



Analysis of Random Fields Using `CompRandFld`

Simone A Padoan
Bocconi University

Moreno Bevilacqua
University of Valparaíso

Abstract

Statistical analysis based on random fields has become a widely used approach in order to better understand real processes in many fields such as engineering, environmental sciences, etc. Data analysis based on random fields can be sometimes problematic to carry out from the inferential prospective. Examples are when dealing with: large dataset, counts or binary responses and extreme values data. This article explains how to perform, with the R package **CompRandFld**, challenging statistical analysis based on Gaussian, binary and max-stable random fields. The software provides tools for performing the statistical inference based on the composite likelihood in complex problems where standard likelihood methods are difficult to apply. The principal features are illustrated by means of simulation examples and an application of Irish daily wind speeds.

Keywords: binary data, composite likelihood, covariance tapering, environmental data analysis, extremal coefficient, Gaussian random field, geostatistics, kriging, large dataset, max-stable random field, wind speed.

1. Introduction

Today, statistical analyses with one or more variables, observed in space and time are useful in different areas, such as environmental sciences and engineering fields (Cressie and Wikle 2011). Random fields (RFs) are the mathematical foundation for statistical spatial analyses; basic introductions to this topic can be found in: Wackernagel (1998), Diggle and Ribeiro (2007), Cressie and Wikle (2011) to name a few. One of the primary aims of spatio-temporal analyses, together with marginal information, is to assess the dependence structure of the underlying random field. With spatial or spatio-temporal applications, the fitting of models with the maximum likelihood (ML) method may be troublesome to apply for several reasons. We discuss two causes: when the analysis concerns large datasets and when the analysis concerns datasets that are not normally distributed.

In the first case, suppose that the dataset of interest, with size m , is coming from a Gaussian

random field. For most of the well-known covariance models, the expression of the ML estimator is unknown and hence the parameter estimation is obtained by numerically maximizing the likelihood. Each likelihood evaluation requires the computation of the determinant and the inverse of the variance-covariance matrix and these tasks are generally performed with algorithms based on the Cholesky factorization, demanding a computational burden of the order $O(m^3)$, see [Bevilacqua, Gaetan, Mateu, and Porcu \(2012\)](#) and the references therein. This can be very time consuming.

In the second case, two interesting types of data are widely discussed in the literature: categorical and extreme values. In order to model spatial binary or categorical data, one way is by dichotomizing or categorizing the outcome of a Gaussian RF into several classes depending on some threshold values, see for instance [Heagerty and Lele \(1998\)](#), [de Leon \(2005\)](#). By doing this, the joint probability model depends on the underlying latent process. For a dataset of size m , the expression of the joint distribution involves a sum of q^m m -dimensional normal integrals, where q is the number of categories. The closed form of the ML estimator is unknown and hence as an alternative the parameter estimation can be done numerically, but with a high computational cost.

In order to model spatial extremes, one way is to work with max-stable RFs (see e.g., [de Haan and Ferreira 2006](#), Chapter 9). These were introduced by [de Haan \(1984\)](#) as an extension of the block maxima approach from finite-dimensional to infinite-dimensional spaces. Also in this case likelihood-based inference is challenging. For most of the known models, (e.g., [Davison, Padoan, and Ribatet 2012](#)) although the closed form of the joint distribution is known, with a dataset of size m the derivation of the likelihood function requires computing a number of derivatives equal to the m th Bell number ([Davison and Gholamrezaee 2012](#)). Thus, for most spatial applications this calculation is not feasible in practice.

Alternative approaches have been proposed in order to perform parametric statistical inference with spatial and spatio-temporal dataset. For instance, in the last decade the strategies proposed are based on simplifications of the covariance structure (e.g., [Cressie and Johannesson 2008](#)) or likelihood approximations (e.g., [Stein, Chi, and Welty 2004](#); [Fuentes 2007](#); [Kaufman, Schervish, and Nychka 2008](#)). Inference based on the composite likelihood function ([Lindsay 1988](#)) has been shown to be a valid competitor. Motivations are the following. Firstly, the composite generalizes the ordinary likelihood and contains many interesting alternatives. Secondly, a number of theoretical results exist about the consistency and the asymptotic distribution of the related estimator and statistical tests. Thirdly, the computational cost is substantially lower than that of the likelihood. Finally, the approach is applicable with different types of data such as Gaussian, binary and extreme values. Examples of statistical analyses based on the composite likelihood are: [Nott and Rydén \(1999\)](#), [Heagerty and Lele \(1998\)](#) for binary data, [Padoan, Ribatet, and Sisson \(2010\)](#), [Davison and Blanchet \(2011\)](#) for extreme values, [Bevilacqua *et al.* \(2012\)](#) for Gaussian responses.

For Gaussian, binary and max-stable RFs, several routines for performing the statistical inference based on the composite likelihood have been put together and implemented in an R ([R Core Team 2014](#)) package named **CompRandFld** ([Padoan and Bevilacqua 2015](#)). This article describes how to use this package to perform statistical analysis of random fields. The article is organized as follows: Section 2 summarizes basic notions on Gaussian, binary and max-stable RFs, Section 3 summarizes the principal concepts of the composite likelihood approach with some simulation examples that explain the main capabilities of **CompRandFld**. Section 4 provides a real data example. Section 5 provides the conclusions.

2. Random fields

2.1. Gaussian random fields

Gaussian RF are useful for describing the random behavior of the mean levels of a real process. Here, a few basic notions are recalled.

Let $\{Y(\mathbf{s}, t), (\mathbf{s}, t) \in \mathcal{I}\}$ be a real-valued random process defined on \mathcal{I} , where $\mathcal{I} = \mathcal{S} \times \mathcal{T}$, with $\mathcal{S} \subseteq \mathbb{R}^d$ and $\mathcal{T} \subseteq \mathbb{R}$. When $\mathcal{T} = \{t_0\}$ then $\mathcal{I} \equiv \mathcal{S}$ and $Y(\mathbf{s}) \equiv Y(\mathbf{s}, t_0)$ is a purely spatial random field. When $\mathcal{S} = \{\mathbf{s}_0\}$ then $\mathcal{I} \equiv \mathcal{T}$ and $Y(t) \equiv Y(\mathbf{s}_0, t)$ is a purely temporal random field. Let $I := \{1, \dots, l\}$ and $J := \{1, \dots, k\}$ be sets of indices and $I^* = I \times J$ be the Cartesian product with cardinality $m = k \cdot l$. It is assumed for simplicity that $Y(\mathbf{s}, t)$ is a second-order stationary spatio-temporal Gaussian RF, that means $\mathbb{E}\{Y(\mathbf{s}, t)\} = \mu$ and $\text{var}\{Y(\mathbf{s}, t)\} = \omega^2$ are finite constants for all $(\mathbf{s}, t) \in \mathcal{I}$ and the covariance $\text{cov}\{Y(\mathbf{s}, t), Y(\mathbf{s}', t')\} = \mathbb{E}\{Y(\mathbf{s}, t), Y(\mathbf{s}', t')\} - \mu^2$, $(\mathbf{s}, t), (\mathbf{s}', t') \in \mathcal{I}$ can be expressed by $c(\mathbf{h}, u)$, namely $c(\cdot, \cdot)$ is a positive definite function that depends only on the separations between the points $\mathbf{h} = \mathbf{s}' - \mathbf{s}$ and $u = t' - t$ (Cressie and Wikle 2011, Chapter 6). The covariance is said to be translation-invariant. Since $c(\mathbf{0}, 0) = \text{var}\{Y(\mathbf{s}, t)\}$, then $\rho(\mathbf{h}, u) = c(\mathbf{h}, u)/c(\mathbf{0}, 0)$ is the correlation function of Y . In order to take into account the local variation of real phenomena the overall process is defined by $Y(\mathbf{s}, t) = Y_0(\mathbf{s}, t) + \varepsilon(\mathbf{s}, t)$, where $\varepsilon(\mathbf{s}, t) \sim N(0, \tau^2)$ is a white noise independent from Y_0 . The covariance takes the following functional form

$$c(\mathbf{h}, u) = \tau^2 \mathbb{I}_{\{(\mathbf{0}, 0)\}}(\mathbf{h}, u) + \sigma^2 \rho_0(\mathbf{h}, u), \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}, \quad (1)$$

where $\mathbb{I}(\cdot)$ is the indicator function, $\sigma^2 > 0$ and ρ_0 are the common variance and correlation of the RF, whereas $\tau^2 > 0$ is the local variance that describes the so-called nugget effect (e.g., Diggle and Ribeiro 2007; Cressie and Wikle 2011, Chapter 4). The variance of $Y(\mathbf{s}, t)$ is equal to $\omega^2 = \tau^2 + \sigma^2$ and the covariance function $\sigma^2 \rho_0(\mathbf{h}, u)$. Since $Y(\mathbf{s}, t)$ is also intrinsically stationary (Cressie and Wikle 2011, Chapter 6), then its variability can equivalently be represented as $2\gamma(\mathbf{h}, u) = \text{var}\{Y(\mathbf{s} + \mathbf{h}, t + u) - Y(\mathbf{s}, t)\}$. The function $\gamma(\cdot, \cdot)$ is called the semivariogram while the quantity $2\gamma(\cdot, \cdot)$ is known as the variogram and under second order stationary $\gamma(\mathbf{h}, u) = c(\mathbf{0}, 0) - c(\mathbf{h}, u)$. A stationary space time covariance is called spatially isotropic if $c(\mathbf{h}, u_0)$ depends only on $(\|\mathbf{h}\|, u_0)$ for any fixed temporal lag u_0 , see Porcu, Montero, and Schlather (2012).

There is extensive literature on parametric spatial covariance models, see Cressie and Wikle (2011, Chapter 4) and the reference therein, and in the past years many parametric models for space-time covariances have been proposed. A possible covariance model is obtained as the product of any valid spatial and temporal covariance. This kind of covariance model, called separable model, has been criticized for its lack of flexibility. For this reason, different classes of non separable covariance models have been proposed recently in order to capture possible space-time interactions (e.g., Porcu, Mateu, and Christakos 2009; Cressie and Wikle 2011, Chapter 6). Some examples of covariance models implemented in **CompRandFld** and used later in Section 3 and 4 are found in Table 1. Observe that the first model is a separable space-time stable or power-exponential covariance, the second is a special case of the class proposed in Gneiting (2002), the third and fourth are special cases of the classes proposed in De Iaco, Myers, and Posa (2002) and Porcu *et al.* (2009). Finally, the fifth and sixth models are spatial correlation models (e.g., Gneiting and Schlather 2004). A list of all the correlation

Model	Functional form	Parameters' range
Double stable	$\rho_{st}(\mathbf{h}, u) = \exp\{-d_s(\mathbf{h}) - d_t(u)\}$	
Gneiting	$\rho_{st}(\mathbf{h}, u) = g_t(u)^{-1} \exp\left\{-\frac{d_s(\mathbf{h})}{g_t(u)^{-\eta\kappa_s/2}}\right\}$	$0 \leq \eta \leq 1$
Iacocesare	$\rho_{st}(\mathbf{h}, u) = \{g_t(u) + d_s(\mathbf{h})\}^{-\eta}$	$\eta > 0$
Porcu	$\rho_{st}(\mathbf{h}, u) = \{0.5g_s(\mathbf{h})^\eta + 0.5g_t(u)^\eta\}^{-1/\eta}$	$0 \leq \eta \leq 1$
Matern	$\rho_s(\mathbf{h}) = 2^{1-\alpha} d_s(\mathbf{h})^{\alpha/\kappa_s} K_\alpha\left(\frac{\ \mathbf{h}\ }{\psi_s}\right) \Gamma(\alpha)^{-1}$	$\alpha > 0$
Gen. Cauchy	$\rho_s(\mathbf{h}) = g_s(\mathbf{h})^{-\alpha/\kappa_s}$	$\alpha > 0$
	$d_s(\mathbf{h}) = \left(\frac{\ \mathbf{h}\ }{\psi_s}\right)^{\kappa_s}, g_s(\mathbf{h}) = 1 + d_s(\mathbf{h}),$	$\psi_s, \psi_t > 0$
	$d_t(u) = \left(\frac{ u }{\psi_t}\right)^{\kappa_t}, g_t(u) = 1 + d_t(u)$	$0 < \kappa_s, \kappa_t \leq 2$

Table 1: Parametric families of isotropic correlation models. The first four models are spatio-temporal and the last two are spatial and K_α is the modified Bessel function of order α , Γ is the gamma function, η is a separability parameter, α , κ_s , κ_t are smoothing parameters and ψ_s , ψ_t are scale parameters. The subscripts s and t stand for space and time.

functions implemented in the **CompRandFld** is available with the online help of the routine **CovMatrix**.

