*Chilean Journal of Statistics*
Vol. 3, No. 2, September 2012, 157–170

Probabilistic and Inferential Aspects of Skew-Symmetric Models
Special Issue: "IV International Workshop in honour of Adelchi Azzalini's 60th Birthday"

# On the existence of some skew-normal stationary processes

Marco Minozzo[1,*] and Laura Ferracuti[1]

[1]Department of Economics, University of Verona, Verona, Italy

(Received: 03 October 2011 · Accepted in final form: 17 July 2012)

**Abstract**

Recently, some authors have introduced in the literature stationary stochastic processes in the time and spatial domains, whose finite-dimensional marginal distributions are multivariate skew-normal. In this paper, we show with a counter-example that the characterizations of these processes are not valid and so these processes do not exist. In particular, we exhibit through a marginalization argument that the set of finite-dimensional marginal distributions of these processes is not self-coherent. In addition, we point our attention to some valid constructions of stationary stochastic processes, which can be used to model skewed data.

**Keywords:** Autocorrelation function · Generalized linear mixed model · Geostatistics · Multivariate skew-normal distribution · Spatial process · Stationary process.

**Mathematics Subject Classification:** Primary 62M30 · Secondary 62H11.

## 1. Introduction

In the recent past, considerable attention has been devoted in the literature to multivariate versions of the skew-normal distribution, first systematically dealt with in the seminal paper by Azzalini (1985). Among the many multivariate versions appeared in the literature, the multivariate skew-normal distribution studied by Azzalini and Dalla Valle (1996) and by Azzalini and Capitanio (1999) seems to be the one that has received so far the widest attention by the statistical community. We say that a random vector $\boldsymbol{Z} = (Z_1, \ldots, Z_n)^\top$ has an extended skew-normal distribution (see, e.g., Azzalini, 2005) with parameters $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, $\boldsymbol{\alpha}$ and $\tau$, and we write $\boldsymbol{Z} \sim \mathrm{ESN}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}, \tau)$, if it has density of the form

$$f(\boldsymbol{z}) = \phi_n(\boldsymbol{z} - \boldsymbol{\mu}; \boldsymbol{\Sigma}) \frac{\Phi(\alpha_0 + \boldsymbol{\alpha}^\top \boldsymbol{D}^{-1}(\boldsymbol{z} - \boldsymbol{\mu}))}{\Phi(\tau)}, \quad \boldsymbol{z} \in \mathbb{R}^n, \tag{1}$$

*Corresponding author. Marco Minozzo, Department of Economics, University of Verona, Via dell'Artigliere 19, 37129, Verona, Italy. Email: marco.minozzo@univr.it

ISSN: 0718-7912 (print)/ISSN: 0718-7920 (online)

where $\boldsymbol{\mu} \in \mathbb{R}^n$ is a vector of location parameters, $\phi_n(\cdot;\boldsymbol{\Sigma})$ is the $n$-dimensional normal density with zero mean and (positive-definite) variance-covariance matrix $\boldsymbol{\Sigma}$ having elements $\sigma_{ij}$, $\Phi(\cdot)$ is the scalar N(0,1) distribution function, $\boldsymbol{D} = \mathrm{diag}(\sigma_{11}, \ldots, \sigma_{nn})^{1/2}$ is the diagonal matrix formed with the standard deviations of the scale matrix $\boldsymbol{\Sigma}$, $\boldsymbol{\alpha} \in \mathbb{R}^n$ is a vector of skewness parameters, and $\tau \in \mathbb{R}$ is an additional parameter. Moreover, $\alpha_0 = \tau(1 + \boldsymbol{\alpha}^\top \boldsymbol{R}\boldsymbol{\alpha})^{1/2}$, where $\boldsymbol{R}$ is the correlation matrix associated with $\boldsymbol{\Sigma}$, that is, $\boldsymbol{R} = \boldsymbol{D}^{-1}\boldsymbol{\Sigma}\boldsymbol{D}^{-1}$. Clearly, this distribution extends the multivariate normal distribution through the parameter vector $\boldsymbol{\alpha}$, and, for $\boldsymbol{\alpha} = 0$, it reduces to the latter. When $\tau = 0$, also $\alpha_0 = 0$ and Equation (1) reduces to

$$f(\boldsymbol{z}) = 2\phi_n(\boldsymbol{z} - \boldsymbol{\mu};\boldsymbol{\Sigma})\Phi(\boldsymbol{\alpha}^\top \boldsymbol{D}^{-1}(\boldsymbol{z} - \boldsymbol{\mu})), \quad \boldsymbol{z} \in \mathbb{R}^n. \tag{2}$$

In this case, we simply say that $\boldsymbol{Z}$ has a skew-normal distribution and we write, more concisely, $\boldsymbol{Z} \sim \mathrm{SN}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha})$.

