Modeling and simulating spatio-temporal multivariate and non-stationary Gaussian Processes: a Gaussian mixtures perspective

Denis Allard, with

Lionel Benoit, Lucia Clarotto, Xavier Emery, Céline Lacaux, Christian Lantuéjoul and Said Obakrim

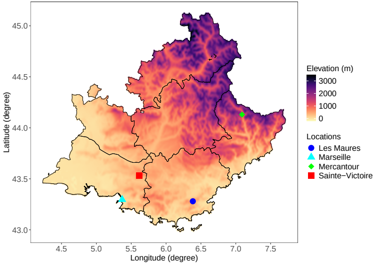
Biostatistique et processus Spatiaux (BioSP), MathNum, INRAE
Avignon, France

Funded by the Geolearning Chair

SWGEN 2025, Grenoble

A multivariate spatio-temporal SWG

- Talk by Saïd Obakrim, after the coffee break
- Region of interest: PACA, highly non-stationary
- 6 daily variables: precipitation, humidity, radiation, wind, min and max temperature
- SAFRAN reanalysis data (8 km $\times$ 8 km), from 2012 to 2021
- 498 pixels $\times$ 6 variables $\times$ 3652 days $\simeq$ 7.5 M data



Obakrim S., Benoit L. & Allard D. (2025) A multivariate and space-time stochastic weather generator using a latent Gaussian framework. *Stochastic Environmental Research and Risk Assessment*. doi.org/10.1007/s00477-024-02897-8

Motivation

We need

1. covariance functions in complex settings: spatio-temporal, multivariate, nonstationary; **sometimes all at once**
2. Together with simulation techniques for very large number of (spatial $\times$ temporal) sites
3. Not necessarily on grids

State of the art:

- Cholesky decomposition is limited to $N < 10^4$
- Sparse approximations, low rank approximations $\rightarrow$ impact on the covariance function; rarely avaliable in complex settings
- Circulant embedding methods using FFT $\rightarrow$ limited to stationary covariances and simulations on regular grids

$\hookrightarrow$ There is a need for more versatile methods

Motivation

The way forward

- Start from the *spectral simulation method*
- Revisit this method with a **Gaussian mixture perspective**
- These simulation algorithms are constructive arguments for defining **new classes of covariance functions** in these complex settings

Allard, D., Benoit, L., & Obakrim, S. (2025). Modeling and simulating spatio-temporal, multivariate and nonstationary Gaussian Random Fields: a Gaussian mixtures perspective. *Preptint* https://hal.inrae.fr/hal-05034982

Outline

1. **Introduction:** reminders on the spectral method and their extensions
2. **Focus on Gaussian mixtures**
3. **Nonstationarity:** a general result generalizing the Paciorek-Shervish construction
4. **The full combo:** new nonstationary, multivariate, spatio-temporal Gaussian Random Fields (GRFs)

Outline

The "vanilla" spectral method

Shinozuka (1971), Matheron (1973)

Use Bochner Theorem,

$$C(\mathbf{h}) = \int_{\mathbb{R}^d} \exp(i\mathbf{h}^t \boldsymbol{\omega}) d\mu(\boldsymbol{\omega}), \quad \forall \mathbf{h} \in \mathbb{R}^d,$$

or

$$C(\mathbf{h}) = \int_{\mathbb{R}^d} \cos(\mathbf{h}^t \boldsymbol{\omega}) d\mu(\boldsymbol{\omega}), \quad \forall \mathbf{h} \in \mathbb{R}^d.$$

Then,

$$\tilde{Z}_L(\mathbf{s}) = \sqrt{\frac{2}{L}} \sum_{l=1}^{L} \cos\left(\boldsymbol{\Omega}_l^t \mathbf{s} + \Phi_l\right), \quad \boldsymbol{\Omega}_l \sim \mu, \quad \Phi_l \sim \mathcal{U}(0, 2\pi), \quad \text{all i.i.d}$$

is approximately a GRF with expectation 0 and covariance function C

The "vanilla" spectral method

Shinozuka (1971), Matheron (1973)

