

A MODEL-BASED GEOSTATISTICAL APPROACH FOR SKEWED RADIOACTIVITY DATA

Luca Bagnato ¹ & Marco Minozzo ²

¹ *Department of Mathematical Disciplines, Mathematical Finance and Econometrics,
Catholic University of the Sacred Heart,
Via Necchi 9, 20123 Milano, Italy, email luca.bagnato@unicatt.it*

² *Department of Economics,
University of Verona,
Via dell'Artigliere 19, 37129 Verona, Italy, email marco.minozzo@univr.it*

Abstract. In the last years, to improve the performance of prediction of radioactive contamination, an increasing number of studies have explored and exploited the potentials of geostatistical methods. However, traditional methods like kriging and cokriging are optimal only in the case in which the data may be assumed Gaussian and do not properly cope with data measurements that are discrete, nonnegative or show some degree of skewness, as in many environmental applications concerned with radioactivity measurements. To deal with geostatistical skewed data, we consider a model-based approach in which measurements are modeled with the help of a latent Gaussian structure and some recent classes of skewed distributions extending the normal one. For our model we investigate the implied spatial autocorrelation structure and the marginal distributions of the observable variables. In particular we show that all finite-dimensional marginal distributions of the observable variables belong to the family of the unified skew-normal distribution. Estimation of some of the unknown parameters of the model can be carried out by employing a Monte Carlo expectation maximization procedure, whereas predictions of both latent and observed (at unsampled sites) variables, can be supplied by Markov chain Monte Carlo algorithms.

Keywords. Latent Gaussian process, Monte Carlo EM, multivariate geostatistics, skew-normal distribution, unified skew-normal distribution.

1 Skewed radioactivity data

The problem of dealing with geostatistical non-Gaussian data, and in particular with skewed data, is common to a large variety of spatial data sets (on radiological monitoring, rainfalls, winds, etc.). However, though this problems had been dealt with in the literature in several papers, the modeling of geostatistical skewed data still remains an issue. For instance, with regard to radiological monitoring, Brenning and Dubois (2008), disregarding any physically-based modeling approach, argue on the necessity of developing mapping algorithms for emergency detection taking into consideration the skewness in the data.

A boost to these developments came from the Spatial Interpolation Comparison 2004 - SIC2004 (see Dubois and Galmarini, 2005) in which, whereas the routine scenario could easily be modeled using a Gaussian random field, the emergency scenario, which mimics an accidental release of radioactivity, needs to be modeled taking properly into account that, due to the presence of extreme measurements, the data are positively skewed. Just to cite a few works, to deal with skewed measurements coming from radioactive monitoring, Kazianka and Pilz (2010) and Gräler (2014) propose copula-based geostatistical approaches, whereas Desnoyers et al. (2011) argue that the structuring of extreme values can be faced in a coherent manner by using the class of Hermitian isofactorial models. Moreover, Bechler et al. (2013) propose a Gaussian anamorphosis transformation to deal with skewed data coming from contaminated facilities, and Kazianka and Pilz (2011) argue in favor of a Bayesian approach pointing out that both the Gaussian copula and the non-Gaussian χ^2 -copula models are inappropriate to model strongly skewed radioactivity measurements. Other works dealing with skewed radiological measurements are those of Oliver and Badr (1995), which is concerned with the estimation of the variogram and the development of optimal sampling plans, De and Faria (2011), which proposes a dynamic spatial Bayesian model for non-Gaussian measurements from radioactivity deposition, as well as the works of Maglione and Diblasi (2004), Pilz and Spöck (2008) and Spöck (2012).

Although many works have been devoted to the modeling and prediction of different types of non-Gaussian data, in particular skewed data, much fewer have been framed in a multivariate context, that is, in presence of more than one regionalized variable. From the point of view of applications, some examples of these contexts are given by the radiological data in Herranz et al. (2011) or by the data related to the Fukushima disaster (data are available from TEPCO at <http://www.tepco.co.jp>) where more than one radiological measurement has been collected for each sampling site. Among the works dealing with multivariate geostatistical non-Gaussian data are those of Minozzo and Fruttini (2004) and Minozzo and Ferrari (2013) (see also Ren and Banerjee, 2013), which extend to a multivariate setting the modeling approach of Diggle, Moyeed and Tawn (1998), that of Chagneau et al. (2011), which propose a hierarchical Bayesian approach to model Gaussian, count, and ordinal variables, and that of Wibrin, Bogaert and Fasbender (2006), which explore the use of the Bayesian maximum entropy approach in presence of both continuous and categorical regionalized variables.