The growing interest in these and other related multivariate families of distributions (see, e.g., Genton, 2004; Azzalini, 2005; Arellano-Valle and Azzalini, 2006) has led some authors to the specification of stochastic processes in the time and spatial domains, that is, with indexing parameter in $\mathbb{R}$ or in $\mathbb{R}^2$ (or in some suitable subset of it), having their univariate marginal distributions, or their multivariate finite-dimensional marginal distributions, to belong to some skew-normal family. Indeed, in many applications, the availability of such skew-normal stochastic processes is potentially of great importance. For example, confining ourselves to the spatial domain, in many environmental or ecological studies, the variable under investigation, observed at, say, $n$ sampling sites, is not Gaussian and may show some degree of skewness. In these cases, together with the spatial autocorrelation structure, it is important also to model the distribution of the data to account for the observed skewness. In particular, this is necessary if we are seeking for minimum mean square error predictions, which cannot be supplied in these non-Gaussian cases by standard kriging predictions. Let us notice that in these last situations, it is not sufficient to model the observed data using some multivariate skew-normal distribution. Although the observed data by itself could indeed be fitted using some multivariate skew-normal distribution of a given finite dimension $n$, the prediction problem would require the adoption of some stochastic process. In fact, for predicting the value assumed by the variable under investigation at an unobserved spatial location, we would need (for carrying out predictions with minimum mean square error) the conditional distribution of our variable at the unobserved spatial location, given the observed data at the $n$ sampling sites, which means that we would need the $(n+1)$-dimensional joint distribution of our variable at the $n+1$ spatial locations in question. Thus, since we usually need to carry out predictions at many (ideally infinite) unobserved spatial locations, the modeling of the observations with a multivariate distribution is not sufficient and we need to assume that the observed data is a partial realization of a stochastic process with its indexing parameter varying in some suitably infinite set.

Among others, skew-normal processes in the time domain have been put forward by Gualtierotti (2004, 2005), Pourahmadi (2007), and Corns and Satchell (2007), whereas skew-normal spatial processes have been defined by Kim and Mallick (2002, 2004, 2005), Kim et al. (2004), Naveau and Allard (2004), Allard and Naveau (2007), Zhang and El-Shaarawi (2010), and Hosseini et al. (2011). Although some of these works contain significant contributions, in this paper, we point our attention to the poor characterization of some of these skew-normal stochastic processes. Though at a first sight some of these characterizations might appear appealing, they are nevertheless not correct. Indeed, in some of these works, the characterization of the underlying skew-normal process mimics, wrongly, the definition of a Gaussian process. In these cases, then, it is possible to show with a counter-example that the characterizations are not valid and so that the advocated stochastic processes do not exist. The reason is that the assumption made in these characterizations that any finite collection of random variables making up the process has a multivariate skew-normal joint distribution is not compatible with the use of a (stationary) autocorrelation function to characterize their scale matrices. To clarify our point, let us underline that our negative result does not have anything to do with the existence of (stationary) stochastic processes having, for instance, univariate marginal distributions that are skew-normal. Apart from the trivial sequence of independent and identically distributed (i.i.d.) univariate skew-normal random variables, a well know example in this direction is given by the particular self exciting threshold autoregressive (SETAR) model

$$Z_t = -\delta_s |Z_{t-1}| + \varepsilon_t, \quad t = 0, \pm 1, \pm 2, \ldots, \tag{3}$$

where $\{\varepsilon_t \colon t = 0, \pm 1, \pm 2, \ldots\}$ are i.i.d. random variables with standard normal distribution, and $\delta_s$ is a real parameter. Here, a sufficient condition for the (strong) stationarity of the stochastic process $\{Z_t \colon t = 0, \pm 1, \pm 2, \ldots\}$ is that $|\delta_s| < 1$, and, following Azzalini (1986) and Tong (1990, p. 140), for $|\delta_s| < 1$, the univariate marginal stationary density given in Equation (3) is

