TASK QUARTERLY 10 No 1, 21–26

CONDITIONAL SIMULATION OF SPATIOTEMPORAL RANDOM FIELDS OF ENVIRONMENTAL CONTAMINATION

ROBERT JANKOWSKI AND HENRYK WALUKIEWICZ

Faculty of Civil and Environmental Engineering,
Gdansk University of Technology,
Narutowicza 11/12, 80-952 Gdansk, Poland
jankowr@pg.gda.pl

(Received 6 December 2005; revised manuscript received 27 December 2005)

Abstract: The paper considers a method of conditional simulation of spatiotemporal scalar random fields of certain environmental phenomena. The method can be used to predict field values at given space points at specified time, on the basis of field values at other locations and data on second order moment functions in the domain. This approach has been applied to a space-time prognosis of soil contamination fields. The assessment of the spatiotemporal variability of heavy metals' concentrations provides the knowledge needed to monitor and control soil contamination. Empirical data of heavy metal (*viz.* chromium) concentration in the soil of northern Poland have been used in the study. The acceptance-rejection method has been applied to generate covariance matrices and vectors of discrete field values, taking into account conditional probability distributions. The results of the study show that the considered method can be successfully used to model conditional, spatiotemporal random fields of contamination with relatively small simulation errors.

Keywords: stochastic modelling, random fields, spatiotemporal covariance function, soil contamination

1. Introduction

In recent years, the theory of random fields has been intensively studied and applied in many areas, including civil and environmental engineering, geotechnics, mechanics, ocean engineering, earth sciences and environmental protection [1–4]. Simulation methods are often used to deal with random processes such as, for example, the propagation of seismic waves [5], fluctuations of wind forces or ocean waves [6], geometrical imperfections in structures [7], or environmental contamination [8]. In particular, methods of modelling random fields of contamination have been shown to be very useful in monitoring contamination levels and predicting unknown contamination values (see *e.g.* [4, 8]).

The present paper is concerned with reconstructing the random field of an environmental (contamination-type) phenomenon in terms of the fragmentary data available. The data include space-time concentrations of some contaminants and

the corresponding second-order functional characteristics of spatiotemporal random fields. The concept of spatiotemporal random fields (see [9]) offers means of adequate representation of natural processes that are irregular, space-nonhomogeneous and time-nonstationary in character. The field reconstruction can be approached in an efficient (practical) or inefficient (*e.g.* too general) manner. The choice of the model and numerical techniques discussed in this paper has been governed by their usefulness in practical stochastic analysis of soil contamination with heavy metals. In the considered approach, the theoretical model and the numerical simulations are considered jointly, which leads to a deeper understanding of the random phenomena of environmental contamination in terms of covariance functions, optimal sampling, *etc.*

2. Theoretical models of random fields of contamination

In the theory of random fields of contamination (see *e.g.* [3, 8]), natural environmental phenomena are simulated by means of spatial random field models. It is usually assumed that $X(\boldsymbol{r})$ represents a scalar, in general a space-nonhomogeneous random field, where $\boldsymbol{r} \in \boldsymbol{R}^2$ denotes a two-dimensional position vector. The so-called second order field is characterised in terms of its mean value function:

$$m(\boldsymbol{r}) = E(X(\boldsymbol{r})), \tag{1}$$

and the spatial covariance function:

$$K(\boldsymbol{r}_1, \boldsymbol{r}_2) = E((X(\boldsymbol{r}_1) - m(\boldsymbol{r}_1)) \cdot (X(\boldsymbol{r}_2) - m(\boldsymbol{r}_2))), \tag{2}$$

where $E(\cdot)$ denotes the expectation operator and $\boldsymbol{r}_1, \boldsymbol{r}_2 \in \boldsymbol{R}^2$.