2 A multivariate geostatistical skew-normal model

Disregarding any physically-based modeling, to tackle skewness in a multivariate (that is, in presence of more than one regionalized variable) geostatistical context, we propose an approach based on the use of the skew-normal distribution (Azzalini and Dalla Valle, 1996; Azzalini, 2005; Azzalini, 2014) and on a latent (unobserved) multivariate Gaussian spatial process, extending some of the ideas in Minozzo and Ferracuti (2012), Minozzo and Ferrari

(2013), and Zhang and El-Shaarawi (2010). Let $y_i(\mathbf{x}_k)$, $i = 1, \dots, m$, $k = 1, \dots, K$, be a set of geo-referenced data measurements relative to m regionalized variables, gathered at K spatial locations \mathbf{x}_k . Each of these m measured variables can be viewed as a partial realization of a particular stochastic process $Y_i(\mathbf{x})$, $i = 1, \dots, m$, $\mathbf{x} \in \mathbb{R}^2$. We assume that these stochastic processes are given by

$$Y_i(\mathbf{x}) = \beta_i + Z_i(\mathbf{x}) + \omega_i S_i(\mathbf{x}), \quad i = 1, \dots, m, \quad (1)$$

where β_i and ω_i are unknown constants, representing, respectively, an intercept and a scale parameter, and $Z_i(\mathbf{x})$ and $S_i(\mathbf{x})$ are latent processes. In particular, for every $i = 1, \dots, m$, $Z_i(\mathbf{x})$ is a mean zero stationary Gaussian process, whereas for every $i = 1, \dots, m$, and for each $\mathbf{x} \in \mathbb{R}^2$, $S_i(\mathbf{x})$ is an independent random variable distributed as a skew-normal (Azzalini, 2005), that is, $S_i(\mathbf{x}) \sim SN(0, 1, \alpha_i)$, which means that, for every $\mathbf{x} \in \mathbb{R}^2$, the density of $S_i(\mathbf{x})$ is given by $f_{S_i}(s) = 2\phi_1(s; 1)\Phi(\alpha_i s)$, for $-\infty < s < \infty$, where $\alpha_i \in \mathbb{R}$, $\phi_1(\cdot; 1)$ is the scalar normal density function with zero mean and unit variance, and $\Phi(\cdot)$ is the scalar $N(0, 1)$ distribution function.

Let us note that, for each $i = 1, \dots, m$, and for every $\mathbf{x} \in \mathbb{R}^2$, conditionally on $Z_i(\mathbf{x})$, the random variable $Y_i(\mathbf{x})$ has a skew-normal distribution, that is,

$$Y_i(\mathbf{x}) | Z_i(\mathbf{x}) \sim SN(\beta_i + Z_i(\mathbf{x}), \omega_i^2, \alpha_i), \quad (2)$$

which means that we can write its density as

$$f(y_i(\mathbf{x}) | z_i(\mathbf{x})) = 2 \phi_1(y_i(\mathbf{x}) - \beta_i - z_i(\mathbf{x}); \omega_i^2) \Phi\left(\frac{\alpha_i}{\omega_i}(y_i(\mathbf{x}) - \beta_i - z_i(\mathbf{x}))\right),$$

where $\phi_1(\cdot; \sigma^2)$ is the scalar normal density function with zero mean and positive variance σ^2 . Moreover, for each $i = 1, \dots, m$, and for every $\mathbf{x} \in \mathbb{R}^2$, the (scalar) random variable $Y_i(\mathbf{x})$ has a (marginal) skew-normal distribution, that is,

$$Y_i(\mathbf{x}) \sim SN\left(\beta_i, \varsigma_i^2 + \omega_i^2, \alpha_i \omega_i / \sqrt{\varsigma_i^2(1 + \alpha_i^2) + \omega_i^2}\right), \quad (3)$$

where $\varsigma_i^2 = \text{Var}[Z_i(\mathbf{x})]$.

For the latent part of the model, that is, for the zero mean stationary Gaussian processes $Z_i(\mathbf{x})$, $i = 1, \dots, m$, we assume that they have covariance and cross-covariance functions

$$\text{Cov}[Z_i(\mathbf{x}), Z_j(\mathbf{x} + \mathbf{h})] = \begin{cases} \varsigma_i^2 \rho(\mathbf{h}), & i = j, \\ \varsigma_{ij} \rho(\mathbf{h}), & i \neq j, \end{cases}$$

where $\mathbf{h} \in \mathbb{R}^2$, $\varsigma_{ij} = \text{Cov}[Z_i(\mathbf{x}), Z_j(\mathbf{x})]$ and $\rho(\mathbf{h})$ is a real spatial autocorrelation function with $\rho(\mathbf{0}) = 1$ and $\rho(\mathbf{h}) \rightarrow 0$, as $\|\mathbf{h}\| \rightarrow \infty$.

