UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Weibull random fields through Clayton spatial copula

An application to mining haul roads

Eloy Alvarado Narváez
eloy.alvarado@usm.cl

Departamento de industrias
Universidad Técnica Federico Santa María

11 de noviembre de 2024

Table of Contents

Spatial statistics

A brief introduction

- **Spatial statistics** is a branch of statistics focusing on datasets with textbfspatial dependence.
- It includes **a set of statistical methods** whose goal is to describe and predict the spatial variability of the data.

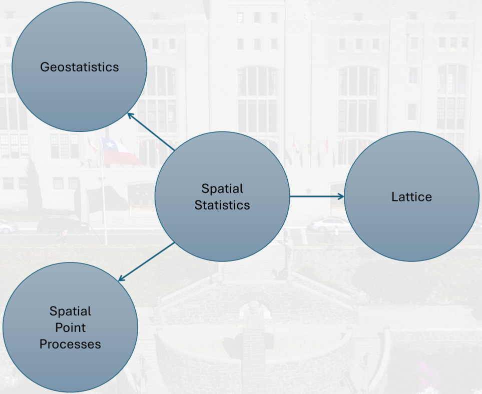


Figure 1: Spatial Statistics subfields

Spatial statistics

Geostatistical data

DEPARTAMENTO DE INDUSTRIAS
UNIVERSIDAD TECNICA FEDERICO SANTA MARIA

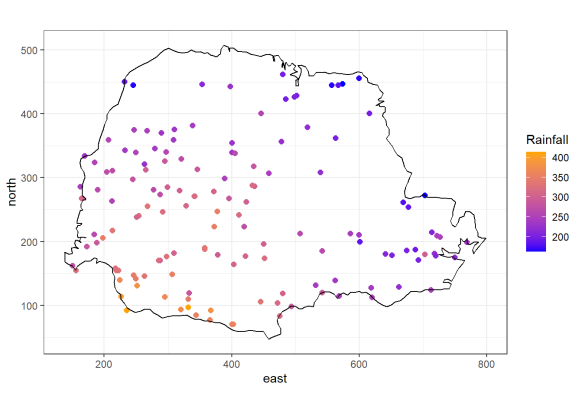


Figure 2: Average rainfall measured at 143 recording stations in Paraná state, Brazil. [Moraga, 2019]

Random fields

DEPARTAMENTO
DE INDUSTRIAS
UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

We briefly review key concepts regarding random fields. Let $\boldsymbol{s} \in \mathbb{R}^d$ be a generic data location in a $d-$dimensional euclidean space.

Definition

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $D \subset \mathbb{R}^d$ be an arbitrary set. For each $\boldsymbol{s} \in D$ the function $Z(\boldsymbol{s}, \cdot) : \Omega \to \mathbb{R}, \omega \to (\boldsymbol{s}, \omega)$ is a random variable, and any collection of random variables $Z = \{Z(\boldsymbol{s}, \cdot), \ \boldsymbol{s} \in D \subset \mathbb{R}^d\}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ is a random field with index set D.

Let assume zero mean $\mathbb{E}(Z(\boldsymbol{s})) = 0$ and unit variance $\mathbb{V}(Z(\boldsymbol{s})) = 1$ and **isotropic correlation function**

$$Corr(Z(\boldsymbol{s}_1), Z(\boldsymbol{s}_2)) = \rho(\boldsymbol{h}), \qquad \boldsymbol{h} = \|\boldsymbol{s}_i - \boldsymbol{s}_j\|$$

i.e, is **weakly stationary**.

