



**HAL**  
open science

# Fully nonseparable Gneiting covariance functions for multivariate space-time data

Denis Allard, Lucia Clarotto, Xavier Emery

► **To cite this version:**

Denis Allard, Lucia Clarotto, Xavier Emery. Fully nonseparable Gneiting covariance functions for multivariate space-time data. 2022. hal-03564931

**HAL Id: hal-03564931**

**<https://hal.science/hal-03564931v1>**

Preprint submitted on 10 Feb 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Fully nonseparable Gneiting covariance functions for multivariate space-time data

Denis Allard<sup>a,\*</sup>, Lucia Clarotto<sup>b</sup>, Xavier Emery<sup>c,d</sup>

<sup>a</sup>Biostatistics and Spatial Processes (BioSP), INRAE, 84914 Avignon, France

<sup>b</sup>Mines Paris, PSL University, Centre de Geosciences, 77300 Fontainebleau, France

<sup>c</sup>Department of Mining Engineering, University of Chile, Avenida Beauchef 850, Santiago 8370448, Chile

<sup>d</sup>Advanced Mining Technology Center, University of Chile, Avenida Beauchef 850, Santiago 8370448, Chile

\*Corresponding author: denis.allard@inrae.fr

## Abstract

We broaden the well-known Gneiting class of space-time covariance functions by introducing a very general parametric class of fully nonseparable direct and cross-covariance functions for multivariate random fields, where each component has a spatial covariance function from the Matérn family with its own smoothness and scale parameters and, unlike all currently available models, its own correlation function in time. We present sufficient conditions that result in valid models with varying degrees of complexity and we discuss the parameterization of those. Continuous-in-space and discrete-in-time simulation algorithms are also given, which are not limited by the number of target spatial coordinates and allow tens of thousands of time coordinates. The application of the proposed model is illustrated on a weather trivariate dataset over France. Our new model yields better fitting and better predictive scores compared to a more parsimonious model with common temporal correlation function.

**Keywords:** Spatio-temporal modeling; matrix-valued covariance function; pseudo-variogram; Matérn covariance; spectral simulation

# 1 Introduction

An increasing amount of multivariate data indexed by space-time coordinates are available in numerous scientific and engineering applications, which range from environmental and climate studies to remote sensing and monitoring systems in industrial operations and process management. For statisticians analyzing these data, one of the key issues is to model the space-time dependence structure not only for each variable separately, but also across the variables. This requires versatile models that can account for different scale and smoothness parameters for each variable, and yet whose parameters can be accurately estimated. [Bourotte et al. \(2016\)](#) proposed valid multivariate space-time classes based on Matérn and Cauchy spatial covariance functions, inspired by the multivariate Matérn model in [Gneiting et al. \(2010\)](#) and [Apanasovich et al. \(2012\)](#). For both classes, each variable has its own scale and own degree of smoothness in space, while allowing for some degree of cross-correlation. However, in this construction, the marginal temporal correlation function is identical for all variables, which, as discussed in [Bourotte et al. \(2016\)](#), was already seen as a restrictive assumption because it was found that the time correlations of the three variables in the analyzed dataset were different from one variable to the other.

To account for different scale parameters in time, space-time linear models of coregionalizations can be proposed, as in [Finazzi et al. \(2013\)](#). However, there are two major drawbacks to this construction. First, the smoothness of any component of the multivariate field is restricted to that of the roughest underlying univariate field ([Gneiting et al., 2010](#)). Second, the number of parameters to be estimated increases quickly as the number of components of the model increases, thus raising issues of robustness and over-fitting. [Porcu et al. \(2021\)](#) discussed several strategies for building valid multivariate space-time covariance functions, but they all lead to models with temporal correlation functions with an identical smoothness parameter in time.

In this work, we propose new classes of multivariate space-time covariance functions with Matérn spatial traces that allow, for each variable, different correlation functions in time and different smoothness and scale parameters in space. Our main contribution is to establish sufficient validity conditions for a very general class of matrix-valued covariance functions of the type

$$C_{ij}(\mathbf{h}, u) = \sigma_{ij}(u)\mathcal{M}(\mathbf{h}; r_{ij}(u), \nu_{ij}(u)), \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}, \quad i, j = 1, \dots, p,$$

where  $\mathcal{M}(\cdot; r, \nu)$  denotes a Matérn covariance function with scale and smoothness parameters  $r$  and  $\nu$ . We offer two main constructions. The first one builds on the substitution approach proposed in [Allard et al. \(2020\)](#) for simulating univariate space-time random fields with nonseparable Gneiting-type covariance functions. The second construction uses results on multivariate spatial covariance models established in [Emery et al. \(2022\)](#). These constructions offer new modeling possibilities, in particular negatively nonseparable covariance functions.

The weather dataset analyzed in [Bourotte et al. \(2016\)](#), consisting of a 10-year record of three daily variables (solar radiation, temperature and humidity) is revisited using a specific model obtained with the second construction. We will show that the advantages of the fully nonseparable model proposed in this work over the less flexible model are twofold. On the one hand, the new model offers a better fit to the data thanks to the additional parameters in the temporal covariance. On the other hand, we will show improved performances for both temporal prediction (i.e. predicting values at existing weather stations for day  $t$  using values from all weather stations at days  $t - 1, t - 2, \dots$ ) and spatio-temporal interpolation (i.e. predicting values at locations without a weather station for day  $t$  using values at days  $t, t - 1, t - 2, \dots$ ).

The outline of the paper is the following. Section 2 provides the necessary background material on matrix-valued space-time covariance functions, on the Gneiting class of space-time covariance functions and on matrix-valued pseudo-variograms. Section 3 proposes a first, parsimonious, construction based on the substitution approach. Section 4 then proposes a versatile, fully multivariate, Gneiting-Matérn class. In both sections, simulation algorithms are provided for the respective class of models. Section 5 illustrates the fully multivariate Gneiting-Matérn model on a multivariate weather dataset. Concluding remarks follow in Section 6. Technical lemmas, proofs of our results and other general forms of the Gneiting-Matérn model are given in appendices.

## 2 Background material

Throughout this work, we will use roman letters for scalars and bold letters for vectors, matrices and matrix-valued functions. In particular,  $\mathbf{0}$  and  $\mathbf{1}$  denote vectors or matrices of all-zeros and all-ones, respectively. Also, all matrix operations (product, ratio, power, square root, exponential, etc.) are understood as elementwise.

### 2.1 Positive semidefinite, conditionally positive semidefinite, and separable matrices

Let  $p$  be a positive integer and  $\mathbf{a} = [a_{ij}]_{i,j=1}^p$  a symmetric real matrix. We write  $\mathbf{a} \succeq 0$  (resp.  $\mathbf{a} \preceq 0$ ) to indicate that  $\mathbf{a}$  is positive (resp. negative) semidefinite, i.e.  $\sum_{i=1}^p \sum_{j=1}^p \omega_i \omega_j a_{ij} \geq 0$  (resp.  $\leq 0$ ),  $\forall \omega_1, \dots, \omega_p \in \mathbb{R}$ . Likewise,  $\mathbf{a} \succeq_c 0$  (resp.  $\mathbf{a} \preceq_c 0$ ) indicates that  $\mathbf{a}$  is conditionally positive (resp. negative) semidefinite, i.e.,  $\sum_{i=1}^p \sum_{j=1}^p \omega_i \omega_j a_{ij} \geq 0$  (resp.  $\leq 0$ ),  $\forall \omega_1, \dots, \omega_p \in \mathbb{R}$  such that  $\sum_{i=1}^p \omega_i = 0$ . A conditionally null definite matrix is a matrix that is both conditionally positive and conditionally negative semidefinite. An addition separable ( $a$ -separable) matrix is such that  $2a_{ij} = a_{ii} + a_{jj}$  for all  $i, j = 1, \dots, p$ , and a product separable ( $p$ -separable) matrix is such that

$a_{ij}^2 = a_{ii}a_{jj}$  for all  $i, j = 1, \dots, p$ . It is straightforward to prove that any  $a$ -separable matrix is conditionally null definite and that any  $p$ -separable matrix with nonnegative entries is positive semidefinite (see Properties 1 and 2 in Appendix A).

## 2.2 Matrix-valued space-time covariance functions

We consider multivariate space-time random fields  $\mathbf{Z}(\mathbf{s}, t) = [Z_i(\mathbf{s}, t)]_{i=1}^p$  defined over  $\mathbb{R}^d \times \mathbb{R}$ , where  $d$  is the space dimension and  $p$  the number of random field components, each being real-valued. A space-time coordinate will be denoted  $(\mathbf{s}, t)$ , with  $\mathbf{s} \in \mathbb{R}^d$  and  $t \in \mathbb{R}$ . For the sake of light notations, the Euclidean norm of  $\mathbf{h}$  will be denoted  $|\mathbf{h}|$ . Without loss of generality, we assume that all random field components are centered, i.e.  $\mathbb{E}[Z_i(\mathbf{s}, t)] = 0$ ,  $\forall (\mathbf{s}, t) \in \mathbb{R}^d \times \mathbb{R}$ ,  $\forall i = 1, \dots, p$ . It will also be assumed that the multivariate random field  $\mathbf{Z}(\mathbf{s}, t)$  is second-order stationary, so that all the covariance functions exist and depend only on the space-time lag  $(\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}$ :

$$\text{Cov}(Z_i(\mathbf{s}, t), Z_j(\mathbf{s} + \mathbf{h}, t + u)) = C_{ij}(\mathbf{h}, u), \quad (1)$$

for any pair  $i, j = 1, \dots, p$ ,  $\forall (\mathbf{s}, \mathbf{s} + \mathbf{h}) \in \mathbb{R}^d \times \mathbb{R}^d$  and  $\forall (t, t + u) \in \mathbb{R} \times \mathbb{R}$ . The functions  $C_{ij}$  are called direct covariance functions when  $i = j$  and cross-covariance functions otherwise. The matrix-valued covariance function  $\mathbf{C} : \mathbb{R}^d \times \mathbb{R} \rightarrow M_{p \times p}$ , where  $M_{p \times p}$  is the set of  $p \times p$  real matrices, associates each space-time lag  $(\mathbf{h}, u)$  with the  $p \times p$  matrix  $\mathbf{C}(\mathbf{h}, u) = [C_{ij}(\mathbf{h}, u)]_{i,j=1}^p$ .  $\mathbf{C}$  must be positive semidefinite in  $\mathbb{R}^d \times \mathbb{R}$ , that is, for any finite collection of space-time coordinates  $(\mathbf{s}_k, t_k)_{k=1}^N$ , the matrix  $[[C_{ij}(\mathbf{s}_l - \mathbf{s}_k, t_l - t_k)]_{i,j=1}^p]_{k,l=1}^N$  is positive semidefinite.

Setting  $t = 0$  in (1) defines the purely spatial matrix-valued covariance function  $\mathbf{C}_S(\mathbf{h}) := \mathbf{C}(\mathbf{h}, 0)$ . Likewise,  $\mathbf{C}_T(u) := \mathbf{C}(\mathbf{0}, u)$  is a purely temporal matrix-valued covariance function.

A matrix-valued space-time covariance model is space-time separable if it is the elementwise product of a matrix-valued spatial covariance function with a matrix-valued temporal covariance matrix, i.e. when  $C_{ij}(\mathbf{h}, u)C_{ij}(\mathbf{0}, 0) = C_{ij}(\mathbf{h}, 0)C_{ij}(\mathbf{0}, u)$  for all  $i, j = 1, \dots, p$ . The direct covariance  $C_{ii}$  is said to be positively (resp. negatively) space-time nonseparable if  $C_{ii}(\mathbf{h}, u) \geq C_{ii}(\mathbf{0}, 0)^{-1}C_{ii}(\mathbf{h}, 0)C_{ii}(\mathbf{0}, u)$  (resp.  $\leq$ ) for all  $(\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}$ .

A matrix-valued space-time covariance model is said to be a proportional model when it is obtained as the product of a  $p \times p$  covariance matrix  $\mathbf{R}$  and a valid univariate space-time correlation function  $\rho$ , i.e.  $\mathbf{C}(\mathbf{h}, u) = \mathbf{R}\rho(\mathbf{h}, u)$ ,  $\forall (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}$ . It is proportional-in-time if  $\mathbf{C}(\mathbf{h}, u) = \mathbf{R}\rho(\mathbf{0}, u)$  and proportional-in-space if  $\mathbf{C}(\mathbf{h}, u) = \mathbf{R}\rho(\mathbf{h}, 0)$ ,  $\forall (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}$ . Bourotte et al. (2016) proposed nonproportional multivariate space-time covariance models in which the purely spatial part is a nonproportional multivariate model but the purely temporal part is a proportional multivariate model, i.e., the model is proportional-in-time but not in space. In this work, we will build valid and flexible parametric classes of matrix-valued space-time covariance functions such that both the spatial and

temporal components are multivariate and nonproportional. Such matrix-valued space-time covariances are referred to as being *fully nonseparable*.

### 2.3 The univariate Gneiting class of space-time covariance functions

We first restrict ourselves to the univariate setting, i.e.,  $p = 1$ . We thus choose to drop temporarily the use of subscripts for the ease of notation. The Gneiting class of spatio-temporal covariance functions was originally presented in [Gneiting \(2002\)](#). For an easier exposition of our results, we follow [Allard et al. \(2020\)](#) and consider the extended class of functions of the form

$$C(\mathbf{h}, u) = \frac{1}{(\gamma(u) + 1)^{d/2}} \varphi\left(\frac{|\mathbf{h}|^2}{\gamma(u) + 1}\right), \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}. \quad (2)$$

[Zastavnyi and Porcu \(2011\)](#) showed that the function in (2) is a second-order stationary covariance function for any continuous and completely monotone function  $\varphi$  on  $[0, \infty)$  if and only if the function  $\gamma : \mathbb{R} \rightarrow [0, \infty)$  is a variogram on  $\mathbb{R}$ . Recall that the variogram of an intrinsically stationary temporal random field  $Z(t)$ ,  $t \in \mathbb{R}$ , is defined as

$$\gamma(u) = \frac{1}{2} \mathbb{E} \left[ (Z(t+u) - Z(t))^2 \right], \quad \forall t, u \in \mathbb{R},$$

and is a conditionally negative semidefinite function ([Chilès and Delfiner, 2012](#)). The Gneiting class of spatio-temporal covariances thus involves two functions  $\varphi$  and  $\gamma$ : the former is associated with the “spatial” structure, since  $C_S(\mathbf{h}) = \varphi(|\mathbf{h}|^2)$ , whereas the latter is associated with the temporal structure, with  $C_T(u) = (\gamma(u) + 1)^{-d/2}$ . Notice that  $C_T(u) \rightarrow 0$  as  $|u| \rightarrow \infty$  if and only if  $\gamma(u)$  is unbounded. [Allard et al. \(2020\)](#) further showed that a Gneiting covariance as in (2) has a spectral density if and only if  $\gamma(u)$  is unbounded and its spectral measure is absolutely continuous. From a modeling point of view, the formulation (2) offers more flexibility than the original formulation in [Gneiting \(2002\)](#) and provides a direct geostatistical interpretation in the temporal dimension.

