Open Journal of Statistics, 2023, 13, 46-60
https://www.scirp.org/journal/ojs
ISSN Online: 2161-7198
ISSN Print: 2161-718X

Sampling Geostatistical Structures in Extremal Framework

Fabrice Ouoba[1], Hay Yoba Talkibing[2], Diakarya Barro[3]

[1]Université de Fada N'Gourma, Fada N'Gourma, Burkina Faso
[2]Université Joseph Ki-Zerbo, Ouagadougou, Burkina Faso
[3]Université Thomas Sankara, Ouagadougou, Burkina Faso
Email: didifab@yahoo.fr, talkibingfils@yahoo.fr, dbarro2@gmail.com

How to cite this paper: Ouoba, F., Talkibing, H.Y. and Barro, D. (2023) Sampling Geostatistical Structures in Extremal Framework. *Open Journal of Statistics*, **13**, 46-60.
https://doi.org/10.4236/ojs.2023.131004

Received: December 5, 2022
Accepted: February 21, 2023
Published: February 24, 2023

Abstract

Geostatistics of extreme values makes it possible to model the asymptotic behavior of random phenomena that depend on time or space. In this paper, we propose new models of the extremal coefficient of a stationary random field where the cumulative distribution is associated with a multivariate copula. More precisely, some models of extensions of the extremogram and these derivatives are built in a spatial framework. Moreover, both these two geostatistical tools are modeled using the extremal variogram which characterizes the asymptotic stochastic behavior of the phenomena.

Keywords

Extremal Index, Extremogram, Variogram, Copulas, Stationary Process, Extreme Values

1. Introduction

In spatial statistical analysis, variograms and covariance functions are technical tools used to describe how spatial continuity changes with a given separation distance between two pairs of stations. Thus, the classical variogram provides a framework for modeling and predicting the variability of a given spatial stochastic process. Geostatistics provides tools for statistical analysis of spatial or spatio temporal datasets. This branch of statistics was developed originally in the year 1930 by a pioneering work of George Matheron [1] to predict the probability distributions of more grades for mining operations. Since, geostatistics became a subdomain of statistics based random fields including petroleum geology, hydrogeology, geochemistry, geometallurgy, geography, forestry, environmental control, landscape ecology, soil science and agriculture.

The family of copulas provides a natural way to construct multivariate distribtutions whose marginals are uniform and not necessarily exchangeable. Let $(X_1,\cdots,X_n)$ be a random vector with multivariate continuous distribution function (c.d.f.) H and c.d.f marginal $H_1,\cdots,H_n$. The copula of X or the c.d.f. H respectively is the multivariate c.d.f. C of the random vector uniform $U=[H_1(X_1),\cdots,H_n(X_n)]$ each component of U is standard uniformly distributed, *i.e.*, $U_i \sim U(0,1)$ for $i=1,\cdots,n$.

More specifically, every n-copula must satisfy the n-increasing property [2]. That means that, for any rectangle $B=[a,b]^n \subseteq \mathbb{R}^n$, the B-volume C_B of C is positive, that is,

$$C_B=\int_B \mathrm{d}C(u)=\sum_{i_1=1}^{2}\cdots\sum_{i_n=1}^{2}(-1)^{i_1+\cdots+i_n}C\left(u_{1i_1},\cdots,u_{1i_n}\right)\geq 0. \quad (1)$$

While modeling the main geostatistical tools Ouoba & *et al.* [3] proposed a family of copulas based variogram, correlogram and madogam and they pointed out that these tools do not take into account the extremal data observed in the different observation sites. However, copulas make it possible to model the extreme data and detect any nonlinear link between different observation sites. It is therefore necessary to express the variogram and the covariogram according to the copula in order to be able to model the spatial structuring even if our database contains extreme values and to be able to detect the presence of some nonlinear dependence.

The variogrm and covariogram $\hat{c}(s_i,s_j)$ are linked by the copula function via the relation:

$$\vartheta(s_i,s_j)=\sigma_Z^2(s_i)+\sigma_Z^2(s_j)-2\int_0^1\int_0^1 F_Z^{-1}(u)F_Z^{-1}(v)c(u,v)\mathrm{d}u\mathrm{d}v-2m_i m_j,$$

and

$$\hat{c}(s_i,s_j)=\int_0^1\int_0^1 F_Z^{-1}(u)F_Z^{-1}(v)c(u,v)\mathrm{d}u\mathrm{d}v-m_i m_j;$$

where m_i and m_j the respective averages of $Z(s_i)$ and $Z(s_j)$; $c(u,v)$ the copula density function attached to $Z(s_i)$ and $Z(s_j)$.

The major contribution of this paper is to provide tools to model the distributional dependence of spatial or temporal extremes in various sites in a study area. Specifically, we offer geostatistical tools such as the extremal coefficient, the extremogram and its derivatives based on copula and unit Frechet processes for modeling spatial extremes. In Section 2, we develop the tools needed to achieve our goals. Our main results are given in Section 3, where we propose new models of the extremal coefficient of a stationary spatial field using multivariate copula and extensions of the extremogram and the crossed extremmogram in a spatial framework using the extremal variogram which characterizes the asymptotic stochastic behavior of phenomena.

2. Materials and Methods

In this section, we collect the necessary definitions and usefull properties on multivariate extremes.

Multivariate extreme values (MEV) theory present the framework of coordinatewise maxima. Towards a multivariate analogue of Fisher-Tippett we are looking for some sort of multivariate limit distribution for conveniently normalized vectors of multivariate maxima. For an arbitrary index of set T denoting generally a space of time, a random vector $Y_t = \{Y_j(t); 1 \le j \le m, t \in T\}$ in $\mathbb{R}^m$ is said to be max-stable if, for all $n \in \mathbb{N}$, every $Y_j(t) = \left(Y_j^{(1)}(t); \cdots; Y_j^{(n)}(t)\right)$ is a n-dimensionnal max-stable vector, that is, there exists suitable and time-varying non-random sequences $\{a_n(t) > 0\}$ and $\{b_n(t) \in \mathbb{R}^d\}$ such as

$$\frac{1}{a_n(t)}\left[M_n(t) - b_n(t)\right] \xrightarrow{f.d.d} X(t); t \in T, \tag{2}$$

where $\xrightarrow{f.d.d}$ denotes the convergence for the finite-dimensional distributions while $M_n(t) = \max_{1 \le i \le n}(X_i(t)); t \in T$ being the component-wise maxima of the time-variying vector $X(t)$.

