

Prediction and Conditional Simulation of a 2D Lognormal Diffusion Random Field

R. Gutiérrez · C. Roldán ·
R. Gutiérrez-Sánchez · J. M. Angulo

Received: 22 September 2005 / Revised: 2 April 2006
Accepted: 9 May 2006 / Published online: 29 March 2007
© Springer Science + Business Media, LLC 2007

Abstract This paper describes techniques for estimation, prediction and conditional simulation of two-parameter lognormal diffusion random fields which are diffusions on each coordinate and satisfy a particular Markov property. The estimates of the drift and diffusion coefficients, which characterize the lognormal diffusion random field under certain conditions, are used for obtaining kriging predictors. The conditional simulations are obtained using the estimates of the drift and diffusion coefficients, kriging prediction and unconditional simulation for the lognormal diffusion random field.

Keywords Conditional simulation · Diffusion random field · Kriging · Lognormal diffusion process

AMS 2000 Subject Classifications 60J60 · 62M05 · 62M30

R. Gutiérrez · R. Gutiérrez-Sánchez · J. M. Angulo
Department of Statistics and Operations Research, University of Granada,
Campus de Fuentenueva s/n, E-18071 Granada, Spain

R. Gutiérrez
e-mail: rgjaimez@ugr.es

R. Gutiérrez-Sánchez
e-mail: ramongs@ugr.es

J. M. Angulo
e-mail: jmangulo@ugr.es

C. Roldán (✉)
Department of Statistics and Operations Research, University of Jaén,
Paraje Las Lagunillas s/n, B-3, Dep. 074, E-23071 Jaén, Spain
e-mail: iroldan@ujaen.es

1 Introduction

Lognormal random fields represent the technically more complex stage of lognormal modelling. Problems as parameter estimation, lognormal kriging, estimation based on lognormal maximum entropy, simulation, among others, are generally undertaken introducing the lognormal random field as a random field whose logarithm is a Gaussian random field, without reference to any specific diffusion structure (see, for example, Christakos 1992; Cressie 1993). When the structure of the lognormal random field is a diffusion, the latter approach is useful in the analysis of first passage through barriers, to obtain associated Itô equations and both Kolmogorov's forward-backward equations, etc.

Nowadays, increased emphasis in the formulation of dynamic models for the transport of chemical or biological agents leads to study diffusions and, particularly, Gaussian and lognormal diffusion random fields. Among the contributions to theoretical foundations for 2D diffusion random fields, see Nualart (1983). In this context, Gutiérrez et al. (2005) have considered lognormal random field models which are diffusions on each coordinate. The drift, involving exogenous factors, and diffusion coefficient, which characterize a 2D lognormal diffusion under certain conditions, are estimated by maximum likelihood. For data on a regular grid, an alternative method is proposed to estimate the diffusion coefficient.

In this work, the estimates of the drift and the diffusion coefficients given in Gutiérrez et al. (2005) are used for obtaining predictions and conditional simulations. The contents are organized as follows. First, the 2D lognormal random field model is introduced. Second, estimation of the drift and diffusion coefficients based on a discrete finite set of data is given. Finally, aspects related to kriging and conditional simulation are addressed and illustrated.

2 Lognormal Diffusion Random Fields

Lognormal diffusion processes are commonly used in the formulation of dynamic models applied in Environmental Sciences and especially in the analysis of economic variables. When the parameter space is a subset of \mathbf{R}_+^2 , Nualart (1983) introduced a class of two-parameter random fields which are diffusions on each coordinate and satisfy a particular Markov property related to partial ordering in \mathbf{R}_+^2 . Using this theory, a 2D lognormal diffusion random field can be introduced as follows.

