Copula-based geostatistical modeling of continuous and discrete data including covariates

Hannes Kazianka · Jürgen Pilz

Received: date / Accepted: date

Abstract It is still common in geostatistics to use the variogram to describe the spatial dependence structure and to use kriging as the spatial prediction methodology. Both methods are sensitive to outlying observations and are strongly influenced by the marginal distribution of the underlying random field. Hence, they lead to unreliable results when applied to extreme value or multimodal data. As an alternative to traditional spatial modeling and interpolation we consider the use of copula functions. We show how trend surface models can be incorporated into copula-based geostatistical models. Moreover, the case where the marginal distributions of the random field are discrete is included in the existing methodology. We propose three different copula-based spatial interpolation methods. By exploiting the relationship between bivariate copulas and indicator covariances, we present indicator kriging and disjunctive kriging. As a second method we present simple kriging of the rank-transformed data. The third method is a plug-in prediction and generalizes the frequently applied trans-Gaussian kriging. Finally, we report on the results obtained for the so-called Joker data set from the spatial interpolation comparison SIC2004.

Keywords Copulas · Trend · Discrete Data

1 Introduction

Copulas describe the dependence between random variables independently from their marginal distributions. They are widely used in financial and actuarial statistics, however, they are just beginning to become popular in geostatistics. Spatial dependence is traditionally described using the variogram which is strongly influenced by the univariate distribution of the random field. Extreme outlying observations adversely affect the empirical and theoretical variogram estimates. Moreover, spatial modeling often relies on the Gaussian assumption which is hardly fulfilled for environmental processes. A number of methods have been suggested to deal with non-Gaussianity, such as log-normal kriging (Journel and Huijbregts 1978), disjunctive kriging (Rivoirard 1994), trans-Gaussian kriging (De Oliveira et al. 1997; Pilz and Spöck 2008) and generalized linear geostatistical models (Diggle et al. 1998). Bárدوssy (2006) in his pioneering work proposed the use of copulas to describe the spatial variability. Kazianka and Pilz (2008), Pilz et al. (2008) and Bárđossy and Li (2008) used this copula-based modeling approach for spatial interpolation of continuous random fields. In the following we adopt this methodology and propose two extensions. The first extension relaxes the assumption of a constant trend and gives the opportunity to consider covariates. The second makes it possible to model random fields with discrete marginal distributions. We present three methods for estimating the values of the random field at unknown locations. The first method we suggest is indicator and disjunctive kriging. The second method is rank-order kriging, originally proposed by Journel and Deutsch (1996), where we calculate the covariance function through its relationship to the Spearman rank correlation. The third method is a plug-in...
predictor and can be used if all multivariate distributions of the random field are modeled using the copula.

The paper is organized as follows. Section 2 reviews the basic properties of copulas, while Section 3 describes the spatial copula methodology, presents a method to build radially asymmetric spatial copulas and suggest how to account for a spatial trend and covariates. Section 4 is concerned with the estimation of the model parameters for both continuous and discrete margins and proposes a method to select a suitable spatial copula for a given data set. The spatial interpolation techniques using copulas are presented in Section 5 and in Section 6 they are used to analyze the Joker data set from the spatial interpolation comparison SIC2004 (Dubois 2005). Section 7 is devoted to conclusions.

2 Copulas

The word “copula” was first used by Sklar (1958) to describe distribution functions on the n-dimensional unit cube, $I^n$, that link multivariate distributions to their one-dimensional margins. To be precise, an n-dimensional copula (often abbreviated n-copula) is an n-dimensional real function $C : I^n \rightarrow I$ which satisfies the following properties:

1. For every $u \in I^n$
   \[ C(u) = 0 \text{ if at least one coordinate of } u \text{ equals } 0, \]
   \[ C(u) = u_k \text{ if all coordinates of } u \text{ are } 1 \text{ except } u_k. \]

2. For every $a, b \in I^n$ with $a \leq b$
   \[ V_C([a, b]) \geq 0, \]
   where $V_C((a, b)) = \Delta^k_0 \Delta^{k-1}_0 \ldots \Delta^k_0 C(u)$ is the n-th order difference of $C$ on $[a, b]$. Thereby, we define $\Delta^k_a C(u) = C(u_1, \ldots, u_{k-1}, b_k, u_{k+1}, \ldots, u_n) - C(u_1, \ldots, u_{k-1}, a_k, u_{k+1}, \ldots, u_n)$ as a first order difference.

From this definition it is clear that a copula is a distribution function on the n-dimensional unit cube with uniformly distributed margins and induces the probability measure $V_C([0, u]) = C(u)$. The most important theoretical result about copulas is known as Sklar’s Theorem and expresses the ability of copulas to describe the dependence between random variables without information about their marginal distributions: If $H$ denotes an n-dimensional distribution function with margins $F_1, \ldots, F_n$, then there exists an n-copula $C$ such that for all $x \in \mathbb{R}^n$,

\[ H(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)). \]  

If $F_1, \ldots, F_n$ are all continuous, then $C$ is unique. Conversely, if $C$ is an n-copula and $F_1, \ldots, F_n$ are distribution functions, then the function $H$ is an n-dimensional distribution function with margins $F_1, \ldots, F_n$. Moreover, if $F_1^{-1}, \ldots, F_n^{-1}$ are the inverse distribution functions of $F_1, \ldots, F_n$, we get

\[ C(u_1, \ldots, u_n) = H(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)). \]  

(2)

Applying the latter equation we obtain that if $C$ is an absolutely continuous copula, its density can be written as

\[ c(u_1, \ldots, u_n) = \frac{h(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n))}{\prod_{i=1}^n f_i(F_i^{-1}(u_i))}, \]  

(3)

where $h$ denotes the density of $H$ and the $f_i$ denote the densities of $F_i$.

Furthermore, Eq. (2) provides a simple way for the construction of copulas. If, for example, $H = \Phi_0, \Sigma$, the distribution function of the multivariate Gaussian distribution with mean 0 and correlation matrix $\Sigma$, and $F_1 = \cdots = F_n = \Phi$, the distribution function of the standard Gaussian distribution, the resulting copula is called the Gaussian copula:

\[ C^G(u_1, \ldots, u_n) = \Phi_0, \Sigma(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n)). \]  

(4)

One of the advantages of working with copulas is that they are invariant under strictly increasing transformations of the random variables. Therefore, typical data transformation methods, such as taking the logarithm or performing a uniform scores transformation, have no impact on the copula. Therefore, if $\alpha_1, \ldots, \alpha_n$ are arbitrary strictly increasing transformations of the random variables $X_1, \ldots, X_n$, then $C_{\alpha_1(X_1), \ldots, \alpha_n(X_n)}(u_1, \ldots, u_n) = C_{X_1, \ldots, X_n}(u_1, \ldots, u_n)$, where $C_{\alpha_1(X_1), \ldots, \alpha_n(X_n)}$ and $C_{X_1, \ldots, X_n}$ denote the copula of the transformed and the original variables, respectively.

