$. Let us assume that it is possible to normalize the SSRF to unit variance by simply dividing the states with the standard deviation. Next, it is possible to express the pseudo-energy functional in terms of an operator notation notation as follows:

$$H[X_\lambda] \equiv \langle X_\lambda(s) | \mathcal{H} | X_\lambda(s) \rangle + S(A) \equiv \int_\Omega ds X_\lambda(s) \mathcal{H} [X_\lambda(s)] + S(A),$$ (3)

where $\mathcal{H}$ is a ‘pseudo-hamiltonian’ operator and $S(A)$ is a surface term. Assuming that the surface term is negligible, the eigenvalue equation becomes:

$$\mathcal{H} |\psi_E(s; b)\rangle = E \psi_E(s; b),$$ (4)

where $\psi_E(s; b)$ is an eigenfunction, $E$ is the corresponding energy and $b$ a degeneracy
vector index, which may include both discrete and continuous components. Since the SSRF has been normalized to unit variance, the eigenfunctions $\psi_E(s; b)$ can also be assumed normalized, i.e., $\int_{\Omega} ds \psi^2_E(s; b) = 1$, and then $H[X\lambda] = E$.

If Eq. (4) admits solutions for non-zero $E$, one can construct eigenfunctions that correspond to positive excitation energies $E$. The realization probability that corresponds to low-lying excitations is high. Hence, the main idea is to consider the observed state or the union of the observations and the predictions as being locally represented by an excited state. This approach can be used for both parameter estimation and prediction (spatial estimation)

A. Eigenfunctions for FGC case

For the FGC functional of Eq. (2), integrating the square-gradient term by parts leads to the following equation:

$$\int_{\Omega} ds \left[ \nabla \psi_E(s; b) \right]^2 = - \int_{\Omega} ds \psi_E(s; b) \nabla^2 \psi_E(s; b) + \int_{A(\Omega)} da \cdot \nabla \psi_E(s; b) \psi_E(s; b).$$

In Eq. (5) $\int_{A(\Omega)} da$ denotes the surface integral on the boundary of the area of interest. Secondly, using Green’s theorem on the square-curvature term one obtains

$$\int_{\Omega} ds \left[ \nabla^2 \psi_E(s; b) \right]^2 = \int_{\Omega} ds \psi_E(s; b) \nabla^4 \psi_E(s; b) + \int_{A(\Omega)} da \cdot \nabla \psi_E(s; b) \nabla^2 \psi_E(s; b)
- \int_{A(\Omega)} da \cdot \nabla \left[ \nabla^2 \psi_E(s; b) \right] \psi_E(s; b).$$

Hence, in the operator notation the FGC functional is expressed as follows:

$$\mathcal{H}_{\text{fgc}} = \frac{1}{2\eta_0 \xi^4} \left[ 1 - \eta_1 \xi^2 \nabla^2 + \xi^4 \nabla^4 \right],$$

and the surface term is given by:

$$S(\Omega) = \frac{1}{2\eta_0 \xi^4} \left[ \eta_1 \xi^2 \int_{A(\Omega)} da \cdot \nabla \psi_E(s; b) \psi_E(s; b)
+ \xi^4 \int_{A(\Omega)} da \cdot \nabla \psi_E(s; b) \nabla^2 \psi_E(s; b)
- \xi^4 \int_{A(\Omega)} da \cdot \nabla \left[ \nabla^2 \psi_E(s; b) \right] \psi_E(s; b) \right].$$
If the units are chosen so that $2\eta_0\xi^d = 1$ and the surface term is ignored, the eigenvalue equation is given by the following partial differential equation (pde):

$$
\psi_E(s; b) - \eta_1 \xi^2 \nabla^2 \psi_E(s; b) + \xi^4 \nabla^4 \psi_E(s; b) = E \psi_E(s; b).
$$

The eigenfunctions $\psi_E(s; b)$ of Eq. (9) are given by the following four plane waves:

$$
\psi_E(s; b) = e^{k_j \cdot s}, \quad k_j = k_j \hat{\theta},
$$

where $\hat{\theta}$ represents the unit direction vector, and $k_j$ the magnitudes of the characteristic wave-vectors that are given by the roots of the fourth-order characteristic polynomial:

$$
\Pi_{bsc}(k\xi) = (1 - E) - \eta_1 \xi^2 k^2 + \xi^4 k^4 = 0.
$$

Thus, the characteristic wavevectors are given by the following expressions:

$$
k_1(\eta_1, \xi, E) = \frac{1}{\sqrt{2\xi}} \sqrt{\eta_1 + \sqrt{\eta_1^2 - 4(1 - E)}}
$$

$$
k_2(\eta_1, \xi, E) = -\frac{1}{\sqrt{2\xi}} \sqrt{\eta_1 + \sqrt{\eta_1^2 - 4(1 - E)}}
$$

$$
k_3(\eta_1, \xi, E) = \frac{1}{\sqrt{2\xi}} \sqrt{\eta_1 - \sqrt{\eta_1^2 - 4(1 - E)}}
$$

$$
k_4(\eta_1, \xi, E) = -\frac{1}{\sqrt{2\xi}} \sqrt{\eta_1 - \sqrt{\eta_1^2 - 4(1 - E)}}.
$$

Note that only the magnitude of the wave-vectors is determined from the pde (9). This is due to the fact that isotropic spatial dependence was assumed in the SSRF model.