Examples of such classes include the Gneiting-Matérn and Gneiting-Cauchy covariance functions, in the cases where  $\varphi(|\mathbf{h}|^2)$  is the Matérn or the Cauchy spatial covariance, respectively. The expression of the Matérn covariance is

$$\mathcal{M}(\mathbf{h}; r, \nu) = \frac{2^{1-\nu}}{\Gamma(\nu)} (r|\mathbf{h}|)^\nu \mathcal{K}_\nu(r|\mathbf{h}|), \quad \mathbf{h} \in \mathbb{R}^d, \quad (3)$$

where  $r > 0$  is a scale parameter ( $1/r$  is called the range),  $\nu > 0$  is a smoothness parameter and  $\mathcal{K}_\nu$  denotes the modified Bessel function of the second kind of order  $\nu$ . A second-order stationary spatial random field  $Z(\mathbf{s})$  with Matérn covariance is  $\kappa$  times mean square differentiable if and only if  $\nu \geq \kappa$ , see [Chilès and Delfiner \(2012\)](#) for a detailed exposition. The Cauchy covariance is

$$\mathcal{C}(\mathbf{h}; r, \alpha) = (1 + r|\mathbf{h}|^2)^{-\alpha}, \quad \mathbf{h} \in \mathbb{R}^d, \quad (4)$$

with  $\alpha, r > 0$ .

## 2.4 Proportional-in-time multivariate Gneiting classes

Going back to the multivariate setting, [Bourotte et al. \(2016\)](#) proposed classes of multivariate space-time covariance models that are extensions of the Gneiting class in (2), where the completely monotone real-valued function  $\varphi$  is replaced by a matrix-valued function  $\boldsymbol{\varphi} = [\varphi_{ij}]_{i,j=1}^p$ , with each component  $\varphi_{ij}$  having specific parameters. In particular, they showed that the multivariate Gneiting-Matérn space-time model  $\mathbf{C}^{\mathcal{M}} = [C_{ij}^{\mathcal{M}}]_{i,j=1}^p$  with

$$C_{ij}^{\mathcal{M}}(\mathbf{h}, u) = \frac{\sigma_i \sigma_j \rho_{ij}}{(\gamma(u) + 1)^{d/2}} \mathcal{M} \left( \mathbf{h}; \frac{r_{ij}}{(\gamma(u) + 1)^{1/2}}, \nu_{ij} \right), \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}, \quad (5)$$

is a valid second-order stationary matrix-valued covariance function if, for all  $i, j = 1, \dots, p$ ,

$$2r_{ij}^2 = r_i^2 + r_j^2; \quad 2\nu_{ij} = \nu_i + \nu_j; \quad \rho_{ij} = \beta_{ij} \frac{\Gamma(\nu_{ij})}{\Gamma(\nu_i)^{1/2} \Gamma(\nu_j)^{1/2}} \frac{r_i^{\nu_i} r_j^{\nu_j}}{r_{ij}^{2\nu_{ij}}}, \quad (6)$$

where  $\Gamma$  is the gamma function,  $\boldsymbol{\beta} = [\beta_{ij}]_{i,j=1}^p$  is a correlation matrix and  $r_i, \nu_i > 0$  for  $i = 1, \dots, p$ . Using matrix notations, the model in (5) is valid if the matrices  $\boldsymbol{\nu} = [\nu_{ij}]_{i,j=1}^p$  and  $\mathbf{r}^2 = [r_{ij}^2]_{i,j=1}^p$  have positive entries and are  $a$ -separable, and if  $\boldsymbol{\rho} \mathbf{r}^{2\nu} / \Gamma(\boldsymbol{\nu}) \succeq 0$ .

Similarly, the multivariate Gneiting-Cauchy space-time model  $\mathbf{C}^{\mathcal{C}} = [C_{ij}^{\mathcal{C}}]_{i,j=1}^p$  with

$$C_{ij}^{\mathcal{C}}(\mathbf{h}, u) = \frac{\sigma_i \sigma_j \rho_{ij}}{(\gamma(u) + 1)^{d/2}} \mathcal{C} \left( \mathbf{h}; \frac{r_{ij}}{(\gamma(u) + 1)^{1/2}}, \alpha_{ij} \right), \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}, \quad (7)$$

is a valid matrix-valued covariance function if, for all  $i, j = 1, \dots, p$ ,

$$2r_{ij}^{-1} = r_i^{-1} + r_j^{-1}, \quad 2\alpha_{ij} = \alpha_i + \alpha_j, \quad \rho_{ij} = \beta_{ij} \frac{\Gamma(\alpha_{ij})}{\Gamma(\alpha_i)^{1/2} \Gamma(\alpha_j)^{1/2}} \frac{r_{ij}^{\alpha_{ij}}}{(r_i^{\alpha_i} r_j^{\alpha_j})^{1/2}}, \quad (8)$$

with  $r_i, \alpha_i > 0$ , and  $\boldsymbol{\beta} = [\beta_{ij}]_{i,j=1}^p$  being a positive semidefinite matrix. The sufficient conditions in (8) are thus that  $\boldsymbol{\alpha}$  and  $\mathbf{r}^{-1}$  are  $a$ -separable and  $\boldsymbol{\rho} \mathbf{r}^{-\alpha} / \Gamma(\boldsymbol{\alpha}) \succeq 0$ . Proofs are given in [Bourotte et al. \(2016\)](#).

Setting  $\mathbf{h} = \mathbf{0}$  in both cases, the entries of the associated temporal matrix-valued covariance functions are

$$C_{ij}^{\mathcal{M}}(\mathbf{0}, u) = C_{ij}^{\mathcal{C}}(\mathbf{0}, u) = \frac{\sigma_i \sigma_j \rho_{ij}}{(\gamma(u) + 1)^{d/2}}, \quad u \in \mathbb{R}.$$

As discussed in Section 2.2, these are proportional-in-time models, since they are the product of a covariance matrix with a single time correlation function  $\rho_T(u) = (\gamma(u) + 1)^{-d/2}$ . In order to build nonproportional multivariate models both in space and time, an essential building block is to define admissible and relevant temporal matrix-valued covariances of the form  $[(\gamma_{ij}(u) + 1)^{-d/2}]_{i,j=1}^p$ . This point is addressed in the next section.

## 2.5 Matrix-valued variograms and pseudo-variograms

Consider a  $p$ -variate temporal random field  $\mathbf{Z}(t) = [Z_i(t)]_{i=1}^p$ ,  $t \in \mathbb{R}$ . Under an assumption of joint intrinsic stationarity, the direct and cross variograms of  $\mathbf{Z}$  are defined as

$$\gamma_{ij}(u) = \frac{1}{2} \text{Cov}(Z_i(t+u) - Z_i(t), Z_j(t+u) - Z_j(t)), \quad (9)$$

for all  $t, u \in \mathbb{R}$  and any pair  $i, j = 1, \dots, p$ . One can define a  $p \times p$  matrix-valued variogram function  $u \mapsto \boldsymbol{\gamma}(u) = [\gamma_{ij}(u)]_{i,j=1}^p$ . For a second-order stationary random field  $\mathbf{Z}$ , the matrix-valued covariance  $\mathbf{C}$  and variogram  $\boldsymbol{\gamma}$  are related by the identity  $\boldsymbol{\gamma}(u) = \mathbf{C}(0) - (\mathbf{C}(u) + \mathbf{C}(-u))/2$ . A necessary and sufficient condition for  $\boldsymbol{\gamma}$  to be a matrix-valued variogram is that  $\boldsymbol{\gamma}(0) = \mathbf{0}$ ,  $\boldsymbol{\gamma}(u) = \boldsymbol{\gamma}(-u) = \boldsymbol{\gamma}(u)^\top$  for all  $u \in \mathbb{R}$  and

$$\sum_{i=1}^p \sum_{j=1}^p \sum_{k=1}^N \sum_{l=1}^N \omega_{ik} \omega_{jl} \gamma_{ij}(t_l - t_k) \leq 0$$

for all choices of the positive integer  $N$ , the time coordinates  $t_1, \dots, t_N$  and the real values  $((\omega_{ik})_{i=1}^p)_{k=1}^N$  such that  $\sum_{k=1}^N \omega_{ik} = 0$  for all  $i = 1, \dots, p$  (Dörr and Schlather, 2021).

The matrix-valued variogram  $\boldsymbol{\gamma}$  is not the only function that can describe the cross-variations between the random field components  $Z_1(t), \dots, Z_p(t)$ . The pseudo-variogram is the  $p \times p$  matrix-valued function  $\boldsymbol{\eta}$  with entries

$$\eta_{ij}(u) = \frac{1}{2} \text{Var}[Z_i(t+u) - Z_j(t)], \quad t, u \in \mathbb{R}, \quad i, j = 1, \dots, p. \quad (10)$$

The right-hand side of (10) does not depend on  $t$  under the assumption that the direct and cross-increments are intrinsically stationary. Notice that all the entries of the pseudo-variogram matrix are nonnegative and that, unlike the variogram in (9), the pseudo-variogram is not necessarily an even function. Notice also that  $\eta_{ii}(u) = \gamma_{ii}(u)$ ,  $\forall u \in \mathbb{R}$  and for any  $i = 1, \dots, p$ .

A necessary and sufficient condition for  $\boldsymbol{\eta}$  to be a matrix-valued pseudo-variogram is that  $\eta_{ii}(0) = 0$  for  $i = 1, \dots, p$ ,  $\boldsymbol{\eta}(u) = \boldsymbol{\eta}(-u)^\top$  for all  $u \in \mathbb{R}$  and

$$\sum_{i=1}^p \sum_{j=1}^p \sum_{k=1}^N \sum_{l=1}^N \omega_{ik} \omega_{jl} \eta_{ij}(t_l - t_k) \leq 0$$

for all choices of the positive integer  $N$ , the time coordinates  $t_1, \dots, t_N$  and the real values  $((\omega_{ik})_{i=1}^p)_{k=1}^N$  such that  $\sum_{i=1}^p \sum_{k=1}^N \omega_{ik} = 0$ . Equivalently,  $\boldsymbol{\eta}$  is a matrix-valued pseudo-variogram if, and only if,  $u \mapsto \exp(-t\boldsymbol{\eta}(u))$  is a matrix-valued correlation function, i.e., a matrix-valued covariance function whose diagonal entries are equal to 1 at  $u = 0$ , for all  $t > 0$  (Dörr and Schlather, 2021).



### 3 First proposal: parsimonious Gneiting-Matérn and Gneiting-Cauchy multivariate models

In this section, we propose an approach to build multivariate space-time models that are neither proportional in space nor in time. It is based on the substitution approach proposed in [Allard et al. \(2020\)](#) for simulating univariate space-time Gaussian random fields characterized with Gneiting-type covariance functions. It also uses the integral representation of completely monotone functions. Compared to the model in (5), the univariate variogram model is generalized into a matrix-valued pseudo-variogram, as defined in (10). This construction is less general than the class that will be presented in Section 4, but it is more parsimonious and the construction is more direct.

#### 3.1 Model construction

**Theorem 1.** *Let  $\boldsymbol{\sigma} = [\sigma_{ij}]_{i,j=1}^p$  be a symmetric positive semidefinite matrix,  $u \mapsto \boldsymbol{\eta}(u) = [\eta_{ij}(u)]_{i,j=1}^p$  be a matrix-valued pseudo-variogram on  $\mathbb{R}$ , and  $t \mapsto \boldsymbol{\varphi}(t) = [\varphi_{ij}(t)]_{i,j=1}^p$  be a matrix of completely monotone functions on  $[0, \infty)$  such that  $\varphi_{ij}(t) = \int_0^\infty e^{-rt} (f_i(r)f_j(r))^{1/2} dr$ , where  $f_i$  and  $f_j$  are probability density functions on  $[0, \infty)$ . Then, the matrix-valued function  $\mathbf{C} : (\mathbf{h}, u) \mapsto [C_{ij}(\mathbf{h}, u)]_{i,j=1}^p$  with*

$$C_{ij}(\mathbf{h}, u) = \frac{\sigma_{ij}}{(\eta_{ij}(u) + 1)^{d/2}} \varphi_{ij} \left( \frac{|\mathbf{h}|^2}{\eta_{ij}(u) + 1} \right), \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}, \quad (11)$$

is positive semidefinite in  $\mathbb{R}^d \times \mathbb{R}$ .

**Example 1** (Gneiting-Matérn model). *Let  $f_i$  be the probability density of an inverse gamma distribution with shape parameter  $\nu_i > 0$  and scale parameter  $r_i^2/4 > 0$ :*

$$f_i(r) = \frac{1}{\Gamma(\nu_i)} \left( \frac{r_i}{2} \right)^{2\nu_i} r^{-\nu_i-1} \exp \left( -\frac{r_i^2}{4r} \right), \quad r > 0.$$

Based on the fact that an inverse gamma mixture of Gaussian covariances is a Matérn covariance ([Emery and Lantuéjoul, 2006](#)), the direct and cross-covariances are found to belong to the Gneiting-Matérn family:

$$C_{ij}(\mathbf{h}, u) = \frac{\Gamma(\nu_{ij})}{\sqrt{\Gamma(\nu_i)\Gamma(\nu_j)}} \frac{r_i^{\nu_i} r_j^{\nu_j}}{r_{ij}^{2\nu_{ij}}} \frac{\sigma_{ij}}{(\eta_{ij}(u) + 1)^{d/2}} \mathcal{M} \left( \mathbf{h}; \frac{r_{ij}}{(\eta_{ij}(u) + 1)^{1/2}}, \nu_{ij} \right), \quad (12)$$

with  $2\nu_{ij} = \nu_i + \nu_j$ ,  $2r_{ij}^2 = r_i^2 + r_j^2$  and  $\mathcal{M}$  the Matérn covariance defined in (3).

**Example 2** (Gneiting-Cauchy model). *Let  $f_i$  be a gamma probability density with shape parameter  $\alpha_i > 0$  and scale parameter  $r_i$*

$$f_i(r) = \frac{1}{\Gamma(\alpha_i)} \exp \left( -\frac{r}{r_i} \right) r_i^{-\alpha_i} r^{\alpha_i-1}, \quad r > 0.$$

Based on the fact that a gamma mixture of Gaussian covariances is a Cauchy covariance (Emery and Lantuéjoul, 2006), the direct and cross-covariances are found to belong to the Gneiting-Cauchy family:

$$C_{ij}(\mathbf{h}, u) = \frac{\Gamma(\alpha_{ij})}{\sqrt{\Gamma(\alpha_i)\Gamma(\alpha_j)}} \frac{r_{ij}^{\alpha_{ij}}}{r_i^{\alpha_i/2} r_j^{\alpha_j/2}} \frac{\sigma_{ij}}{(\eta_{ij}(u) + 1)^{d/2}} \mathcal{C} \left( \frac{\mathbf{h}}{(\eta_{ij}(u) + 1)^{1/2}}; r_{ij}, \alpha_{ij} \right), \quad (13)$$

with  $2\alpha_{ij} = \alpha_i + \alpha_j$ ,  $2r_{ij}^{-1} = r_i^{-1} + r_j^{-1}$  and  $\mathcal{C}$  the Cauchy covariance defined in (4).