Let $\{X(\mathbf{z}) : \mathbf{z} = (s, t) \in I = [0, S] \times [0, T] \subset \mathbf{R}_+^2\}$ be a positive-valued Markov random field, defined on a probability space (Ω, \mathcal{A}, P) , where $X(0, 0)$ is assumed to be constant or a lognormal random variable with $E[\ln X(0, 0)] = \phi_0$ and $\text{var}(\ln X(0, 0)) = \sigma_0^2$. The distribution of the random field is determined by the following transition probabilities:

$$\begin{aligned} &P(B, (s+h, t+k) \mid (x_1, x, x_2), \mathbf{z}) \\ &= P[X(s+h, t+k) \in B \mid X(s, t+k) = x_1, X(\mathbf{z}) = x, X(s+h, k) = x_2], \end{aligned}$$

where $\mathbf{z} = (s, t) \in I, h, k > 0, (x_1, x, x_2) \in \mathbf{R}_+^3$ and B is a Borel subset. We suppose that the transition densities exist and are given by

$$g(y, (s + h, t + k) \mid (x_1, x, x_2), \mathbf{z}) = \frac{1}{y\sqrt{2\pi\sigma_{\mathbf{z},h,k}^2}} \exp \left\{ -\frac{1}{2} \left(\frac{\ln \left(\frac{yx}{x_1x_2} \right) - m_{\mathbf{z},h,k}}{\sigma_{\mathbf{z},h,k}} \right)^2 \right\},$$

for $y \in \mathbf{R}_+$, with

$$m_{\mathbf{z},h,k} = \int_s^{s+h} \int_t^{t+k} \tilde{a}(\sigma, \tau) d\sigma d\tau, \quad \sigma_{\mathbf{z},h,k}^2 = \int_s^{s+h} \int_t^{t+k} \tilde{B}(\sigma, \tau) d\sigma d\tau,$$

and with \tilde{a} and \tilde{B} being continuous functions on I .

The random field $\{Y(\mathbf{z}) : \mathbf{z} \in I\}$ defined as $Y(\mathbf{z}) = \ln X(\mathbf{z})$ is then a Gaussian diffusion random field, with \tilde{a} and \tilde{B} being, respectively, the drift and diffusion coefficients, and $\tilde{a}_1, \tilde{a}_2, \tilde{B}_1$ and \tilde{B}_2 being the corresponding one-parameter drift and diffusion coefficients, where

$$\begin{aligned} \tilde{a}_1(s, t) &= \int_0^t \tilde{a}(s, \tau) d\tau, & \tilde{B}_1(s, t) &= \int_0^t \tilde{B}(s, \tau) d\tau, \\ \tilde{a}_2(s, t) &= \int_0^s \tilde{a}(\sigma, t) d\sigma, & \tilde{B}_2(s, t) &= \int_0^s \tilde{B}(\sigma, t) d\sigma, \end{aligned}$$

for all $\mathbf{z} = (s, t) \in I, x \in \mathbf{R}_+$. Furthermore, if $\mathbf{z} = (s, t), \mathbf{z}' = (s', t') \in I$, using the Markov property and the series expansion of the characteristic function we can obtain the distribution of the random vector $(Y(s', t'), Y(s, t'), Y(s', t), Y(s, t))$ and conclude that

$$\begin{aligned} m_Y(\mathbf{z}) &:= E[Y(\mathbf{z})] = \phi_0 + \int_0^s \int_0^t \tilde{a}(\sigma, \tau) d\sigma d\tau, \\ \sigma_Y^2(\mathbf{z}) &:= var(Y(\mathbf{z})) = \sigma_0^2 + \int_0^s \int_0^t \tilde{B}(\sigma, \tau) d\sigma d\tau, \\ c_Y(\mathbf{z}, \mathbf{z}') &:= cov(Y(\mathbf{z}), Y(\mathbf{z}')) = \sigma_Y^2(\mathbf{z} \wedge \mathbf{z}'), \end{aligned}$$

where we write $\mathbf{z} \wedge \mathbf{z}'$ for $(s \wedge s', t \wedge t')$, with ‘ \wedge ’ denoting the minimum.

Therefore we can assert that $\{X(\mathbf{z}) : \mathbf{z} \in I\}$ is a lognormal diffusion random field, and the one-parameter processes which appear fixing each coordinate of the parameter space are lognormal diffusions with one-parameter drift and diffusion coefficients given by

$$\begin{aligned} a_1(\mathbf{z})x &:= \left(\tilde{a}_1(\mathbf{z}) + \frac{1}{2} \tilde{B}_1(\mathbf{z}) \right) x, & B_1(\mathbf{z})x^2 &:= \tilde{B}_1(\mathbf{z})x^2, \\ a_2(\mathbf{z})x &:= \left(\tilde{a}_2(\mathbf{z}) + \frac{1}{2} \tilde{B}_2(\mathbf{z}) \right) x, & B_2(\mathbf{z})x^2 &:= \tilde{B}_2(\mathbf{z})x^2. \end{aligned}$$

These coefficients can be calculated, following their definition, as the infinitesimal moments of first and second order.