2.2. Binary random fields

In some applications the focus is on modeling the presence or absence of a certain phenomenon, observable across space and time. To describe the stochastic behavior of such a process a binary RF is needed.

An approach for defining a spatio-temporal binary random field is the following (e.g., [Heagerty and Lele 1998](#)). Let $\{Y(\mathbf{s}, t), (\mathbf{s}, t) \in \mathcal{I}\}$ be a second-order stationary spatio-temporal Gaussian RF with correlation function $\rho(\mathbf{h}, u)$ that we assume known up to a vector of parameters $\boldsymbol{\vartheta}$. Then, a binary spatio-temporal RF is defined by

$$Z(\mathbf{s}, t) = \mathbb{I}_{(b, +\infty)}(Y(\mathbf{s}, t) + \varepsilon(\mathbf{s}, t)), \quad (2)$$

where for all $(\mathbf{s}, t) \in \mathcal{I}$, $\varepsilon(\mathbf{s}, t) \sim N(0, \tau^2)$ is a white noise and $b \in \mathbb{R}$ is a threshold. We assume that ε is independent from Y . With this definition, for all $(\mathbf{s}, t) \in \mathcal{I}$, the univariate marginal probability of Z is

$$p_1 \equiv \text{pr}(Z(\mathbf{s}, t) = 1) = \Phi\left(\frac{\mu - b}{\omega}\right),$$

where μ is the mean of Z and Φ is the univariate standard Gaussian distribution. In addition, for any $(\mathbf{s}, t), (\mathbf{s}', t') \in \mathcal{I}$ with $(\mathbf{s}, t) \neq (\mathbf{s}', t')$, the bivariate joint probabilities are equal to

$$\begin{aligned} p_{11}(\mathbf{h}, u) &\equiv \text{pr}\{Z(\mathbf{s}, t) = 1, Z(\mathbf{s}', t') = 1\} = \Phi_2\left(\frac{\mu - b}{\omega}, \frac{\mu - b}{\omega}; \frac{\sigma^2 \rho(\mathbf{h}, u; \boldsymbol{\vartheta})}{\omega^2}\right), \\ p_{00}(\mathbf{h}, u) &= 1 - 2p_1 + p_{11}(\mathbf{h}, u), \\ p_{10}(\mathbf{h}, u) &\equiv p_{01}(\mathbf{h}, u) = p_1 - p_{11}(\mathbf{h}, u), \end{aligned} \quad (3)$$

where Φ_2 is the bivariate standard Gaussian distribution. In our package, Φ_2 is computed using a Fortran routine written by Alan Genz, see [Genz \(1992\)](#). Note that the marginal

and the joint probability of success p_1 and $p_{11}(\mathbf{h}, u)$ completely characterize the bivariate distribution. Higher order joint probabilities are similarly derived, by using the inclusion-exclusion principle. Thus given m spatio-temporal points, the computation of the probabilities for all the related success and failure events requires the evaluation of high-dimensional normal integrals, up to the order m .

2.3. Max-stable random fields

In some applications the focus is on modeling exceptionally high (low) levels of a real process. One way for describing the random behavior of spatial or spatio-temporal extremes is provided by max-stable RFs. For simplicity let us focus on the spatial case. Let $\{Y(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$ be a stationary RF with continuous sample path and Y_1, \dots, Y_n be n independent copies of it. Consider the pointwise maximum process $M_n(\mathbf{s}) := \max_{i=1, \dots, n} Y_i(\mathbf{s})$, $\mathbf{s} \in \mathcal{S}$. If there exist continuous functions $a_n(\mathbf{s}) > 0$ and $b_n(\mathbf{s}) \in \mathbb{R}$, for all $\mathbf{s} \in \mathcal{S}$, such that $a_n^{-1}(\mathbf{s})\{M_n(\mathbf{s}) - b_n(\mathbf{s})\}$ converge weakly to a process $Z(\mathbf{s})$ with non-degenerate marginal distributions for all $\mathbf{s} \in \mathcal{S}$, then Z is a max-stable RF (e.g., [de Haan and Ferreira 2006](#), Chapter 9). For any $\mathbf{s} \in \mathcal{S}$, max-stable random fields have univariate marginal generalized extreme value distributions and for a finite set of locations the joint distribution

$$\text{pr}\{Z(\mathbf{s}_j) \leq z_j, j \in J\} = \exp\{-V(\mathbf{z})\}, \quad V(\mathbf{z}) = k \int_{\Delta_{k-1}} \max_{j \in J} \left(\frac{w_j}{y_j} \right) v(d\mathbf{w}), \quad \mathbf{z} \in \mathbb{R}_+^k \quad (4)$$

with unit Fréchet marginal distributions, where v is a probability distribution defined on the $(k-1)$ -dimensional unit simplex Δ_{k-1} and V is the so-called exponent dependence function (e.g., [de Haan and Ferreira 2006](#), Chapter 1.6). Examples of max-stable RF are the Brown-Resnick ([Kabluchko, Schlather, and de Haan 2009](#); [Kabluchko 2010](#)), the extremal-Gaussian ([Schlather 2002](#)) and the extremal-t ([Davison *et al.* 2012](#)). In the first case the exponent dependence function of the bivariate distribution is

$$V_1(z_1, z_2; \mathbf{h}) = \frac{1}{z_1} \Phi \left(\frac{\lambda(\mathbf{h})}{2} + \frac{\log z_2/z_1}{\lambda(\mathbf{h})} \right) + \frac{1}{z_2} \Phi \left(\frac{\lambda(\mathbf{h})}{2} + \frac{\log z_1/z_2}{\lambda(\mathbf{h})} \right), \quad (5)$$

where $\lambda(\mathbf{h}) = \sqrt{2\gamma(\mathbf{h})}$ with γ denoting the semivariogram of the underlying Gaussian RF. (5) was initially derived by [Hüsler and Reiss \(1989\)](#) and is also the exponent function of the max-stable RF described by [Smith \(1990\)](#) but with a different form for $\lambda(\mathbf{h})$. For these processes, extensions to the spatio-temporal case with dependence function $\lambda(\mathbf{h}, u)$ exist, see [Kabluchko \(2009\)](#), [Davis, Klüppelberg, and Steinkohl \(2013\)](#). In the second case the bivariate distribution depends on the exponent function

$$V_2(z_1, z_2; \mathbf{h}) = \frac{1}{2} \left(\frac{1}{z_1} + \frac{1}{z_2} \right) \left(1 + \left[1 - \frac{2z_1z_2\{\rho(\mathbf{h}) + 1\}}{(z_1 + z_2)^2} \right]^{1/2} \right), \quad (6)$$

where $\rho(\mathbf{h})$ is the correlation function $\rho(\mathbf{h})$ of the underlying Gaussian RF. Finally, in the third case, the bivariate distribution depends on the exponent function

$$V_3(z_1, z_2; \mathbf{h}) = \frac{1}{z_1} T_{\nu+1} \left\{ \frac{(z_2/z_1)^{1/\nu} - \rho(\mathbf{h})}{\sqrt{\{1 - \rho(\mathbf{h})^2\}/(\nu + 1)}} \right\} + \frac{1}{z_2} T_{\nu+1} \left\{ \frac{(z_1/z_2)^{1/\nu} - \rho(\mathbf{h})}{\sqrt{\{1 - \rho(\mathbf{h})^2\}/(\nu + 1)}} \right\}, \quad (7)$$

where $T_{\nu+1}$ is a Student- t distribution with $\nu + 1$ degrees of freedom, $\rho(\mathbf{h})$ is the scaling function and ν the degrees of freedom of the underlying Student- t random field.

A summary for describing the extremal dependence is $\xi(\mathbf{h}) = V(1, 1; \mathbf{h})$, named the pairwise extremal coefficient function (Smith 1990). It holds that $1 \leq \xi(\mathbf{h}) \leq 2$, where the lower bound represents the complete dependence case and the upper bound the independence case. For the above mentioned processes the extremal coefficients are equal to

$$\xi_1(\mathbf{h}) = 2\Phi\left(\sqrt{\frac{\gamma(\mathbf{h})}{2}}\right), \quad \xi_2(\mathbf{h}) = 1 + \sqrt{\frac{1 - \rho(\mathbf{h})}{2}}, \quad \xi_3(\mathbf{h}) = 2T_{\nu+1} \left\{ \sqrt{\frac{(\nu+1)(1 - \rho(\mathbf{h}))}{1 + \rho(\mathbf{h})}} \right\}, \quad (8)$$

respectively.

