MODELLING OF THRESHOLD-CROSSING PROBLEM FOR RANDOM FIELDS OF CONTAMINATION

Robert Jankowski[1], Henryk Walukiewicz[1]

[1]Centre of Excellence for Urban Construction and Rehabilitation CURE
Faculty of Civil Engineering, Gdańsk University of Technology
ul. Narutowicza 11/12, 80-952 Gdańsk, POLAND
phone/fax: (4858) 3472200
e-mail: jankowr@pg.gda.pl

Abstract
The present paper deals with assessing the probability, that a random field of contamination does not exceed a fixed level in a certain two-dimensional (2-D) spatial domain. A real-valued, homogeneous random field described by the mean value and the covariance (differentiable) function is assumed as the basic theoretical model of the contamination field. In the numerical simulation, a suitable discrete model defined on a regular or irregular grid has been developed and tested by the conditional simulation method. The main concept of this work is the mean number of upcrossings of the field level, calculated for some intervals in one-dimensional subspaces of the 2-D domain. This number depends on the second derivative of the covariance function and has been established for the simulated realisations. The practical example is concern with a case study of heavy metals concentration in soil of the northern part of Poland (Gdańsk region).

Introduction
In recent years, the theory of random fields has been intensively studied and applied to a number of randomly occurring engineering and environmental processes (Ref.(1,2)). Methods of modelling of random fields of contamination have been shown to be very useful for the purposes of monitoring contamination level as well as for predicting unknown contamination values (Ref.(3-5)). This paper is dedicated to the development of a new approach to analyse and control 2-D fields of environmental contamination. Theoretical model and numerical simulations are treated jointly, which leads to a deeper understanding of the random phenomena in terms of covariance functions, optimal sampling and probabilities of threshold crossings. Much of the material is based on the present authors research (Ref.(5-8)) in related domains.

Theoretical model of random fields of contamination
The results of the study presented in this paper deal with the spatial behaviour of some natural environmental phenomena, e.g. contamination of soil. The approach is based on the interpretation of the natural process by means of the spatial random field models (Ref.(1-3)). It is assumed that $X(\mathbf{r})$ represents a scalar, in general space-nonhomogeneous random field, where $\mathbf{r} \in \mathbf{R}^2$ denotes a two-dimensional position vector. The so-called second order field is characterised in terms of its mean value function:

$$m_x(\mathbf{r}) = E(X(\mathbf{r})) \tag{1}$$

and spatial covariance function:

$$K_x(\mathbf{r}_1,\mathbf{r}_2) = E\left(\left(X(\mathbf{r}_1) - m_x(\mathbf{r}_1)\right)\cdot\left(X(\mathbf{r}_2) - m_x(\mathbf{r}_2)\right)\right), \tag{2}$$

where $E(\cdot)$ denotes the expectation operator and $\mathbf{r}_1,\mathbf{r}_2 \in \mathbf{R}^2$. The following three concepts: homogeneity, ergodicity and isotropy serve as some useful hypotheses. The second-order field $X(\mathbf{r})$ is called space-homogeneous if its mean and covariance functions do not change under a shift of the vector arguments:

$$m_x(\mathbf{r}) = \text{const}, \tag{3}$$

$$K_x(\mathbf{r}_1,\mathbf{r}_2) = K_x(\boldsymbol{\rho}), \tag{4}$$

where $\boldsymbol{\rho} = \mathbf{r}_2 - \mathbf{r}_1$ is the distance vector. The homogeneous field is ergodic if all the statistical information is included in the single realisation available. A special case of a homogeneous random field is an isotropic field. In this case, the covariance function depends only on the length $\|\boldsymbol{\rho}\|$ of the distance vector:

$$K_x(\boldsymbol{\rho}) = K_x(\|\boldsymbol{\rho}\|) \equiv K_x(\rho). \tag{5}$$

