A Covariance Conversion Approach of Gamma Random Field Simulation

Jun-Jih Liou ¹, Yuan-Fong Su ¹, Jie-Lun Chiang ² and Ke-Sheng Cheng ¹ +

¹ Department of Bioenvironmental Systems Engineering, National Taiwan University, Taiwan
² Dept. of Soil and Water Conservation, National Pingtung University of Science and Technology, Taiwan

Abstract. In studies involving environmental risk assessment, random field generators such as the sequential Gaussian simulation are often used to yield realizations of a Gaussian random field, and then realizations of the non-Gaussian target random field are obtained by an inverse-normal transformation. Such simulation requires a set of observed data for estimation of the empirical cumulative distribution function and covariance function of the random field under investigation. However, such observed-data-based simulation will not work if no observed data are available and realizations of a non-Gaussian random field with specific probability density and covariance function are needed. In this paper we present details of a gamma random field simulation approach which does not require a set of observed data. A key element of the approach lies on the theoretical relationship between the covariance functions of a gamma random field and its corresponding standard normal random field. The proposed gamma random field simulation technique is composed of three sequential components: (1) covariance function conversion between a gamma random field and a corresponding Gaussian random field, (2) generating realizations of a Gaussian random field with standard normal density and the desired covariance function, and (3) transforming Gaussian realizations to corresponding gamma realizations. Through a set of devised simulation scenarios, the proposed technique is shown to be capable of generating realizations of the given gamma random fields. The approximation function of the gamma-Gaussian covariance conversion works well for coefficient of skewness of the gamma density not exceeding 3.0.

Keywords: stochastic simulation, gamma distribution, geostatistics, random field simulation

1. Introduction

Random field simulation has been widely applied to studies involving environmental risk assessment such as soil contamination, flood inundation, solutes transport in a porous medium, etc. Many of these applications use random field generators to generate large sets of realizations which can be characterized by the desired Gaussian distribution and spatial variation structure (or semi-variogram). However, many environmental variables are non-Gaussian and asymmetric. For practical applications involving uncertainty or risk mapping of these variables, observed data are transformed to normal scores in the beginning, and sequential Gaussian simulation (SGS) is then performed in normal space. Finally, data generated by SGS are back-transformed to realizations of the target variables. In such practices, normal transformation of the observed data and inverse-normal transformation of the simulated data are generally based on the empirical cumulative distribution function (ECDF) of the observed data.

From a theoretical point of view, if no observed data are available and we want to generate realizations of a non-Gaussian random field with specific probability density and covariance function, the above observed-data-based SGS process will not work since the covariance function of the Gaussian random field cannot be estimated without a set of normal scores transformed from observed data. In this paper a gamma random field simulation approach which does not require observed data is presented.

+ Corresponding author. Tel.: +886-2-33663465; fax: +886-2-23635854.
E-mail address: rslab@ntu.edu.tw

2. Conceptual description of the proposed approach

Based on the bivariate gamma simulation approach proposed by Cheng et al. (2008), simulation of a gamma random field can be achieved in a similar, yet more complicated, manner. Through a theoretical conversion between \( \rho_{\text{UV}} \) and \( \rho_{XY} \), random samples of a pair of bivariate gamma random variates \((X, Y)\) with desired marginal densities and correlation coefficient \( \rho_{XY} \) can be obtained by first generating random samples of a corresponding pair of bivariate variates \((U, V)\) with standard normal density and correlation coefficient \( \rho_{\text{UV}} \), followed by a Gaussian-to-gamma transformation. In contrast to stochastic simulation of bivariate random variables, stochastic simulation of a random field \( Z(x) \) involves correlations among a set of random variables which are expressed in terms of a multivariate covariance (or correlation) matrix. For random field simulation, the whole process is composed of three sequential components:

- Converting the covariance function \( C_Z(h) \) of a gamma random field \( Z(x) \) to the covariance function \( C_W(h) \) of a corresponding Gaussian random field \( W(x) \). In a sequential random field simulation process, the covariance function appears as a covariance matrix \( \Sigma \) which involves a point for random number generation and its neighboring points. Thus, conversion between the covariance function \( C_Z(h) \) and \( C_W(h) \) is equivalent to conversion between the covariance matrices \( \Sigma_Z \) and \( \Sigma_W \).
- Generating realizations of the Gaussian random field with covariance function \( C_W(h) \), and
- Transforming realizations of \( W(x) \) to corresponding realizations of the gamma random field \( Z(x) \).

3. Methodology and detailed procedures

3.1. Sequential Gaussian simulation

Given a homogeneous and isotropic random field \( \{Z(x), x \in \Omega\} \) with known probability density function \( f_Z(z) \) and covariance function \( C_Z(h) \) (or semi-variogram \( \gamma_Z(h) \)), we want to generate as many random samples (realizations) of the random field. For a set of \((p+q)\) normally distributed random variables, say \( W' = (W_1, \ldots, W_{p+q}) \), the multivariate joint density is given by (Morrison, 1990)

\[
f_W(w) = \frac{1}{(2\pi)^{(p+q)/2} |\Sigma_W|^{1/2}} e^{-\frac{1}{2}(w-\mu)^T \Sigma_w^{-1} (w-\mu)}
\]

where \( \Sigma_W \) is the covariance matrix with \((p+q) \times (p+q)\) dimension and \( \mu \) is the \((p+q)\) dimensional mean vector. Let \( W \) be divided into two subsets \( W'_1 = (W_1, \ldots, W_p) \) and \( W'_2 = (W_{p+1}, \ldots, W_{p+q}) \), the mean vector and the covariance matrix can thus be respectively expressed as

\[
\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}
\]

and

\[
\Sigma_W = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{bmatrix}
\]

where \( \mu_1 \) and \( \mu_2 \) are respectively the mean vectors of \( W_1 \) and \( W_2 \), and \( \Sigma_y \) is the covariance matrix of \( W_i \) and \( W_j \) (\( i, j = 1 \) or \( 2 \)).

Suppose that values of \( W_2 \) are known, i.e. \( W'_2 = (w_{p+1}, \ldots, w_{p+q}) \), the conditional multivariate density of \( W'_1 \) given \( W'_2 = w_2 \) can then be expressed by (Morrison, 1990)

\[
f_{W'_1|W'_2}(w'_1 \mid w_2) = \frac{1}{(2\pi)^{p/2} |\Sigma^*|^{1/2}} e^{-\frac{1}{2}(w'_1-\mu)^T \Sigma^*-1 (w'_1-\mu)}
\]

\[
\mu^* = \mu_{W'_1|W'_2} = \mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (w_2 - \mu_2)
\]

\[
\Sigma^* = \Sigma_{W'_1|W'_2} = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12}^T
\]
Equation (4) lays the foundation for stochastic simulation of a Gaussian random field.

### 3.2. Covariance matrices conversion

In the previous section we have implicitly assumed the covariance function $C_W(h)$ is given or known. This unfortunately raises a question of how can we be sure that our stochastic simulation process using the assumed $C_W(h)$ will eventually yield realizations of $Z(x)$ which are associated with the desired covariance function $C_Z(h)$. Apparently, the covariance function $C_W(h)$ cannot be arbitrarily chosen and a conversion between $C_W(h)$ and $C_Z(h)$ (or between $\Sigma_Z$ and $\Sigma_W$) is necessitated.