As regard to the applicability of our model to skewed data, in particular to radiological measurements, the flexibility of our model, which includes the Gaussian random field as a

particular case, allows to naturally model situations in which extreme measurements either are present or not. In some applications, since the observable $Y_i(\mathbf{x})$ can take values on the whole real line, when the actual measurements are nonnegative and their distribution is too close to the origin, our model might have to be applied to transformed data (using, for instance, a logarithmic transformation) and not to the raw data $y_i(\mathbf{x}_k)$. However, contrary to usual practice, we do not pretend with such a transformation to eliminate the skewness in the data and to resemble a Gaussian realization, but just intend to use it to change the range of the observed measurements. Let us note that the modeling construction proposed here is substantially different from some of the most popular constructions based on the skew-normal distribution that have recently appeared in the literature to model univariate skewed spatial data (for a critical discussion on some of these constructions see Minozzo and Ferracuti, 2012).

Assuming to know the spatial autocorrelation function $\rho(\mathbf{h})$, the model depends on the parameter vector $\boldsymbol{\vartheta}^* = (\boldsymbol{\beta}, \boldsymbol{\Sigma}_Z, \boldsymbol{\omega}, \boldsymbol{\alpha})$, where $\boldsymbol{\beta} = (\beta_1, \dots, \beta_m)^T$, $\boldsymbol{\Sigma}_Z$ is the $m \times m$ matrix with elements ς_i^2 and ς_{ij} , $\boldsymbol{\omega} = (\omega_1, \dots, \omega_m)^T$, and $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_m)^T$. As shown by Minozzo and Bagnato (2014), assuming to know also the parameters $\boldsymbol{\omega}$ and $\boldsymbol{\alpha}$ characterizing the shape of the skew-normal (conditional) distributions, it is possible to estimate the parameter $\boldsymbol{\vartheta} = (\boldsymbol{\beta}, \boldsymbol{\Sigma}_Z)$ with a likelihood based estimation procedure by implementing a computationally intensive Monte Carlo expectation maximization (MCEM) algorithm and exploiting Markov chain Monte Carlo (MCMC) techniques. On the other hand, the other parameters of the model as well as the functional form of the spatial autocorrelation function $\rho(\mathbf{h})$ can previously be calibrated comparing the theoretical marginal distributions and the theoretical variograms with their corresponding empirical counterparts.

Given as known all parameters of the model, prediction of the observable processes $Y_i(\mathbf{x})$ at an unobserved spatial location (or at an unobserved set of spatial locations) can be carried out either by exploiting some of the properties of the finite-dimensional marginal distributions of the model, or by implementing some MCMC algorithm. On the other hand, for the prediction of the unobserved processes $Z_i(\mathbf{x})$, we need to resort to MCMC algorithms. In the case in which we are interested in predicting a latent process on a large set of spatial locations (maybe on a grid), instead of carrying out an MCMC run at each spatial location, we can carry out an MCMC run only at the sampling points (that is, only at those points for which we gathered observations), and then exploit a linear property (see Minozzo and Bagnato, 2014) similar to Kriging, and also similar to that found by Zhang (2002) in a univariate framework, to obtain predictions at all other spatial locations

3 Variograms and cross-variograms

Let us consider here the correlation structure of the observable processes, induced by the latent Gaussian spatial processes. For the observable stochastic processes $Y_i(\mathbf{x})$,

$i = 1, \dots, m$, we have that

$$\mathbb{E}[Y_i(\mathbf{x})] = \beta_i + \omega_i \delta_i \left(\frac{2}{\pi}\right)^{\frac{1}{2}}, \quad \text{Var}[Y_i(\mathbf{x})] = \varsigma_i^2 + \omega_i^2 \left[1 - \frac{2}{\pi} \delta_i^2\right],$$

where $\delta_i = \alpha_i / \sqrt{1 + \alpha_i^2}$. Moreover, for $\mathbf{h} \neq \mathbf{0}$, the variogram of the observable $Y_i(\mathbf{x})$ takes the form

$$\gamma_{ii}(\mathbf{h}) = \frac{1}{2} \text{Var}[Y_i(\mathbf{x} + \mathbf{h}) - Y_i(\mathbf{x})] = \omega_i^2 \left[1 - \frac{2}{\pi} \delta_i^2\right] + \varsigma_i^2 [1 - \rho(\mathbf{h})], \quad (4)$$

and, since $\gamma_{ii}(\mathbf{0}) = 0$ and $\gamma_{ii}(\mathbf{0}^+) = \omega_i^2 [1 - (2/\pi) \delta_i^2]$, it is discontinuous in zero. To visually assess Formula (4), Figure 1 shows the form taken by the variogram $\gamma_{ii}(\mathbf{h})$ for different values of the parameters, in the case of a Cauchy spatial autocorrelation function $\rho(\mathbf{h}) = [1 + (\|\mathbf{h}\|/\gamma)^2]^{-\eta}$, with $\gamma = 1$ and $\eta = 1$, and of a powered exponential (stable) spatial autocorrelation function $\rho(\mathbf{h}) = \exp[-(\gamma \|\mathbf{h}\|)^\eta]$, with $\gamma = 1$ and $\eta = 1$. As we can see, the nugget of the variogram decreases for decreasing values of ω and for values of the skewness parameter α departing from zero.