Random fields

Non Gaussian Random Fields

- Unfortunately, the Gaussian assumption is rarely met in practice.
- A very flexible class of non-Gaussian RFs can be obtained through a suitable **transformation of one or independent copies** $Z_1, Z_2, \ldots$ of Z a standard Gaussian RF sharing a **common isotropic parametric correlation** $\rho(r, \boldsymbol{\tau})$.
- Specifically, let $Z = \{Z(\boldsymbol{s}), \boldsymbol{s} \in A\}$, $A \subset \mathbb{R}^d$ a Gaussian RF and let $Y = \{Y(\boldsymbol{s}), \boldsymbol{s} \in A\}$ a RF defined through the transformation

$$Y(\boldsymbol{s}) = f(g_1(Z_1(\boldsymbol{s})), g_2(Z_2(\boldsymbol{s})), \ldots.. g_q(Z_q(\boldsymbol{s}))), \quad q \geqslant 1$$

where $f : \mathbb{R}^q \to \mathbb{R}$, $g_i : \mathbb{R} \to \mathbb{R}$, for $i = 1, \ldots, q$ are suitable functions and $Z_1, \ldots, Z_q$, are independent copies of Z.

- Tukey h RF [Xu and Genton, 2017], Skew-Gaussian RF [Zhang and El-Shaarawi, 2010], t-student RF [Bevilacqua et al., 2021], **Weibull RF** [Bevilacqua et al., 2020], among others.

Random fields

Gamma Random Fields

DEPARTAMENTO
DE INDUSTRIAS
UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Gamma Random Fields

Let $\{Z(\boldsymbol{s}),\ \boldsymbol{s} \in S\}$ a zero mean with unit variance weakly stationary Gaussian random process where s represent a location in the domain S and $Z_1, \ldots, Z_m,\ m = 1, 2, \ldots$ independent copies of Z, then the RF $X_m = \{X_m(\boldsymbol{s}),\ \boldsymbol{s} \in S\}$ defined as:

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

is a scaled version of a χ^2 random process [Ma, 2009] with marginal distribution $Gamma(m/2, m/2)$, where the pairs $m/2$, $m/2$ are the shape and rate parameters.

DEPARTAMENTO DE INDUSTRIAS
UNIVERSIDAD TECNICA FEDERICO SANTA MARIA

Weibull Random Fields

Following [Bevilacqua et al., 2020], a random field with Weibull marginal distribution $\{W(\mathbf{s}), \mathbf{s} \in S\}$ can be obtained as:

$$W_\kappa(\mathbf{s}) = \nu(\kappa) X_2(\mathbf{s})^{1/\kappa},$$

where $\nu(\kappa) = \Gamma^{-1}(1 + 1/k)$ and $\kappa > 0$ is a shape parameter. Under this specific parametrization $W_\kappa(s) \sim \mathit{Weibull}(\kappa, \nu(\kappa)), \mathbb{E}(W_\kappa(s)) = 1$ and $\mathbb{V}(W_\kappa(s)) = (\Gamma(1 + 2/\kappa)\nu^2(\kappa) - 1)$.

Non-stationary version of W_κ

A non-stationary version of $W_\kappa(s)$ can be easily obtained through the multiplicative model:

$$Y_\kappa(s) = \mu(s) W_\kappa(s)$$

where $\mu(s) > 0$ is a non random function that specify the mean of Y, *i.e.*$\mathbb{E}(Y_\kappa(s)) = \mu(s)$ and affects its variance $\mathbb{V}(Y_\kappa(s)) = \mu(s)^2(\Gamma(1 + 2/\kappa)\nu^2(\kappa) - 1)$.
A useful parametric specification for $\mu(s)$ is given through a log-linear function.

Spatial Copulas

A **general powerful modeling tool to obtain non-Gaussian random fields with arbitrary marginal distribution** can be obtained under the **copula framework** [Joe, 2014] adapted to the spatial setting.

Definition: Copula

For $d \geqslant 2$ a d−dimensional copula is a distribution function on $[0,1]^d$ with standard uniform marginal distributions. For the vector $\boldsymbol{u} = (U_1, \ldots, U_d)^T$ where $U_i \sim U(0,1)$ then the distribution function

$$C : [0,1]^d \to [0,1]$$

We use the notation $C(\boldsymbol{u}) = C(u_1, \ldots, u_d)$.