$$\tilde{f}_{Z_t}(z) = \sqrt{\frac{2(1-\delta_s^2)}{\pi}} \, \exp\left( -\frac{1}{2}(1-\delta_s^2) z^2 \right) \Phi(-\delta_s z), \qquad -\infty < z < \infty,$$

that is, by $\tilde{f}_{Z_t}(z) = 2\phi_1(z; (1-\delta_s^2)^{-1})\Phi(-\delta_s z)$, where $\phi_1(\cdot; (1-\delta_s^2)^{-1})$ is the one-dimensional normal density with zero mean and variance $(1-\delta_s^2)^{-1}$. In other words, for each $t = 0, \pm 1, \pm 2, \ldots$, the random variables $Z_t$ are marginally distributed as skew-normals, precisely as $\mathrm{SN}_1(0, (1-\delta_s^2)^{-1}, -\delta_s/\sqrt{1-\delta_s^2})$. However, this fact does not imply that also the multivariate finite-dimensional marginal distributions given in Equation (3) are (multivariate) skew-normal. For instance, without loss of generality, consider the case in which $\delta_s = 1/\sqrt{2}$. Then,

$$\tilde{f}_{Z_t}(z) = 2\frac{1}{\sqrt{2\pi}} \, \frac{1}{\sqrt{2}} \, \exp\left( -\frac{z^2}{4} \right) \Phi\left( -\frac{z}{\sqrt{2}} \right),$$

and the bivariate marginal density of $Z_{t-1}$ and $Z_t$ is given, for each $t = 0, \pm 1, \pm 2, \ldots$, by

$$\begin{aligned}
\tilde{f}_{Z_{t-1}, Z_t}(z_1, z_2) &= \tilde{f}_{Z_t | Z_{t-1}}(z_2 | z_1) \tilde{f}_{Z_{t-1}}(z_1) \\
&= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{1}{\pi}} \, \exp\left( -\frac{1}{2}\left( z_2 + \frac{|z_1|}{\sqrt{2}} \right)^2 \right) \, \exp\left( -\frac{z_1^2}{4} \right) \, \Phi\left( -\frac{z_1}{\sqrt{2}} \right),
\end{aligned}$$

which does not belong to any of the commonly considered multivariate skew-normal families. Thus, Equation (3) specifies a stationary stochastic process having univariate, but not multivariate, marginal distributions that are skew-normal, according, for instance, to the definition in Azzalini and Dalla Valle (1996).

The paper is organized as follows. In Section 2, we first consider a counter-example showing the wrong characterization of a particular spatial skew-normal stationary stochastic process that appears in the literature and then provide some general discussion. In Section 3, in a geostatistical setting and following a hierarchical approach, we consider a simple way to characterize a stationary stochastic processes having univariate skew-normal conditional and marginal distributions, for which we can derive some of its moments. Lastly, in Section 4 we conclude with some remarks.

## 2. Skew-Normal Stationary Processes

In this section, to start our discussion on the existence of stationary stochastic processes having all their finite-dimensional marginal distributions (multivariate) skew-normal. Let us first focus on a particular characterization (which somehow mimics the characterization of a Gaussian process) appeared in the literature and show with a counter-example that it is faulty.

### 2.1 A spatial skew-normal stationary process that does not exist

Consider the following characterization of a spatial skew-normal stationary stochastic process as appeared in Kim and Mallick (2004). Indicating with $\{Z(\boldsymbol{x})\colon \boldsymbol{x} \in \mathbb{R}^2\}$ a spatial random function (let us assume here, without loss of generality, that this random function is defined all over the plane, and so that it is not restricted to a subregion of it), they assume that, for every fixed $n$, the vector $\boldsymbol{Z} = (Z(\boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_n))^\top$, where $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n$ are $n$ fixed spatial locations, has the following skew-normal density (Formula (3) of their paper)

$$f(\boldsymbol{z}) = 2\ \phi_n(\boldsymbol{z} - \boldsymbol{F}\boldsymbol{\beta}; \sigma^2\boldsymbol{K}_\theta)\,\Phi\left(\frac{\alpha}{\sigma}\ \mathbf{1}_n^\top\ (\boldsymbol{z} - \boldsymbol{F}\boldsymbol{\beta})\right), \quad \boldsymbol{z} \in \mathbb{R}^n, \tag{4}$$