3. Composite likelihood approach

Composite likelihood (CL) is a general class of estimating functions based on the likelihood of marginal or conditional events (e.g., Varin, Reid, and Firth 2011). Let $\mathbf{Y} = (Y_1, \dots, Y_m)^\top$ be an m -dimensional random vector with joint distribution $f(\mathbf{Y}; \boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^d$. Let $\mathcal{A}_1, \dots, \mathcal{A}_R$ be marginal or conditional sets of \mathbf{Y} , then CL is defined as

$$\ell_c(\boldsymbol{\theta}) = \sum_{r=1}^R \ell(\boldsymbol{\theta}; \mathcal{A}_r) w_r, \quad (9)$$

where $\ell(\boldsymbol{\theta}; \mathcal{A}_r)$ is the log-likelihood related to the density $f(\mathbf{Y} \in \mathcal{A}_r; \boldsymbol{\theta})$ obtained by considering only the random variables in \mathcal{A}_r and w_r are suitable non negative weights that do not depend on $\boldsymbol{\theta}$. The maximum composite likelihood (MCL) estimator is given by $\hat{\boldsymbol{\theta}} = \operatorname{argmax}_{\boldsymbol{\theta}} \ell_c(\boldsymbol{\theta})$. The CL estimation method has nice properties (e.g., Varin *et al.* 2011) among which, if the sample size n is such that $n \gg m$ and n goes to infinity, then $\hat{\boldsymbol{\theta}}$ is weakly consistent and asymptotically normally distributed with variance-covariance matrix $G(\boldsymbol{\theta})^{-1}$, where

$$G(\boldsymbol{\theta}) = H(\boldsymbol{\theta})J(\boldsymbol{\theta})^{-1}H(\boldsymbol{\theta})^\top \quad (10)$$

is the Godambe information matrix and

$$H(\boldsymbol{\theta}) = -\mathbb{E}[\nabla^2 \ell_c(\boldsymbol{\theta})], \quad J(\boldsymbol{\theta}) = \mathbb{E}[\nabla \ell_c(\boldsymbol{\theta}) \nabla \ell_c(\boldsymbol{\theta})^\top],$$

where ∇ denotes the vector differential operator. On the other hand, if $n \ll m$ or dependent data are available such as a single spatial sequence or a univariate time series, then the above mentioned asymptotic results do not necessary hold, see Cox and Reid (2004). Then, consistency and asymptotic normality of the estimator, as m goes to infinity, must be shown case by case. Likelihood-based statistics such as Wald statistic, score statistic and likelihood ratio statistic and their asymptotic distributions are also available in the CL setting (Chandler and Bate 2007; Varin *et al.* 2011; Pace, Salvan, and Sartori 2011). In addition, the Akaike information criterion used for model selection in the CL context assumes the following form:

$$CLIC = -2\ell_c(\boldsymbol{\theta}) + 2\operatorname{tr}\{H(\boldsymbol{\theta})G^{-1}(\boldsymbol{\theta})\}, \quad (11)$$

where $\operatorname{tr}(A)$ is the trace of the matrix A , see Varin and Vidoni (2005). The CL is a wide class of functions that also includes the standard likelihood and other alternatives as special cases. For example, the former is obtained letting $\mathcal{A}_r = \mathbf{Y}$, $R = 1$ and $w_r = 1$. Depending on the type of likelihood that is specified on the right hand side of (9), different types of

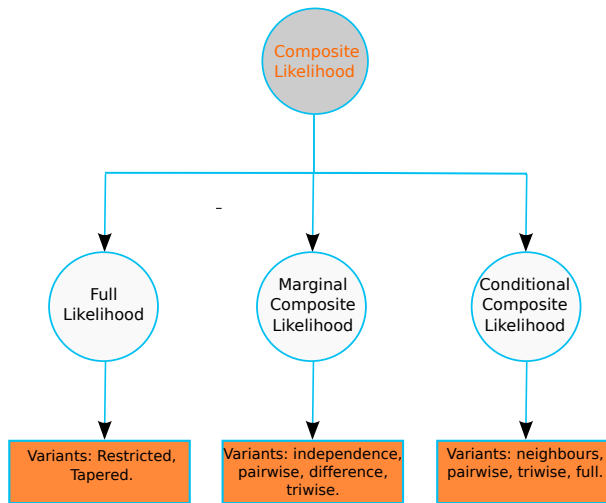


Figure 1: Hierarchical structure involved with the composite likelihood approach.

CL can be obtained. One of the most explored cases is the pairwise likelihood. Considering the marginal pairs $\mathcal{A}_r = (Y_i, Y_j)$, then we obtain the pairwise marginal log-likelihood. If we let $\mathcal{A}_r = (Y_i|Y_j)$, then we obtain the pairwise conditional log-likelihood and finally setting $\mathcal{A}_r = (Y_i - Y_j)$, then we obtain the pairwise difference log-likelihood. In these special cases, log-likelihoods, for a single data contribution are respectively:

$$\ell_{pm}(\boldsymbol{\theta}) = \sum_{i=1}^m \sum_{j>i}^m \log f(y_i, y_j; \boldsymbol{\theta}) w_{ij}, \quad (12)$$

$$\ell_{pc}(\boldsymbol{\theta}) = \sum_{i=1}^m \sum_{j \neq i}^m \log f(y_i | y_j; \boldsymbol{\theta}) w_{ij}, \quad (13)$$

$$\ell_{pd}(\boldsymbol{\theta}) = \sum_{i=1}^m \sum_{j>i}^m \log f(y_i - y_j; \boldsymbol{\theta}) w_{ij}. \quad (14)$$

For these likelihoods the computational burden is $O(m^2)$. In principle, the role of the weights is to improve the statistical efficiency and save computational time, see for example Padoan *et al.* (2010), Bevilacqua *et al.* (2012), Bevilacqua and Gaetan (2014) and Sang and Genton (2014). The diagram in Figure 1 summarizes the types of likelihoods belonging to the CL class that have been implemented in the package. For comparison purposes we have also implemented in the package other variants of the ordinary likelihood, such as the restricted likelihood of Harville (1977) and the tapered likelihood of Kaufman *et al.* (2008).

3.1. The Gaussian case

Assume that $\mathbf{y} = \{y(\mathbf{s}_1, t_1), \dots, y(\mathbf{s}_k, t_l)\}^\top$ is a realisation from a spatio-temporal Gaussian RF (see Section 2.1). \mathbf{y} has an m -dimensional normal density with the parameter vector $\boldsymbol{\theta} = (\mu, \tau^2, \sigma^2, \boldsymbol{\vartheta})^\top$. $\boldsymbol{\vartheta}$ is a vector of correlation parameters (see for example the models in Table 1). In the Gaussian case, there are three estimation methods available. Standard ML fitting of the Gaussian model can be performed setting the parameters `model = "Gaussian"`,

Model	Functional form	Parameters' range
Wendland1	$\rho_s(\mathbf{h}; c_s) = \{1 - d_s(\mathbf{h})\}_+^2 \{1 + d_s(\mathbf{h})/2\}$	
Wendland2	$\rho_s(\mathbf{h}; c_s) = \{1 - d_s(\mathbf{h})\}_+^4 \{1 + 4d_s(\mathbf{h})\}$	
Double Wendland2	$\rho_{st}(\mathbf{h}, u; c_s, c_t) = \{1 - d_s(\mathbf{h})\}_+^4 \{1 + 4d_s(\mathbf{h})\} \times \{1 - d_t(u)\}_+^4 \{1 + 4d_t(u)\}$	
$d_s(\mathbf{h}) = \frac{\ \mathbf{h}\ }{c_s}, d_t(u) = \frac{ u }{c_t}$		$c_s, c_t > 0$

Table 2: Parametric families of isotropic taper functions.

`likelihood = "Full"` and `type = "Standard"` of the `FitComposite` routine. This method is recommended for small dataset. $\boldsymbol{\theta}$ is estimated by maximization of the full likelihood

$$\ell(\boldsymbol{\theta}) = -\frac{1}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} (\mathbf{y} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\mu}), \quad (15)$$

where $\boldsymbol{\mu} = \mu \mathbf{1}_m$ and $\boldsymbol{\Sigma} = \{\tau^2 \mathbb{I}_{\{(r,q)\}}((j,i)) + \sigma^2 \rho_{st}(\|\mathbf{s}_j - \mathbf{s}_r\|, |t_i - t_q|); (i,j), (q,r) \in I^*\}$ is the $m \times m$ covariance matrix defined by (1). The restricted version is also available and it can be used setting `type = "Restricted"`. Under increasing domain the asymptotic variance-covariance matrix of the ML estimator is given by the inverse of the Fisher information matrix $I(\boldsymbol{\theta})$, see [Mardia and Marshall \(1984\)](#).

The second estimation method is based on the tapering likelihood ([Kaufman et al. 2008](#)), recommended for large datasets. This method is performed setting the parameter `type = "Tapering"`. The tapering likelihood is given by

$$\ell_{tap}(\boldsymbol{\theta}, \mathbf{c}) = -\frac{1}{2} \log |\tilde{\boldsymbol{\Sigma}}| - \frac{1}{2} (\mathbf{y} - \boldsymbol{\mu})^\top \{\tilde{\boldsymbol{\Sigma}}^{-1} \circ \boldsymbol{\Omega}_{\mathbf{c}}\} (\mathbf{y} - \boldsymbol{\mu}), \quad (16)$$

where $\boldsymbol{\Omega}_{\mathbf{c}} = \{\rho_{st}(\|\mathbf{s}_j - \mathbf{s}_r\|, |t_i - t_q|; \mathbf{c}); (i,j), (q,r) \in I^*\}$ is the $m \times m$ taper matrix, $\tilde{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma} \circ \boldsymbol{\Omega}_{\mathbf{c}}$ is an $m \times m$ tapered covariance matrix and \circ is the Schur product. Specifically, $\rho(\cdot, \cdot; \mathbf{c})$ is a correlation function (taper function) depending on a parameter's vector \mathbf{c} , which defines the compact support in space, or in time, or both ([Wendland 1995](#)). Hence, tapering in likelihood (15) both the covariance matrix $\boldsymbol{\Sigma}$ and the matrix $\tilde{\boldsymbol{\Sigma}}^{-1}$, the tapering likelihood (16) is obtained. The parameter `taper` of `FitComposite` allows us to set different taper functions. A list of models implemented in the package is reported in Table 2. The first two are spatial tapers, while the last is a spatio-temporal taper. The compact support parameters for space and time can be set providing two positive values c_s and c_t at `maxdist` and `maxtime` (see Table 2). Likelihood (16) is implemented in **CompRandFld** and in the optimization procedure it calls for specific routines of the package **spam** ([Furrer and Sain 2010](#); [Furrer 2014](#)) for handling the sparse matrix $\tilde{\boldsymbol{\Sigma}}$. Under increasing domain, the maximizer of (16) is a consistent estimator and has asymptotic Gaussian distribution, with variance given by the inverse of the Godambe matrix which is given as in (10) with

$$H(\boldsymbol{\theta}) = -\mathbb{E}[\nabla^2 \ell_{tap}(\boldsymbol{\theta}, \mathbf{c})], \quad J(\boldsymbol{\theta}) = \mathbb{E}[\nabla \ell_{tap}(\boldsymbol{\theta}, \mathbf{c}) \nabla \ell_{tap}(\boldsymbol{\theta}, \mathbf{c})^\top],$$

see [Shaby and Ruppert \(2012\)](#). The tapering can have an effect, depending on the choice of the compact support, when assessing a correlation structure. For example, the top-left panel of Figure 2 reports a simulation from a zero-mean, unit-variance, spatial Gaussian random field

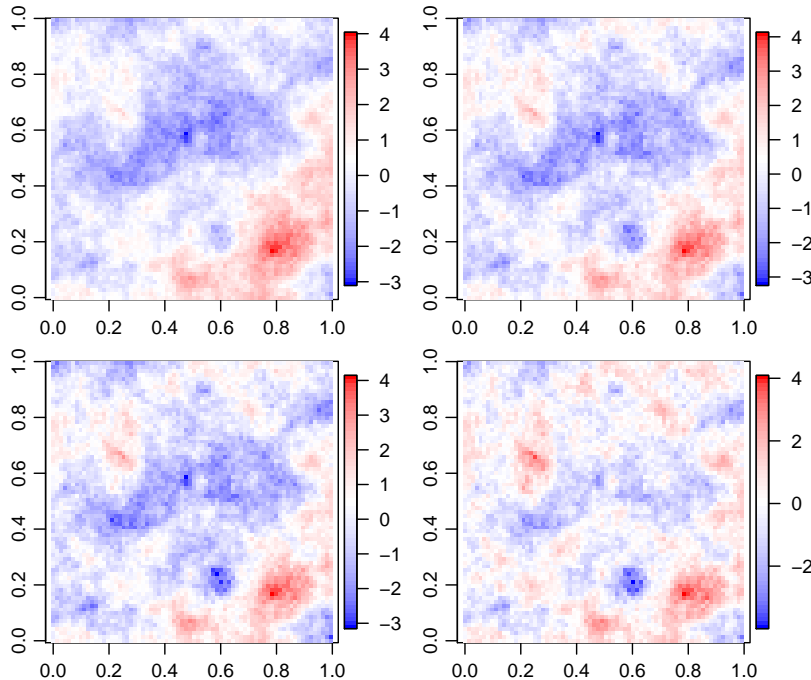


Figure 2: A simulation from a zero-mean, unit-variance spatial Gaussian random field with exponential correlation and scale parameter $c_s = 0.2$ is reported in the top-left panel. The same simulation is shown from the top-right to the bottom-right panels, tapering the correlation using a Wendland function with compact supports 0.6, 0.4 and 0.2.

with exponential correlation and scale parameter $c_s = 0.2$ on the unit square. Simulations are repeated by tapering the covariance matrix with the first Wendland function of Table 2 using the compact supports: 0.6, 0.4 and 0.2. When the compact support is greater or equal to the practical range (0.6), the simulation in the top-right panel shows that there is no difference with the original simulation. While for decreasing values of compact support, the simulations on the two bottom panels show a decreasing range of the correlation structure.