On the other hand, for any two stochastic processes $Y_i(\mathbf{x})$ and $Y_j(\mathbf{x})$, with $i \neq j$, the cross-variogram is given by

$$\gamma_{ij}(\mathbf{h}) = \frac{1}{2} \text{Cov}[Y_i(\mathbf{x} + \mathbf{h}) - Y_i(\mathbf{x}), Y_j(\mathbf{x} + \mathbf{h}) - Y_j(\mathbf{x})] = \varsigma_{ij} [1 - \rho(\mathbf{h})]. \quad (5)$$

To somehow evaluate the degree of flexibility of the spatial autocorrelation structure of our model, we can compare it with that implied by other geostatistical approaches proposed in the literature, in particular with copula-based approaches, for the modeling of skewed data, coming in particular from radiological measurements. For instance, in the univariate case, that is, considering just one (strongly) stationary spatial process $Y(\mathbf{x})$, for $\mathbf{x} \in \mathbb{R}^2$, and with regard to copula-based approaches, Bárdossy (2006) showed that, for any cutoff value ξ , having defined the indicator variable $I_\xi(\mathbf{x})$ taking value 1, if $Y(\mathbf{x}) < \xi$, and value 0, otherwise, the indicator variogram is given by

$$\gamma_\xi(\mathbf{h}) = \frac{1}{2} \mathbb{E}[(I_\xi(\mathbf{x} + \mathbf{h}) - I_\xi(\mathbf{x}))^2] = F_Y(\xi) - C_{\mathbf{h}}(F_Y(\xi), F_Y(\xi)),$$

where $F_Y(\cdot)$ is the cumulative distribution function of the univariate marginal distribution of the spatial process $Y(\mathbf{x})$, and, for any two selected quantiles ξ_1 and ξ_2 , the spatial copula $C_{\mathbf{h}}(\cdot, \cdot)$ is given by

$$C_{\mathbf{h}}(\xi_1, \xi_2) = P(F_Y(Y(\mathbf{x})) < \xi_1, F_Y(Y(\mathbf{x} + \mathbf{h})) < \xi_2) = C(F_Y(Y(\mathbf{x})), F_Y(Y(\mathbf{x} + \mathbf{h}))),$$

where $C(\cdot, \cdot)$ is a standard bivariate copula model which may belong, for instance, to the Gaussian or to the Archimedean class. A detailed comparison of the spatial autocorrelation structure of our model with that implied by other approaches, and in particular by copula-based approaches, shows that our approach is flexible enough to model a large variety of situations.