where $\boldsymbol{F}$ is a known design matrix (of dimension $n \times q$) with full column rank, $\boldsymbol{\beta} \in \mathbb{R}^q$ are unknown regression parameters, $\sigma \in \mathbb{R}^+$ is a scale parameter, $\alpha \in \mathbb{R}$ is a skewness parameter, and $\mathbf{1}_n$ is the $n$-dimensional column vector of ones. Moreover, they also assume that $\boldsymbol{K}_\theta$ is a positive definite matrix (of dimension $n \times n$) with each entry given by $K_\theta(||\boldsymbol{x}_i - \boldsymbol{x}_j||)$, where $||\boldsymbol{x}_i - \boldsymbol{x}_j||$ denotes the Euclidean distance between $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, and $K_\theta(\cdot)$ is an isotropic spatial (stationary) autocorrelation function, depending on some (in general multivariate) parameter $\boldsymbol{\theta} \in \Theta$. This autocorrelation function $K_\theta(d)$, for $d \geq 0$, where $d$ is the (Euclidean) distance between two given and generic locations, is nonnegative, decreases monotonically with $d$. For $d = 0$, we have $K_\theta(0) = 1$, and $\lim_{d\to\infty} K_\theta(d) = 0$. In particular, they consider the power (general) exponential autocorrelation function

$$K_\theta(d) = \exp(-\nu d^{\theta_2}), \qquad d \geq 0,$$

where $\nu > 0$ and $\theta_2 \in (0, 2]$, which can also be expressed as $K_\theta(d) = \theta_1^{d^{\theta_2}}$, putting $\theta_1 = \exp(-\nu)$. In passing, note that for the stochastic process $\{Z(\boldsymbol{x})\colon \boldsymbol{x} \in \mathbb{R}^2\}$ to be stationary, we must have at least $\boldsymbol{\beta} = 0$; and that the scale matrix $\sigma^2\boldsymbol{K}_\theta$ is not, in general, the variance-covariance matrix of $\boldsymbol{Z} = (Z(\boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_n))^\top$. Indeed, Kim and Mallick (2004) used the function $K_\theta(d)$ to characterize a matrix of parameters of the process. The function $K_\theta(d)$ would represent the spatial autocorrelation function of the process $\{Z(\boldsymbol{x})\colon \boldsymbol{x} \in \mathbb{R}^2\}$ only in the case in which $\alpha = 0$.

Though this definition of a skew-normal spatial stochastic process might appear appealing, it is nevertheless not correct. The reason is that, although the marginal densities of the above density in Equation (4) are still skew-normal, they are not of the same form. Indeed, the assumption that any finite collection of random variables making up the process should have a multivariate skew-normal joint distribution is at clash with the adoption of a spatial (stationary) autocorrelation function to characterize their scale matrices.

To show our point, consider three distinct and fixed spatial locations $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3 \in \mathbb{R}^2$. Let $K_\theta(\cdot)$ be a given (isotropic) spatial autocorrelation function as before, and let $k_{12} \doteq K_\theta(||\boldsymbol{x}_1 - \boldsymbol{x}_2||)$, $k_{13} \doteq K_\theta(||\boldsymbol{x}_1 - \boldsymbol{x}_3||)$ and $k_{23} \doteq K_\theta(||\boldsymbol{x}_2 - \boldsymbol{x}_3||)$. Then, being $\{Z(\boldsymbol{x})\colon \boldsymbol{x} \in \mathbb{R}^2\}$ a spatial random function, applying Formula (3) of Kim and Mallick (2004), where we consider for simplicity $\boldsymbol{\beta} = 0$ and $\sigma^2 = 1$, the joint distribution of the vector $(Z(\boldsymbol{x}_1), Z(\boldsymbol{x}_2), Z(\boldsymbol{x}_3))^\top$ should be multivariate skew-normal with density

$$f(z_1, z_2, z_3) = 2\ \phi_3\left(\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}; \begin{bmatrix} 1 & k_{12} & k_{13} \\ k_{12} & 1 & k_{23} \\ k_{13} & k_{23} & 1 \end{bmatrix}\right) \Phi\big(\alpha(z_1 + z_2 + z_3)\big), \tag{5}$$

whereas the joint distribution of $(Z(\boldsymbol{x}_1), Z(\boldsymbol{x}_2))^\top$ should be multivariate skew-normal with density

$$f(z_1, z_2) = 2\ \phi_2\left(\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}; \begin{bmatrix} 1 & k_{12} \\ k_{12} & 1 \end{bmatrix}\right) \Phi\big(\alpha(z_1 + z_2)\big). \tag{6}$$