As in the non-spatial analysis, several canonical representations of max-stable processes have been suggested in spatial extreme values context. In the same vain, Barro *et al.* (see [4]) have proposed a general form of the one-dimensional marginal of the max-stable ST process $\{Y_t\}$ where $Y_t(x) = Y(x_t)$; $x_t \in \chi_D \times T \subset R^3$. $(Y_j^s(t)); j \ge 0; t \in T; s \in S$ such that, for each fixed couple (t, s), the sequence is independent and identically distributed according to a joint cumulative function G_t^s. Under the assumption that this function is max-stable, every univariate margins $G_{t,i}^s$ lies its own domain of attraction and is expressed by on the space of interest $S^+_{\xi_i,t,s} = \{z \in \mathbb{R}; \sigma_{i,t,s} + \xi_{i,t,s}(y_{t,i}^s - \mu_{i,t,s}) > 0; 1 \le i \le n\}$ by

$$G_i(y_i(s)) = \begin{cases} \exp\left\{-\left[1 + \xi_i(s)\left(\frac{y_i(s) - \mu_i(s)}{\sigma_i(s)}\right)\right]^{\frac{-1}{\xi_i(s_i)}}\right\} & \text{if } \xi_i(s) \ne 0 \\ \exp\left\{-\exp\left\{-\left(\frac{y_i(s) - \mu_i(s)}{\sigma_i(s)}\right)\right\}\right\} & \text{if } \xi_i(s) = 0 \end{cases}; \tag{3}$$

and for all site s, the parameters $\{\mu_{i,t,s} \in \mathbb{R}\}$, $\{\sigma_{i,t,s} > 0\}$ and $\{\xi_{i,t,s} \in \mathbb{R}\}$ are referred to as the location, the scale and the shape parameters respectively. Particularly, the different values of $\xi_i(s) \in \mathbb{R}$ allows 3 to be a spatial EV model, that is, to belong either to Frechet family, the Weibull one or Gumbel one.

In multivariate case, if the one-dimensional margins of F are unit-Fréchet distributed let M be a non-empty subset of $N = \{1, \cdots, n\}$ and c_M the n-dimensional vector of which the jth coordinate is one or zero according to $j \in M$ or $j \notin M$. Then, the multivariate, θ_M is defined on the n-dimensional unit simplex, $S_n = \left\{(t_1, \cdots, t_n) \in [0,1]^n, \sum_{i=1}^n t_i \le 1\right\}$, such as,

$$\theta_M = V(c_M) = \int_{S_n} \max_{j \in M}\left(\frac{w_j}{\|w\|_1}\right) \mathrm{d}H(w_j), \tag{4}$$

where H is a finite non-negative measure of probability and $\|.\|_1$, the 1-norm, see

[5] [6] [7]. Particularly

$$P\left[F_1(x_1) \le p, \cdots, F_n(x_n) \le p\right] = p^\theta \text{ for all } 0 < p < 1. \quad (5)$$

In spatial study, a natural way to measure dependence among spatial maxima stems from considering the distribution of the largest value that might be observed on domain of study.

Our main results are summaried by the following sections.

3. Main Results

Extreme value analysis is frequently used to model spatio-temporal data, for which the phenomenon of dependence is often intrinsic. In this article, this dependence is measured for a process studied on different space stations. Authors such as Ledford *et al.* [8] and Coles and *et al.* [9], proposed tools to measure lower and upper tail dependence for a bivariate couple (X, Y) of distribution function (F_X, F_Y). We propose new tools for measuring the dependence of a spatial process $Z = \left\{Z(s), s \in \mathbb{R}^d\right\}$ at two measurement points s and $s+h$.

More precisely, we first model the extremal coefficient using the copula in a spatial framework. Then, we propose other versions of the spatio-temporal extreme modeling tools such as the extremogram and the crossed extremogram using the copulas. Finally, we propose tools to model the asymptotic dependence of extremes in a spatial context.

3.1. Extremal Dependence Index and Copulas

The study of MEV theory have been extended both to spatial and multivariate contexts these last years. The extremal coefficient, a natural dependence measures for extreme value models which provides the magnitude of the asymptotic dependence of a random field at two points of the domain. This section gives the relationship between the extremal coefficient via copula.

Proposition 1. Let $\left\{Z(s), s \in \mathbb{R}^2\right\}$ be stationary max-stable random process with Fréchet marginal. Then, the extremal copula-based coefficient is given by:

$$\theta(h) = \begin{cases} u_\beta(z)\left[\mu + \dfrac{\int_0^1 F_Z^{-1}(u)\,\mathrm{d}C_h(u,u) - \mu}{\Gamma(1-\xi)}\right] & \text{if } \xi \ne 0 \\ \exp\left\{\dfrac{\int_0^1 F_Z^{-1}(u)\,\mathrm{d}C_h(u,u) - \mu}{\sigma}\right\} & \text{if } \xi = 0 \end{cases}, \quad (6)$$

where

$$u_\beta(z) = \begin{cases} \left[1 + \xi\left(\dfrac{z-\mu}{\sigma}\right)\right]^{1/\xi} & \text{if } 1 + \xi\left(\dfrac{z-\mu}{\sigma}\right) > 0 \\ 0 & \text{if } 1 + \xi\left(\dfrac{z-\mu}{\sigma}\right) \le 0 \end{cases}; \text{ for all } z \in \mathbb{R},$$

