Archives of Hydro-Engineering and Environmental Mechanics
Vol. 53 (2006), No. 2, pp. 127–136

Application of the Moving Average Method of Simplifying Simulation of Random Fields

Stanisław Maciejewski, Wioletta Gorczewska-Langner

Institute of Hydro-Engineering of the Polish Academy of Sciences, ul. Kościerska 7, 80-328 Gdańsk, Poland, e-mails: macstan@ibwpan.gda.pl, wiwi@ibwpan.gda.pl

(Received September 29, 2005; revised March 23, 2006)

Abstract

A method of simulating random fields using a moving average is proposed in this paper. Random fields were simulated for different sizes of sub-field and different numbers of cycles of calculations of the moving average. For the fields obtained the covariance function was analyzed. In order to estimate the efficiency of the proposed simulation method of random field based on the method of diagonal covariance matrix was performed. It is shown that two of the simulation methods presented are able to generate a multidimensional random variable with required correlation function. However, the method of diagonal covariance matrix has some limitations caused by the size of the simulated random field, which result from the necessity of converting a relatively large matrix. Using the proposed simulation method it is possible to simulate, in a comparatively quick and simple manner, a random field with a large number of nodes on PC-s. The presented method can be useful in the stochastic analysis of transport phenomena in soil.

Key words: stochastic model, multidimensional random variable, moving average, transport phenomena in soil

1. Introduction

The water flow in soil is determined by external conditions and the hydraulic properties of the medium. Since such hydraulic parameters as suction pressure, diffusivity, water capacity, hydraulic conductivity or flux intensity, are heterogeneous and of random character, the soil can be considered as a random field. The water flow and pollutant transport in soil can be described using a stochastic computational model. Let us consider a homogeneous and isotropic random field, where the covariance is a direction-independent function of distance. The solutions of the stochastic model for such a field are non-random time-dependent characteristics of unknown random variables (hydraulic parameters), Maciejewski (1998).

Consequently, in order to take into account the heterogeneity of the soil, it is necessary to simulate the multidimensional random variables first. This paper presents a simple method of doing this.

The group of n random variables $Y_1, Y_2, \ldots, Y_n$ is treated as a vector of random variables $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ in n-dimensional vector space. It is characterized by the expected value $\boldsymbol{\eta}$ and covariance matrix $\mathbf{C}$ (Brandt 1998, Rozanow 1974, Węglarczyk 1999):

$$\boldsymbol{\eta} = (\eta_1, \eta_2, \ldots, \eta_n)\,, \tag{1}$$

$$\mathbf{C} = E\left(\mathbf{Y} \cdot \mathbf{Y}^T\right) - \boldsymbol{\eta} \cdot \boldsymbol{\eta}^T. \tag{2}$$

It is assumed that the covariance function of the local fluctuation f can be described by an exponential expression in the form:

$$C\left(|\mathbf{r}|\right) = E[f\ (\mathbf{x} + \mathbf{r})\ f\ (\mathbf{x})] = \alpha \cdot \exp\left(-\beta \cdot |\mathbf{r}|\right), \tag{3}$$

where:

- $|\mathbf{r}|$ – distance between points $\mathbf{x}$ and $\mathbf{x} + \mathbf{r}$,
- α – variance of local fluctuation,
- β – decay coefficient.

Next, one can introduce the n-dimensional vector of deviations $\mathbf{U} = (U_1, U_2, \ldots, U_n)$ defined as:

$$\mathbf{U} = \mathbf{Y} - \boldsymbol{\eta}, \tag{4}$$

the expected value of which is equal to zero and covariance matrix $\mathbf{C}$ is equal to:

$$\mathbf{C} = E\left(\mathbf{U} \cdot \mathbf{U}^T\right). \tag{5}$$

2. Simulation of Random Fields using the Moving Average Method

The proposed simulation method lies in generating n independent random variables $Y_1, Y_2, \ldots, Y_n$ at nodes of the grid (Fig. 1). The obtained independent variables are then transformed into correlated random variables.