In the case where $\alpha_1, \ldots, \alpha_n$ are strictly decreasing transformations the copula is not invariant. For the marginal distribution of $\alpha_i(X_i)$ we get

\[ F_{\alpha_i}(x_i) = P(\alpha_i(X_i) \leq x_i) = P(X_i \geq \alpha_i^{-1}(x_i)) = 1 - F_i(\alpha_i^{-1}(x_i)). \]

By using $C_{\alpha_1(X_1), \ldots, \alpha_n(X_n)}(F_{\alpha_1}(x_1), \ldots, F_{\alpha_n}(x_n)) = P(X_1 \geq \alpha^{-1}(x_1), \ldots, X_n \geq \alpha^{-1}(x_n))$, $u_i = F_{\alpha_i}(x_i)$ and Poincaré’s formula we find out that

\[ C_{\alpha_1(X_1), \ldots, \alpha_n(X_n)}(u_1, \ldots, u_n) = 1 + \sum_{i=1}^n (-1)^i S_i, \]  

(5)
where \( S_1 = \sum_{j=1}^{n} (1 - u_j) \) and for all \( i = 2, \ldots, n \)
\[
S_i = \sum_{j_i < \cdots < j_{i-1}} C_{X_{j_1} \cdots X_{j_i}} (1 - u_{j_1}, \ldots, 1 - u_{j_i}) .
\]

From Eq. (1) we immediately get that the random variables are independent if and only if their copula is the so-called product copula, \( \Pi^n (u) = \prod_{i=1}^{n} u_i \). Likewise, one obtains that each of the random variables is a strictly increasing function of any of the others if and only if their copula equals the Frechet upper bound, \( M^n (u) = \min (u_1, \ldots, u_n) \).

Bivariate copulas are directly linked to several scale free measures of association, such as Spearman’s rho or Kendall’s tau. For our purposes we need the relation to the Spearman rank correlation between two random variables \( X_1 \) and \( X_2 \) with copula \( C \) which can be calculated as
\[
\rho_{X_1, X_2} = 12 \iint_{\mathbb{T}^2} u_1 u_2 dC(u_1, u_2) - 3 \\
= 12 \iint_{\mathbb{T}^2} C(u_1, u_2) du_1 du_2 - 3 .
\]

For a thorough introduction to copulas the reader is referred to Nelsen (2006).

### 3 Spatial modeling using copulas

Although copulas are widely used for describing the dependence between random variables, for example in financial statistics, there are only a few papers about incorporating copulas into the geostatistical framework so far. In the following assume that we have a continuous, second-order stationary random field \( \{ Z(x) | x \in S \} \), where \( S \subseteq \mathbb{R}^2 \) is the area of interest. Bondar et al. (2005) use copulas for describing the dependence between \( X = \| h \| \) and \( Y = \frac{Z(x) - Z(x+h)}{\sqrt{\| h \|}} \), where \( h \) denotes a vector separating two points. Performing a median regression of \( Y \) on \( X \) yields a robust semivariogram estimator without using any frequently applied variogram model. Archimedeian copulas are used in Porcu et al. (2004) for building nonseparable anisotropic classes of spatio-temporal variograms.

#### 3.1 Describing the random field using copulas

Bárdossy (2006) presented a different method for spatial modeling using copulas that aims at describing all multivariate distributions of the random field using copulas. Let \( F_Z \) denote the univariate distribution of the random process. The stationarity of the field ensures that \( F_Z \) is the same for each location \( x \in S \). With the help of Sklar’s Theorem we are able to model all multivariate distributions of the field by setting \( F_Z = F_1 = \cdots = F_n \) in Eq. (1). For instance, the relation between two locations separated by the vector \( h \) is characterized by the bivariate distribution
\[
P(Z(x) \leq z_1, Z(x + h) \leq z_2) = C_h(F_Z(z_1), F_Z(z_2))
\]
whose dependence structure is described by the spatial copula \( C_h \). The copula becomes a function of the separating vector \( h \) or of the separating distance \( h := \| h \| \) in the case of an isotropic field. Hence, spatial copulas describe spatial dependence over the whole range of quantiles for a given separating vector \( h \) and not only the mean dependence as the variogram does.

Not all continuous copulas are suitable for geostatistical modeling. A natural assumption for a spatial copula is symmetry, implying that, for example, the dependence between locations \( x_1 \) and \( x_2 \) is the same as the dependence between \( x_2 \) and \( x_1 \). In general it means that \( C_h(u_1, \ldots, u_n) = C_h(u_{\sigma(1)}, \ldots, u_{\sigma(n)}) \) for an arbitrary permutation \( \sigma \) and \( n \geq 2 \). Another requirement is easy parameterization of the correlation structure to make it a function of \( h \). Moreover, we want to add the following two restrictions:

- as \( \| h \| \to \infty \) we require \( C_h(u) \to \Pi^n \) which implies that far distant observations are nearly independent, and
- as \( \| h \| \to 0 \) we require in the absence of a nugget effect that \( C_h(u) \to M^n \) ensuring that observations that are very close to each other have a strong dependence.

These constraints and the fact that constructing multivariate copulas is far more difficult than building bivariate copulas are the reasons why many of the well-known copula families, such as the Farlie-Gumbel-Morgenstern family, are not useful in our approach.

The most important class of random fields are the Gaussian random fields in which all multivariate distributions follow a Gaussian distribution. The copula-based model includes the Gaussian random field as a special case where the copula is equal to the Gaussian copula \( C^G_{\Sigma} \), defined in Eq. (4), and the marginal distribution is \( F_Z = \Phi_{m, \sigma^2} \), where \( m \) and \( \sigma^2 \) denote mean and variance, respectively. The aforementioned restrictions for spatial copulas are all satisfied and it is even possible to describe negative spatial dependence. The Gaussian copula becomes a function of \( h \) by assuming that its correlation function follows one of the well-known parametric geostatistical models, e.g. the Matern model.