A very important for our analysis is the fact that for a homogeneous random field, the behaviour of its covariance function in the neighbourhood of $\boldsymbol{\rho} = \mathbf{0}$ may be a determining factor in regard to differentiability (in the so-called mean square sense) of the field. For example, the homogeneous 1-D field is differentiable (in the mean square sense) if, and only if, its covariance function K_x has the second derivative at $\rho = 0$. Moreover, the covariance function of the first derivative of the process equals:

$$K_{x'}(\rho) = -\frac{d^2}{d\rho^2} K_x(\rho) \equiv -K_x''(\rho). \tag{6}$$

Examples

The following covariance function describes the differentiable, homogeneous, isotropic, 2-D random field (named by the authors as Shinozuka field- see Ref. (5)):

$$K_x(\mathbf{r}_2 - \mathbf{r}_1) = \sigma_x^2 \cdot \exp\left(-\alpha\left((r_{2x} - r_{1x})^2 + (r_{2y} - r_{1y})^2\right)\right) \equiv \sigma_x^2 \exp\left(-\alpha\left(\rho_x^2 + \rho_y^2\right)\right), \tag{7}$$

where σ_x is a standard deviation of the field, α is a scale parameter describing the degree of space correlation, $\alpha > 0$, and the indices *x*, *y* on the right hand side denote the orthogonal axes. The process along the line $x = 0$ has the covariance function:

$$K_x(\rho_y) \equiv K_x(\rho) = \sigma_x^2 \cdot \exp(-\alpha\rho^2). \tag{8}$$

According to eq.(6) one obtains:

$$K_{x'}(\rho) = -\frac{d}{d\rho}\left(-2\alpha\rho\sigma_x^2 \exp(-\alpha\rho^2)\right) = -\left(-2\alpha\sigma_x^2 \exp(-\alpha\rho^2) + 4\alpha^2\rho^2\sigma_x^2 \exp(-\alpha\rho^2)\right) \tag{9}$$

and finally:

$$K_{x'}(0) = 2\alpha\sigma_x^2. \tag{10}$$

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

$$K_x(\boldsymbol{\rho}) = \sigma_x^2 \cdot \exp\left(-\alpha\rho_x^2 - \beta\rho_y^2\right), \tag{11}$$

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 the non-differentiable (in the mean square sense) although homogeneous and isotropic 2-D field is the so-called white-noise field defined by the covariance function:

$$K_x(\rho) = \begin{cases} \sigma_x^2 & \text{for } \rho = 0 \\ 0 & \text{for } \rho \neq 0 \end{cases} \tag{12}$$

and the zero-mean value function.

Upcrossing problem in 2-D contamination fields with continuous arguments

First principles allow to derive an upper bound on the probability P_u of upcrossing some deterministic level $u(r)$ in 1-D homogeneous random field $X(r)$, where: $r \in \mathbf{R}$. Let $N_u(S)$ denotes the number of upcrossings in the space interval $[0, S]$. The probability is expressed as:

$$\begin{aligned} &P_u\left(X(r) \geq u(r) \text{ for some } r \in [0,S]\right) \;=\; P_u\left(\text{upcrossing at } r = 0 \text{ or } N_u(S) \geq 1\right) \;= \\ &P_u\left(\text{upcrossing at } r = 0\right) \;+\; P_u\left(N_u(S) \geq 1\right) \;-\; P_u\left(\text{upcrossing at } r = 0 \text{ and } N_u(S) \geq 1\right). \end{aligned} \tag{13}$$

It should be noticed that the last negative term of eq.(13) is smaller than the smallest positive one. Therefore, from the theorem of the total event, the upper bound on P_u is found:

$$P_u\left(X(r)\geq u(r)\text{ for some } r\in[0,S]\right) \quad\leq\quad P_u\left(\mathrm{X}(0)\geq \mathrm{u}(0)\right)+P_u\left(N_u(S)\geq 1\right). \tag{14}$$

