On modeling positive continuous data with spatio-temporal dependence

Moreno Bevilacqua

Department of Statistics

Universidad de Valparaiso, Valparaiso, Chile

and

Millennium Nucleus Center

for the Discovery of Structures in Complex Data, Chile

Christian Caamaño-Carrillo

Department of Statistics,

Universidad del Bío-Bío, Concepcion, Chile

Carlo Gaetan

Dipartimento di Scienze Ambientali, Informatica e Statistica,

Università Ca' Foscari di Venezia, Venice, Italy.

May 19, 2020

Abstract

In this paper, we concentrate on an alternative modeling strategy for positive data that exhibit spatial or spatio-temporal dependence. Specifically, we propose to consider stochastic processes obtained through a monotone transformation of scaled version of χ^2 random processes. The latter are well known in the specialized literature and originates by summing independent copies of a squared Gaussian process. However, their use as stochastic models and related inference has not been much considered. Motivated by a spatio-temporal analysis of wind speed data from a network of meteorological stations in the Netherlands, we exemplify our modeling strategy by means of a non-stationary process with Weibull marginal distributions. For the proposed Weibull process we study the second-order and geometrical properties and we provide analytic expressions for the bivariate distribution. Since the likelihood is intractable, even for a relatively small data-set, we suggest adopting the pairwise likelihood as a tool for inference. Moreover, we tackle the prediction problem and we propose to use a linear prediction. The effectiveness of our modeling strategy is illustrated by analyzing the aforementioned Netherland wind speed data that we integrate with a simulation study. The proposed method is implemented in the R package `GeoModels`.

Keywords: Copula; Linear Prediction; Non-Gaussian data; Pairwise likelihood; Regression model; Wind speed data.

1. INTRODUCTION

Climatology, Environmental Sciences and Engineering, to name some fields, show an increasing interest in the statistical analysis of spatial and/or temporal data. In order to model the inherent uncertainty of the data, Gaussian random processes play a fundamental role (see Cressie and Wikle, 2011, for instance). Indeed, the Gaussian random processes have to offer marginal and dependence modelling in terms of mean and covariance functions, methods of inference well studied and scalable for large dataset (Heaton et al., 2018) and optimality in the prediction (Stein, 1999).

However, data collected in a range of studies such as wind speeds (Pryor and Barthelmie, 2010), ocean surface currents (Galanis et al., 2012) and rainfalls (Neykov et al., 2014) take continuous positive values and exhibit skewed sampling distributions. In this case the Gaussian probability model becomes unrealistic.

Transformations of a Gaussian process, i.e. trans-Gaussian kriging (Cressie, 1993), is a general approach to model this kind of data by applying a nonlinear transformations to the original data. In the literature the most common transformations, the square root and the natural logarithm, belong to the Box-Cox power transformation (see Haslett and Raftery, 1989; Allcroft and Glasbey, 2003; Bessac et al., 2015, for instance). In particular, Log-Gaussian processes have been broadly used for the analysis of positive dependent data due to their well known mathematical properties (De Oliveira et al., 1997; De Oliveira, 2006). For an alternative (parametric) family of transformations of Gaussian processes, the Tukey g-and-h transformation, see Xu and Genton (2017), Yan and Genton (2019).

Nevertheless, it can be difficult to find an adequate non linear transformation and some appealing properties of the Gaussian process may not be inherited by the transformed process (Wallin and Bolin, 2015).

Another possibility is to resort to Gaussian copulas. Copula theory (Joe, 2014) allows joint distributions to be constructed from specified marginal continuous distributions for positive data. The role of the copula is to describe the spatio-temporal dependence structure between random variables without information on the marginal distributions. Even though which copula model to use for a given analysis is not generally known a priori, the

copula based on the multivariate Gaussian distribution (Kazianka and Pilz, 2010; Masarotto and Varin, 2012; Gräler, 2014) has gained a general consensus since the definition of the multivariate dependence relies again on the specification of the pairwise dependence, i.e. on the covariance function.

Actually the two aforementioned approach are strongly related since monotone transformations of a Gaussian process share the same copula model. As we will see in our real data example, the kind of dependence described by the Gaussian copula could be restrictive or unsuitable. In fact the Gaussian copula expresses a symmetrical dependence (see Section 2), i.e. high values exhibit a spatial/temporal dependence similar to low ones. Copula-based model using symmetrical dependence is still used in a recent paper (Tang et al., 2019) on spatio-temporal modelling wind speed data.

Concluding this short review we mention that Wallin and Bolin (2015) proposed recently non-Gaussian processes derived from stochastic partial differential equations. Nevertheless, their method is restricted to the spatial Matérn covariance model and its statistical properties are much less understood that the Gaussian process.

In this paper, we shall look at processes that are derived by Gaussian processes but differently from the trans-Gaussian random processes and the copula models we do not consider just one copy of the Gaussian process. We suggest to model positive continuous data by transforming χ^2 processes (Adler, 1981; Ma, 2010) i.e. a sum of squared of independent copies of a standard Gaussian process. Even though probabilistic properties of a sum of squared Gaussian processes have been studied several years ago, less attention has been paid to use this for statistical modelling of dependent positive data. We are convinced that the Gaussian processes offer an incomparable tool case for those who want to model the dependence between observations. However, we aim to overcome some aforementioned restrictions.

Motivated by a spatio-temporal analysis of daily wind speed data from a network of meteorological stations in the Netherlands, we exemplify our construction by proposing a non-stationary spatio-temporal process with asymmetric dependence and Weibull marginal distribution even though other stochastic processes with different marginal distributions

could be studied starting from transformations of χ^2 processes. In fact, in scientific literature a variety of probability distribution has been suggested to describe wind speed distributions and the Weibull model constitutes one of the most widely accepted (see Carta et al., 2009, for a review).

The proposed Weibull process is parametrized in such a way that both regression and dependence analysis can be jointly performed. Additionally the process inherits the geometrical properties of the underlying Gaussian process. This implies that mean-square continuity and differentiability, as in the Gaussian processes, can be modeled using suitable parametric correlation models such as the Matérn (Stein, 1999) or the Generalized Wendland (Bevilacqua et al., 2019) models, in the spatial case.

It must be said that it is the difficult to evaluate the multivariate density for the proposed model and this fact prevents the inference based on the full likelihood and the derivation of the analytical form of the predictor that minimizes the mean square prediction error. For this reason, we propose the use of a weighted version of the pairwise likelihood (Lindsay, 1988; Varin et al., 2011) for estimating the unknown parameters. Moreover a linear and unbiased predictor is proposed following the approach detailed in De Oliveira (2014).

In addition, we study a specific example where both the multivariate density and the optimal predictor have an explicit closed-form (see Section 5). For this example, we perform a simulation study with the goal of investigating the relative efficiency of the maximum weighted pairwise likelihood method with respect to the maximum likelihood method and the relative efficiency of the proposed linear predictor with respect to the optimal predictor. Simulation, estimation, and prediction of the Weibull process are implemented in a R package `GeoModels` (Bevilacqua and Morales-Oñate, 2019).

The remainder of the paper is organized as follows. In Section 2 we introduce the random processes and we describe their features. In Section 3 we concentrate on a random process with Weibull marginal distributions. Section 4 starts with a short description of the estimation method and ends with tackling the prediction problem. In Section 5 we report the numerical results of a simulation study and in Section 6 we apply our method to the daily wind speed data measurements from a network of meteorological stations in the Netherlands

using the Log-Gaussian process as the benchmark. Finally some concluding remarks are consigned to Section 7.

2. SCALED χ^2 RANDOM PROCESSES

2.1 Definition

We start by considering a 'parent' Gaussian random process $Z := \{Z(s), s \in S\}$, where s represents a location in the domain S. Spatial ($S \subseteq \mathbb{R}^d$) or spatio-temporal examples ($S \subseteq \mathbb{R}^d \times \mathbb{R}_+$) will be considered indifferently. We also assume that Z is stationary with zero mean, unit variance and correlation function $\rho(h) := \mathbb{C}or(Z(s+h), Z(s))$ where $s + h \in S$. Let $Z_1, \ldots, Z_m$ be $m = 1, 2, \ldots$ independent copies of Z and define the random process $X_m := \{X_m(s), s \in S\}$ as

$$X_m(s) := \sum_{k=1}^{m} Z_k(s)^2/m. \tag{1}$$

The stationary process X_m is a scaled version of a χ^2 random process (Adler, 1981; Ma, 2010) with marginal distribution Gamma$(m/2, m/2)$ where the pairs $m/2, m/2$ are the shape and rate parameters. By definition, $\mathbb{E}(X_m(s)) = 1$ and $\mathbb{V}ar(X_m(s)) = 2/m$ for all s.

The analytical expressions of the multivariate density of a vector of n observations $X_m(s_1) = x_1, \ldots, X_m(s_n) = x_n$ can be derived only in some special cases (Krishnamoorthy and Parthasarathy, 1951; Royen, 2004). An interesting example is made up for $s_1 < s_2 < \ldots < s_n$ locations on $S = \mathbb{R}$ and for a Gaussian process Z with exponential covariance function. In this case the multivariate density can be derived as

$$\begin{aligned} f_{X_m}(x_1, \ldots, x_n) \;=\; & \frac{m^{m/2-1+n} 2^{-m/2+1-n} (x_1 x_n)^{m/4-1/2}}{\Gamma(m/2) \prod_{i=1}^{n-1} \{(1-\rho_{i,i+1}^2)\rho_{i,i+1}^{m/2-1}\}} \\ & \times \exp\left[-\frac{m x_1}{2(1-\rho_{1,2}^2)} - \frac{m x_n}{2(1-\rho_{n-1,n}^2)} - \sum_{i=2}^{n-1} \frac{m(1-\rho_{i-1,i}^2\rho_{i,i+1}^2)x_i}{2(1-\rho_{i-1,i}^2)(1-\rho_{i,i+1}^2)} \right] \\ & \times \prod_{i=1}^{n-1} I_{m/2-1}\left(\frac{m\rho_{i,i+1}\sqrt{x_i x_{i+1}}}{(1-\rho_{i,i+1}^2)} \right) \end{aligned} \tag{2}$$

with $\rho_{ij} := \exp\{-|s_i - s_j|/\phi\}$, $\phi > 0$ and $I_a(x)$ the modified Bessel function of the first kind of order a.

The evaluation of the bivariate densities of a pair of observations $X_m(s_1)$ and $X_m(s_2)$ can be derived irrespective of the dimension of the space S and the correlation function (Vere-Jones, 1967). The bivariate distribution of X_m is known as the Kibble bivariate Gamma distribution (Kibble, 1941) with density

$$\begin{aligned} f_{X_m}(x_1, x_2) &= \frac{2^{-m} m^m (x_1 x_2)^{m/2-1}}{\Gamma(m/2)(1-\rho^2)^{m/2}} \left(\frac{m|\rho|\sqrt{x_1 x_2}}{2(1-\rho^2)} \right)^{1-m/2} \exp\left\{ -\frac{m(x_1+x_2)}{2(1-\rho^2)} \right\} \\ &\quad \times I_{m/2-1}\left(\frac{m|\rho|\sqrt{x_1 x_2}}{(1-\rho^2)} \right), \end{aligned} \tag{3}$$

where $\rho = \rho(s_1 - s_2)$.

2.2 Dependence structure

It is easy to show that the correlation function of X_m, $\rho_{X_m}(h)$, is equal to $\rho^2(h)$, the squared of the correlation function of the 'parent' Gaussian random process. However, a way for looking more deeply into the dependence structure between random variables regardless of the marginal distributions is inspecting their copulas (Joe, 2014). For a n-variate cumulative distribution function (cdf) $F(y_1, \ldots, y_n) := \Pr(Y_1 \leq y_1, \ldots, Y_n \leq y_n)$ with i-th univariate margin $F_i(y_i) := \Pr(Y_i \leq y_i)$, the copula associated with F is a cdf function $C : [0,1]^n \to [0,1]$ with $\mathcal{U}(0,1)$ margins that satisfies $F(y_1, \ldots, y_n) = C(F_1(y_1), \ldots, F_n(y_n))$. If F is a continuous cdf, with quantile functions F_i^{-1}, $i = 1, \ldots, n$, then Sklar's theorem (Sklar, 1959) guarantees that the cdf on the n-hypercube $C(u_1, \ldots, u_n) = F(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n))$ is the unique choice. The corresponding copula density, obtained by differentiation, is denoted by $c(u_1, \ldots, u_n)$.