In our case the multidimensional random variable $\mathbf{Y}$ is the hydraulic conductivity $\mathbf{K}\,(Y_1 = K_{11}, Y_2 = K_{12}, \ldots Y_n = K_{lm})$.

The independent random variables are generated using the following formula:

$$K_{ij}^0 = \overline{K} + \Delta K \cdot N\,(0, 1)\,, \tag{6}$$

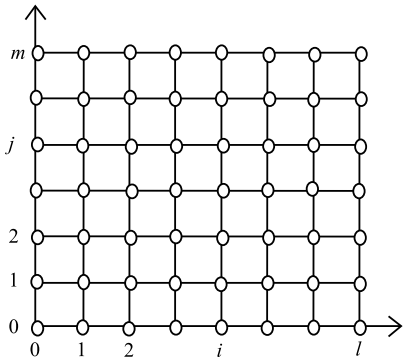

Fig. 1. The nodes of the grid (random field)

where:

$N(0, 1)$ – normalized random variable,
ΔK – assumed standard deviation,
$\overline{K}$ – assumed mean value,
K_{ij}^{0} – independent random variable at the grid nodes.

The generated random variables represent white noise which, as is well known, does not have any correlation. In the second step the random field is smoothed using central moving average $\overline{K_{ij}}$ (Fig. 2):

$$\overline{K_{ij}} = \frac{\sum_{l=i-s}^{i+s} \sum_{m=j-s}^{j+s} K_{lm}}{(2s+1)^2}. \tag{7}$$

The size of the sub-field used to calculate the moving average is characterized by s and can be arbitrary. In the case shown in Fig. 2, the sub-field consists of 9 points ($s = 1$). For $s = 2$ the sub-field consists of 25 points etc. The smooth process can be repeated ξ times ($\xi \geq 1$). ξ is number of smoothing cycle. The number of smoothing cycles, in which the moving average is calculated can also be arbitrary. The influence of the sub-field size and number of smoothing cycles on the results of the simulation is presented in the Section 3 "Results and Discussion".

In the next step, the average of all nodal values is calculated. It represents the average value of the variable for the whole area considered:

$$\overline{K}^{area} = \frac{\sum_{i=0}^{l} \sum_{j=0}^{m} \overline{K_{ij}}}{n}, \tag{8}$$

where n – total number of nodes.

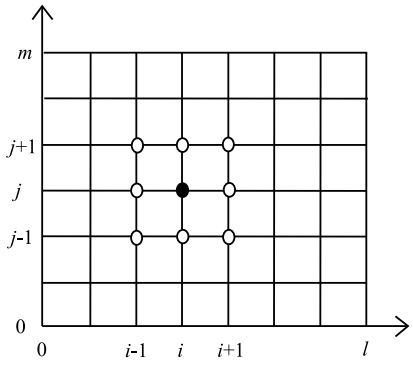

Fig. 2. Random field with marked subfield used to calculate the moving average ($s = 1$)

Also, the new value of the standard deviation for the area ΔK^* is calculated. The resulting value ΔK^* is smaller than the value assumed initially. This is due to the smoothing of random field. The required value of standard deviation is obtained using the following transformation:

$$K_{ij} = \frac{\overline{K_{ij}} - \overline{K}^{area}}{\Delta K^*} \cdot \Delta K + \overline{K}^{area}, \tag{9}$$

where the random variable K_{ij} is already a correlated variable, i.e. there exists a correlation between the neighbouring points of grid.

Now in the literature, it is possible to find several methods of simulating random fields proposed by researchers (Dietrich, Newsam 1993, Mantoglou, Wilson 1982, Przewłócki 1998, Robin et al 1993). For comparison, below we present another method of simulating a multidimensional random variable. It is described in detail in Wilde's work (Wilde 1981). This method was also used by Przewłócki (2003) to determine the ultimate bearing resistance of soil loaded by continuous footing. Here only a brief outline of the procedure is presented.