(a) If $\eta_1 > 0 \land 1 - \eta_1^2/4 < E < 1$ all the roots are real. (b) If $\eta_1 > 0 \land E > 1$, then $k_1, k_2$ are real, while $k_3, k_4$ are purely imaginary. (c) If $\eta_1 > 0 \land 1 - \eta_1^2/4 > E$, then all the roots are complex. (d) If $\eta_1 < 0 \land 1 - \eta_1^2/4 < E < 1$, then all the roots are imaginary. (e) If $\eta_1 < 0 \land E > 1$, then $k_1, k_2$ are real, while $k_3, k_4$ are imaginary. (f) If $\eta_1 < 0 \land 0 < E < 1 - \eta_1^2/4$ all the roots are complex.

In general, an excited state formed by the linear superposition of degenerate eigenstates of energy $E$ is given by the expression:

$$
Z_E(s; c_b) = \sum_{j=1}^{4} u(k_c - ||k_j||) \int d\hat{\theta} c_j(\hat{\theta}) \exp \left( k_j \hat{\theta} \cdot s \right),
$$

(16)
where \( c_j(\hat{\theta}) \) is a direction-dependent (possibly complex-valued) function, \( \|k_j\| \) is the modulus of the characteristic wavevector, and \( u(.) \) is the unit step function, used to guarantee that the fluctuations in the excited state do not exceed the cutoff ‘frequency’. For the estimation of real-valued processes, the coefficients \( c_j(\hat{\theta}) \) are constrained to give real values for the excited state \( Z_E(s; c_b) \). If \( c_j(\hat{\theta}) = c_j \), an isotropic excited state is obtained, which can be expressed as \( Z_E(s; c_1, \ldots, c_4) = \sum_{j=1}^4 c_j u(k_c - \|k_j\|) \psi_E(s; j) \), where \( \psi_E(s; j) = \int d\hat{\theta} \exp\left(k_j \hat{\theta} \cdot s\right) \).

### B. Eigenstates in \( d = 1 \)

We examine in more detail the real-valued eigenstates that are trigonometric or hyperbolic functions in the one-dimensional domain \([0, L] \in \mathbb{R}\).

1. **Exponential Eigenstates**

For characteristic wave-vectors \( k \) that are real numbers, the normalized eigenfunctions and the corresponding energies of Eq. (9) are given by

\[
X(s) = e^{-ks} \sqrt{\frac{2k}{1 - e^{-2kL}}} , \quad E = 1 - \eta_1(k \xi)^2 + (k \xi)^4 .
\]

However, if the exponential function is inserted in Eq. (2), the resulting energy is given by

\[
H[X(s)] = 1 + \eta_1(k \xi)^2 + (k \xi)^4 .
\]

The difference between the energy given by Eq. (17) and the correct energy, given by Eq. (19) is due to the fact that the boundary term can not be ignored for the localized exponential excitation.

2. **Trigonometric Eigenstates**

If \( k \) is an imaginary number, the eigenfunctions are trigonometric functions. A normalized cosine eigenfunction and the corresponding energy are given by:
\[ X(s) = \cos(k s) \sqrt{\frac{2}{L \left[ 1 + \text{sinc}(2kL) \right]}}, \]  
(20)

\[ E = 1 + \eta_1 (k \xi)^2 \frac{1 - \text{sinc}(2kL)}{1 + \text{sinc}(2kL)} + (k \xi)^4. \]  
(21)

For large domains, \( kL \gg 1 \), Eq. (21) is practically equivalent to Eq. (19). As expected, in the case of an extended eigenstate (as the cosine) the boundary term can be ignored.

