Title
Random Fields Generation by the Source Point Method

Authors
Pozdniakov, S.P.
Tsang, Chin-Fu

Publication Date
1994-07-01

LBL-37100
UC-603
Preprint

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

EARTH SCIENCES DIVISION

Submitted to Journal of Mathematical Geology

Random Fields Generation by the Source Point Method

S.P. Pozdniakov and C.-F. Tsang

July 1994

Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098

LBL-37100
UC-603
RAC-8

Random Fields Generation by the Source Point Method

Sergey P. Pozdniakov

Hydrogeological Division Department of Geology
Moscow State University
Vorobyovy Hills, Moscow, 119899, Russia

Chin-Fu Tsang

Earth Sciences Division
Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

July 1994

This work was supported by the Department of Energy, Office of Environmental Management, Office of Technology Development (DOE/EM-OTD), and the Department of Energy, Office of Energy Research, Office of Basic Energy Sciences (DOE/ER-BES), under Contract No. DE-AC03-76SF00098.

Random Fields Generation by the Source Point Method

Sergey P. Pozdniakov* and Chin-Fu Tsang

Earth Sciences Division
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

ABSTRACT

A semi-empirical method of generating random 2-D fields with exponential correlation is developed. The most important features of this method are the relatively small number of calculations at each point of the field and the simplicity of the programming algorithm. In the present work, the method was improved for modeling anisotropic fields and fields where the dimension of the modeling domain is comparable to the correlation scale. An integral method of evaluating the results of the generated large field was developed by dividing the field into blocks and calculating the statistics of the generated data averaged over the blocks.

Key Words: heterogeneity, random field, anisotropy, correlation length

*Current Address: Hydrogeological Division Department of Geology, Moscow State University, Vorobyovy Hills, Moscow, 119899, Russia.

INTRODUCTION

Within the framework of stochastic hydrology theory, the estimation of aquifer properties and calculation of flow and contaminants transport requires the generation of random autocorrelated fields. Examples of the use of randomly generated fields in the investigation of ground water flow and transport can be found in the works of several authors, including Shvidler (1964), Tsang et al. (1988), Tompson and Gelhar (1990), and Desbarats (1992). In many cases the heterogeneity of the natural hydrogeological parameters (e.g., aquifer transmissivity, porosity, and thickness) can be accurately described by normal or normalized random fields with an exponential correlation function. There are several methods of generating such fields, and they are presented by Heller (1972), Smith and Freeze (1979), and Mantoglou and Wilson (1982). The present paper examines and extends the Source Point Methods (SPM) of Heller (1972) and Ghori et al. (1993). An important feature of this method is the relatively small number of calculations at each point of the generating field. The method also provides users with simple programming algorithms. However, a limitation is associated with this method—it lacks a strong theoretical background for the correlation function. The goal of this paper is to improve and extend the method for generating anisotropic 2-D random fields and fields with dimensions comparable to the correlation scale.

BACKGROUND OF METHOD

According to Ghori et al. (1993), the main idea of the SPM is to calculate a random set of function values $f(x_i,y_i)$ at given points by the values at the sources, whose coordinates are being randomly distributed as:

$$f(x_i,y_i)=\sum_{j=1}^{N}\alpha_{i,j}\cdot q_j \quad , \tag{1}$$

where q_i is the normally distributed value and $\alpha_{i,j}$ is the weighting function determined from

$$\alpha_{i,j} = \left[(x_j - x_i)^2 + (y_j - y_i)^2\right]^{-1} / \sum_{j=1}^{N}\left[(x_j - x_i)^2 + (y_i - y_j)^2\right]^{-1} \quad , \tag{2}$$

where N is the number of sources involved in the calculation of the function at a given point with x_i and y_i coordinates and x_j and y_j are the random coordinates of the j-source. Evidently,

$$\sum_{j=1}^{N}\alpha_{i,j} = 1 \quad , \quad \sum_{j=1}^{N}\alpha_{i,j}^2 \neq 1 \quad . \tag{3}$$

The performance of this method is exceptionally good because it only needs two generations of the uniformly distributed numbers for the coordinates of the sources, one generation of normally distributed numbers for source values, and a small number of calculations for each point of field according to eq. (1). It was shown in the work of Ghori et al. (1993) that by using some empirical relations between the correlation length and the number of involved sources one can generate a field with the form of the correlation function close to that of the exponential function.