Use Bochner Theorem,

$$C(\mathbf{h}) = \int_{\mathbb{R}^d} \exp(i\mathbf{h}^t\boldsymbol{\omega})d\mu(\boldsymbol{\omega}), \quad \forall \mathbf{h} \in \mathbb{R}^d,$$

or

$$C(\mathbf{h}) = \int_{\mathbb{R}^d} \cos(\mathbf{h}^t\boldsymbol{\omega})d\mu(\boldsymbol{\omega}), \quad \forall \mathbf{h} \in \mathbb{R}^d.$$

Then,

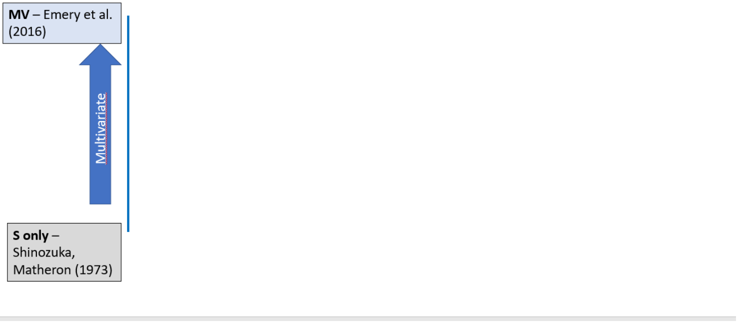
$$\tilde{Z}_L(\mathbf{s}) = \sqrt{\frac{2}{L}} \sum_{l=1}^{L} \cos\left(\boldsymbol{\Omega}_l^t\mathbf{s} + \Phi_l\right), \quad \boldsymbol{\Omega}_l \sim \mu, \quad \Phi_l \sim \mathcal{U}(0, 2\pi), \quad \text{all i.i.d}$$

is approximately a GRF with expectation 0 and covariance function C

Extensions of the spectral method

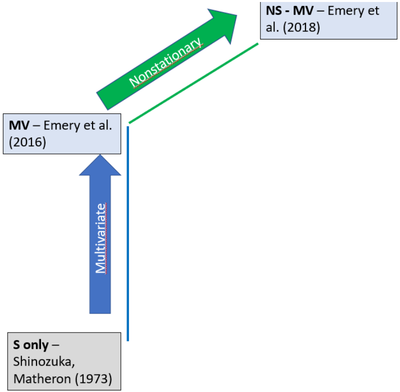
S only –
Shinozuka,
Matheron (1973)

Extensions of the spectral method



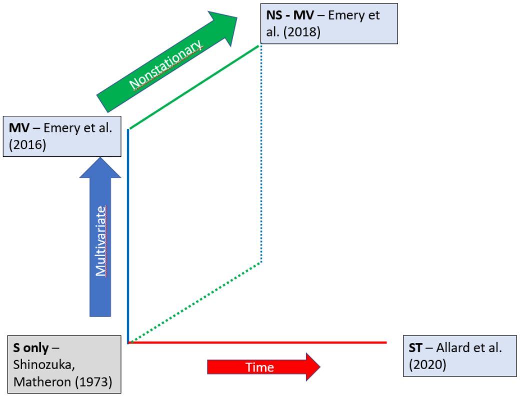
See Emery et al. (2016)

Extensions of the spectral method



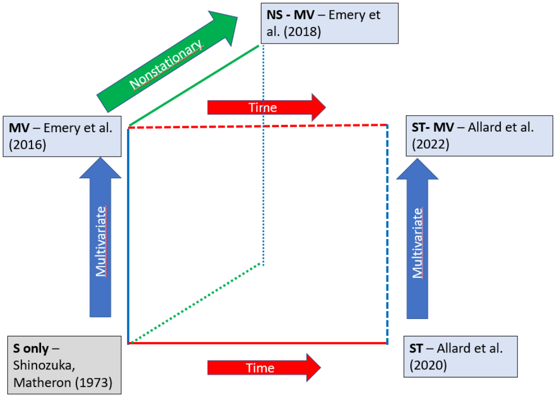
See Emery et al. (2016) and Emery and Arroyo (2018)