The third estimation method is based on the pairwise likelihoods (12)–(14), also recommended for large dataset. Bevilacqua and Gaetan (2014) have shown by means of a simulation study, that under an increasing-domain asymptotic and for a fixed given compact support, the pairwise likelihood, compared to the other estimation methods, is the fastest. For instance, Figure 3 shows that the ratio between the computational time required for a single numerical evaluation of the full and the pairwise, and the tapering and the pairwise likelihoods is greater than one, stressing that the pairwise likelihood is the fastest estimation method. The ratios have been computed for different sample sizes. We can see that for increasing sample size the pairwise likelihood provide a computational convenience that is increasingly prominent. CL estimation is performed setting `likelihood = "Marginal"` and `type = "Difference"` in `FitComposite`. This is based on the difference pairwise likelihood

$$\ell_{pd}(\boldsymbol{\theta}) = \sum_{(i,j,q,r) \in D} \left\{ \frac{\log \gamma_{ijqr}(\sigma^2, \tau^2, \boldsymbol{\vartheta})}{2} + \frac{\{y(\mathbf{s}_j, t_i) - y(\mathbf{s}_r, t_q)\}^2}{4\gamma_{ijqr}(\sigma^2, \tau^2, \boldsymbol{\vartheta})} \right\} w_{ijqr},$$

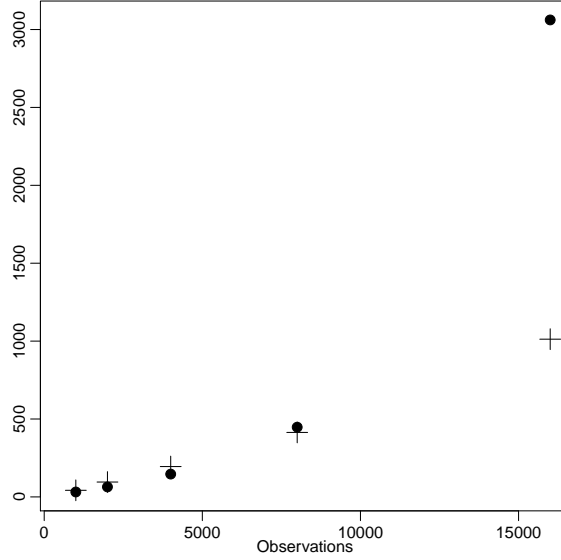


Figure 3: Points depict the ratio between the times required to numerically evaluate the full and the pairwise likelihood. The crosses report the comparison between the tapering and the pairwise likelihood.

where

$$D = \begin{cases} (i, j, q, r) : & q > i, i \in I, \quad \text{if } r = j, j \in J \\ & i, q \in I, \quad \quad \quad \text{if } r > j, j \in J \end{cases} \quad (17)$$

and $\gamma_{ijqr}(\sigma^2, \tau^2, \boldsymbol{\vartheta}) = \tau^2 \mathbb{I}_{\{(r,q)\}}((j, i)) + \sigma^2 \{1 - \rho_{st}(\|\mathbf{s}_j - \mathbf{s}_r\|, |t_i - t_q|; \boldsymbol{\vartheta})\}$ is the variogram. Pairwise or conditional likelihoods (12) and (13) can be used setting `likelihood = "Marginal"` and `type = "Pairwise"` or `likelihood = "Conditional"`. Binary weights

$$w_{ijqr} = \begin{cases} 1 & \|\mathbf{s}_j - \mathbf{s}_r\| \leq c_s, |t_i - t_q| \leq c_t \\ 0 & \text{otherwise} \end{cases}$$

can be set by providing two positive values c_s and c_t at `maxdist` and `maxtime`. Under increasing domain the maximizer of the CL is consistent and asymptotically normal with variance-covariance matrix as in (10) (Bevilacqua *et al.* 2012). Standard errors can be obtained setting `varest=TRUE`. These are obtained, for the three likelihood methods, as the square roots of the diagonal elements of the inverse of the Fisher and Godambe information. In the Gaussian case the analytical expressions of such matrices are known.

In the composite likelihood case however, since the expression of J is unwieldy, its numerical evaluation can be computationally demanding. So to quickly obtain standard errors, two options are implemented. If a single data replication is available, then the matrix is computed with a sub-sampling technique and this is obtained specifying `vartype = "subsamp"`. If data replications are available, then the Godambe matrix is computed using its sampling version (Varin *et al.* 2011) and this is obtained specifying `vartype = "sampling"`. Here is a practical example.

```
R> library("CompRandFld")
R> set.seed(13)
```

```
R> coords <- cbind(runif(25, 0, 1), runif(25, 0, 1))
R> times <- 1:20
R> param <- list(nugget = 0, mean = 1, scale_s = 2, scale_t = 3, sill = 1)
R> data <- RFsim(coordx = coords, coordt = times, corrmmodel = "exp_exp",
+   param = param)$data
```

This code shows how to use the routine `RFsim` to simulate a spatio-temporal random field in 25 locations and for 20 times. The correlation is specified with `corrmmodel` and the parameters with `param`. The model used is the double exponential, the first of Table 1 with $\kappa_s = \kappa_t = 1$. Note that simulations from Gaussian random fields are performed using the Cholesky decomposition method. The next code shows the use of the `FitComposite` routine to perform MCL estimations of the parameters.

```
R> pml <- FitComposite(data, coordx = coords, coordt = times,
+   corrmmodel = "exp_exp", likelihood = "Marginal", type = "Pairwise",
+   start = list(scale_s = 2, scale_t = 3, sill = 1),
+   fixed = list(mean = 1, nugget = 0), varest = TRUE, maxdist = 0.5,
+   maxtime = 3)
R> pml$param
```

```
scale_s  scale_t    sill
2.244840 5.739743 1.392082
```

With `start`, users can specify a starting value for any parameter that needs to be estimated. While with `fixed`, users can specify a value for a parameter that does not need to be estimated. Starting values are admitted only for parameters that are not fixed. Note that the routine `CorrelationParam` returns the list of parameters for any correlation model. Estimates are obtained by fitting the marginal pairwise likelihood. Binary weights are used so that only pairs of observations with spatial distances and temporal intervals smaller or equal to 0.5 and 3 are retained. `print(pml)` provides the summary of the fitting, omitted here for brevity. The collected results are: the parameters' estimate with related standard errors and variance-covariance matrix, the value that maximizes the log pairwise likelihood and the value of the CLIC (11) and other information. The fitting results can be used for estimating the correlation or variogram function, computing the practical range (Diggle and Ribeiro 2007) with the `Covariogram` routine, or as the following code shows, for predicting the process on ungauged locations for future times.

```
R> times_to_pred <- 21:23
R> x <- seq(0,1, 0.02)
R> loc_to_pred <- as.matrix(expand.grid(x, x))
R> param <- as.list(c(pml$param, pml$fixed))
R> pr <- Kri(data, coords, coordt = times, loc = loc_to_pred,
+   time = times_to_pred, corrmmodel = "exp_exp", param = param)
```

The `Kri` routine performs simple kriging (or ordinary kriging, Diggle and Ribeiro 2007) in the new set of locations `loc_to_pred` and times `times_to_pred`. A dataset, spatio-temporal coordinates and a correlation model with the estimated parameters need to be specified. The

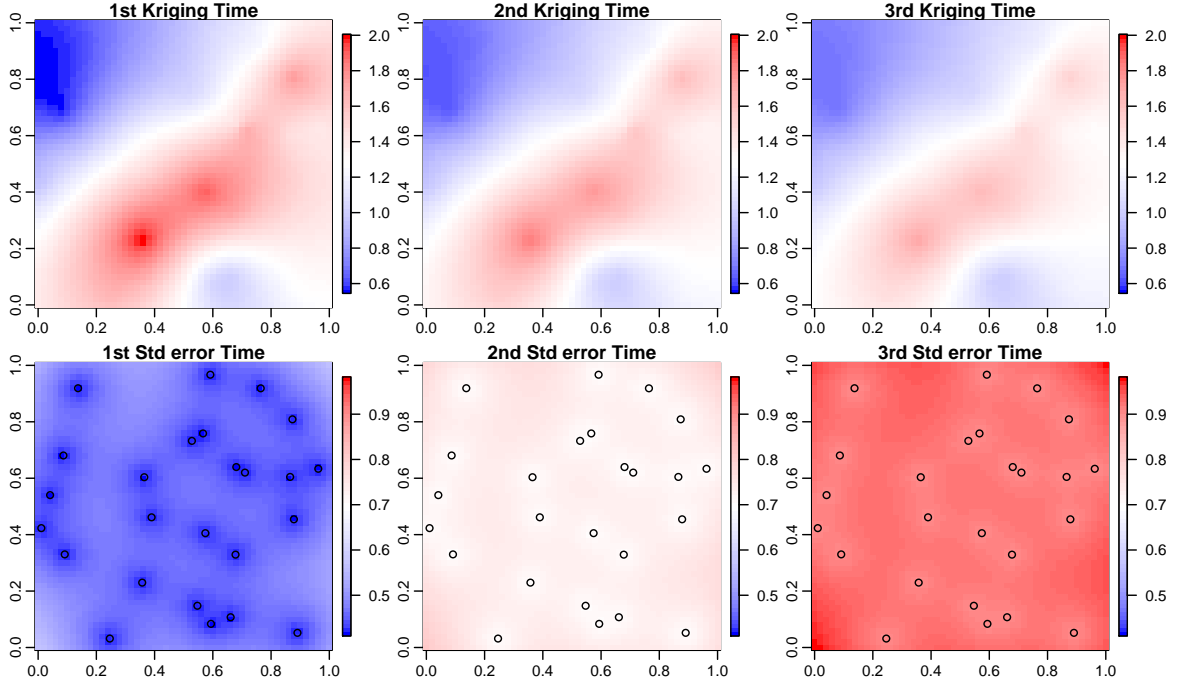


Figure 4: Simple kriging on a 51×51 grid of points on $[0,1]$ and for 3 times (from left to right). The top panels report the prediction and the bottom the standard error of the prediction. The points are the locations where the data has been generated.

routine returns two matrices (together with other information), one with the predicted values and one with their associated standard errors. With standard R code, the results can be easily visualized. Figure 4 shows, with the top and bottom panels, the predicted values and their standard errors, on a 51×51 grid of points in $[0, 1]$ and three consecutive times (left-right). Due to the strong dependence structure, we see that a noticeable dependence remains on time (although decreasing) and that the lower variabilities are recorded on the points where the data has been simulated. In **CompRandFld** simple and tapered kriging are implemented. The default option is simple kriging. This is performed working with the full covariance matrix, which is computationally intensive for large spatial or spatio-temporal data. A remedy is to use the tapered kriging proposed by Furrer, Genton, and Nychka (2006). Tapered kriging is an approximation of the classical kriging, faster to compute and hence useful when dealing with large datasets. With tapered kriging, the tapered covariance matrix is used in the simple kriging systems instead of the full covariance matrix, so that the solution is obtained using fast algorithms for sparse matrices. The routine `Kri` calls specific routines of the package `spam` (Furrer and Sain 2010) in order to handle sparse matrices. Tapered kriging can be done setting `type = "Tapering"` and specifying values for the parameters `taper`, `maxdist` and `maxtime`. For performing tapered kriging see also Nychka, Furrer, and Sain (2014).