The following three properties must hold:

1. $C(u_1, \ldots, u_d)$ is increasing in each component u_i.
2. $C(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i$ for all $i \in \{1, \ldots, d\}$, $u_i \in [0,1]$
3. For all $(a_1, \ldots, a_d), (b_1, \ldots, b_d) \in [0,1]^d$ with $a_i \leqslant b_i$ we have:

$$\sum_{i_1=1}^{2} \cdots \sum_{i_d=1}^{2} (-1)^{i_1+\cdots+i_d} C(u_{1i_1}, \ldots, u_{di_d}) \geqslant 0,$$

where $u_{j1} = a_j$ and $u_{j2} = b_j$ for all $j \in \{1, \ldots, d\}$.

Spatial Copulas

Archimedean copulas

Archimedean generator

We call Archimedean generator any continuous and decreasing function $\Psi : [0, \infty] \to [0, 1]$ that satisfies the conditions

1. $\Psi(0) = 1$
2. $\lim_{t\to\infty} \Psi(t) = 0$
3. Strictly decreasing on $[0, \inf\{t|\Psi(t) = 0\}]$

Archimedean copula

A $d-$dimensional copula C is called Archimedean if it admits the representation

$$C(\boldsymbol{u}) = \Psi(\Psi^{-1}(u_1) + \Psi^{-1}(u_2) + \cdots + \Psi^{-1}(u_d)),$$

for all $\boldsymbol{u}$ in $[0, 1]^d$ and for some Archimedean generator Ψ.

Spatial Copulas

Archimedean copulas

$\Psi_\theta(x)$	$\Psi_\theta^{-1}(y)$	$\theta \in$	Name of copula
$(1+x)^{-1/\theta}$	$y^{-\theta}-1$	$(0,\infty)$	Clayton
$e^{-x^{1/\theta}}$	$(-\log(y))^{\theta}$	$[1,\infty)$	Gumbel
$\frac{1-\theta}{e^x-\theta}$	$\log\left(\frac{1-\theta}{y}+\theta\right)$	$[0,1)$	Ali–Mikhail–Haq

Table 1: Popular completely monotone generators of Archimedean copulas [Mai and Scherer, 2014].

Representation of a subclass of Archimedean copula

Let $\mathbf{E} = (E_1, \dots, E_d)$ a vector of i.i.d. random variables with exponential distribution and M an independent positive random variable with Laplace transform Ψ, then the random vector:

$$\mathbf{U} = \Psi\left(\frac{\mathbf{E}}{M}\right),$$

has the copula C_Ψ as joint distribution function. [Mai and Scherer, 2014]

Spatial Copulas

Archimedean copulas

Archimedean-like spatial copula [Bevilacqua et al., 2024]

Let $\{U(\boldsymbol{s}), \boldsymbol{s} \in A\}$ be a random field with uniform marginals defined as:

$$U(\boldsymbol{s}) = \Psi\left(\frac{E(\boldsymbol{s})}{M(\boldsymbol{s})}\right),$$

where E is a random field with standard exponential marginal distribution, M is a positive random field and $\Psi(\cdot)$ is the Laplace transform of $M(\boldsymbol{s})$.

- This general construction allows classical Archimedean copulas to be extended into the spatial setting.
- [Bevilacqua et al., 2024] focus in the special case where $E \equiv G_2$ and M is a copy of a Gamma random field, that is $M \equiv G_\nu$ with marginal distribution $Gamma(\nu/2, 1)$. This construction can be seen as the spatial extension of the Clayton copula because it follow same the choosing of M and Ψ as in classical copula theory.

Spatial Copulas

Gaussian copula random field

- Specifically, given

$$U = \{U(\boldsymbol{s}), s \in A\}$$

a RF with **standard uniform marginals**, a random field $Y = \{Y(\boldsymbol{s}), s \in A\}$ with **an arbitrary marginal cumulative distribution function (cdf)** $F_Y(\theta)$ can be obtained trough the transformation

$$Y(\boldsymbol{s}) = F_Y^{-1}(U(\boldsymbol{s}), \theta),$$

where F_Y^{-1} is the generalized inverse of $F_Y(\theta)$.