The method, presented by Ghori et al. (1993), performs well. However, there are some areas that can be improved, and they can be identified by asking the following questions:

- How can one assume the mean and variance of the modeling field to be the same as preset values?
- Is it necessary to use every source number to calculate each point value of the field?
- How can one simulate a field that has dimensions comparable to the scale of correlation?

The Ghori et al. (1993) article also contains an empirical, nondimensional equation (see eq. (6) of this article). Such an approach makes it difficult to generate an anisotropic field with a different form from that of Ghori et al. (1993). This paper will present some efforts to improve the SPM's modeling of random fields along these lines.

IMPROVEMENTS TO SPM

Let us simulate a normal distributed stationary field $f(x_i, y_i)$, with a mean M and standard derivation σ, having the antisotropic correlation function $R(\bar{r})$ in the domain $\{X,Y\}$:

$$R(\vec{r}) = \exp\left(-\sqrt{(x / L_x)^2 + (y / \lambda L_x)^2}\right) \quad , \quad \lambda = L_y / L_x \geq 1 \quad , \tag{4}$$

where X, Y are the dimension of a domain and L_x, L_y, are correlation scales along the x and y axes, respectively.

Now let us transform and include the modeled domain into a unity square so that the point coordinates (x_i', y_i') in the latter are connected with the actual area coordinates (x_i, y_i) by the ratios:

$$y_i' = (y_i / \lambda + L_x) / D_{\max} \quad , \quad x_i' = (x_i + L_x) / D_{\max} \quad , \tag{5}$$

where $D_{\max} = \max\{X,Y\} + 2 \cdot L_x$. Having the simulated isotropic field $f^0(x_i', y_i')$ on the unity square with a zero mean and unit variance, one can obtain the actual anisotropic field by using the inverse coordinates transformation and scaling the field on the nonzero mean and actual variance:

$$f(x_i, y_i) = M + \sigma \cdot f^0(x_i, y_i) \quad . \tag{6}$$

Variation of the Generated Random Field Realization

Only in the case when $X >> L_x$ and $Y >> L_y$ will the variance of the mean σ_m^2 of random field $f^0(x_i', y_i')$ equal to 0 and the field variance σ_f^2 equal to 1. In practice the size of the modeled field often does not exceed ten times that of the correlation scale, so it is necessary to estimate beforehand the influence of the ratio of the field dimension to the correlation length on modeling the field statistics.

The variance of the field mean can be calculated as:

$$\sigma_m^2 = \frac{1}{X^2Y^2}\left\langle \iint f^0(x,y)dxdy \right\rangle = \frac{1}{X^2Y^2}\iiiint R(\vec{r})dxdydx'dy' \quad . \tag{7}$$

The variance of field can be calculated as:

$$\sigma_f^2 = \left\langle \frac{1}{XY}\iint (f^0(x,y))^2 dxdy - \frac{1}{X^2Y^2}\left(\iint f^0(x,y)dxdy\right)^2 \right\rangle = 1 - \sigma_m^2 \quad . \tag{7a}$$

The four-fold integral in eq. (7) was solved (for another problem) for an exponential isotopic correlated field by Shvidler (1964). The final equation for the variance of the field mean is:

$$\sigma_m^2 \approx \varphi'\left(\frac{Y}{L_y}\right)\varphi'\left(\frac{Y}{L_x}\right) \varphi'(z) = 0.5\left(\varphi(z) + \varphi\left(z / \sqrt{2}\right)\right) \quad , \tag{8}$$

where the function φ is

$$\varphi(z) = 2z^{-2}(e^{-2} + z - 1) \quad . \tag{9}$$

Another widely used function in stochastic theory is the Gaussian correlation function:

$$R(\vec{r}) = \exp(-x^2 / L_x^2 - y^2 / L_y^2) \quad . \tag{10}$$

The structure of the means's variance for this function is the same as for eq. (8)

$$\sigma_m^2 = \psi\left(\frac{Y}{L_y}\right)\psi\left(\frac{X}{L_x}\right) \quad , \tag{10a}$$

where the function ψ is

$$\psi(z) = z^{-2}\left(z\sqrt{\pi}erf(z) + e^{-z^2} - 1\right) \quad . \tag{11}$$

For a general n-dimensional case, the equation for the variation of the field's mean has the form

$$\sigma_m^2 = \prod_{i=1}^{n} \varphi(\lambda_i) \quad \text{or} \quad \sigma_m^2 = \prod_{i=1}^{n} \psi(\lambda_i) \quad , \quad \lambda_i = \frac{D_i}{L_i} \quad , \tag{12}$$

where n = 1, 2, 3 denotes the dimension of modeled field and D_i/L_i gives the ratio of field size to correlation length for each axis.