3.2. Binary case

Assume that $z(\mathbf{s}_1, t_1), \dots, z(\mathbf{s}_k, t_l)$ is a sample from a spatio-temporal binary random field (see Section 2.2), where $z(\mathbf{s}_j, t_i) = 1$ ($z(\mathbf{s}_j, t_i) = 0$) represents the success (failure) events. In this case, the ML estimator is unknown in closed form and the numerical maximization

of the full likelihood is too computationally demanding, also for moderate sample sizes. For this reason, [Heagerty and Lele \(1998\)](#) proposed performing the parameter estimation of the probit model (2), using the pairwise marginal likelihood

$$\begin{aligned} \ell_{pm}(\boldsymbol{\theta}) = & \sum_{(i,j,q,r) \in D} z(\mathbf{s}_j, t_i) z(\mathbf{s}_r, t_q) \log(p_{11}(\|\mathbf{s}_j - \mathbf{s}_r\|, |t_i - t_q|)) \\ & + \{z(\mathbf{s}_j, t_i) + z(\mathbf{s}_r, t_q) - 2z(\mathbf{s}_j, t_i)z(\mathbf{s}_r, t_q)\} \log(p_{01}(\|\mathbf{s}_j - \mathbf{s}_r\|, |t_i - t_q|)) \\ & + \{1 - z(\mathbf{s}_j, t_i)\} \{1 - z(\mathbf{s}_r, t_q)\} \log(p_{00}(\|\mathbf{s}_j - \mathbf{s}_r\|, |t_i - t_q|)), \end{aligned} \quad (18)$$

where D is as in (17) and the probabilities $p_{xy}(\cdot, \cdot)$, $x, y = 0, 1$, are described in (3). MCL fitting can be obtained with the routine `FitComposite` specifying the values `model = "BinaryGauss"`, `likelihood = "Marginal"` and `type = "Pairwise"`. The conditional likelihood (12) is also available and it can be used by setting `likelihood = "Conditional"`. Observe that, for simplicity and identifiability we follow the parametrization of [Heagerty and Lele \(1998\)](#), where a constant mean is assumed, $\tau^2 + \sigma^2 = 1$, $0 < \sigma^2 < 1$, $\tau^2 = 1 - \sigma^2$ and $b = 0$. Here is a practical example.

```
R> set.seed(11)
R> x <- seq(0, 10, 1)
R> y <- seq(0, 10, 1)
R> times <- seq(1, 15, 1)
R> param <- list(nugget = 0, mean = 0, scale_t = 1.5, scale_s = 2,
+   sill = 0.6)
R> data <- RFsim(x, y, times, corrmmodel = "exp_exp", model = "BinaryGauss",
+   param = param, grid = TRUE)$data
```

This code shows how to simulate a binary random field with the routine `RFsim`. Since the process is based on a latent Gaussian RF, a correlation function and its parameters need to be specified. Figure 5 shows the simulation for the first four temporal instants (top-left to bottom-right), performed on an 11×11 grid on $[0, 10]^2$ and for 15 times. The dependence can be estimated using the lorelogram function ([Heagerty and Zeger 1998](#)). The lorelogram estimator is defined as logarithm of the pairwise odds ratio,

$$\pi(\|\mathbf{s}_j - \mathbf{s}_r\|, |t_i - t_q|) = \log \left(\frac{\{p_{11}(\|\mathbf{s}_j - \mathbf{s}_r\|, |t_i - t_q|)\} \{p_{00}(\|\mathbf{s}_j - \mathbf{s}_r\|, |t_i - t_q|)\}}{\{p_{10}(\|\mathbf{s}_j - \mathbf{s}_r\|, |t_i - t_q|)\} \{p_{01}(\|\mathbf{s}_j - \mathbf{s}_r\|, |t_i - t_q|)\}} \right). \quad (19)$$

In the binary case, it provides a better measure of dependence than the covariance function, for instance it allows ease of interpretation and its upper and lower bounds do not depend on the mean. The following code shows how to compute the empirical estimator of (19) using the routine `EVariogram`.

```
R> lor <- EVariogram(data, x, y, times, type = "lorelogram", grid = TRUE,
+   maxdist = 10, maxtime = 10)

R> pml <- FitComposite(data, x, y, times, corrmmodel = "exp_exp",
+   grid = TRUE, maxdist = 2, maxtime = 4, model = "BinaryGauss",
+   fixed = list(nugget = 0), start = list(mean = 0, scale_s = 2,
+   scale_t = 1.5, sill = 0.5))
R> pml$param
```

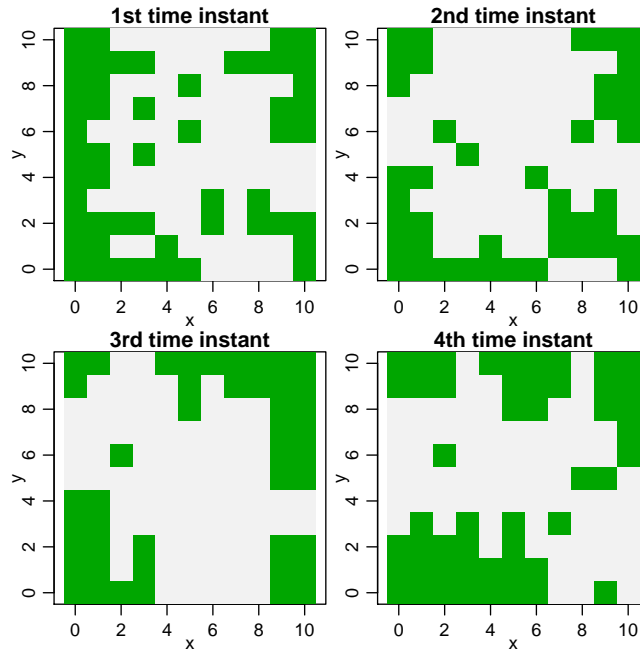


Figure 5: Simulations of spatio-temporal binary random fields on an 11×11 grid. From left to right panels, the simulations of 4 consecutive times are reported. The green and gray colors represent the cases $Y(\mathbf{s}, t) = 1$ and $Y(\mathbf{s}, t) = 0$.

```

      mean  scale_s  scale_t  sill
0.0162018 2.1896897 1.8202424 0.9560474

```

```

R> library("scatterplot3d")
R> Covariogram(pml, vario = lor, show.vario = TRUE)

```

The above code shows how to compute a parametric estimate of the lorelogram. Firstly, the `FitComposite` routine computes the parameter estimates of the underlying latent Gaussian random field. With `print(pml)`, the summary of the fitting results are returned (omitted here for brevity) and with `pml$param`, only the parameter estimates are displayed. Secondly, to compute a smooth estimate of the lorelogram, based on the output obtained from `FitComposite`, the routine `Covariogram` is used. This also displays the estimated function (using `scatterplot3d`, Ligges and Mächler 2003). Figure 6 shows the results. The left and right panels display the non-parametric and parametric estimates of the spatio-temporal lorelogram versus spatial distances and temporal intervals.

3.3. Max-stable case

Assume that independent copies, $\mathbf{z}_1, \dots, \mathbf{z}_m$, of the vector of spatial componentwise maxima $\mathbf{z} = \{z(\mathbf{s}_1), \dots, z(\mathbf{s}_k)\}^\top$, are available. Examples of real data analyses using this type of extremes are Padoan *et al.* (2010), Davison and Blanchet (2011) and Davison *et al.* (2012). Suppose for simplicity that the random RF behavior of the spatial maxima is well described by one of the classes of spatial max-stable RFs described in Section 2.3. For those families, although the k -dimensional distribution is known (Hüsler and Reiss 1989; Opitz 2013), the derivation of

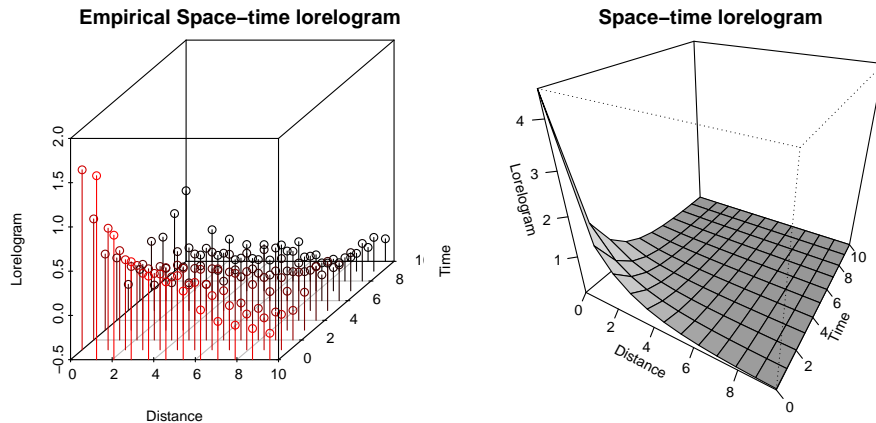


Figure 6: Spatio-temporal lorelogram versus spatial distances and time intervals. The left and right panels report the empirical and parametric estimates of the lorelogram

their relative density functions is unfeasible in practice. Indeed, given the particular form of a max-stable distribution, see (4), for deriving the density, a k th Bell number of derivatives is required to compute it (Davison and Gholamrezaee 2012). Nevertheless, the bivariate density is still feasible to compute in practice and it has the form

$$\exp\{-V(z_1, z_2)\}[\{\partial V(z_1, z_2)/\partial z_1\}\{\partial V(z_1, z_2)/\partial z_2\} - \partial^2 V(z_1, z_2)/\partial z_1 \partial z_2].$$

Taking the first and second order derivatives of (5)–(7) allows us to obtain the Brown–Resnick, the extremal-Gaussian and extremal- t bivariate densities. Therefore, for inference the full likelihood can be replaced by the pairwise marginal likelihood (12). Details on CL inference for max-stable random fields are provided by Padoan *et al.* (2010) and Davison *et al.* (2012). Here, practical examples are illustrated.

```
R> library("RandomFields")
R> set.seed(3132)
R> coords <- cbind(runif(100, 0, 10), runif(100, 0, 10))
R> param <- list(mean = 0, sill = 1, nugget = 0, scale = 1.5, power = 1.2)
R> extg <- RFsim(coords, corrmodel = "stable", model = "ExtGauss",
+   replicates = 30, param = param)$data
```