- The choice $U(\boldsymbol{s}) = \Phi(Z(\boldsymbol{s}))$ where Φ is the *cdf* of the Gaussian distribution **leads to the so-called Gaussian copula random field**.

$$Y(\boldsymbol{s}) = F_Y^{-1}(\Phi(Z(\boldsymbol{s})), \theta),$$

Spatial Copulas

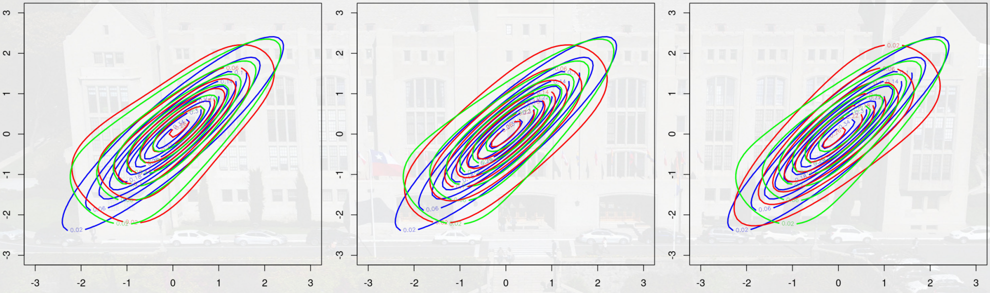
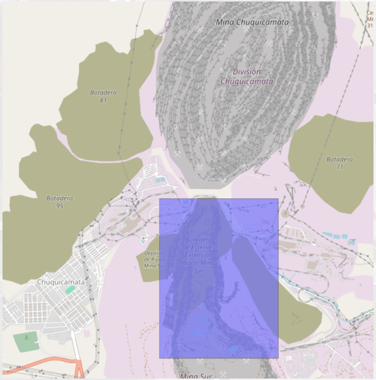


Figure 3: Contours of the Gaussian copula (blue), $W_{\kappa}^{RF}(\boldsymbol{s})$ (green), and $W_{\kappa,\nu}^{C}(\boldsymbol{s})$ (red). From left to right, for correlation $\rho(\boldsymbol{h}) = 0.9$, and symmetry parameter $\nu = \{1, 2, 6\}$, from left to right.

Application

Mining haul road data



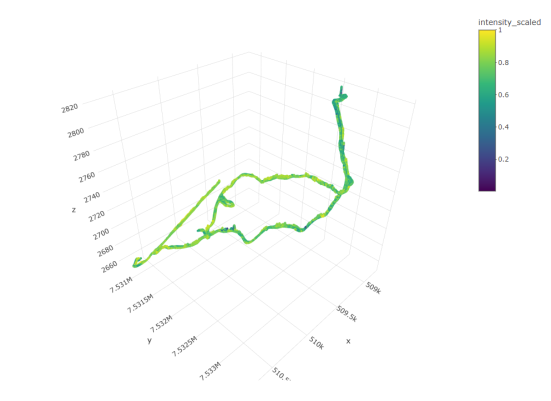


Figure 4: Mining haul road data, Mina Sur, Chile. The left panel shows the geographic location of the mine with the selected haul road section highlighted in blue. The right panel illustrates the corresponding road trajectory captured by the data collection system, visualizing the road's layout and path.

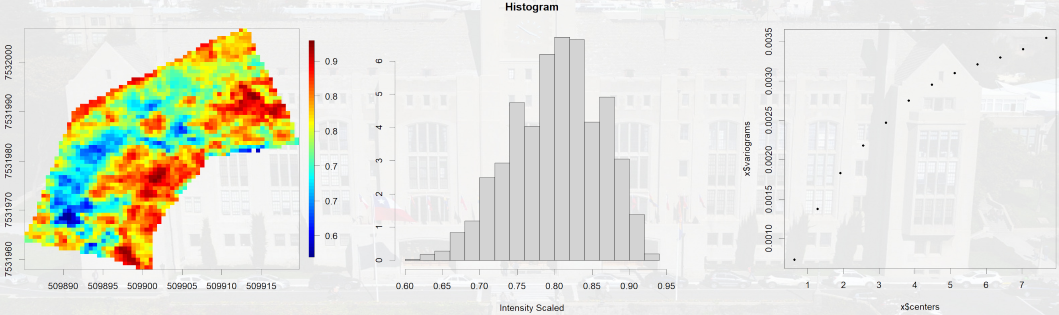


Figure 5: Exploratory Data Analysis of the LiDAR Intensity in the selected cluster.