This upper bound is further developed as:

$$\begin{aligned} P_u\left(\mathrm{X}(0)\geq \mathrm{u}(0)\right)+P_u\left(N_u(S)\right) \quad&\leq\quad P_u\left(\mathrm{X}(0)\geq \mathrm{u}(0)\right)+\sum_{n=1}^{\infty}P_u\left(N_u(S)=n\right) \quad\leq \\ P_u\left(\mathrm{X}(0)\geq \mathrm{u}(0)\right)+\sum_{n=1}^{\infty}nP_u\left(N_u(S)=n\right) \quad&=\quad P_u\left(\mathrm{X}(0)\geq \mathrm{u}(0)\right)+E\left(N_u(S)\right). \end{aligned} \tag{15}$$

The last approximation is proper if:

$$P_u\left(N_u(S)=1\right) \quad\gg\quad \sum_{n=2}^{\infty}nP_u\left(N_u(S)=n\right). \tag{16}$$

The above inequality is valid in practical situations with the high level u, when clustering of crossings can be neglected. The classical theory of Rice (Ref.(1)) gives the mean value of $N_u(S)$ in terms of the covariance function K_x of the underlying process. If $X(r)$ is a zero-mean, homogeneous, differentiable (in the mean square sense) Gaussian process on $[0,S]$ then:

$$E\left(N_u(S)\right)=\frac{S}{2\pi}\sqrt{\frac{-K_x^{''}(0)}{K_x(0)}}\exp\left(-\frac{u^2}{2K_x(0)}\right). \tag{17}$$

For example, in the case of covariance function (8), from eq.(9-10) one may obtain:

$$E\left(N_u(S)\right)=\frac{S}{2\pi}\sqrt{2\alpha}\exp\left(-\frac{u^2}{2\sigma_x^2}\right). \tag{18}$$

<u>Example</u>

It is useful to evaluate the numerical values of the two terms on the right-hand side of basic formula (15). Let us consider a practical environmental random field of soil contamination by a heavy metal (chromium) in the northern part of Poland – Gdańsk region (see also Ref.(5,8)). The measured chromium concentration values in the soil are as follows: lower bound $a=11.3$ ppm, upper bound $b=26.1$ ppm, mean value $m=18.7$ ppm. The upper bound of the space interval and the scale parameter are equal to: $S=92$ km, $\alpha=0.12$ km^{-2}. The three constant u levels above the mean value: $u=m+2\sigma_x$, $u=m+3\sigma_x$, $u=m+4\sigma_x$ $(\sigma_x=1.48 \text{ ppm})$ for Gaussian 1-D random field with covariance function (8) are taken into consideration. The probability values for different levels of u, obtained using eq.(15) and eq.(18), are presented in Table 1. It can be clearly seen from Table 1 that the second term is the dominating one in the probability upper bound.

Table 1: Theoretical probability of upcrossing for different levels of u

u level	$P_u\left(X(0)\geq u\right)$	$E\left(N_u(S)\right)$	$P_u\left(X(0)\geq u\right)+E\left(N_u(S)\right)$
$m+2\sigma_x$	0.0228	0.9684	0.9912
$m+3\sigma_x$	0.0014	0.0796	0.0810
$m+4\sigma_x$	0.0001	0.0024	0.0025

The next model connected with the upcrossing level in 2-D fields deals with certain non-differentiable covariances. We make use of the so-called Slepian inequality (Ref.(9)). If $X(\mathbf{r})$ and $Y(\mathbf{r})$ are two zero-mean Gaussian fields such that for all $\mathbf{r}_1,\mathbf{r}_2\in C\subset\mathbf{R}^2$, where C is a compact domain, and moreover if:

$$E\left(X(\mathbf{r}_1)X(\mathbf{r}_2)\right)\geq E\left(Y(\mathbf{r}_1)Y(\mathbf{r}_2)\right) \tag{19}$$