Extensions of the spectral method



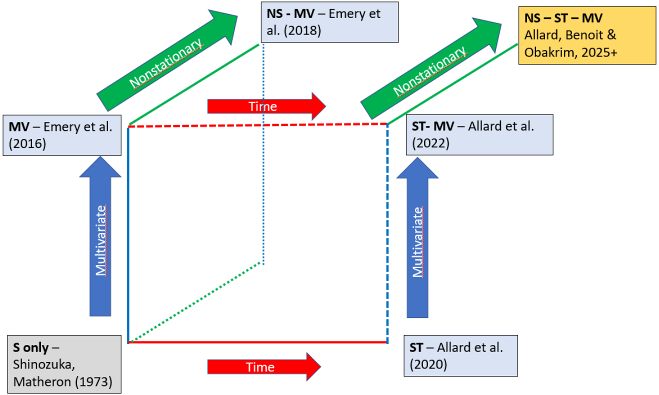
See Allard et al. (2020)

Extensions of the spectral method



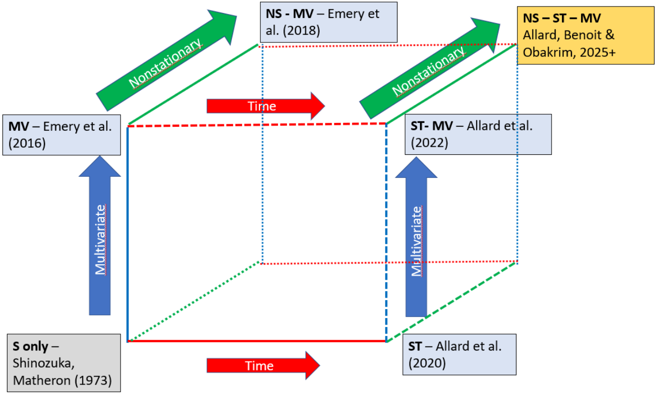
See Allard et al. (2020) and Allard et al. (2022)

Extensions of the spectral method



This work

Extensions of the spectral method



This work

Outline

Gaussian mixtures

Schoenberg (1938)

Define $\mathcal{C}_\infty$ the class of continuous isotropic covariance functions valid on $\mathbb{R}^d$, $\forall d \geq 1$. Then, $\phi \in \mathcal{C}_\infty$ if and only if

$$\phi(\mathbf{h}) = \int_{\mathbb{R}^+} \exp(-||\mathbf{h}||^2 \xi) f(\xi) d\xi$$

$f(\xi)$ is the **Gaussian scale mixture**

Consequences

- A Gaussian mixture for the covariance function entails the same Gaussian mixture of the spectral density.
- Use Gaussian mixtures in spectral simulations.

Simulation algorithms for stationary univariate spatial GRFs

`Spectral simulation`

Require: $C \in \mathcal{C}_\infty$, $\mathbf{\Sigma}^{-1/2}$ and μ
Require: A set of points, $\mathcal{S} \in \mathbb{R}^d$
Require: A large number L

for $l = 1$ to L **do**
    Simulate $\mathbf{\Omega}_l \sim \mu$
    Simulate $\Phi_l \sim \mathcal{U}(0, 2\pi)$
end for
For each $\mathbf{s} \in \mathcal{S}$ return

$$\tilde{Z}(\mathbf{s}) = \sqrt{\frac{2}{L}} \sum_{l=1}^{L} \cos\left(\mathbf{\Sigma}^{-1/2} \mathbf{\Omega}_l^t \mathbf{s} + \Phi_l\right)$$