The influence of the ratio of the field size to the correlation length on the mean's variation is shown in Table 1 for the case of the 2-D square isotopic field with an exponential correlation. One can find from this table that the variance of the field's mean tends to 0 when the field size is ten times or more the correlation scale.

The influence of the ratio of the field size to the correlation length on the mean's variation is shown in Fig. 1 where n = 1,2,3 denotes the dimension of the isotropic field with an exponential correlation. Using the equation for mean variance and the normal distributed number, one can estimate the expected values of the mean and variance of the generated field.

Procedure for Field Calculation

To generate a field $f^0(x_i', y_i')$ using (1) and (2), it is necessary to estimate the total number N_s of sources in a unit square, the parameters of the q source value distribution, and the number of sources N for calculations at each point of the field. According to eqs. (1) and (2) the distribution of field values will be of the same type as the distribution of sources values, i.e., a normal distribution. The mean of the field value distribution will be equal to the mean of the source values distribution. There is no way a priori to estimate the variance of distribution of source values that would give unit variance of the

Table 1. The influence of the ratio of the field size, D, to the correlation length, L, on the variation of the mean.

D/L	0.5	1	2	4	6	8	10	20
σ_m^2	0.726	0.541	0.322	0.142	0.072	0.048	0.032	0.009

field. Thus, we use a distribution of source values with a zero mean and unit variance. In this case the distribution of values for the generated field $f^0(x_i', y_i')$ has nonunit (less than one) variance, and it has a mean close to zero. This distribution is then transformed to zero mean and unit variance $N(0,1)$ distribution by using the estimated values of mean and variance obtained during this field generation. The $N(0,1)$ distribution could easily be transformed onto another normal distribution with expected values of variance and mean.

For an estimation of N_s and N, we will use the following procedure. It is clear that the correlation length of the generated field is comparable to the search radius of the nearest source. It means that for search radius R_{search} it is possible to assume an empirical equation:

$$R_{search} \approx \beta \cdot L_x / D_{\max} \quad , \quad 1 \le \beta \le 2 \quad . \tag{13}$$

Judging from the structure equation of $\alpha_{i,j}$, these coefficients are a rapidly increasing function, and hence, the total number N of members of the series in eq. (1) have to be kept below 10–20.

From the above, it is possible to assume an empirical equation for N_s:

$$N_s = N \cdot \frac{D_{\max}^2}{L_x^2 \pi \beta^2} \quad ; \quad 1 \le \beta \le 2 \quad , \quad 10 \le N \le 20 \quad . \tag{13a}$$

After generating a series of fields for various values of N and β, it is found that the best values of N is 16 and the best value of β is 1.25.

In summary, the procedure for field generation consists of the following steps:

- Transform the field domain into a unit square.
- Calculate the mean variance and the expected field variance according to eqs. (7) and (8).
- Simulate the expected value of the field mean, using its estimated variance.
- Calculate the number of sources according to eq. (13a).
- Generate uniform random coordinates of sources on the unit square.

- Generate a normal distributed sources value with $N(0.1)$ distribution.
- Calculate field values on the unit square according to eqs. (1) and (2) using N nearest sources.
- Transform field values to an $N(0,1)$ distribution.
- Inverse the coordinates transformation.
- Find the field values using the expected values of mean and variance according to eq. (6).
- Transform the field as needed by exponentiation, etc.