and

$$E\left(X^2(\mathbf{r})\right) = E\left(Y^2(\mathbf{r})\right), \tag{20}$$

then for any level *u*:

$$P\left(\sup_C X(\mathbf{r}) \ge u\right) \le P\left(\sup_C Y(\mathbf{r}) \ge u\right). \tag{21}$$

where "sup" denotes a least upper bound. In the special 1-D case described by the covariance of the white-noise field type (see eq.(12)), it follows from inequality (21) that the probability of the upcrossing is greater than in the case of the covariance (8), since the conditions (19-20) are fulfilled. The same conclusion is valid in the 2-D case described by eq.(7) and eq.(12).

Model of conditional simulations of discretised random fields

For numerical simulation of the threshold-crossing problem we have to consider the discrete parameter random field in the form of the multi-dimensional continuous random variables. The variables are defined at every node of the regular or irregular spatial grid. An important question that arises is: at what points in the parameter space should we sample a random field? Starting from the informational-theoretic approach Vanmarcke (Ref.(1)) concludes that the length of the optimal sampling interval Δr may be expected to be proportional to the scale of the fluctuation θ and he proposes:

$$\Delta r = \frac{1}{2}\theta , \tag{22}$$

where:

$$\theta = \frac{2}{\sigma_x^2}\int_0^\infty K_x(\rho)\, d\rho , \tag{23}$$

for a 1-D homogeneous, ergodic in the mean field. In the case of the squared exponential covariance function (8), one obtains:

$$\theta = 2\int_0^\infty \exp\left(-\alpha\rho^2\right) d\rho = \sqrt{\frac{\pi}{\alpha}} . \tag{24}$$

Therefore, the length of the sampling interval should be equal to:

$$\Delta r = \frac{1}{2}\sqrt{\frac{\pi}{\alpha}} . \tag{25}$$

As the first step of presented approach, an exploratory (experimental) data such as: lower bounds, upper bounds, mean values, standard deviations, correlation coefficients are collected from the small number of places. From the assumed theoretical form of the covariance function we select the best function in the mean-square sense.

<u>Example</u>

As the example, let us consider once again the random field of soil contamination by a heavy metal (chromium) in the northern part of Poland (Gdańsk region) with properties specified above. Treating a generated value at a chosen point of the field as the known one, let us generate values of contamination at other locations along one axis of the space interval $[0,S]$. For $\alpha = 0.12\ \text{km}^{-2}$, using eq.(25), the length of the optimal sampling interval is calculated as equal to: $\Delta r \simeq 2.6\ \text{km}$ what means that the distance of $S = 92\ \text{km}$ should be divided into 35 intervals. For the generation purposes, the method of the conditional random fields simulation with the acceptance-rejection algorithm is used (see Ref.(5,8) for details). In the analysis, the simplified cumulative simulation procedure (see Ref.(7)) is chosen. In this procedure, the field value at every next location is generated independently based on all so far generated field values at previous locations. The example of generations of contamination values at all 36 field points is presented in Figure1. In the graph, the lines indicating the mean level $(18.7\ \text{ppm})$ and three different *u* levels above the mean: $u = m + 2\sigma_x = 21.66\ \text{ppm}$, $u = m + 3\sigma_x = 23.14\ \text{ppm}$, $u = m + 4\sigma_x = 24.62\ \text{ppm}$ are also plotted. The probability of upcrossing of different *u* levels calculated based on 100 numerical realisations is shown in Table 2.