The Gaussian copula, besides having zero tail dependence, does not only express a symmetric but also a radially symmetric dependence where the copula density satisfies \( c(u_1, \ldots, u_n) = c(1 - u_1, \ldots, 1 - u_n) \). In
this case high and low quantiles of the distribution have equal dependence properties. To allow for more flexibility it is therefore often desirable to use radially asymmetric copula families. Bárdossy (2006) introduced a non-Gaussian copula family which is constructed from a multivariate non-central $\chi^2$-distribution. Squaring the entries of a Gaussian random vector $Y \sim \mathcal{N}(m, \Sigma)$, where $m = (m, \ldots, m)$ and $\Sigma$ denote the mean vector and the correlation matrix respectively, leads to a multivariate distribution with margins having a non-central $\chi^2$-distribution with 1 degree of freedom and non-centrality parameter $\lambda = m^2$. The density $d$ and the distribution function $D$ can be calculated as

$$d(z_1, \ldots, z_n) = \frac{\sum_{i=0}^{n-1} \phi_{m, \Sigma}(\varepsilon_i)}{2^n \sqrt{\prod_{i=1}^n z_i}},$$

$$D(z_1, \ldots, z_n) = \frac{\sum_{i=0}^{n-1} (-1)^{n+1} \phi_{m, \Sigma}(\varepsilon_i)}{2^n \sqrt{\prod_{i=1}^n z_i}},$$

where $\varepsilon_i = \left((-1)^i \sqrt{z_1}, \ldots, (-1)^i \sqrt{z_n}\right)$, $i_j \in \{0, 1\}$, $i = \sum_{j=1}^n i_j 2^{j-1}$ and $\phi_{m, \Sigma}$ denotes the Gaussian density function. Using Eqs. (2) and (3) the copula and its density can be evaluated. A generalization of this copula family is presented by Bárdossy and Li (2008).

The obvious choice for a non-Gaussian copula would be a copula generated from an elliptically contoured distribution, $EC(\mu, \Sigma, \varphi)$, as was advertised in Bárdossy and Li (2008). An example would be the t-copula which is generated out of the multivariate generalization of Student’s t-distribution. However, the elliptical copulas are not a suitable candidate for spatial modeling since there is no simple way for making them a function of $h$. Parameterizing $\Sigma$ similar to the Gaussian case fails because for $\Sigma$ being diagonal the margins are uncorrelated but not independent (Fang and Zhang 1990) which violates the restrictions for spatial copulas.

3.2 Radially asymmetric theoretical spatial copula

The non-central $\chi^2$-copula is derived by non-monotonic transformations of the marginals of a Gaussian distribution. Here we use a different construction principle for radially asymmetric copulas which was proved by Liescher (2008): Let $C_1, \ldots, C_k : \Gamma^n \rightarrow \Gamma$ be copulas and $g_{ji} : \Gamma \rightarrow \Gamma$ for $j = 1, \ldots, k$ and $i = 1, \ldots, n$ be functions that are continuous, strictly increasing (or identically equal to 1) so that $\prod_{j=1}^k g_{ji}(u) = u$. Then

$$C(u_1, \ldots, u_n) = \prod_{j=1}^k C_j(g_{j1}(u_1), \ldots, g_{jn}(u_n))$$

is also a copula.

A special case of the above theorem would be $k = 2$, $C_1 = C_{\Sigma_1}^G$, $C_2 = C_{\Sigma_2}^G$, $g_{i1}(u) = u^\lambda$, $g_{i2}(u) = u^1-\lambda$ and $0 \leq \lambda \leq 0.5$. The copula density can be written as

$$c(u_1, \ldots, u_n) = \sum_{i=0}^{n} \sum_{(u_{j1}, \ldots, u_{jn}) \in \psi((u_1, \ldots, u_n))} \frac{\partial^n C_{\Sigma_1}^G(u_1^\lambda, \ldots, u_n^\lambda)}{\partial u_{j1} \ldots \partial u_{jn}} \times \frac{\partial^n C_{\Sigma_2}^G(u_1^{1-\lambda}, \ldots, u_n^{1-\lambda})}{\partial u_{j1+1} \ldots \partial u_{jn}},$$

where $\psi$ denotes the power set. All the properties of a spatial copula are inherited from the Gaussian copula which arises if $\lambda = 0$. Furthermore, for $\lambda > 0$ we get a non-Gaussian radially asymmetric copula. Figure 1 shows the bivariate copula density for $\lambda = 0.5$, $\rho_1 = 0.1$ and $\rho_2 = 0.9$, where $\rho_1$ and $\rho_2$ are the off-diagonal elements of $\Sigma_1$ and $\Sigma_2$ respectively. Even more difficult dependence structures could be modeled for $k > 2$, however, this would make the evaluation of the copula density very difficult. If $k = 2$, one has to sum already over $2^n$ terms which is also the case for the non-central $\chi^2$-copula density.

3.3 Incorporating spatial trend

Up to now copulas are used for flexible modeling the multivariate distribution of a random field with univariate marginal distribution $F_Z$. The assumption that all univariate margins are equal implies the absence of a spatial trend which is unrealistic for most applications. Moreover, it is often desired to include additional predictor variables into the analysis. As a natural extension to incorporate trend models we propose to further parameterize $F_Z$. If $F_Z$ belongs to the class of exponential
dispersion models (Song 2007), ED \( (\mu, \sigma) \), the mean \( \mu \) appears explicitly inside the analytical expression of the density \( f_Z \) and we can set
\[
g(\mu_i) = \eta(y_i) = y_i^T \beta, \quad i = 1, \ldots, n
\]
where \( g \) is the link function, \( \eta(y_i) \) is the response surface, \( y_i \) is the vector of covariates corresponding to location \( x_i \), and \( \beta \) is the vector of regression parameters. If the margins are Gaussian, the link function would be the identity while for gamma or Poisson distributed margins the log-link is appropriate. If the margins are Gaussian, the link function would be parameterized in the same way. However, the mean of the distribution could include other parameters as well or may not exist at all.

Note, that in this way we have created an alternative to the well-known generalized linear geostatistical model (GLGM) introduced by Diggle et al. (1998). In their approach, which is an extension of generalized linear mixed models, it is
\[
g(\mu_i) = y_i^T \beta + S(x_i), \quad i = 1, \ldots, n
\]
where \( S \) is a stationary Gaussian process with mean zero, variance \( \sigma^2 \) and correlation function \( \rho \). The main difference between GLGM and our approach is that in GLGM only \( \mu_i \) is modeled while we build a complete multivariate distribution for the response variable \( Z(x_i) \). Poisson-GLGMs have the advantage that accounting for overdispersion is done by adding a spatially unstructured random effect to Eq. (8). In our model one has to completely change the marginal distribution e.g. from Poisson to a zero-inflated generalized Poisson distribution (Famoye and Singh 2003).

4 Parameter estimation for continuous and discrete margins

In the spatial copula model we have mainly three different types of parameters. We have parameters \( \Theta \) defining the correlation structure of the copula, copula parameters \( \lambda \) and parameters \( \eta \), including the trend parameters \( \beta \), for the family of marginal distributions \( F_Z \). Therefore, we write the spatial copula \( C_{\theta, \lambda} \) and the margin \( F_Z \equiv F_\theta \eta \) as a function of their parameters. Báródy and Li (2008) presented a maximum likelihood approach that estimates \( \lambda \) and \( \theta \) in the case of continuous marginal distributions. The parameters \( \eta \) for the margins are not considered in their approach. In the following we propose a joint estimation of all model parameters using maximum likelihood for both continuous and discrete margins. The subsequent analysis can be extended to account for geometric anisotropy by a coordinate transformation and the introduction of appropriate anisotropy parameters.

