Non-Stationary Covariance Functions for Spatial Data on Linear Networks

Alfredo Alegría

Department of Statistics, Pontificia Universidad Católica de Chile, Santiago, Chile

Supplementary Material

S1 Stochastic representation

The following theorem provides an explicit construction of an elementary random field whose covariance function is given by (3.3). Throughout this section, we assume without loss of generality that F is a probability measure on $(0, \infty)$.

Theorem S1. *Let Z be a real-valued random variable with probability density function g that is strictly positive on $\mathbb{R}$. Let W be a positive random variable with distribution F. Let ϵ be a random variable with zero mean and unit variance. Let $\{A(s) : s \in \mathcal{L}\}$ denote the auxiliary GRF defined in Equation (2.1). Assume that Z, W, ϵ, and A are mutually independent. Let*

$\phi(\cdot;\mu,\sigma^2)$ denote the probability density function of a normal distribution with mean μ and variance σ^2. Then, the random field

$$X(s) = \epsilon \sqrt{\frac{b(s)}{g(Z)}} \left(\frac{\pi}{W}\right)^{1/4} \phi\left(Z; \frac{A(s)}{\sqrt{2}}, \frac{a(s)}{4W}\right), \qquad s \in \mathcal{L}, \tag{S1.1}$$

has zero mean and covariance function given by (3.3).

Proof. The proof consists of three steps.

Step 1: zero mean. Since $\mathrm{E}[\epsilon] = 0$ and ϵ is independent of the remaining ingredients, the expectation of the product factorizes and the random field has zero mean.

Step 2: computation of the covariance. By independence and noting that $\mathrm{E}\big[\epsilon^2\big] = 1$, the covariance function is given by

$$\begin{aligned}
\operatorname{cov}\big[X(s), X(t)\big] &= \mathrm{E}\big[\epsilon^2\big]\, \mathrm{E}_{Z,W,A}\left[\frac{\sqrt{b(s)b(t)}}{g(Z)} \left(\frac{\pi}{W}\right)^{1/2} \phi\left(Z; \frac{A(s)}{\sqrt{2}}, \frac{a(s)}{4W}\right) \phi\left(Z; \frac{A(t)}{\sqrt{2}}, \frac{a(t)}{4W}\right)\right] \\
&= \mathrm{E}_A\left[\int_0^\infty \int_{\mathbb{R}} \frac{\sqrt{b(s)b(t)}}{g(z)} \left(\frac{\pi}{w}\right)^{1/2} \phi\left(z; \frac{A(s)}{\sqrt{2}}, \frac{a(s)}{4w}\right) \phi\left(z; \frac{A(t)}{\sqrt{2}}, \frac{a(t)}{4w}\right) g(z)\, \mathrm{d}z\, \mathrm{d}F(w)\right] \\
&= \mathrm{E}_A\left[\int_0^\infty \int_{\mathbb{R}} \sqrt{b(s)b(t)} \left(\frac{\pi}{w}\right)^{1/2} \phi\left(z; \frac{A(s)}{\sqrt{2}}, \frac{a(s)}{4w}\right) \phi\left(z; \frac{A(t)}{\sqrt{2}}, \frac{a(t)}{4w}\right) \mathrm{d}z\, \mathrm{d}F(w)\right]
\end{aligned} \tag{S1.2}$$

where the second equality follows by conditioning on the auxiliary field and integrating out Z and W, whose distributions are g and F, respectively, and the third one because the density g introduced by the integration over Z cancels the factor $1/g(z)$.

Step 3: identification. The right hand side of (S1.2) coincides with the representation (A.5) obtained in the proof of Theorem 1. Therefore, we recover the covariance function (3.3). □

The stochastic representation introduced in Equation (S1.1) extends Proposition 3.4 of Alegría et al. (2026) from the isotropic setting to the non-stationary case.

The importance of this representation lies in the simulation of GRFs with covariance function (3.3), as an alternative to the Cholesky decomposition of the covariance matrix, whose computational complexity is cubic (see, e.g., Lantuéjoul, 2001). Since the random field in Theorem S1 is not Gaussian, one strategy is to add several independent copies of this elementary random field, rescale appropriately, and then invoke a central limit approximation. More precisely, an approximately Gaussian random field with covariance function (3.3) is given by

$$X(s) = \frac{1}{\sqrt{J}} \sum_{j=1}^{J} \epsilon_j \sqrt{\frac{b(s)}{g(Z_j)}} \left(\frac{\pi}{W_j}\right)^{1/4} \phi\left(Z_j; \frac{A^{(j)}(s)}{\sqrt{2}}, \frac{a(s)}{4W_j}\right), \qquad s \in \mathcal{L},$$

where J is a large positive integer, and $\epsilon_1, \ldots, \epsilon_J$, $Z_1, \ldots, Z_J$, $W_1, \ldots, W_J$, and $A^{(1)}, \ldots, A^{(J)}$ are independent copies of the random ingredients involved in Theorem S1.