`Gaussian mixture simulation`

Require: $C \in \mathcal{C}_\infty$ $\mathbf{\Sigma}^{-1/2}$ and f
Require: A set of points, $\mathcal{S} \in \mathbb{R}^d$
Require: A large number L

for $l = 1$ to L **do**
    Simulate $\xi_l \sim f$
    Simulate $\mathbf{\Omega}_l \sim \sqrt{2\xi_l} \mathcal{N}_d(0, \mathbf{I}_d)$
    Simulate $\Phi_l \sim \mathcal{U}(0, 2\pi)$
end for
For each $\mathbf{s} \in \mathcal{S}$ return

$$\tilde{Z}(\mathbf{s}) = \sqrt{\frac{2}{L}} \sum_{l=1}^{L} \cos\left(\mathbf{\Sigma}^{-1/2} \mathbf{\Omega}_l^t \mathbf{s} + \Phi_l\right)$$

Simulation algorithms for stationary univariate spatial GRFs

`Spectral simulation`

Require: $C \in \mathcal{C}_\infty$, $\mathbf{\Sigma}^{-1/2}$ and μ
Require: A set of points, $\mathcal{S} \in \mathbb{R}^d$
Require: A large number L
1: **for** $l = 1$ to L **do**
2:     **Simulate** $\mathbf{\Omega}_l \sim \mu$
3:     Simulate $\Phi_l \sim \mathcal{U}(0, 2\pi)$
4: **end for**
5: For each $\mathbf{s} \in \mathcal{S}$ return

$$\tilde{Z}(\mathbf{s}) = \sqrt{\frac{2}{L}} \sum_{l=1}^{L} \cos\left(\mathbf{\Sigma}^{-1/2}\mathbf{\Omega}_l^t \mathbf{s} + \Phi_l\right)$$

`Gaussian mixture simulation`

Require: $C \in \mathcal{C}_\infty$ $\mathbf{\Sigma}^{-1/2}$ and f
Require: A set of points, $\mathcal{S} \in \mathbb{R}^d$
Require: A large number L
1: **for** $l = 1$ to L **do**
2:     **Simulate** $\xi_l \sim f$
3:     **Simulate** $\mathbf{\Omega}_l \sim \sqrt{2\xi_l}\mathcal{N}_d(0, \mathbf{I}_d)$
4:     Simulate $\Phi_l \sim \mathcal{U}(0, 2\pi)$
5: **end for**
6: For each $\mathbf{s} \in \mathcal{S}$ return

$$\tilde{Z}(\mathbf{s}) = \sqrt{\frac{2}{L}} \sum_{l=1}^{L} \cos\left(\mathbf{\Sigma}^{-1/2}\mathbf{\Omega}_l^t \mathbf{s} + \Phi_l\right)$$

Some covariance functions

Matérn covariance

$$C_{\mathcal{M}}(\mathbf{h}) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)}(\kappa||\mathbf{h}||)^{\nu} K_{\nu}(\kappa||\mathbf{h}||)$$

$$\mu_{\mathcal{M}}(\boldsymbol{\omega}) \propto \frac{1}{(1+||\boldsymbol{\omega}||^2/\kappa^2)^{\nu+d/2}}$$

$$f_{\mathcal{M}}(\xi) = \left(\frac{\kappa^2}{4}\right)^{\nu} \frac{\xi^{-1-\nu}}{\Gamma(\nu)} e^{-\kappa^2/4\xi}.$$