Analogously, the multivariate survival function $\overline{F}(y_1, \ldots, y_n) := \Pr(Y_1 > y_1, \ldots, Y_n > y_n)$ could be expressed using the univariate survival functions $\overline{F}_i = 1 - F_i$ and the survival copula $\overline{F}(y_1, \ldots, y_n) = \overline{C}\left(\overline{F}_1(y_1), \ldots, \overline{F}_n(y_n)\right)$. The survival copula is in this way a distribution function on the n-hypercube $\overline{C}(u_1, \ldots, u_n) = \overline{F}(\overline{F}^{-1}(u_1), \ldots, \overline{F}^{-1}(u_n))$.

Among the copulas, the Gaussian copula is a convenient model for spatial data (Masarotto and Varin, 2012) as it offers a parametrization in terms of a correlation function. Let Φ^{-1} denote the quantile function of Φ the cdf of a standard Gaussian variable. The Gaussian copula with correlation matrix R is defined by $C(u_1, \ldots, u_n) = \Phi_R(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n))$, where Φ_R denotes the joint cumulative distribution function of an n-variate Gaussian random

vector with zero means and correlation matrix R.

The Gaussian copula is reflection symmetric, $C(u_1, \ldots, u_n) = \overline{C}(u_1, \ldots, u_n)$, that is the probability of having all variables less than their respective u-th quantile is equal to the probability of having all the variables greater than the complementary marginal quantile. Such property is a potential issue for data in which upper quantiles might exhibit a different pairwise spatial dependence than lower quantiles.

Although copula theory uses transformations to $\mathcal{U}(0,1)$ margins, it is better to consider $\mathcal{N}(0,1)$ margins (Joe, 2014, p. 9) for identifying the copula and for diagnostic purpose. In particular, the plot of the bivariate copula density can be compared with the Gaussian bivariate density and with the scatter-plot of pairs of observations on a normal scale, the *normal scores*. The bivariate copula density with $\mathcal{N}(0,1)$ margins is given by

$$c_N(z_1, z_2) = \frac{c(\Phi^{-1}(z_1), \Phi^{-1}(z_2))}{\phi(z_2)\phi(z_2)}$$

where $\Phi(z)$ (resp. $\phi(z)$) is the cdf (resp. pdf) of the standardized Gaussian distribution. Under this transform reflection symmetry means that the bivariate density contour plot is symmetric to the $(z_1, z_2) \to (-z_1, -z_2)$ reflection, i.e. $c_N(z_1, z_2) = c_N(-z_1, -z_2)$.

In Figure 1 we compare the contour plots of the bivariate copula density function entailed by (3) with the elliptical contours of the bivariate Gaussian density. Note that the copula for $m = 1$ is the copula introduced in Bárdossy (2006). Sharper corners (relative to ellipse) indicate more tail dependence of X_m than the Gaussian process and we notice also reflection asymmetries.

Dependence among extremal events can be summarized by the *upper tail dependence coefficient* (Sibuya, 1960; Coles et al., 1999) that looks at the limit behavior of the conditional probability of two random variables Y_1 and Y_2

$$\tau := \lim_{u \to 1^-} \frac{\Pr(F_1(Y_1) > u, F_2(Y_1) > u)}{\Pr(F_2(Y_1) > u)} = \lim_{u \to 1^-} \frac{1 - 2u + C(u,u)}{1-u} = \lim_{u \to 1^-} \frac{\overline{C}(1-u, 1-u)}{1-u}.$$

The value of the coefficient helps to distinguish between asymptotic dependence and asymptotic independence of the observations as the quantile increases. Under spatial asymptotic dependence, the likelihood of a large event happening in one location is tightly related to high values being recorded at a location nearby; the opposite is true under asymptotic indepen-

dence, in which large events might be recorded at one location only and not in neighboring locations (Wadsworth and Tawn, 2012).

We say that Y_1 and Y_2 are asymptotically dependent if $\tau > 0$ is positive. The case $\tau = 0$ characterizes asymptotic independence. Simply adapting Theorem 2.1 in Hashorva et al. (2014) we can prove that the X_m is asymptotically independent for all m, i.e. $\tau = 0$.

[Figure 1 about here.]

3. A RANDOM PROCESS WITH WEIBULL MARGINAL DISTRIBUTION

We focus our attention on stochastic modeling of wind speed data. In scientific literature, a variety of probability distribution has been suggested to describe wind speed frequency distributions (see Carta et al., 2009, for a review). Among them, the Weibull distribution constitutes one of the most widely accepted distribution for wind speed and it can be derived from a physical argument.

Suppose that the two orthogonal wind components (Z_1, Z_2) are assumed to be individually Gaussian with zero mean and independent, isotropic fluctuations. The distribution of the speed $V = \sqrt{Z_1^2 + Z_2^2}$ is the Rayleigh distribution i.e. the distribution V^2 is the exponential distribution (Johnson et al., 1995, pag. 417). The Weibull distribution can be obtained from the Rayleigh distribution through the power law transformation of V that has been shown to fit better wind speed samples due to its flexible form induced by the additional shape parameter.

Following this idea, let X_2 a special case of the rescaled χ^2 random process defined in Equation (1), with standard exponential marginal distribution. To obtain a stationary positive random process $W = \{W(s), s \in S\}$ with marginal Weibull distribution we consider the power transformation

$$W(s) := \nu(\kappa) X_2(s)^{1/\kappa}, \tag{4}$$

where $\nu(\kappa) = \Gamma^{-1}(1 + 1/\kappa)$ and $\kappa > 0$ is a shape parameter. Note that under this specific parametrization $W(s) \sim \text{Weibull}(\kappa, \nu(\kappa))$, $\mathbb{E}(W(s)) = 1$ and $\mathbb{V}ar(W(s)) = (\Gamma(1 + 2/\kappa)\nu^2(\kappa) - 1)$. In addition, the density of a pair of observations $W(s_1) = w_1$ and $W(s_2) = w_2$ is easily

obtained from (3) and (4), namely

$$f_W(w_1, w_2) = \frac{\kappa^2 (w_1 w_2)^{\kappa-1}}{\nu^{2\kappa}(\kappa)(1-\rho^2)} \exp\left[-\frac{w_1^\kappa + w_2^\kappa}{\nu^\kappa(\kappa)(1-\rho^2)}\right] I_0\left(\frac{2|\rho|(w_1 w_2)^{\kappa/2}}{\nu^\kappa(\kappa)(1-\rho^2)}\right). \quad (5)$$

Using Proposition 1 in Appendix we can also obtain the correlation function of W, namely

$$\rho_W(h) = \frac{\nu^{-2}(\kappa)}{\left[\Gamma\left(1+2/\kappa\right) - \nu^{-2}(\kappa)\right]} \left[{}_2F_1\left(-1/\kappa, -1/\kappa; 1; \rho^2(h)\right) - 1\right], \quad (6)$$

where the function

$${}_pF_q(a_1, a_2, \ldots, a_p; b_1, b_2, \ldots, b_q; x) := \sum_{k=0}^{\infty} \frac{(a_1)_k, (a_2)_k, \ldots, (a_p)_k}{(b_1)_k, (b_2)_k, \ldots, (b_q)_k} \frac{x^k}{k!} \quad \text{for} \quad p, q = 0, 1, 2, \ldots$$

is the generalized hypergeometric function (Gradshteyn and Ryzhik, 2007) and $(a)_k := \Gamma(a+k)/\Gamma(a)$, for $k \in \mathbb{N} \cup \{0\}$, is the Pochhammer symbol. Note that $\rho(h) = 0$ implies pairwise independence, as in the Gaussian case since (5) can be factorized in the product of two Weibull densities. Additionally, since ${}_2F_1(\cdot, \cdot, \cdot; 0) = 1$, $\rho(h) = 0$ implies $\rho_W(h) = 0$ that is if a compactly supported correlation function (Bevilacqua et al., 2019) is used as underlying correlation model, then also the correlation of the Weibull process is compactly supported. This feature is particularly appealing from the computational point of view since algorithms for sparse matrices can be used to handle the correlation matrix associated with ρ_W (see Section 4.2).

More important, it can be shown that some nice properties such as stationarity, mean-square continuity, degrees of mean-square differentiability and long-range dependence can be inherited from the 'parent' Gaussian process Z. In particular, using the results in Stein (1999, Section 2.4) linking the behavior of the correlation at the origin and the geometrical properties of the associated process, we can prove that W is mean square continuous if Z is mean square continuous and it is k-times mean-square differentiable if Z is k-times mean-square differentiable. Finally, it is trivial to see that the sample path continuity and differentiability are inherited from the 'parent' Gaussian process. As a consequence, mean-square continuity and differentiability of the sample paths of the Weibull process can be modeled using suitable flexible parametric correlation functions as in the case of the Gaussian processes.

As an illustrative example, Figure 2 collects three simulations of W on a fine grid of $S = [0,1]^2$ with Matérn correlation function $\rho(h) = 2^{1-\psi}\Gamma(\nu)^{-1}\left(\|h\|/\phi\right)^{\nu}\mathcal{K}_{\nu}\left(\|h\|/\phi\right)$, where $\phi, \nu > 0$ and $\mathcal{K}_a$ is a modified Bessel function of the second kind of order $a > 0$.

We have considered three different parametrization for the smoothness parameter $\nu = 0.5, 1.5, 2.5$. Under this setting, the paths of the 'parent' Gaussian process is $0, 1, 2-$times mean square differentiable, respectively. The values, $\phi = 0.067, 0.042, 0.034$, of the range parameter have been chosen in order to obtain a practical range, i.e. the distance at which the correlation $\rho(h)$, equal to 0.05 equal to 0.2. Additionally we have fixed the shape parameter of the Weibull distribution as $\kappa = 10, 3, 1$. The corresponding correlation functions $\rho_W(h)$ are plotted (from left to right) in the top panel of Figure 2 and the bottom panel reports the histograms of the observations.

It is apparent that the correlation $\rho_W(h)$ inherits the change of the differentiability at the origin from $\rho(h)$ when increasing ν. This changes have consequences on the geometrical properties of the associated random processes. In fact the smoothness of the realizations (central panel of Figure 2) increase with ν. Note also the flexibility of the Weibull model when modeling positive data in the bottom panel of Figure 2 since both positive and negative skewness can be achieved with different values of κ.

[Figure 2 about here.]

Finally, a non-stationary version of W can be easily obtained trough a multiplicative model:

$$Y(s) := \mu(s)W(s), \tag{7}$$

where $\mu(s) > 0$ is a non random function that specify the mean of Y *i.e.* $\mathbb{E}(Y(s)) = \mu(s)$ and affects its variance $\mathbb{V}ar(Y(s)) = \mu(s)^2(\Gamma\left(1 + 2/\kappa\right)\nu^2(\kappa) - 1)$.

A useful parametric specification for $\mu(s)$ is given through log-linear link function

$$\log(\mu(s)) = \beta_0 + \beta_1 v_1(s) + \cdots + \beta_p v_p(s)$$

where $v_i(s)$, $i = 1, \ldots, p$, are covariates and $\beta = (\beta_0, \ldots, \beta_p)^{\top}$ is a vector of regression parameters but other types of parametric or nonparametric functions can be considered.

Finally, note that given the observations $y(s_1), \ldots, y(s_n)$ and an estimation of the mean function $\hat{\mu}(s_i)$, the estimated residuals $\widehat{w}(s_i) = y(s_i)/\hat{\mu}(s_i)$ can be treated as a realization of the stationary Weibull process W with marginal distribution Weibull$(\kappa, \nu(\kappa))$ and correlation $\rho_W(h)$. This can be used to check the agreement between the distribution of the residuals and the estimated theoretical marginal distribution or to check the agreement between the theoretical estimated semi-variogram model obtained from (6) and its empirical counterpart (see Section 6).

