Simulation of Stationary Gaussian Random Fields with a Gneiting Spatio-Temporal Covariance

Denis Allard, Xavier Emery, Céline Lacaux, and Christian Lantuéjoul

Abstract The nonseparable Gneiting covariance has become a standard to model spatio-temporal random fields. Its definition relies on a completely monotone function associated with the spatial structure and a conditionally negative semidefinite function associated with the temporal structure. This work addresses the problem of simulating stationary Gaussian random fields with a Gneiting-type covariance. Two algorithms, in which the simulated field is obtained through a combination of cosine waves are presented and illustrated with synthetic examples. In the first algorithm, the temporal frequency is defined on the basis of a temporal random field with stationary Gaussian increments, whereas in the second algorithm the temporal frequency is drawn from the spectral measure of the covariance conditioned to the spatial frequency. Both algorithms perfectly reproduce the correlation structure with minimal computational cost and memory footprint.

Keywords Substitution random fields · Spectral simulation · Spectral measure · Central limit approximation

1 Introduction

The modeling, prediction and simulation of stationary random fields defined on Euclidean spaces crossed with the time axis, $\mathbb{R}^k \times \mathbb{R}$ with, in general, $k = 2$ or 3, is widespread in hydrology, environment, climate, ecology and epidemiology appli-

D. Allard (✉)
INRAE (BioSP), 84914 Avignon, France
e-mail: denis.allard@inrae.fr

X. Emery
University of Chile, Santiago, Chile

C. Lacaux
Avignon University, LMA EA 2151, 84000 Avignon, France

C. Lantuéjoul
Mines Paris, PSL University, Fontainebleau, France

cations. The representation of their correlation structure via traditional covariance models in $\mathbb{R}^{k+1}$ is often unsuitable to capture space-time interactions, reason for which specific models need to be developed. One of these, the Gneiting covariance, is widely used in climate studies due to its versatility, and is defined as

$$C(\boldsymbol{h}, u) = \frac{\sigma^2}{\left(1 + \gamma(u)\right)^{k/2}} \, \varphi\left(\frac{|\boldsymbol{h}|^2}{1 + \gamma(u)}\right), \qquad \boldsymbol{h} \in \mathbb{R}^k, \, u \in \mathbb{R}, \qquad (1)$$

where $\sigma > 0$, γ is a variogram (i.e., a conditionally negative semidefinite function) on $\mathbb{R}$ and φ is a completely monotone function on $\mathbb{R}_+$. A subclass originally proposed by Gneiting in [1] is obtained by considering a variogram γ of the form $\gamma(u) = \psi(u^2) - 1$, where ψ is a Bernstein function, i.e., a positive primitive of a completely monotone function. The general formulation (1), in which γ can be any variogram on $\mathbb{R}$, is due to Zastavnyi and Porcu in [2]. Hereinafter, without loss of generality, we assume $\sigma = 1$ and $\varphi(0) = 1$.

This work deals with the problem of simulating a stationary Gaussian random field with zero mean and Gneiting covariance on a (structured or unstructured) grid of $\mathbb{R}^k \times \mathbb{R}$. The following section presents some theoretical results, which will be used in Sects. 3 and 4 to design two simulation algorithms, which will be illustrated on synthetic examples. Concluding remarks follow in Sect. 5.

2 Theoretical Results

The completely monotone function φ can be written as a nonnegative mixture of decreasing exponential functions on $\mathbb{R}_+$:

$$\varphi(t) = \varphi(0) \int_{\mathbb{R}_+} \exp(-r\,t)\,\mu(dr), \qquad t \in \mathbb{R}_+, \qquad (2)$$

where μ is a probability measure. Also, the continuous Fourier transform of a squared exponential function in $\mathbb{R}^k$ is another squared exponential function:

$$\int_{\mathbb{R}^k} \cos(\langle \boldsymbol{\omega}, \boldsymbol{h} \rangle) \, \exp(-a\,|\boldsymbol{\omega}|^2) d\boldsymbol{\omega} = \left(\frac{\pi}{a}\right)^{k/2} \exp\left(-\frac{|\boldsymbol{h}|^2}{4a}\right) \qquad \boldsymbol{h} \in \mathbb{R}^k, \, a > 0, \qquad (3)$$

where $\langle \cdot, \cdot \rangle$ stands for the usual scalar product in $\mathbb{R}^k$.

Proposition 1 *By combining Eqs. (2) and (3), one can rewrite the Gneiting covariance (1) as follows:*

$$C(\boldsymbol{h}, u) = \int_{\mathbb{R}} \int_{\mathbb{R}^k} \cos(\sqrt{2r}\,\langle \boldsymbol{\omega}, \boldsymbol{h} \rangle) \, \exp\left(-\frac{\gamma(u)}{2}\,|\boldsymbol{\omega}|^2\right) g_{I_k}(\boldsymbol{\omega}) \, d\boldsymbol{\omega}\,\mu(dr), \qquad (4)$$

with I_k the identity matrix of order k and g_{I_k} the probability density of a k-dimensional Gaussian random vector with zero mean and covariance matrix I_k.