Under suitable regularity conditions, it is possible to obtain a SPDE (stochastic partial differential equation) formulation for a 2D diffusion random field. In fact, we need hypotheses *I–V* stated in Nualart (1983) to be satisfied, in order to apply *Theorem 2.8* established in that paper. These hypotheses and the uniqueness of solution have been proved by the authors to hold for the lognormal diffusion random field considered. Thus, there exists a 2D Wiener random field $\{W(\mathbf{z}) : \mathbf{z} \in I\}$ (adjoining, if it is necessary, a new probability space) such that $\{X(\mathbf{z}) : \mathbf{z} \in I\}$ is the only diffusion random field satisfying the following SPDE:

$$\begin{aligned} & \frac{\partial^2 X(s, t)}{\partial s \partial t} - X^{-1}(s, t) \frac{\partial X(s, t)}{\partial s} \frac{\partial X(s, t)}{\partial t} - \frac{\partial a_2(s, t)}{\partial s} X(s, t) \\ & = \left(\frac{\partial B_2(s, t)}{\partial s} + B_1(s, t) B_2(s, t) \right)^{1/2} X(s, t) \frac{\partial^2 W(s, t)}{\partial s \partial t}. \end{aligned}$$

This aspect is not essential for the approach considered in this work, although it provides an alternative interesting interpretation of the random field formulation considered.

Henceforth we will assume that the conditions usually considered for estimation of the drift and diffusion coefficients in the one-parameter case hold; that is, $P[\ln X(0, 0) = \phi_0] = 1$ (i.e. $\sigma_0^2 = 0$) and $\sigma_Y^2(\mathbf{z}) = \tilde{B}st, \mathbf{z} = (s, t) \in I$.

3 Estimation of the Drift and Diffusion Coefficients

Let $\{X(\mathbf{z}) : \mathbf{z} \in I\}$ be a lognormal diffusion random field and $\{Y(\mathbf{z}) : \mathbf{z} \in I\}$ the associated Gaussian diffusion random field. Let us denote

$$\mathbf{m}_Y = (m_Y(\mathbf{z}_1), \dots, m_Y(\mathbf{z}_n))^t, \quad \Sigma_Y = (\sigma_Y^2(\mathbf{z}_i \wedge \mathbf{z}_j))_{i, j=1, \dots, n}$$

with $\sigma_Y^2(\mathbf{z}_i \wedge \mathbf{z}_j) = \tilde{B}(s_i \wedge s_j)(t_i \wedge t_j)$.

Data $X(\mathbf{z}_1), \dots, X(\mathbf{z}_n)$ are assumed to be observed at known spatial locations $\mathbf{z}_1 = (s_1, t_1), \mathbf{z}_2 = (s_2, t_2), \dots, \mathbf{z}_n = (s_n, t_n) \in I$ and let $\mathbf{x} = (x_1, x_2, \dots, x_n)^t$ be a sample.

3.1 MLEs for the Drift and Diffusion Coefficients Using Exogenous Factors

In this subsection it will be assumed that $m_Y(\mathbf{z}) = E[Y(\mathbf{z})]$ is an unknown linear combination of known functions, set $\{f_0(\mathbf{z}), f_1(\mathbf{z}), \dots, f_p(\mathbf{z}) : \mathbf{z} \in I\}$. For example, considering polynomial representations for the trend surface

$$m_Y(\mathbf{z}) = \sum_{0 \leq k+l \leq r} \phi_{kl} s^k t^l, \quad \mathbf{z} = (s, t),$$

for some appropriate choice of r .