4. ESTIMATION AND PREDICTION

4.1 Pairwise likelihood inference

Suppose that we have observed $y_1, \ldots, y_n$ at the locations $s_1, \ldots, s_n$ and let θ be the vector of unknown parameters for the Weibull random process (7). The evaluation of the full likelihood for θ is impracticable since, as outlined in Section 2, the multivariate density can be derived only in some special cases. A possible alternative (Lindsay, 1988; Varin et al., 2011) combines the bivariate distributions of all possible distinct pairs of observations (y_i, y_j). The weighted pairwise likelihood (WPL) function is given by

$$\mathrm{pl}(\theta) := \sum_{i=1}^{n} \sum_{j>i}^{n} \log f(y_i, y_j; \theta) c_{ij} \tag{8}$$

where $f(y_i, y_j; \theta)$ is the bivariate densities of (7) and c_{ij} are non-negative weights. The choice of cut-off weights, namely $c_{ij} = 1$ if $\|s_i - s_j\| \leq \Delta$, and 0 otherwise, for a positive value of Δ, can be motivated by its simplicity and by observing that that dependence between observations which are distant is weak. Therefore, the use of all pairs may skew the information confined in pairs of near observations (Davis and Yau, 2011; Bevilacqua and Gaetan, 2015).

Under the increasing domain asymptotics framework (Cressie, 1993) and arguing as in Bevilacqua and Gaetan (2015), it can be shown that the maximum weighted pairwise likelihood (MWPL) estimator $\widehat{\theta} := \mathrm{argmax}_\theta \, \mathrm{pl}(\theta)$ is consistent and asymptotically Gaussian. The asymptotic covariance matrix of the estimator is given by the inverse of the Godambe information

$$\mathcal{G}_n(\theta) := \mathcal{H}_n(\theta)^\top \mathcal{J}_n(\theta)^{-1} \mathcal{H}_n(\theta),$$

where $\mathcal{H}_n(\theta) := \mathbb{E}[-\nabla^2 \, \mathrm{pl}(\theta)]$ and $\mathcal{J}_n(\theta) := \mathbb{V}ar[\nabla \, \mathrm{pl}(\theta)]$.

The matrix $\mathcal{H}_n(\theta)$ can be estimated by $\widehat{\mathcal{H}} = -\nabla^2 \,\mathrm{pl}(\widehat{\theta})$ and the estimate $\widehat{\mathcal{J}}$ of $\mathcal{J}_n(\theta)$ can be calculated with a sub-sampling technique (Heagerty and Lele, 1998; Bevilacqua et al., 2012). Additionally, model selection can be performed by considering the pairwise likelihood information criterion (PLIC) (Varin and Vidoni, 2005)

$$\mathrm{PLIC} := -2\,\mathrm{pl}(\hat{\theta}) + 2\mathrm{tr}(\widehat{\mathcal{J}}\widehat{\mathcal{H}}^{-1})$$

which is the composite likelihood version of the Akaike information criterion (AIC).

4.2 Linear prediction

The lack of workable multivariate densities forestalls the use of the conditional distributions for the prediction. Therefore we choose a sub optimal solution, i.e. a linear predictor for the random variable $Y(s_0)$ at some unobserved location s_0 based on the data at locations $s_1, \ldots, s_n$, following a suggestion in Bellier et al. (2010) and De Oliveira (2014).

The predictor for the non-stationary Weibull process is given by

$$\widehat{Y}(s_0) := \mu(s_0) \left\{ 1 + \sum_{i=1}^{n} \lambda_i (W(s_i) - 1) \right\}, \tag{9}$$

where $W(s_i) = Y(s_i)/\mu(s_i)$. It is a linear predictor and unbiased predictor of $Y(s_0)$ for any vector of weights $\lambda = (\lambda_1, \ldots, \lambda_n)'$. The vector of weights $\lambda = (\lambda_1, \ldots, \lambda_n)'$ is set by minimizing the mean square error $\mathbb{E}[Y(s_0) - \hat{Y}(s_0)]^2$ with respect to λ. Note that this is a classical geostatistical approach applied to a multiplicative model instead of the classical additive model. It turns out that the solution for the predictor is given by the equations of the simple kriging, (Cressie, 1993, Chapter 3) i.e. $\lambda = C_W^{-1}\, c_W(s_0)$ and the associated mean square prediction error is given by

$$\mathbb{V}ar(\widehat{Y}(s_0)) := \mu^2(s_0)\sigma_W^2 \left\{ 1 - c_W(s_0)' C_W^{-1} c_W(s_0) \right\},$$

where $\sigma_W^2 := (\Gamma\,(1 + 2/\kappa)\,\nu^2(\kappa) - 1)$, $c_W(s_0) = (\rho_W(s_0 - s_i), \cdots, \rho_W(s_0 - s_n))'$ and C_W is the $n \times n$ correlation matrix whose (i, j)th element is $\rho_W(s_i - s_j)$ with $\rho_W(h)$ given in (6).

In practice the predictor cannot be evaluated since $\mu(s)$ and $\rho_W(h)$ are unknown. For this reason we suggest to use a plug-in estimate for $\mu(s)$ and $\rho_W(h)$ using the pairwise likelihood estimates.

5. SIMULATION RESULTS

In this section we investigate, through some numerical experiments, the relative efficiency of the MWPL estimator with respect to the maximum likelihood (ML) estimator and the relative efficiency of the linear predictor (9) with respect to the optimal predictor, under a specific setting of simulation where the comparisons can be explicitly performed. This specific setting is when the process is defined on $\mathbb{R}$ and the underlying correlation function is exponential. Even though this setting may seem artificial, the simulation study gives an idea of the relative efficiency of the MWPL estimation method and the proposed linear predictor under more general settings.

We have considered a non-stationary Weibull model (7) observed at 150 locations of a regular grid $0 = s_1 < s_2 < \cdots < s_{150} = 1$ where the 'parent' Gaussian random process has exponential correlation function $\rho_{i,j} := \rho(s_i - s_j) = \exp(-|s_i - s_j|/\phi)$.

In this case, the multivariate density function associated with the Weibull process is easily obtained from (2), namely

$$f_Y(y_1, \ldots, y_n) = \left\{ \frac{\kappa}{\nu(\kappa)^\kappa} \right\}^n f_{X_2}(x_1, \ldots x_n) \prod_{i=1}^{n} \frac{y_i}{\mu_i} \tag{10}$$

where $x_i := \{y_i/(\nu(\kappa)\mu_i)\}^\kappa$, $\mu_i := \mu(s_i)$ and f_{X_2} can be obtained from (2). Therefore this setup allows a comparison of the MWPL and ML estimation methods.

We set $\mu(s) = \exp\{\beta_0 + \beta_1 v_1(s)\}$ where $v_1(s)$ is a value from the $(0, 1)$-uniform distribution and $\beta_0 = 0.25$ and $\beta_1 = -0.15$. Three choices of the shape parameter $\kappa = 1, 3, 10$ are coupled with three values of the range parameter $\phi = a/3$, $a = 0.1, 0.2, 0.3$.

We simulate $1,000$ realizations from each model setting and for each realization, we calculate θ_k^a, $k = 1, \ldots, 1000$, $a = ML, MWPL$, estimates of $\theta = (\beta_0, \beta_1, \phi, \kappa)'$. We set Δ equal to the minimum distance among the points in (8) and we use the true value of the parameters as starting value for the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm implemented in the `optim` function of `R` software (R Core Team, 2019).

Table 1 reports the relative efficiency of the MWPL estimates with respect to the ML estimates for each parameter in terms of mean squared error. Additionally, as an overall

measure of relative efficiency for the multi-parameter case we have considered

$$\text{RE} = \left(\frac{\det[F^{MWPL}]}{\det[F^{ML}]} \right)^{1/p},$$

where $p = 4$ is the number of unknown parameters in θ and the matrix F^a is the sample mean squared error matrix $F^a = 1000^{-1} \sum_{k=1}^{1000} \left(\hat{\theta}_k^a - \theta \right) \left(\hat{\theta}_k^a - \theta \right)'$. In this experiment, using the WPL instead of the likelihood function, we loose about 13% of the overall efficiency in the worst case which is an encouraging result. It is interesting to note that the relative efficiency of each parameter is different, but only the relative efficiency of the shape parameter κ is affected when we consider different strengths of the spatial dependence, i.e. different values of ϕ.

[Table 1 about here.]

We modify slightly our example to illustrate the quality of the linear predictor (9) in terms of the mean squared prediction error (MSPE). Suppose that we have observed $Y(s_1) = y_1, \ldots, Y(s_n) = y_n$ and we want to predict $Y(s_{n+1})$ with $s_{n+1} > s_n$. In such case the conditional expectation of $Y(s_{n+1})$, i.e. the predictor the minimizes the MSPE, can be derived in closed form (see Appendix), namely

$$\begin{aligned} Y^*(s_{n+1}) :=& \Gamma\left(\frac{1}{\kappa} + 1\right) (1 - \rho_{n,n+1}^2)^{1/\kappa} \mu_{n+1} \nu(\kappa) \\ & \times \exp\left\{ -\frac{[y_n/(\mu_n \nu(\kappa))]^\kappa}{(1 - \rho_{n-1,n}^2)} \left[\frac{(1 - \rho_{n-1,n}^2 \rho_{n,n+1}^2)}{(1 - \rho_{n,n+1}^2)} - 1 \right] \right\} \\ & \times {}_1F_1\left(\frac{1}{\kappa} + 1; 1; \frac{[y_n/(\mu_n \nu(\kappa))]^\kappa}{(1 - \rho_{n,n+1}^2)} \rho_{n,n+1}^2 \right). \end{aligned}$$

Having collected $n = 21$ observations at locations $s_1 = 0, s_2 = 0.05, \ldots, s_n = 1$, we predict the random variable $Y(s_{n+1})$ at $s_{n+1} = 1.05$ by means of the optimal predictor $Y^*(s_{n+1})$, and the linear predictor $\widehat{Y}(s_{n+1})$ as in (9).

We simulate $1,000$ realizations from the stationary Weibull model, i.e. $\mu(s_i) = 1$, with the same dependence structure as before. Then we compute the average of the squared prediction errors $[Y(s_{21}) - Y^*(s_{21})]^2$ and $[Y(s_{21}) - \widehat{Y}(s_{21})]^2$ and their ratio. Table 2 shows the ratio between the linear and the optimal predictor. This ratio deteriorates when the strength of the dependence increases as expected, but the loss of the relative efficiency does not exceed thirty-two percent.

[Table 2 about here.]

6. WIND SPEED DATA EXAMPLE

Our motivating example is a dataset of daily average wind speeds from a network of meteorological stations in the Netherlands. The dataset is stored on the website of the KNMI Data Centre (`https://data.knmi.nl/about`) and its access is provided under the OpenData policy of the Dutch government.

Among the fifty stations in the dataset, we extracted thirty stations (Figure 3-a) that do not contain missing data in the period from 01/01/2000 to 31/12/2008 .

Figures 3-(b,c,d) show the time series plots of daily mean wind speeds at four different locations (Cabauw, Nieuw Beerta, Hoek Van Holland and Rotterdam) in 2000-2004. The seasonal trend is clearly recognizable and the heteroscedasticity seems related to this trend. Furthermore if we consider the wind speed box-plots for each station (Figure 4-a), it is clear that the distribution also depends on the location. To avoid a complicated spatial trend specification, we transform $Y(s,t)$, the observation of location s and time t, to $\widetilde{Y}(s,t) = Y(s,t)/a(s)$ where $a(s)$ is the average of the observations at site s. The transformation seems to have an effect of reducing the differences in distribution, see (Figure 4-b).

We specify a multiplicative model for the transformed data, namely

$$\widetilde{Y}(s,t) = \mu(t)E(s,t), \tag{11}$$

in which we conveys the seasonal pattern in the deterministic positive function $\mu(t)$ and $E = \{E(s,t)\}$ is a stationary positive process with unit mean. In particular we specify a harmonic model for the temporal trend, i.e.

$$\log \mu(t) \;\; = \;\; \beta_0 + \sum_{k=1}^{q} \left\{ \beta_{1,k} \cos\left(\frac{2\pi kt}{P}\right) + \beta_{2,k} \sin\left(\frac{2\pi kt}{P}\right) \right\} \tag{12}$$

where we set $P = 365.25$ days to handle leap years.