AN EXAMPLE OF RANDOM FIELD GENERATION

Random fields of hydraulic transmissivity were generated for Monte Carlo simulations of transport in the heterogeneous aquifer. From the empirical data, the estimated value of the variance of log transmissivity was 1,4. Generating a random field of transmissivity with such a large variance could lead to abnormally high (on the order of $n^*10^4 m^2/day$) and abnormally low (less than $10^{-3}\ m^2/day$) transmissivity values. Thus a distribution was sought that would limit the extreme values near the small and large probability values but give values that correspond to the lognormal distribution near the median. The Johnson's S-V distribution was chosen—which is a normal distribution for the transformed function F_T

$$F_T = \log \frac{T - T_{\min}}{T_{\max} - T} \tag{14}$$

where T is transmissivity, and $T_{\min}$ are its minimum and maximum values. Parameters of the transmissivity distribution are:

$$M_F = -2.73 \quad , \quad \sigma_F^2 = 2.0 \quad , \quad T_{\min} = 0.4 \frac{m^2}{day} \quad , \quad T_{\max} = 870 \frac{m^2}{day} \quad .$$

The horizontal exponential correlation scales of F_T were $L_x = 30$ and $L_y = 90$ meters. The transformation used to obtain the random transmissivity field is in two steps:

a) Generation of $F_T(x,y)$ field,

b) transform $F_T(x,y)$ into $T(x,y)$ according to

$$T = \frac{T_{\max} \exp(F_T) + T_{\min}}{1 + \exp(F_T)} \quad .$$

The transmissivity field was generated on a square mesh 560 × 560 meters with a constant grid of 10 × 10 meters.

The estimation of field and correlation scale size interaction according to eqs. (7) and (8) gave the next values for variances: $\sigma_m^2 \approx 0.037$; $\sigma_f^2 \approx 0.963$.

More than 50 field realizations were simulated. For each realization, the empirical distribution of F_T and three variograms (for all points, X-X and Y-Y directions) were calculated. For each variogram the parameters of the exponential curve were found. In all variants the distribution shape was normal and the estimated correlation lengths were close to the modeled values. The standard derivation of the estimated correlation lengths was less than 30%.

An example of the generated field of transmissivity is shown in Fig. 2. Empirical variograms of the transformed field of transmissivity calculated for this realization of field according to eq. (13), are shown in Fig. 3. The less than expected value of 3 for the ratio of correlation lengths can be explained by the variogram method in the X-X and Y-Y calculations. For the calculation of variograms with the dependence on direction, a search of nearest points is performed for over an angle 90 quadrant.

Table 2 compares computing time on an i486/33DX machine for this method with that for the Turning Bands Method (TBM) of Tompson and Gelhar (1990).

Table 2. Comparison of computing time for SPM vs TBM methods on an i486/33DX computer.

Method	Compiler	cpu time for one simulation
SPM	VBDOX 1.0	55 sec
TMB	FORTRAN Power Station 1.0	2 min and 15 sec

The table indicates that the SPM is an efficient method of generating random fields with regard to computer time. This method could be used when many realizations of the field are needed, but when an accurate reproduction of the theoretical correlation function is not required. It is easy to extend this method for 3-D field generation.

VERIFICATION OF RESULTS OF FIELD GENERATION

The generated field must satisfy the predetermined correlation structure independent of the method of field generation. In this section we will give the integral method of verification of results of generation without regard to the SPM. We also will consider the method of verifying the general case of an *n*-dimensional field.

The best verification measure is to calculate the variogram or the correlation function of the field that has been generated. Often the number of generated points of the field within the modeled domain exceeds $10^5 - 10^6$ (Tompson and Gelhar, 1990). For such a large field the calculation of the variogram *n* using all points and searching in all directions can take a long time. The main idea of the integral method is to consider the right-angled blocks oriented along the main axes of correlation having sides of length B_k, where $k = 1,...,n$. Let us consider the covariation function averaged over the block value f_{ac} of random function f that has variance σ^2 and the Gaussian correlation function. This covariation function $Cov(\tau,B)$ can be represented in the form:

$$Cov(\tau, B) = \sigma^2 \prod_{k=1}^{n} Cov_0(\tau_k / L_k, B_k / L_k) \quad , \qquad (16)$$

where the partial covariation $Cov_0(\tau/L, B/L)$ is

$$Cov_0(\tau/L, B/L) = B^{-2} \int_{x-0.5B}^{x+0.5B} \int_{x+\tau-0.5B}^{x+\tau+0.5B} \exp(-(\theta_1-\theta_2)^2/L^2)\,d\theta_1 d\theta_2 \quad .$$