In this case, the drift coefficient \tilde{a} of Y is a linear combination of several known functions, set $\{h_1(\mathbf{z}), \dots, h_p(\mathbf{z}) : \mathbf{z} \in I\}$, with real coefficients ϕ_1, \dots, ϕ_p :

$$\tilde{a}(\mathbf{z}) = \sum_{\alpha=1}^p \phi_\alpha h_\alpha(\mathbf{z}), \quad \mathbf{z} \in I.$$

Considering, for $\mathbf{z} = (s, t) \in I$,

$$f_0(\mathbf{z}) = 1, \quad f_\alpha(\mathbf{z}) = \int_0^s \int_0^t h_\alpha(\sigma, \tau) d\sigma d\tau, \quad \alpha = 1, \dots, p,$$

the mean of Y is given by

$$m_Y(s, t) = \phi_0 + \sum_{\alpha=1}^p \phi_\alpha \int_0^s \int_0^t h_\alpha(\sigma, \tau) d\sigma d\tau = \sum_{\alpha=0}^p \phi_\alpha f_\alpha(\mathbf{z}).$$

Let us denote $\mathbf{F} = (\mathbf{f}_0, \mathbf{f}_1, \dots, \mathbf{f}_p)$, with $\mathbf{f}_\alpha = (f_\alpha(\mathbf{z}_1), f_\alpha(\mathbf{z}_2), \dots, f_\alpha(\mathbf{z}_n))^t$, for $\alpha = 0, 1, \dots, p$, and $\boldsymbol{\phi} = (\phi_0, \phi_1, \dots, \phi_p)^t$. Thus, we have

$$\mathbf{m}_Y = (\phi_0 \mathbf{f}_0 + \phi_1 \mathbf{f}_1 + \dots + \phi_p \mathbf{f}_p) = \mathbf{F}\boldsymbol{\phi}.$$

Writing

$$\Sigma_Y = \tilde{\mathbf{B}}\mathbf{M} := \tilde{\mathbf{B}} \begin{pmatrix} s_1 t_1 & (s_1 \wedge s_2)(t_1 \wedge t_2) & \cdots & (s_1 \wedge s_n)(t_1 \wedge t_n) \\ (s_1 \wedge s_2)(t_1 \wedge t_2) & s_2 t_2 & \cdots & (s_2 \wedge s_n)(t_2 \wedge t_n) \\ \vdots & \vdots & \ddots & \vdots \\ (s_1 \wedge s_n)(t_1 \wedge t_n) & (s_2 \wedge s_n)(t_2 \wedge t_n) & \cdots & s_n t_n \end{pmatrix},$$

and differentiating the appropriate log-likelihood function with respect to $\boldsymbol{\phi}$ and $\tilde{\mathbf{B}}$ (see Gutiérrez et al. 2005), the MLEs for the drift and diffusion coefficients are, respectively,

$$\boldsymbol{\phi}^* = (\phi_0^*, \phi_1^*, \dots, \phi_p^*)^t = (\mathbf{F}^t \mathbf{M}^{-1} \mathbf{F})^{-1} \mathbf{F}^t \mathbf{M}^{-1} \ln \mathbf{x} \tag{1}$$

and

$$\tilde{\mathbf{B}}^* = \frac{1}{n} (\ln \mathbf{x} - \mathbf{m}_Y^*)^t \mathbf{M}^{-1} (\ln \mathbf{x} - \mathbf{m}_Y^*), \tag{2}$$

where $\mathbf{m}_Y^* = \mathbf{F}\boldsymbol{\phi}^*$.

3.2 Estimation of the Diffusion Coefficient from Data on a Regular Grid

Let us expound an alternative method to estimate the diffusion coefficient for data obtained on a regular grid in \mathbf{R}_+^2 . Let $\mathbf{z} = (s, t)$ be a point in a set S of locations included in the regular grid and let us denote the 2D four-point increment of Y given by

$$Y(\Delta_{hk}(\mathbf{z})) = Y(s + h, t + k) - Y(s, t + k) - Y(s + h, t) + Y(s, t),$$

for $h, k > 0$. Taking into account that the variance of this increment,

$$var(Y(\Delta_{hk}(\mathbf{z}))) = \sigma_{\mathbf{z},h,k}^2 = \int_s^{s+h} \int_t^{t+k} \tilde{B}(\sigma, \tau) d\sigma d\tau = \tilde{B}hk,$$

does not depend on the location \mathbf{z} , but only on the area hk , the diffusion coefficient \tilde{B} can be estimated using a similar approach to Matheron’s estimator for the variogram (see, for example, Cressie 1993), considering here 2D four-point increments, as follows.