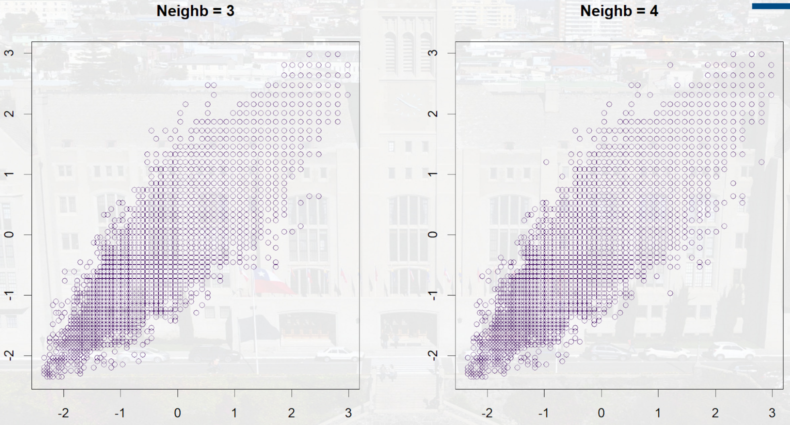


Figure 6: Spatial scatterplots of LiDAR Intensity data for two neighborhood orders after normal score transformation.

Application

Methodology

We considered three Weibull random fields obtained as:

$$W_{\kappa}^{(RF)}(\boldsymbol{s}) = Y_{\kappa}(\boldsymbol{s}) \qquad W_{\kappa}^{(G)}(\boldsymbol{s}) = F_W^{-1}[\Phi\{Z(\boldsymbol{s})\}] \qquad W_{\nu,\kappa}^{(C)}(\boldsymbol{s}) = F_W^{-1}(U_{\nu}(\boldsymbol{s})),$$

considering a log-linear link function:

$$\log(\mu(\boldsymbol{s})) = \beta_0 + \beta_1 B(\boldsymbol{s}),$$

where $\mu(s)$ represent the mean of the LiDAR Intensity, and $B(s)$ the blue component from the RGB spectrum.

Application

Results

Estimation	$W_{\kappa}^{RF}(\boldsymbol{s})$	$W_{\kappa}^{G}(\boldsymbol{s})$	$W_{\kappa,2}^{C}(\boldsymbol{s})$	$W_{\kappa,4}^{C}(\boldsymbol{s})$	$W_{\kappa,6}^{C}(\boldsymbol{s})$
$\widehat{\beta}_0$	−0.8983 (0.0094)	−0.8925 (0.0109)	−0.8965 (0.0107)	−0.9003 (0.0092)	−0.8996 (0.0097)
$\widehat{\beta}_1$	1.0167 (0.0116)	1.0152 (0.0149)	1.0149 (0.0149)	1.0202 (0.0133)	1.0201 (0.0140)
$\widehat{\alpha}$	4.3078 (2.7169)	5.8307 (1.0411)	7.7469 (0.6535)	4.8704 (0.1116)	5.8930 (0.2132)
$\widehat{\kappa}$	76.1658 (9.1958)	79.9035 (10.1454)	77.9041 (6.6301)	76.7420 (4.5341)	78.9485 (7.0045)
CLAIC	−1216644	−1175458	−1226867	−1225818	**−1227133**

Table 2: Estimations of the random fields $W_{\kappa}^{RF}(\boldsymbol{s})$, $W_{\kappa}^{G}(\boldsymbol{s})$, $W_{\kappa,2}^{C}(\boldsymbol{s})$, $W_{\kappa,4}^{C}(\boldsymbol{s})$, $W_{\kappa,6}^{C}(\boldsymbol{s})$ and the correlation model $GW_{0,0.25,\alpha}(\boldsymbol{h})$. In parentheses (for each parameter), the associated standard error is provided. The last line reports the CLAIC information criterion.