The concepts of homogeneity, ergodicity and isotropy serve as useful hypotheses. The $X(\boldsymbol{r})$ second-order field is considered to be space-homogeneous if its mean and covariance functions are unaffected by an argument shift. A homogeneous field is ergodic if all the statistical information is included in the single realisation available. An isotropic field is a special instance of a homogeneous random field. In this case, the covariance function depends only on the $|\boldsymbol{r}_2 - \boldsymbol{r}_1|$ length of the distance vector. Importantly, the behaviour of the covariance function of a homogeneous random field in the neighbourhood of $\boldsymbol{r}_2 - \boldsymbol{r}_1 = \boldsymbol{0}$ may be a determining factor with regard to the field's differentiability (in the so-called mean-square sense). For example, a homogeneous 1-D field is differentiable (in the mean-square sense) if, and only if, its covariance function K has its second derivative at $|\boldsymbol{r}_2 - \boldsymbol{r}_1| = 0$.

An example of a differentiable, homogeneous, isotropic, 2-D random field is the so-called Shinozuka field. Its covariance function is described by [10]:

$$K(\boldsymbol{r}_1, \boldsymbol{r}_2) = \sigma^2 \cdot \exp\left(-\alpha\left((r_{2x} - r_{1x})^2 + (r_{2y} - r_{1y})^2\right)\right), \tag{3}$$

where σ is the field's standard deviation, α ($\alpha > 0$) is a scale parameter describing the degree of space correlation, and r_{ix}, r_{iy} ($i = 1,2$) are the components of the distance vector, $\boldsymbol{r}_i$.

If the covariance function of the 2-D field is of the form:

$$K(\boldsymbol{r}_1, \boldsymbol{r}_2) = \sigma^2 \cdot \exp\left(-\alpha(r_{2x} - r_{1x})^2 - \beta(r_{2y} - r_{1y})^2\right), \tag{4}$$

where $\alpha \neq \beta$ ($\alpha > 0, \beta > 0$), then the field is anisotropic but homogeneous and differentiable (in the mean-square sense). An example of a non-differentiable (in the

mean-square sense) but homogeneous and isotropic 2-D field is the so-called white-noise field (see [8, 11]) defined by the covariance function:

$$K(\boldsymbol{r}_1,\boldsymbol{r}_2)=\begin{cases}\sigma^2 & \text{for } |\boldsymbol{r}_2-\boldsymbol{r}_1|=0\\ 0 & \text{for } |\boldsymbol{r}_2-\boldsymbol{r}_1|\neq 0\end{cases} \tag{5}$$

and the zero-mean value function.

3. Spatiotemporal random fields

Let us now denote a scalar, space-nonhomogeneous, time-nonstationary random field by $X(\boldsymbol{r},t)$, where $(\boldsymbol{r},t)\in\boldsymbol{R}^2\times T$ denotes space-time coordinates such that $\|(\boldsymbol{r},t)\|^2=\|\boldsymbol{r}\|^2+t^2$ with $\|\boldsymbol{r}\|^2=r_1^2+r_2^2$. The second-order field, $X(\boldsymbol{r},t)$, $\boldsymbol{r}\in\boldsymbol{R}^2$, $t\in T=[0,\infty)$, is characterised in terms of its spatiotemporal mean value function:

$$m(\boldsymbol{r},t)=E(X(\boldsymbol{r},t)) \tag{6}$$

and the spatiotemporal covariance function:

$$K(\boldsymbol{r}_1,t_1;\boldsymbol{r}_2,t_2)=E((X(\boldsymbol{r}_1,t_1)-m(\boldsymbol{r}_1,t_1))\cdot(X(\boldsymbol{r}_2,t_2)-m(\boldsymbol{r}_2,t_2))). \tag{7}$$

We can interpret $X(\boldsymbol{r},t)$, $\boldsymbol{r}\in\boldsymbol{R}^2$, $t\in T=[0,\infty)$ either as a random field in a three-dimensional space or as a time-dependent random field in a two-dimensional space. The latter interpretation is more convenient.