### III. SPATIAL ESTIMATION WITH SSRF’S

Assume \( S_m = (s_1, \ldots, s_N) \) is a set of data points with the respective vector of measurements denoted by \( \mathbf{X}^* = (X_1^*, \ldots, X_N^*) \); let \( s_0 \notin S_m \) be the estimation point and \( \hat{X}_\lambda(s_0) \) the estimate (spatial prediction). The local neighborhood of \( s_0 \) is the set \( S_0 \equiv B(s_0; r_c) \) of all the data points \( s_j, j = 1, \ldots, M \) inside a ‘sphere’ of radius equal to one correlation range from \( s_0 \). In geostatistics, \( \hat{X}(s_0) \) is determined by optimal linear filters (kriging estimators) \[9, 13\], which form the estimate as a superposition of the data values inside the local neighborhood, and there is no explicit resolution scale. The coefficients of the superposition are selected to make the estimate unbiased and to minimize the mean square error. Kriging is an exact interpolator, meaning that for any \( s_i \in S_m, \hat{X}(s_i) = \mathbf{X}^*(s_i) \). Exactitude is not always desirable, since it ignores measurement errors and leads to excessive smoothing of the fluctuations. Hence, different estimation methods are useful. The SSRF models can be used in kriging algorithms to provide new, differentiable covariance functions. In addition, within the SSRF framework it is possible to define a new type of estimator.

#### A. Low Local Energy Estimators

The central idea is that a ‘good’ estimate should correspond to a state with significant probability of realization. If the energy functional is non-negative, as in Eq. (2), the highest probability is associated with the uniform state \( X_\lambda(s) = 0 \), which is not physically interesting. Other states with high probability correspond to low-energy excitations. Let us superimpose the degenerate eigenstates with energy \( E \) to form a mixed state \( Z_E(s; c) = \sum_{i=1}^{D} c_i \psi_E(s; b_i) \); \( c = (c_1, \ldots, c_D) \) is a \( D \)-dimensional vector of linear coefficients.
that correspond to the degeneracy indices. In principle $D$ can be infinite since the directional dependence given by Eq. (10) is continuous. However, in practice it may be simplest to restrict the search to one ‘optimal’ direction. The energy $H[Z_E(s; c)]$ of the mixed state is not necessarily equal to $E$. In fact, for orthonormal eigenstates $H[Z_E(s; c)] = \mu E$, where $\mu = \sum_{i=1}^{D} c_i$. This reflects the fact that the ‘energy level’ of the observed process is set by the measurements (i.e., the coefficients $c_i$). Since the scale coefficient $\eta_0$ is inversely proportional to the magnitude of the fluctuations, it follows that $\mu^{-1} \propto \eta_0$. It should also be noted that if two mixed states $(c_1, E_1)$ and $(c_2, E_2)$ are energetically equivalent, i.e., $\mu_1 E_1 = \mu_2 E_2$, they are not in general linearly related, since according to Eqs. (10), (12)-(15) and (16), the dependence of $Z_E(s; c)$ on $E$ is nonlinear.

We propose that the observations for $s_j \in B(s_0; r_c)$ be expressed as $X^*(s_j) = Z_E(s_j; c_0) + \varepsilon(s_j)$, where $Z_E(s; c_0)$ is a ‘local’ excitation and $\varepsilon(s_j)$ is the local excitation residual. Local dependence stems from the fact that the coefficients $c_0$ depend on $s_0$, in contrast with the solution of Eq. (16), in which the coefficient vector is global. The LLEE estimator is then given by $\hat{X}_\lambda(s_0) = Z_E(s_0; c_0)$. Since $Z_E(s; c_0)$ is an estimate of the underlying process $X_\lambda(s)$, the excitation residual $\varepsilon(s_j)$ is not in general the same as the noise $\epsilon(s)$. The coefficients $c_0$, follow from minimizing the mean square excitation residual inside $B(s_0; r_c)$, i.e.,

$$c_0 = \arg \min_c \sum_{j=1}^{M} [X^*(s_j) - Z_E(s_j; c)]^2. \quad (22)$$

The above is a typical problem of multiple linear regression, where the regressors are the functions $\psi_E(s_i; b_j)$. If we define the $M \times D$ matrix $\psi_{E,ij} \equiv \psi_E(s_i; b_j)$, the solutions for $c_{0,i}$ and the LLEE are given by:

$$\alpha_{ik} = \sum_{j=1}^{M} \psi_{E,ji} \psi_{E,jk}, \quad i = 1, \ldots, M; \quad k = 1, \ldots, D, \quad (23)$$

$$c_{0,i} = \sum_{k=1}^{D} [\alpha]_{ik}^{-1} \sum_{l=1}^{M} \psi_{E,ik} X^*_l, \quad i = 1, \ldots, M, \quad (24)$$

$$\hat{X}_\lambda(s_0) = w_0 \cdot \mathbf{X}^*, \quad (25)$$
where \( w_0 \) is a weight vector given by:

\[
 w_{0,i} = \sum_{k=1}^{D} \psi_{E,0k} \sum_{j=1}^{D} \left[ \alpha \right]^{-1}_{kj} \psi_{E,ij}, \quad i = 1, \ldots, M. \tag{26}
\]