The simulation of a correlated random variable is performed in the following steps:

- choose n independent random variables $(V_1, V_2, \ldots, V_n) = \mathbf{V}$ with normal distribution and expected value equal to zero,
- diagonalize the covariance matrix, using the transformation of independent random variable $\mathbf{V}$ to random variable $\mathbf{U}$ in the following form

$$\mathbf{U} = \mathbf{p} \cdot \mathbf{V}, \tag{10}$$

$\mathbf{p}$ – a lower triangular matrix with diagonal elements equal to one; the matrix is non-singular as its determinant is equal to one:

$$\mathbf{p} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ p_{21} & 1 & 0 & 0 & 0 & 0 \\ p_{31} & p_{32} & 1 & 0 & 0 & 0 \\ p_{41} & p_{42} & p_{43} & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & 1 & 0 \\ p_{n1} & p_{n2} & p_{n3} & p_{n4} & \cdots & 1 \end{bmatrix}. \tag{11}$$

Random variables **V** are independent, hence are uncorrelated and their covariance matrix **d** diagonal:

$$\mathbf{d} = \begin{bmatrix} d_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & d_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & d_4 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & d_n \end{bmatrix}. \tag{12}$$

Using matrices $\mathbf{p}, \mathbf{V}, \mathbf{d}$ the covariance matrix of the random variable **U** can be written in the following form:

$$\mathbf{C} = E\left(\mathbf{U} \cdot \mathbf{U}^T\right) = E\left(\mathbf{p} \cdot \mathbf{V} \cdot \mathbf{V}^T \cdot \mathbf{p}^T\right) = \mathbf{p} \cdot E\left(\mathbf{V} \cdot \mathbf{V}^T\right) \cdot \mathbf{p}^T = \mathbf{p} \cdot \mathbf{d} \cdot \mathbf{p}^T. \tag{13}$$

From this equation one can determine the all elements of matrix **d** and **p**:

$$d_1 = C_{11}, \tag{14}$$

$$d_i = C_{ii} - \sum_{l=1}^{i-1} p_{il}^2 d_l \quad \text{for} \quad i = 2, \ldots, n, \tag{15}$$

$$d_1 p_{j1} = C_{1j} \quad \text{for} \quad j = 2, \ldots, n, \tag{16}$$

$$p_{ji} = \frac{C_{ij}}{d_i} - \frac{\sum_{l=1}^{i-1} p_{il} p_{jl} d_l}{d_i} \qquad \text{for} \qquad j > i \qquad i, j = 2, \ldots n. \tag{17}$$

Using the elements of matrix **p** it is possible to determine the correlated random variable **U** from Eq. (10). Elements of matrix **U** are equal to:

$$U_j = \sum p_{ji} \cdot V_i + V_j. \tag{18}$$

3. Results and Discussion

The random fields were simulated with the method given by Eqs. (6)–(9) for various sizes of the sub-fields and various numbers of the cycles of calculations of the moving average. In all presented examples the field consisted of 10251 nodes. Next, the random field was investigated through the analysis of covariance function. As shown in Fig. 3, the covariance function of the simulated multidimensional random variable can be approximated with the formula Eq. (3). As the number of points in the sub-field increases this function has a smoother course, thus, the decay coefficient β assumes smaller values.