This code shows how to simulate an extremal-Gaussian RF with unit Fréchet margins with the routine `RFsim`. The simulation is performed using the routine `MaxStableRF` developed in the R package **RandomFields** (Schlather, Malinowski, Menck, Oesting, and Strokorb 2015). The correlation model and the parameters for the extremal-Gaussian (Schlather 2002) can be specified with `corrmodel` and `param`. The next code shows how to perform the composite likelihood-fitting of the extremal-Gaussian, using the routine `FitComposite`.

```
R> fextg <- FitComposite(extg, coords, corrmodel = "stable",
+   model = "ExtGauss", replicates = 30, varest = TRUE,
+   vartype = "Sampling", margins = "Frechet", fixed = list(sill = 1),
+   start = list(scale = 1.5, power = 1.2))
R> fextg$param
```



```

power    scale
0.9156696 1.7135199

```

With `print(fextg)` the summary of the fitting results is returned (omitted here for brevity) and with `fextg$param` only the parameter estimates are displayed. The standard errors of the estimates are computed and specifying `vartype = "Sampling"` the matrices H and J in (10) are computed using their sampling version. Note that other correlation models, such as the Whittle-Matérn or the generalized Cauchy (see Table 1) can also be fitted. The next code shows how to simulate and fit also the Brown-Resnick and Extremal- t models.

```

R> bres <- RFSim(coords, corrmodel = "stable", model = "BrowResn",
+   param = param, numblock = 1000, replicates = 30)$data
R> extt <- RFSim(coords, corrmodel = "stable", model = "ExtT",
+   param = c(param, df = 3), numblock = 1000, replicates = 30)$data
R> fbres <- FitComposite(bres, coords, corrmodel = "stable",
+   model = "BrowResn", replicates = 30, varest = TRUE,
+   vartype = "Sampling", margins = "Frechet",
+   start = list(scale = 1.5, power = 1.2))
R> fextt <- FitComposite(extt, coords, corrmodel = "stable",
+   model = "ExtT", replicates = 30, varest = TRUE,
+   vartype = "Sampling", margins = "Frechet",
+   fixed = list(sill = 1), start = list(df = 3, scale = 1.5, power = 1.2))

```

The first two commands perform simulations from the two respective random fields. Simulations are obtained computing the pointwise maximum (appropriately rescaled) over a number of replications of the Gaussian or Student- t field. The number of replications is specified with `numblock`.

Max-stable RF can also be fitted using the least squares method proposed by [Smith \(1990\)](#). The following code shows an example with the Brown-Resnick case.

```

R> wlbr <- WLeastSquare(bres, coords, corrmodel = "stable",
+   model = "BrowResn", replicates = 30, fixed = list(sill = 1))

```

The extremal coefficient functions in (8) can be computed on the basis of the fitting results. These are useful for summarizing the extremal dependence. The next code shows how to use the routine `Covariogram` for computing the extremal coefficients for the case of the Brown-Resnick RF.

```

R> extcbr <- Covariogram(fbres, answer.extc = TRUE)
R> extcbrw <- Covariogram(wlbr, answer.extc = TRUE)

```

Non-parametric estimators of the extremal coefficient are also available. A useful estimator is obtained by means of the madogram and F -madogram introduced by [Cooley, Neveau, and Poncet \(2006\)](#). For instance, the empirical version of F -madogram is

$$\hat{\nu}(h) = \frac{1}{2m|\mathcal{P}_h|} \sum_{\{z(\mathbf{s}_j), z(\mathbf{s}_r)\} \in \mathcal{P}_h} \sum_{i=1}^m |F\{z^{(i)}(\mathbf{s}_j)\} - F\{z^{(i)}(\mathbf{s}_r)\}|, \quad j, r \in J, h > 0,$$

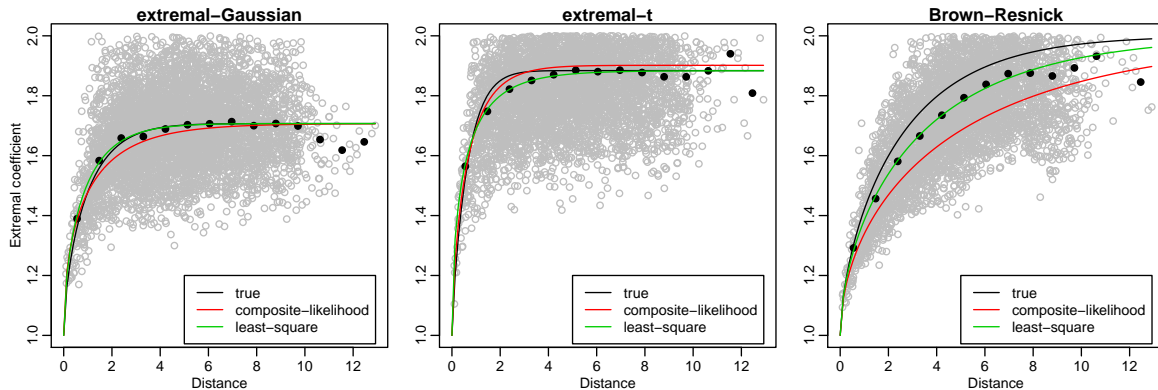


Figure 7: Spatial extremal coefficients versus distances. Gray circles are the estimates of the pairwise extremal coefficients obtained with the F-madogram. Black points are estimates obtained with the binned F-madogram. The black line is the true extremal coefficient function. The red and green lines are estimates of the extremal coefficient obtained with the composite-likelihood and least-square methods.

where F denotes the unit Fréchet distribution, $z^{(i)}(\mathbf{s}_j)$ is the i th sample at location \mathbf{s}_j (similarly for $z^{(i)}(\mathbf{s}_r)$), \mathcal{P}_h is the set of all the sample pairs $\{z(\mathbf{s}_j), z(\mathbf{s}_r)\}$ whose distances $\|\mathbf{h}\|$ fall in the interval $(h - \epsilon, h + \epsilon)$ with $\epsilon > 0$ and $|\mathcal{P}_h|$ is the cardinality of this set. The binned version of the estimator is computed for a sequence of distances, h , that are the centers of a series of equally spaced intervals that divide the range of the pairwise distances. The unbinned or cloud version is computed with the sequence of observed distances, in this case $|\mathcal{P}_{\|\mathbf{s}_j - \mathbf{s}_r\|}| = 1$ for every observed distance $\|\mathbf{s}_j - \mathbf{s}_r\|$. An estimate of the extremal coefficient is obtained by $\hat{\xi}(h) = \{1 + 2\hat{\nu}(h)\} / \{1 - 2\hat{\nu}(h)\}$. The next code shows how to obtain a non-parametric estimate of the extremal coefficient with the routine `EVariogram`.

```
R> fmcbr <- EVariogram(bres, coords, type = "Fmadogram",
+   replicates = 30, cloud = TRUE)
R> fmbbr <- EVariogram(bres, coords, type = "Fmadogram",
+   replicates = 30, numbins = 10)
```

The pairwise extremal coefficients (first line) and the binned version (second line) with 10 equally spaced bins are computed. With standard R commands, the parametric and non-parametric estimates of the extremal coefficient can be graphically represented. Figure 7 presents the results. The panels report, for the three max-stable RFs, the evolution of the estimated extremal coefficients versus the spatial distances. We see that estimates provided by the binned F -madogram, the CL and the least squares methods are quite similar, as expected.

4. Real data example: Irish wind speed data

We consider daily wind speeds collected over 18 years, from 1961 to 1978, at 12 sites in Ireland (Figure 8). These have originally been analyzed in [Haslett and Raftery \(1989\)](#). Following [Gneiting, Genton, and Guttorp \(2007\)](#) we omit the Rosslane station, we consider the square root transformation of the data and we remove the seasonal component. The latter is estimated by calculating the average of the square roots of the daily means over all years and

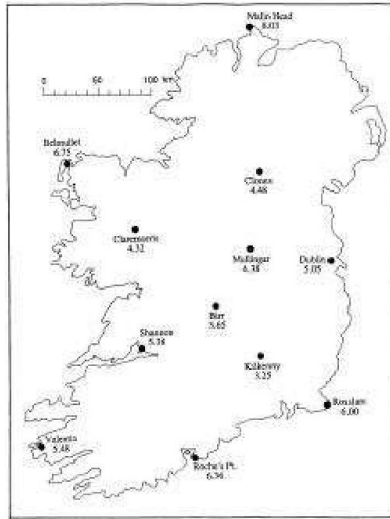


Figure 8: Irish map. The points are the locations where the wind speed data are recorded.

stations and regressing this on a set of annual harmonics. The resulting transformed data, $\{z(s_j, t_i), j = 1, \dots, 11, i = 1 \dots 6574\}$, is assumed to be a realization from a zero-mean Gaussian spatio-temporal RF. For comparison purposes we also perform ML estimation, then, in order to reduce the computational time, we focus our analysis on considering the first half of the year 1962. So, we have $11 \times 183 = 2013$ observations and ML estimation is still feasible, although very slow. We start the analysis, first estimating non-parametrically the dependence structure. The next code shows how to estimate the empirical spatio-temporal variogram with the routine `EVariogram`.

```
R> data("irishwinds")
R> coords <- as.matrix(winds.coords[-3, 5:4], ncol = 2)
R> times <- (365 + 1):(365 + 183)
R> data <- winds[times, ]
R> evar <- EVariogram(data, coords, coordt = times, maxtime = 7,
+   distance = "Chor")
```

If the coordinates of the location sites are given in latitude and longitude, as in this case, it is possible to compute the chordal or the geodesic distance in km ([Banerjee 2005](#); [Nychka et al. 2014](#)) setting `distance = "Chor"` or `distance = "Geod"`. Note that only some of the correlations models described in [Table 1](#) are still valid using the geodesic distance instead of the euclidean distance. For instance the Matern and the Cauchy models are still valid using the geodesic distance under some restrictions on the parameters ([Gneiting 2013](#)).