The two previous examples generalize (5) and (7), with a matrix-valued pseudo-variogram  $\boldsymbol{\eta}$  instead of a usual variogram  $\gamma$ .

The advantage of this construction is that it is easy to understand and that it is straightforward to simulate realizations of the multivariate random field  $\mathbf{Z}$ , as shown in the next subsection. However, an implicit assumption is that  $\varphi_{ij}(t) = \int_0^\infty e^{-tr} f_{ij}(r) dr$ , with  $f_{ij}$  being  $p$ -separable:  $f_{ij} = f_i^{1/2} f_j^{1/2}$ . This is a parsimonious parameterization, but also a restrictive condition. As shown in Emery et al. (2022) and in Section 4.1, this condition may entail weak correlations between  $Z_i$  and  $Z_j$ . In Section 4, we will propose a more general class of multivariate Gneiting-Matérn models that encompasses the model presented in Example 1.

### 3.2 Substitution simulation algorithm

A realization of the  $p$ -variate random field  $\mathbf{Z} = [Z_i]_{i=1}^p$  with approximate Gaussian distribution can be constructed from the standardized sum of a large number  $M$  of independent copies of non-Gaussian random fields as in Equation (23) (Appendix B), each of which depending only on two random variables  $R$  and  $\Phi$ , two random vectors  $\mathbf{V}$  and  $\boldsymbol{\Omega}$  and a temporal  $p$ -variate random field  $\mathbf{W}$ . This is summarized in Algorithm 1 hereinafter, where  $\langle \cdot, \cdot \rangle$  stands for the usual scalar product in  $\mathbb{R}^d$ ,  $R \sim f$  is the short notation for “ $R$  is distributed according to the distribution  $f$ ”,  $\mathcal{U}(0, 2\pi)$  is the uniform distribution on  $(0, 2\pi)$ ,  $\mathcal{N}_d(\mathbf{0}, \boldsymbol{\Sigma})$  is the  $d$ -dimensional normal distribution with expectation  $\mathbf{0}$  and variance-covariance matrix  $\boldsymbol{\Sigma}$ , and  $\mathbf{I}_d$  is the  $d \times d$  identity matrix.

The intrinsic random field  $\mathbf{W} = [W_i]_{i=1}^p$  with Gaussian direct and cross-increments can be simulated with the covariance matrix decomposition method (Algorithm 2 below), by setting  $W_1(0) = 0$  and using the nonstationary covariance function (Papritz et al., 1993, Equation (6))

$$\text{Cov}(W_i(t), W_j(t')) = \text{Cov}(W_i(t) - W_1(0), W_j(t') - W_1(0)) = -\eta_{ij}(t - t') + \eta_{i1}(t) + \eta_{1j}(-t'). \quad (14)$$

This construction (as well as the more general class presented in Section 4) is limited to multivariate random fields  $\mathbf{W}$  with Gaussian direct and cross-increments that are characterized by the pseudo-variogram  $\boldsymbol{\eta}$ . The discussion on valid classes of pseudo-variograms is deferred to Section 4.3.

The method is applicable as long as the number of time coordinates targeted for simulation is not too large (in practice, up to a few tens of thousands), so that the square root matrix can be

---

**Algorithm 1** Substitution simulation algorithm

---

**Require:** Admissible Matérn or Cauchy parameters as given in Theorem 1

**Require:** Admissible pseudo-variogram  $\eta$

**Require:** A probability density function  $f$  on  $[0, \infty)$  whose support contains the supports of  $f_1, \dots, f_p$

**Require:** A large integer  $M$

- 1: Initialize  $Z_i(\mathbf{s}, t) = 0$ , for all  $i$  and at each target space-time coordinate  $(\mathbf{s}, t) \in \mathbb{R}^d \times \mathbb{R}$
- 2: **for**  $m = 1$  to  $M$  **do**
- 3:   Simulate  $R \sim f$
- 4:   Simulate  $\boldsymbol{\Omega} \sim \mathcal{N}_d(\mathbf{0}, \mathbf{I}_d)$
- 5:   Simulate  $\Phi \sim \mathcal{U}(0, 2\pi)$
- 6:   Simulate  $\mathbf{V} \sim \mathcal{N}_p(\mathbf{0}, \boldsymbol{\sigma})$
- 7:   Simulate a  $p$ -variate intrinsic random field  $\mathbf{W} = [W_i]_{i=1, \dots, p}$  with zero-mean Gaussian increments and pseudo-variogram  $\eta$  at each target time coordinate  $t$
- 8:   **for** each target space-time coordinate  $(\mathbf{s}, t)$  **do**
- 9:     Compute

$$Z_i(\mathbf{s}, t) = Z_i(\mathbf{s}, t) + \sqrt{\frac{2}{M}} \sqrt{\frac{f_i(R)}{f(R)}} V_i \cos\left(\sqrt{2R} \langle \boldsymbol{\Omega}, \mathbf{s} \rangle + \frac{|\boldsymbol{\Omega}|}{\sqrt{2}} W_i(t) + \Phi\right), \quad i = 1, \dots, p$$

10:   **end for**

11: **end for**

---

---

**Algorithm 2** Matrix decomposition algorithm

---

**Require:** Admissible pseudo-variogram  $\eta$

**Require:** Set of time coordinates  $t_1, \dots, t_{n_T}$  targeted for simulation

- 1: Calculate the  $(p n_T) \times (p n_T)$  variance-covariance matrix  $\mathbf{C}_W = [[\text{Cov}(W_i(t_m), W_j(t_n))]_{i,j=1}^p]_{m,n=1}^{n_T}$ , as per (14)
  - 2: Simulate  $\mathbf{X} \sim \mathcal{N}_{p n_T}(\mathbf{0}, \mathbf{I}_{p n_T})$
  - 3: Compute  $\mathbf{W} = [[W_i(t_n)]_{i=1}^p]_{n=1}^{n_T} = \mathbf{C}_W^{1/2} \mathbf{X}$ , with  $\mathbf{C}_W^{1/2}$  the principal square root (Horn and Johnson, 2013, Theorem 7.2.6) of  $\mathbf{C}_W$
-

computed. For larger numbers, iterative simulation methods based on Gibbs sampling can be used, still with the nonstationary covariance function (14) (Lantuéjoul and Desassis, 2012; Arroyo and Emery, 2020). Also, for specific classes of pseudo-variograms such as the one presented in Section 4.3, the multivariate random field  $\mathbf{W}$  can be constructed by combinations of stationary Gaussian random fields and intrinsic random fields with Gaussian increments, for which discrete (Dietrich and Newsam, 1997; Stein, 2002) and continuous (Emery et al., 2016; Arroyo and Emery, 2017) spectral simulation algorithms are available.

The distribution  $f$  in Algorithm 1 can be any positive probability density function on  $\mathbb{R}^d$ . From a practical viewpoint,  $f$  should be chosen so as to quicken the convergence of the simulated random field to a Gaussian random field as  $M$  tends to infinity. Since in Examples 1 and 2 all  $f_i$ s belong to the same family of distributions, it is reasonable to choose  $f$  in that same family. Following Arroyo and Emery (2021), it is recommendable to set  $f$  as the distribution  $f_i$  with the heaviest tail. Accordingly, in the Gneiting-Matérn case,  $f$  will be an inverse gamma distribution with shape parameter  $\nu = \min_{1 \leq i} \nu_i$ . In the Gneiting-Cauchy case,  $f$  will be a gamma distribution with shape parameter  $\alpha = \max_{1 \leq i} \alpha_i$ .

## 4 Second proposal: fully multivariate Gneiting-Matérn classes

We now present a very general class of multivariate space-time covariance models for which the purely spatial part belongs to the Matérn family. This construction uses results on multivariate spatial covariance models established in Emery et al. (2022).

### 4.1 Main result

**Theorem 2.** *Let  $\boldsymbol{\nu}$ ,  $\mathbf{b}^2$  and  $\mathbf{a}^2$  be  $p \times p$  symmetric conditionally negative semidefinite matrices, all with positive entries, such that  $\mathbf{a}^2 - \boldsymbol{\nu} \preceq_c 0$ . Let  $u \mapsto \boldsymbol{\rho}(u)$  be a  $p \times p$  matrix-valued covariance function on  $\mathbb{R}$  and let  $u \mapsto \boldsymbol{\eta}(u)$  be a  $p \times p$  matrix-valued pseudo-variogram on  $\mathbb{R}$ . Then, the matrix-valued function  $\mathbf{C} : (\mathbf{h}, u) \mapsto [C_{ij}(\mathbf{h}, u)]_{i,j=1}^p$  with*

$$C_{ij}(\mathbf{h}, u) = \sigma_{ij}(u) \mathcal{M}(\mathbf{h}; r_{ij}(u), \nu_{ij}), \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R},$$

is positive semidefinite in  $\mathbb{R}^d \times \mathbb{R}$  for

$$\mathbf{r}(u) = [r_{ij}(u)]_{i,j=1}^p = \sqrt{\frac{\alpha \boldsymbol{\eta}(u) + \mathbf{a}^2}{\beta \boldsymbol{\eta}(u) + \mathbf{b}^2}} \quad \text{and} \quad \boldsymbol{\sigma}(u) = [\sigma_{ij}(u)]_{i,j=1}^p = \frac{\boldsymbol{\rho}(u) \Gamma(\boldsymbol{\nu}) \exp(\boldsymbol{\nu})}{(\alpha \boldsymbol{\eta}(u) + \mathbf{a}^2)^\nu (\beta \boldsymbol{\eta}(u) + \mathbf{b}^2)^{d/2}}, \quad (15)$$

where  $\alpha$  and  $\beta$  are nonnegative.

**Corollary 1.** Let  $\boldsymbol{\nu}$ ,  $\mathbf{b}^2$ ,  $\mathbf{a}^2$ ,  $\boldsymbol{\eta}$ ,  $\alpha$  and  $\beta$  be as in Theorem 2. Let  $\boldsymbol{\tau}$  be a  $p \times p$  symmetric real matrix. The matrix-valued function  $\mathbf{C} : (\mathbf{h}, u) \mapsto [C_{ij}(\mathbf{h}, u)]_{i,j=1}^p$  with

$$C_{ij}(\mathbf{h}, u) = \frac{\tau_{ij}}{(\beta\eta_{ij}(u) + b_{ij}^2)^{d/2}(\alpha\eta_{ij}(u) + a_{ij}^2)^{\nu_{ij}}} \mathcal{M} \left( \mathbf{h}; \sqrt{\frac{\alpha\eta_{ij}(u) + a_{ij}^2}{\beta\eta_{ij}(u) + b_{ij}^2}}, \nu_{ij} \right), \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}, \quad (16)$$

is positive semidefinite in  $\mathbb{R}^d \times \mathbb{R}$  if the matrix  $\boldsymbol{\tau}e^{-\boldsymbol{\nu}}/\Gamma(\boldsymbol{\nu})$  is positive semidefinite.

The multivariate space-time covariance model (16) is of Gneiting-Matérn type. The sufficient conditions on the parameters call for several comments.

1. A sufficient condition for  $\mathbf{a}^2 - \boldsymbol{\nu} \preceq_c 0$  is that  $\mathbf{a}^2 \preceq_c 0$  and  $\boldsymbol{\nu}$  is  $a$ -separable (see Property 4 in Appendix A). This more restrictive condition is also found in [Apanasovich et al. \(2012\)](#), [Genton and Kleiber \(2015\)](#) and [Bourotte et al. \(2016\)](#). Even though  $a$ -separability is not necessary, it is convenient because it is easy to satisfy. From a statistical point of view, it reduces the number of parameters from  $p(p+1)/2$  to  $p$  and therefore leads to a more parsimonious model.
2. The conditional negative semidefiniteness condition on  $\mathbf{a}^2$  is similar to that in [Apanasovich et al. \(2012\)](#) in a spatial context. It is weaker than the  $a$ -separability condition in (6) that applies to the Gneiting-Matérn model in [Bourotte et al. \(2016\)](#).
3. Since  $\boldsymbol{\nu} \preceq_c 0$ , the matrix  $\exp(-\boldsymbol{\nu})$  is positive semidefinite (see Property 3 in Appendix A). Hence, based on Schur's product theorem, a sufficient condition is that  $\boldsymbol{\tau}/\Gamma(\boldsymbol{\nu})$  is positive semidefinite.
4. The positive semidefiniteness condition on  $\boldsymbol{\tau}$  in Corollary 1 only depends on the smoothness parameters  $\boldsymbol{\nu}$ . In contrast, the positive semidefiniteness condition implied by the Gneiting-Matérn example for the substitution approach in (12) (which is similar to the condition (6) from [Bourotte et al. \(2016\)](#)) depends on both  $\boldsymbol{\nu}$  and  $\mathbf{r}(u)$ , which is a stronger condition. Theorem 2 therefore offers much more flexibility for modeling multivariate space-time data. As a striking illustration of this point, let us consider a bivariate Matérn spatio-temporal field as in (16), with  $\alpha = 0, \beta = 1$  and  $\mathbf{b}$  being an all-ones matrix. According to (6), the absolute value of the collocated correlation between the two field components is bounded from above by

$$\frac{\Gamma(\nu_{12})a_{11}^{\nu_{11}}a_{22}^{\nu_{22}}}{\sqrt{\Gamma(\nu_{11})\Gamma(\nu_{22})}a_{ij}^{2\nu_{ij}}}.$$

With the conditions of Theorem 2, the upper bound is  $\Gamma(\nu_{12})/\sqrt{\Gamma(\nu_{11})\Gamma(\nu_{22})}$  when  $\boldsymbol{\nu}$  is  $a$ -separable. For example, with  $\nu_{11} = 0.5, \nu_{12} = 1.5, \nu_{22} = 2.5, r_{11} = 1.25, r_{22} = 0.66$ , we get  $|\rho_{12}| \leq 0.229$  with condition (6) corresponding to [Bourotte et al. \(2016\)](#) and  $|\rho_{12}| \leq 0.577$  according to Theorem 2. The gain is substantial.

This Gneiting-Matérn class is quite versatile, even though more general classes, with the smoothness parameters  $\nu$  depending on the temporal lag  $u$ , can be elaborated (see Theorems 3 and 4 in Appendix C). The space-time properties depend very much upon the values of  $\alpha$  and  $\beta$ . Obviously, if  $\alpha = \beta = 0$ , the model is purely spatial. In this context, setting  $\mathbf{b}$  to be the all-ones matrix, the condition on  $\sigma$  reduces to the condition of Theorem 3B in Emery et al. (2022). Hence, as a side-result, Theorem 2 also provides new classes of multivariate Matérn spatial models.