and Γ, the well-known gamma function, such as for all

$$z > 0, \quad \Gamma(z) = \int_0^{+\infty} t^{z-1} e^{-t} dt.$$

Proof. Let Z be a stationary random field of the second order of form parameter $\xi < 1$. The extremal coefficient is given using the underlying madogram (see [9]) by:

$$\theta(h) = \begin{cases} u_\beta(z)\left[\mu + \frac{M(h)}{\Gamma(1-\xi)}\right] & \text{if } \xi \neq 0 \\ \exp\left\{\frac{M(h)}{\sigma}\right\} & \text{if } \xi = 0 \end{cases}.$$

where M_h is the semi-variogram given by:

$$M(h) = \frac{E(|Z(x+h) - Z(x)|)}{2}. \quad (7)$$

For all $x \in \mathbb{R}^2$ and by taking into account the fact that

$$|Z(x+h) - Z(x)| = 2\max[Z(x+h), Z(x)] - Z(x+h) - Z(x);$$

then, the relation 7 provides:

$$M(h) = \frac{E(2\max[Z(x+h), Z(x)] - Z(x+h) - Z(x))}{2}.$$

So, it follows that:

$$M(h) = E(\max[Z(x+h), Z(x)]) - \frac{1}{2}[E(Z(x+h)) + E(Z(x))]$$

Then, for a stricly continous context,

$$M(h) = E(\max[Z(x+h), Z(x)]) - \mu, \quad (8)$$

where $\mu = E(Z(x+h)) = E(Z(x))$ is the means of $Z(.)$, stationary at second order. Thus, we have

$$E(\max[Z(x+h), Z(x)]) = \int_{-\infty}^{+\infty} z dC_h(F_Z(z), F_Z(z)).$$

Which gives

$$E(\max[Z(x+h), Z(x)]) = \int_0^1 F_Z^{-1}(u) dC_h(u,u). \quad (9)$$

Then, using the Formula (9) in (8), one obtain

$$M(h) = \int_0^1 F_Z^{-1}(u) dC_h(u,u) - \mu. \quad (10)$$

By using the relation (10) in the expression of the coefficient extremal we get

$$\theta(h) = \begin{cases} u_\beta(z)\left[\mu + \frac{\int_0^1 F_Z^{-1}(u) dC_h(u,u) - \mu}{\Gamma(1-\xi)}\right] & \text{if } \xi \neq 0 \\ \exp\left\{\frac{\int_0^1 F_Z^{-1}(u) dC_h(u,u) - \mu}{\sigma}\right\} & \text{if } \xi = 0 \end{cases}.$$

Finally, it yields the relation (6) as disserted.

Let Z be a max-stable random field. The extremal coefficient (**Figure 1**) and the copula function are related differently depending on the marginal distribution of the Z process.

Proposition 2. Let Z be a spatial domain of a stationary max-stable model G with either or Gumbel or Weibull univariate marginal. Then, the extremal coefficient is given by:

$$\theta(h)=\begin{cases}\left[1-G\left(F_Z,C_h,u\right)+\mu\right]^{-1} & \text{for standard Weibull margin}\\ \exp\left(G\left(F_Z,C_h,u\right)-\mu\right) & \text{for standard Gumbel margin}\end{cases} \tag{11}$$

where $G\left(F_Z,C_h,u\right)=\int_0^1 F_Z^{-1}(u)\mathrm{d}C_h(u,u)$.

Proof. Dealing with the case where the margins of Z are distributed according the Weibull model, it is well known that the extremal coefficient and the madogram are associated by the relation

$$\theta(h)=\left[1-M(h)\right]^{-1}.$$

So, using (10) in this relation, it comes, under the existence, that

$$\theta(h)=\left[1-\int_0^1 F_Z^{-1}(u)\mathrm{d}C_h(u,u)+\mu\right]^{-1}.$$

Hence the first result of (11).

Similarly, if the margins of Z are Brown-Resnick model (see [3]), then $\theta(h)=\exp\left(M(h)\right)$. Furthermore, by using (10) in this relationship, it comes that

$$\theta(h)=\exp\left(\int_0^1 F_Z^{-1}(u)\mathrm{d}C_h(u,u)-\mu\right).$$

Hence the last result of (11).

The following section allows us to construct a model of the extremogram and the cross-extremogram via copula function to determine the distributional dependence of the random variables of the random field Z depending on the inter-site distance.

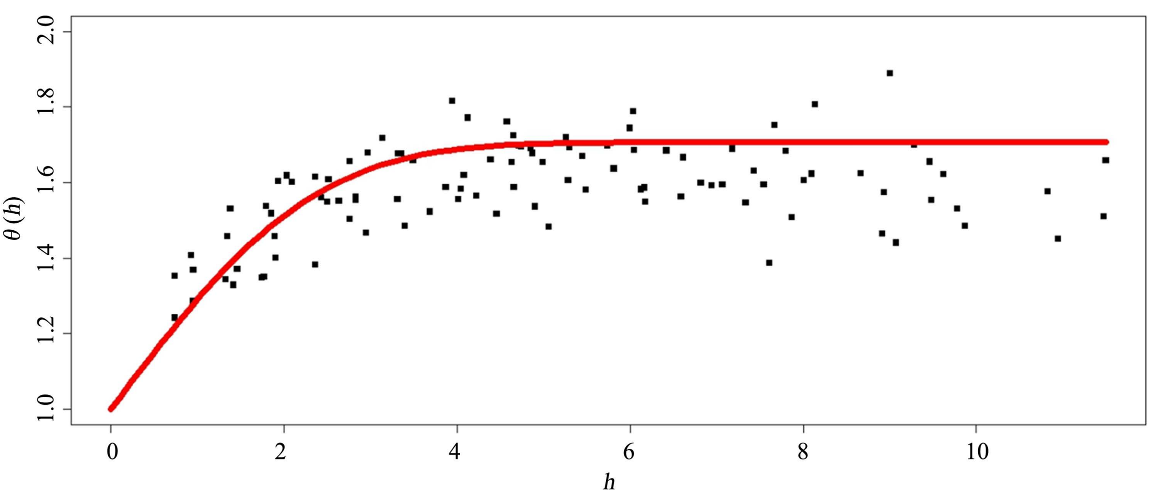

Figure 1. Graph of Extremal coefficient for a Brown-Resnick process 1.