Let $X$ be a gamma random variable, i.e., $X \sim \Gamma(\alpha, \lambda)$, with the following density

$$f_X(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \quad \alpha, \lambda > 0 \text{ and } 0 \leq x < +\infty.$$  \hspace{1cm} (7)

We proved that the covariance of two gamma random variables $X_1$ and $X_2$ can be related to covariance of two corresponding standard normal random variables $W_1$ and $W_2$, i.e., $\rho_{W_1,W_2}$, by the following equation:

$$\text{COV}(X_1, X_2) \approx \left(\frac{\alpha_1}{\lambda_1}\right)\left(\frac{\alpha_2}{\lambda_2}\right)(A + B\rho_{W_1,W_2} + C\rho_{W_1,W_2}^2 + D\rho_{W_1,W_2}^3) - \left(\frac{\alpha_1}{\lambda_1}\right)\left(\frac{\alpha_2}{\lambda_2}\right)$$  \hspace{1cm} (8)

where

$$\tilde{\xi}_1 = \frac{1}{9\alpha_1}, \quad \tilde{\xi}_2 = \frac{1}{9\alpha_2},$$

$$A = (1 - \tilde{\xi}_1)^3 (1 - \tilde{\xi}_1) + 3(1 - \tilde{\xi}_1)\tilde{\xi}_2 (1 - \tilde{\xi}_2)^3 + 3(1 - \tilde{\xi}_2)^3 (1 - \tilde{\xi}_2)\tilde{\xi}_1 (1 - \tilde{\xi}_1)\tilde{\xi}_2 + 9(1 - \tilde{\xi}_2)^5 (1 - \tilde{\xi}_1)\tilde{\xi}_2$$

$$B = 9(1 - \tilde{\xi}_1)^2 \tilde{\xi}_1 + 0.5 (1 - \tilde{\xi}_2)^2 \tilde{\xi}_2^0.5 + 9\tilde{\xi}_1 1.5 (1 - \tilde{\xi}_2)^2 \tilde{\xi}_2^0.5 + 9(1 - \tilde{\xi}_1)^2 \tilde{\xi}_1^0.5 \tilde{\xi}_2^1.5 + 9\tilde{\xi}_1^{1.5} \tilde{\xi}_2^{1.5}$$

$$C = 18(1 - \tilde{\xi}_1)\tilde{\xi}_1 (1 - \tilde{\xi}_2)\tilde{\xi}_2$$

$$D = 6\tilde{\xi}_1^{1.5} \tilde{\xi}_2^{1.5}$$

Approximation by Equation (8) works well for $\alpha_1, \alpha_2 \geq 0.5$, or equivalently, coefficients of skewness of $X_1$ and $X_2$ being less than 3.0. Equation (8) can be used to convert entry values in $\Sigma_Z$ to their corresponding values in $\Sigma_W$. It should be emphasized that the covariance matrices conversion is unique since the conversional relationship in Equation (8) is also a single-value function.

### 3.3. Transforming Gaussian realizations to gamma realizations

It can be shown that if $X$ is a gamma random variable with density function of Eq.(7), then a new random variable $Y = g(X) = 2\lambda X$ is chi-square distributed with degree of freedom $k = 2\alpha$. An approximation of the chi-squared distribution by the standard normal distribution, known as the Wilson-Hilferty approximation, is given as follows (Patel and Read, 1996)

$$y \approx 2\alpha \left[1 - \frac{1}{9\alpha} + w \sqrt{\frac{1}{9\alpha}}\right]^3$$  \hspace{1cm} (9)

where $w$ represents the standard normal deviate and $y$ is the corresponding chi-squared variate. Thus, the transformation of the standard normal deviate $w$ to gamma variate $x$ can be derived as

$$x = \frac{y}{2\lambda} \approx \frac{\alpha}{\lambda} \left[1 - \frac{1}{9\alpha} + w \sqrt{\frac{1}{9\alpha}}\right]^3.$$  \hspace{1cm} (10)

We end this section with the following summary of the simulation procedures:

- Generate a standard Gaussian random number at the initial node,
- Determine neighboring nodes to be involved in the subsequent simulation by considering range of the random field $Z(x)$, and use the covariance function $C_Z(h)$ to establish the covariance matrix $\Sigma_Z$,
- Transform $\Sigma_Z$ to $\Sigma_W$ using Equation (8),
- Generate a standard Gaussian number at the target node using the conditional Gaussian density of Equation (4),
- Repeat the above procedures until a realization of the standard Gaussian random field is established,
Perform point-to-point Gaussian-to-gamma transformation using Equation (10), and it yields a realization of the gamma random field with desired properties.