4.1 Inference for continuous margins

In the case of continuous margins the multivariate distribution is absolutely continuous and possesses a density with respect to the Lebesgue measure. Hence, the likelihood given a realization \( D = \{ z(x_1), \ldots, z(x_n) \} \) can be evaluated as
\[
l(\Theta; D) = \prod_{i=1}^n f_\theta \eta \left( z(x_i) \right),
\]
where \( \Theta = (\theta, \lambda, \eta) \) is the parameter vector, \( c_{\theta, \lambda} \) is the copula density and \( f_\eta \) denotes the marginal density. Hence, the maximum likelihood estimate (MLE) is 
\[
\hat{\Theta} = \arg \max_{\Theta} l(\Theta; D).
\]

Optimal parameters are typically found by numerical optimization routines. If the number of parameters is large, we suggest a profile-likelihood approach. For the Gaussian copula \( \lambda \) does not occur and \( \theta \) are the parameters of the correlation model. Note, that we do not need to estimate an overall sill for this correlation function since it is equal to 1. The reason is that the overall variance of the random field is a property of the marginal distribution and the copula describes the dependence structure without information about the margins. In general, there is no parameter in \( \eta \) that is equivalent to the sill. If the margins are from the class of continuous exponential dispersion models, \( ED \ (\mu, \sigma^2) \), and there is no geometric anisotropy, a maximization by parts algorithm is proposed that converges to the maximum likelihood estimate under the information dominance condition and gives more reliable estimates especially for small sample sizes (Song 2007; Song et al. 2005): Partition the log-likelihood into two additive parts using
\[
\ell(\Theta; D) = \ell_w(\eta; D) + \ell_c(\Theta; D),
\]
\[
\ell_w(\eta; D) = \sum_{i=1}^n \log f_\eta \left( z(x_i) \right),
\]
\[
\ell_c(\Theta; D) = \frac{1}{2} u^T (I^n - \Sigma^{-1}) u - \frac{1}{2} \log \text{det} \Sigma,
\]
with \( u_\eta = \left( \Phi^{-1}(F_\eta \left(z(x_1)\right)), \ldots, \Phi^{-1}(F_\eta \left(z(x_n)\right)) \right) \) and \( \eta = (\beta, \sigma^2) \). The algorithm proceeds as follows:

- Step 1: To obtain \( \theta^1 \), solve \( \frac{\partial \ell_c(\theta; D, \eta)}{\partial \eta} = 0 \) for \( \eta^1 \) and \( \frac{\partial \ell_c(\theta, \lambda; D, \eta^1)}{\partial \lambda} = 0 \) for \( \lambda^1 \) and \( \theta^1 \).
The derivatives of $\ell$ in our geostatistical context are

\[
\begin{align*}
\frac{\partial \ell_c(\Theta; D)}{\partial \beta} &= \frac{\partial u_n^T}{\partial \beta} \left( I^n - \Sigma^{-1} \right) u_n, \\
\frac{\partial \ell_c(\Theta; D)}{\partial \sigma^2} &= \frac{\partial u_n^T}{\partial \sigma^2} \left( I^n - \Sigma^{-1} \right) u_n, \\
\frac{\partial \ell_c(\Theta; D)}{\partial \theta_j} &= \frac{1}{2} \text{tr} \left[ \frac{\partial \Sigma}{\partial \theta_j} \right] \left( \Sigma^{-1} u_n u_n^T \Sigma^{-1} - \Sigma^{-1} \right),
\end{align*}
\]

where the $\theta_j$ usually stand for the nugget, range, and smoothness parameter of the specific correlation function model.

For certain copula families such as the non-central $\chi^2$-copula or the multiplied Gaussian copula presented in Sect. 3.2 it happens that the evaluation of the density is infeasible in higher dimensions. Here we proceed to perform composite maximum likelihood estimation only with the bivariate copula densities. Under the assumption that different pairs of observations are treated as independent we have to maximize the pairwise likelihood

\[
l(\Theta; D) = \prod_{i,j \in \{1, \ldots, n\}} c_{\theta, \lambda}(F_{\eta}(z(x_i)), F_{\eta}(z(x_j))) \times \prod_{k \in \{i,j\}} f_{\eta}(z(x_k)).
\]

Composite likelihood methods are studied by Zhao and Joe (2005) and we found that they work well as long as there is no intention to estimate geometric anisotropy.

4.2 Inference for discrete marginals

If the univariate margins are all discrete, the density with respect to the counting measure can be written as

\[
l(\Theta; D) = \sum_{j_1=0}^{1} \cdots \sum_{j_n=0}^{1} (-1)^l C_{\theta, \lambda}(u_{1j_1}, \ldots, u_{nj_n}),
\]

where $u_{00} = F_{\eta}(z(x_i))$, $u_{i1} = \lim_{z(x) \to z(x_i)-} F_{\eta}(z(x)) = F_{\eta}(z(x_i))$ denotes the left-sided limit of $F_{\eta}$ at $z(x_i)$ for all $i = 1, \ldots, n$ and $l = \sum_{k=1}^{n} j_k$. In Song (2007) a Gauss-Newton type algorithm is proposed to obtain the maximum likelihood estimate. However, it is clear that this only works for a small number of locations since we need to sum over $2^n$ terms to evaluate the likelihood.

Another problem arises when the copula itself is difficult to compute in higher dimensions. For example, the evaluation of the Gaussian copula requires to calculate the multivariate Gaussian distribution function which is numerically unstable when $n$ is large.

To circumvent these problems we suggest a composite likelihood approach similar to Eq. (9). In this case the estimates $\hat{\Theta}$ maximize the pairwise likelihood

\[
l(\hat{\Theta}; D) = \prod_{i,j \in \{1, \ldots, n\}} \sum_{j_1=0}^{1} \sum_{j_2=0}^{1} (-1)^l C_{\theta, \lambda}(u_{ij_1}, u_{jj_2}).
\]

Another possibility for approximate inference is to use the so-called generalized quantile transform. If $V$ is uniformly distributed on the unit interval, it follows that $G(V, z(x)) = (1 - V) F_{\eta}(z(x)) - V F_{\eta}(z(x))$ is also uniform and $F_{\eta}^{-1}(G(V, z(x)))$ has the same distribution as $z(x)$ (Pflug and Römisch 2007). By setting $V = \frac{1}{n}$ the likelihood is approximated by

\[
l(\hat{\Theta}; D) \approx c_{\theta, \lambda}(u_1, \ldots, u_n) \prod_{i=1}^{n} P(Z(x) = z(x_i)),
\]

(11)

where $u_i = \frac{u_{i1} + u_{i2}}{2}$. In the case of independent observations the approximation leads to the exact MLE. The approximation works well as long as the marginal variance is not too small, similar to a Gaussian approximation. However, it shows significantly better performance than a Gaussian approximation for skewed distributions like the geometric distribution.