3.2. Sampling Extremogram with Copulas

In this subsection, we model the extremogram function (**Figure 2**) using a copula function for all $A \subset \mathbb{R}_*^+$ and $a \in A$. By using the tools proposed in ([10] [11] [12]) and by considering that the variables $Z(s_i)$ and $Z(s_j)$ have uniform margins, we obtain the following proposition.

Proposition 3. Consider $\{F_{i,Z}; i = 1, \cdots, n\}$ the distribution function of the random variable Z_i and $U_i = F_{i,Z}(Z)$ the uniform transformation of $F_{i,Z}$. Then, a copula-based extremogram is given, for all $x_i, x_j \in \mathbb{R}^d$, by:

$$\rho_{AA}(h_{ij}) = \rho_{(a,+\infty)}(h_{ij}) = 2 - \lim_{u \to 1^-} \frac{1 - C_{h_{ij}}(u,u)}{1-u},$$

where h_{ij} is the separating distance between x_i, x_j.

Proof. It is well known that $\rho_{AA}(h_{ij}) = \lim_{z \to \infty} P\left(\frac{Z(x_j)}{z} \in A / \frac{Z(x_i)}{z} \in A\right)$. Such as: $A = (a, +\infty)$, this expression can be written as,

$$\rho_{AA}(h_{ij}) = \lim_{z \to +\infty} \frac{P\left(\frac{Z(x_j)}{z} > a, \frac{Z(x_i)}{z} > a\right)}{P\left(\frac{Z(x_i)}{z} > a\right)}.$$

Then, it is easy to show that,

$$\rho_{AA}(h_{ij}) = \lim_{z \to +\infty} \frac{P\left(Z(x_j) > az, Z(x_i) > az\right)}{P\left(Z(x_i) > az\right)},$$

Z being a stationary random field. Under the assumption that $F_i(az) = F_j(az) = u$.

Then, it follows that,

$$\rho_{AA}(h_{ij}) = \lim_{u \to 1^-} \frac{P\left(U_j > u, U_i > u\right)}{P\left(U_i > u\right)}.$$

Nevertheless, using the survival copula, when have:

$$P\left(U_j > u, U_i > u\right) = 1 - u - u + C_{h_{ij}}(u,u).$$

Therefore,

$$\rho_{AA}(h_{ij}) = \lim_{u \to 1^-} \frac{1 - 2u + C_{h_{ij}}(u,u)}{1-u}.$$

Then, based on a result of Cooley & *et al.* [9], it follows that:

$$\rho_{AA}(h_{ij}) = \lim_{u \to 1^-} \left[2 - \frac{1 - C_{h_{ij}}(u,u)}{1-u}\right].$$

So, as disserted

$$\rho_{AA}(h_{ij}) = \rho_{(a,+\infty)}(h) = 2 - \lim_{u \to 1^-} \frac{1 - C_{h_{ij}}(u,u)}{1-u}.$$

In the particular case, where $a=1$, that is $F_i(az)=F_i(z)$ the extremogram merges with the upper tail dependence measure. So,

$$\rho_{AA}(h_{ij})=\lim_{u\to1^-}2-\frac{1-C_{h_{ij}}(u,u)}{1-u}=2-\lim_{u\to1^-}\frac{1-C_{h_{ij}}(u,u)}{1-u}=\chi(h_{ij}).$$

For the particular case where $a=1$. Moreover If $\rho_{(1,+\infty)}(h_{ij})=0$, then the random variables Z_i and Z_j are asymptotically independent.

In a second case, considering that $A\subset\mathbb{R}_*^-$ and $a\in A$, we obtain next relation of the extremogram via the underlying copula. In particular, if χ_h is reduced to a single site *x*, the law of Y^* is either the Frechet distribution, the Gumbel or the Weibull distribution.

The following result provides a copula-based extension of the extremogram of the process.

Proposition 4. The extremogram ρ_{AA} and the copula function $C_{h_{ij}}$ are linked by the relation:

$$\rho_{AA}(h_{ij})=\lim_{u\to0^+}\frac{C_{h_{ij}}(u,u)}{u},u\in[0,1]. \tag{12}$$

Proof. It is well known that $\rho_{AA}(h_{ij})=\lim_{z\to-\infty}P\left(\frac{Z(x_j)}{z}\in A/\frac{Z(x_i)}{z}\in A\right)$.

Since $A=(-\infty,a)$, it follows that:

$$\rho_{AA}(h_{ij})=\lim_{z\to-\infty}P\left(\frac{Z(x_j)}{z}\le a/\frac{Z(x_i)}{z}\le a\right).$$

Then,

$$\rho_{AA}(h_{ij})=\lim_{z\to-\infty}P\left(Z(x_j)\le az/Z(x_i)\le az\right).$$

Thus,

$$\rho_{AA}(h_{ij})=\lim_{u\to0^+}P\left(U_j\le u/U_i\le u\right).$$

Therefore,

$$\rho_{AA}(h_{ij})=\lim_{u\to0^+}\frac{P\left(U_j\le u,U_i\le u\right)}{P(U_i\le u)}.$$

Hence the result (12) as disserted.

The following subsection gives a relation between the cross-extremogram and the copula function.

3.3. Cross-Extremogram Sampling with Spatial Copulas

The following result provides a characterization of the cross extremogram in a copula contex, for two given sites s_i and s_j.