In the sequel we want to compare two specifications of E, namely the proposed Weibull model $E(s,t) = W(s,t)$ and a log-Gaussian model $E(s,t) = \exp(\sigma Z(s,t) - \sigma^2/2)$, $\sigma > 0$ where Z is a standard space-time Gaussian process.

We first get a preliminary estimate of the seasonal effect $\mu(t)$ assuming space-time independence and by using least squares and regressing $q = 4$ annual harmonics on the logarithm of the observations

$$\log \widetilde{Y}(s,t) = \beta_0 + \sum_{k=1}^{4} \left\{ \beta_{1,k} \cos\left(\frac{2\pi kt}{P}\right) + \beta_{2,k} \sin\left(\frac{2\pi kt}{P}\right) \right\} + \varepsilon(s,t) \tag{13}$$

with $\mathbb{E}(\varepsilon(s,t)) = 0$ and $\mathbb{V}ar(\varepsilon(s,t)) = \sigma_\varepsilon^2 < \infty$. Under the Weibull marginal distribution for $E(s,t)$ we identify β_0 with $\beta_0 + \log(\nu(\kappa)) - \gamma/\kappa$ since $-\log W(s,t)$ is a Gumbel random variable with mean $-\log \nu(\kappa) + \gamma/\kappa$ and $\gamma \approx 0.5772$ is the Euler–Mascheroni constant. Instead, under the log-Gaussian marginal distribution for $E(s,t)$ we identify β_0 with $\beta_0 - \sigma^2/2$.

We have used the values $e(s,t) = \exp(\widehat{\varepsilon}(s,t))$, where $\widehat{\varepsilon}(s,t)$ are the estimated residuals of the fitted regression model (13), for getting more insight about soundness of the model specification (11) for the wind speed data. Since parameter $\mu(t)$ controls at the same time the expectation and the variance of $\tilde{Y}(s,t)$, it is expected that the residuals $e(s,t)$ are homoschedastic. This is confirmed by grouping $e(s,t)$ by month and looking at the corresponding boxplots (Figure 5-(a)). Moreover comparing the overall qq-plots of $e(s,t)$ (Figure 5-(b)) there is convincing evidence that a Weibull distribution is more appropriate with respect to the log-Gaussian one.

In addition, if we transform the residuals of each location to the normal scores by means of the empirical transform, the scatter-plots of the normal scores of Rotterdam station versus the normal scores of three other stations (Figure 6) point out that there is more dependence in the upper corner, i.e. the lack of symmetry. This implies that the Weibull model seems more appropriate for modeling the pairwise dependence with respect to a log-Gaussian model since its copula is reflection symmetric.

Finally, the spatial and temporal marginal empirical semi-variograms of the residuals exhibit a strong and long decay dependence for the spatial margin and a weak dependence for the temporal margin. This suggests the use of the following space-time correlation (Porcu et al., 2019):

$$\rho(h,u) = \frac{1}{(1 + \|h\|/\phi_S)^{2.5}} \left(1 - \frac{|u|}{\phi_T (1 + \|h\|/\phi_S)^{-\phi_{ST}}}\right)_+^{3.5}, \tag{14}$$

with $\phi_S > 0$, $\phi_T > 0$, $0 \leq \phi_{ST} \leq 1$. When the space-time interaction parameter ϕ_{ST} is zero, then the space-time correlation is simply the product of a spatial Cauchy correlation function and a temporal Wendland correlation function (Bevilacqua et al., 2019), i.e. a separable model for the underlying spatio temporal Gaussian process. However, from (6) it is apparent that separability is not inherited for the Weibull and Log-Gaussian models. In this application, we have considered three different degree of space-time interaction by fixing $\phi_{ST} = 0, 0.5, 1$.

[Figure 3 about here.]

[Figure 4 about here.]

[Figure 5 about here.]

[Figure 6 about here.]

[Figure 7 about here.]

Using the preliminary estimates of the regression parameters as starting values in the BFGS (Fletcher, 1987) optimization algorithm we have fitted the Weibull and Log-Gaussian models with WPL using seven years (2000-2007). The last year has been used for the evaluation of the prediction performance using time-forward predictions. Note that the sample size (87630 observations) prevents the use of a full likelihood approach even for the Log-Gaussian model.

The estimation of regression and dependence parameters for the six models is based on WPL, with a cut-off weight c_{ij} equal to one if $|t_i - t_j| \leq 1$ and zero otherwise. Table 3 collects the results of the estimation stage including the standard error estimates obtained with a sub-sampling technique as in Bevilacqua et al. (2012). As one could expect, there is no big difference in trend estimates among different models and correlation functions. However considering the PLIC criterion, our preference goes to the Weibull model with $\phi_{ST} = 0$.

For the Weibull case with $\phi_{ST} = 0$, using the MWPL estimates of the regression parameters we first compute the estimated residuals and then we compute the empirical spatio-temporal semi-variogram of the residuals. We compare it with the estimated theoretical

semi-variogram obtained plugging the MWPL estimates into the theoretical spatio-temporal semi-variogram i.e. $\gamma_W(h,u) = \sigma_W^2(1-\rho_W(h,u))$ with ρ_W given by:

$$\rho_W(h,u) = \frac{\nu^{-2}(\kappa)}{\left[\Gamma\left(1+2/\kappa\right) - \nu^{-2}(\kappa)\right]}\left[{}_2F_1\left(-1/\kappa,-1/\kappa;1;\rho^2(h,u)\right)-1\right]. \tag{15}$$

Figure 7 shows the good agreement of the estimated theoretical spatial and temporal marginal semi-variograms (i.e. the estimation of $\gamma_W(h,0)$ and $\gamma_W(0,u)$ respectively) with the empirical counterparts.

[Table 3 about here.]

We want to further evaluate the predictive performances of the proposed model by considering one-day ahead predictions for the wind speed at the thirty meteorological stations but we have limited the number of predictor variables due to the computational load. Specifically, the predictor variables are the 150 wind speeds observed during the past five days at the stations.

For the Weibull models, we used the simple kriging predictor (9). For the Log-Gaussian models, we have chosen the conditional expectation given the past observations (De Oliveira, 2006, formula 2). In both cases, the predictions are obtained by plugging in the estimated parameters in the formulas. As benchmark, we have also considered the naïve predictor $\widehat{Y}(s_i,t) = y(s_i,t-1)$, that uses the observation recorded the day before at the station. The prediction performances are compared looking to the root-mean-square prediction error (RMSE) and the mean absolute prediction error (MAE).

In addition, we considered the sample mean of the continuous ranked probability score (CRPS) to evaluate the marginal predictive distribution performance (Gneiting and Raftery, 2007). For a single predictive cumulative distribution function F and a verifying observation y, the score is defined as

$$\mathrm{CRPS}(F,y) = \int_{-\infty}^{\infty}(F(t)-\mathbf{1}_{[y,\infty]}(t))^2dt.$$

For a Weibull distribution, we have derived an analytical expression of the corresponding score (see the Appendix) which turns out to be, under our parametrization:

$$\mathrm{CRPS}_W(F_{\kappa,\mu\nu(\kappa)},y) = y\left\{2\left(1-\exp\{-(y/\mu\nu(\kappa))^\kappa\}-1\right\}+2\mu\left[2^{-1/k}-\nu(\kappa)\gamma\left(1+\frac{1}{\kappa},\frac{y^\kappa}{(\mu\nu(\kappa))^\kappa}\right)\right],$$

where $\gamma(s,z) = \int_0^z t^{s-1}\,\mathrm{e}^{-t}\,\mathrm{d}t$ is the lower Gamma incomplete function.

Baran and Lerch (2015) derived the corresponding one for a Log-Gaussian random variable $Y = \exp(\alpha + \beta Z)$ with cdf $F_{\alpha,\beta}$, where Z is standard Gaussian random variable. Under our parametrization:

$$\mathrm{CRPS}_{LG}(F_{\mu-\sigma^2/2,\sigma}, y) = y\,[2\Phi(l(y)-1)] + 2e^{\mu}\left[1 - \Phi\left(\sigma/\sqrt{2}\right) - \Phi\left(l(y)-\sigma\right)\right],$$

where $\Phi(\cdot)$ is the CDF of the standard Gaussian distribution and $l(y) = [\log(y) - (\mu - \sigma^2/2)]/\sigma$.

As a general consideration the prediction based on a model (Weibull or Log-Gaussian) outclasses always the naïve prediction (see Table 4). Moreover, even though the simple kriging predictor is a suboptimal solution, the Weibull model outperforms the Log-Gaussian model in terms of RMSE, MAE. Finally also the CRPS of the Weibull model outperform the Log-Gaussian model. Note that CRPS values do not dependent on ϕ_{ST}. This is not surprising since the estimated marginal parameters for both models are very similar for $\phi_{ST} = 0, 0.5, 1$.

Among the fitted covariance models, we give again a preference to the correlation function with $\phi_{ST} = 0$.

[Table 4 about here.]

7. CONCLUDING REMARKS

Motivated by a spatio-temporal analysis of daily wind speed data from a network of meteorological stations in the Netherlands, we proposed a non-stationary stochastic process with Weibull marginal distributions for regression and dependence analysis when dealing with positive continuous data. In contrast to a Gaussian copula or, more generally, to monotonic transformations of a Gaussian process, our model offers a workable solution in the presence of different dependence in the lower and upper distribution tails, *i.e.* reflection asymmetry.

Additionally, we have shown that nice properties such as stationarity, mean-square continuity and degrees of mean-square differentiability are inherited from the 'parent' Gaussian

random process. However, discontinuity of the paths can be easily induced by choosing a discontinuous correlation function for the 'parent' Gaussian process.

We also remark that even though we have limited ourselves to continuous Euclidean space, our models can be extended to a spherical domain (Gneiting, 2013; Porcu et al., 2016) or to a network space. In this respect the X_m random process should represent a generalization of the model in Warren (1992).

A common drawback for the proposed model is the lack of an amenable expression of the density outside of the bivariate case that prevents an inference approach based on likelihood methods and the derivation of an optimal predictor that minimizes the mean square prediction error. We have shown with some numerical experiments that an inferential approach based on the pairwise likelihood is an effective solution for estimating the unknown parameter. On the other hand probabilities of multivariate events could be evaluated by Monte Carlo method since the random processes can be quickly simulated. However, our solution to the conditional prediction, based on a linear predictor, is limited and deserves further consideration even if, in our simulations and real data example, it has been performed well.

ACKNOWLEDGEMENTS

Partial support was provided by FONDECYT grant 1200068, Chile and by Millennium Science Initiative of the Ministry of Economy, Development, and Tourism, grant "Millenium Nucleus Center for the Discovery of Structures in Complex Data" for Moreno Bevilacqua and by Proyecto de Iniciación Interno DIUBB 173408 2/I de la Universidad del Bío-Bío for Christian Caamaño.

REFERENCES

Adler, R. J. (1981). *The Geometry of Random Fields*. Wiley, New York.

Allcroft, D. and Glasbey, C. (2003). A latent Gaussian Markov random field model for spatio-temporal rainfall disaggregation. *Journal of the Royal Statistical Society: Series C* **52,** 487–498.

Baran, S. and Lerch, S. (2015). Log-normal distribution based ensemble model output statistics models for probabilistic wind-speed forecasting. *Quarterly Journal of the Royal Meteorological Society* **141,** 2289–2299.

Bárdossy, A. (2006). Copula-based geostatistical models for groundwater quality parameters. *Water Resources Research* **42,**.

Bellier, E., Monestiez, P., and Guinet, C. (2010). Geostatistical modelling of wildlife populations: a non-stationary hierarchical model for count data. In Atkinson, P. and Lloyd, C., editors, *geoENV VII Geostatistics for Environmental Applications. Springer*, pages 1–12.

Bessac, J., Ailliot, P., and Monbet, V. (2015). Gaussian linear state-space model for wind fields in the North-East atlantic. *Environmetrics* **26,** 29–38.