The mappings $\boldsymbol{h} \mapsto \cos(\sqrt{2r} \langle \boldsymbol{\omega}, \boldsymbol{h} \rangle)$ and $u \mapsto \exp(-\gamma(u) |\boldsymbol{\omega}|^2/2)$ are covariances functions in $\mathbb{R}^k$ and $\mathbb{R}$, respectively. Their product is therefore a covariance function in $\mathbb{R}^k \times \mathbb{R}$, and so is $C(\boldsymbol{h}, u)$ as a positive mixture of covariances functions in $\mathbb{R}^k \times \mathbb{R}$. This result proves that every member of the Gneiting class (1) is a valid space-time covariance. In particular C is a positive semidefinite function in $\mathbb{R}^k \times \mathbb{R}$.

Proposition 2 *One can further decompose the Gneiting covariance as follows:*

$$C(\boldsymbol{h}, u) = \mathbb{E} \left\{ \cos \left(\sqrt{2R} \langle \boldsymbol{\Omega}, \boldsymbol{h} \rangle + Y \sqrt{\gamma(u)} |\boldsymbol{\Omega}| \right) \right\}, \tag{5}$$

where $\mathbb{E}\{\cdot\}$ the mathematical expectation, R a nonnegative random variable with distribution μ, Y a standard normal random variable, $\boldsymbol{\Omega}$ a k-dimensional standardized Gaussian random vector, and where $R, Y, \boldsymbol{\Omega}$ are independent.

Proof One uses (3) to write the squared exponential function in (4) as a Fourier transform on $\mathbb{R}$:

$$C(\boldsymbol{h}, u) = \int_{\mathbb{R}} \int_{\mathbb{R}^k} \int_{\mathbb{R}} \cos(\sqrt{2r} \langle \boldsymbol{\omega}, \boldsymbol{h} \rangle) \cos(y \sqrt{\gamma(u)} |\boldsymbol{\omega}|) \, g_1(y) \, g_{I_k}(\boldsymbol{\omega}) \, dy \, d\boldsymbol{\omega} \, \mu(dr),$$

where g_1 is the standard normal univariate probability density function. Owing to the parity of this function and to the product-to-sum trigonometric identity, this expression simplifies into

$$C(\boldsymbol{h}, u) = \int_{\mathbb{R}} \int_{\mathbb{R}^k} \int_{\mathbb{R}} \cos(\sqrt{2r} \langle \boldsymbol{\omega}, \boldsymbol{h} \rangle + y \sqrt{\gamma(u)} |\boldsymbol{\omega}|) \, g_1(y) \, g_{I_k}(\boldsymbol{\omega}) \, dy \, d\boldsymbol{\omega} \, \mu(dr),$$

which yields the claim.

3 A Discrete-in-Time and Continuous-in-Space Substitution Algorithm

Consider a space-time cosine wave of the following form:

$$Z(\boldsymbol{x}, t) = \sqrt{2} \cos \left(\sqrt{2R} \langle \boldsymbol{\Omega}, \boldsymbol{x} \rangle + W(t) \frac{|\boldsymbol{\Omega}|}{\sqrt{2}} + \Phi \right), \qquad \boldsymbol{x} \in \mathbb{R}^k, t \in \mathbb{R}, \tag{6}$$

where

- R and $\boldsymbol{\Omega}$ are a random variable and a random vector as defined in (5);
- $\{W(t) : t \in \mathbb{R}\}$ is a strictly intrinsic random field with variogram γ and Gaussian increments;

- Φ is a uniform random variable on $]0, 2\pi[$;
- R, $\boldsymbol{\Omega}$, W and Φ are independent.

Because Φ is uniform on $]0, 2\pi[$ and is independent of $(R, \boldsymbol{\Omega}, W)$, this cosine wave is centered. Moreover, the covariance between $Z(\boldsymbol{x} + \boldsymbol{h}, t + u)$ and $Z(\boldsymbol{x}, t)$ is found to be equal to the expectation in (5), that is, $C(\boldsymbol{h}, u)$ (see [3]). The random field in (6) is a particular case of substitution random field, consisting of the composition of a stationary coding process on $\mathbb{R}$ and an intrinsic directing function on $\mathbb{R}^k \times \mathbb{R}$ (see [4]).