Theorem 5. For two given sites s_i and s_j separated by h_{ij}, the extremal coefficient it given by:

$$\rho_{AB}(h_{ij})=1-\lim_{(u_{1i},u_{2j})\to(1^-,1^-)}\frac{u_{2j}-C_{h_{ij}}(u_{1i},u_{2j})}{1-u_{1i}}. \tag{13}$$

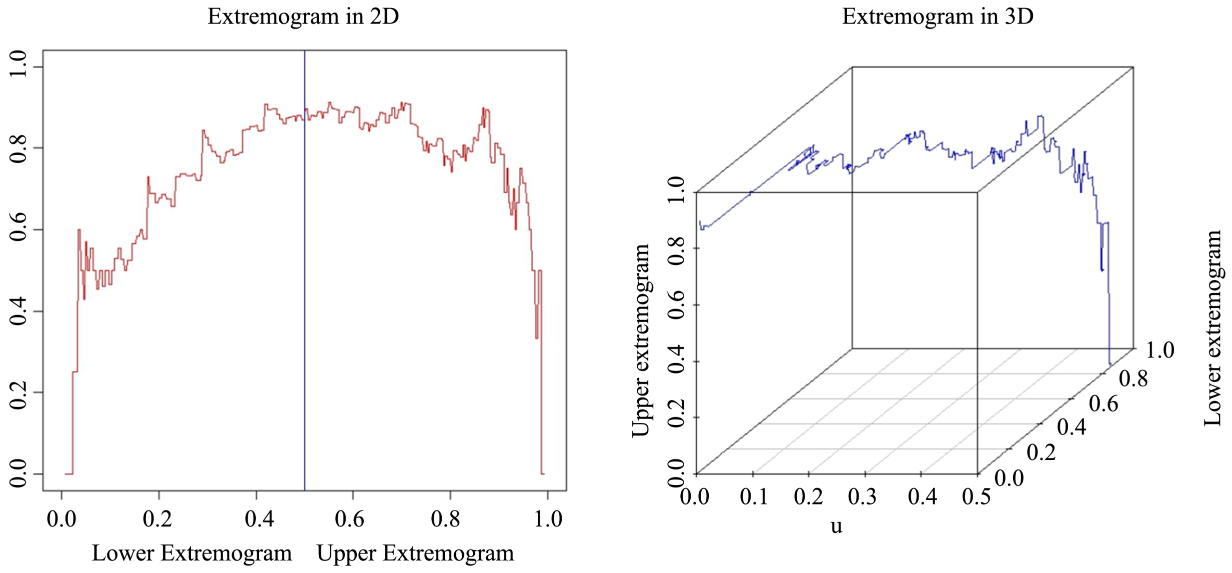

Figure 2. Graph of theoretical extremogram. This figures gives the representation in 2D and 3D for $A=(-\infty,1)$ and $A=(1,+\infty)$. We denote by upper extremogram for $A=(1,+\infty)$ and lower extremogram for $A=(-\infty,1)$. above the figure.

Proof. Let $F_{ji}(Z(x_i))=U_{ji}$ be the univariate distribution functions obtained by integral transforms to the variables $Z_j(x_i)$ with $x_j-x_i=h_{ij}$, $1\le i,j\le n$, $i\ne j$. It is well known that

$$\rho_{AB}(h_{ij})=\lim_{z\to\infty}P\big(Z_2(s_j)\in zB/Z_1(s_i)\in zA\big).$$

Since $A=(a,\infty)$ and $B=(b,\infty)$, it follows that

$$\rho_{AB}(h_{ij})=\lim_{z\to\infty}P\big(Z_2(s_j)\in zB/Z_1(s_i)\in zA\big).$$

It follow that,

$$\rho_{AB}(h_{ij})=\lim_{z\to\infty}P\big(Z_2(s_j)\ge bz/Z_1(s_i)\ge az\big).$$

So,

$$\rho_{AB}(h_{ij})=\lim_{z\to\infty}\frac{P\big(Z_2(s_j)\ge bz,Z_1(s_i)\ge az\big)}{P\big(Z_1(s_i)\ge az\big)}.$$

$$\rho_{AB}(h_{ij})=\lim_{z\to\infty}\frac{\hat{H}_{h_{ij}}(bz,az)}{P\big(Z_1(s_i)\ge az\big)}$$

with $\hat{H}_{h_{ij}}(bz,az)$ the survival function of the variables $Z_2(s_j)$ and $Z_1(s_i)$.

Moreover, if $C_{h_{ij}}$ is the jointed copula underlying the distribution of $Z_1(s_i)$ and $Z_2(s_j)$, then, it follows that:

$$\hat{H}_{h_{ij}}(bz,az)=1-F_{1i}(az)-F_{2j}(bz)+H_{h_{ij}}(bz,az)$$

so,

$$\hat{H}_{h_{ij}}(bz,az)=1-F_{1i}(az)-F_{2j}(bz)+C_{h_{ij}}\big(F_{2j}(bz),F_{1i}(az)\big),$$

Likewise

$$P\left(Z_1\left(s_i\right) \geq az\right) = 1 - P\left(Z_1\left(s_i\right) < az\right) = 1 - F_{1i}\left(az\right).$$

By replacing these two last relations in (8), we obtain the following result:

$$\rho_{AB}\left(h_{ij}\right) = \lim_{z\to\infty} \frac{1 - F_{1i}\left(az\right) - F_{2j}\left(bz\right) + C_{h_{ij}}\left(F_{1i}\left(az\right), F_{2j}\left(bz\right)\right)}{1 - F_{1i}\left(az\right)}. \tag{14}$$

Let us consider $u_{1i} = F_{1i}\left(az\right)$ et $u_{2j} = F_{2j}\left(bz\right)$. When $z \to +\infty$ then $u_{1i} \to 1^-$ and $u_{2j} \to 1^-$. By using these transformations in the relation (14), it follows that:

$$\rho_{AB}\left(h_{ij}\right) = \lim_{\left(u_{1i},u_{2j}\right)\to\left(1^-,1^-\right)} \frac{1 - u_{1i} - u_{2j} + C_{h_{ij}}\left(u_{1i}, u_{2j}\right)}{1 - u_{1i}}.$$

So

$$\rho_{AB}\left(h_{ij}\right) = 1 - \lim_{\left(u_{1i},u_{2j}\right)\to\left(1^-,1^-\right)} \frac{u_{2j} - C_{h_{ij}}\left(u_{1i}, u_{2j}\right)}{1 - u_{1i}}.$$

Hence the result (13) as disserted.