In the present paper, let us propose the spatiotemporal covariance function in the following form:

$$K(\boldsymbol{r}_1,t_1;\boldsymbol{r}_2,t_2)=\sigma^2\cdot\exp\left(-\alpha\left((r_{2x}-r_{1x})^2+(r_{2y}-r_{1y})^2\right)\right)\cdot\exp(-\beta(t_2-t_1)), \tag{8}$$

where α, β $(\alpha>0,\ \beta>0)$ are the scale parameters, which are essential for describing the degree of space and time correlation of the field, respectively. Notably, the above covariance function can be considered as an extension of the spatial Shinozuka field (see Equation (3)) into the space-time domain.

4. Numerical simulations of spatiotemporal random fields

For simulation purposes, let us consider the discrete random field in the form of 2-D random variables defined at every node of the spatiotemporal grid. Then, the covariance matrix, $\boldsymbol{K}$, assumes the role of the covariance function, K. Let us now denote the number of points in a plane 2-D space domain as M and the number of points in a 1-D time domain as N. Then, the covariance matrix, $\boldsymbol{K}$, of dimension $(M\cdot N)\times(M\cdot N)$ is given in the following block form:

$$\boldsymbol{K}=\begin{bmatrix}\boldsymbol{k}_{11} & \boldsymbol{k}_{12} & \dots & \boldsymbol{k}_{1N}\\ \boldsymbol{k}_{21} & \boldsymbol{k}_{22} & \dots & \boldsymbol{k}_{2N}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{k}_{N1} & \boldsymbol{k}_{N2} & \dots & \boldsymbol{k}_{NN}\end{bmatrix}, \tag{9}$$

where $\boldsymbol{k}_{ij}=E\left((\boldsymbol{X}_i-\boldsymbol{m}_i)\cdot(\boldsymbol{X}_j-\boldsymbol{m}_j)^T\right)$ is a matrix of dimension $M\times M$ and $\boldsymbol{m}_i$ is a vector of mean values at the space points at time i (i, j =1,2,...,N). Let us now divide the random variable vector $\boldsymbol{X}$ (of dimension $M\cdot N\times 1$) with an assumed joint Gaussian

truncated (with high truncation level) probability density, $f(\boldsymbol{X})$, into two blocks consisting of unknown, $\boldsymbol{X}_u\,(n\times 1)$, and known, $\boldsymbol{X}_k\,(p\times 1)$, elements $(M\cdot N=n+p)$:

$$\boldsymbol{X}=\left\{\begin{array}{c}\boldsymbol{X}_u\\ \boldsymbol{X}_k\end{array}\right\}. \tag{10}$$

Due to this division, the covariance matrix $\boldsymbol{K}\,(M\cdot N\times M\cdot N)$ and the mean values vector $\boldsymbol{m}\,(M\cdot N\times 1)$ can be parted into the following blocks:

$$\boldsymbol{K}=\begin{bmatrix}\boldsymbol{K}_{11} & \boldsymbol{K}_{12}\\ \boldsymbol{K}_{21} & \boldsymbol{K}_{22}\end{bmatrix},\quad \boldsymbol{m}=\left\{\begin{array}{c}\boldsymbol{m}_u\\ \boldsymbol{m}_k\end{array}\right\}. \tag{11}$$

Then, the conditional covariance matrix $\boldsymbol{K}_c\,(n\times n)$ and the conditional mean value vector $\boldsymbol{m}_c\,(n\times 1)$ are calculated from the conditional distribution [8]:

$$f(\boldsymbol{X}_u|\,\boldsymbol{X}_k)=(\det\boldsymbol{K}_c)^{-\frac{1}{2}}\cdot(2\pi)^{-\frac{n}{2}}\cdot\exp\left(-\frac{1}{2}(\boldsymbol{X}_u-\boldsymbol{m}_c)^T\boldsymbol{K}_c^{-1}(\boldsymbol{X}_u-\boldsymbol{m}_c)\right) \tag{12}$$

and are equal to:

$$\boldsymbol{K}_c=\boldsymbol{K}_{11}-\boldsymbol{K}_{12}\boldsymbol{K}_{22}^{-1}\boldsymbol{K}_{21}, \tag{13}$$

$$\boldsymbol{m}_c=\boldsymbol{m}_u+\boldsymbol{K}_{12}\boldsymbol{K}_{22}^{-1}(\boldsymbol{X}_k-\boldsymbol{m}_k). \tag{14}$$

It can be proved that vector $\boldsymbol{m}_c$ is the best approximation (in the mean-square sense) of vector $\boldsymbol{X}_u$ of the unknown discrete field values.