Bevilacqua, M., Faouzi, T., Furrer, R., and Porcu, E. (2019). Estimation and prediction using generalized Wendland functions under fixed domain asymptotics. *The Annals of Statistics* **47,** 828–856.

Bevilacqua, M. and Gaetan, C. (2015). Comparing composite likelihood methods based on pairs for spatial Gaussian random fields. *Statistics and Computing* **25,** 877–892.

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,** 268–280.

Bevilacqua, M. and Morales-Oñate, V. (2019). Geomodels: a package for geostatistical Gaussian and non Gaussian data analysis. https://vmoprojs.github.io/GeoModels-page/. R package version 1.0.3-4.

Carta, J., Ramírez, P., and Velázquez, S. (2009). A review of wind speed probability distributions used in wind energy analysis: Case studies in the Canary Islands. *Renewable and Sustainable Energy Reviews* **13,** 933 – 955.

Coles, S., Heffernan, J., and Tawn, J. (1999). Dependence measures for extreme value analyses. *Extremes* **2,** 339–365.

Cressie, N. (1993). *Statistics for Spatial Data.* Wiley, New York, revised edition.

Cressie, N. and Wikle, C. (2011). *Statistics for Spatio-Temporal Data.* Wiley, New York.

Davis, R. and Yau, C.-Y. (2011). Comments on pairwise likelihood in time series models. *Statistica Sinica* **21,** 255–277.

De Oliveira, V. (2006). On optimal point and block prediction in log-Gaussian random fields. *Scandinavian Journal of Statistics* **33,** 523–540.

De Oliveira, V. (2014). Poisson kriging: A closer investigation. *Spatial Statistics* **7,** 1–20.

De Oliveira, V., Kedem, B., and Short, D. A. (1997). Bayesian prediction of transformed Gaussian random fields. *Journal of the American Statistical Association* **92,** 1422–1433.

Fletcher, R. (1987). *Practical Methods of Optimization.* Wiley, New York, second edition.

Galanis, G., Hayes, D., Zodiatis, G., Chu, P., Kuo, Y.-H., and Kallos, G. (2012). Wave height characteristics in the Mediterranean Sea by means of numerical modeling, satellite data, statistical and geometrical techniques. *Marine Geophysical Research* **33,** 1–15.

Gneiting, T. (2013). Strictly and non-strictly positive definite functions on spheres. *Bernoulli* **19,** 1327–1349.

Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association* **102,** 359–378.

Gradshteyn, I. and Ryzhik, I. (2007). *Table of Integrals, Series, and Products.* Academic Press, Cambridge, MA, eight edition.

Gräler, B. (2014). Modelling skewed spatial random fields through the spatial vine copula. *Spatial Statistics* **10,** 87 –102.

Hashorva, E., Nadarajah, S., and Pogány, T. K. (2014). Extremes of perturbed bivariate rayleigh risks. *REVSTAT-Statistical Journal* **12,** 157–168.

Haslett, J. and Raftery, A. (1989). Space-time modelling with long-memory dependence: assessing ireland's wind-power resource (with discussion). *Applied Statistics* **38,** 1–50.

Heagerty, P. and Lele, S. (1998). A composite likelihood approach to binary spatial data. *Journal of the American Statistical Association* **93,** 1099 –1111.

Heaton, M. J., Datta, A., Finley, A. O., Furrer, R., Guinness, J., Guhaniyogi, R., Gerber, F., Gramacy, R. B., Hammerling, D., Katzfuss, M., Lindgren, F., Nychka, D. W., Sun, F., and Zammit-Mangion, A. (2018). A case study competition among methods for analyzing large spatial data. *Journal of Agricultural, Biological and Environmental Statistics* .

Joe, H. (2014). *Dependence Modeling with Copulas*. Chapman and Hall/CRC, Boca Raton, FL.

Johnson, N., Kotz, S., and Balakrishnan, N. (1995). *Continuous Univariate Distributions*. Wiley, New York.

Kazianka, H. and Pilz, J. (2010). Copula-based geostatistical modeling of continuous and discrete data including covariates. *Stochastic Environmental Research and Risk Assessment* **24,** 661–673.

Kibble, W. F. (1941). A two-variate gamma type distribution. *Sankhyā: The Indian Journal of Statistics (1933-1960)* **5,** 137–150.

Krishnamoorthy, A. S. and Parthasarathy, M. (1951). A multivariate Gamma-type distribution. *The Annals of Mathematical Statistics* **22,** 549–557.

Lindsay, B. (1988). Composite likelihood methods. *Contemporary Mathematics* **80,** 221–239.

Ma, C. (2010). χ^2 random fields in space and time. *IEEE Transactions on Signal Processing* **43,** 378–383.

Masarotto, G. and Varin, C. (2012). Gaussian copula marginal regression. *Electronic Journal of Statistics* **6,** 1517–1549.

Neykov, N. M., Neytchev, P. N., and Zucchini, W. (2014). Stochastic daily precipitation model with a heavy-tailed component. *Natural Hazards and Earth System Sciences* **14,** 2321–2335.

Porcu, E., Bevilacqua, M., and Genton, M. (2016). Spatio-temporal covariance and cross covariance functions of the great circle distance on a sphere. *Journal of the American Statistical Association* **111,** 888–898.

Porcu, E., Bevilacqua, M., and Genton, M. G. (2019). Nonseparable space-time covariance functions with dynamical compact supports. *Statistica Sinica* to appear.

Pryor, S. C. and Barthelmie, R. J. (2010). Climate change impacts on wind energy: a review. *Renewable and Sustainability Energy Reviews* **14,** 430–437.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Royen, T. (2004). Multivariate Gamma distributions II. In *Encyclopedia of Statistical Sciences*, pages 419–425. New York: John Wiley & Sons.

Sibuya, M. (1960). Bivariate extreme statistics. *Annals of the Institute of Statistical Mathematics* **11,** 195–210.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris* **8,** 229–231.

Stein, M. (1999). *Interpolation of Spatial Data. Some Theory of Kriging*. Springer-Verlag, New York.

Tang, Y., Wang, H. J., Sun, Y., and Hering, A. S. (2019). Copula-based semiparametric models for spatiotemporal data. *Biometrics* **75,** 1156–1167.

Varin, C., Reid, N., and Firth, D. (2011). An overview of composite likelihood methods. *Statistica Sinica* **21,** 5–42.

Varin, C. and Vidoni, P. (2005). A note on composite likelihood inference and model selection. *Biometrika* **52,** 519–528.

Vere-Jones, D. (1967). The infinite divisibility of a bivariate gamma distribution. *Sankhyā: The Indian Journal of Statistics* **29,** 421–422.

Wadsworth, J. and Tawn, J. (2012). Dependence modelling for spatial extremes. *Biometrika* **99,** 253–272.

Wallin, J. and Bolin, D. (2015). Geostatistical modelling using non-gaussian matérn fields. *Scandinavian Journal of Statistics* **42,** 872 –890.

Warren, D. (1992). A multivariate gamma distribution arising from a Markov model. *Stochastic Hydrology and Hydraulics* **6,** 183–190.

Xu, G. and Genton, M. G. (2017). Tukey g-and-h random fields. *Journal of the American Statistical Association* **112,** 1236–1249.

Yan, Y. and Genton, M. G. (2019). Non-gaussian autoregressive processes with tukey -and-h transformations. *Environmetrics* **30,** e2503.

APPENDIX

In the sequel we will exploit the identity for the hypergeometric function ${}_0F_1$,

$$ {}_0F_1(;b;x) = \Gamma(b) x^{(1-b)/2} I_{b-1}(2\sqrt{x}). $$

where $I_a(x)$ is the modified Bessel function of the first kind of order a.

Proposition 1. *The $(a,b)-th$ product moment of any pairs $W_1 := W(s_1)$ and $W_2 := W(s_2)$ is given by*

$$ \mathbb{E}(W_1^a\, W_2^b) = \frac{\Gamma\left(1+a/\kappa\right)\Gamma\left(1+b/\kappa\right)}{\Gamma\left(1+1/\kappa\right)^{a+b}} {}_2F_1\left(-a/\kappa, -b/\kappa; 1; \rho^2\right) \tag{A.1} $$

where $\rho = \rho(s_1 - s_2)$

Proof. Using the series expansion of hypergeometric function ${}_0F_1$, we have:

$$ \begin{aligned} \mathbb{E}(W_1^a\, W_2^b) &= \frac{\kappa^2\Gamma\left(1+1/\kappa\right)^{2\kappa}}{1-\rho^2} \int_0^\infty\int_0^\infty u^{\kappa+a-1} v^{\kappa+b-1} \exp\left\{-\frac{\Gamma\left(1+1/\kappa\right)^\kappa}{(1-\rho^2)}\left(u^\kappa+v^\kappa\right)\right\} \\ &\quad \times {}_0F_1\left(1; \frac{\rho^2(uv)^\kappa\Gamma\left(1+1/\kappa\right)^{2\kappa}}{(1-\rho^2)^2}\right) dudv \\ &= \frac{\kappa^2\Gamma\left(1+1/\kappa\right)^{2\kappa}}{1-\rho^2} \sum_{m=0}^\infty \frac{1}{(m!)^2}\left(\frac{\rho^2\Gamma\left(1+1/\kappa\right)^{2\kappa}}{(1-\rho^2)^2}\right)^m \\ &\quad \times \int_0^\infty\int_0^\infty \exp\left\{-\frac{\Gamma\left(1+1/\kappa\right)^\kappa}{(1-\rho^2)}\left(u^\kappa+v^\kappa\right)\right\} dudv \\ &= \frac{\kappa^2\Gamma\left(1+1/\kappa\right)^{2\kappa}}{1-\rho^2} \sum_{m=0}^\infty \frac{I(m)}{(m!)^2}\left(\frac{\rho^2\Gamma\left(1+1/\kappa\right)^{2\kappa}}{(1-\rho^2)^2}\right)^m \end{aligned} \tag{A.2} $$

Using Fubini's Theorem and (3.381.4) in Gradshteyn and Ryzhik (2007), we obtain

$$ \begin{aligned} I(m) &= \int_0^\infty u^{\kappa+a+m\kappa-1} \exp\left\{-\frac{\Gamma\left(1+1/\kappa\right)^\kappa}{(1-\rho^2)}u^\kappa\right\} du \\ &\quad \times \int_0^\infty v^{\kappa+b+m\kappa-1} \exp\left\{-\frac{\Gamma\left(1+1/\kappa\right)^\kappa}{(1-\rho^2)}v^\kappa\right\} dv \\ &= \kappa^{-2}\Gamma\left(1+a/\kappa+m\right)\Gamma\left(1+b/\kappa+m\right) \\ &\quad \times \left(\frac{1-\rho^2}{\Gamma\left(1+1/\kappa\right)^\kappa}\right)^{1+a/\kappa+m} \left(\frac{1-\rho^2}{\Gamma\left(1+1/\kappa\right)^\kappa}\right)^{1+b/\kappa+m} \end{aligned} \tag{A.3} $$

Combining equations (A.2) and (A.3), we obtain

$$\begin{aligned}\mathbb{E}(W_1^a\, W_2^b) &= \frac{(1-\rho^2)^{1+(a+b)/\kappa}\Gamma\left(1+a/\kappa\right)\Gamma\left(1+b/\kappa\right)}{\Gamma\left(1+1/\kappa\right)^{a+b}} \\ &\quad \times {}_2F_1\left(1+a/\kappa, 1+b/\kappa; 1; \rho^2\right)\end{aligned}$$