4.3 Goodness-of-fit testing for spatial copulas

For selecting a spatial copula model that suits the given data we have to perform a goodness-of-fit test. We use a blanket test that was recently presented (Genest and Remillard 2008) and validated (Genest et al. 2008) for arbitrary copulas. To use it in the geostatistical framework we apply it to all the different lag classes $h_1, \ldots, h_r$, in the spirit of Bădăossy (2006). Although the test is originally designed for $n$-copulas, we recommend to work only with bivariate copulas for simplicity. The test is based on a parametric bootstrapping procedure and makes use of the Kolmogorov-Smirnov statistic, $T_n$, or the Cramér-von Mises statistic, $S_n$:

\[
S_n = \int_{[0,1]^2} C_n(u)^2 dC_n(u), \quad T_n = \sup_{u \in [0,1]^2} |C_n(u)|,
\]

where $C_n = \sqrt{n} \left( C_n - C_\Phi \right)$, $C_n$ is the empirical copula calculated using the $n$ data points and $C_\Phi$ is the estimation under the null hypothesis. The steps of the algorithm are as follows:
1. For each of the lags $h_1, \ldots, h_r$, compute the empirical copula $C_{n_{h_1}}, \ldots, C_{n_{h_r}}$.
2. Estimate the parameters of the theoretical copula using the maximum likelihood approach outlined in Sect. 4.1-4.2. Denote the estimated parameters by $\hat{\Theta}$. For every lag class there is a corresponding theoretical copula $C_{\hat{\Theta}_1}, \ldots, C_{\hat{\Theta}_h}$.
3. Calculate the Cramér-von Mises or the Kolmogorov-Smirnov statistic for every lag class, $T_{n_{h_1}}, \ldots, T_{n_{h_r}}$ or $S_{h_1}, \ldots, S_{h_r}$.
4. For a large integer $N$, repeat the following steps for every $k \in \{1, \ldots, N\}$
   a) Simulate a random field whose copula is exactly the estimated theoretical copula from step 2.
   b) Compute the empirical copula for every lag class, $C_{n_{h_1}}, \ldots, C_{n_{h_r}}$.
   c) Estimate the parameters of the theoretical copula of the simulated field and denote them by $\hat{\Theta}_k$. For every lag class there is a corresponding theoretical copula, $C_{\hat{\Theta}_1}, \ldots, C_{\hat{\Theta}_h}$.
   d) Evaluate $T_{n_{h_1}}, \ldots, T_{n_{h_r}}$ or $S_{h_1}, \ldots, S_{h_r}$.
5. In the case of the Kolmogorov-Smirnov statistic, for every lag class $h_1, \ldots, h_r$, an approximate p-value for the test is given by
   \[ p_{h_j} = \frac{1}{N} \sum_{k=1}^{N} I( T_{n_{h_j}}^{h_j} > T_{n_{h_j}}^{h_j} ) \]  
   where $I$ is an indicator function and $j = 1, \ldots, r$. To get an approximate p-value in the case of the Cramér-von Mises statistic, just replace the corresponding expressions in Eq. (12).

In the case where the spatial copula is constructed from a multivariate distribution, simulation of a random field with a predefined copula means simulating from this distribution and subsequently transforming the margins according to Eq. (2).

5 Spatial interpolation using copulas

After having estimated the parameters $\Theta$ of the spatial copula model we are interested in predicting the values of the random field at unknown locations $x_0$. In the following we propose three different methods for performing spatial interpolation using copulas.

5.1 Indicator kriging and disjunctive kriging

Indicator kriging is used to estimate the conditional distribution of the random field given the data. This is done by cokriging of indicator variables $I(Z(x) \leq z_j)$, where the $z_j$ are certain thresholds e.g. quantiles. Bárdozy (2006) showed that bivariate copulas are related to indicator variograms and cross-variograms. Therefore, it is simple to derive the following relationships between bivariate copulas and indicator covariances and cross-covariances:

\[ \gamma_{z_j}(h) = C_{\hat{\Theta}_j}(F_{\hat{\eta}}(z_j), F_{\hat{\eta}}(z_j) - F_{\hat{\eta}}(z_j))^2, \] (13)

\[ \gamma_{z_k, z_k}(h) = C_{\hat{\Theta}_j}(F_{\hat{\eta}}(z_j), F_{\hat{\eta}}(z_k)) - F_{\hat{\eta}}(z_j) F_{\hat{\eta}}(z_k). \]

Plugging these equations in the cokriging procedure, we arrive at an indicator kriging that is based on the spatial copula model. The fact that only bivariate copulas are needed makes it possible to use one of the numerous flexible copulas that do not have multivariate extensions or that have too few parameters for using them in a multivariate approach, e.g. the Gumbel-Hougaard extreme value copulas.

Note that Eq. (13) and, hence, the copula-based indicator kriging approach is valid for both continuous and discrete margins. Moreover, if the random field is discretized and takes only a finite number of values, say $1$ to $m$, every function $f(Z(x))$ can be written as a weighted sum of indicator variables,

\[ f(Z(x)) = f_1 I(Z(x) \leq 1) + \cdots + f_m I(Z(x) \leq m). \]

The disjunctive kriging (DK) estimator can be calculated by cokriging (CK) of the indicators

\[ [f(Z(x))]^{DK} = f_1 [f(Z(x) \leq 1)]^{CK} + \cdots + f_m [f(Z(x) \leq m)]^{CK}. \]

Again, the relationships (13) are used in the cokriging system. Rivoirard (1994) concluded that "in the same way that kriging is based on the variogram, so disjunctive kriging is based on the bivariate distributions". In our case the bivariate distribution of the random field is defined in terms of the bivariate copula and so is disjunctive kriging.

5.2 Rank-order kriging

Assume we have an isotropic random field with estimated univariate distribution $F_{\hat{\eta}}$ and the bivariate distributions can be described by the copula $C_{\hat{\Theta}_j}$. Moreover, we have a realization $\{ z(x_i) \mid x_i \in S \}$ of the random field and we want to predict the values at locations $x_{n+1}, \ldots, x_{n+m}$. Applying Eq. (6) we can calculate the Spearman rank correlation curve $\rho$ as a function of $h$ which is exactly the correlation function for the rank-transformed variable $V(x) = F_{\hat{\eta}}(Z(x))$. Since $V(x)$ is a uniform distribution on $[0,1]$, $\frac{\rho}{\sqrt{2}}$ gives the corresponding covariance function. Journel and Deutsch (1996) proposed to apply a simple kriging of ranks to
account for non-Gaussianity. Thereby, the linear predictor at the unknown locations \( x_j \) is given by

\[
\hat{\bar{V}}(x_j) = \sum_{i=1}^{n} \lambda_i v(x_i) + \frac{1}{2} \left( 1 - \sum_{i=1}^{n} \lambda_i \right),
\]

where \( j = n + 1, \ldots, n + m \). Since back-transforming \( \hat{\bar{V}}(x_j) \) using \( F^{-1}_\eta \) would lead to a biased estimate for \( Z(x_j) \), a bias correction is introduced

\[
\tilde{Z}(x_j) = F^{-1}_\eta \left( \hat{\bar{V}}(x_j) \right) + \lambda(x_j)
\times \left[ F^{-1}_\eta \left( L \left( \hat{\bar{V}}(x_j) \right) \right) - F^{-1}_\eta \left( \hat{\bar{V}}(x_j) \right) \right],
\]

where \( L \) is the distribution function of all kriged values \( \hat{\bar{V}}(x_j) \) and \( \lambda(x_j) = \left( \frac{\sigma^2_K(x_j)}{\sigma^2_{\text{max}}} \right)^{1/2} \) with \( \sigma^2_K \) being the kriging variance, \( \sigma^2_{\text{max}} \) being the maximal kriging variance of all estimations and \( \omega > 0 \) a correction level parameter.