We refer the reader to (Alegría et al., 2026, Section 4.5) for a more detailed treatment of the computational aspects of these types of algorithms in the linear network scenario, as well as for implementation guidelines.

S2 Vector-valued GRFs

In multivariate spatial statistics, several georeferenced variables are analyzed simultaneously (Wackernagel, 2003). Interest focuses not only on the marginal covariance functions of each variable but also on the cross-covariance functions between them.

Formally, consider a vector-valued GRF, namely

$$\{\boldsymbol{X}(s) = (X_1(s), \ldots, X_q(s))^\top : s \in \mathcal{L}\},$$

where q is the number of components, and $\text{var}[X_i(s)] < \infty$, for all $s \in \mathcal{L}$ and $i = 1, \ldots, q$. In this setting, the covariance function is matrix-valued, $\boldsymbol{K} : \mathcal{L} \times \mathcal{L} \to \mathbb{R}^{q \times q}$, with entries

$$K_{ij}(s,t) = \text{cov}\big[X_i(s), X_j(t)\big], \qquad s,t \in \mathcal{L},\ i,j = 1, \ldots, q.$$

The covariance function must be positive semidefinite in the sense that, for each $N \in \mathbb{N}$ and any collection of locations $s_1, \ldots, s_N \in \mathcal{L}$, the covariance matrix of the Nq-dimensional vector $(\boldsymbol{X}(s_1)^\top, \ldots, \boldsymbol{X}(s_N)^\top)^\top$ is positive semidefinite.

The following theorem provides a construction of non-stationary matrix-valued covariance functions, extending Theorem 1 to the vector-valued setting.

Theorem S2. *Let $a_i, b_i : \mathcal{L} \to (0, \infty)$, for $i = 1, \ldots, q$, and define $\alpha_{ij}(s,t) = [a_i(s) + a_j(t)]/2$. Let $\boldsymbol{\beta} = [\beta_{ij}]_{i,j=1}^q$ be a symmetric positive semidefinite matrix. Let $d_R(s,t)$ denote the resistance metric between s and t, and let F be a positive finite measure on $(0, \infty)$. Then, the matrix-valued function $\boldsymbol{K} : \mathcal{L} \times \mathcal{L} \to \mathbb{R}^{q \times q}$, with entries*

$$K_{ij}(s,t) = \int_0^\infty \frac{\beta_{ij}\sqrt{b_i(s)b_j(t)}}{\sqrt{\alpha_{ij}(s,t) + d_R(s,t)w}} dF(w), \qquad s,t \in \mathcal{L},\ i,j = 1,\ldots,q, \tag{S2.1}$$

is positive semidefinite.

Proof. We first consider the simplified case where $\beta_{ij} = 1$ for all $i, j = 1, \ldots, q$. In this case, the function in Equation (S2.1) reduces to $\tilde{K}_{ij}$ defined by

$$\tilde{K}_{ij}(s,t) = \int_0^\infty \frac{\sqrt{b_i(s)b_j(t)}}{\sqrt{\alpha_{ij}(s,t) + \mathrm{d}_R(s,t)w}} \mathrm{d}F(w).$$

Using the same arguments as in the proof of Theorem 1, we obtain

$$\tilde{K}_{ij}(s,t) = \sqrt{\pi}\,\mathrm{E}\bigg[\int_0^\infty \int_{\mathbb{R}} \sqrt{b_i(s)b_j(t)}\,\phi\left(z; \frac{A(s)}{\sqrt{2}}, \frac{a_i(s)}{4w}\right) \phi\left(z; \frac{A(t)}{\sqrt{2}}, \frac{a_j(t)}{4w}\right) \mathrm{d}z\, w^{-1/2}\, \mathrm{d}F(w)\bigg]. \tag{S2.2}$$

Note that this is a multivariate version of Equation (A.5). Let $N \in \mathbb{N}$,