In order to check the error of this prediction, let us generate realisations of discrete spatiotemporal random fields in the acceptance-rejection approach (see [5, 8, 11] for details). The known statistical formulae yield estimators of the mean value $\overline{\boldsymbol{m}}$ and the global covariance matrix $\overline{\boldsymbol{K}}$ of the generated set of realisations, $\boldsymbol{X}_i$:

$$\overline{\boldsymbol{m}}=\frac{1}{NR}\sum_{i=1}^{NR}\boldsymbol{X}_i, \tag{15}$$

$$\overline{\boldsymbol{K}}=\frac{1}{NR-1}\sum_{i=1}^{NR}(\boldsymbol{X}_i-\overline{\boldsymbol{m}})(\boldsymbol{X}_i-\overline{\boldsymbol{m}})^T, \tag{16}$$

where NR is the number of realisations in a set. Two types of errors can be used. Global error, GE, is a useful measure of quality of the generation method and can be expressed as:

$$GE=\left|\frac{\|\boldsymbol{K}\|-\|\overline{\boldsymbol{K}}\|}{\|\boldsymbol{K}\|}\right|\cdot 100\%, \tag{17}$$

where $\|\boldsymbol{K}\|$ and $\|\overline{\boldsymbol{K}}\|$ are the Euclidean norms for the assumed (theoretical) and the generated covariance matrix, respectively, and can be calculated as:

$$\|\boldsymbol{K}\|=\sqrt{\mathrm{tr}(\boldsymbol{K}^2)}=\sqrt{\sum_{i,j=1}^{M\cdot N}K_{ij}}, \tag{18}$$

Local error, LE, is a measure of quality of the spatiotemporal prediction by the conditional mean value, m_{ci}:

$$LE_i=\left|\frac{X_i-m_{ci}}{X_i}\right|\cdot 100\%, \tag{19}$$

where $i=1,2,\ldots,M$ for a fixed time $t\in[1,2,\ldots,N]$.

5. Example of modelling of spatiotemporal contamination field

In this study, an exemplary spatiotemporal random field is considered described by a heavy metal (*viz.* chromium) concentration in the soil of northern Poland. Five locations have been chosen for analysis. The data concerning chromium concentration in the soil at these points are presented in Table 1. The mean values for all points are used as the parameters to formulae (12)–(14) at time t_1. The objective of the analysis is to predict the soil contamination value at location no. 1 at time t_2, if, for example, $t_2 - t_1 = 1$ year.

Table 1. Data concerning chromium concentration in the soil

No.	Place	Position		Chromium concentration (ppm)			
		r_x (km)	r_y (km)	Lower bound, a	Upper bound, b	Mean value, m	Standard deviation, σ
1	Kościerzyna	10.00	10.00	9.6	26.2	17.9	1.66
2	Strzebielino	14.10	61.00	12.8	26.8	19.8	1.40
3	Rębiechowo	40.85	38.30	11.1	20.7	15.9	0.96
4	Wejherowo	26.45	63.10	13.7	30.7	22.2	1.70
5	Gdynia-Obłuże	43.60	57.60	9.3	26.1	17.7	1.68
			Mean values:	11.3	26.1	18.7	1.48

The spatiotemporal covariance function defined by Equation (8) has been used in the study. For the numerical simulation purposes, the acceptance-rejection method has been employed (see [12]).