The following results provides an asymptotic statement.

Proposition 6. Consider $F_{ij}\left(az\right) = u_{ij} \in [0,1]$ the distribution function of the variable $Z_j\left(x_i\right)$. If $z \mapsto +\infty$, then $u_{ij} \mapsto 1^-$ The relation (16) is written according to the copula by the relation:

$$\begin{pmatrix} \rho_{AA}^{11}\left(h_{ij}\right) \\ \rho_{BB}^{22}\left(h_{ij}\right) \\ \rho_{AB}^{12}\left(h_{ij}\right) \\ \rho_{BA}^{21}\left(h_{ij}\right) \end{pmatrix} = \begin{pmatrix} 1 - \lim\limits_{u_{11}\to 1^-} \dfrac{u_{11} - C_{h_{ij}}\left(u_{11}, u_{11}\right)}{1 - u_{11}} \\ 1 - \lim\limits_{u_{22}\to 1^-} \dfrac{u_{22} - C_{h_{ij}}\left(u_{22}, u_{22}\right)}{1 - u_{22}} \\ 1 - \lim\limits_{\left(u_{11},u_{22}\right)\to\left(1^-,1^-\right)} \dfrac{u_{22} - C_{h_{ij}}\left(u_{22}, u_{11}\right)}{1 - u_{11}} \\ 1 - \lim\limits_{\left(u_{11},u_{22}\right)\to\left(1^-,1^-\right)} \dfrac{u_{11} - C_{h_{ij}}\left(u_{11}, u_{22}\right)}{1 - u_{22}} \end{pmatrix} \tag{15}$$

Proof. In matrix form, the extremogram and the crossed extremogram can be written, (see Muneya *et al.* [13]), for all $x, x+h \in \mathbb{R}^d$, such as,

$$\begin{pmatrix} \rho_{AA}^{11}(h) \\ \rho_{BB}^{22}(h) \\ \rho_{AB}^{12}(h) \\ \rho_{BA}^{21}(h) \end{pmatrix} = \lim_{z\to+\infty} \begin{pmatrix} P\left(Z_1\left(x+h\right) \in zA / Z_1\left(x\right) \in zA\right) \\ P\left(Z_2\left(x+h\right) \in zB / Z_2\left(x\right) \in zB\right) \\ P\left(Z_2\left(x+h\right) \in zB / Z_1\left(x\right) \in zA\right) \\ P\left(Z_1\left(x+h\right) \in zA / Z_2\left(x\right) \in zB\right) \end{pmatrix} \tag{16}$$

Consider $B = A = \left(a, +\infty\right)$. Z being stationary, let $u_{1i} = u_{1j} = u_{11}$. With these transformations the relation (13) is written in the form,

$$\rho_{AA}^{11}\left(h_{ij}\right) = \lim_{u_{11}\to 1^-} \frac{1 - 2u_{11} + C_{h_{ij}}\left(u_{11}, u_{11}\right)}{1 - u_{11}}.$$

Hence the first expression of (15).

In the same way, let us consider that $A = B = \left(b, +\infty\right)$ and $u_{2j} = u_{2i} = u_{22}$, the relation (13) is written in the form,

$$\rho_{BB}^{22}\left(h_{ij}\right)=\lim_{u_{22}\to 1^-}\frac{1-2u_{22}+C_{h_{ij}}\left(u_{22},u_{22}\right)}{1-u_{22}}.$$

Hence the second expression of (15).

Similarly for $A=\left(a,+\infty\right)$ and $B=\left(b,+\infty\right)$, let $u_{1i}=u_{1j}=u_{11}$ and $u_{2j}=u_{2i}=u_{22}$. The relation (13) is written in the form,

$$\rho_{AB}^{12}\left(h_{ij}\right)=\lim_{\left(u_{11},u_{22}\right)\to\left(1^-,1^-\right)}\frac{1-u_{22}-u_{11}+C_{h_{ij}}\left(u_{22},u_{11}\right)}{1-u_{11}}.$$

By swapping *A* and *B*, u_{11} and u_{22} will change location. So this new relationship is still written in the form,

$$\rho_{BA}^{21}\left(h_{ij}\right)=\lim_{\left(u_{22},u_{11}\right)\to\left(1^-,1^-\right)}\frac{1-u_{11}-u_{22}+C_{h_{ij}}\left(u_{11},u_{22}\right)}{1-u_{22}}.$$

Hence the third and fourth expressions of (15).

The following section is used to characterize the asymptotic dependence of extremes through the extremogram.

3.4. Asymptotic Dependence and Extremogram Model

Consider a random variable $T\left(x\right)$ of a spatial process $T=\left\{T\left(x\right),x\in\mathbb{R}^d\right\}$ of standardized marginalized $F_T\left(T\left(x\right)\right)$ (see [2] [14]). The following result gives another approach to quantify the asymptotic dependence of a random field at two sites x and $x+h$.

Proposition 7. Let $Z=\left\{Z\left(x\right),x\in\mathbb{R}^d\right\}$ be a spatial stationary process such that

$$Z\left(.\right)=-\left[\log\left(F_T\left(T\left(.\right)\right)\right)\right]^{-1}.$$

The marginal distribution of Z are Fréchet standard marginal. The extremogram of random field $Z(.)$ in two sites $x,x+h\in\mathbb{R}^d$ is define such as,

$$\rho_{AA}\left(h\right)=\mathcal{L}_h\left(u\right)u^{1-\frac{1}{\eta_h}};\tag{17}$$

where $\eta\left(h\right)\in\left(0,1\right]$ is the tail dependence coefficient, $A=\left(a;\infty\right)$ with $a\in\left(0,1\right]$ and $\mathcal{L}\left(.\right)$ a slowly varying function.