Finally, using Euler transformation, we obtain (A.1). □

Proposition 2. *Let $s_1 < s_2 < \cdots < s_n < s_{n+1}$, with $s_i \in \mathbb{R}$. For the Weibull process Y with underlying exponential correlation function, the conditional expectation of $Y^a(s_{n+1})$, $a > 0$, given $Y(s_1) = y_1, \ldots, Y(s_n) = y_n)$ is*

$$\begin{aligned}\mathbb{E}(Y^a(s_{n+1})|Y(s_1) = y_1, \ldots, Y(s_n) = y_n) &= \Gamma\left(\frac{a}{\kappa}+1\right)(1-\rho_{n,n+1}^2)^{a/\kappa}[\nu(\kappa)\mu_{n+1}]^a \\ &\quad \times \exp\left\{-\frac{y_n^\kappa}{(1-\rho_{n-1,n}^2)[\nu(\kappa)\mu_n]^\kappa}\left[\frac{(1-\rho_{n-1,n}^2\rho_{n,n+1}^2)}{(1-\rho_{n,n+1}^2)}-1\right]\right\} \\ &\quad \times {}_1F_1\left(\frac{a}{\kappa}+1; 1; \frac{\rho_{n,n+1}^2 y_n^\kappa}{[\nu(\kappa)\mu_n]^\kappa(1-\rho_{n,n+1}^2)}\right)\end{aligned}$$

Proof. First, note that using (10), the density of the random variable $Y(s_{n+1})|(Y(s_1) = y_1, \ldots, Y(s_n) = y_n)$ is easily obtained as:

$$\begin{aligned}f(y_{n+1}|y_1, \ldots, y_n) &= \frac{\kappa y_{n+1}^{\kappa-1}}{\nu^\kappa(\kappa)\mu_{n+1}^\kappa(1-\rho_{n,n+1}^2)}\exp\left\{-\frac{1}{(1-\rho_{n,n+1}^2)}\left[\frac{y_{n+1}}{\nu(\kappa)\mu_{n+1}}\right]^\kappa\right\} \\ &\quad \times \exp\left\{-\frac{y_n^\kappa}{(1-\rho_{n-1,n}^2)[\nu(\kappa)\mu_n]^\kappa}\left[\frac{(1-\rho_{n-1,n}^2\rho_{n,n+1}^2)}{(1-\rho_{n,n+1}^2)}-1\right]\right\} \\ &\quad \times I_0\left(\frac{2|\rho_{n,n+1}|(y_n y_{n+1})^{\kappa/2}}{\nu^\kappa(\kappa)(\mu_n\mu_{n+1})^{\kappa/2}(1-\rho_{n,n+1}^2)}\right).\end{aligned}$$

Using the series expansion of hypergeometric function ${}_0F_1$, we obtain:

$$\begin{aligned}
\mathbb{E}(Y^a(s_{n+1})|Y(s_1)=y_1,\ldots,Y(s_n)=y_n) &= \frac{\kappa}{\nu^\kappa(\kappa)\mu_{n+1}^\kappa(1-\rho_{n,n+1}^2)} \\
&\times \exp\left\{-\nu^{-\kappa}(\kappa)\left[\frac{(1-\rho_{n-1,n}^2\rho_{n,n+1}^2)y_n^\kappa}{\mu_n^\kappa(1-\rho_{n-1,n}^2)(1-\rho_{n,n+1}^2)}-\frac{y_n^\kappa}{\mu_n^\kappa(1-\rho_{n-1,n}^2)}\right]\right\} \\
&\times \int_0^\infty y_{n+1}^{\kappa+a-1} e^{-\frac{1}{(1-\rho_{n,n+1}^2)}\left[\frac{y_{n+1}}{\nu(\kappa)\mu_{n+1}}\right]^\kappa} {}_0F_1\left(;1;\frac{\rho_{n,n+1}^2(y_n y_{n+1})^\kappa}{\nu^{2\kappa}(\kappa)(\mu_n\mu_{n+1})^\kappa(1-\rho_{n,n+1}^2)^2}\right) dy_{n+1} \\
&= \frac{\kappa}{\nu^\kappa(\kappa)\mu_{n+1}^\kappa(1-\rho_{n,n+1}^2)} \\
&\times \exp\left\{-\nu^{-\kappa}(\kappa)\left[\frac{(1-\rho_{n-1,n}^2\rho_{n,n+1}^2)y_n^\kappa}{\mu_n^\kappa(1-\rho_{n-1,n}^2)(1-\rho_{n,n+1}^2)}-\frac{y_n^\kappa}{\mu_n^\kappa(1-\rho_{n-1,n}^2)}\right]\right\} \\
&\times \sum_{m=0}^\infty \frac{I(m)}{(m!)^2}\left(\frac{\rho_{n,n+1}^2 y_n^\kappa}{\nu^{2\kappa}(\kappa)(\mu_n\mu_{n+1})^\kappa(1-\rho_{n,n+1}^2)^2}\right)^m
\end{aligned} \tag{A.4}$$

where

$$\begin{aligned}
I(m) &= \int_0^\infty y_{n+1}^{\kappa+a+\kappa m-1}\exp\left\{-\frac{1}{(1-\rho_{n,n+1}^2)}\left[\frac{y_{n+1}}{\nu(\kappa)\mu_{n+1}}\right]^\kappa\right\} dy_{n+1} \\
&= \kappa^{-1}[(1-\rho_{n,n+1}^2)\nu^\kappa(\kappa)\mu_{n+1}^\kappa]^{a/\kappa+m+1}\Gamma\left(\frac{a}{\kappa}+m+1\right)
\end{aligned} \tag{A.5}$$

Combining equations (A.4) and (A.5), we obtain the conditional expectation in proposition 2. □

Proposition 3. *The* CRPS *associated with the* $Weibull(\alpha,\beta)$ *distribution is given by*

$$\mathrm{CRPS}(F_{\alpha,\beta},y) = y[2F_{\alpha,\beta}(y)-1] - 2\beta\gamma\left(1+\frac{1}{\alpha},\frac{y^\alpha}{\beta^\alpha}\right) + 2^{-1/\alpha}\beta\Gamma\left(1+\frac{1}{\alpha}\right) \tag{A.6}$$

where $F_{\alpha,\beta}(y) = 1-\exp^{-(y/\beta)^\alpha}$ *and* $\gamma(s,z) = \int_0^z t^{s-1}\,\mathrm{e}^{-t}\,\mathrm{d}t$, $s>0$ *is the lower incomplete gamma function.*

Proof. We first note that the CRPS can also be written as

$$\mathrm{CRPS}(F,y) = \mathbb{E}_F|Y-y| - \frac{1}{2}\mathbb{E}_F|Y-Y'|$$

where Y and Y' are independent random variables with cumulative distribution function F and finite first moment. The first term can be integrated out using the properties of the

Weibull density, yielding

$$\begin{aligned}
\mathbb{E}_F|Y-y| &= \int_{-\infty}^{y}(y-t)f_{\alpha,\beta}(t)dt - \int_{y}^{\infty}(y-t)f_{\alpha,\beta}(t)dt \\
&= yF_{\alpha,\beta}(y) - \int_{-\infty}^{y}\frac{\alpha}{\beta^\alpha}t^\alpha \exp\left[-\left(\frac{t}{\beta}\right)^\alpha\right]dt - y[1-F_{\alpha,\beta}(y)] \\
&+ \int_{y}^{\infty}\frac{\alpha}{\beta^\alpha}t^\alpha \exp\left[-\left(\frac{t}{\beta}\right)^\alpha\right]dt \\
&= y[2F_{\alpha,\beta}(y)-1] - \beta\gamma\left(1+\frac{1}{\alpha},\frac{y^\alpha}{\beta^\alpha}\right) + \beta\Gamma\left(1+\frac{1}{\alpha},\frac{y^\alpha}{\beta^\alpha}\right)
\end{aligned}$$

where $\gamma(s,z) = \int_0^z t^{s-1}\,\mathrm{e}^{-t}\,\mathrm{d}t$, $s>0$ is the lower incomplete gamma function and $\Gamma(s,z) = \Gamma(s) - \gamma(s,z)$ is the upper incomplete gamma function. We have:

$$\mathbb{E}_F|X-y| = y[2F_{\alpha,\beta}(y)-1] - \beta\left[2\gamma\left(1+\frac{1}{\alpha},\frac{y^\alpha}{\beta^\alpha}\right) - \Gamma\left(1+\frac{1}{\alpha}\right)\right]$$

The second term can be calculated using its relation to the Gini concentration ratio G:

$$\mathbb{E}_F|Y-Y'| = \int_{\mathbb{R}_+^2}|y-y'|f_{\alpha,\beta}(y)f_{\alpha,\beta}(y')dy\,dy' = 2\mathbb{E}(Y)G = 2\beta\Gamma\left(1+\frac{1}{\alpha}\right)(1-2^{-1/\alpha})$$

Putting both terms together, we obtain

$$\mathrm{CRPS}(F_{\alpha,\beta},y) = y[2F_{\alpha,\beta}(y)-1] - 2\beta\gamma\left(1+\frac{1}{\alpha},\frac{y^\alpha}{\beta^\alpha}\right) + 2^{-1/k}\beta\Gamma\left(1+\frac{1}{\alpha}\right).$$

□

List of Figures

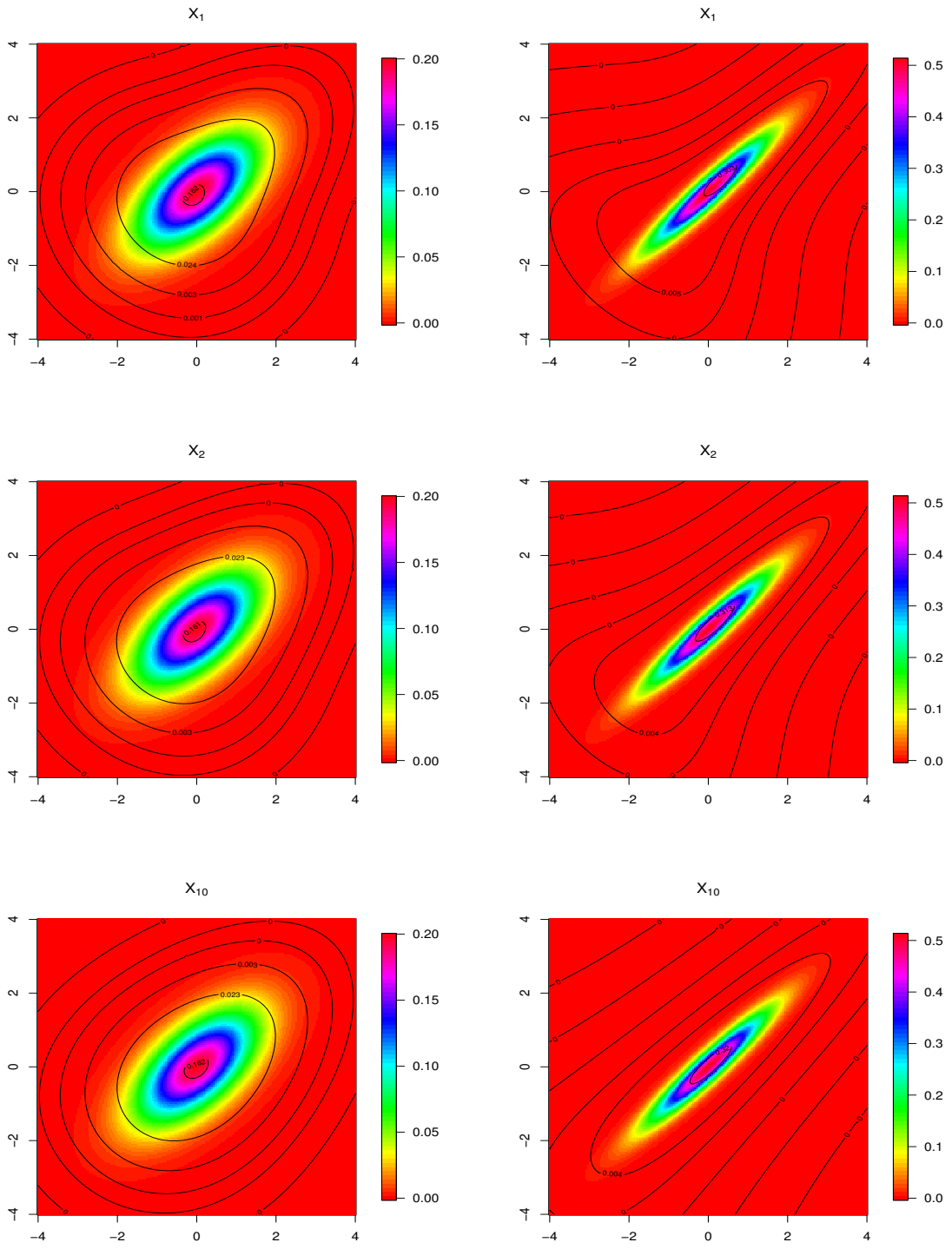


Figure 1: X_m process: bivariate density contour plots for different values of m and ρ after transforms to $\mathcal{N}(0,1)$ margins with $m = 1, 2, 10$ from top to bottom and $\rho = 0.6, 0.95$ from left to right. The background image is a grid of colored pixels with colors corresponding to the values of the standard bivariate Gaussian density with correlation ρ.