Now, if Formula (3) of Kim and Mallick (2004) was indeed characterizing the finite-dimensional marginal distributions of a stochastic process, we should be able (at least) to obtain the joint distribution given in Equation (6) of $(Z(\boldsymbol{x}_1), Z(\boldsymbol{x}_2))^\top$ by marginalization of the joint distribution given in Equation (5) of $(Z(\boldsymbol{x}_1), Z(\boldsymbol{x}_2), Z(\boldsymbol{x}_3))^\top$ with respect to $Z(\boldsymbol{x}_3)$. Starting from Equation (5), using, for instance, the marginalization Formulas (19) and (20) of Azzalini (2005), with $\tau = 0$, $\boldsymbol{\xi} = 0$ and $\boldsymbol{\Omega} = \bar{\boldsymbol{\Omega}}$ (due to the presence of some misprints, we cannot use Formula (7) of Kim and Mallick, 2004), the distribution of $(Z(\boldsymbol{x}_1), Z(\boldsymbol{x}_2))^\top$ is given by

$$f(z_1, z_2) = 2\ \phi_2\left(\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}; \begin{bmatrix} 1 & k_{12} \\ k_{12} & 1 \end{bmatrix}\right) \Phi\left(\boldsymbol{\alpha}_{1(2)}^\top \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}\right), \tag{7}$$

where

$$\begin{aligned} \boldsymbol{\alpha}_{1(2)} &= \frac{\begin{bmatrix} \alpha \\ \alpha \end{bmatrix} + \dfrac{\alpha}{1 - k_{12}^2} \begin{bmatrix} 1 & -k_{12} \\ -k_{12} & 1 \end{bmatrix} \begin{bmatrix} k_{13} \\ k_{23} \end{bmatrix}}{\sqrt{1 + \alpha^2\left(1 - \dfrac{k_{13}^2 - 2k_{13}k_{12}k_{23} + k_{23}^2}{1 - k_{12}^2}\right)}} \\ &= \frac{1}{\sqrt{\dfrac{1 - k_{12}^2 + \alpha^2(1 - k_{12}^2 - k_{13}^2 - k_{23}^2 + 2k_{13}k_{12}k_{23})}{1 - k_{12}^2}}} \begin{bmatrix} \alpha + \dfrac{\alpha(k_{13} - k_{12}k_{23})}{1 - k_{12}^2} \\ \alpha + \dfrac{\alpha(k_{23} - k_{12}k_{13})}{1 - k_{12}^2} \end{bmatrix}. \end{aligned}$$

So, for Equations (6) and (7) to be equal, we should have

$$\alpha(z_1 + z_2) = \boldsymbol{\alpha}_{1(2)}^\top [z_1 \quad z_2]^\top, \tag{8}$$

and, with a little algebra, we see that a necessary condition for Equation (8) to be true is that the elements of $\boldsymbol{\alpha}_{1(2)}$ must be equal, that is,

$$\alpha + \frac{\alpha(k_{13} - k_{12}k_{23})}{1 - k_{12}^2} = \alpha + \frac{\alpha(k_{23} - k_{12}k_{13})}{1 - k_{12}^2},$$

which is guaranteed whenever

$$(k_{13} - k_{23})(1 + k_{12}) = 0, \tag{9}$$

that is, whenever either $k_{12} = -1$ or $k_{13} = k_{23}$. Since (one or both of) these two conditions should be satisfied for every choice of $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$, it follows that 'admissible' isotropic spatial autocorrelation functions $K_\theta(d)$ are found among the constant functions of the Euclidean distance $d$. Thus, there are no spatial autocorrelation functions $K_\theta(d)$ which are nonnegative and monotonically decreasing with $d$, with $K_\theta(0) = 1$ and $\lim_{d\to\infty} K_\theta(d) = 0$, for which Equation (9), and so Equation (8), hold true.

This counter-example shows that the characterization of Kim and Mallick (2004) is improper and so the advocated spatial skew-normal stationary process does not exist. In the context of their paper, this means, among other things, that the predicted values of the process at unobserved spatial locations (obtained through a Bayesian Markov chain Monte Carlo (MCMC) algorithm) are not self-coherent.