Hence

$$\text{Step 2 : Simulate} \quad \xi_l \sim IG(\nu, \kappa^2/4)$$

Cauchy covariance

$$C_{\mathcal{C}}(\mathbf{h}) = \left(1 + a||\mathbf{h}||^2\right)^{-\nu}$$

$$\mu_{\mathcal{C}} = \text{Unknown}$$

$$f_{\mathcal{C}}(\xi) = a^{-\nu}\Gamma(\nu)^{-1}\xi^{\nu-1}e^{-\xi/a}$$

Hence

$$\text{Step 2 : Simulate} \quad \xi_l \sim G(\nu, a).$$

Some covariance functions

Matérn covariance

$$\begin{aligned} C_{\mathcal{M}}(\mathbf{h}) &= \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)}(\kappa||\mathbf{h}||)^{\nu} K_{\nu}(\kappa||\mathbf{h}||) \\ \mu_{\mathcal{M}}(\boldsymbol{\omega}) &\propto \frac{1}{(1+||\boldsymbol{\omega}||^2/\kappa^2)^{\nu+d/2}} \\ f_{\mathcal{M}}(\xi) &= \left(\frac{\kappa^2}{4}\right)^{\nu} \frac{\xi^{-1-\nu}}{\Gamma(\nu)} e^{-\kappa^2/4\xi}. \end{aligned}$$

Hence

$$\text{Step 2: Simulate} \quad \xi_l \sim IG(\nu, \kappa^2/4)$$

Cauchy covariance

$$\begin{aligned} C_{\mathcal{C}}(\mathbf{h}) &= \left(1 + a||\mathbf{h}||^2\right)^{-\nu} \\ \mu_{\mathcal{C}} &= \text{Unknown} \\ f_{\mathcal{C}}(\xi) &= a^{-\nu}\Gamma(\nu)^{-1}\xi^{\nu-1}e^{-\xi/a} \end{aligned}$$

Hence

$$\text{Step 2: Simulate} \quad \xi_l \sim G(\nu, a).$$

Main take aways

Use Gaussian mixtures

- Almost identical simulation algorithm
- Restricted to kernels in $\mathcal{C}_\infty$
- Paves the way to many extensions : temporal, multivariate, non-stationary

Outline

General result

Allard et al. (2025+)

- Let $\phi \in \mathcal{C}_\infty$, with a Gaussian mixture belonging to the exponential family of pdfs

$$f(\xi;\boldsymbol{\theta}) = h(\boldsymbol{\theta}) \exp\left(-\boldsymbol{\ell}(\boldsymbol{\theta})^t \boldsymbol{T}(\xi)\right) \tag{1}$$

  Includes Gamma (Cauchy cov.), Inverse Gamma (Matérn cov.), Beta, Gaussian, Inverse Gaussian, etc.
- Let $\boldsymbol{\Sigma}_{\mathbf{s}}^{-1/2}$ be anisotropy matrices, $\forall \mathbf{s} \in \mathbb{R}^d$
- Let $f(\cdot, \boldsymbol{\theta}_{\mathbf{s}})$ be a family of mixtures as in (1)
- Set f_1 be an **instrumental density**: any pdf whose support is $\mathbb{R}^+$. One can set $f_1 = f(\cdot, \boldsymbol{\theta} = \mathbf{1})$

General result

Allard et al. (2025+)

Proposition

Under the condition above, define:

$$Z(\mathbf{s}) = \sqrt{\frac{2f(\xi;\boldsymbol{\theta}_{\mathbf{s}})}{f_1(\xi)}}\sqrt{\frac{\mu^G_{\boldsymbol{\Sigma}_{\mathbf{s}}}(\boldsymbol{\Omega})}{\mu^G_{\boldsymbol{I}_d}(\boldsymbol{\Omega})}}\cos\left(\boldsymbol{\Omega}^t\mathbf{s}+\Phi\right).$$

Then, its non-stationary covariance function is

$$C^*(\mathbf{s},\mathbf{s}') = |\boldsymbol{\Sigma}_{\mathbf{s}}|^{1/4}|\boldsymbol{\Sigma}_{\mathbf{s}'}|^{1/4}|\boldsymbol{\Sigma}_{\mathbf{s},\mathbf{s}'}|^{-1/2}C(\boldsymbol{\Sigma}_{\mathbf{s},\mathbf{s}'}{}^{-1/2}(\mathbf{s}-\mathbf{s}');\boldsymbol{\theta}_{\mathbf{s},\mathbf{s}'}),$$