When  $\alpha = 0$ , accounting for the fact that  $\exp(-\nu)$  and all  $p$ -separable matrices with nonnegative entries are positive semidefinite (see Properties 2 and 3 in Appendix A), the models in (15) and (16) generalize the Gneiting-Matérn models (5) and (12). If, furthermore, the direct variograms  $\eta_{ii}$  are increasing functions for  $i = 1, \dots, p$ , the scale parameters  $r_{ii}(u) = a_{ii}(\beta\eta_{ii}(u) + b_{ii}^2)^{-1/2}$  decrease as  $u$  increases. The space-time direct correlations are thus higher than what would happen for a separable space-time covariance function, i.e., they are positively nonseparable.

When  $\beta = 0$  and under the same monotonicity assumption for the direct variograms, the scale parameters  $r_{ii}(u) = \sqrt{\alpha\eta_{ii}(u) + a_{ii}^2}/b_{ii}$  increase with  $u$ . This is a very original feature, since for all Gneiting-type spatio-temporal models of our knowledge, the scale parameter decreases with  $u$ . Applying the same reasoning as above, the resulting space-time direct covariances are thus negatively nonseparable.

When  $\alpha$  and  $\beta$  are both positive, the model is positively or negatively space-time nonseparable, depending on the relative values of the other parameters involved in  $\mathbf{r}(u)$ . However, in the common case when  $\eta_{ij}(u) \rightarrow \infty$  as  $u \rightarrow \infty$ ,  $\mathbf{r}(u)$  asymptotically tends to a matrix with all elements equal to  $\sqrt{\alpha/\beta}$ . The space-time covariance is thus asymptotically separable as  $|u| \rightarrow \infty$ .

## 4.2 A spectral simulation algorithm

It is of interest to simulate a  $p$ -variate space-time random field  $\mathbf{Z}$  with zero mean and with a matrix-valued covariance function of the form  $\mathbf{C} : (\mathbf{h}, u) \mapsto \sigma(u) \mathcal{M}(\mathbf{h}; \mathbf{r}(u), \nu)$ , as stated in Theorem 2, on a grid of  $\mathbb{R}^d$  with  $n_S$  nodes  $\mathbf{s}_1, \dots, \mathbf{s}_{n_S}$  crossed with a grid of  $\mathbb{R}$  with  $n_T$  nodes  $t_1, \dots, t_{n_T}$ . Let  $q = pn_T$  and, for  $\boldsymbol{\omega} \in \mathbb{R}^d$ , define the  $q \times q$  matrix  $\mathbf{F}(\boldsymbol{\omega}) = [[F_{i,m;j,n}(\boldsymbol{\omega})]_{i,j=1}^p]_{m,n=1}^{n_T}$  whose generic entry is the spectral density of the Matérn covariance  $\mathbf{h} \mapsto \sigma_{ij}(u_{mn}) \mathcal{M}(\mathbf{h}; r_{ij}(u_{mn}), \nu_{ij})$ , with  $u_{mn} = t_m - t_n$ , evaluated at  $\boldsymbol{\omega}$  (Lantuéjoul, 2002):

$$F_{i,m;j,n}(\boldsymbol{\omega}) = \sigma_{ij}(u_{mn}) \frac{\pi^{\frac{d}{2}} 2^d \Gamma(\nu_{ij} + \frac{d}{2})}{r_{ij}(u_{mn})^d \Gamma(\nu_{ij})} \left( 1 + \frac{|2\pi\boldsymbol{\omega}|^2}{r_{ij}^2(u_{mn})} \right)^{-\frac{d}{2} - \nu_{ij}}.$$

In the same way as in the proof of Theorem 2 given in Appendix B, the  $p$ -variate random field  $\mathbf{Z}$  to simulate at the  $n_S \times n_T$  target nodes of  $\mathbb{R}^d \times \mathbb{R}$  can be viewed as a  $q$ -variate random field  $\mathbf{Y}$  at  $n_S$  nodes of  $\mathbb{R}^d$  with a multivariate Matérn covariance with parameters  $[[r_{ij}(u_{mn})]_{i,j=1}^p]_{m,n=1}^{n_T}$ ,

$[[\nu_{ij}]_{i,j=1}^p]_{m,n=1}^{n_T}$  and  $[[\sigma_{ij}(u_{mn})]_{i,j=1}^p]_{m,n=1}^{n_T}$ . The simulation can be done by means of the continuous spectral algorithm presented in [Emery et al. \(2016\)](#), which relies on importance sampling and on the square root decomposition of the  $q \times q$  positive semidefinite matrix  $\mathbf{F}(\boldsymbol{\Omega})$  evaluated at a random vector  $\boldsymbol{\Omega}$ . In addition to parameters verifying the conditions of [Theorem 2](#), a probability density function  $g$  that does not vanish on  $\mathbb{R}^d$  is required to define  $\boldsymbol{\Omega}$ . Owing to the central limit theorem, a random field with an approximately Gaussian distribution is obtained by rescaling properly the sum of a large number  $M$  of independent copies of  $\mathbf{Y}$ .

---

**Algorithm 3** Spectral simulation algorithm

---

**Require:** Admissible parameters for  $\boldsymbol{\nu}$ ,  $\mathbf{b}^2$  and  $\mathbf{a}^2$

**Require:** Admissible pseudo-variogram  $\boldsymbol{\eta}$

**Require:** A probability density function  $g$  not vanishing on  $\mathbb{R}^d$

**Require:** A large integer  $M$

- 1: Initialize  $\mathbf{Y}(\mathbf{s}_k) = \mathbf{0}$  for each target spatial coordinate  $\mathbf{s}_k$ ,  $k = 1, \dots, n_S$
- 2: **for**  $m = 1$  to  $M$  **do**
- 3:   Simulate  $\boldsymbol{\Omega} \sim g$
- 4:   Compute the principal square root  $\mathbf{H}(\boldsymbol{\Omega})$  of the matrix  $2\mathbf{F}(\boldsymbol{\Omega})/g(\boldsymbol{\Omega})$
- 5:   Simulate  $\Phi \sim \mathcal{U}(0, 2\pi)$
- 6:   Simulate a random integer  $P$  uniformly distributed in  $\{1, \dots, p\}$
- 7:   **for** each target spatial coordinate  $\mathbf{s}_k$  **do**
- 8:     Compute

$$\mathbf{Y}(\mathbf{s}_k) = \mathbf{Y}(\mathbf{s}_k) + \sqrt{\frac{p}{M}} \mathbf{H}(\boldsymbol{\Omega}; P) \cos(2\pi\langle \boldsymbol{\Omega}, \mathbf{s}_k \rangle + \Phi)$$

where  $\mathbf{H}(\boldsymbol{\Omega}; P)$  denotes the  $P$ -th column of  $\mathbf{H}(\boldsymbol{\Omega})$

- 9:   **end for**
  - 10: **end for**
  - 11:  $\mathbf{Z} :=$  re-ordering of  $\mathbf{Y}$  into a  $p$ -variate random field at  $n_S \times n_T$  space-time coordinates
- 

The spectral simulation algorithm is applicable when  $q$  is not too large, say, less than a few tens of thousands, so that the square root decomposition of the spectral density matrix  $\mathbf{F}(\boldsymbol{\Omega})$  can be achieved. There is no such restriction on the number of spatial nodes  $n_S$ , which can be much higher, insofar as the simulation in space is achieved in a continuous manner, the spatial random field being a mixture of cosine waves that can be computed at as many coordinates as desired.

Following the reasoning of [Section 3.2](#), the density  $g$  can be chosen as a multivariate distribution with  $p$  independent marginal gamma distributions with shape parameter  $\nu = \min_{1 \leq i} \nu_{ii}$ .

### 4.3 Models for unbounded pseudo-variograms

Recall that the temporal structure is described by a matrix-valued pseudo-variogram  $\boldsymbol{\eta} : u \mapsto [\eta_{ij}(u)]_{i,j=1}^p$ , which must be unbounded as discussed in Section 2.3. Matrix-valued variograms can easily be constructed by use of the Linear Model of Coregionalization (LMC) (Chilès and Delfiner, 2012; Genton and Kleiber, 2015). However, as shown in Chen and Genton (2019) the LMC-based approach cannot be used to construct matrix-valued pseudo-variograms with different diagonal entries, which is necessary for a fully nonseparable multivariate model. In Papritz et al. (1993) it is shown that for a matrix-valued pseudo-variogram  $\boldsymbol{\eta}$ , the following large distance behavior holds:

$$\lim_{|u| \rightarrow \infty} \frac{\eta_{ij}(u)}{\eta_{ii}(u)} = 1, \quad (17)$$

for any pair  $i, j = 1, \dots, p$ . In other words, all entries of a matrix-valued pseudo-variogram must have the same behavior for very large lag separation  $u$ .

As a simple model for an unbounded matrix-valued pseudo-variogram with different diagonal entries (direct variograms), we propose the following construction. Let us define

$$W_i(u) = Y_0(u) + Y_i(u), \quad u \in \mathbb{R}, \quad (18)$$

where  $Y_1, \dots, Y_p$  are second-order stationary random fields with direct and cross-covariance functions  $\mathbf{R} : u \mapsto [R_{ij}(u)]_{i,j=1}^p$  and  $Y_0$  is an independent intrinsic random field with unbounded variogram  $\gamma_0$ . Then, the pseudo-variogram of  $\mathbf{W} = [W_i]_{i=1}^p$  is such that

$$\eta_{ij}(u) = \gamma_0(u) + \frac{R_{ii}(0) + R_{jj}(0)}{2} - R_{ij}(u), \quad u \in \mathbb{R}.$$

In matrix notation, a valid class of unbounded matrix-valued pseudo-variogram is thus

$$\boldsymbol{\eta}(u) = \gamma_0(u)\mathbf{1}_p + \mathbf{R}^0 - \mathbf{R}(u), \quad u \in \mathbb{R}, \quad (19)$$

where  $\mathbf{1}_p$  is the all-ones matrix of size  $p \times p$ ,  $\mathbf{R}$  is any matrix-valued covariance function and  $\mathbf{R}^0$  is the  $a$ -separable matrix with entries  $R_{ij}^0 = (R_{ii}(0) + R_{jj}(0))/2$ . Notice that  $\mathbf{1}_p$  and  $\mathbf{R}^0$  are conditionally negative semidefinite matrices, whilst  $\mathbf{R}$  is a positive semidefinite matrix-valued function.

A more general model consists in replacing  $Y_0(u)$  in (18) by  $Y_0(u + \delta_i)$ , with  $\delta_i \in \mathbb{R}$ , which leads to the matrix-valued pseudo-variogram  $\boldsymbol{\eta}(u) = [\gamma_0(u + \delta_i - \delta_j)]_{i,j=1}^p + \mathbf{R}^0 - \mathbf{R}(u)$ ,  $u \in \mathbb{R}$ .

## 5 Application to a weather dataset

### 5.1 Introduction

We now illustrate the use of the fully nonseparable multivariate Gneiting-Matérn model proposed in (16) on a weather dataset consisting of three daily variables (solar radiation  $R$ , temperature  $T$ , and



humidity H) recorded at 13 stations in Western France from 2003 to 2012. A previous analysis of this dataset (Bourotte et al., 2016) has shown that a proportional-in-time multivariate Gneiting-Matérn improved the modeling and the prediction scores as compared either to a model with a separable spatio-temporal covariance function or to univariate models. In this analysis, we shall compare a specific fully nonseparable multivariate model (FULL), as proposed in (16), to a Proportional-In-Time (PIT) model proposed in Bourotte et al. (2016). These models will be applied to residuals, after centering in time and space, in order to filter out any seasonal and regional effect.

An exploratory analysis was performed to assess the nonseparability of those residuals. For each of the three variables, the separability index  $S(\mathbf{h}, u) = C(\mathbf{h}, u) - C_S(\mathbf{h})C_T(u)$  was computed from the space-time empirical variogram obtained with the function `EVariogram` from the package `CompRandFld` (Padoan and Bevilacqua, 2015), using spatial bins of 50 m. Recall that, since the variograms have been computed on residuals, their sill are equal to 1 and the separability index must be in the interval  $[-1, 1]$ , with  $S(\mathbf{h}, u) = 0$  indicating separability of  $C(\mathbf{h}, u)$ . Figure 1 shows the separability index of all variables at the first three time steps. Without having to compute a formal test for separability as for example in Mitchell et al. (2006), it is clear from these plots that the residuals are uniformly positively space-time nonseparable for  $|\mathbf{h}| \leq 500$  m and  $u \leq 3$ , which is actually often the case for climate variables. The same analysis was performed using other values for the spatial bins and similar results were consistently obtained (not shown).

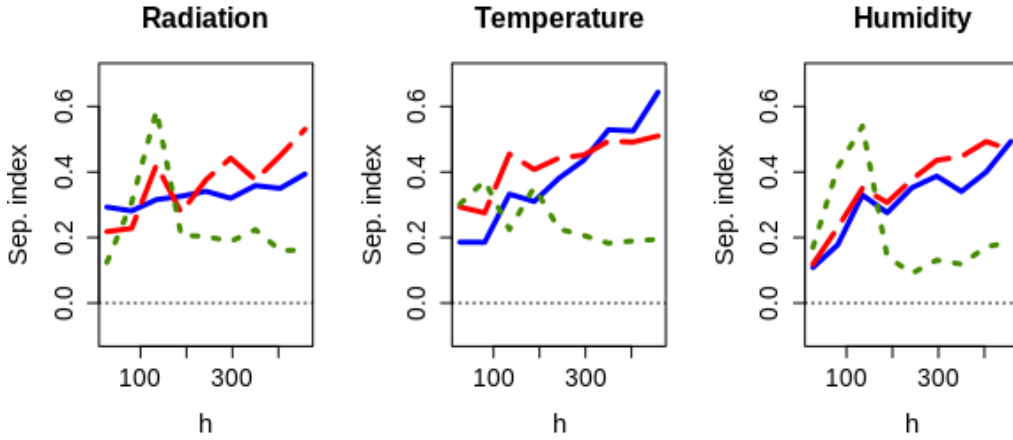


Figure 1: Separability index  $S(\mathbf{h}, u)$  computed at  $u = 1$  (blue solid lines),  $u = 2$  (red dashed lines) and  $u = 3$  (green dotted lines).