$\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N \in \mathbb{R}^q$, and $s_1, \ldots, s_N \in \mathcal{L}$, where $\boldsymbol{r}_k = (r_{k1}, \ldots, r_{kq})^\top$. Then,

$$\sum_{k=1}^{N}\sum_{l=1}^{N} \boldsymbol{r}_k^\top \tilde{\boldsymbol{K}}(s_k, s_l)\boldsymbol{r}_l = \sum_{k=1}^{N}\sum_{l=1}^{N}\sum_{i=1}^{q}\sum_{j=1}^{q} r_{ki} r_{lj} \tilde{K}_{ij}(s_k, s_l)$$

$$= \sqrt{\pi}\,\mathrm{E}\left[\int_0^\infty \int_{\mathbb{R}} \left\{\sum_{k=1}^{N}\sum_{i=1}^{q} r_{ki}\sqrt{b_i(s_k)}\,\phi\left(z; \frac{A(s_k)}{\sqrt{2}}, \frac{a_i(s_k)}{4w}\right)\right\}^2 \mathrm{d}z\, w^{-1/2}\,\mathrm{d}F(w)\right] \geq 0.$$

In conclusion, the $Nq \times Nq$ matrix

$$\begin{pmatrix} \tilde{\boldsymbol{K}}(s_1, s_1) & \cdots & \tilde{\boldsymbol{K}}(s_1, s_N) \\ \vdots & \ddots & \vdots \\ \tilde{\boldsymbol{K}}(s_N, s_1) & \cdots & \tilde{\boldsymbol{K}}(s_N, s_N) \end{pmatrix} \tag{S2.3}$$

is positive semidefinite, as expected.

For the general case, where the parameters β_{ij} are not necessarily identically equal to one, we observe that

$$\begin{pmatrix} \boldsymbol{K}(s_1, s_1) & \cdots & \boldsymbol{K}(s_1, s_N) \\ \vdots & \ddots & \vdots \\ \boldsymbol{K}(s_N, s_1) & \cdots & \boldsymbol{K}(s_N, s_N) \end{pmatrix} = \left[\mathbf{1}_N \mathbf{1}_N^\top \otimes \boldsymbol{\beta}\right] \circ \begin{pmatrix} \tilde{\boldsymbol{K}}(s_1, s_1) & \cdots & \tilde{\boldsymbol{K}}(s_1, s_N) \\ \vdots & \ddots & \vdots \\ \tilde{\boldsymbol{K}}(s_N, s_1) & \cdots & \tilde{\boldsymbol{K}}(s_N, s_N) \end{pmatrix}, \tag{S2.4}$$

where $\circ$ and $\otimes$ denote the Hadamard and Kronecker products, respectively, and $\mathbf{1}_N$ is an N-dimensional vector of ones.

Since $\boldsymbol{\beta}$, $\mathbf{1}_N \mathbf{1}_N^\top$, and the matrix in (S2.3) are positive semidefinite, and positive semidefiniteness is preserved under both Kronecker and Hadamard products, it follows that (S2.4) is positive semidefinite. □

This theorem is the network counterpart of a result presented in Kleiber and Nychka (2012) for Euclidean spaces.

In this construction, the functions b_i and a_i allow us to control, respectively, the local variance and correlation range of the i-th component of the vector-valued GRF, in a similar fashion to the univariate ($q = 1$) case. The parameters β_{ij}, in turn, allow us to modulate the colocated correlation coefficient between the i-th and j-th components. More precisely, the effect of β_{ij} on the colocated correlation is given by

$$\operatorname{cor}[X_i(s), X_j(s)] = \frac{\operatorname{cov}[X_i(s), X_j(s)]}{\sqrt{\operatorname{var}[X_i(s)]\operatorname{var}[X_j(s)]}} = \frac{\beta_{ij}}{\sqrt{\beta_{ii}\beta_{jj}}} \frac{[a_i(s)a_j(s)]^{1/4}}{\sqrt{\alpha_{ij}(s,s)}},$$

for each $s \in \mathcal{L}$ and $i \neq j$.

Using the representation in terms of generalized Stieltjes functions of order 1/2, as described in Section 3.2, the function (S2.1) can be expressed as

$$K_{ij}(s,t) = \begin{cases} \beta_{ij}\sqrt{\dfrac{b_i(s)b_j(s)}{\alpha_{ij}(s,s)}}F((0,\infty)), & \text{if } s = t, \\ \\ \beta_{ij}\sqrt{\dfrac{b_i(s)b_j(t)}{\mathrm{d}_R(s,t)}}\psi\left(\dfrac{\alpha_{ij}(s,t)}{\mathrm{d}_R(s,t)}\right), & \text{if } s \neq t, \end{cases}$$

with ψ defined according to Equation (3.4). Thus, analogous examples to those presented in Section 3.2 can be constructed for the vector-valued case.