The uncertainty of the LLEE estimate is determined from the ensemble variance of the local excitation residual \( \sigma^2_\varepsilon(s_0) = E \left[ X^*(s_0) - \hat{X}(s_0) \right]^2 \), i.e.:

\[
 \sigma^2_\varepsilon(s_0) = \sigma^2_{x^*} + \sum_{i=1}^{M} \sum_{j=1}^{M} w_{0,i} w_{0,j} G_{x^*,ij} - 2 \sum_{i=1}^{M} w_{0,i} G_{x^*,0i}, \tag{27}
\]

where \( G_{x^*,ij} = E \left[ X_i^* X_j^* \right] \) is the covariance matrix at the observation points, \( G_{x^*,0i} = E \left[ X_i^* X_0^* \right] \), is the covariance vector of the fluctuations between \( s_0 \) and the estimation point, and \( \sigma^2_{x^*} = E \left[ X_i^* X_i^* \right] \) is the variance of the observed process.

**B. Properties of the LLEE**

It follows from Eqs. (25) and (26) that the LLEE is linear in the fluctuations. Hence, the estimates are unbiased and follow the Gaussian law (if the observations are normally distributed). Kriging methods are based on minimization of the (ensemble) mean square error, which is a global optimality criterion. In contrast, the LLEE criterion is local (i.e., minimum of the average squared excitation residual in the neighbourhood of the estimation point). Another difference with kriging is that low local energy estimates do not match exactly the measurements at observation points. The property of non-exactitude is maintained even when the noise can be ignored. Finally, unlike kriging predictions, the LLEE provides multiple estimates, since different energy levels lead to different excited states. In this respect the LLEE is similar to a simulation method. However, simulations involve the generation of random numbers, in contrast with the LLEE method. It should also be noted that the energy of local excitations is not necessarily the energy of the estimated state, because the locality of the coefficient vector \( c_0 \) means that the operators \( \nabla \) and \( \nabla^2 \) contribute to the overall energy when they act on the coefficients of the mixed state in Eq. (2).

**IV. CONCLUSIONS**

A spatial estimation method for applications in the geosciences is presented. The method is based on the use of ‘pseudo-energy’ functionals, motivated by explicit constraints or
heuristic physical arguments, to capture the spatial heterogeneity of the observed process. Estimates of the process at unmeasured points (predictions) are based on local interpolating functions that represent low-energy excitations of the pseudo-energy. Multiple estimates of the process can be generated by considering local interpolating functions that correspond to different excitation energies.

[1] G. Christakos, *Random Field Models in Earth Sciences*, (Academic Press, San Diego, 1992).
[2] G. Christakos, and D. T. Hristopulos, *Spatiotemporal Environmental Health Modelling*, (Kluwer, Boston, 1998).
[3] P. Goovaerts, *Geostatistics for Natural Resources Evaluation*, (Oxford, NY, 1997).
[4] M. E. Hohn, *Geostatistics and Petroleum Geology*, (Kluwer, Dordrecht, 1999).
[5] D. T. Hristopulos, Stoch. Environ. Res. Risk Assessment, 16, 43 (2002).
[6] D. T. Hristopulos, SIAM J. Sci. Comput. 24 2125 (2003).
[7] D. T. Hristopulos, in *Proceedings of the International Conference of Computational Methods in Sciences and Engineering, Kastoria, Greece, 2003*, edited by T. E. Simos, (World Scientific, London, UK), p. 242; J. Comput. Methods Sci. Eng., 5(2) (2005) in print.
[8] D. T. Hristopulos, in *Proceedings of the 1st International Conference on Advances in Mineral Resources Management and Environmental Geotechnology, Chania, Greece, 2004*, edited by Z. Agioutantis and K. Komnitsas, (Heliotopos Conferences, Greece, 2004), p. 127; M. Galetakis and D. T. Hristopulos, *ibid.*, p. 133.
[9] P. K. Kitanidis, *Introduction to Geostatistics: Applications to Hydrogeology*, (Cambridge, 1997).
[10] C. Lantuejoul, *Geostatistical Simulation: Models and Algorithms*, (Springer, Berlin, 2002).
[11] Y. Rubin, *Applied Stochastic Hydrogeology*, (Oxford University Press, New York, 2003).
[12] D. Sornette, *Critical Phenomena in Natural Sciences*, (Springer, Berlin, 2004).
[13] H. Wackernagel, *Multivariate Geostatistics*, (Springer, Berlin, 2003).
[14] G. Winkler, *Image Analysis, Random Fields and Dynamic Monte Carlo Methods: A Mathematical Introduction*, (Springer, New York, 1995).
[15] M. Yaglom, *Correlation Theory of Stationary and Related Random Functions I*, (Springer, New York, 1987).