Maximum likelihood estimation requires computational costs of the order of  $pN$ , where  $N$  is the total number of space-time records, and  $p$  the number of variables. Maximizing the likelihood is thus computationally demanding, when possible. Composite likelihood methods have proven efficient in

the context of spatio-temporal data, with less time-consuming steps and good asymptotic properties (Bevilacqua et al., 2012). Pairwise Likelihood (PL) is the product of marginal Gaussian likelihoods computed on all pairs  $\{Z_i(\mathbf{s}_\alpha, t_\alpha), Z_j(\mathbf{s}_\beta, t_\beta)\}$  such that  $\|\mathbf{s}_\alpha - \mathbf{s}_\beta\| \leq d_m$  and  $|t_\alpha - t_\beta| \leq t_m$ , where  $i, j = 1, \dots, p$  and  $\alpha, \beta = 1, \dots, N$ . On a simulation study mimicking the sampling design of the dataset, it was shown in Bourotte et al. (2016) that, for the smoothness and scale parameters, Pairwise Likelihood (PL) provides estimates with only a slight loss in efficiency as compared to a Full Likelihood (FL) approach, with a significant gain in terms of computation. Moreover, for prediction, the difference between PL and FL is negligible on all tested prediction scores. We have thus decided to estimate the parameters of our model using PL in this same way, since it has already been proven efficient. We have set  $d_m = 250$  km and  $t_m = 2$  days because these values have been shown to minimize the trace of the estimated covariance matrix, thus providing optimal estimation for all parameters (Bourotte et al., 2016). The maximization algorithm will be detailed in Section 5.3 hereinafter.

Two validation stations were selected, neither too close, nor too far from the other stations (Le Rheu and Bourran). The other 11 stations were used to estimate the parameters of both the FULL and the PIT models, reported in Table 1. In order to restrict ourselves to data being stationary in time, we have selected data recorded in January, from 2003 to 2012.

## 5.2 Model specification

Given the positive nonseparability of the residuals, and in light of the discussion in Section 4, the parameter  $\alpha$  in (16) is set to  $\alpha = 0$ , which results in a significant reduction of parameters. The temporal structure is described by a matrix-valued unbounded pseudo-variogram as proposed in (19). Since  $d = 2$  here, the entries of the matrix-valued covariance function thus read:

$$C_{ij}(\mathbf{h}, u) = \frac{\tau_{ij} a_{ij}^{-2\nu_{ij}}}{\beta \eta_{ij}(u) + b_{ij}^2} \mathcal{M} \left( \mathbf{h}; \frac{a_{ij}}{\sqrt{\beta \eta_{ij}(u) + b_{ij}^2}}, \nu_{ij} \right), \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}, \quad (20)$$

with the condition that  $\tau e^{-\nu} / \Gamma(\nu) \succeq 0$ . Setting  $\mathbf{h} = \mathbf{0}$  and  $u = 0$  in (20), one gets the generic entry of the covariance matrix

$$C_{ij}(\mathbf{0}, 0) = \tau_{ij} a_{ij}^{-2\nu_{ij}} (\beta [R_{ij}^0 - R_{ij}(0)] + b_{ij}^2).$$

For identifiability reasons, we further set  $\beta = 1$  and  $b_{ij}^2 = 1 - R_{ij}^0$  with  $\max_i \{R_{ii}(0)\} < 1$ . According to Property 4 in Appendix A, the matrix  $\mathbf{b}^2$  is conditionally definite negative, as required. As a consequence, we get the simpler expression  $C_{ij}(\mathbf{0}, 0) = \tau_{ij} a_{ij}^{-2\nu_{ij}} / [1 - R_{ij}(0)]$ .

The pseudo-variogram  $\boldsymbol{\eta}$  needs to be specified. For ease of comparison with Bourotte et al. (2016), we choose the parametric form also advocated in Gneiting (2002)

$$\gamma_0(u) = (1 + |cu|^{2a})^b - 1, \quad u \in \mathbb{R},$$

with  $c > 0, 0 < a \leq 1, 0 \leq b \leq 1$ . This variogram is reminiscent of the generalized Cauchy covariance model which will also be used for  $\mathbf{R}$ :

$$R_{ij}(u) = A_i A_j (1 + |ru|^{2\lambda})^{-b}, \quad u \in \mathbb{R},$$

with  $r > 0, 0 < \lambda < 1$  and  $0 \leq A_i < 1$  for  $i = 1, \dots, p$ , and where for the sake of parsimony the parameter  $b$  is set identical to the parameter  $b$  in  $\gamma_0$ . Notice that the matrix  $[A_i A_j]_{i,j=1}^p$  is  $p$ -separable and thus positive semidefinite (see Appendix A).  $\mathbf{R}$  is therefore a valid matrix-valued covariance function. The covariance model that we thus obtain is:

$$C_{ij}(\mathbf{h}, u) = \frac{\tau_{ij} a_{ij}^{-2\nu_{ij}}}{(1 + |cu|^{2a})^b - A_i A_j (1 + |ru|^{2\lambda})^{-b}} \times \mathcal{M} \left( \mathbf{h}; \frac{a_{ij}}{[(1 + |cu|^{2a})^b - A_i A_j (1 + |ru|^{2\lambda})^{-b}]^{1/2}}, \nu_{ij} \right).$$

In the above parameterization, the space-time nonseparability parameter  $b$  acts both inside and outside the Matérn function  $\mathcal{M}$ , implying a constant perfect correlation in time if  $b = 0$ . Following Gneiting (2002) and Bourotte et al. (2016), a reparameterization is thus useful. Multiplying the above equation by the temporal matrix-valued covariance function  $\tilde{\rho}(u) = [(1 + c|u|^{2a})^\delta - A_i A_j (1 + r|u|^{2\lambda})^{-\delta}]^{-1}$  with  $0 \leq \delta \leq 1$  leads to the parametric family that will be used in the rest of this section. The parameterization is further simplified by setting  $\tau_{ii} = \sigma_i^2 a_{ii}^{2\nu_{ii}} (1 - A_i^2)^2$  and, for  $i \neq j$ ,  $\tau_{ij} = \rho_{ij} \sqrt{\tau_{ii} \tau_{jj}}$ , with  $[\rho_{ij}]_{i,j=1,\dots,p}$  being a correlation matrix. With this parameterization, the matrix  $\boldsymbol{\tau}$  is positive definite by construction. Finally, the fully multivariate Gneiting-Matérn covariance model reads

$$C_{ij}(\mathbf{h}, u) = \frac{1}{(1 + |cu|^{2a})^\delta - A_i A_j (1 + |ru|^{2\lambda})^{-\delta}} \frac{\rho_{ij} a_{ij}^{-2\nu_{ij}} \sqrt{\tau_{ii} \tau_{jj}}}{(1 + |cu|^{2a})^b - A_i A_j (1 + |ru|^{2\lambda})^{-b}} \times \mathcal{M} \left( \mathbf{h}; \frac{a_{ij}}{[(1 + |cu|^{2a})^b - A_i A_j (1 + |ru|^{2\lambda})^{-b}]^{1/2}}, \nu_{ij} \right), \quad (21)$$

with  $(\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}$  and  $i, j = 1, \dots, p$ .

The model in (21) provides interpretable parameters for a fully nonseparable, i.e. a multivariate Gneiting-Matérn covariance function, which is space-time nonseparable and neither proportional-in-time nor proportional-in-space:

- The collocated covariance matrix  $[C_{ij}(\mathbf{0}, 0)]_{i,j=1,\dots,p}$  has elements

$$C_{ij}(\mathbf{0}, 0) = \rho_{ij} \sigma_i \sigma_j \frac{a_{ii}^{\nu_{ii}} a_{jj}^{\nu_{jj}} (1 - A_i^2)(1 - A_j^2)}{a_{ij}^{2\nu_{ij}} (1 - A_i A_j)^2}, \quad (22)$$

which simplify to  $C_{ii}(\mathbf{0}, 0) = \sigma_i^2$  if  $i = j$ . This is the elementwise product of a covariance matrix  $[\rho_{ij} \sigma_i \sigma_j]_{i,j=1,\dots,p}$  with a weight matrix whose elements are ratios combining the other

parameters. Notice that the quantities  $\sigma_i^2 a_{ii}^{2\nu_{ii}}$  appearing in these ratios are precisely the micro-ergodic quantities that can be consistently estimated by Maximum Likelihood under Gaussian assumption in an infill asymptotics setting (Zhang, 2004).

- The vector  $\mathbf{A} = [A_i]_{i=1}^p$  describes the multivariate aspect of the temporal covariance function. When  $A_i = 0$  for  $i = 1, \dots, p$ , the above model reduces to the model in Bourotte et al. (2016, Equation 9), up to a slight reparameterization.
- The temporal parameters are, on the one hand,  $(a, c)$  that characterize the common part of the temporal covariance function and, on the other hand,  $(r, \lambda, \mathbf{A})$  that characterize the multivariate part of it. In addition, the parameters  $(b, \delta)$  act on both parts of the temporal covariance function.
- The spatial parameters are the scale matrix  $\mathbf{a}$ , with  $\mathbf{a}^2 \preceq_c 0$ , and the smoothness matrix  $\boldsymbol{\nu}$ , which must verify  $\mathbf{a}^2 - \boldsymbol{\nu} \preceq_c 0$ . For simplicity, and for a fair comparison to Bourotte et al. (2016),  $\mathbf{a}^2$  and  $\boldsymbol{\nu}$  will be chosen as  $a$ -separable. The two above conditions on  $\mathbf{a}$  and  $\mathbf{a}^2 - \boldsymbol{\nu}$  are thus always verified.
- The parameter  $b$ , which appears in both the spatial and temporal terms, is the separability parameter. It can vary from  $b = 0$  (separability) to  $b = 1$  (full spatio-temporal nonseparability).

### 5.3 Estimation of the parameters

The data analyzed here are standardized residuals after centering for spatial and temporal trends. We thus set  $\sigma_i = 1$ , for  $i = 1, \dots, p$ . There is a total of  $p(p + 2) + 6$  parameters to estimate, where  $p$  is the number of variables. Here  $p = 3$ , leading to a total of 21 parameters. Compared to the model in Bourotte et al. (2016), our more flexible model contains only 6 additional parameters describing the multivariate nature of the temporal covariance. Let us denote  $\theta$  the vector of all parameters, and let  $\Theta$  be the space of parameters for which the model (21) is valid. The optimization of the PL is not straightforward. First, the space  $\Theta$  has a complex shape owing to the semidefinite positiveness condition on  $\boldsymbol{\tau}e^{-\boldsymbol{\nu}}/\Gamma(\boldsymbol{\nu})$ . Checking this condition at every call of the function computing the PL prevents the optimization functions to work properly. It was thus decided to consider a hyper-box shape for  $\Theta$  and check the condition at convergence. It was found that the semidefinite positiveness condition was always verified. Second, employing blindly an optimization function to maximize the PL in the high-dimensional space  $\Theta$  is likely to fail. As a way to alleviate this problem, PL was maximized sequentially in subspaces of  $\Theta$  corresponding to blocks of related parameters, while keeping all other parameters fixed to the values previously attained.

The following procedure builds on the procedure presented in [Bourotte et al. \(2016\)](#), which was proven to be efficient.

1. The  $p \times p$  marginal empirical covariance matrix with elements  $\hat{c}_{ij}$ ,  $i, j = 1, \dots, p$  is computed.
2. The separability parameter  $b$ , known to be difficult to estimate, is successively fixed to  $0, 0.1, \dots, 1$ .
3. For every fixed value of  $b$ :
  - (a) Initial temporal parameters  $(c, a, r, \lambda, \delta, \mathbf{A})$  are estimated by maximizing the PL for the temporal covariance model  $\mathbf{C}_T$ . Using (21) and (22), one gets

$$C_{T,ij}(u) = \frac{1}{(1 + |cu|^{2a})^\delta - A_i A_j (1 + |ru|^{2\lambda})^{-\delta}} \frac{\hat{c}_{ij} (1 - A_i A_j)^2}{(1 + |cu|^{2a})^b - A_i A_j (1 + |ru|^{2\lambda})^{-b}}.$$

- (b) Using the estimates  $\hat{\mathbf{A}}$  from the previous step, initial spatial parameters  $(\mathbf{a}, \boldsymbol{\nu})$  are estimated by maximizing the PL for the spatial matrix-valued covariance  $\mathbf{C}_S$  with entries

$$C_{S,ij}(\mathbf{h}) = \hat{c}_{ij} \mathcal{M} \left( \mathbf{h}; \frac{a_{ij}}{(1 - \hat{A}_i \hat{A}_j)^{1/2}}, \nu_{ij} \right).$$

- (c) The estimates  $\hat{\rho}_{ij}$  are computed using (22), given  $\hat{\mathbf{a}}, \hat{\boldsymbol{\nu}}, \hat{\mathbf{A}}$ .
  - (d) The temporal parameters are updated, considering all other parameters fixed, by maximizing the PL of the multivariate spatio-temporal model (21).
  - (e) The spatial parameters are updated, considering all other parameters fixed by maximizing the PL of the multivariate spatio-temporal model (21).
  - (f) Step (c)-(e) are iterated until a stopping criterion is reached and the positive semi-definiteness of  $\boldsymbol{\tau} e^{-\boldsymbol{\nu}} / \Gamma(\boldsymbol{\nu})$  is checked.

Iterations are stopped when PL is increased by less than 0.5 unit after one iteration from (c) to (e). Finally  $\hat{b}$  is the value of  $b$  among  $\{0, 0.1, \dots, 0.9, 1\}$  corresponding to the highest maximized PL. To perform the maximization in the subspaces of  $\Theta$ , we used the package `nlmminb` implemented in R with lower and upper bounds for the parameters when mathematically necessary. In addition, for the sake of numerical stability, the smoothness parameter  $\nu$  was upper bounded at 6.0.

Figure 2(left) shows the maximum log-PL as a function of the separability parameter  $b$  for the fully multivariate Gneiting-Matérn model, referred to as FULL. The maximum is reached for  $\hat{b} = 0.1$  with  $\log \text{PL}_{\text{FULL}}(\hat{\theta}_{b=0.1}) = -615135.3$ . For comparison purpose, PL was also maximized for the PIT model, i.e. when  $\mathbf{A} = \mathbf{0}$ . For this simpler model,  $\log \text{PL}_{\text{PIT}}(\hat{\theta}_b)$  is monotonically decreasing from  $-615635.2$  to  $-615701.5$  as  $b$  varies from 0 to 1. The difference between the maximum PL for both models

		Temporal parameters									
		$c$	$a$	$r$	$\lambda$	$\delta$	$A_1$	$A_2$	$A_3$		
FULL		0.098	0.999	0.686	0.796	1.000	0.946	0.822	0.802		
PIT		2.016	0.699	N/A	N/A	0.971	0	0	0		
		Spatial parameters									
		$a_1$	$a_2$	$a_3$	$\nu_1$	$\nu_2$	$\nu_3$				
FULL		13.5 $10^{-3}$	4.0 $10^{-3}$	16.7 $10^{-3}$	6.000	0.566	0.722				
PIT		37.0 $10^{-3}$	7.8 $10^{-3}$	34.3 $10^{-3}$	4.774	0.671	1.009				
		Correlation parameters						PL			
		$\rho_{12}$	$\rho_{13}$	$\rho_{14}$							
FULL		-0.066	-0.926	-0.112				-615 135.3			
PIT		-0.082	-0.278	-0.114				-615 635.2			

Table 1: PL parameter estimates for the weather dataset with  $d_m = 250$  km,  $t_m = 2$  days. FULL: fully nonseparable model as in (21) with  $b = 0.1$ . PIT: simplified Proportional-In-Time model with  $b = 0$ .

is thus approximately equal to 500 units, indicating strong evidence in favor of the more complex model, considering that the number of data is moderate with a total of 10 independent repetitions of 1023 correlated data. As a point of comparison, if all the data were independent, the Bayesian Information Criterion penalization to the full likelihood of the more complex model would be equal to  $2.5 \ln(10230) \simeq 23$ .