For simplicity, setting `maxtime` we restrict the computation of the spatio-temporal variogram focusing only on times whose pairwise intervals are smaller or equal to 7 days. [Figure 9](#) reports an illustration of the estimated variogram. The code to reproduce the plots will be discussed in the next paragraph for comparison purposes with the parametric estimate of the variogram. The top-left panel shows, with points, the estimated empirical spatio-temporal variogram (vertical axis) versus spatial distances (horizontal axis) and time intervals (oblique

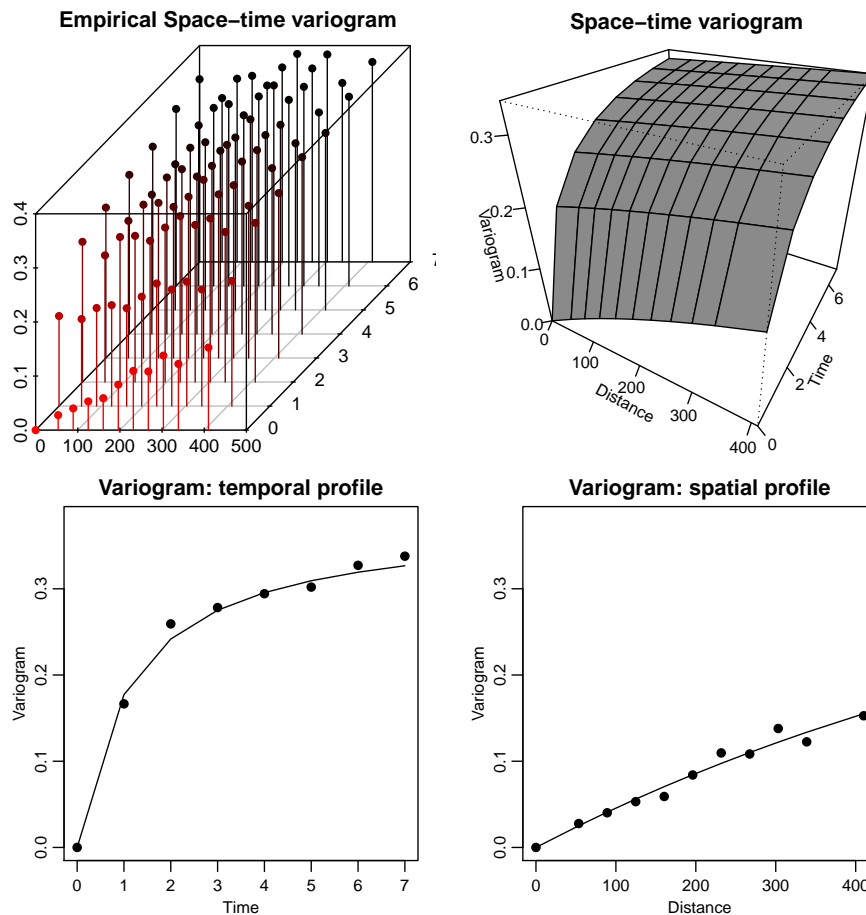


Figure 9: Spatio-temporal variogram versus distances and times. The top left and right panels report the empirical and parametric estimates. The bottom panels show their marginal temporal and spatial sections for a fixed distance and time. Points represent the empirical variogram and the lines, the parametric estimate.

axis). Estimates are in matrix format, where each row (column) is a temporal (spatial) variogram for a given spatial distance (temporal instant). Bottom-left and right panels show, with dots, respectively a spatial and temporal variogram for the fixed temporal interval equal to 0 days and the spatial distance equal to 0 km. Both are simply a section of the spatio-temporal variogram.

We consider now a parametric estimate of the variogram. We focus on the Gneiting (Gneiting 2002) and the double exponential model, see Table 1. The first is a non-separable model for $0 < \eta \leq 1$, while it is separable for $\eta = 0$ and the second is separable. The aim is to investigate if, for these data, the dependence structure is separable. Observe that, when jointly estimating all the parameters, the likelihood surface is multimodal, so in order to avoid this we consider fixed η . However, we estimate the correlation parameters for different values of the separability parameter η in $[0, 1]$. Using the `FitComposite` routine, the covariance models are fitted with maximum composite, tapering and standard likelihood estimators. The next code illustrates the case $\eta = 0.5$. The code for the other values of η and the double exponential covariance is similar.

```

R> cormod <- "gneiting"
R> fixed <- list(mean = 0, nugget = 0, power_s = 1, power_t = 1, sep = 0.5)
R> start <- list(scale_s = 600, scale_t = 1, sill = var(c(data)))
R> wls <- WLeastSquare(data, coords, coordt = times, corrmmodel = cormod,
+   maxdist = 400, maxtime = 7, fixed = fixed, start = start,
+   distance = "Chor", weighted = TRUE)
R> start <- as.list(wls$param)
R> CL <- FitComposite(data, coords, coordt = times, corrmmodel = cormod,
+   likelihood = "Marginal", type = "Pairwise", maxdist = 400, maxtime = 4,
+   fixed = fixed, start = start, distance = "Chor", varest = TRUE)
R> start <- as.list(CL$param)
R> TP <- FitComposite(data, coords, coordt = times, corrmmodel = cormod,
+   likelihood = "Full", type = "Tapering", taper = "Wendland2_Wendland2",
+   maxtime = 4, maxdist = 400, fixed = fixed, start = start,
+   distance = "Chor", varest = TRUE)
R> start <- as.list(TP$param)
R> ML <- FitComposite(data, coords, coordt = times, corrmmodel = cormod,
+   likelihood = "Full", type = "Standard", fixed = fixed, start = start,
+   distance = "Chor", varest = TRUE)

```

On the fourth line a preliminary estimate with the weighted least square method (Diggle and Ribeiro 2007) is obtained. This is used as a starting value for the optimization of the likelihood-based methods. On the sixth line the marginal pairwise likelihood estimation is performed. With `maxdist` and `maxtime` pairs of observations that are more than 400 km far apart and separated by more than four days, are excluded. On the eighth line the tapering likelihood estimation, using a Wendland space-time separable taper function (Table 2), is performed. The setting is the same as the CL. This leads to a tapered covariance matrix with 4.4% of no-zero values. On the tenth line ordinary likelihood estimation is performed.

Table 3 presents the estimation results. The three estimation methods provide similar results and the standard ML method exhibits the lower standard errors as expected. However, the CL method is computationally the fastest. Indeed, with these data the numerical evaluation of the standard, tapering and composite likelihood requires respectively an elapsed time of 3.4, 1.1 and 0.05 (the time is computed with the `system.time` routine of R installed in a machine with a 2.27 GHz processor and 6 GB of memory). Therefore, it is 68 (22) times faster to evaluate the composite than the standard (tapering) likelihood. In the sixth column of Table 3, the AIC/CLIC values suggest that with a sample of data, the separable model performs the best fitting. This stresses that the underlying RF can be considered as the product of a purely spatial and temporal RF. The next code shows how to compute and graphically represent the spatio-temporal variogram with the `Covariogram` routine.

```

R> Covariogram(CL, vario = evar, fix.lagt = 1, fix.lags = 1,
+   show.vario = TRUE, pch = 19)

```

Both parametric or nonparametric estimates of the variogram are represented. The top-right panel of Figure 9 displays the variogram estimated with the CL. The bottom-left and right panels show, with the solid lines, an estimate of the spatial and temporal variogram conditionally to the fixed temporal interval equal to 0 days and the spatial distance equal to

Method	σ^2	ψ_s	ψ_t	AIC/CLIC
$\eta = 0$				
Likelihood	0.35 (0.03)	674 (20)	0.71 (0.02)	223
Composite	0.38 (0.02)	787 (37)	1.22 (0.18)	332979
Tapering	0.38 (0.04)	685 (28)	1.18 (0.08)	
$\eta = 0.5$				
Likelihood	0.37 (0.03)	716 (27)	0.96 (0.05)	231
Composite	0.38 (0.02)	782 (48)	1.15 (0.16)	332991
Tapering	0.39 (0.05)	711 (39)	1.25 (0.10)	
$\eta = 1$				
Likelihood	0.39 (0.04)	774 (52)	1.22 (0.08)	241
Composite	0.38 (0.02)	780 (68)	1.09 (0.15)	333004
Tapering	0.39 (0.05)	736 (58)	1.31 (0.11)	

Table 3: Fits of the spatio-temporal Gaussian process with Gneiting models to Irish winds data. Results are reported for the separability parameter $\eta = 0, 0.5, 1$. The first two columns give the estimation method and the correlation function used. From the third to the sixth columns the parameter estimates (standard errors) are reported; σ^2 , ψ_s , ψ_t are the sill, the spatial scale and temporal scale parameters. AIC/CLIC is the information criterion or composite information criterion.

0 km. From the top and bottom panels, we can see by comparing the dots with the lines that empirical and model-based variogram estimators provide similar results.

Finally the spatio-temporal simple kriging is performed using the estimated Gneiting correlation model (with $\eta = 0$). The prediction sets are the ungauged sites within the coastlines of Ireland for two consecutive times, in particular the days 549 and 550. The next code shows how to obtain the prediction with the `Kri` routine.

```
R> lon <- seq(min(coords[, 1]) - 0.3, max(coords[, 1]) + 0.85, 0.06)
R> lat <- seq(min(coords[, 2]) - 0.3, max(coords[, 2]) + 0.13, 0.06)
R> pr.s <- as.matrix(expand.grid(lon, lat))
R> pr.t <- (times[length(times)] + 1):(times[length(times)] + 2)
R> par <- as.list(c(CL$param, fixed))
R> pr <- Kri(data, coords, coordt = times, corrmodel = cormod,
+   distance = "Chor", loc = pr.s, param = par, time = pr.t)
```

Figure 10 displays the prediction maps for the two selected days. The plots have been done with standard R code and using the `map` routine of the package `mapproj` (McIlroy, Brownrigg, Minka, and Bivand 2014). Specifically, the top panels show the prediction maps for the two consecutive days. We can see that high wind speeds are concentrated between the north-east to the center of the island, decreasing when approaching the south and west-coast. The bottom panels show the standard error prediction maps. As expected, we can firstly see that the standard errors decrease in the parts of the island where the gauging weather stations are located, second, the standard errors increase from the first to the second predicted day.

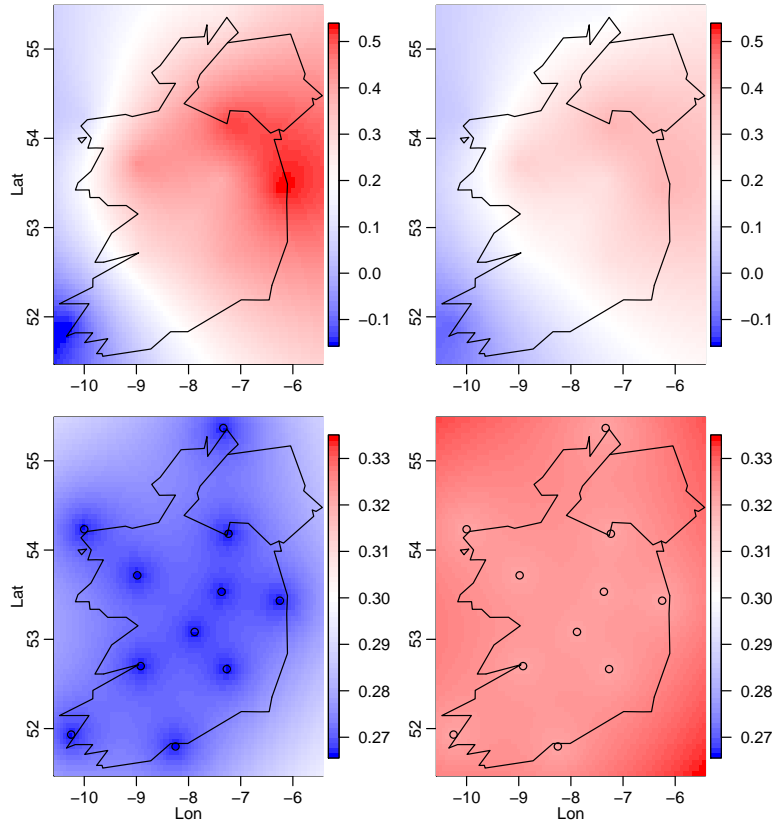


Figure 10: Ireland wind speed prediction (top-panels) and standard error prediction (bottom-panels) maps for two consecutive temporal instants 549, 550 (left-right). Points denote the gauging stations where the wind speed is recorded.

5. Conclusion

The composite likelihood is a wide class of estimating functions that generalizes the usual standard likelihood and includes different alternatives of the latter. The wide range likelihood versions that it provides makes it a very flexible estimation method that in recent years has attracted growing interest. There are many statistical fields and applications where it has been successfully used, see [Varin *et al.* \(2011\)](#). Spatial and spatio-temporal applications are among these.