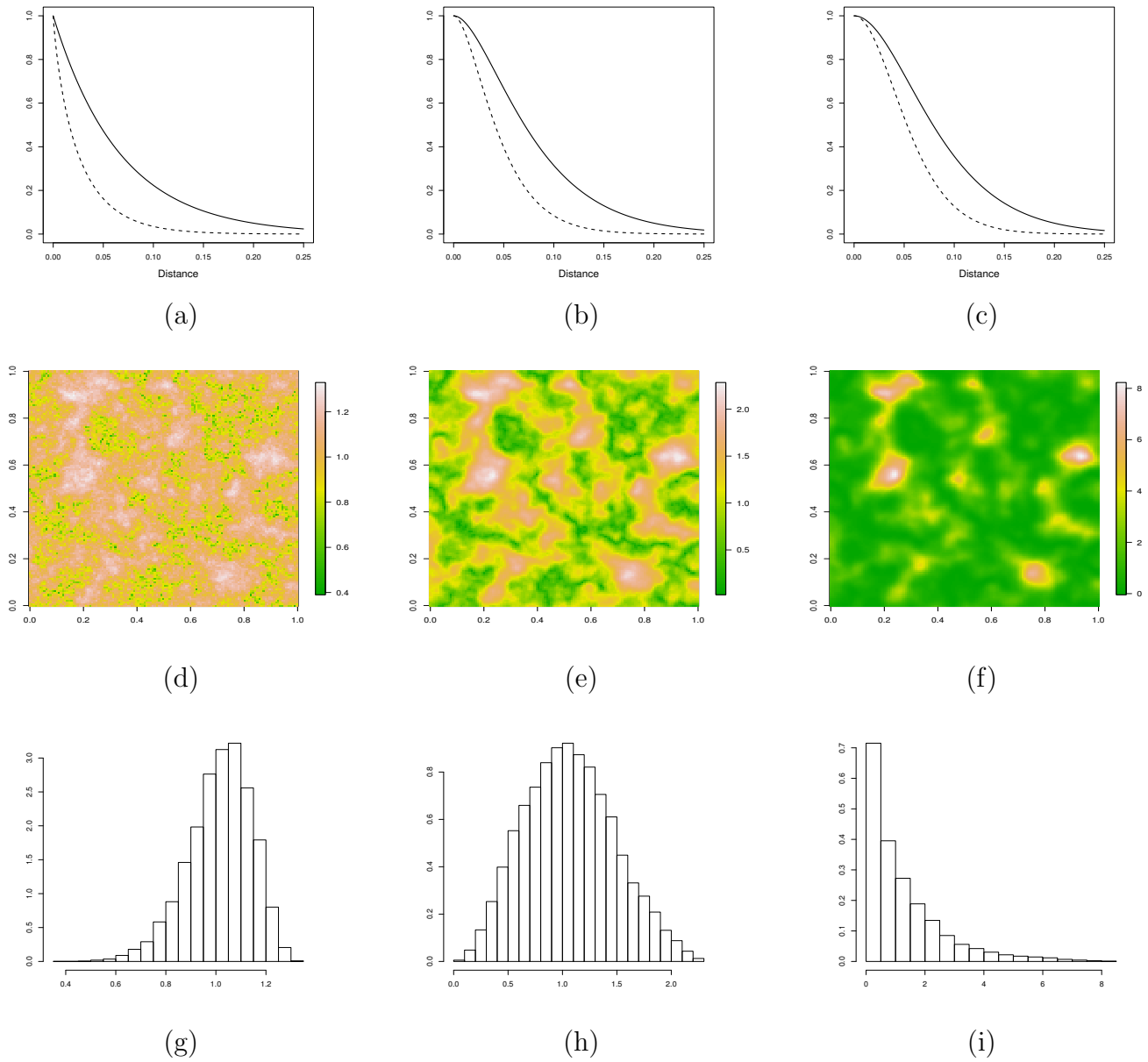


Figure 2: Top: comparison between $\rho(h)$, the correlation function of the 'parent' Gaussian process (a Matérn correlation function with smoothness parameter $\nu = 0.5, 1.5, 2.5$ and practical range approximately equal to 0.2), with the associated correlation of the Weibull model $\rho_W(h)$ (dashed line) with shape parameter κ for $(\kappa, \nu) = (10, 0.5), (3, 1.5), (1, 2.5)$, from left to right. Center: three realizations of the Weibull model W under the setting (a),(b), and (c). Bottom: histograms of the realizations in (d),(e), and (f).

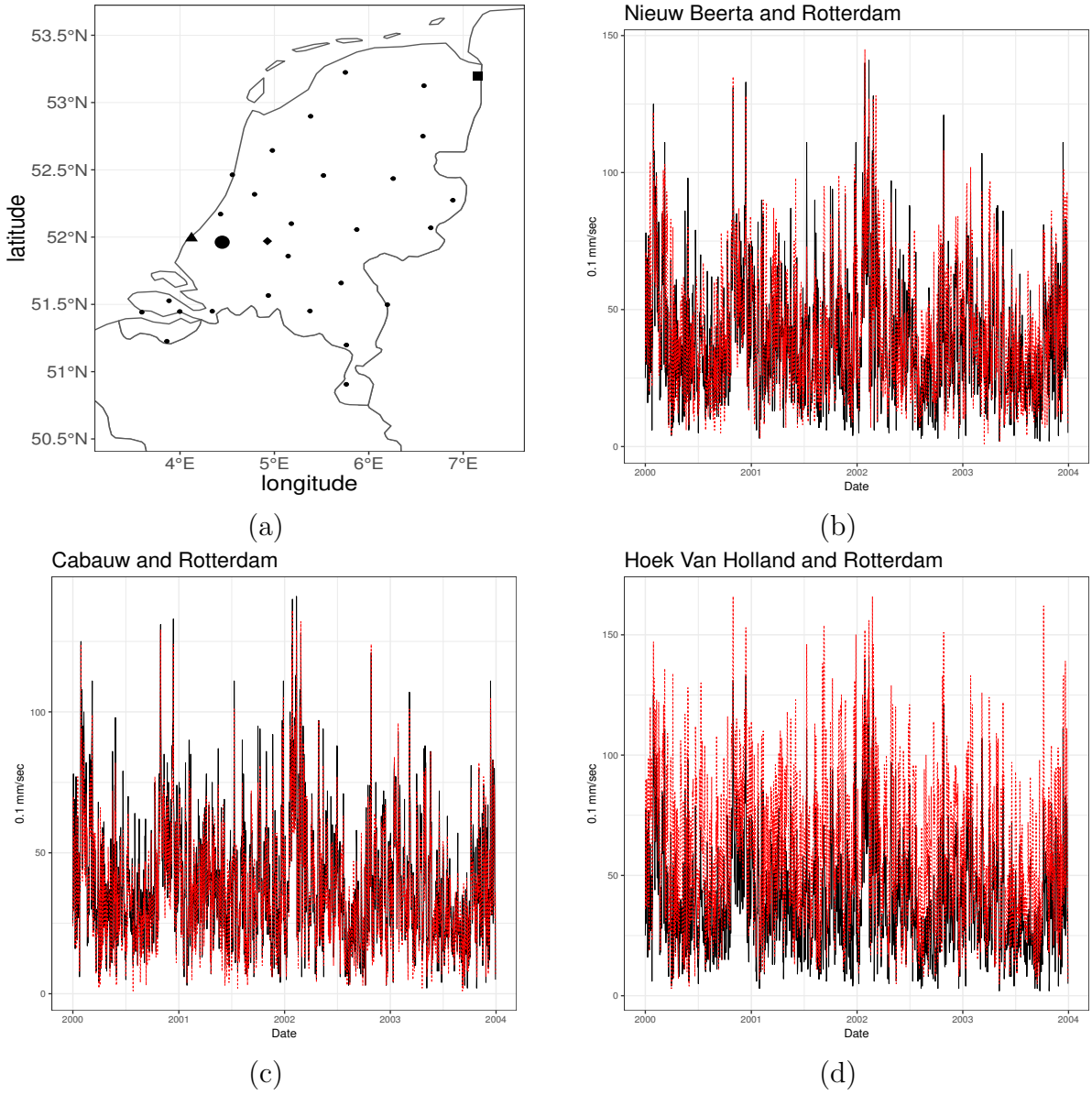
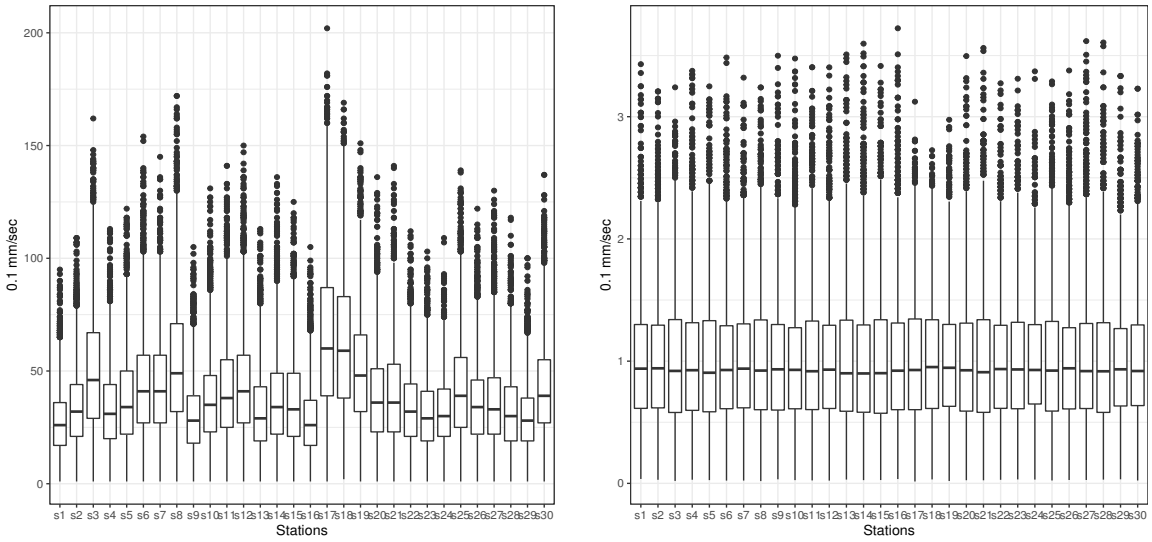
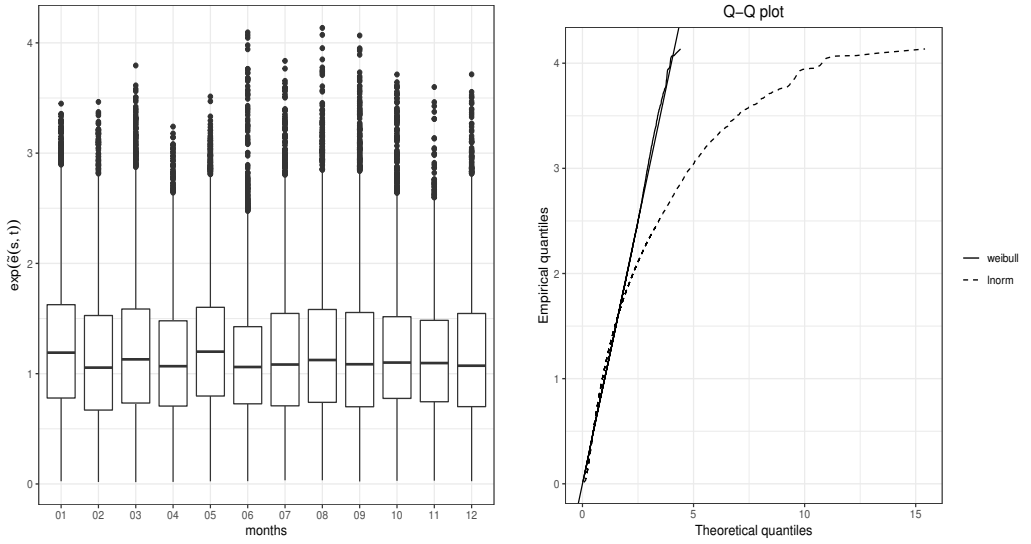


Figure 3: Wind speed data of Netherlands. (a) Map of the meteorological stations selected for our case study. Symbols ▲, ♦, ■, ● correspond to Cabauw, Hoek Van Holland, Nieuw Beerta and Rotterdam stations; (b-c-d) Time series plots (black lines¡) of the daily wind speed data (01/01/2000-31/12/2004) at Cabauw, Hoek Van Holland and Nieuw Beerta stations versus Rotterdam stations (red line).