### 2.2 Other examples of skew-normal stationary processes

A problem very much similar to the poor characterization of a spatial stationary skew-normal process by Kim and Mallick (2004) can be found in the paper by Allard and Naveau (2007); see also Naveau and Allard (2004) and Allard and Soubeyrand (2012). Basically, in Allard and Naveau (2007), to characterize a spatial stochastic process $\{Z(\boldsymbol{x})\colon \boldsymbol{x} \in \mathbb{R}^2\}$, the authors assume that for any given set of $n$ spatial locations $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ in $\mathbb{R}^2$, the random variables $Z(\boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_n)$ are jointly distributed (see Formula (5) of Allard and Naveau, 2007) as a multivariate closed skew-normal distribution as defined by González-Farías et al. (2004a) and González-Farías et al. (2004b), with the elements of the scale matrix specified through a (stationary) covariance function. In particular, given $n$ spatial locations $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$, they assume that the $n$-dimensional random vector $\boldsymbol{Z} = (Z(\boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_n))^\top$ has a multivariate closed skew-normal distribution, denoted by $\mathrm{CSN}_{n,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{D}_c, \boldsymbol{\nu}, \boldsymbol{\Delta})$, with density

$$f(\boldsymbol{z}) = \frac{1}{\Phi_m(0; \boldsymbol{\nu}, \boldsymbol{\Delta} + \boldsymbol{D}_c^\top \boldsymbol{\Sigma} \boldsymbol{D}_c)} \phi_n(\boldsymbol{z}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \Phi_m(\boldsymbol{D}_c^\top (\boldsymbol{z} - \boldsymbol{\mu}); \boldsymbol{\nu}, \boldsymbol{\Delta}), \quad \boldsymbol{z} \in \mathbb{R}^n,$$

where $m$ is an integer greater than 0; $\phi_n(\cdot; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\Phi_n(\cdot; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ are the density and the distribution function, respectively, of the $n$-dimensional normal distribution with mean vector $\boldsymbol{\mu}$ and variance-covariance matrix $\boldsymbol{\Sigma}$; $\boldsymbol{\mu} = \mu_0 \mathbf{1}_n$, where $\mu_0 \in \mathbb{R}$; $\boldsymbol{\nu} = 0$ is the null vector in $\mathbb{R}^m$; and $\boldsymbol{D}_c \in \mathbb{R}^{n\times m}$ is given by $\boldsymbol{D}_c = d\boldsymbol{A}$, where $d \in \mathbb{R}$ and $\boldsymbol{A} \in \mathbb{R}^{n\times m}$ is a non null matrix; $\boldsymbol{\Sigma} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{\Delta} \in \mathbb{R}^{m\times m}$ are positive-definite matrices both built using some (real valued, stationary and ergodic) spatial covariance function $c(\boldsymbol{h}) = \mathrm{Cov}\big(Z(\boldsymbol{x}), Z(\boldsymbol{x} + \boldsymbol{h})\big)$. Of course, in this case, we can write $c(\boldsymbol{h}) = c(0)\rho(\boldsymbol{h})$, where $c(0) = \mathrm{Var}\big[Z(\boldsymbol{x})\big]$ and $\rho(\boldsymbol{h}) = \mathrm{Corr}\big(Z(\boldsymbol{x}), Z(\boldsymbol{x}+\boldsymbol{h})\big)$ is some real valued spatial autocorrelation function for which $\rho(0) = 1$ and $\rho(\boldsymbol{h}) \to 0$, as $||\boldsymbol{h}|| \to \infty$.

To simplify our discussion, consider the case in which $m = 1$, $\boldsymbol{\mu} = 0$, $\boldsymbol{\Sigma}$ is a correlation matrix with ones along the diagonal, $\boldsymbol{\Delta} = 1$, $\boldsymbol{A} = \mathbf{1}_n$, and the autocorrelation function $\rho(\boldsymbol{h})$ is isotropic (actually, Allard and Naveau, 2007, consider the constraint $\boldsymbol{A}^\top \boldsymbol{\Sigma} \boldsymbol{A} = \boldsymbol{\Delta}$, but this does not change the rationale of our discussion). In this case, it is immediate to see that the $n$-dimensional random vector $\boldsymbol{Z} = (Z(\boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_n))^\top$ is distributed as a $\mathrm{CSN}_{n,1}(0, \boldsymbol{\Sigma}, d\mathbf{1}_n, 0, 1)$ with density

$$f(\boldsymbol{z}) = 2\phi_n(\boldsymbol{z}; 0, \boldsymbol{\Sigma}) \Phi_1(d\mathbf{1}_n^\top \boldsymbol{z}; 0, 1), \quad \boldsymbol{z} \in \mathbb{R}^n,$$