Although this method reproduces the original distribution of the data and \( \tilde{Z} \) is an unbiased estimate, the covariance structures of \( V(x) \) and \( \tilde{Z}(x) \) are not reproduced. Another disadvantage of rank-order kriging is that there is no guarantee for the estimated ranks to be in the interval \([0, 1]\). To ensure that all estimates are between 0 and 1 it is sufficient to force all kriging weights to be non-negative, however, this is accompanied by a loss in accuracy. Moreover, \( \tilde{Z} \) has no minimum kriging variance, only \( \hat{\bar{V}} \) has that property.

To partially overcome these drawbacks a direct sequential simulation of the ranks at the kriging locations \( x_j, j = n + 1, \ldots, n + m \), can be performed. The simulated ranks are drawn from a uniform distribution with mean equal to \( \hat{\bar{V}}(x_j) \) and variance equal to \( \sigma^2_K(x_j) \).

\[
\tilde{Z}(x_j) = \sqrt{3}\sigma_K(x_j), \tilde{Z}(x_j) = \sqrt{3}\sigma_K(x_j) \text{.}
\]

In all simulation steps, the kriging system consists of the original data and the previously sampled data. It may occur that the endpoints of the uniform distribution are outside the unit interval which could lead to simulated ranks outside \([0, 1]\). In this case they have to be reset to 0 or 1, depending on whether they are < 0 or > 1. After simulation the bias correction described in Eq. (15) is applied to the estimated ranks. For a large number \( N \) the sequential simulation is repeated \( N \) times and the resulting predictors are back-transformed using \( F^{-1}_\eta \) and averaged. This procedure yields estimates that are exact at known data locations, unbiased, follow the univariate distribution \( F_\eta \) and reproduce the covariance of the random field.

Sequential simulation is a time-consuming method for large data sets. Hence, we adapt a method proposed by Saito and Goovaerts (2000) who used it in the case of a normal-score transformation. Again, the simple kriging predictor, \( \hat{\bar{V}}(x_j) \), and the simple kriging variance, \( \sigma^2_K(x_j) \), are calculated. The conditional distribution of \( V(x_j) \) given the data is modeled as a uniform distribution with mean equal to the kriging predictor and variance equal to the kriging variance. If the endpoints \( a \) and \( b \) of the uniform distribution are outside the \([0, 1]\) interval, they are reset to 0 and 1, respectively, and the density of the local distribution changes to

\[
\begin{align*}
\frac{1}{2} \min \left\{ \frac{1}{V(x_j) - a}, \frac{1}{b - V(x_j)} \right\}, & \quad x \in [\max \{0, a\}, V(x_j)], \\
\frac{1}{2} \max \left\{ \frac{1}{b - V(x_j)}, \frac{1}{V(x_j) - a} \right\}, & \quad x \in [V(x_j), \min \{b, 1\}].
\end{align*}
\]

The e.g. 100 percentiles, \( v_p(x_j) \), of this local distribution are calculated, where \( p = \frac{k}{100} - 0.5 \) and \( k = 1, \ldots, 100 \). The back-transformed percentiles \( z_p(x_j) = F^{-1}_\eta(v_p(x_j)) \) are unbiased estimators for the quantiles of the local distribution of \( Z(x) \). Their average is an unbiased estimator for the mean, hence, the kriging estimate is defined as

\[
\tilde{Z}(x_j) = \frac{1}{100} \sum_{k=1}^{100} z_p(x_j) \quad \text{with} \quad p = \frac{k}{100} - 0.5.
\]

5.3 Plug-in estimation

The copula enters the rank order kriging procedure only through the Spearman rank correlation. Furthermore, both rank order kriging and disjunctive kriging only use bivariate copulas. On the one hand these facts may be useful since flexible bivariate copula families can be applied, but on the other hand these methods do not fully exploit the spatial copula model presented in Sect. 3.1. Here we calculate the full predictive distribution and plug-in estimators for this model in the case of continuous and discrete margins.

5.3.1 Continuous margins

When we go the Bayesian way, we can take account of the uncertainty of parameter estimation. Moreover, there exists a full predictive distribution for every rank-transformed variable \( V(x_0) \) at an unknown location \( x_0 \),

\[
p(v(x_0) \mid D) = \int p(v(x_0) \mid \Theta, D) p(\Theta \mid D) d\Theta.
\]

When we falsely assume that the maximum likelihood estimates, \( \hat{\Theta} \), of all the parameters are the true values, we get that \( p(v(x_0) \mid D) = \hat{c}_{\hat{\Theta}, \hat{\Lambda}}(v(x_0), v(x_1), \ldots, v(x_n)) \). In the spatial copula model this conditional density is exactly the density of the conditional copula of \( V(x_0) \) given the rank-transformed data and the estimated parameters

\[
c_{\hat{\Theta}, \hat{\Lambda}}(v(x_0), v(x_1), \ldots, v(x_n)) = \frac{c_{\hat{\Theta}, \hat{\Lambda}}(v(x_0), v(x_1), \ldots, v(x_n))}{c_{\hat{\Theta}, \hat{\Lambda}}(v(x_1), \ldots, v(x_n))},
\]

where \( c_{\hat{\Theta}, \hat{\Lambda}}(v(x_0), v(x_1), \ldots, v(x_n)) \) is the density of the conditional copula of \( V(x) \), given the rank-order kriging procedure and disjunctive kriging only use bivariate copulas. On the one hand these facts may be useful since flexible bivariate copula families can be applied, but on the other hand these methods do not fully exploit the spatial copula model presented in Sect. 3.1. Here we calculate the full predictive distribution and plug-in estimators for this model in the case of continuous and discrete margins.
where \( v(x_i) = F_\eta(z(x_i)) \) and \( i = 1, \ldots, n \). Since the predictive density of \( V(x_0) \) is defined on \([0, 1] \), we avoid estimated ranks outside the unit interval that occur in the rank order kriging approach. If the copula is constructed from a multivariate distribution with conditional density \( d \) and marginal distribution \( F \) with density \( f \), Eq. (3) tells us that the conditional copula can be written as

\[
c_{\hat{\theta}, \hat{\lambda}}(v(x_0) \mid D) = \frac{d \left( F^{-1}(v(x_0)) \right) \mid \hat{\Theta}, \hat{D}}{f \left( F^{-1}(v(x_0)) \right)}.
\]

In the case of a Gaussian copula \( F = \Phi \), \( f = \phi \) and \( d = \phi_{\mu, \sigma^2} \) is a Gaussian density with mean \( \mu = \Sigma_{12} \Sigma_{22}^{-1} a \) and variance \( \sigma^2 = 1 - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \), where \( \Sigma_{22} \) is the correlation matrix of the known locations, \( \Sigma_{12} = \Sigma_{21}^T \) is the vector of correlations between the known locations and the location where prediction should take place and \( a = (\Phi^{-1}(v(x_1)), \ldots, \Phi^{-1}(v(x_n)))^T \).