In particular, when a_i and b_i are constant functions for all $i = 1, \ldots, q$,

Equation (S2.1) defines a family of matrix-valued isotropic covariance functions. Although this is a special case, it provides a useful multivariate extension of a subclass of the isotropic covariance models introduced by Anderes et al. (2020). It also complements previous results on vector-valued random fields on linear networks by Filosi et al. (2025), where isotropic matrix-valued covariance functions on trees were derived.

S3 Sensitivity to the bandwidth

We complement the study of Section 4.2 with a brief analysis of the sensitivity of the estimates to the bandwidth of the compactly supported weights. We work on the same network and with the same random field, and we estimate a at two of the six sites of Figure 4, namely site 2, which lies near the boundary of the network, and site 5, which lies in its interior. We take τ from 195 to 220 in steps of five, and for each of 100 independent realizations we estimate a at both sites with every bandwidth, so that the comparison across bandwidths is paired. The accuracy of the estimates is summarized through the median of $|\hat{a} - a|$ over the realizations, a measure that reflects both the location and the dispersion of the estimates and is not affected by a small proportion of extreme values.

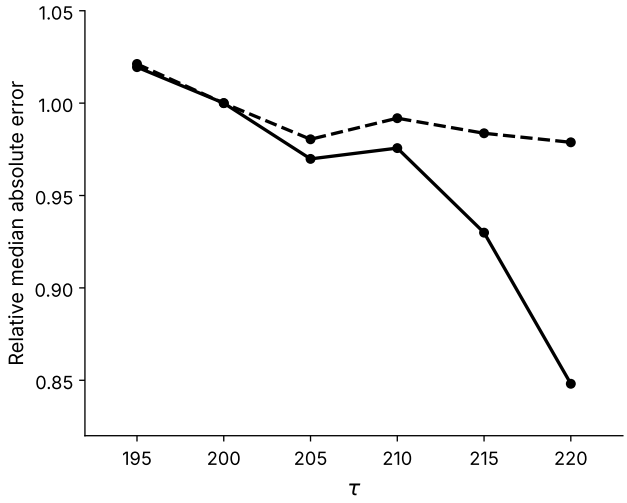


Figure S1: Median absolute error of the estimates of a over 100 realizations, as a function of the bandwidth τ of the compactly supported weights, divided by its value at $\tau = 200$. The solid line corresponds to site 2, which lies near the boundary of the network, and the dashed line to site 5, which lies in its interior.

Figure S1 shows this quantity for each site, divided by its value at $\tau = 200$, which is the bandwidth used in Section 4.2, so that each curve gives the relative change in accuracy that would result from a different choice.

Over this range of bandwidths the two curves remain close to one. At site 5 the departure from the value at $\tau = 200$ never exceeds 2%, and at site 2 it reaches 15% at the largest bandwidth. The sensitivity is larger at the site near the boundary, where fewer observations contribute to the

estimate, which is consistent with the reduced precision reported at site 2 in Section 4.2.

References

Alegría, A., Emery, X., Filosi, T., and Porcu, E. (2026). Computationally efficient algorithms for simulating isotropic Gaussian random fields on graphs with Euclidean edges. *Journal of Computational and Graphical Statistics*, 35(2):937–950.

Anderes, E., Møller, J., and Rasmussen, J. G. (2020). Isotropic covariance functions on graphs and their edges. *The Annals of Statistics*, 48(4):2478–2503.

Filosi, T., Porcu, E., Emery, X., Agostinelli, C., and Alegría, A. (2025). Vector-valued Gaussian processes and their kernels on a class of metric graphs. *arXiv preprint arXiv:2501.10208*.

Kleiber, W. and Nychka, D. (2012). Nonstationary modeling for multivariate spatial processes. *Journal of Multivariate Analysis*, 112:76–91.

Lantuéjoul, C. (2001). *Geostatistical simulation: models and algorithms*. Number 1139. Springer Science & Business Media.

Wackernagel, H. (2003). *Multivariate geostatistics: an introduction with applications*. Springer Science & Business Media.