Upon integrating eq. (17) we obtain the equation for a partial covariation as a function of dimensionless block side length $\overline{B} = B/L$ and dimensionless displacement $\overline{\tau} = \tau/L$:

$$\begin{aligned} &Cov(\overline{B}, \overline{\tau}) = \\ &\quad 0.5\overline{B}^{-1}\left\{\sqrt{\pi}\left[(\tau_0+1)erf(\overline{\tau}+\overline{B}) - 2\tau_0 erf(\overline{\tau}) + (\tau_0-1)erf(\overline{\tau}-\overline{B})\right]\right. \\ &\quad \left.+\overline{B}^{-1}\left[\exp(-(\overline{\tau}+\overline{B})^2) - 2\exp(-\overline{\tau}) + \exp(-(\overline{\tau}-\overline{B})^2)\right]\right\} \quad ; \quad \tau_0 = \tau/B \quad . \end{aligned} \tag{18}$$

For zero displacement, partial covariation equals the function $\psi(\overline{B})$ defined by eq. (11). It means that the dimensionless variance $\overline{\sigma}_b^2 = \sigma_b^2/\sigma^2$ of the block values of random function can then be calculated from eq. (12) with the parameter λ_i, the ratio of the length of each side of the correlation scale L in this direction. For a zero block size, eq. (18) approaches the equation for the correlation function of the field eq. (10).

One can divide the field into blocks and calculate the variance of averaged block value as a function of block size and compare the results with theoretical numbers from eq. (12). For the widely used lognormal distribution, an effective verification of the generation is to calculate the coefficient of variation of the block value. It is easy to show that for the Gaussian model of log correlation the equation for the coefficient of variation ξ has the form:

$$\xi^2 = \exp\left(\sigma^2 \prod_{k=1}^{n} \psi\left(\frac{B_k}{L_k}\right)\right) - 1 \quad . \tag{19}$$

This equation depends on the size of the blocks, the correlation scale, and the type of field correlation. If the empirical data agree well with the various block sizes, they should agree well for the zero size block, i.e., for the original field.

As an example of verification of the generation in Fig. 4, one can compare the theoretical values calculated from the coefficient of variation from eq. (19) and field generation data from the work of Dykaar and Kitanidis (1993). The experimental data are obtained by averaging values of the lognormal Gaussian correlated field generated on a 3-D fine mesh with the length of the cubic cell equal to L/π, increasing step by step the size of blocks in a horizontal direction for averaging. Each experimental point is computed from a sample size of 10 realizations, and each realization is generated with the variance equal to two. Theoretical curves are calculated using eq. (19). The good agreement between the theoretical and experimental values shows that the method of field generation in this paper gives the tolerance realizations of the expected field.

ACKNOWLEDGMENTS

This paper was prepared under the auspices of the Russian-American Center for Contaminant Transport Studies at the Lawrence Berkeley Laboratory. We appreciate funding support from the Office for Environmental Management, Office of Technology Development and from the Office of Energy Research, Engineering and Geosciences Division of the U.S. Department of Energy, under Contract No. DE-AC03-76SF00098. We would like also to acknowledge support from the Russian Academy of Science through Grant N93-05-89-21.

REFERENCES

Desbarats, A.J. (1992), "Spatial Averaging of the Transmissivity in Heterogeneous fields with Flow Toward Well," *Water Resources Research 3*, 757–767.

Dykaar, B.B. and Kitanidis, P.K. (1993), "Transmissivity of Heterogeneous Formation," *Water Resources Research 4*, 985–1001.

Gelhar, L.J. (1993), Stochastic Subsurface Hydrology, Prentice-Hall, Inc., New Jersey.

Ghori, S.G., Heller, J.P., and Singh, A.K. (1993), "An Efficient Method of Generating Random Permeability Fields by the Source Point Method," *Mathematical Geology 5*, 559–572.

Heller, J.P. (1972), "Observations of Mixing and Diffusion in Porous Media," *Proc. Second Symposium, Fundamentals of Transport in Phenomena in Porous Media*, AHR-ISSS, Ontario, Canada, pp. 1–26.