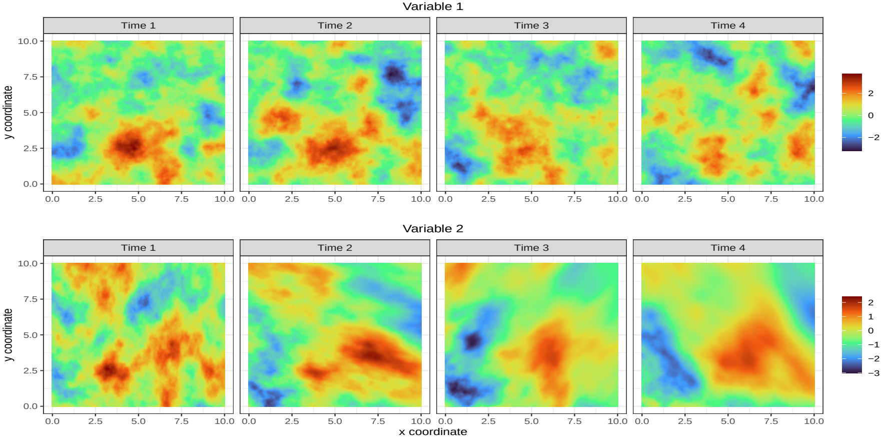
with $\boldsymbol{\Sigma}_{\mathbf{s},\mathbf{s}'} = (\boldsymbol{\Sigma}_{\mathbf{s}}+\boldsymbol{\Sigma}_{\mathbf{s}'})/2$, and where $\boldsymbol{\theta}_{\mathbf{s},\mathbf{s}'}$ is such that

$$\ell(\boldsymbol{\theta}_{\mathbf{s},\mathbf{s}'}) = \frac{\ell(\boldsymbol{\theta}_{\mathbf{s}})+\ell(\boldsymbol{\theta}_{\mathbf{s}'})}{2}$$

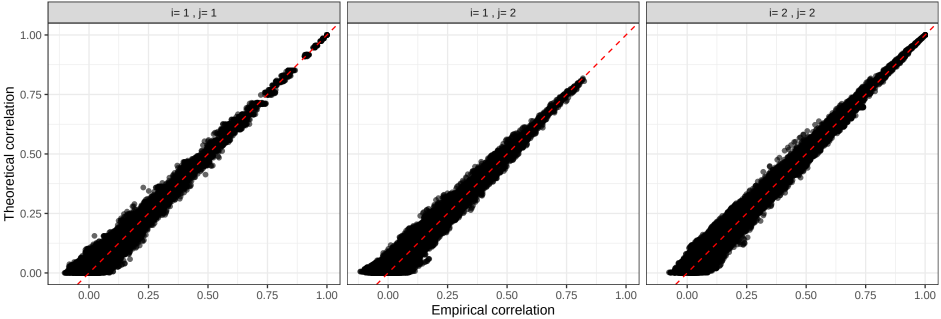
$\Rightarrow$ Generalizes the construction in Paciorek and Schervish (2006) and Emery and Arroyo (2018)

Outline

Illustration



Illustration



A simulation algorithm for NS MV S-T GRFs

Require: A family of **scale mixtures**, $f(\cdot;\boldsymbol{\theta})$, belonging to the exponential family
Require: Parameters $\boldsymbol{\theta}_{ii,\mathbf{x}}$ and anisotropy matrices $\boldsymbol{\Sigma}_{ii,\mathbf{x}}^{-1/2}$; covariance matrices $\boldsymbol{\sigma}_{\mathbf{x}} = \boldsymbol{L}_{\mathbf{x}}\boldsymbol{L}_{\mathbf{x}}^t$
Require: Pseudo variogram γ. Non separability parameter b ; $\delta > 0$