Looking at the temporal marginals, the models show very different behaviors – as we expected. Figure 2(right) shows the temporal empirical covariances for variables R and H with their fitted models. Thanks to the flexibility of the FULL model, the temporal covariance functions are able to fit the empirical values for both variables. In contrast, the unique covariance function of the PIT model lies somewhere between the empirical covariance of the two variables, thus being unable to provide a satisfactory fit to any of the two variables. Table 1 reports the estimated parameters maximizing PL for both models. It is interesting to note that the parameter  $c$  of  $\gamma_0(h)$  is close to 0 in the FULL model, thereby showing that the pseudo-variogram of the FULL model is mainly driven by its multivariate part. It must also be noted that the parameter  $\delta$  is equal or close to 1 in both cases. Figure 3 shows the spatial (at time lag  $u = 0$ ) and spatio-temporal covariances (at time lags  $u = 1, 2$ ) for all variables and pairs of variables. Experimental covariances between all pairs of stations and the FULL model with estimated parameters from Table 1 are shown. Overall, the fit is good. On the top left panel representing the covariance function for variable R, the very high smoothness of the covariance function near the origin is clearly visible. The spatial smoothness parameter is indeed larger for R, and even equal to the upper bound for the FULL model.

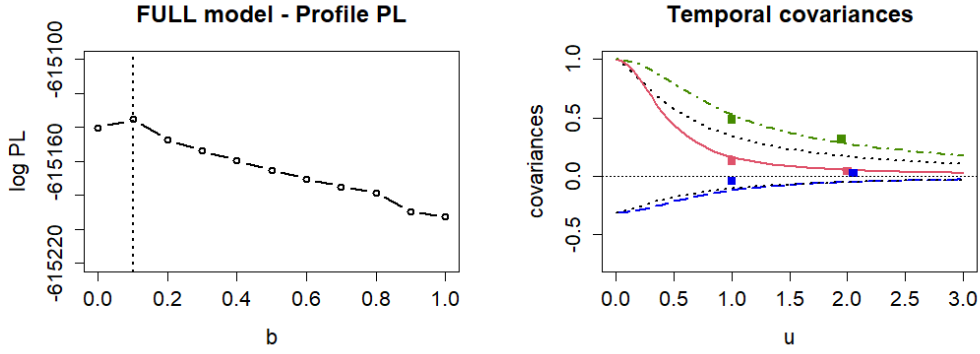


Figure 2: Left: PL as a function of the space-time separability parameter  $b$ . Right: (R,H) temporal correlation functions. PIT model with unique direct correlation and cross-correlation in dotted black lines. FULL model with direct correlation for R and H in solid and longdashed lines, respectively, and cross-correlation in dotdashed line. Squares: empirical correlations.

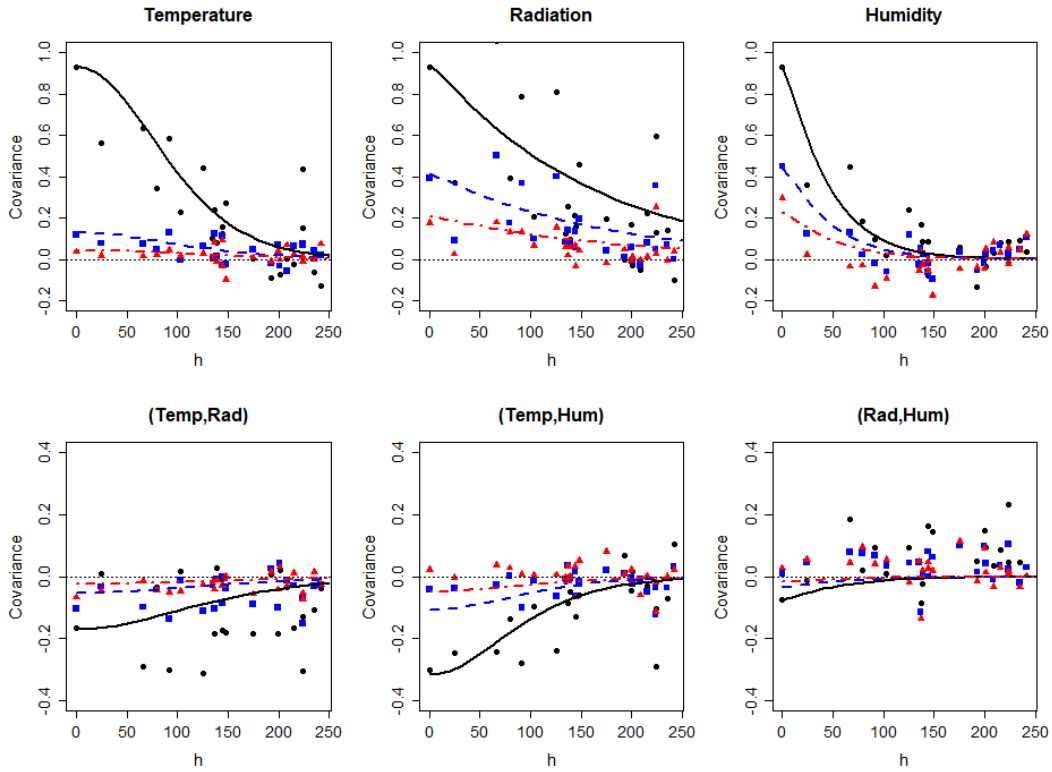


Figure 3: Empirical and fitted spatial-temporal correlation at  $u = 0$  (black solid lines and circles),  $u = 1$  (blue dashed lines and squares) and  $u = 2$  (red dotdashed lines and triangles).

## 5.4 Prediction

We considered two different prediction settings. In the first setting, prediction of the three variables at a validation station at day  $t$  is made knowing the trivariate data at the 11 estimation stations at day  $t$ ,  $t - 1$  and  $t - 2$ . This setting is called *Spatial Interpolation* because data from the same day at different location are used (as well as data from previous days). In the second setting, prediction is made with data from days  $t - 1$  and  $t - 2$  only at all 13 stations (estimation and validation station). In this second setting, called *Temporal Prediction*, purely temporal covariances are thus introduced. The first setting corresponds to interpolation at an ungauged location while the second setting corresponds to the classical one-day ahead forecast. Predictions are computed from January 3rd to January 31st since the two previous days must be used as conditioning data.

In the Gaussian framework, the conditional expectation is the best linear unbiased prediction, also called (co)kriging in the geostatistics literature, while the conditional covariance matrix is the covariance matrix of the cokriging errors. The reader is referred to (Chilès and Delfiner, 2012) for a detailed exposition.

The two models (FULL and PIT) were compared by means of four different scores: the Root Mean Square Error (RMSE), the Mean Absolute Error (MAE), the Continuous Ranked Probability Score (CRPS) and the Logarithmic Score (LogS) (Gneiting and Raftery, 2007). The first two scores, RMSE and MAE, compare the conditional expectation to the true value. The other two scores, CRPS and LogS, assess not only the prediction but its variance as well. The CRPS measures the discrepancy between the predictive cumulative distribution function and the true value. The marginal logarithmic score (LogS) is the opposite of the logarithm of the marginal predictive density at the true value. They are easily computed in the case of a normal predictive distribution and details can be found in Bourotte et al. (2016). Notice that the conditional variances are identical for all days since the configuration of the conditioning data remains identical.

Table 2 reports our results. As a benchmark, we also report the scores obtained with a trivial prediction where all predicted values are set to the expectation (equal to 0) and all variances set to the theoretical variance (equal to 1). Recall that lower scores indicate a better adequacy between the model and the data, that the first three scores are bounded from below by 0 whilst LogS is unbounded.

Overall, both settings show better (sometimes comparable) performances for the FULL model as compared to the PIT model. Specifically, for variables R and T, Spatial Interpolation is more accurate than Temporal Prediction, which in turn is more accurate than the trivial prediction. This is in accordance with the fact that more conditioning data are available in the Spatial Interpolation setting. Moreover, the FULL model brings an interesting improvement over PIT for the scores involving the conditional variance, particularly in the Temporal Prediction setting. In contrast, for variable H,



the situation is reverse: Temporal Prediction is more accurate than Spatial Interpolation, which indicates that the temporal dependencies are much more important than the spatial ones for this variable. Notice also that FULL improves over PIT for all scores in the Temporal Prediction setting, in accordance with the fact that the FULL model is able to better fit the data thanks to the extra parameters in the temporal matrix-valued covariance function.

Finally, Figure 4 shows the predictions of the three variables from January 3<sup>rd</sup> to January 31<sup>st</sup> 2003 following the Spatial Interpolation setting. The predicted residuals were back-transformed to the original scales to be compared to the real values. An envelope of  $\pm 2$  conditional standard deviations was added to the plot.

	RMSE			MAE			CRPS			LogS		
	R	T	H	R	T	H	R	T	H	R	T	H
Spatial interpolation												
FULL	<b>0.746</b>	<b>0.582</b>	<b>0.864</b>	<b>0.587</b>	0.464	0.732	<b>0.427</b>	<b>0.335</b>	<b>0.509</b>	<b>0.268</b>	0.015	<b>0.393</b>
PIT	0.751	0.584	<b>0.864</b>	0.593	<b>0.462</b>	<b>0.731</b>	0.430	<b>0.335</b>	<b>0.509</b>	0.271	<b>0.010</b>	<b>0.393</b>
Temporal prediction												
FULL	<b>0.919</b>	0.837	<b>0.813</b>	0.784	<b>0.706</b>	<b>0.667</b>	<b>0.531</b>	<b>0.482</b>	<b>0.470</b>	<b>0.421</b>	<b>0.329</b>	<b>0.311</b>
PIT	0.933	<b>0.835</b>	0.825	<b>0.780</b>	0.714	0.684	0.537	0.483	0.479	0.437	0.331	0.330
Trivial interpolation												
	0.933	0.893	0.919	0.801	0.777	0.782	0.540	0.521	0.538	0.439	0.403	0.443

Table 2: RMSE, MAE, CRPS and LogS of predicted values at the validation stations at day  $t$ , using ML estimates. FULL: fully nonseparable model as in (21) with  $b = 0.1$ . PIT: simplified Proportional-In-Time model with  $b = 0$ . The best scores among FULL or PIT are shown in bold font.

## 6 Concluding remarks

We have proposed new parametric classes of matrix-valued covariance functions for multivariate spatio-temporal random fields, where each component has its own smoothness and scale parameter in space and its own correlation function in time. We have also designed computationally efficient algorithms for simulating Gaussian random fields with these covariance functions, which are not restricted in the number of target spatial coordinates and allow tens of thousands of time coordinates, hence are applicable to large-scale space-time prediction problems. Our constructions generalize the Gneiting class of space-time covariance functions and are fully nonseparable, in the sense that they are space-time nonseparable and that they are neither proportional-in-space nor proportional-in-time. This is a major improvement to the models proposed up to date where the marginal temporal correlation function is identical for all the variables, which is overly restrictive.

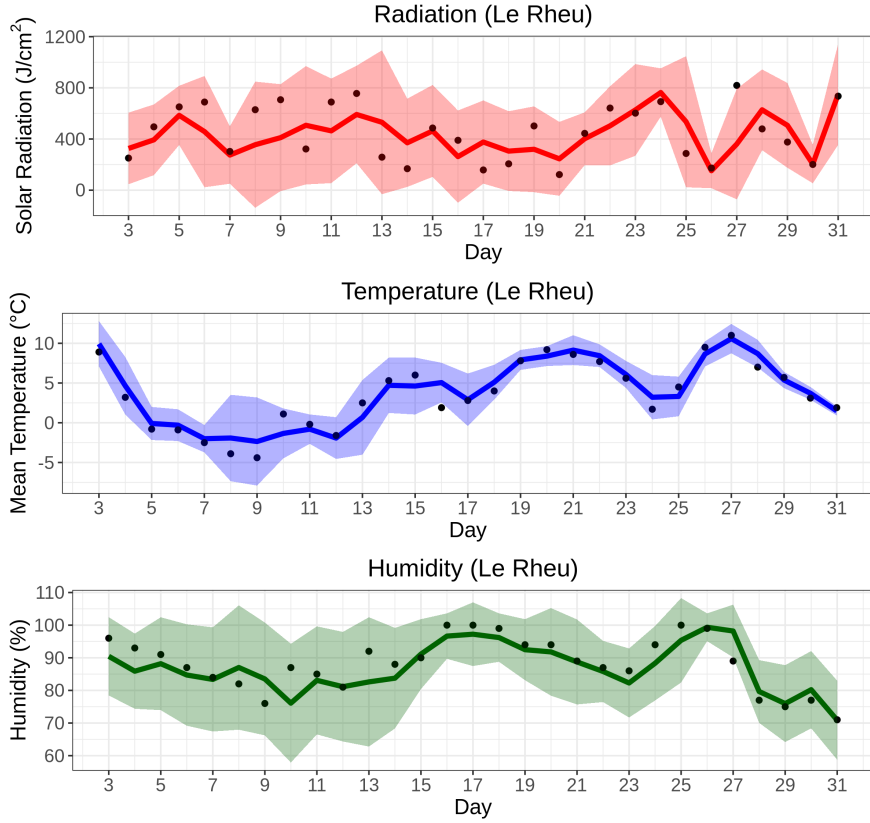


Figure 4: Predictions of R, T and H at Le Rheu from January 3<sup>rd</sup> to January 31<sup>st</sup> 2003 following the *Spatial Interpolation* setting. Predictions (continuous line), real values (dots) and envelope of  $\pm 2$  conditional standard deviations (colored ribbon).

The main class proposed in Section 4 allows for a general behavior for the spatial scale parameter as a function of the time lag  $u$ , depending on the values taken by the parameters. It is well known that the Gneiting class of models is positively nonseparable. A very original feature of our construction is that the models can be negatively nonseparable or asymptotically separable as  $|u| \rightarrow \infty$ .

Since these models are continuous over  $\mathbb{R}^d \times \mathbb{R}$ , their application is not limited to data organized as regular records at measurement stations. More complex designs are in principle possible, such as irregular or incomplete records, heterotopic designs or even mobile measurement devices. We hope to see applications of our models to such dataset in a near future.