To obtain an approximately Gaussian random field with zero mean and Gneiting covariance, one can (i) multiply the cosine by a Rayleigh random variable with scale parameter $2^{-1/2}$, which makes the marginal distribution of $Z(\boldsymbol{x}, t)$ be standard Gaussian, and (ii) sum and standardize many of such independent cosine waves, so that the finite-dimensional distributions of $Z(\boldsymbol{x}, t)$ become approximately multivariate Gaussian due to the central limit theorem. The simulated random field thus takes the form:

$$Z(\boldsymbol{x}, t) = \sum_{j=1}^{p} \sqrt{\frac{-2 \ln U_j}{p}} \cos\left(\sqrt{2R_j} \langle \boldsymbol{\Omega}_j, \boldsymbol{x} \rangle + W_j(t) \frac{|\boldsymbol{\Omega}_j|}{\sqrt{2}} + \Phi_j\right), \quad (7)$$

where p is a large integer, $\{(R_j, \boldsymbol{\Omega}_j, W_j, \Phi_j) : j = 1, ..., p\}$ are independent copies of $(R, \boldsymbol{\Omega}, W, \Phi)$, and $\{U_j : j = 1, ..., p\}$ are independent random variables uniformly distributed on $]0, 1[$ and are independent of $\{(R_j, \boldsymbol{\Omega}_j, W_j, \Phi_j) : j = 1, ..., p\}$.

As an illustration, consider the simulation of a random field on a regular grid of $\mathbb{R}^1 \times \mathbb{R}$ with 500×500 nodes and mesh 1×0.2, with the following parameters:

- $k = 1$;
- $\varphi(r) = \exp(-0.001r)$
- $\gamma(u) = \sqrt{1 + |u|} - 1$;
- $p = 10, 100$ or 1000.

The intrinsic random field W is simulated by using the covariance matrix decomposition algorithm with the nonstationary covariance function $(t, t') \mapsto \gamma(t) + \gamma(t') - \gamma(t' - t)$. The simulation obtained with $p = 10$ cosine waves exhibits an apparent periodicity in space, which indicates that the central limit approximation is poor, which is no longer the case when using 100 or more cosine waves (Fig. 1, left). In contrast, since the simulated process W is ergodic, the time variations are well reproduced, irrespective of the number of cosine waves. Note that the simulated random field has a Gaussian spatial covariance and a gamma temporal covariance with parameter 0.5, hence it is smooth in space but not in time.
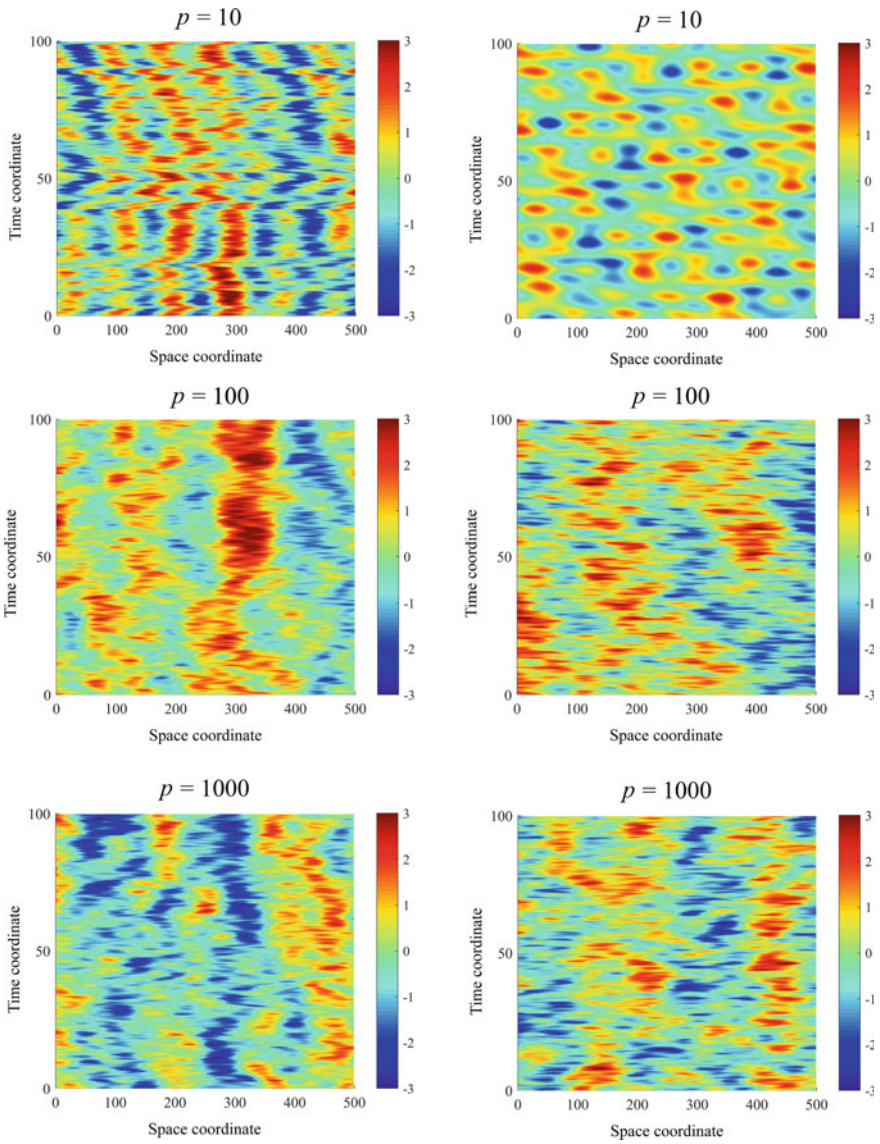
Fig. 1 Simulation of a random field with Gneiting covariance function (with an exponential function for φ and a square root for γ) using 10 (top), 100 (center) and 1000 (bottom) cosine waves. Left: substitution algorithm. Right: spectral algorithm