Set $f_1 := f(\theta)$, e.g. for $\theta = 1$
for $l = 1$ to L **do**
  Simulate a p-variate RF $\boldsymbol{Z}_{T,l}$ with matrix-valued covariance function $\boldsymbol{C}_T(t) = (1+\gamma(t))^{-\delta}$
  Simulate a p-variate RF $\boldsymbol{W}_l = [W_{l,i}]_{i=1}^p$ with pseudo-variogram γ_b
  Simulate $\xi_l \sim f_1$
  Simulate $\boldsymbol{V}_l \sim \mathcal{N}_d(0, \mathbf{I}_d)$; set $\boldsymbol{\Omega}_l = \sqrt{2\xi_l}\mathbf{V}_l$
  Simulate $\Phi_l \sim \mathcal{U}(0, 2\pi)$; Simulate $\boldsymbol{A}_l \sim \mathcal{N}_p(0, \boldsymbol{I}_p)$
end for
For each $\mathbf{x} = (\mathbf{s}, t) \in \mathcal{S}$, and for $i = 1, \dots, p$, return

$$\tilde{Z}_{L,i}(\mathbf{s},t) = \sqrt{\frac{2}{L}} \sum_{l=1}^{L} Z_{T,l,i}(t) \sqrt{\frac{f_{ii,\mathbf{x}}(\xi_l)}{f_1(\xi_l)}} \sqrt{\frac{\mu^G_{\boldsymbol{\Sigma}_{ii,\mathbf{x}}}(\sqrt{2}\mathbf{V}_l)}{\mu^G_{\mathbf{I}_d}(\sqrt{2}\mathbf{V}_l)}} (\boldsymbol{L}_{\mathbf{x}}\boldsymbol{A}_l)_i \cos\left(\boldsymbol{\Omega}_l^t\mathbf{s} + \Phi_l + \frac{\|\mathbf{V}_l\|}{\sqrt{2}} W_i(t)\right)$$

pointwise correlation
non-stationary importance weights
non sep. space-time

A simulation algorithm for NS MV S-T GRFs

Require: A family of **scale mixtures**, $f(\cdot;\boldsymbol{\theta})$, belonging to the exponential family
Require: Parameters $\boldsymbol{\theta}_{ii,\mathbf{x}}$ and anisotropy matrices $\boldsymbol{\Sigma}_{ii,\mathbf{x}}^{-1/2}$; covariance matrices $\boldsymbol{\sigma}_{\mathbf{x}} = \boldsymbol{L}_{\mathbf{x}}\boldsymbol{L}_{\mathbf{x}}^t$
Require: Pseudo variogram γ. Non separability parameter b ; $\delta > 0$

1: Set $f_1 := f(\boldsymbol{\theta})$, e.g. for $\boldsymbol{\theta} = \mathbf{1}$
2: **for** $l = 1$ to L **do**
3: Simulate a p-variate RF $\boldsymbol{Z}_{T,l}$ with matrix-valued covariance function $\boldsymbol{C}_T(t) = (1+\gamma(t))^{-\delta}$
4: Simulate a p-variate RF $\boldsymbol{W}_l = [W_{l,i}]_{i=1}^p$ with pseudo-variogram γ_b
5: Simulate $\xi_l \sim f_1$
6: Simulate $\boldsymbol{V}_l \sim \mathcal{N}_d(0,\mathbf{I}_d)$; set $\boldsymbol{\Omega}_l = \sqrt{2\xi_l}\mathbf{V}_l$
7: Simulate $\Phi_l \sim \mathcal{U}(0,2\pi)$; Simulate $\boldsymbol{A}_l \sim \mathcal{N}_p(0,\boldsymbol{I}_p)$
8: **end for**
9: For each $\mathbf{x} = (\mathbf{s},t) \in \mathcal{S}$, and for $i = 1,\ldots,p$, return

$$\tilde{Z}_{L,i}(\mathbf{s},t) = \sqrt{\frac{2}{L}}\sum_{l=1}^{L} Z_{T,l,i}(t)\sqrt{\frac{f_{ii,\mathbf{x}}(\xi_l)}{f_1(\xi_l)}}\sqrt{\frac{\mu^G_{\boldsymbol{\Sigma}_{ii,\mathbf{x}}}(\sqrt{2}\mathbf{V}_l)}{\mu^G_{\mathbf{I}_d}(\sqrt{2}\mathbf{V}_l)}}\,(\boldsymbol{L}_{\mathbf{x}}\boldsymbol{A}_l)_i\cos\left(\boldsymbol{\Omega}_l^t\mathbf{s}+\Phi_l+\frac{\|\mathbf{V}_l\|}{\sqrt{2}}W_i(t)\right)$$

pointwise correlation
non-stationary importance weights
non sep. space-time

Nonstationary multivariate space-time model

Theorem (Allard et al., 2025+)