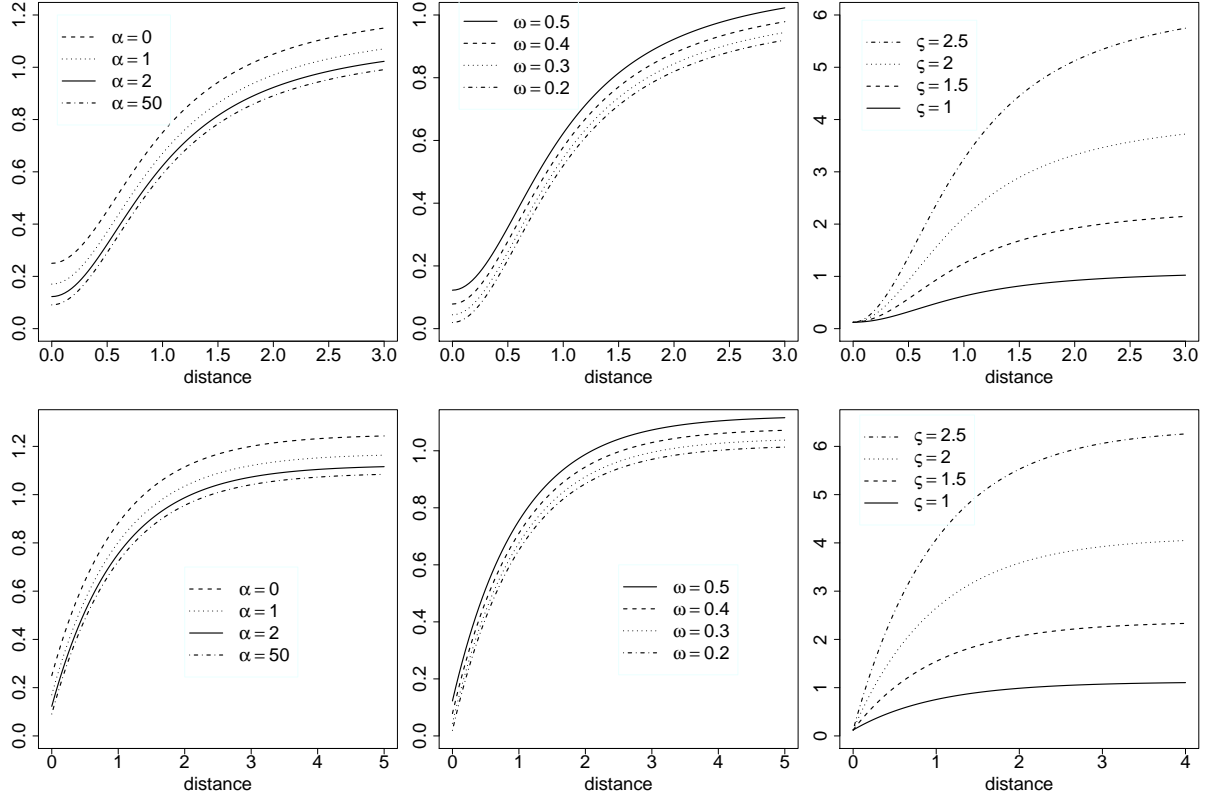


Figure 1: The graphs show the shape of the theoretical variogram $\gamma_{ii}(\mathbf{h})$ given in Formula 4, for a Cauchy autocorrelation function with both parameters equal to 1 (top), and for a powered exponential autocorrelation function with both parameters equal to 1 (bottom). The other parameters have been set equal to: (left) $\omega = 0.5$, $\zeta = 1$; (middle) $\alpha = 2$, $\zeta = 1$; (right) $\alpha = 2$, $\omega = 0.5$. The solid lines in the graphs correspond to the same set of parameter values. The lines in the graphs on the left corresponding to $\alpha = 0$ give the variograms in the case of a Gaussian process.

4 Finite-dimensional marginal distributions

Before stating our results, we need to recall some definitions. Following Azzalini (2005), a random vector $\mathbf{Y} = (Y_1, \dots, Y_d)^T$ has a (multivariate) *skew-normal distribution* with parameters $\boldsymbol{\xi}$, $\boldsymbol{\Omega}$ and $\boldsymbol{\alpha}$, and we write $\mathbf{Y} \sim \text{SN}_d(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha})$, if it has probability density function

$$f(\mathbf{y}) = 2 \cdot \phi_d(\mathbf{y} - \boldsymbol{\xi}; \boldsymbol{\Omega}) \cdot \Phi(\boldsymbol{\alpha}^T \text{diag}(\boldsymbol{\omega})^{-1}(\mathbf{y} - \boldsymbol{\xi})), \quad (6)$$

for $\mathbf{y} \in \mathbb{R}^d$, where $\boldsymbol{\xi} \in \mathbb{R}^d$ is a vector of location parameters, $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_d)^T$, with $\alpha_i \in \mathbb{R}$, $i = 1, \dots, d$, is a vector of skewness parameters, $\boldsymbol{\Omega}$ is a (positive-definite) covariance matrix, $\boldsymbol{\omega} = (\omega_1, \dots, \omega_d)^T$ is the vector formed with the standard deviations of the matrix

Ω , that is, with the square roots of the diagonal elements of Ω , and $\text{diag}(\omega)$ is the $d \times d$ diagonal matrix with elements $\omega_1, \dots, \omega_d$ on the main diagonal and zeros elsewhere; moreover, $\phi_d(\cdot; \Omega)$ is the density function of the d -dimensional multivariate normal distribution with zero mean and (positive-definite) variance-covariance matrix Ω , and $\Phi(\cdot)$ is the cumulative distribution function of the (scalar) standard normal distribution $N(0, 1)$. Note that $\Omega = \text{diag}(\omega) \bar{\Omega} \text{diag}(\omega)$, where the matrix $\bar{\Omega}$ is a correlation matrix. We remember that ξ and Ω are not the mean vector and the variance-covariance matrix of the random vector \mathbf{Y} . This distribution extends the multivariate normal distribution and when $\alpha = \mathbf{0}$ it reduces to the latter.

In accordance with Arellano-Valle and Azzalini (2006), we say that a random vector $\mathbf{Y} = (Y_1, \dots, Y_d)^T$ has a *unified skew-normal distribution* (SUN) with parameters q, ξ, γ, ω and Ω^* , in symbols $\mathbf{Y} \sim \text{SUN}_{d,q}(\xi, \gamma, \omega, \Omega^*)$, if it has probability density function

$$f(\mathbf{y}) = \phi_d(\mathbf{y} - \xi; \Omega) \frac{\Phi_q(\gamma + \Delta^T \bar{\Omega}^{-1} \text{diag}(\omega)^{-1}(\mathbf{y} - \xi); \Gamma - \Delta^T \bar{\Omega}^{-1} \Delta)}{\Phi_q(\gamma; \Gamma)}, \quad (7)$$

for $\mathbf{y} \in \mathbb{R}^d$, where ξ, Ω and ω are as before, q is a positive integer, $\gamma \in \mathbb{R}^q$ and the matrices Γ, Δ and $\bar{\Omega}$ are defined by the following partitioning of the (positive-definite) correlation matrix Ω^* of dimension $(q + d) \times (q + d)$:

$$\Omega^* = \begin{bmatrix} \Gamma & \Delta^T \\ \Delta & \bar{\Omega} \end{bmatrix};$$

furthermore, $\Phi_d(\cdot; \Sigma)$ is the cumulative distribution function of the d -dimensional multivariate normal distribution with zero mean and (positive-definite) variance-covariance matrix Σ .