Before giving the proof of the above theorem, let's note that, even in a spatial study, there no loss of generality in dealing with Fréchet marginal, for any continuous function *f*, the transformation $f\left(Y_i\left(x\right)\right)=\frac{-1}{\log\left(Y_i\left(x\right)\right)}$ gives approximatively this distribution. Indeed, the parameters of the GEV in (2) as smooth function of the explanatory variables (longitude, altitude, elevation etc.) such as:

$$Y\left(x\right)=\mu\left(x\right)+\frac{\sigma\left(x\right)}{\xi\left(x\right)}\left[Z\left(x\right)^{\xi\left(x\right)}-1\right]\text{ where } Z(x)\sim\text{Unit-Fréchet}$$

for some partially correlation. That needs to model both spatial behaviour of marginal parameters and spatial joint dependence.

Proof. Considering $A=(a;\infty)$, $a\in(0;1]$, the extremogram is written

$$\rho_{AA}(h)=\lim_{z\to+\infty}\frac{P(Z(x+h)>az, Z(x)>az)}{P(Z(x)>az)}.$$

According to Ledford and Tawn [15], when z tends towards infinity,

$$P(Z_2>r, Z_1>r)\sim\mathcal{L}(r)r^{1-\frac{1}{\eta}}. \tag{18}$$

Using (18), for any spatial process Z at two sites x and $x+h$ when z tends towards infinity, we can write

$$P(Z(x+h)>az/Z(x)>az)\sim\mathcal{L}_h(az)(az)^{1-\frac{1}{\eta_h}}. \tag{19}$$

Thus, let $F(Z(x))$ be the distribution function of $Z(x)$ and $F(Z(x+h))$ the distribution function of $Z(x+h)$. According to the above, when z tends towards infinity, it follows that:

$$P(F(Z(x+h))>F(az)/F(Z(x))>F(az))\sim\mathcal{L}_h(az)(az)^{1-\frac{1}{\eta_h}}. \tag{20}$$

Considering $U=F(Z(x))$, $V=F(Z(x+h))$ and $u=F(az)$, it follows that:

$$P(V>u/U>u)\sim\mathcal{L}_h(u)u^{1-\frac{1}{\eta_h}}, \tag{21}$$

when z tends towards infinity.

Using (21) in the expression of the extremogram, it follows that:

$$\rho_{AA}(h)=\mathcal{L}_h(u)u^{1-\frac{1}{\eta_h}}$$

Hence (17) as disserted.

Ancona and Tawn [16] proposed a measure of extreme dependence called extreme variogram. This measure of dependence is expressed as a function of the dependence of tail by the relation:

$$\gamma_E(h)=2(1-\eta(h)). \tag{22}$$

Thus, the extremogram is modeled according to the extreme variogram by the following result.

Corollary 1. Let $\gamma_E(h)$ be the extreme variogram of two stationary random variables. The extremogram is linked to the extreme variogram by the relation:

$$\rho_{AA}(h)=\mathcal{L}_h(u)u^{-\frac{\gamma_E(h)}{2-\gamma_E(h)}}; \tag{23}$$

with $\gamma_E(h)\in[0;2)$.

Proof. From the relation (22), we can say that $\eta(h)=1-\frac{1}{2}\gamma_E(h)$. Using this relation in (17), it follows that:

$$\rho_{AA}(h)=\mathcal{L}_h(u)u^{1-\frac{2}{2-\gamma_E(h)}}.$$

Hence the expression,

$$\rho_{AA}(h) = \mathcal{L}_h(u) u^{\frac{-\gamma_E(h)}{2-\gamma_E(h)}}.$$

In the following, we estimate the extremogram using the relation (17). In this relationship, estimation of the extremogram requires estimation of the slowly varying function and the tail dependence coefficient. The following result gives the estimate of the extremogram.

Proposition 8. Consider two spatial random variables $Z(x)$ and $Z(x+h); x, x+h \in \mathbb{R}^d$ of respective marginal distribution function $F_Z(Z(x))$ and $F_Z(Z(x+h))$. Let $W(.)$ considering

$$W(h) = \min\left\{\frac{-1}{\log\left(F_Z(Z(x))\right)}; \frac{-1}{\log\left(F_Z(Z(x+h))\right)}\right\},$$

the estimated extremogram is written, for a fixed threshold u_h, in the form,

$$\hat{\rho}_{AA}(h) = \hat{c}_h(u) u^{1-\frac{1}{\hat{\eta}(h)}}. \tag{24}$$

where

$$\hat{c}_h(u) = \frac{n_{u_h}}{n} u_h^{\frac{1}{\hat{\eta}(h)}}; \quad \hat{\eta}(h) = \frac{1}{n_{u_h}} \sum_{k=1}^{n_{u_h}} \log\left\{\frac{\hat{w}_k(h) - u_h}{u_h}\right\},$$

with $\hat{w}_k(h), k = 1, \cdots, n_{u_h}$ are the observations $\hat{W}(h)$ exceeding the threshold u_h.

Proof. The extremogram is expressed by the relation,

$$\rho_{AA}(h) = \lim_{z\to+\infty} \frac{P\left(Z(x+h) > z, Z(x) > z\right)}{P\left(Z(x) > z\right)} = \lim_{z\to+\infty} \frac{P\left(W(h) > z\right)}{P\left(Z(x) > z\right)}.$$