Under the implicit condition that $\mathbf{z}_i = (s_i, t_i) < \mathbf{z}_j = (s_j, t_j)$, we denote

$$[\mathbf{z}_i, \mathbf{z}_j] = \{(s_i, t_i), (s_i, t_j), (s_j, s_i), (s_j, t_j)\}.$$

The estimator, for $\mathbf{z} = (s, t)$, is

$$\begin{aligned} \widehat{var}(Y(\Delta_{hk}(\mathbf{z}))) &= \frac{1}{|N(hk)|} \sum_{N(hk)} (Y(s+h, t+k) - Y(s, t+k) - Y(s+h, t) + Y(s, t) \\ &\quad - m_Y(s+h, t+k) + m_Y(s, t+k) + m_Y(s+h, t) - m_Y(s, t))^2, \end{aligned}$$

where

$$N(hk) \equiv \{(\mathbf{z}_i, \mathbf{z}_j) : [\mathbf{z}_i, \mathbf{z}_j] \in S, (s_j - s_i)(t_j - t_i) = hk, i, j = 1, \dots, n\}$$

and $|N(hk)|$ is the number of different elements in $N(hk)$. If the mean is unknown, it can be estimated using (1) by $m_Y^*(\mathbf{z}) = \sum_{\alpha=0}^p \phi_\alpha^* f_\alpha(\mathbf{z})$. Finally, an estimate for \tilde{B} , denoted \tilde{B}^{**} , can be obtained from the slope of the ‘best’ straight line starting at the origin and approximating the set of points $(hk, \widehat{var}(Y(\Delta_{hk}(\mathbf{z}))))$ calculated (see Fig. 2). In this work the usual least-squares criterion will be used for optimality.

Fig. 1 Contour-level plot of 49 values generated (simulation 1) for the lognormal diffusion random field (known non constant mean case)

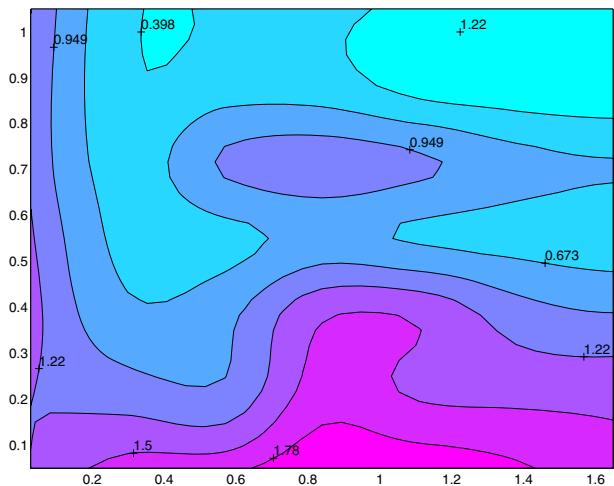


Table 1 Estimates of \bar{B} by the two methods considered (* denotes de MLE, and ** denotes the fitting slope estimator), for 16 simulations of the lognormal diffusion random field (known non constant mean case)

Sim. no.	\bar{B}^*	\bar{B}^{**}	Sim. no.	\bar{B}^*	\bar{B}^{**}
1	1.1115	0.8199	9	0.9911	0.8236
2	1.0605	0.5584	10	0.9909	0.4597
3	1.2060	1.0004	11	1.0107	0.4792
4	1.2153	0.5457	12	0.9016	0.3990
5	1.1324	0.8595	13	0.8914	1.5703
6	0.8309	0.6103	14	1.1850	0.9163
7	0.6138	0.3944	15	0.9870	1.2419
8	1.3243	0.4456	16	1.0684	1.1750

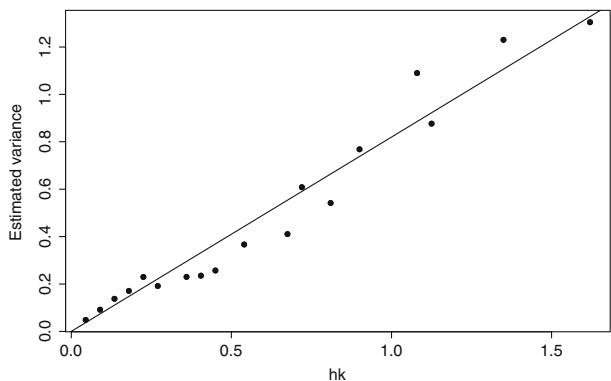
4 Numerical Examples

We describe some numerical examples illustrating prediction and conditional simulation for a lognormal diffusion random field under the approaches considered. First, using simulated data on a regular grid, the two estimation methods for the diffusion coefficient respectively described in Sections 3.1 and 3.2 are compared, considering the case of known non constant mean (for the associated Gaussian random field). Second, we obtain a prediction and a conditional simulation for a lognormal diffusion random field (see Fig. 1).