Now, considering our (weakly and strongly stationary) multivariate spatial process $(Y_1(\mathbf{x}), \dots, Y_m(\mathbf{x}))^T$, for $\mathbf{x} \in \mathbb{R}^2$, given in Equation (1), though its multivariate finite-dimensional marginal distributions are not skew-normal (in the sense of Azzalini, 2005), it is possible to show (Minozzo and Bagnato, 2014) that they all belong to the family of the SUN distribution of Arellano-Valle and Azzalini (2006). This implies, for instance, that, for each $i = 1, \dots, m$, the univariate spatial process $Y_i(\mathbf{x})$, for $\mathbf{x} \in \mathbb{R}^2$, has all its finite-dimensional marginal distributions belonging to the SUN family, and that, for any fixed site $\mathbf{x} \in \mathbb{R}^2$, the random vector $(Y_1(\mathbf{x}), \dots, Y_m(\mathbf{x}))^T$ has a SUN distribution. This last result is reported, without proof, in Proposition 1.

Proposition 1 *For any fixed site $\mathbf{x} \in \mathbb{R}^2$, the random vector $\mathbf{Y}_x = (Y_1(\mathbf{x}), \dots, Y_m(\mathbf{x}))^T$, where $Y_i(\mathbf{x}), \mathbf{x} \in \mathbb{R}^2, i = 1, \dots, m$, are the spatial processes defined in Equation (1), has a distribution in the SUN family, that is,*

$$\mathbf{Y}_x \sim \text{SUN}_{m,m+1}(\beta, \mathbf{0}_{m+1}, \mathbf{v}, \Omega_m^*), \quad (8)$$

where $\beta = (\beta_1, \dots, \beta_m)^T$, $\mathbf{0}_n$ is the vector of length n of all zeros,

$$\Omega_m^* = \begin{bmatrix} \mathbf{I}_{m+1} & \Delta_m^T \\ \Delta_m & \text{diag}(\mathbf{v})^{-1}(\Sigma_Z + \text{diag}(\omega)^2) \text{diag}(\mathbf{v})^{-1} \end{bmatrix}, \quad \Delta_m^T = \begin{bmatrix} \mathbf{0}_m^T \\ \text{diag}(\mathbf{v})^{-1} \text{diag}(\omega) \text{diag}(\delta) \end{bmatrix},$$