From a statistical point of view, the main challenge is to find the right balance between the flexibility of the model, which must be able to account for the complex interactions between space and time across several variables, and the number of parameters which should remain interpretable and not too numerous. More complex models should only be preferred if strongly supported by the data. The dataset analyzed in this work is of moderate size with 10 independent repetitions

of  $11 \times 31 = 341$  space-time trivariate data. Its analysis has shown that our model improves the fitting and the predictive performances over the more parsimonious model in [Bourotte et al. \(2016\)](#), in particular when the time dimension is important for prediction. In [Section 5.3](#), it was shown that the maximum PL difference is around 500. Combined with the improved prediction scores, this provides strong evidence in favor of our new model for this dataset. The analysis of other datasets, in particular with more spatial locations, is necessary for confirming this first result.

Estimating the parameters of a parametric model such as the one in [Section 5](#) is a challenge. Here, we have used Pairwise Likelihood because it was proven efficient on the proportional-in-time model in [Bourotte et al. \(2016\)](#). It is however relatively slow since on a recent laptop computer reaching convergence with our optimization algorithm and the function `nlmbinb` takes typically a couple of hours. There is thus certainly room for improvement for a an easier use of such complex models. At the cost of even larger computing times, composite likelihood with larger sets of elements than pairs might be an interesting option for a more accurate estimation of the parameters, considering the complex interactions between space, time and the variables. Conducting a thorough comparison of several composite likelihoods was out of the scope of this work and left for future research. In our opinion, more interesting improvements can be expected by using Vecchia’s Gaussian process approximation ([Katzfuss and Guinness, 2021](#)). We are looking forward to (multivariate) spatio-temporal versions of the `GPVecchia` package ([Katzfuss et al., 2020](#)).

In [Appendix C](#) we have proposed even more general constructions where both the spatial scale and spatial smoothness parameters can vary with the time lag  $u$ . Also, all the models presented in this work can easily be generalized to  $\mathbb{R}^d \times \mathbb{R}^\ell$ , where  $\ell$  is an integer greater than one. Other constructions based on characterizations established in [Porcu et al. \(2022\)](#) are also possible. Extensions of our models to other spaces that are of interest in environmental sciences, atmospheric sciences and geosciences, such as a sphere crossed with an Euclidean space ( $\mathbb{S}^d \times \mathbb{R}^\ell$ ), are left to further research.

## Acknowledgments

X. Emery acknowledges the funding of the National Agency for Research and Development of Chile, through grants ANID/FONDECYT/REGULAR/No. 1210050 and ANID PIA AFB180004.

## A Technical lemmas and properties

**Lemma 1.** *Let  $\mathbf{a} = [a_{ij}]_{i,j=1}^p$  be a real, symmetric conditionally negative semidefinite matrix. Then, there exist random variables  $Y_1, \dots, Y_p$  such that:*

$$a_{ij} = \frac{a_i + a_j}{2} + \frac{1}{2} \text{Var}(Y_i - Y_j), \quad i, j = 1, \dots, p.$$

*Proof.* The matrix  $\mathbf{b} = [b_{ij}]_{i,j=1}^p$  with  $b_{ij} = a_{ip} + a_{pj} - a_{ij} - a_{pp}$  is positive semidefinite (Berg et al., 1984, Chapter 3, Lemma 2.1). Let  $Y_1, \dots, Y_p$  be Gaussian random variables with variance-covariance matrix  $\mathbf{b}$ . Then

$$\frac{1}{2} \text{Var}(Y_i - Y_j) = \frac{b_{ii} + b_{jj}}{2} - b_{ij} = a_{ij} - \frac{a_{ii} + a_{jj}}{2}, \quad i, j = 1, \dots, p.$$

□

**Property 1.** *A matrix is conditionally null definite if and only if it is a-separable.*

**Property 2.** *All p-separable matrices with nonnegative entries are positive semidefinite.*

**Property 3.** *If  $\mathbf{a} \preceq_c 0$ , then  $\exp(-t\mathbf{a}) \succeq 0$  for all  $t \geq 0$ .*

**Property 4.** *If  $\mathbf{a} \preceq_c 0$  and  $\mathbf{b}$  is a-separable, with  $\mathbf{a}$  and  $\mathbf{b}$  of size  $p \times p$ , then  $\mathbf{a} + \mathbf{b} \preceq_c 0$  and  $\mathbf{a} - \mathbf{b} \preceq_c 0$ . In particular,  $\mathbf{1} - \mathbf{b} \preceq_c 0$ , where  $\mathbf{1}$  is the all-ones matrix of size  $p \times p$ .*

*Proof.* The fourth property derives from the first one, which can be proven by using Lemma 1, and from the fact that the sum of two conditionally negative (positive) semidefinite matrices is still conditionally negative (positive) semidefinite. For a proof of the second and third properties, see Berg et al. (1984, Chapter 3, Property 1.9 and Theorem 2.2). □

**Lemma 2.** *Let  $\mathbf{C} : u \mapsto [C_{ij}(u)]_{i,j=1}^p$ ,  $u \in \mathbb{R}$ , be the matrix-valued covariance function of a p-variate second-order stationary random field  $\mathbf{Y}$  in  $\mathbb{R}$ . Then,  $\forall n_T \in \mathbb{N}^*$ ,  $\forall t_1, \dots, t_{n_T} \in \mathbb{R}$ , the matrix  $\mathbf{C}_0 = [[C_{ij}(t_m - t_n)]_{i,j=1}^p]_{m,n=1}^{n_T}$  is positive semidefinite.*

*Proof.* It is immediate, as  $\mathbf{C}_0$  is the variance-covariance matrix of the random vector  $[\mathbf{Y}(t_1), \dots, \mathbf{Y}(t_{n_T})]$ . □

**Lemma 3.** *Let  $\mathbf{a} = [a_{ij}]_{i,j=1}^p$  be a conditionally negative semidefinite matrix and  $\boldsymbol{\eta} : u \mapsto [\eta_{ij}(u)]_{i,j=1}^p$ ,  $u \in \mathbb{R}$ , be the matrix-valued pseudo-variogram of an intrinsically stationary multivariate random field  $\mathbf{Y} = [Y_i]_{i=1}^p$  in  $\mathbb{R}$ . Then,  $\forall n_T \in \mathbb{N}^*$ ,  $\forall t_1, \dots, t_{n_T} \in \mathbb{R}$ , the matrix  $\boldsymbol{\eta}_0 = [[a_{ij} + \eta_{ij}(t_m - t_n)]_{i,j=1}^p]_{m,n=1}^{n_T}$  is conditionally negative semidefinite.*

*Proof.* Let  $\boldsymbol{\Lambda} = [[\lambda_{m,i}]_{i=1}^p]_{m=1}^{n_T}$  be a real vector with entries adding to zero. Define  $\tilde{\boldsymbol{\Lambda}} = [\tilde{\lambda}_i]_{i=1}^p$  with  $\tilde{\lambda}_i = \sum_{m=1}^{n_T} \lambda_{m,i}$  and  $\tilde{\mathbf{C}}_0 = [[C_{i,j;m,n}]_{i,j=1}^p]_{m,n=1}^{n_T}$  with  $C_{i,j;m,n} = \text{Cov}(Y_i(t_m) - Y_1(t_1), Y_j(t_n) - Y_1(t_1))$ .

Then:

$$\begin{aligned}
& \sum_{m=1}^{n_T} \sum_{n=1}^{n_T} \sum_{i=1}^p \sum_{j=1}^p \lambda_{m,i} \lambda_{n,j} [a_{ij} + \eta_{ij}(t_m - t_n)] \\
&= \sum_{m=1}^{n_T} \sum_{n=1}^{n_T} \sum_{i=1}^p \sum_{j=1}^p \lambda_{m,i} \lambda_{n,j} [a_{ij} + \frac{1}{2} \text{Var}(Y_i(t_m) - Y_1(t_1) + Y_1(t_1) - Y_j(t_n))] \\
&= \sum_{m=1}^{n_T} \sum_{n=1}^{n_T} \sum_{i=1}^p \sum_{j=1}^p \lambda_{m,i} \lambda_{n,j} [a_{ij} + \eta_{i1}(t_m - t_1) + \eta_{j1}(t_n - t_1) - \text{Cov}(Y_i(t_m) - Y_1(t_1), Y_j(t_n) - Y_1(t_1))] \\
&= \left( \sum_{n=1}^{n_T} \sum_{j=1}^p \lambda_{n,j} \right) \left( \sum_{m=1}^{n_T} \sum_{i=1}^p \lambda_{m,i} \eta_{i1}(t_m - t_1) \right) + \left( \sum_{m=1}^{n_T} \sum_{i=1}^p \lambda_{m,i} \right) \left( \sum_{n=1}^{n_T} \sum_{j=1}^p \lambda_{n,j} \eta_{j1}(t_n - t_1) \right) \\
&+ \sum_{i=1}^p \sum_{j=1}^p \tilde{\lambda}_i \tilde{\lambda}_j a_{ij} - \sum_{m=1}^{n_T} \sum_{n=1}^{n_T} \sum_{i=1}^p \sum_{j=1}^p \lambda_{m,i} \lambda_{n,j} C_{i,j;m,n}.
\end{aligned}$$

Since the entries of  $\mathbf{A}$  add to zero, the first two terms in the last equality are zero, while the last two terms are negative or zero since, on the one hand, the entries of  $\tilde{\mathbf{A}}$  add to zero and  $\mathbf{a}$  is conditionally negative semidefinite and, on the other hand,  $\tilde{\mathbf{C}}_0$  is a variance-covariance matrix, hence positive semidefinite.  $\square$

## B Proofs

*Proof of Theorem 1.* The proof is constructive and based on an extension of the substitution approach proposed in Allard et al. (2020). Consider the  $p$ -variate space-time random field  $\mathbf{Z} = [Z_i]_{i=1}^p$  defined as

$$Z_i(\mathbf{s}, t) = \sqrt{\frac{2f_i(R)}{f(R)}} V_i \cos \left( \sqrt{2R} \langle \mathbf{\Omega}, \mathbf{s} \rangle + \frac{|\mathbf{\Omega}|}{\sqrt{2}} W_i(t) + \Phi \right), \quad (\mathbf{s}, t) \in \mathbb{R}^d \times \mathbb{R}, \quad (23)$$

where  $R$  is a nonnegative random variable with a probability density  $f$  whose support contains that of  $f_1, \dots, f_p$ ,  $\mathbf{V} = [V_i]_{i=1}^p$  is a centered Gaussian random vector with covariance matrix  $\boldsymbol{\sigma} = [\sigma_{ij}]_{i,j=1}^p$ ,  $\mathbf{\Omega}$  is a standard Gaussian random vector with  $d$  independent components,  $\Phi$  is a uniform random variable in  $(0, 2\pi)$ , and  $W_i, i = 1, \dots, p$ , are random fields on  $\mathbb{R}$  whose direct and cross-increments are second-order stationary as in (1), characterized by the matrix-valued pseudo-variogram  $\boldsymbol{\eta}$ , and Gaussian. The random variable  $R$ , the random vectors  $\mathbf{\Omega}$ ,  $\Phi$  and  $\mathbf{V}$  and the random field  $\mathbf{W} = [W_i]_{i=1}^p$  are independent.

Following Allard et al. (2020, Theorem 3), it is straightforward to show that, for any  $i, j = 1, \dots, p$ ,  $Z_i$  and  $Z_j$  have a zero mean and that the covariance function between  $Z_i$  and  $Z_j$  is

$$\begin{aligned}
\mathbb{E}[Z_i(\mathbf{s}, t) Z_j(\mathbf{s}', t')] &= \frac{\sigma_{ij}}{(2\pi)^{d/2}} \int_0^\infty \int_{\mathbb{R}^d} \sqrt{f_i(r) f_j(r)} \cos(\sqrt{2r} \langle \boldsymbol{\omega}, \mathbf{s} - \mathbf{s}' \rangle) \exp \left( -|\boldsymbol{\omega}|^2 \frac{\eta_{ij}(t - t') + 1}{2} \right) d\boldsymbol{\omega} dr \\
&= \frac{\sigma_{ij}}{(\eta_{ij}(t - t') + 1)^{d/2}} \varphi_{ij} \left( \frac{|\mathbf{s} - \mathbf{s}'|^2}{\eta_{ij}(t - t') + 1} \right), \quad (\mathbf{s}, t) \in \mathbb{R}^d \times \mathbb{R}, (\mathbf{s}', t') \in \mathbb{R}^d \times \mathbb{R}.
\end{aligned}$$

The cross-covariance between  $Z_i$  and  $Z_j$  depends only on the spatial and temporal lags  $\mathbf{h} = \mathbf{s} - \mathbf{s}'$  and  $u = t - t'$ . Accordingly, the random field  $\mathbf{Z}$  is second-order stationary with zero mean and covariance function  $\mathbf{C} : (\mathbf{h}, u) \mapsto [C_{ij}(\mathbf{h}, u)]_{i,j=1}^p$ , with  $C_{ij}(\mathbf{h}, u)$  given as in (11).  $\square$

*Proof of Theorem 2.* Without loss of generality, let us consider a regular design of  $n_S \times n_T$  space-time coordinates  $(\mathbf{s}_1, t_1), \dots, (\mathbf{s}_{n_S}, t_1), \dots, (\mathbf{s}_{n_S}, t_{n_T}) \in \mathbb{R}^d \times \mathbb{R}$ . Define  $u_{mn} = t_m - t_n$ ,  $\mathbf{h}_{kl} = \mathbf{s}_k - \mathbf{s}_l$  and  $\mathbf{M} = [[[\sigma_{ij}(u_{mn}) \mathcal{M}(\mathbf{h}_{kl}; r_{ij}(u_{mn}), \nu_{ij})]_{k,l=1}^{n_S}]_{i,j=1}^p]_{m,n=1}^{n_T}$ . For  $\mathbf{C}$  to be a positive semidefinite function, we need to show that  $\mathbf{M}$  is positive semidefinite for any such design in  $\mathbb{R}^d \times \mathbb{R}$ . This can be done by viewing the  $p$ -variate spatio-temporal random field at  $n_T$  time coordinates as a  $q$ -variate spatial random field, with  $q = p n_T$ , and using the sufficient validity conditions for multivariate spatial Matérn models given in Emery et al. (2022). Specifically, if one sets  $\boldsymbol{\psi}(u) = \beta \boldsymbol{\eta}(u) + \mathbf{b}^2$ , under the conditions of Theorem 2, the  $q \times q$  matrices  $[[\psi_{ij}(u_{mn})]_{i,j=1}^p]_{m,n=1}^{n_T}$  and  $[[r_{ij}^2(u_{mn}) \psi_{ij}(u_{mn}) - \nu_{ij}]_{i,j=1}^p]_{m,n=1}^{n_T}$  are conditionally negative semidefinite (see Lemma 3 in Appendix A). Furthermore, the matrix

$$\left[ [\sigma_{ij}(u_{mn}) e^{-\nu_{ij}} r_{ij}(u_{mn})^{2\nu_{ij}} \psi_{ij}(u_{mn})^{\nu_{ij}+d/2} / \Gamma(\nu_{ij})]_{i,j=1}^p \right]_{m,n=1}^{n_T} = [[\rho_{ij}(u_{mn})]_{i,j=1}^p]_{m,n=1}^{n_T}$$

is positive semidefinite (see Lemma 2 in Appendix A). Hence, according to Emery et al. (2022, Theorem 2B), the matrix-valued function  $\mathbf{h} \mapsto \sigma_{ij}(u_{mn}) \mathcal{M}(\mathbf{h}; r_{ij}(u_{mn}), \nu_{ij})$  is the covariance function of a  $q$ -variate random field  $\mathbf{Y} = [Y_v]_{v=1}^q$  in  $\mathbb{R}^d$ .  $\mathbf{M}$  is therefore positive semidefinite since it is the variance-covariance matrix of the random vector  $[[Y_v(\mathbf{s}_k)]_{k=1}^{n_S}]_{v=1}^q$ .  $\square$