The parameter space considered is $I = [0, 1.65] \times [0, 1.05]$. Parameter estimates are obtained based on data \mathbf{X} , consisting of the values corresponding to the 7×7 regular grid, G , with SW corner at the origin $(0, 0)$ and NE corner at point $(1.65, 1.05)$. Kriging predictions and simulations are obtained on a regular 19×19 grid, S , set determined by the same corner points. Note that, for simplicity, in this example $G \subset S$. Unconditionally simulated realizations are obtained by the method of unconstrained simulation described in Christakos (1992).

We consider a lognormal diffusion random field with non constant mean, with $\phi_0 = 0.25$, $\tilde{a}(\mathbf{z}) = -2$, for all $\mathbf{z} \in I$, and $\bar{B} = 1$. Assuming that the mean of the associated Gaussian random field is known, Table 1 gives the estimates of the diffusion

Fig. 2 Straight line fitted by least squares using data obtained in simulation 1. The slope is 0.8199832



coefficient, \tilde{B}^* and \tilde{B}^{**} , obtained for 16 independent unconstrained simulations for this random field. Figure 2 shows, as a example, the regression line from which \tilde{B}^{**} was obtained in simulation 1.

In this case and in several other cases studied by the authors, the maximum likelihood estimation method overall provides more accurate estimates for the diffusion coefficient than the alternative method based on evaluation of 2D four-point increments. Lack of stability in the estimate \tilde{B}^{**} can be possibly overcome by robust estimation of the slope of $\widehat{var}(Y(\Delta_{hk}(\mathbf{z})))$ vs. hk instead of using the least-squares approach.

For the example of prediction and conditional simulation we consider the previous diffusion and data obtained in simulation 1. The mean of the associated Gaussian random field is known, with $\phi_0 = 0.25$, $\tilde{a}(\mathbf{z}) = -2$, for all $\mathbf{z} \in I$, and following the above remarks, in this work we will use the maximum likelihood method for the estimation of the diffusion coefficient, that is, $\tilde{B}^* = 1.1115$. Thus, for $\mathbf{z} = (s, t) \in I$,

$$m_Y(\mathbf{z}) = 0.25 - 2st,$$

$$\hat{\sigma}_Y^2(\mathbf{z}) = 1.1115st.$$

Let $\mathbf{z} \in S$ and $\{\mathbf{z}_1, \dots, \mathbf{z}_n\} = G \subset I$. We denote

$$\hat{\mathbf{c}}_Y = (\hat{\sigma}_Y^2(\mathbf{z} \wedge \mathbf{z}_1), \dots, \hat{\sigma}_Y^2(\mathbf{z} \wedge \mathbf{z}_n))^t,$$

$$\hat{\Sigma}_Y = \tilde{B}^* \mathbf{M}. \tag{3}$$

Using Eq. 3 we have calculated 19×19 predictions by simple lognormal kriging. An unbiased predictor of $X(\mathbf{z})$ is given by

$$\hat{X}(\mathbf{z}) = \exp \left\{ \hat{\mathbf{c}}_Y^t \hat{\Sigma}_Y^{-1} (\mathbf{Y} - \mathbf{m}_Y) + m_Y(\mathbf{z}) + \frac{\hat{\sigma}_Y^2(\mathbf{z}) - \hat{\mathbf{c}}_Y^t \hat{\Sigma}_Y^{-1} \hat{\mathbf{c}}_Y}{2} \right\}$$

$$= \exp \left\{ \hat{y}(\mathbf{z}_i) + \frac{\hat{\sigma}_{Ysk}^2(\mathbf{z})}{2} \right\},$$

where $\hat{y}(\mathbf{z}) = \hat{\mathbf{c}}_Y \hat{\Sigma}_Y^{-1} (\mathbf{Y} - \mathbf{m}_Y) + m_Y(\mathbf{z})$. The results are plotted in Fig. 3.