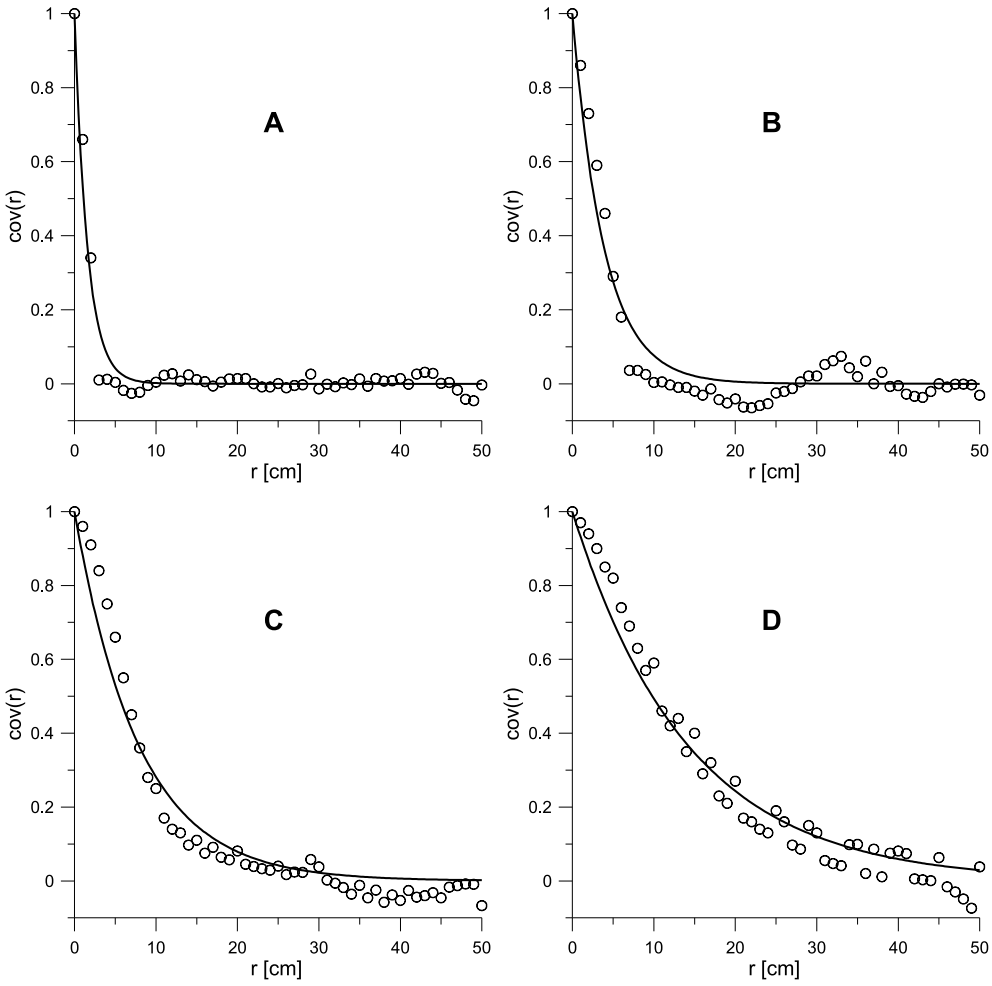

Fig. 3. Relationships between covariance and distance r for various sizes of the sub-field: A – sub-field consisting of 9 points, B – 49 points, C – 81 points, D – 121 points. Sign: ○ – points obtained from simulation, solid line – fitted curve Eq. (3)

The relation between the decay coefficient β and the number of cycles of calculations of the moving average is presented in Fig. 4. One can observe that increasing the number of cycles of calculations influences significantly the shape of the covariance function. It can be seen that for the same number of points in the sub-field, there is an increase in the number of cycles of calculations determining shape of the covariance function.