The predictive density of \( Z(x_0) \) can be calculated by using a Jacobian transformation (Kazianka and Pilz 2008; Bárdoossy and Li 2008). To get back from the ranks to the original scale the transformation is \( F^{-1} \). The corresponding Jacobian determinant is exactly the marginal density \( f_\eta \). Hence,

\[
p \left( z(x_0) \mid \hat{\Theta}, D \right) = c_{\hat{\theta}, \hat{\lambda}}(F_\eta(z(x_0)) \mid D) f_\eta(z(x_0)).
\]

The Bayes estimator for \( Z(x_0) \) under the quadratic loss is the mean of the predictive distribution, \( \hat{Z}(x_0) = E\left( Z(x_0) \mid \hat{\Theta}, D \right) \). With the help of an integral transformation we derive an estimator that is easy to evaluate due to the finite integral boundaries:

\[
\hat{Z}(x_0) = \int_\infty^{-\infty} z(x_0) c_{\hat{\theta}, \hat{\lambda}}(F_\eta(z(x_0)) \mid D) F_\eta(z(x_0)) \, dz = \int_0^1 F^{-1}_\eta(v(x_0)) c_{\hat{\theta}, \hat{\lambda}}(v(x_0) \mid D) \, dv.
\]

Analogously, the prediction variance \( \hat{\sigma}^2(x_0) \) is

\[
\hat{\sigma}^2 = \int_0^1 \left( F^{-1}_\eta(v(x_0)) - \hat{Z}(x_0) \right)^2 c_{\hat{\theta}, \hat{\lambda}}(v(x_0) \mid D) \, dv.
\]

If the mean \( \hat{Z}(x_0) \) (and, hence, the variance) does not exist, the Bayes estimator under the absolute loss, the median of the predictive distribution, can be calculated. Estimates for the quantiles \( z_{p}(x_0) \) and especially also for the median \( z_{0.5}(x_0) \) can be calculated by solving \( \int_{0}^{z}(x_0) c_{\hat{\theta}, \hat{\lambda}}(v(x_0) \mid D) \, dv = p \) for \( \hat{v}(x_0) \) and transforming to the original scale, \( \hat{z}_{p}(x_0) = F^{-1}_\eta\left( \hat{v}_p(x_0) \right) \).

Similarly to copula kriging the frequently applied trans-Gaussian kriging (De Oliveira et al. 1997) also works with a marginal transformation of the random field. The aim of trans-Gaussian kriging is to deal with non-Gaussian random fields by assuming that the transformed random field, \( Y(x) = g(Z(x)) \), is Gaussian and \( g \) is a suitable transformation that has to be determined. In most applications the transformation \( g \) is chosen from the Box-Cox family of transformations. In the following we show that there is a direct relationship between the trans-Gaussian kriging model and the spatial copula model.

**Theorem 1** The trans-Gaussian kriging model using strictly monotone, differentiable transformation functions is equivalent to the Gaussian spatial copula model.

**Proof** Assume that we have a trans-Gaussian random field with a strictly monotone, differentiable transformation \( g \). Hence, \( Y(x) = g(Z(x)) \sim N(\mu, \sigma^2) \). If \( g \) is increasing, the invariance theorem mentioned in Sect. 2 yields that the copula corresponding to the multivariate distribution of \( Z(x) \) must be the Gaussian copula corresponding to \( Y(x) \). If \( g \) is decreasing, we deduce the same from Eq. (5) and from the radial-symmetry of the Gaussian copula:

\[
c_{g(Z_1), \ldots, g(Z_n)}(u_1, \ldots, u_n) = c_{Z_1, \ldots, Z_n}(1 - u_1, \ldots, 1 - u_n)
\]

Using \( Z(x) = g^{-1}(Y(x)) \) we obtain the univariate marginal distribution of \( Z(x) \) as

\[
F_Z(z) = \int_{-\infty}^{z} \phi_{\mu, \sigma^2}(g(t)) |g'(t)| \, dt,
\]

and the Gaussian spatial copula model is fully determined. If we assume that the random field follows the Gaussian spatial copula model with known \( F_Z \), then \( g(z) = \Phi^{-1}(F_Z(z)) \) is a suitable transformation.

Since we can also use any other copula different from the Gaussian copula in our approach, we get that the spatial copula model is a generalization of the trans-Gaussian model. Even if we want to stay within the Gaussian framework, it is more convenient to use the copula methodology because it is easier to specify the univariate distribution of the random field than to determine a suitable transformation function. Especially when we work with multimodal or extreme value data this fact gets obvious.

For certain copula families not all data values can be used to build the predictive distribution. For the non-central \( \chi^2 \)-copula and the multiplied Gaussian copula mentioned in Sects. 3.1 and 3.2 this happens because one needs to evaluate \( 2^n \) terms for the calculation of the conditional copula. In these cases we propose a local prediction with the nearest, say 15, data points to have a numerically tractable method.
5.3.2 Discrete margins

When a continuous approximation (cf. Eq. (11)) has been applied for parameter estimation, one can simply follow the steps in Sect. 5.3.1 and change integrals to sums where necessary. In general, the full predictive density \( p \left( z(x_0) \mid \Theta, D \right) \) is of the form

\[
\sum_{j_1=0}^{n} \cdots \sum_{j_n=0}^{1} (-1)^{l_1} C_{\Theta, \lambda} \left( u_{0j_1}, u_{1j_1}, \ldots, u_{nj_1} \right) \\
\sum_{j_1=1}^{n} \cdots \sum_{j_n=0}^{1} (-1)^{l_2} C_{\Theta, \lambda} \left( u_{1j_2}, \ldots, u_{nj_2} \right)
\]

where \( l_1 = \sum_{i=0}^{n} j_i \), \( l_2 = \sum_{i=1}^{n} j_i^2 \) and the other denotations are taken from Eq. (10). Due to the numerical complexity in the evaluation of the copula, only local predictions can be performed. Estimates for prediction mean and variance can be obtained easily once the predictive density is evaluated.