Fig. 3 Contour-level plot of the 361 predictions obtained by simple lognormal kriging using the 49 values plotted in Fig. 1

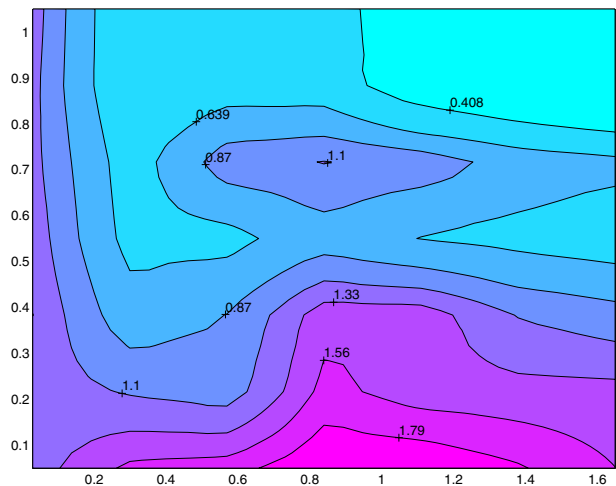
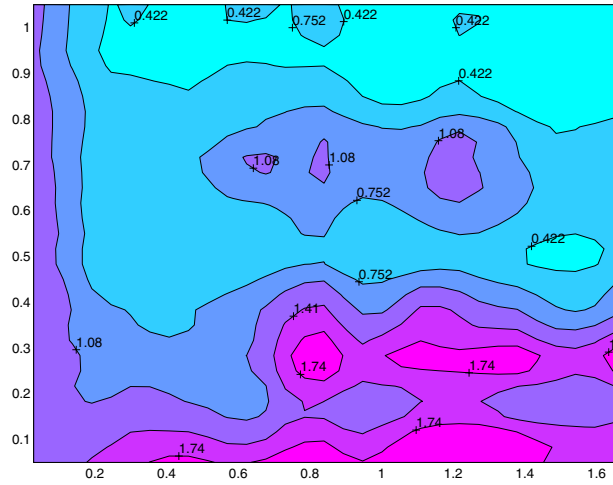


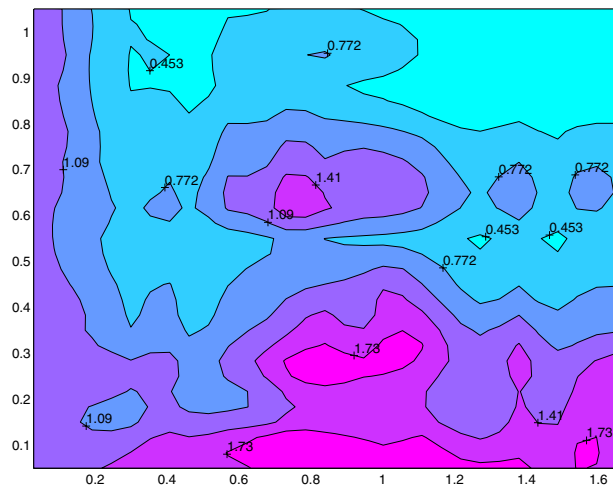
Fig. 4 Contour-level plot of the 361 simulations obtained by conditional simulation using the 49 values plotted in Fig. 1



As for simulation, we have considered a practical method for generating conditional simulations that combines unconditional simulation and kriging, described in Yuh-Ming and Hugh Ellis (1997). This technique yields an unbiased conditional simulation (with respect to sample data) and reproduces conditional variances. Basically, this method provides conditional simulations correcting unconditional simulations through the comparison of the kriging prediction based on the raw data and the kriging prediction based on the unconditional simulation data on G . Let us consider the log-transformed n -dimensional random vector, $\mathbf{Y} = (\ln X(\mathbf{z}_1), \ln X(\mathbf{z}_2), \dots, \ln X(\mathbf{z}_n))^t = \ln \mathbf{X}$. We can summarize the procedure as follows:

- Step 1 Estimate the mean and the variance of the associated Gaussian random field, if they are unknown, using the estimates given in Eqs. 1 and 2 with raw data $\mathbf{Y} = \ln \mathbf{X}$.