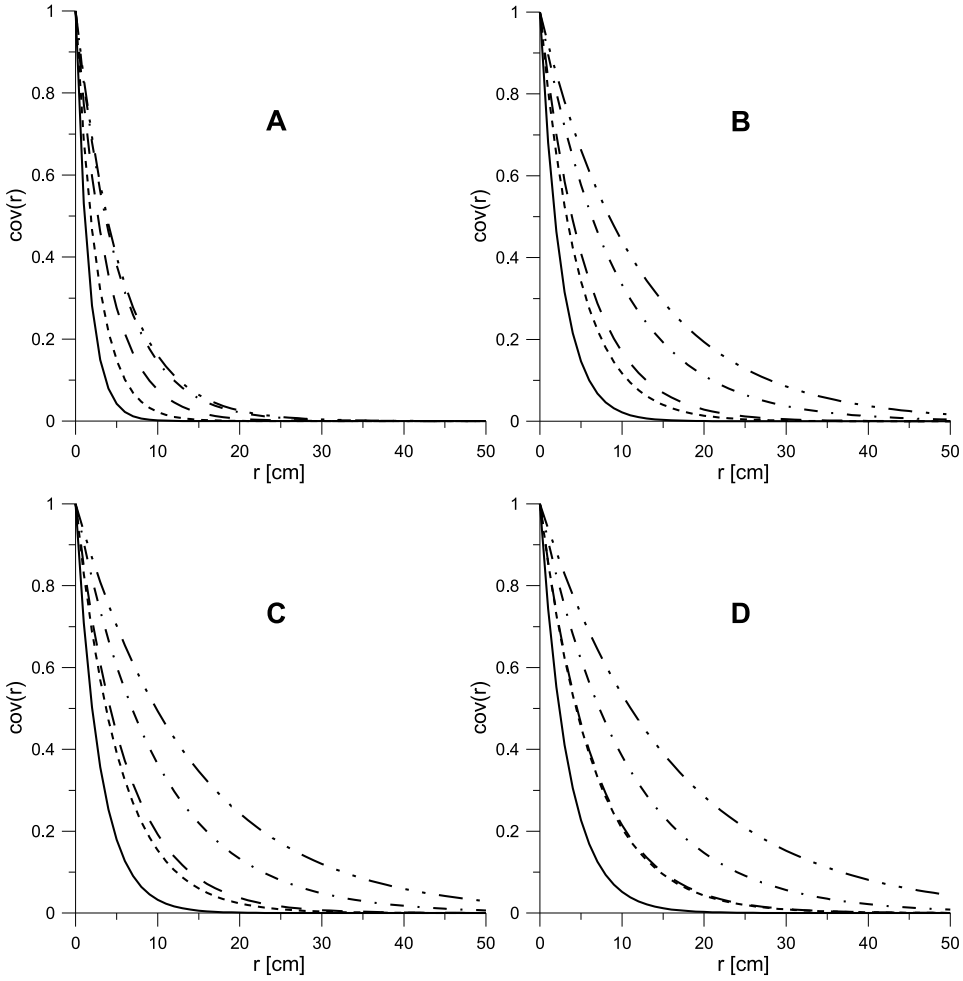

Fig. 4. Relationships between covariance and distance r for various numbers of cycles of calculations of the moving average and various sizes of the sub-field: A – 1 cycle of smoothing, B – 3 cycles, C – 4 cycles, D – 6 cycles,
—— sub-field consisting of 9 points, - - - - - 25 points, – – – – 49 points,
– · – –· – 81 points, – ·· – ·· – 121 points

The relation between the decay coefficient and the size of the sub-field is shown in Fig. 5. The results are shown for various numbers of the calculation cycles. The value of the decay coefficient β decreases as the number of the calculation cycles increases. Using this plot one can determine the simulation parameters such as the size of sub-field and the number of cycles in order to obtain a multidimensional random variable with required covariance function. Note that the initial estimation of the decay β can be made by choosing appropriate sub-field size and number of cycles of calculations (Tab. 1).

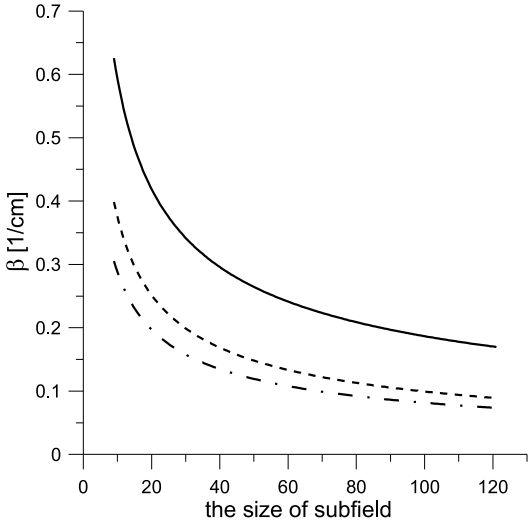

Fig. 5. Relationships between the decay coefficient β and size of sub-field for various numbers of cycles of calculations of the moving average:
——— 1 cycle, – – – – 3 cycles, – · – · – · – 6 cycles