Mantoglou, A. and Wilson, J.L. (1982), "The Turning Bands Method for Simulation of Random Field Using Line Generation by a Spectral Method," *Water Resources Research 18*, 1379–1394.

Smith, L. and Freeze, R.A. (1979), "Stochastic Analysis of Steady State Ground Water Flow in a Bounded Domain, 1. One Dimension Simulation," *Water Resources Research 3*, 521–535.

Shvidler, M.I. (1964), *Filtration Flow in Heterogeneous Media*, Cons. Bureau, New York.

Tsang, Y.W., Tsang, C.F., Neretnieks, I., and Moreno, L. (1988), "Flow and Tracer Transport in Fractured Media: A Variable Aperture Channel Model and Its Properties," *Water Resources Research 12*, 2049–2060.

Tompson, A.B. and Gelhar, L.W. (1990), "Numerical Simulation of Solute Transport in Three-Dimensional, Randomly Heterogeneous Porous Media," *Water Resources Research 10*, 2541–2562.

FIGURE CAPTIONS

Figure 1. Relation between the variation of the mean and ratio of the field size, D, to the correlation length, L, for 1, 2, 3 dimensional cases.

Figure 2. An example of random anisotropic field of transmissivity.

Figure 3. An experimental variogram of random anisotropic field.

Figure 4. The comparison of theoretical values of coefficient of variation of transmissivity averaged into the blocks and data from numerically generated fields (Dykaar and Kitanidis, 1993), where h is the vertical length of the block in units of integral correlation scale.

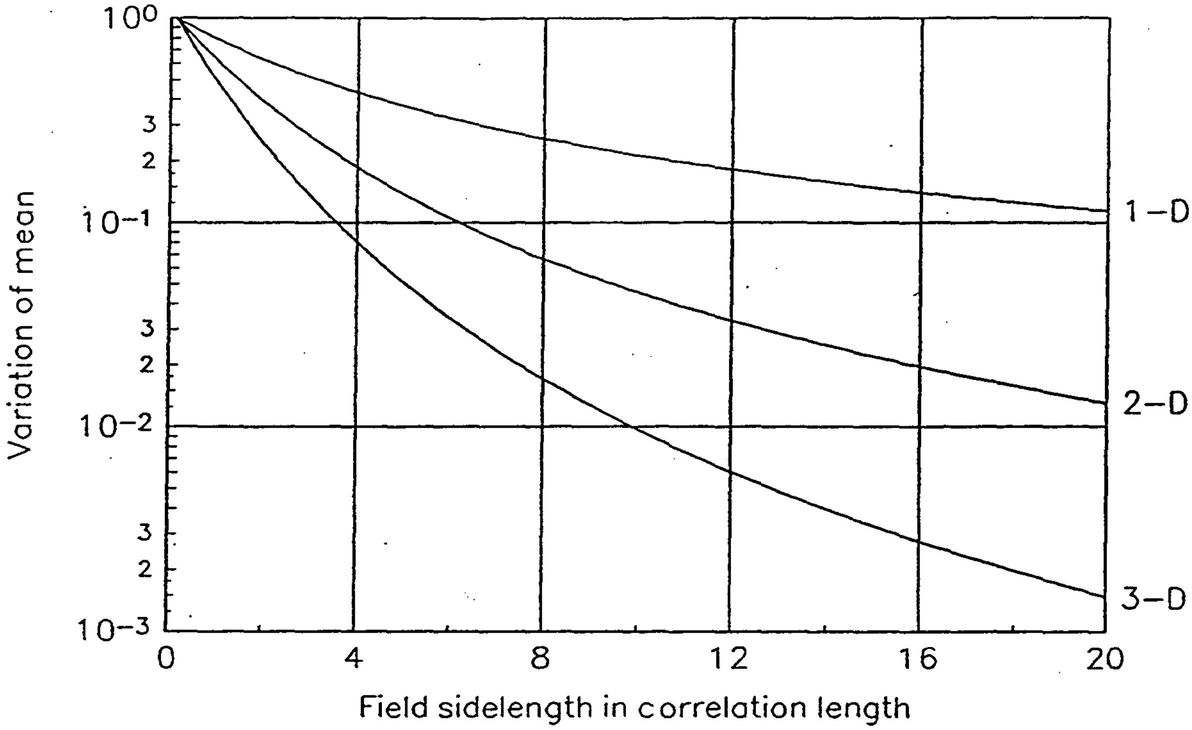

Figure 1. Relation between the variation of the mean and ratio of the field size, D, to the correlation length, L, for 1, 2, 3 dimensional cases.