First, the values of chromium concentration at five locations at time t_1 were generated with the Gaussian distribution. Then, by treating them as known values (X_{ki}, $i = 1,2,\ldots,5$), the chromium concentration value at location no. 1 at time t_2 was generated (X_u). Subsequently, the conditional mean value, m_c, was determined from Equation (14). Finally, *GE* and *LE* errors were calculated using Equations (17) and (19), respectively.

Examples of the generated field values and the calculated local errors, *LE*, for the scale parameters of $\alpha = 0.004$ and $\beta = 0.693$, obtained with the help of the least squares error method, are shown in Table 2. The errors of the generated covariance matrices are presented in Table 3 as global errors, *GE*, and global errors of variances (calculated similarly to *GE*).

6. Concluding remarks

In modelling contamination fields, the general theory of spatiotemporal random fields can be restricted to specific classes of moment functions. In the present paper, a practical example of soil contamination prediction in northern Poland has been solved with the hypothesis of a space-time separable correlation structure. This assumption, however, does not restrict the generality of the method. The results of the study show that the considered method allow us to model conditional spatiotemporal random fields of contamination with relatively small simulation errors.

Table 2. Examples of generated field values

X_{k1}	18.1672	17.0747	19.4393	17.3650	19.4787	20.9898	17.8201	19.2872
X_{k2}	18.4993	20.3615	17.1380	15.6612	18.4270	19.6545	18.8928	17.9199
X_{k3}	16.7030	19.5609	18.0329	18.8318	18.9942	15.5204	18.7518	19.2100
X_{k4}	16.8197	18.2984	20.5802	21.7889	21.0538	17.1773	17.7755	17.6295
X_{k5}	20.1572	19.3130	20.0301	19.2404	18.7476	18.4905	17.9011	20.5769
X_u	18.9874	19.0985	18.3704	18.7669	18.6109	18.8965	18.2252	19.7605
m_c	19.4287	19.0077	19.3637	18.9675	18.7233	18.5961	18.3008	19.6381
LE (%)	2.32	0.48	5.41	1.07	0.60	1.59	0.41	0.62
							Mean error LE (%):	1.56

Table 3. Global errors with respect to the number of realisations (NR)

	$NR = 500$	$NR = 2000$	$NR = 10\,000$
Global error (%)	5.48	5.05	5.70
Global error of variances (%)	2.01	1.41	0.37

In the standard methods of random field simulation of various environmental phenomena, stochastic estimation and simulation are considered as two separate problems. In the presented modelling approach, simulation and estimation are treated jointly, with computational advantages such as unified conditional methodology, evaluation of errors and simple algorithms.

References

[1] Vanmarcke E H 1983 *Random Fields: Analysis and Synthesis*, MIT Press, Cambridge, USA
[2] Shinozuka M 1987 *Stochastic Methods of Structural Dynamics*, Martinus Nijhoff Publ., Boston, USA, pp. 93–133
[3] Christakos G 1992 *Random Field Models in Earth Sciences*, Academic Press Inc., San Diego, USA
[4] De Wispelaere C 1981 *Air Pollution Modelling and its Application* Vol. 1, Plenum Press, New York, USA
[5] Jankowski R and Wilde K 2000 *Engng Structures* **22** 552
[6] Borgman L E 1990 *The Sea, Ocean Engineering Science* Vol. 9, John Wiley & Sons, New York, USA, pp. 121–168
[7] Bielewicz E, Górski J, Schmidt R and Walukiewicz H 1994 *Computers and Structures* **51** 267
[8] Jankowski R and Walukiewicz H 1997 *Prob. Engng Mech.* **12** 115
[9] Szczepankiewicz E 1985 *Applications of Random Fields*, PWN, Warsaw, Poland (in Polish)
[10] Yamazaki F and Shinozuka M 1990 *J. Eng. Mech.* **116** 268
[11] Walukiewicz H, Bielewicz E and Górski J 1997 *Computers and Structures* **64** 491
[12] Flury B D 1990 *SIAM Review* **32** 474