Fig. 5 Contour-level plot of the 361 values (including the 49 values used for estimating \bar{B}) generated (simulation 1) for the 2D lognormal diffusion considered



- Step 2 Predict $\{\widehat{y}(\mathbf{z}_i) : \mathbf{z}_i \in S\}$ based on the raw data \mathbf{Y} and using the predictor of simple kriging (which uses the estimates calculated in step 1).
- Step 3 Calculate unconditionally simulated realizations $\{y^u(\mathbf{z}_i) : \mathbf{z}_i \in S\}$ based on the method of unconstrained simulation and using the estimates calculated in step 1.
- Step 4 Calculate the set of predictions $\{\widehat{y}^u(\mathbf{z}_i) : \mathbf{z}_i \in S\}$ based on the data $\{y^u(\mathbf{z}_i) : \mathbf{z}_i \in G\}$ and using the predictor of simple kriging (which is calculated using the estimates obtained in step 1).
- Step 5 Calculate conditional simulation realizations of Y by

$$y^c(\mathbf{z}_i) = y^u(\mathbf{z}_i) + [\widehat{y}(\mathbf{z}_i) - \widehat{y}^u(\mathbf{z}_i)], \quad \forall \mathbf{z}_i \in S.$$

- Step 6 Calculate conditional simulation realizations of X by

$$x^c(\mathbf{z}_i) = \frac{\exp\{y^u(\mathbf{z}_i)\} \exp\{\widehat{y}(\mathbf{z}_i)\}}{\exp\{\widehat{y}^u(\mathbf{z}_i)\}} \equiv \frac{x^u(\mathbf{z}_i) \widehat{x}(\mathbf{z}_i)}{\widehat{x}^u(\mathbf{z}_i)}, \quad \forall \mathbf{z}_i \in S.$$

For the example considered (here the mean of the associated Gaussian random field is known), using Eq. 3, if $\mathbf{z}_i \in S$ we obtain in step 2

$$\widehat{y}(\mathbf{z}_i) = \widehat{c}_Y \widehat{\Sigma}_Y^{-1} (\mathbf{Y} - \mathbf{m}_Y) + m_Y(\mathbf{z}_i),$$

where \mathbf{Y} is the log-transformed raw data. Similarly in step 4, if $\mathbf{z}_i \in S$,

$$\widehat{y}^u(\mathbf{z}_i) = \widehat{c}_Y \widehat{\Sigma}_Y^{-1} (\mathbf{Y}^u - \mathbf{m}_Y) + m_Y(\mathbf{z}_i),$$

where $\mathbf{Y}^u = (y^u(z_1), \dots, y^u(z_n))$ is the vector of unconditional simulations obtained in step 3 but only using the data simulated on G . Note that the remaining unconditional simulated data are used in step 5.

Figure 4 displays a contour-level plot for the 19×19 conditional simulation realization based on the data of simulation 1, and Fig. 5 displays the original contour-level plot.

5 Conclusions

We have studied prediction and conditional simulation for a 2D lognormal diffusion random field, including exogenous factors in its formulation. These random fields are not intrinsically stationary, and then well-known related techniques cannot be applied. The methods considered here allows us to use well-known techniques for estimation and prediction, such as simple kriging, and for conditional simulation.

For real applications, we are considering possible extensions like the study of non-constant diffusion-type values at the boundary axes as well as higher-dimension spatial, spatio-temporal formulations and the development of validation techniques.

Acknowledgements We are grateful to the referee for his/her constructive comments which lead us to improving the final version of the paper. This work has been partially supported by projects MTM2005-09209 and MTM2005-08597 of the DGI, Ministerio de Educación y Ciencia, Spain.

References

- G. Christakos, *Random Field Models in Earth Sciences*, Academic: San Diego, 1992.
- N. Cressie, *Statistics for Spatial Data*, Wiley: New York, 1993 (rev. ed.).
- R. Gutiérrez, C. Roldán, R. Gutiérrez-Sánchez, and J. M. Angulo, “Estimation and prediction of a 2D lognormal diffusion random field,” *Stochastic Environmental Research and Risk Assessment* vol. 19(4) pp. 258–265, 2005.
- D. Nualart, “Two-parameter diffusion processes and martingales,” *Stochastic Processes and Their Applications* vol. 15 pp. 31–57, 1983.
- L. Yuh-Ming and J. Hugh Ellis, “Estimation and simulation of lognormal random fields,” *Computers and Geosciences* vol. 23(1) pp. 19–31, 1997.