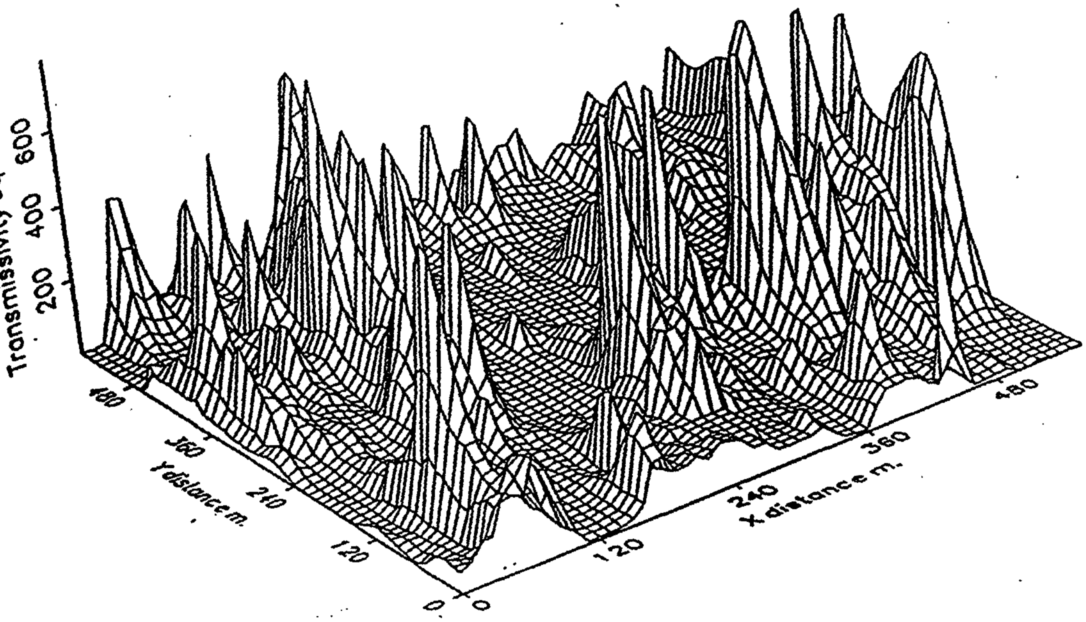

Figure 2. An example of random anisotropic field of transmissivity.