(a) (b)

Figure 4: (a) boxplots of the daily wind speed data for each meteorological stations over the period 2000-2008; (b) boxplots of the daily wind speed data rescaled by the average over the considered period.



(a) (b)

Figure 5: Preliminary analysis of residuals, $r(s,t)$ obtained in fitting model (13) by least-squares: (a) boxplots of the exponential of the residuals for each month. (b) qq-plot of the exponential of the residuals against the Weibull and Log-Gaussian distribution.

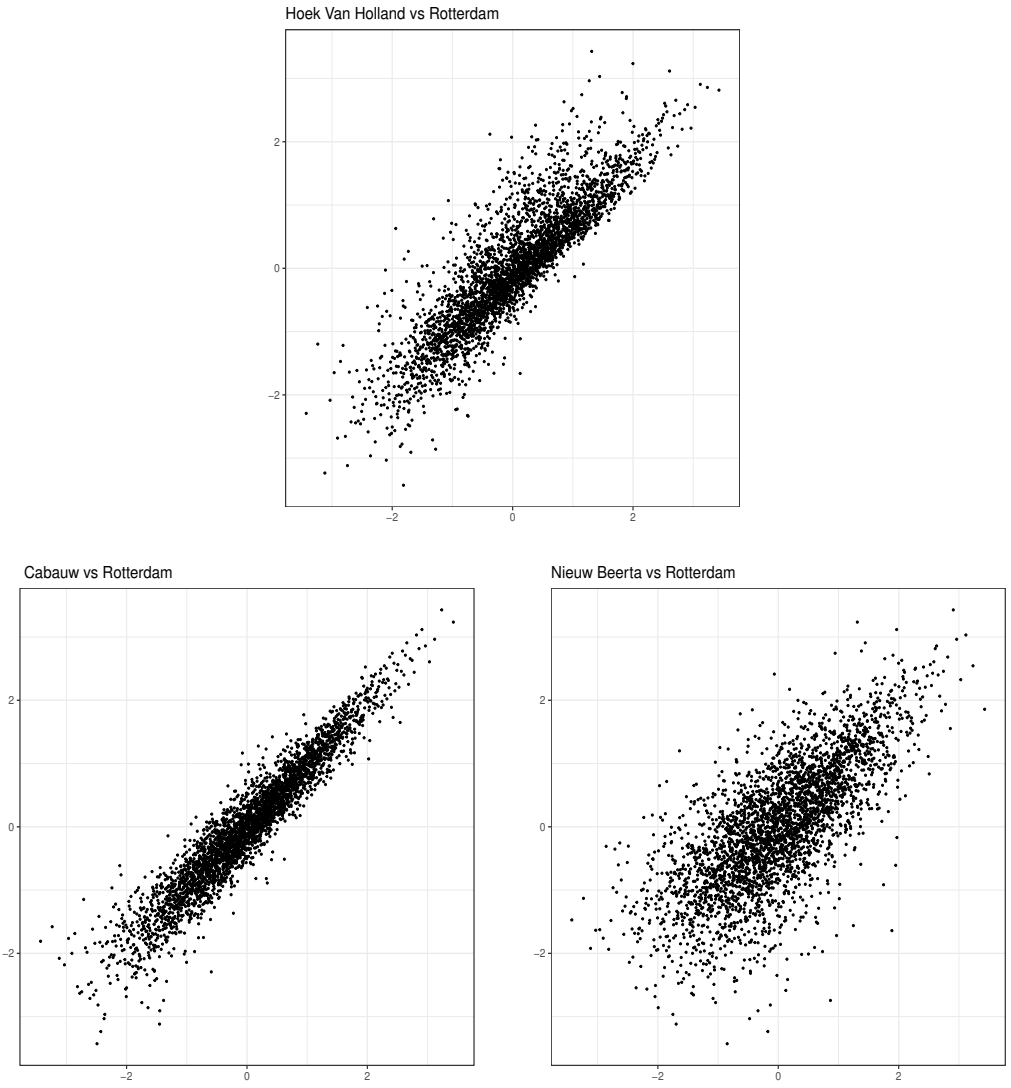


Figure 6: Preliminary analysis of residuals, $r(s,t)$ obtained in fitting model (13) by least-squares: the scatterplots of the normal scores of Rotterdam station vs the normal scores of three other stations.

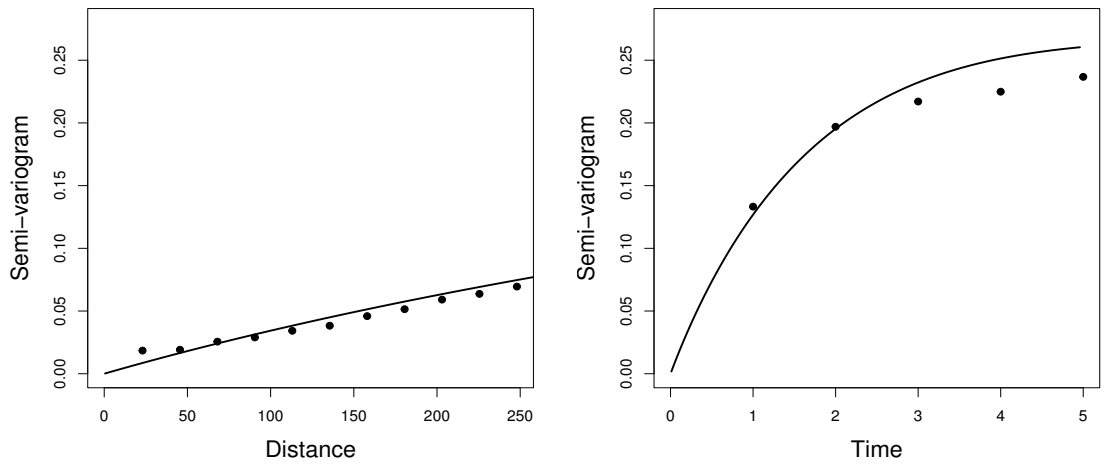


Figure 7: From left to right: Empirical spatial and temporal marginal semi-variograms of the residuals (dotted points) and estimated theoretical counterparts (solid line).

List of Tables

κ		$\phi = 0.1/3$	$\phi = 0.2/3$	$\phi = 0.3/3$
1	β_0	0.964	0.967	0.956
	β_1	0.862	0.860	0.874
	ϕ	1.045	1.037	1.034
	κ	0.885	0.703	0.550
	RE	0.954	0.913	0.884
3	β_0	0.953	0.947	0.930
	β_1	0.862	0.860	0.874
	ϕ	1.045	1.036	1.034
	κ	0.885	0.703	0.550
	RE	0.955	0.914	0.886
10	β_0	0.947	0.933	0.911
	β_1	0.862	0.860	0.874
	ϕ	1.044	1.036	1.034
	κ	0.885	0.703	0.550
	RE	0.955	0.914	0.886

Table 1: Mean squared error relative efficiency for each parameter and overall relative efficiency (RE) of MWPL vs ML.

	$\phi = 0.1/3$	$\phi = 0.2/3$	$\phi = 0.3/3$
$\kappa = 1$	0.953	0.805	0.687
$\kappa = 3$	0.960	0.825	0.721
$\kappa = 10$	0.967	0.851	0.764

Table 2: Relative efficiency of the linear predictor versus the optimal predictor for a stationary Weibull model defined on $S = [0, 1]$ with underlying exponential correlation $\rho(h) = \exp\{-|h|/\phi\}$.

	$\phi_{ST}=0$		$\phi_{ST}=0.5$		$\phi_{ST}=1$	
	Weibull	Log-Gaussian	Weibull	Log-Gaussian	Weibull	Log-Gaussian
β_0	−0.0222	0.0166	−0.0222	0.0168	−0.0221	0.0170
	(0.0026)	(0.0015)	(0.0026)	(0.0015)	(0.0026)	(0.0016)
$\beta_{1,1}$	0.0747	0.0787	0.0747	0.0787	0.0747	0.0787
	(0.0025)	(0.0024)	(0.0025)	(0.0026)	(0.0025)	(0.0029)
$\beta_{2,1}$	0.1822	0.1995	0.1822	0.1996	0.1822	0.1996
	(0.0030)	(0.0028)	(0.0029)	(0.0028)	(0.0029)	(0.0030)
$\beta_{1,2}$	−0.0087	−0.0270	−0.0087	−0.0270	−0.0087	−0.0270
	(0.0567)	(0.0192)	(0.0566)	(0.0190)	(0.0566)	(0.0200)
$\beta_{2,2}$	0.0138	0.0107	0.0138	0.0107	0.0137	0.0107
	(0.0306)	(0.0489)	(0.0306)	(0.0484)	(0.0306)	(0.0509)
$\beta_{1,3}$	0.0274	0.0237	0.0274	0.0237	0.0274	0.0237
	(0.0192)	(0.0229)	(0.0192)	(0.0224)	(0.0192)	(0.0234)
$\beta_{2,3}$	−0.0339	−0.0519	−0.0338	−0.0519	−0.0338	−0.0519
	(0.0101)	(0.0110)	(0.0100)	(0.0110)	(0.0100)	(0.0116)
$\beta_{1,4}$	0.0093	0.0273	0.0093	0.0273	0.0093	0.0273
	(0.0548)	(0.0215)	(0.0548)	(0.0213)	(0.0548)	(0.0224)
$\beta_{2,4}$	0.0042	0.0110	0.0042	0.0110	0.0042	0.0110
	(0.1238)	(0.0526)	(0.1238)	(0.0522)	(0.1238)	(0.0549)
κ	2.0265		2.0264		2.0263	
	(0.0264)		(0.0257)		(0.0255)	
σ^2		0.3855		0.3858		0.3862
		(0.0009)		(0.0009)		(0.0011)
ϕ_S	4067.21	1066.277	4071.738	1072.0239	4076.578	1078.6964
	(89.2924)	(3.4777)	(61.9349)	(3.4782)	(50.1251)	(3.3496)
ϕ_T	12.2794	4.9687	12.4249	5.1731	12.5715	5.3820
	(0.4035)	(0.0480)	(0.4057)	(0.0529)	(0.4080)	(0.0532)
PLIC	8864239	10392021	8864463	10392832	8864821	10405428

Table 3: MWPL estimates for Weibull and Log-Gaussian models for the correlation model (14). The standard error of the estimates are reported between the parentheses.

	$\phi_{ST} = 0$			$\phi_{ST} = 0.5$			$\phi_{ST} = 1$		
	RMSE	MAE	CRPS	RMSE	MAE	CRPS	RMSE	MAE	CRPS
W	0.4461	0.3486	0.3057	0.4469	0.3491	0.3057	0.4502	0.3503	0.3057
LG	0.4517	0.3555	0.3068	0.4555	0.3585	0.3068	0.4611	0.3629	0.3068
Naïve	MAE= 0.5137, RMSE= 0.4021								

Table 4: Prediticton performances for the Weibull and Log-Gaussian models for different space time interaction.