Figure 1: Example of generated contamination values

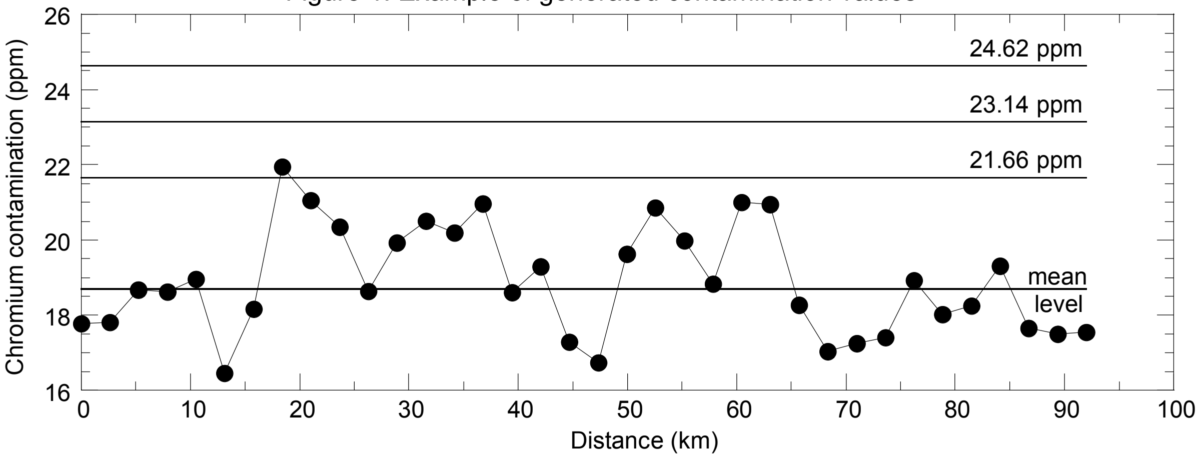

Table 2: Probability of upcrossing for different levels of *u* based on numerical simulations

u level	Probability
$m+2\sigma_x$	0.9444
$m+3\sigma_x$	0.0015
$m+4\sigma_x$	0.0001

Concluding remarks

- Theoretical modelling of the level crossings in 2-D random fields with the continuous parameter (Table 1) has shown a good agreement with the numerical simulations of the fields with the discretised parameter (Table 2).
- A challenging task is modelling of the level crossings in the continuous parameterised field defined on more complex sets than the line intervals.

Acknowledgements

The study was supported by the European Community under the FP5 Programme, key-action "City of Tomorrow and Cultural Heritage" (Contract No. EVK4-CT-2002-80005). This support is greatly acknowledged.

References

(1) E.H. Vanmarcke, Random Fields: Analysis and Synthesis, MIT Press, Cambridge, Massachusetts, USA (1983)
(2) M. Shinozuka, "Stochastic fields and their digital simulation", Stochastic Methods of Structural Dynamics, Martinus Nijhoff Publ., Boston, USA, 93-133 (1987)
(3) G. Christakos, Random Field Models in Earth Sciences, Academic Press Inc., San Diego - Toronto, USA (1992)
(4) J. Namieśnik, W. Chrzanowski and P. Żmijewska (Editors), New horizons and challenges in environmental analysis and monitoring, Centre of Excellence in Environmental Analysis and Monitoring (CEEAM), Gdańsk University of Technology, Gdańsk (2003)
(5) R. Jankowski and H. Walukiewicz, Modeling of two-dimensional random fields, Prob. Engng Mech., **12(2)**, 115-121 (1997)
(6) H. Walukiewicz, E. Bielewicz and J. Górski, Simulation of nonhomogeneous random fields for structural applications, Computers & Structures, **64**, 491-498 (1997)
(7) R. Jankowski and K. Wilde, A simple method of conditional random field simulation of ground motions for long structures, Engng. Structures, **22**, 552-561 (2000)
(8) R. Jankowski and H. Walukiewicz, Modelling of conditional spatiotemporal contamination fields, 5th International Symposium & Exhibition on Environmental Contamination in Central & Eastern Europe, Praque, Czech Repablic, (September 12-14, 2000)
(9) J.D. Slepian, The one-sided barrier problem for Gaussian noise, Bell System Tech. J., **41**, 463-501 (1962)