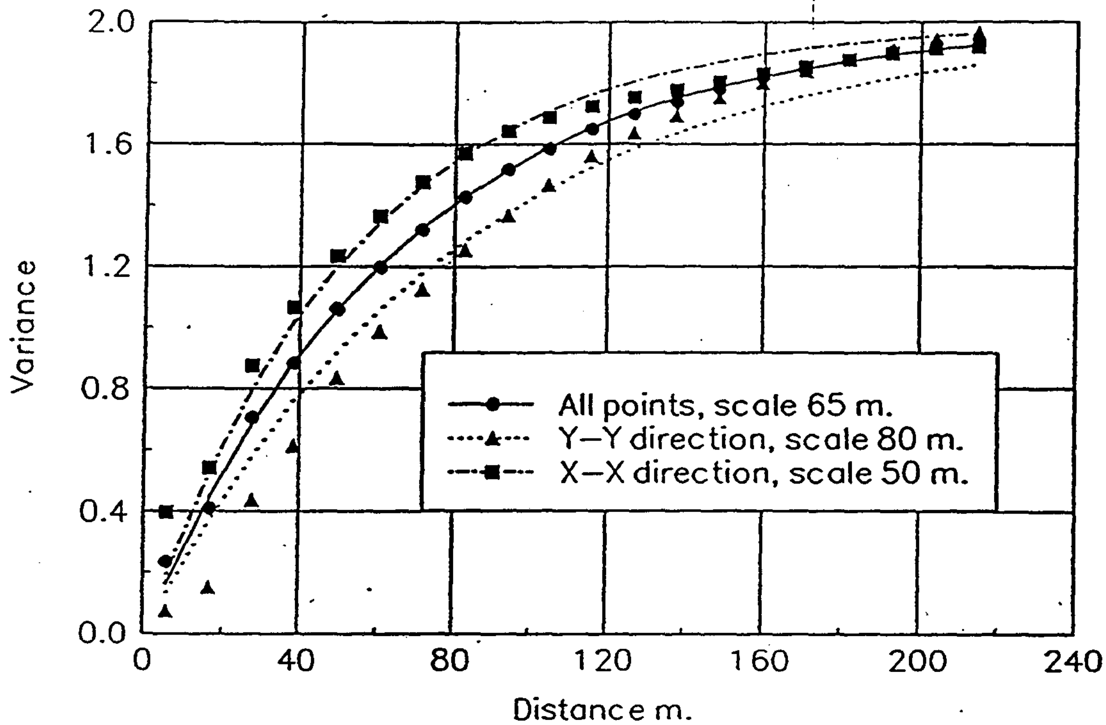

Figure 3. An experimental variogram of random anisotropic field.

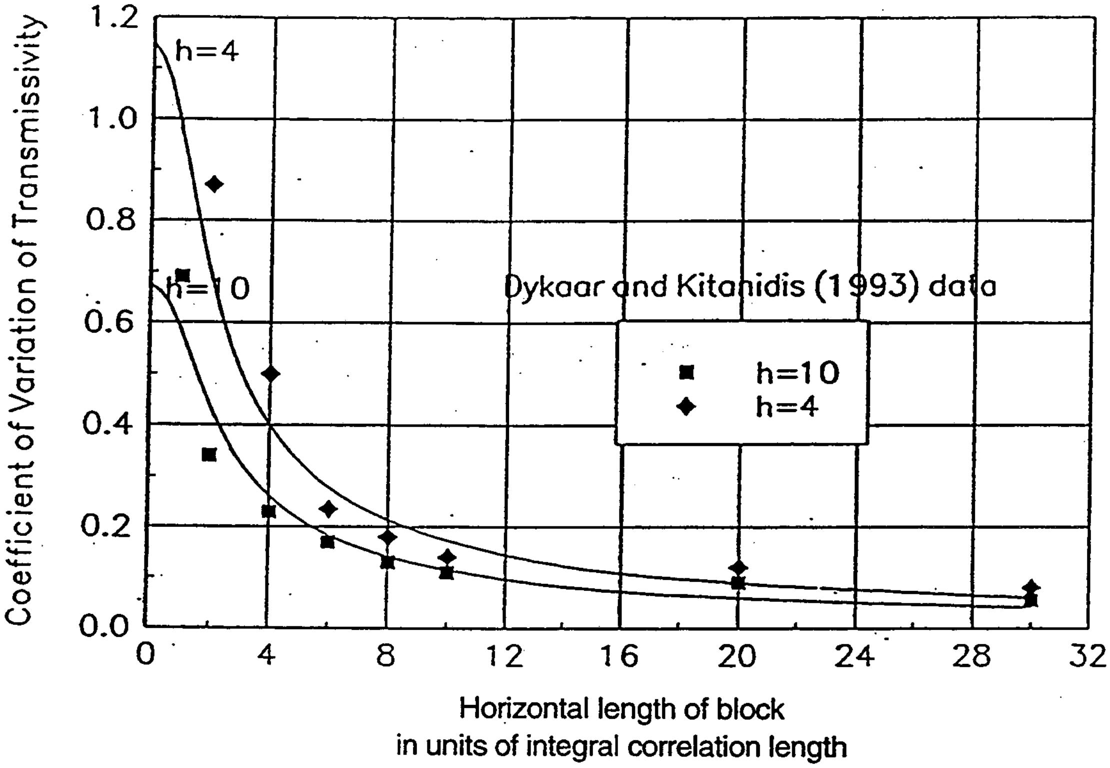

Figure 4. The comparison of theoretical values of coefficient of variation of transmissivity averaged into the blocks and data from numerically generated fields (Dykaar and Kitanidis, 1993), where h is the vertical length of the block in units of integral correlation scale.

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
TECHNICAL AND ELECTRONIC
INFORMATION DEPARTMENT
BERKELEY, CALIFORNIA 94720