4. Simulation and verification

In order to demonstrate the implementation and verification of the proposed gamma random field simulation approach, several simulation scenarios (see Table 1) using the spherical semi-variogram model with range $a$ and sill $\omega$ were adopted for generation of realizations.

Table 1. Parameters of the gamma density and spherical semi-variogram model designated for random field stochastic simulation.

<table>
<thead>
<tr>
<th>ParameterEs</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>3.9698</td>
<td>3.8720</td>
<td>3.8765</td>
<td>4.0265</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>1.9863</td>
<td>1.9347</td>
<td>1.9393</td>
<td>1.9902</td>
</tr>
<tr>
<td>$\omega$</td>
<td>0.9976</td>
<td>1.0014</td>
<td>1.0098</td>
<td>1.0372</td>
</tr>
<tr>
<td>$a$</td>
<td>NA</td>
<td>2.0115</td>
<td>3.0631</td>
<td>6.0735</td>
</tr>
</tbody>
</table>

Table 2. Sample mean of the parameter estimators of the density function and the semi-variogram model.

One hundred simulation runs were conducted for a specific simulation scenario, and density and variogram parameters were estimated from each of the 100 realizations. Parameters (sill $\omega$ and range $a$) of the spherical semi-variogram model were estimated by the ordinary least squares fitting method. The method of ordinary least squares assumes that different data pairs involved in semi-variogram fitting are uncorrelated. Except for realizations with very short range as compared to the size of simulation, such assumption is a clear violation of the simulated realizations which are associated with certain spatial covariance matrix. Thus, it inevitably introduces higher degree of uncertainty in the process of parameter estimation. Other semi-variogram fitting methods such as the weighted least squares (Cressie, 1985; Gotway, 1991; Pardo-Iguzquiza, 1999) and the generalized least squares (Pardo-Iguzquiza and Dowd, 2001) have also been proposed, and the uncertainty of semi-variogram parameter estimation using these methods has also been addressed (Pardo-Iguzquiza and Dowd, 2001).

These estimates vary among realizations. Therefore, the sample mean of these realization-specific parameter estimates were calculated and tabulated in Table 2. Generally speaking, the estimated values are in very good agreement with the true parameter values. It is worthy to note that since a random field with one-pixel range is technically pure random, no attempt was made to fit the spherical semi-variogram model to the corresponding experimental semi-variograms. Finally, images of several simulated realizations are shown in Figure 1. As can be expected, when the range $a$ increases, higher degree of spatial correlation can be clearly observed.

5. Conclusions

In this study we present a technique for gamma random field simulation. The proposed gamma random field simulation technique is composed of three sequential components: (1) covariance function conversion between a gamma random field and a corresponding Gaussian random field, (2) generating realizations of a Gaussian random field with standard normal density and the desired covariance function, and (3) transforming Gaussian realizations to corresponding gamma realizations. Through a set of devised
simulation scenarios, the proposed technique is shown to be capable of generating realizations of given gamma random fields. The proposed gamma random field simulation technique may be very useful for risk assessment and spatial modeling of non-negative and asymmetric environmental variables.

Fig. 1: Images of simulated realizations of a gamma random field ($\alpha = 4, \lambda = 2, \omega = 1$, size of simulation 80×80).
Grey levels represent simulated values of the random field.

6. Acknowledgements

This research was originally initiated from a project funded by the Council of Agriculture, Taiwan, R.O.C. We also thank the National Science Council for subsequent funding support.

7. References