Let us denote $\mathbf{x} = (\mathbf{s}, t)$. Then,

$$C_{ij}(\mathbf{s}_1, \mathbf{s}_2; t_1, t_2) = |\boldsymbol{\Sigma}_{ii,\mathbf{x}_1}|^{1/4} |\boldsymbol{\Sigma}_{jj,\mathbf{x}_2}|^{1/4} \frac{\sigma_{ij,\mathbf{x}_1\mathbf{x}_2}}{|\boldsymbol{\Lambda}_{ij,\mathbf{x}_1,\mathbf{x}_2}|^{1/2}} \phi_{ij}\left(\boldsymbol{\Lambda}^{-1/2}_{ij,\mathbf{x}_1,\mathbf{x}_2}(\mathbf{s}_1 - \mathbf{s}_2); \boldsymbol{\theta}_{\mathbf{x}_1,\mathbf{x}_2}\right)$$

where

$$\boldsymbol{\Lambda}_{ij,\mathbf{x}_1,\mathbf{x}_2} = (\boldsymbol{\Sigma}_{ii,\mathbf{x}_1} + \boldsymbol{\Sigma}_{jj,\mathbf{x}_2})/2 + \gamma_{ij}(t_1 - t_2)\boldsymbol{I}_d$$

- Proof: it is the covariance resulting from the Algorithm above

Final words

- We propose a change of perspective: from spectral representation to Gaussian mixture representation
- It paves the way to general theorem allowing for the construction of a new and wide class of nonstationary covariance functions
- Two well separated steps: i) stochastic generation; ii) projection onto $\mathcal{S}$
- The second step is massively parallelizable
- Many possible extensions: non-stationarity in time, including transport and advection, non Euclidean spaces, etc.

Preprint

Allard, D., Benoit, L., & Obakrim, S. (2025). Modeling and simulating spatio-temporal, multivariate and nonstationary Gaussian Random Fields: a Gaussian mixtures perspective. *Preprint* https://hal.inrae.fr/hal-05034982. With *Statistical Science* (minor revisions submitted)

References

Allard, D., Clarotto, L., and Emery, X. (2022). Fully nonseparable gneiting covariance functions for multivariate space–time data. *Spatial Statistics*, 52:100706.

Allard, D., Emery, X., Lacaux, C., and Lantuéjoul, C. (2020). Simulating space-time random fields with nonseparable gneiting-type covariance functions. *Statistics and Computing*, 30(5):1479–1495.

Emery, X. and Arroyo, D. (2018). On a continuous spectral algorithm for simulating non-stationary gaussian random fields. *Stochastic Environmental Research and Risk Assessment*, 32:905–919.

Emery, X., Arroyo, D., and Porcu, E. (2016). An improved spectral turning-bands algorithm for simulating stationary vector Gaussian random fields. *Stochastic Environmental Research and Risk Assessment*, 30(7):1863–1873.

Matheron, G. (1973). The intrinsic random functions and their applications. *Advances in applied probability*, 5(3):439–468.

Paciorek, C. J. and Schervish, M. J. (2006). Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics*, 17.

Schoenberg, I. J. (1938). Metric spaces and completely monotone functions. *Annals of Mathematics*, 39(4):811–841.

Shinozuka, M. (1971). Simulation of multivariate and multidimensional random processes. *The Journal of the Acoustical Society of America*, 49(1B):357–368.