*Proof of Corollary 1.* In Theorem 2,  $\boldsymbol{\rho}(u)$  must be a matrix-valued covariance function. Here, we set  $\boldsymbol{\rho}(u) = \boldsymbol{\rho}$ , for all  $u \in \mathbb{R}$ . By comparing (16) with (15), it is thus required that  $\boldsymbol{\tau} e^{-\nu} / \Gamma(\nu) = \boldsymbol{\rho}$  is a positive semidefinite matrix.  $\square$

## C General forms of Gneiting-Matérn covariances with temporal-lag dependent parameters

The following theorems provide sufficient validity conditions for a nonseparable multivariate space-time covariance model in which the spatial direct and cross-covariances belong to the Matérn family, with collocated correlation coefficients, scale and/or smoothness factors that depend on the temporal lag. The proof follows the line of that of Theorem 2 and are based on the sufficient validity conditions for multivariate spatial Matérn models given in Emery et al. (2022, Theorems 1, 2, 3), Du et al. (2012, Theorem 3), Apanasovich et al. (2012, Theorem 1) and Gneiting et al. (2010, Theorem 1). Also, the spectral algorithm 3 can be adapted to simulate space-time random fields having such a covariance.

**Theorem 3.** *Let  $d, p$  be positive integers,  $r$  and  $\nu$  be positive real numbers. Let  $\boldsymbol{\mu}_0, \boldsymbol{\nu}_0$  and  $\boldsymbol{\psi}_0$  be  $p \times p$  symmetric conditionally negative semidefinite matrices, the first two ones with positive entries*

and the last one with nonnegative entries. Also, let  $u \mapsto \boldsymbol{\mu}(u) - \boldsymbol{\mu}_0$ ,  $u \mapsto \boldsymbol{\nu}(u) - \boldsymbol{\nu}_0$  and  $u \mapsto \boldsymbol{\psi}(u) - \boldsymbol{\psi}_0$  be  $p \times p$  matrix-valued pseudo-variograms in  $\mathbb{R}$  and  $u \mapsto \boldsymbol{\rho}(u)$  be a  $p \times p$  matrix-valued covariance in  $\mathbb{R}$ . Then, the matrix-valued function  $\mathbf{C} : (\mathbf{h}, u) \mapsto \boldsymbol{\sigma}(u) \mathcal{M}(\mathbf{h}; \mathbf{r}(u), \boldsymbol{\nu}(u))$  with generic entry

$$C_{ij}(\mathbf{h}, u) = \sigma_{ij}(u) \mathcal{M}(\mathbf{h}; r_{ij}(u), \nu_{ij}(u)), \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}, \quad i, j = 1, \dots, p,$$

with  $\mathcal{M}$  the Matérn covariance defined in (3), is positive semidefinite in  $\mathbb{R}^d \times \mathbb{R}$  if one of the following sets of sufficient conditions holds:

$$(A) \quad \mathbf{r}(u) = \sqrt{\frac{\boldsymbol{\psi}(u) + \boldsymbol{\nu}(u)}{\boldsymbol{\mu}(u)}}} \quad \text{and} \quad \boldsymbol{\sigma}(u) = \boldsymbol{\rho}(u) \mathbf{r}(u)^{-2\nu(u)} \Gamma(\boldsymbol{\nu}(u)) \exp(\boldsymbol{\nu}(u)) \boldsymbol{\mu}(u)^{-\nu(u) - d/2}$$

or

$$(B) \quad \mathbf{r}(u) = r \sqrt{\boldsymbol{\psi}(u) + \boldsymbol{\nu}(u)} \quad \text{and} \quad \boldsymbol{\sigma}(u) = \boldsymbol{\rho}(u) \Gamma(\boldsymbol{\nu}(u)) \exp(\boldsymbol{\nu}(u)) \left( \frac{\mathbf{r}(u)}{r} \right)^{-2\nu(u)}$$

or

$$(C) \quad \mathbf{r}(u) = \sqrt{\frac{\boldsymbol{\nu}(u)}{\boldsymbol{\mu}(u)}}} \quad \text{and} \quad \boldsymbol{\sigma}(u) = \boldsymbol{\rho}(u) \mathbf{r}(u)^d \Gamma(\boldsymbol{\nu}(u)) \exp(\boldsymbol{\nu}(u)) \boldsymbol{\nu}(u)^{-\nu(u) - d/2}$$

or

$$(D) \quad \boldsymbol{\nu}(u) \leq \boldsymbol{\nu}, \quad \mathbf{r}(u) = \sqrt{\boldsymbol{\mu}(u)} \quad \text{and} \quad \boldsymbol{\sigma}(u) = \boldsymbol{\rho}(u) \frac{\Gamma(\boldsymbol{\nu}(u))}{\Gamma(\boldsymbol{\nu}(u) + d/2)} \left[ \frac{\Gamma(\nu_{ii}(u) + \nu_{jj}(u) + d/2)}{r_{ij}(u)^{2\nu + \nu_{ii}(u) + \nu_{jj}(u)}} \right]_{i,j=1}^p$$

or

$$(E) \quad \boldsymbol{\nu}(u) = \boldsymbol{\nu} \mathbf{1}, \quad \mathbf{r}(u) = \boldsymbol{\mu}(u) \quad \text{and} \quad \boldsymbol{\sigma}(u) = \boldsymbol{\rho}(u) \mathbf{r}(u)^{-\lfloor \frac{d+1+3\lceil 2\nu \rceil}{2} \rfloor}$$

or

$$(F) \quad \mathbf{r}(u) = r \mathbf{1} \quad \text{and} \quad \boldsymbol{\sigma}(u) = \boldsymbol{\rho}(u) \frac{\Gamma(\boldsymbol{\nu}(u))}{\Gamma(\boldsymbol{\nu}(u) + d/2)}$$

or

$$(G) \quad \nu_{ij}(u) = \frac{\nu_{ii}(u) + \nu_{jj}(u)}{2}, \quad \mathbf{r}(u) = r \mathbf{1} \quad \text{and} \quad \boldsymbol{\sigma}(u) = \frac{\boldsymbol{\rho}(u) \Gamma(\boldsymbol{\nu}(u))}{\Gamma(\boldsymbol{\nu}(u) + d/2)} \left[ \sqrt{\frac{\Gamma(\nu_{ii}(u) + d/2) \Gamma(\nu_{jj}(u) + d/2)}{\Gamma(\nu_{ii}(u)) \Gamma(\nu_{jj}(u))}} \right]_{i,j=1}^p,$$

with  $\mathbf{1}$  standing for the all-ones matrix of size  $p \times p$ , and  $\lfloor \cdot \rfloor$  and  $\lceil \cdot \rceil$  for the floor and ceil functions.

**Theorem 4.** Let  $q \in \mathbb{N}^*$  and  $\mathbf{x}_1, \dots, \mathbf{x}_p \in \mathbb{R}^q$ . Let  $B_1$  and  $B_2$  be Bernstein functions (i.e., positive primitives of completely monotone functions) in  $[0, \infty)$  and  $u \mapsto \boldsymbol{\rho}(u)$  a  $p \times p$  matrix-valued covariance in  $\mathbb{R}$ . Then, the matrix-valued function  $\mathbf{C} : (\mathbf{h}, u) \mapsto \boldsymbol{\sigma}(u) \mathcal{M}(\mathbf{h}; \mathbf{r}(u), \boldsymbol{\nu}(u))$  with

$$\boldsymbol{\nu}(u) = \left[ B_1(\sqrt{|u|^2 + |\mathbf{x}_i - \mathbf{x}_j|^2}) \right]_{i,j=1}^p,$$

$$\mathbf{r}(u)^2 = \left[ \frac{1}{B_2(\sqrt{|u|^2 + |\mathbf{x}_i - \mathbf{x}_j|^2})} \right]_{i,j=1}^p$$

and

$$\boldsymbol{\sigma}(u) = \frac{\boldsymbol{\rho}(u) \mathbf{r}(u)^d \Gamma(\boldsymbol{\nu}(u))}{\Gamma(\boldsymbol{\nu}(u) + d/2)}$$

is positive semidefinite in  $\mathbb{R}^d \times \mathbb{R}$ .

## References

- 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.
- Apanasovich, T. V., Genton, M. G., and Sun, Y. (2012). A valid Matérn class of cross-covariance functions for multivariate random fields with any number of components. *Journal of the American Statistical Association*, 107(497):180–193.
- Arroyo, D. and Emery, X. (2017). Spectral simulation of vector random fields with stationary Gaussian increments in  $d$ -dimensional Euclidean spaces. *Stochastic Environmental Research and Risk Assessment*, 31(7):1583–1592.
- Arroyo, D. and Emery, X. (2020). Iterative algorithms for non-conditional and conditional simulation of Gaussian random vectors. *Stochastic Environmental Research and Risk Assessment*, 34(10):1523–1541.
- Arroyo, D. and Emery, X. (2021). Algorithm 1013: An R implementation of a continuous spectral algorithm for simulating vector Gaussian random fields in Euclidean spaces. *Transactions on Mathematical software*, 47(1):8.
- Berg, C., Christensen, J. P. R., and Ressel, P. (1984). *Harmonic Analysis on Semigroups: Theory of Positive Definite and Related Functions*. Springer-Verlag.
- Bevilacqua, M., Gaetan, C., Mateu, J., and Porcu, E. (2012). Estimating space and space-time covariance functions for large data sets: a weighted composite likelihood approach. *Journal of the American Statistical Association*, 107(497):268–280.
- Bourotte, M., Allard, D., and Porcu, E. (2016). A flexible class of non-separable cross-covariance functions for multivariate space–time data. *Spatial Statistics*, 18:125–146.
- Chen, W. and Genton, M. G. (2019). Parametric variogram matrices incorporating both bounded and unbounded functions. *Stochastic Environmental Research and Risk Assessment*, 33(10):1669–1679.
- Chilès, J.-P. and Delfiner, P. (2012). *Geostatistics: Modeling Spatial Uncertainty, Second Edition*. John Wiley & Sons.
- Dietrich, C. and Newsam, G. (1997). Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix. *SIAM Journal of Scientific Computing*, 18(4):1088–1107.



- Dörr, K. and Schlather, M. (2021). Characterization theorems for pseudo-variograms. *ArXiv*, arXiv:2112.04026.
- Du, J., Leonenko, N., Ma, C., and Shu, H. (2012). Hyperbolic vector random fields with hyperbolic direct and cross covariance functions. *Stochastic Analysis and Applications*, 30(4):662–674.
- 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.
- Emery, X. and Lantuéjoul, C. (2006). TBSIM: A computer program for conditional simulation of three-dimensional Gaussian random fields via the turning bands method. *Computers & Geosciences*, 32(10):1615–1628.
- Emery, X., Porcu, E., and White, P. (2022). Flexible validity conditions for the multivariate Matérn covariance in any spatial dimension and for any number of components. *arXiv:2101.04235v2 [stat.ME]*.
- Finazzi, F., Scott, E. M., and Fassò, A. (2013). A model-based framework for air quality indices and population risk evaluation, with an application to the analysis of scottish air quality data. *Journal of the Royal Statistical Society. Series C, Applied Statistics*, 62(2):287.
- Genton, M. G. and Kleiber, W. (2015). Cross-covariance functions for multivariate geostatistics. *Statistical Science*, pages 147–163.
- Gneiting, T. (2002). Nonseparable, stationary covariance functions for space–time data. *Journal of the American Statistical Association*, 97(458):590–600.
- Gneiting, T., Kleiber, W., and Schlather, M. (2010). Matérn cross-covariance functions for multivariate random fields. *Journal of the American Statistical Association*, 105(491):1167–1177.
- Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American statistical Association*, 102(477):359–378.
- Horn, R. A. and Johnson, C. R. (2013). *Matrix Analysis*. Cambridge University Press.
- Katzfuss, M. and Guinness, J. (2021). A general framework for Vecchia approximations of Gaussian processes. *Statistical Science*, 36(1):124–141.
- Katzfuss, M., Jurek, M., Zilber, D., Gong, W., Guinness, J., Zhang, J., and Schäfer, F. (2020). GPvecchia: Fast Gaussian-process inference using Vecchia approximations. *R package version 0.1, 3*.

- Lantuéjoul, C. (2002). *Geostatistical Simulation: Models and Algorithms*. Springer.
- Lantuéjoul, C. and Desassis, N. (2012). Simulation of a Gaussian random vector: a propagative version of the Gibbs sampler. In *Presented at: Ninth International Geostatistics Congress*, pages 1–8.
- Mitchell, M. W., Genton, M. G., and Gumpertz, M. L. (2006). A likelihood ratio test for separability of covariances. *Journal of Multivariate Analysis*, 97(5):1025–1043.
- Padoan, S. A. and Bevilacqua, M. (2015). Analysis of random fields using CompRandFld. *Journal of Statistical Software*, 63:1–27.
- Papritz, A., Künsch, H., and Webster, R. (1993). On the pseudo cross-variogram. *Mathematical Geology*, 25(8):1015–1026.
- Porcu, E., Emery, X., and Mery, N. (2022). Criteria and characterizations for spatially isotropic and temporally symmetric matrix-valued covariance functions. *HAL*, HAL-0351604.
- Porcu, E., Furrer, R., and Nychka, D. (2021). 30 years of space–time covariance functions. *WIREs Computational Statistics*, 13(2):e1512.
- Stein, M. (2002). Fast and exact simulation of fractional Brownian surfaces. *Journal of Computational and Graphical Statistics*, 11(3):587–599.
- Zastavnyi, V. P. and Porcu, E. (2011). Characterization theorems for the Gneiting class of space–time covariances. *Bernoulli*, 17(1):456–465.
- Zhang, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *Journal of the American Statistical Association*, 99(465):250–261.