Application

Results

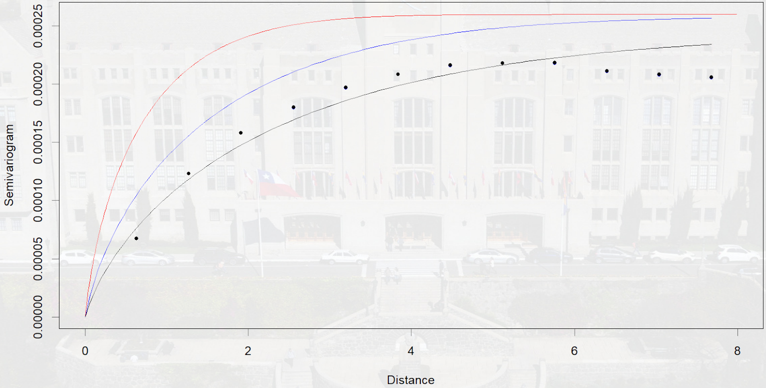


Figure 7: Empirical and estimated semivariogram of LiDAR Intensity of $W_{\kappa}^{G}(\boldsymbol{s})$ (the Gaussian copula, red line), $W_{\kappa}^{RF}(\boldsymbol{s})$ (the Weibull random field, blue line) and $W_{\kappa,6}^{C}(\boldsymbol{s})$ (the Clayton copula, black line).

Concluding Remarks

- **Weibull Random Fields (RFs)** can be constructed through several approaches:
  1. As transformations of underlying Gaussian RFs [Bevilacqua et al., 2020]
  2. Using Gaussian spatial copulas
  3. Using Clayton-like spatial copulas
- **Clayton-like spatial copulas** offer flexible dependence structures, enabling the construction of RFs with arbitrary marginal distributions.
- **Mining haul road LiDAR Intensity data:** The proposed methodology outperforms our benchmark models (W^{RF} and W^{G}) based on the Composite Likelihood Akaike Information Criterion, providing a practical demonstration of its effectiveness.
- This study serves as an initial step towards creating input models for asset survival, with implications for reliability and maintenance engineering.

Open Research Questions:

- How does the performance of the Clayton spatial copula vary with different marginal models?
- How does the Clayton spatial copula compare to the Gumbel and AMH spatial copulas?

References

Bevilacqua, M., Alvarado, E., and Caamaño-Carrillo, C. (2024).
A flexible clayton-like spatial copula with application to bounded support data.
Journal of Multivariate Analysis, 201:105277.

Bevilacqua, M., Caamaño-Carrillo, C., Arellano-Valle, R. B., and Morales-Oñate, V. (2021).
Non-gaussian geostatistical modeling using (skew) t processes.
Scandinavian Journal of Statistics, 48(1):212–245.

Bevilacqua, M., Caamaño-Carrillo, C., and Gaetan, C. (2020).
On modeling positive continuous data with spatiotemporal dependence.
Environmetrics, 31(7):e2632.

Joe, H. (2014).
Dependence modeling with copulas.
CRC press.

Ma, C. (2009).
χ^2 random fields in space and time.
IEEE transactions on signal processing, 58(1):378–383.

References (cont.)

DEPARTAMENTO
DE INDUSTRIAS
UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Mai, J.-F. and Scherer, M. (2014).
Financial engineering with copulas explained.
Springer.

Moraga, P. (2019).
Geospatial health data: Modeling and visualization with R-INLA and shiny.
Chapman and Hall/CRC.

Xu, G. and Genton, M. G. (2017).
Tukey g-and-h random fields.
Journal of the American Statistical Association, 112(519):1236–1249.

Zhang, H. and El-Shaarawi, A. (2010).
On spatial skew-gaussian processes and applications.
Environmetrics: The official journal of the International Environmetrics Society, 21(1):33–47.

Questions?