Ledford ([17] [18]) proposed to consider $\mathcal{L}_h(u)$ as constant that is, $\mathcal{L}_h(u) = c_h$ for all values z exceeding the threshold u_h. Using the observations of the independent replications of the spatial process approximate independent observations on $\hat{W}(h)$ are obtained, where $\hat{W}(h)$ is the approximation to the variable $W(h)$. From model (20) and n independent observations, the log-likelihood is

$$l(c_h, \eta_h) = \left(n - n_{u_h}\right) \log\left(1 - \frac{c_h}{u_h^{1/\eta_h}}\right) + n_{u_h} \log\left(\frac{c_h}{\eta_h} - c_h\right) - \frac{1}{\eta_h} \sum_{i=1}^{n_{u_h}} \hat{w}_i(h),$$

where $\{\hat{w}_i(h)\}, i = 1, \cdots, n_{u_h}$ are the observations of $\hat{W}(h)$ above the threshold u_h. Using the maximum likelihood method, the estimate of c_h is written,

$$c_h = \frac{n_{u_h}}{n} u_h^{\frac{1}{\hat{\eta}(h)}},$$

and using the Hill estimator method, the estimate of η_h is written,

$$\hat{\eta}(h) = \frac{1}{n_{u_h}} \sum_{k=1}^{n_{u_h}} \log\left\{\frac{\hat{w}_k(h) - u_h}{u_h}\right\}.$$

where $\hat{w}_k(h), k=1,\cdots,n_{u_h}$ are the observations $\hat{W}(h)$ exceeding the threshold u_h. Hence the result,

$$\hat{\rho}_{AA}(h)=\hat{c}_h(u)u^{1-\frac{1}{\hat{\eta}(h)}}. \tag{25}$$

Therefore Formula (24) is verified.

4. Conclusions

In this study, we have been modeling some technical tools of spatial prediction within a copula-based space. Thus, the extremal coefficient and the extremogram have been expressed via the underlying copulas. These results are important insofar as we want to determine the inter-site distribution dependence of a definite area. The results of this paper make it possible to find a relation between the extremal coefficient and the extremogram using the copula function. These new models are very crucial since the copula is a parametrization of the number of variables that do not deal with the marginal distribution. Hence, they allow not only determining the distributional dependence of spatial or temporal extremes, but also, and above all, the conditional distributional dependence between these extremes in various observation sites.

In our further work, we will use these tools to determine the extreme distribution of metals in a mining context.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Matheron, G. (1976) Les concepts de base et l'évolution de la géostatistique minière. In: Guarascio, M., David, M. and Huijbregts, C., Eds., *Advanced Geostatistics in the Mining Industry*, NATO Advanced Study Institutes Series, Vol. 24, Springer, Dordrecht, 3-10. https://doi.org/10.1007/978-94-010-1470-0_1

[2] Frahm, G. (2006) On the Extremal Dependence Coefficient of Multivariate Distributions. *Statistics & Probability Letters*, **76**, 1470-1481. https://doi.org/10.1016/j.spl.2006.03.006

[3] Ouoba, F., Diakarya, B. and Talkibing, H.Y. (2019) Geostatistical Analysis with Copula-Based Models of Madograms, Correlograms and Variograms. *European Journal of Pure and Applied Mathematics*, **12**, 1052-1068. https://doi.org/10.29020/nybg.ejpam.v12i3.3389

[4] Barro, D., Koté, B. and Moussa, S. (2012) Spatial Stochastic Framework for Sampling Time Parametric Max-Stable Processes. *International Journal of Statistics and Probability*, **1**, 203. https://doi.org/10.5539/ijsp.v1n2p203

[5] Crisci, C. and Perera, G. (2022) Distribution extrême asymptotique pour les données non stationnaires et fortement dépendantes. *Advances in Pure Mathematics*, **12**, 479-489. https://doi.org/10.4236/apm.2022.128036

[6] Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J.L. (2004) Statistics of Extremes: Theory and Applications. John Wiley & Sons, Hoboken.

https://doi.org/10.1002/0470012382

[7] Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J.L. (2005) Statistics of Extremes: Theory and Application. Wiley, Chichester. https://doi.org/10.1002/0470012382

[8] Ledford, A.W. and Tawn, J.A. (1997) Modelling Dependence within Joint Tail Region. *Journal of the Royal Statistical Society*, *Series B*, **59**, 475-499. https://doi.org/10.1111/1467-9868.00080

[9] Cooley, D., Naveau, P. and Poncet, P. (2006) Variograms for Spatial Max-Stable Random Fields. In: Bertail, P., Soulier, P. and Doukhan, P., Eds., *Dependence in 8 Probability and Statistics*, Lecture Notes in Statistics, Vol. 187, Springer, New York, 373-390. https://doi.org/10.1007/0-387-36062-X_17

[10] Larsson, M. and Resnick, S.I. (2009) Long-Range Tail Dependence: EDM vs. Extremogram. w 911NF-07-1-0078 at Cornell University.

[11] Muneyamatsui and Mikosch, T. (2015) The Extremogram and the Cross-Extremogram for a Bivariate GARCH(1,1) Process.

[12] Davis, R.A. and Mikosch, T. (2009) The Extremogram: A Correlogram for Extreme Events. *Bernoulli*, **15**, 977-1009. https://doi.org/10.3150/09-BEJ213

[13] Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J.L. (2005) Statistics of Extremes: Theory and Application. Wiley, Chichester. https://doi.org/10.1002/0470012382

[14] Cho, Y.B., Davis, R.A. and Ghosh, S. (2016) Asymptotic Properties of the Empirical Spatial Extremogram. *Scandinavian Journal of Statistics*, **43**, 757-773.

[15] Bondár, I., McLaughlin, K. and Israelsson, H. (2005) Improved Event Location Uncertainty Estimates. 27*th Seismic Research Review*: *Ground-Based Nuclear Explosion Monitoring Technologies*, Rancho Mirage, 20-22 September 2005, 299-307.

[16] Davison, A., Huser, R. and Thibaud, E. (2013) Geostatistics of Dependence and Asymptotically Independent Extremes. *Mathematical Geosciences*, **45**, 511-529. https://doi.org/10.1007/s11004-013-9469-y

[17] Ledford, A.W. and Tawn, J.A. (1997) Modelling Dependence within Joint Tail Region. *Journal of the Royal Statistical Society*, *Series B*, **59**, 475-499. https://doi.org/10.1111/1467-9868.00080

[18] Ledford, A.W. and Tawn, J.A. (1996) Statistics for Near Independence in Multivariate Extreme Values. *Biometrika*, **83**, 169-187. https://doi.org/10.1093/biomet/83.1.169