In this paper we have discussed the R package **CompRandFld** for applying composite likelihood inferential methods for the analysis of spatial and spatio-temporal data. At the moment, the code available in this package can serve the community that is interested in statistical analysis of spatio-temporal Gaussian and binary data, and spatial extremes. We aim to expand the package in different directions. For example, including other types of composite likelihoods such as the triwise marginal likelihood, mixture of pairwise marginal and univariate likelihoods and marginal likelihoods of blocks. In addition, we intend to implement composite likelihoods for categorical and counting dependent data and extend the code in order to estimate multivariate Gaussian RFs.

Acknowledgments

Research work of Moreno Bevilacqua was partially supported by grant FONDECYT 11121408 from the Chilean government. We thank the Associate Editors and two anonymous referees for their careful reading of the entire article and their valuable suggestions. The authors are grateful to the CRAN team for their help.

References

- Banerjee S (2005). “On Geodetic Distance Computations in Spatial Modeling.” *Biometrics*, **61**, 617–625.
- Bevilacqua M, Gaetan C (2014). “Comparing Composite Likelihood Methods Based on Pairs for Spatial Gaussian Random Fields.” *Statistics and Computing*. doi:10.1007/s11222-014-9460-6. Forthcoming.
- Bevilacqua M, Gaetan C, Mateu J, Porcu E (2012). “Estimating Space and Space-Time Covariance Functions: a Weighted Composite Likelihood Approach.” *Journal of the American Statistical Association*, **107**(497), 268–280.
- Chandler RE, Bate S (2007). “Inference for Clustered Data Using the Independence Loglikelihood.” *Biometrika*, **94**, 167–183.
- Cooley D, Neveau P, Poncet P (2006). *Variograms for Spatial Max-Stable Random Fields*, volume 187 of *Lecture Notes in Statistics*. Springer-Verlag.
- Cox DR, Reid N (2004). “A Note on Pseudolikelihood Constructed from Marginal Densities.” *Biometrika*, **91**, 729–737.
- Cressie N, Johannesson G (2008). “Fixed Rank Kriging for Very Large Spatial Data Sets.” *Journal of the Royal Statistical Society B*, **70**, 209–226.
- Cressie NA, Wikle CK (2011). *Statistics for Spatial-Temporal Data*. Probability and Statistics. John Wiley & Sons.
- Davis RA, Klüppelberg C, Steinkohl C (2013). “Max-Stable Processes for Modeling Extremes Observed in Space and Time.” *Journal of the Korean Statistical Society*, **42**(3), 399–414.
- Davison AC, Blanchet J (2011). “Spatial Modelling of Extreme Snow Depth.” *The Annals of Applied Statistics*, **5**(1699–1725).
- Davison AC, Gholamrezaee MM (2012). “Geostatistics of Extremes.” *Proceedings of the Royal Society of London A*, **468**(581–608).
- Davison AC, Padoan SA, Ribatet M (2012). “Statistical Modelling of Spatial Extremes.” *Statistical Science*, **27**(2), 161–186.
- de Haan L (1984). “A Spectral Representation for Max-Stable Processes.” *The Annals of Probability*, **12**, 1194–1204.

- de Haan L, Ferreira A (2006). *Extreme Value Theory: An Introduction*. Springer-Verlag.
- De Iaco S, Myers DE, Posa T (2002). “Nonseparable Space-Time Covariance Models: Some Parametric Families.” *Mathematical Geology*, **34**, 23–42.
- de Leon AR (2005). “Pairwise Likelihood Approach to Grouped Continuous Model and Its Extension.” *Statistics and Probability Letters*, **75**(49–57).
- Diggle PJ, Ribeiro PJ (2007). *Model-Based Geostatistics*. Springer Series in Statistics. Springer-Verlag.
- Fuentes M (2007). “Approximate Likelihood for Large Irregularly Spaced Spatial Data.” *Journal of the American Statistical Association*, **102**, 321–331.
- Furrer R (2014). *spam: Sparse Matrix*. R package version 1.0-1, URL <http://CRAN.R-project.org/package=spam>.
- Furrer R, Genton MG, Nychka D (2006). “Covariance Tapering for Interpolation of Large Spatial Datasetslarge Spatial Datasets.” *Journal of Computational and Graphical Statistics*, **15**, 502–523.
- Furrer R, Sain SR (2010). “**spam**: a Sparse Matrix R Package with Emphasis on MCMC Methods for Gaussian Markov Random Fields.” *Journal of Statistical Software*, **36**(10), 1–25. URL <http://www.jstatsoft.org/v36/i10/>.
- Gentz A (1992). “Numerical Computation of Multivariate Normal Probabilities.” *Journal of Computational and Graphical Statistics*, **1**, 141–150.
- Gneiting T (2002). “Nonseparable, Stationary Covariance Functions for Space-Time Data.” *Journal of the American Statistical Association*, **97**(458), 590–600.
- Gneiting T (2013). “Strictly and Non-Strictly Positive Definite Functions on Spheres.” *Bernoulli*, **19**, 1087–1500.
- Gneiting T, Genton MG, Guttorp P (2007). “Geostatistical Space-Time Models, Stationarity, Separability and Full Symmetry.” In B Finkenstadt, L Held, V Isham (eds.), *Statistical Methods for Spatio-Temporal Systems*, pp. 151–175. Chapman and Hall/CRC, Boca Raton.
- Gneiting T, Schlather M (2004). “Stochastic Models that Separate Fractal Dimension and the Hurst Effect.” *SIAM Review*, **46**, 269–282.
- Harville DA (1977). “Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems.” *Journal of the American Statistical Association*, **72**, 320–338.
- Haslett J, Raftery AE (1989). “Space-Time Modelling with Long-Memory Dependence: Assessing Ireland’s Wind Power Resource.” *Applied Statistics*, **38**(1–50).
- Heagerty P, Zeger SL (1998). “Lorelogram: A Regression Approach to Exploring Association in Longitudinal Categorical Responses.” *Journal of the American Statistical Association*, **93**(441), 150–162.
- Heagerty PJ, Lele SR (1998). “A Composite Likelihood Approach to Binary Spatial Data.” *Journal of the American Statistical Association*, **93**, 1099–1111.

- Hüsler J, Reiss R (1989). “Maxima of Normal Random Vectors: Between Independence and Complete Dependence.” *Statistics and Probability Letters*, **7**, 283–286.
- Kabluchko Z (2009). “Extremes of Space-Time Gaussian Processes.” *Stochastic Processes and Their Applications*, **119**, 3962–3980.
- Kabluchko Z (2010). “Extremes of Independent Gaussian Processes.” *Extremes*, **14**, 285–310.
- Kabluchko Z, Schlather M, de Haan L (2009). “Stationary Max-Stable Fields Associated to Negative Definite Functions.” *The Annals of Probability*, **37**(5), 2042–2065.
- Kaufman CG, Schervish MJ, Nychka DW (2008). “Covariance Tapering for Likelihood-Based Estimation in Large Spatial Data Sets.” *Journal of the Royal Statistical Society B*, **103**(484), 1545–1555.
- Ligges U, Mächler M (2003). “**scatterplot3d** – An R Package for Visualizing Multivariate Data.” *Journal of Statistical Software*, **8**(11), 1–20. URL <http://www.jstatsoft.org/v08/i11/>.
- Lindsay BG (1988). “Composite Likelihood Methods.” *Contemporary Mathematics*, **80**, 221–239.
- Mardia KV, Marshall RJ (1984). “Maximum Likelihood Estimation of Models for Residual Covariance in Spatial Regression.” *Biometrika*, **71**(1), 135–146.
- McIlroy D, Brownrigg R, Minka TP, Bivand R (2014). **mapproj**: *Map Projections*. R package version 1.2-2, URL <http://CRAN.R-project.org/package=mapproj>.
- Nott DJ, Rydén T (1999). “Pairwise Likelihood Methods for Inference in Image Models.” *Biometrika*, **86**, 661–676.
- Nychka D, Furrer R, Sain S (2014). **fields**: *Tools for Spatial Data*. R package version 7.1, URL <http://CRAN.R-project.org/package=fields>.
- Opitz T (2013). “Extremal t Processes: Elliptical Domain of Attraction and a Spectral Representation.” *Journal of Multivariate Analysis*, **122**, 409–413.
- Pace L, Salvan A, Sartori N (2011). “Adjusting Composite Likelihood Ratio Statistic.” *Statistica Sinica*, **21**, 129–148.
- Padoan SA, Bevilacqua M (2015). **CompRandFld**: *Composite likelihood based Analysis of Random Fields*. R package version 1.0.3-4, URL <http://CRAN.R-project.org/package=CompRandFld>.
- Padoan SA, Ribatet M, Sisson SA (2010). “Likelihood-Based Inference for Max-Stable Processes.” *Journal of the American Statistical Association*, **105**, 263–277.
- Porcu E, Mateu J, Christakos G (2009). “Quasi-Arithmetic Means of Covariance Functions with Potential Applications to Space-Time Data.” *Journal of Multivariate Analysis*, **8**(100), 1830–1844.
- Porcu E, Montero JM, Schlather M (eds.) (2012). *Advances and Challenges in Space-Time Modelling of Natural Events*. Lecture Notes in Statistics. Springer-Verlag.

- R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Sang H, Genton MG (2014). “Tapered Composite Likelihood for Spatial Max-Stable Models.” *Spatial Statistics*, **8**, 86–103.
- Schlather M (2002). “Models for Stationary Max-Stable Random Fields.” *Extremes*, **5**(33–44).
- Schlather M, Malinowski A, Menck PJ, Oesting M, Strokorb K (2015). “Analysis, Simulation and Prediction of Multivariate Random Fields with Package **RandomFields**.” *Journal of Statistical Software*, **63**(8), 1–25. URL <http://www.jstatsoft.org/v63/i08/>.
- Shaby B, Ruppert D (2012). “Tapered Covariance: Bayesian Estimation and Asymptotics.” *Journal of Computational and Graphical Statistics*, **21**, 433–452.
- Smith RL (1990). “Max-Stable Processes and Spatial Extremes.” Unpublished manuscript.
- Stein ML, Chi Z, Welty LJ (2004). “Approximating Likelihoods for Large Spatial Data Sets.” *Journal of the Royal Statistical Society B*, **66**, 275–296.
- Varin C, Reid N, Firth D (2011). “An Overview of Composite Likelihood Methods.” *Statistica Sinica*, **21**(1), 5–42.
- Varin C, Vidoni P (2005). “A Note on Composite Likelihood Inference and Model Selection.” *Biometrika*, **92**, 519–528.
- Wackernagel H (1998). *Multivariate Geostatistics: An Introduction with Applications*. Springer-Verlag.
- Wendland H (1995). “Piecewise Polynomial, Positive Definite and Compactly Supported Radial Functions of Minimal Degree.” *Advances in Computational Mathematics*, **4**, 389–396.

Affiliation:

Simone A. Padoan
Department of Decision Sciences
Bocconi University of Milan
20136 Milan, Italy
E-mail: simone.padoan@unibocconi.it
URL: <http://faculty.unibocconi.it/simonepadoan>

Moreno Bevilacqua
Department of Stastics
University of Valparaíso
Avenida Gran Bretaña, Playa Ancha, 1091, Chile
E-mail: moreno.bevilacqua@uv.cl