6 Application: SIC2004 Joker data

In this section we test our methodology for modeling and interpolation in the case of continuous margins by means of the Joker data set, which was investigated in detail during the spatial interpolation comparison SIC2004 (Dubois 2005) and in Smok et al. (2009) and Kazianka and Pilz (2008). This extreme value data set simulates an accidental release of radioactivity using a dispersion process. There are 200 training locations \( x_i = (x_{1i}, x_{2i}) \) with values \( z(x_i) \), \( i = 1, \ldots, 200 \), that have a mean of 108.99, a standard deviation of 121.96 and a skewness of 9.92. Figure 2 displays the training data as gray dots and the 808 test data as gray circles. The two extreme observations (1070.4 and 1499) are indicated by the black dots. A surface plot of the training data is shown in Figure 3.

We employ the Gaussian spatial copula model and use the three-parameter family of generalized extreme value distributions as the margins, similar to Kazianka and Pilz (2008). However, to account for the trend, they only fitted a quadratic trend surface using ordinary least squares. Here we use the methodology presented in Sect. 3.3 and specify the marginal location parameters as

\[
\mu_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}.
\]

The correlation matrix of the Gaussian distribution is parameterized (cf. Sect. 3.1) by a mixture of a Gaussian and an exponential correlation model,

\[
\rho(h) = \varepsilon I(h = 0) + (1 - \varepsilon) \left( ae^{-\frac{h^2}{c^2}} + (1 - \alpha) e^{-\frac{h}{r^2}} \right),
\]

where \( \varepsilon \in [0, 1] \) denotes the nugget effect, \( \alpha \in [0, 1] \) is a mixing parameter and \( r_1 \) and \( r_2 \) are range parameters. Geometric anisotropy, which is a crucial ingredient when analyzing this data set, is considered by the coordinate transformation

\[
(x_{1i}^{\text{trans}}, x_{2i}^{\text{trans}})^T = \begin{pmatrix}
\cos(\varphi) & \sin(\varphi) \\
-t \sin(\varphi) & t \cos(\varphi)
\end{pmatrix}
(x_{1i}, x_{2i})^T,
\]

with \( t \) being the ratio and \( \varphi \) being the angle parameter. All eleven model parameters are estimated using the maximum likelihood approach.

At first we have to test whether it is appropriate to use the Gaussian copula. This is done by applying the goodness-of-fit test presented in Sect. 4.3. For lag distances \( h > 150 \) the spatial dependence is small and the Gaussian copula model is adequate since it models almost independence. Approximated p-values for different lag classes with \( h \leq 150 \) and \( N = 2000 \) simulations are given in Tab. 1. As there is no significantly small p-value, we can assume that the Gaussian spatial copula model fits well to the data.

Spatial prediction is performed using the plug-in approach. The map with the predicted values is displayed...
in Fig. 6(a) and one can clearly identify the region with large values, i.e. the region where the simulated emergency scenario takes place. The estimates are plotted against the true values in Fig. 4. As can be seen, small values are predicted accurately and large values are underestimated. However, the predictive densities at the hotspots (cf. Fig. 5) are heavily tailed due to the use of the generalized extreme value distribution. Therefore, estimated values around 1500 are still contained in a 95% confidence interval. A contour plot with the 95% predictive quantiles for all test data is given in Fig. 6(b). To compare our results with those of more than 30 participants of the SIC2004, we calculate root mean square error (RMSE=64.47), mean absolute error (MAE=15.83), mean error (ME=−2.56) and Pearson correlation (r=0.72) for the test data. This would be the third smallest RMSE, the second smallest MAE and the third largest Pearson correlation. In addition, these results are slightly better than those presented by Kazianka and Pilz (2008) which arises from the flexibility that is obtained by using the trend model proposed in Sect. 3.3. It is worth mentioning that trans-Gaussian kriging using a Box-Cox transformation does not work for such an extreme value data set. Moreover, as is reported in Spöck et al. (2009) who used the Bayesian bootstrap approach by Pilz and Spöck (2008), even trans-Gaussian kriging using a special log-log transformation leads to inferior results. This emphasizes that for environmental mapping and monitoring it is advisable to specify the marginal distribution of the underlying process rather than a family of transformations.

To estimate the exceedance probabilities for certain thresholds we can employ the copula-based indicator kriging presented in Sect. 5.1. In this example we use 70, 100, 120, 140, 200 and 1000 as the thresholds. The probabilities that the random field takes values below these thresholds are displayed in Fig. 7(a)-(f). It can be deduced from Fig. 7(f) that in the region with extremely large observations there is still a probability of about 0.1 for exceeding a value of 1000.

### 7 Conclusion

Copulas can be used to flexibly describe spatial dependence and to perform spatial interpolation. In the case of continuous marginal distributions the proposed approach generalizes the trans-Gaussian kriging method and is therefore a valuable tool for working with non-Gaussian, multimodal and extreme value data. Specifying the univariate distribution of the random field is more convenient than finding a suitable transformation for trans-Gaussian kriging, which makes the spatial copula model attractive even if the Gaussian copula is used. By assuming that the marginal mean or location

<table>
<thead>
<tr>
<th>h</th>
<th>0-30</th>
<th>30-60</th>
<th>60-90</th>
<th>90-120</th>
<th>120-150</th>
</tr>
</thead>
<tbody>
<tr>
<td>T_n</td>
<td>0.50</td>
<td>0.37</td>
<td>0.78</td>
<td>0.69</td>
<td>0.20</td>
</tr>
<tr>
<td>S_n</td>
<td>0.89</td>
<td>0.18</td>
<td>0.89</td>
<td>0.71</td>
<td>0.47</td>
</tr>
</tbody>
</table>

Table 1 Approximated p-values for different lag classes using the Kolmogorov-Smirnov ($T_n$) and Cramer-von Mises ($S_n$) statistics.
Fig. 7 Results for copula-based indicator kriging.

parameter follows a generalized linear model, we are able to take account of a spatial trend and other covariates such as elevation. In this way the model provides an alternative to the well-known GLGM and enables us to work with random fields that have discrete margins. Estimation of the parameters and subsequent spatial interpolation can be easily extended to a Bayesian approach by assigning priors to the model parameters and applying MCMC techniques. Results on the SIC2004 Joker data demonstrate that copula-based spatial modeling can be applied for emergency monitoring and for estimating exceedance probabilities for certain emergency thresholds in environmental monitoring systems.

Acknowledgements This work was partially funded by the European Commission, under the Sixth Framework Programme, by the Contract N. 033811 with DG INFSO, Action Line IST-2005-2.5.12 ICT for Environmental Risk Management. The views expressed herein are those of the authors and are not necessarily those of the European Commission.

References


Sklar A (1959) Fonctions de repartition a n dimensions et leurs marges. Publ Inst Statist Univ Paris, 8:229-231