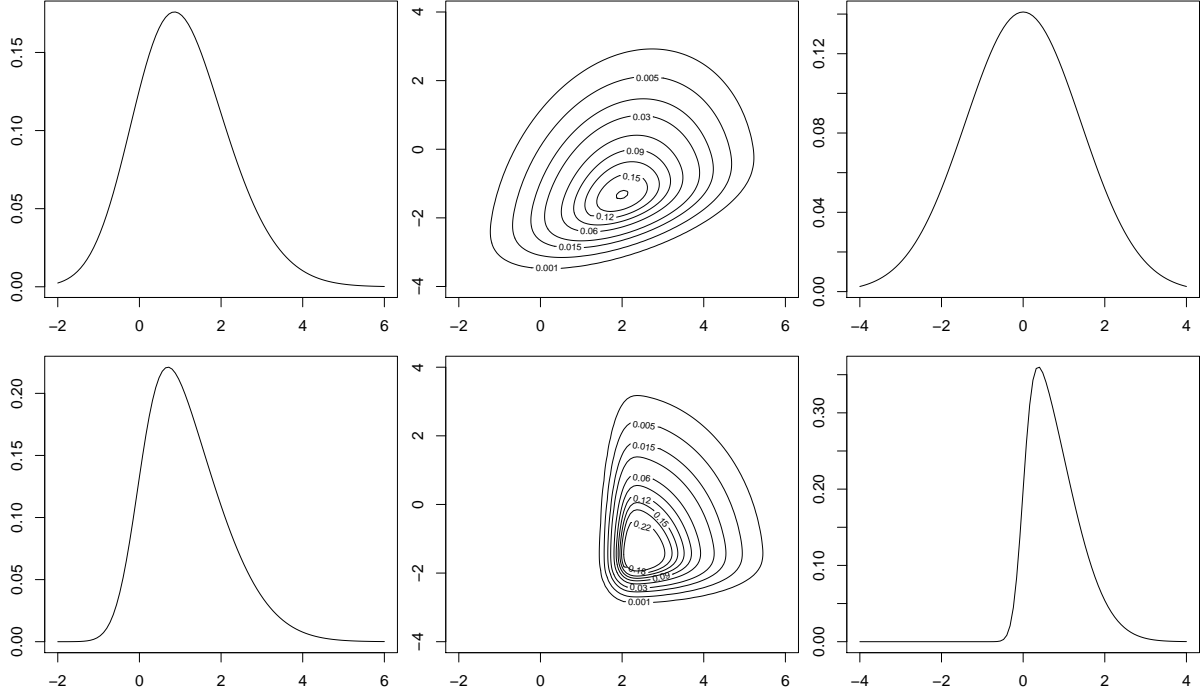


Figure 2: The graphs show the univariate ($Y_1(\mathbf{x})$ left; $Y_2(\mathbf{x})$ right) and the bivariate (middle) marginal distributions (densities) given by Proposition 1 for a model with $m = 2$. The distributions at the top are relative to the set of parameter values: $\beta_1 = 0$, $\beta_2 = 0$, $\alpha_1 = 5$, $\alpha_2 = 0$, $\omega_1 = 1.5$, $\omega_2 = 1$, $\varsigma_1^2 = 0.25$, $\varsigma_2^2 = 1$, $\varsigma_{12} = 0.5$. The distributions at the bottom are relative to the set of parameter values: $\beta_1 = 0$, $\beta_2 = 0$, $\alpha_1 = 5$, $\alpha_2 = 5$, $\omega_1 = 1.5$, $\omega_2 = 1$, $\varsigma_1^2 = 0.01$, $\varsigma_2^2 = 0.0001$, $\varsigma_{12} = 0.001$.

$$\mathbf{v} = (v_1, \dots, v_m)^T, \quad v_i = \sqrt{\varsigma_i^2 + \omega_i^2}, \quad i = 1, \dots, m,$$

$$\boldsymbol{\delta} = (\delta_1, \dots, \delta_m)^T, \quad \delta_i = \alpha_i / \sqrt{1 - \alpha_i^2}, \quad i = 1, \dots, m,$$

$\boldsymbol{\omega} = (\omega_1, \dots, \omega_m)^T$, \mathbf{I}_n is the identity matrix of size $n \times n$, ς_i^2 are the variances of the stationary processes $Z_i(\mathbf{x})$, $i = 1, \dots, m$, and $\boldsymbol{\Sigma}_Z$ is the $m \times m$ matrix with elements $\{Cov[Z_i(\mathbf{x}), Z_j(\mathbf{x})]\}_{i,j=1,\dots,m}$.

Let us note that, as a consequence of the assumptions of our model, the marginal distributions given in Proposition 1 belong to a proper subset of the SUN family. To supply some examples, Figure 2 shows the univariate and bivariate marginal distributions given by Proposition 1 in the case of $m = 2$ and for two different sets of parameter values. To our knowledge, our model seems to be the first geostatistical model for which this or similar properties involving some family of skew-normal distributions have been proved. Under the assumption of stationarity of the observable process, this property

allows a direct check of the model through the empirical marginal (not just univariate) distributions of the data. For instance, for a given set of observations, the empirical distribution of $y_i(\mathbf{x}_k)$, $k = 1, \dots, K$, for any given $i = 1, \dots, m$, can be compared with its theoretical univariate marginal distribution, or the empirical distribution of the couples $(y_i(\mathbf{x}_k), y_j(\mathbf{x}_k))$, $k = 1, \dots, K$, for any given couple $i, j = 1, \dots, m$, can be compared with the corresponding theoretical bivariate marginal distribution.

Let us remark that whereas for our model the marginal distributions are known, these are largely unknown for many other approaches proposed for the modeling of skewed geostatistical data. Moreover, regarding the use of copulas, albeit these have been extensively exploited to deal with univariate non-Gaussian spatial processes, directly modeling some of the finite-dimensional marginal distributions of the process, in presence of more than one regionalized variable their use is much less straightforward.