Table 1. The values of the decay coefficient β [1/cm] as a function of the number of cycles of calculations and the size of the sub-field

The number of cycles	The size of subfield				
	9	25	49	81	121
1	0.6337	0.3815	0.2577	0.1931	0.1833
2	0.4515	0.2625	0.1966	0.1269	0.1122
3	0.3840	0.2145	0.1777	0.1098	0.0821
4	0.3436	0.1877	0.1650	0.1010	0.0707
5	0.3160	0.1694	0.1582	0.0975	0.0657
6	0.2964	0.1575	0.1537	0.0959	0.0627

In order to estimate the efficiency of the proposed method the simulation of random field was performed using the method of diagonalizing the covariance metric described by Eqs. (10)–(18). Due to the large number of elements of the matrix **p** the simulations were performed for a field containing 208 nodes only.

The assumed and calculated values of the decay coefficient β are presented in Table 2.

Table 2. The values of the decay coefficient β

Coefficient β assumed [1/cm]	Coefficient β calculated after simulation [1/cm]
0.3868	0.3229
0.2784	0.2514
0.1923	0.2038
0.1349	0.1804
0.1193	0.1747

For a single simulation the real decay coefficient β calculated for a generated random field differs from the assumed coefficient (Tab. 2). One of the possible reasons for this discrepancy is the small number of nodes (208).

The results of comparison of execution times on PC using the Delphi 5 implementation are presented in Table 3.

Table 3. The times for the generation method

Number of nodes in the random field mesh	Time generation for	
	moving average method	method described by Wilde
10^2	0.4 sec.	0.6 sec.
10^3	1 sec.	1 min.
10^4	1 min 38 sec.	1 h 40 min

4. Conclusions

Two methods presented in this paper enable to generate a multidimensional random variable for an assumed correlation function. In the case of the proposed method based on the moving average the simulation process is very simple and relatively quick. With this method one can simulate large random fields on PC-s, which is also possible (but needs time-consuming operations) with the method described by Wilde due to excessive dimensions of matrix **p**. For a random field of n nodes there are approximately $n^2/2$ non-zero elements in the matrix **p**. Therefore the simple method of generating the random field presented in this paper should be useful for the stochastic analysis of transport phenomena in soil.

References

Brandt S. (1998), *Data Analysis*, PWN Warszawa (in Polish).

Dietrich C. R., Newsam G. N. (1993), A fast and exact method for multidimensional gaussian stochastic simulations, *Water Resour. Res.*, 29(8), 2861–2869.

Maciejewski S. (1998), *Solute Flow Processes in Unsaturated Soil*, IBW PAN Gdańsk (in Polish).

Mantoglou A., Wilson J. L. (1982), The turning bands method for the simulation of random fields using line generation by a spectral method, *Water Resour. Res.* 18(5), 1379–1394.

Przewłócki J. (1998), *Randomness of Selected Problems in Soil Mechanics*, Biblioteka Naukowa Hydrotechnika, IBW PAN Gdańsk (in Polish).

Przewłócki J. (2003), Simulation and discretization of random field in the slip-line method, *Archives of Hydro-Engineering and Environmental Mechanics*, Vol. 50, No. 4, 229–244.

Robin M. J. L., Gutjahr A. L., Sudicky E. A., Wilson J. L. (1993), Crosscorrelated random field generation with the direct Fourier transform method, *Water Resour. Res.*, 29(7), 2385–2397.

Rozanow J. A. (1974), *Introduction to Stochastic Process Theory*, PWN Warszawa (in Polish).

Węglarczyk S. (1999), *Statistical Methods*, Politechnika Krakowska Kraków (in Polish).

Wilde P. (1981), *The Discretization of Random Fields in Engineering Calculations*, PWN Warszawa-Poznań (in Polish).