4   A Fully Continuous Spectral Algorithm

The continuous spectral method relies on the fact that the continuous covariance C is the Fourier transform of a symmetric probability measure F on $\mathbb{R}^k \times \mathbb{R}$. Let $(\boldsymbol{\Omega}, \mathrm{T})$ be a spectral vector distributed according to F, and Φ a random phase uniform on $]0, 2\pi[$ and independent of $(\boldsymbol{\Omega}, \mathrm{T})$. Then, the random field defined by

$$Z(\boldsymbol{x}, t) = \sqrt{2} \cos\big(\langle \boldsymbol{\Omega}, \boldsymbol{x} \rangle + \mathrm{T}t + \Phi\big), \qquad (\boldsymbol{x}, t) \in \mathbb{R}^k \times \mathbb{R}, \qquad (8)$$

is second-order stationary with covariance C. A standard approach for simulating the spectral vector is to simulate at first the spatial component $\boldsymbol{\Omega}$, then the temporal component T given $\boldsymbol{\Omega}$, which requires explicitly knowing the spectral measure. For instance, consider the covariance function with $\varphi(r) = \exp(-ar)$ $(a > 0)$ and $\gamma(u) = \sqrt{1 + |u|} - 1$. In this case, it can be shown (see [3]) that $\boldsymbol{\Omega}$ is a Gaussian random vector with independent components and that T given $\boldsymbol{\Omega} = \omega$ follows a Cauchy distribution whose scale parameter follows an inverse Gaussian distribution, all these distributions being simulatable.

In practice, the multiplication of the cosine wave (8) by a Rayleigh random variable and the independent replication technique of (7) provide a random field with Gaussian marginal distributions and approximately multivariate Gaussian finite-dimensional distributions. Figure 1 (right) displays realizations obtained by using between $p = 10$ and $p = 1000$ cosine waves. The spatial variations are similar to those observed with the substitution algorithm. However, the temporal variations differ when using few cosine waves, exhibiting a smooth and periodic behavior when $p = 10$ or $p = 100$. This suggests that the convergence to a multivariate-Gaussian distribution in time is slower with the spectral algorithm than with the substitution algorithm.

5   Concluding Remarks

The two presented approaches construct the simulated random field as a weighted sum of cosine waves with random frequencies and phases. Their main difference lies in the way to simulate the temporal frequency: from its distribution conditional to the spatial frequency (spectral approach), or from an intrinsic time-dependent random field (substitution approach). Both algorithms have a computational complexity in $\mathcal{O}(n)$, considerably cheaper than generic algorithms such as the covariance matrix decomposition and sequential algorithms, are parallelizable and require minimal memory storage space, which makes them affordable for large-scale problems. The substitution approach is general and only requires the knowledge of the temporal variogram γ and the probability measure μ specifying the completely monotone function φ (2) associated with the spatial structure. In contrast, the spectral approach is more specific, as it also requires the knowledge of the time frequency distribution conditional to the spatial frequency, which has to be solved on a case-by-case basis.

References

1. Gneiting, T.: Nonseparable, stationary covariance functions for space-time data. J. Am. Stat. Assoc. **97**(458), 590–600 (2002)
2. Zastavnyi, V.P., Porcu, E.: Characterization theorems for the Gneiting class of space-time covariances. Bernoulli **17**(1), 456–465 (2011)
3. Allard, D., Emery, X., Lacaux, C., Lantuéjoul, C.: Simulating space-time random fields with nonseparable Gneiting-type covariance functions. Stat. Comput. **30**(5), 1479–1495 (2020)
4. Lantuéjoul, C.: Geostatistical Simulation: Models and Algorithms. Springer, Berlin (2002)