References

- [1] Arellano-Valle, R. B. and Azzalini, A. (2006), On the unification of families of skew-normal distributions, *Scand. J. Statist.*, 33, 561–574.
- [2] Azzalini, A. (2005), The skew-normal distribution and related multivariate families, *Scand. J. Statist.*, 32, 159–188.
- [3] Azzalini, A. (2014), *The Skew-Normal and Related Families*, Cambridge University Press, New York.
- [4] Azzalini, A. and Dalla Valle, A. (1996), The multivariate skew-normal distribution, *Biometrika*, 83, 715–726.
- [5] Bárdossy, A. (2006), Copula-based geostatistical models for groundwater quality parameters, *Water Resources Research*, 42, W11416.
- [6] Bechler, A., Romary, T., Jeannée, N. and Desnoyers, Y. (2013), Geostatistical sampling optimization of contaminated facilities, *Stoch. Environ. Res. Risk Assess.*, 27, 1967–1974.
- [7] Brenning, A. and Dubois, G. (2008), Towards generic real-time mapping algorithms for environmental monitoring and emergency detection, *Stoch. Environ. Res. Risk Assess.*, 22, 601–611.
- [8] Chagneau, P., Mortier, F., Picard, N. and Bacro, J. -N. (2011), A hierarchical Bayesian model for spatial prediction of multivariate non-Gaussian random fields, *Biometrics*, 67, 97–105.
- [9] De, S. and Faria, Á. E. (2011), Dynamic spatial Bayesian models for radioactivity deposition, *J. Time Ser. Anal.*, 32, 607–617.
- [10] Desnoyers, Y., Chilès, J. -P., Dubot, D., Jeannée, N. and Idasiak, J. -M. (2011), Geostatistics for radiological evaluation: study of structuring of extreme values, *Stoch. Environ. Res. Risk Assess.*, 25, 1031–1037.
- [11] Diggle, P. J., Moyeed, R. A. and Tawn, J. A. (1998), Model-based geostatistics (with discussion), *Appl. Stat.*, 47, 299–350.

- [12] Dubois, G. and Galmarini, S. (2005), Spatial interpolation comparison (SIC) 2004: introduction to the exercise and overview of results, in: Dubois, G. (Ed.), *Automatic Mapping Algorithms for Routine and Emergency Monitoring Data - Spatial Interpolation Comparison 2004*, Office for Official Publication of the European Communities.
- [13] Gräler, B. (2014), Modelling skewed spatial random fields through the spatial vine copula, *Spatial Statistics*, <http://dx.doi.org/10.1016/j.spasta.2014.01.001>
- [14] Herranz, M., Romero, L. M., Idoeta, R., Olondo, C., Valiño, F. and Legarda, F. (2011), Inventory and vertical migration of ^{90}Sr fallout and $^{137}\text{Cs}/^{90}\text{Sr}$ ratio in Spanish mainland soils, *J. Environ. Radioactiv.*, 102, 987–994.
- [15] Kazianka, H. and Pilz, J. (2010), Copula-based geostatistical modeling of continuous and discrete data including covariates, *Stoch. Environ. Res. Risk Assess.*, 24, 661–673.
- [16] Kazianka, H. and Pilz, J. (2011), Bayesian spatial modeling and interpolation using copulas, *Computers & Geosciences*, 37, 310–319.
- [17] Maglione, D. S. and Diblasi, A. M. (2004), Exploring a valid model for the variogram of an isotropic spatial process, *Stoch. Environ. Res. Risk Assess.*, 18, 366–376.
- [18] Minozzo, M. and Bagnato, L. (2014), A skew-normal multivariate geostatistical model with latent factors, Manuscript submitted for publication, 42 pages.
- [19] Minozzo, M. and Ferracuti, L. (2012), On the existence of some skew-normal stationary processes, *Ch. J. Stat.*, 3, 157–170.
- [20] Minozzo, M. and Ferrari, C. (2013), Multivariate geostatistical mapping of radioactive contamination in the Maddalena Archipelago (Sardinia, Italy), *AStA Adv. Stat. Anal.*, 97, 195–213.
- [21] Minozzo, M. and Fruttini, D. (2004), Loglinear spatial factor analysis: an application to diabetes mellitus complications, *Environmetrics*, 15, 423–434.
- [22] Oliver, M. A. and Badr, I. (1995), Determining the spatial scale of variation in soil radon concentration, *Mathematical Geology*, 27, 893–922.
- [23] Pilz, J. and Spöck, G. (2008), Why do we need and how should we implement Bayesian kriging methods, *Stoch. Environ. Res. Risk Assess.*, 22, 621–632.
- [24] Ren, Q. and Banerjee, S. (2013), Hierarchical factor models for large spatially misaligned data: a low-rank predictive process approach, *Biometrics*, 69, 19–30.
- [25] Spöck, G. (2012), Spatial sampling design with skew distributions: the special case of trans-Gaussian kriging, *Ninth International Geostatistical Congress*, Oslo, Norway, June 11–15, 2012, 20 pages.
- [26] Wibrin, M., Bogaert, P. and Fusbender, D. (2006), Combining categorical and continuous spatial information within the Bayesian maximum entropy paradigm, *Stochastic Environmental Research and Risk Assessment*, 20, 423–433.
- [27] Zhang, H. (2002), On estimation and prediction for spatial generalized linear mixed models, *Biometrics*, 58, 129–136.
- [28] Zhang, H. and El-Shaarawi, A. (2010), On spatial skew-Gaussian processes and applications, *Environmetrics*, 21, 33–47.