that is, as the $\mathrm{SN}_n(0, \boldsymbol{\Sigma}, d\mathbf{1}_n)$ distribution of Azzalini (2005) and Azzalini and Capitanio (1999) as in Equation (2). Now, by considering three distinct spatial locations $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ and $\boldsymbol{x}_3$, the joint distribution of the random vector $(Z(\boldsymbol{x}_1), Z(\boldsymbol{x}_2), Z(\boldsymbol{x}_3))^\top$ should be multivariate skew-normal with density as in Equation (5). So, following exactly the same argument and using the same marginalization formulas considered in Section 2.1, we can conclude that, for $m = 1$, also some instances of the spatial skew-normal processes advocated by Allard and Naveau (2007), in which the matrices $\boldsymbol{\Sigma}$ and $\boldsymbol{\Delta}$ are built using a spatial autocorrelation function $\rho(\boldsymbol{h})$, such that $\rho(0) = 1$ and $\rho(\boldsymbol{h}) \to 0$, as $||\boldsymbol{h}|| \to \infty$, do not exist. That is, we can conclude that the class of processes advocated by these authors is not properly defined. It is not known whether, for $m = 1$ or $m > 1$, a proper choice of the parameter values might lead to valid coherent processes (apart from the trivial case in which the finite dimensional marginal distributions are multivariate normal), in particular for $m$ equal to the (fixed) number of sampling sites where observations have been gathered, as used in the application by Allard and Naveau (2007).

In a spatial hierarchical framework, adopting the closed skew-normal distribution, a similar approach to the definition of the latent random field has been taken by Hosseini et al. (2011). Also in this case, it is possible to show that this latent random field is not properly defined and so that it does not exist.

Another characterization somehow similar in spirit to the characterization of Kim and Mallick (2004) has been put forward by Gualtierotti (2005) (see also Gualtierotti, 2004, who introduced a skew-normal stochastic process) with indexing parameter varying in the real line, in the context of statistical communication theory. Essentially, this stochastic process is characterized assuming that all its finite-dimensional marginal distributions belong to a particular family of multivariate skew-normal distributions, which, in turn, is characterized following a construction similar to that of the multivariate skew-normal distribution of Arellano-Valle et al. (2002). Though in Gualtierotti (2005) no claim is made about the stationarity of the process, if we tried to build the scale matrices of its finite-dimensional marginal distributions using a (stationary) covariance function decaying to zero, as the separation distance goes to infinity, then we would still get in trouble. In fact, it is easy to see that the family of skew-normal distributions put forward by Gualtierotti (2005) overlaps with the family of skew-normal distributions of Azzalini and Capitanio (1999), and so, following exactly the same argument used in Section 2.1, that, in general, in the class of skew-normal processes proposed by Gualtierotti the scale matrices cannot be constructed using a stationary covariance function.

### 2.3 Skew-normal stationary processes: some negative results

Consider now the following general question. Are there strictly stationary stochastic processes having an autocorrelation function decaying to zero, as the separation distance goes to infinity, for which all the finite-dimensional marginal distributions are (multivariate) skew-normal? Somehow, similar queries were posed by Pourahmadi (2007), which in the context of autoregressive and moving average models tries to argue that there is a considerable trade-off between stationarity and skewness. As far as we are concerned, here, by recalling and adapting to our case some of the arguments of Pourahmadi (2007), we argue that we might encounter some problems.

To this aim, consider a real-valued stochastic process $Z$, with indexing parameter $\boldsymbol{x}$ varying in some indexing set (which might be $\mathbb{R}$, $\mathbb{R}^2$, or some subset of it), for which, for every integer $n \geq 1$, and every set of indexing values $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$, the $n$-dimensional random vector $\boldsymbol{Z} = (Z(\boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_n))^\top$ has an extended skew-normal distribution $\mathrm{ESN}_n(\boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n, \boldsymbol{\alpha}_n, \tau)$ as in Equation (1), where $\boldsymbol{\mu}_n$, $\boldsymbol{\Sigma}_n$ and $\boldsymbol{\alpha}_n$ depend on $n$, and for which the hypothesis of strict stationarity holds. That is, for every integer $n \geq 1$, and every set of indexing values $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ and $\boldsymbol{x}_1 + \boldsymbol{h}, \ldots, \boldsymbol{x}_n + \boldsymbol{h}$, the $n$-dimensional random vector $(Z(\boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_n))^\top$ has the same distribution as $(Z(\boldsymbol{x}_1 + \boldsymbol{h}), \ldots, Z(\boldsymbol{x}_n + \boldsymbol{h}))^\top$. In this case, it follows that $Z$ is also second-order stationary, that is, for all couples of indexing values $\boldsymbol{x}$ and $\boldsymbol{x} + \boldsymbol{h}$:

i) $\mathrm{E}\big[Z(\boldsymbol{x})\big]$ is a constant that does not depend on $\boldsymbol{x}$;
ii) $\mathrm{Cov}\big(Z(\boldsymbol{x}+\boldsymbol{h}), Z(\boldsymbol{x})\big)$ is a function of $\boldsymbol{h}$ that does not depend on $\boldsymbol{x}$.

Remember that for any given random vector $\boldsymbol{Z} = (Z(\boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_n))^\top$ of dimension $n$, since it must have an extended skew-normal distribution, it is well known that we can express $\mathrm{E}\big[\boldsymbol{Z}\big]$ and $\mathrm{Cov}\big(\boldsymbol{Z}\big)$ in terms of $\boldsymbol{\mu}_n$, $\boldsymbol{\Sigma}_n$, $\boldsymbol{\alpha}_n$ and $\tau$ through the formulas

$$\mathrm{E}\big[\boldsymbol{Z}\big] = \boldsymbol{\mu}_n + \zeta_1(\tau)\boldsymbol{D}_n\boldsymbol{\delta}_n \qquad \text{and} \qquad \mathrm{Cov}\big(\boldsymbol{Z}\big) = \boldsymbol{\Sigma}_n + \zeta_2(\tau)\boldsymbol{D}_n\boldsymbol{\delta}_n\boldsymbol{\delta}_n^\top\boldsymbol{D}_n,$$

where $\boldsymbol{\delta}_n = (1 + \boldsymbol{\alpha}_n^\top\boldsymbol{R}_n\boldsymbol{\alpha}_n)^{-1/2}\boldsymbol{R}_n\boldsymbol{\alpha}_n$, and

$$\zeta_1(\tau) = \frac{\phi(\tau)}{\Phi(\tau)}, \qquad \zeta_2(\tau) = -\zeta_1(\tau)\{\tau + \zeta_1(\tau)\},$$

where, as usual, $\phi(\cdot)$ is the one-dimensional standard normal density of mean zero and variance one.

Now, following Pourahmadi (2007), for the stochastic process $Z$ to be strictly stationary, it must be that at least all the shape parameters and variances of its univariate marginal distributions should be the same. Then, for each $n$-dimensional random vector $\boldsymbol{Z} = (Z(\boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_n))^\top$, this would force $\boldsymbol{\delta}_n$ and $\boldsymbol{D}_n$ to be of the form $\boldsymbol{\delta}_n = \delta_0\boldsymbol{1}_n$, $\boldsymbol{D}_n = \gamma_0\boldsymbol{I}_n$, where $\boldsymbol{I}_n$ is the identity matrix of dimension $n$, and hence would force the scale matrix $\boldsymbol{\Sigma}_n$ and the covariance matrix $\mathrm{Cov}\big(\boldsymbol{Z}\big)$ of the random vector $\boldsymbol{Z}$ to be equal, up to an additive constant (which is equal to zero only when $\boldsymbol{\delta}_n$ and so $\boldsymbol{\alpha}_n$ are equal to zero). For instance, if the indexing parameter of $Z$ varies in the integers and for some $n$ the random vector $\boldsymbol{Z} = (Z(\boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_n))^\top$ is such that the indexing values $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are equally spaced, then, since the covariance matrix of $\boldsymbol{Z}$ should have a Toeplitz structure (with constant entries along diagonals), also the scale matrix $\boldsymbol{\Sigma}_n$ should have a Toeplitz structure. Thus, since $\mathrm{Corr}\big(Z(\boldsymbol{x}+\boldsymbol{h}), Z(\boldsymbol{x})\big)$ and $\mathrm{Cov}\big(Z(\boldsymbol{x}+\boldsymbol{h}), Z(\boldsymbol{x})\big)$ differ only by a multiplicative constant, we can conclude that there cannot be strictly stationary stochastic processes for which all the finite-dimensional marginal distributions are skew-normal, the autocorrelation $\mathrm{Corr}\big(Z(\boldsymbol{x}+\boldsymbol{h}), Z(\boldsymbol{x})\big)$ decays to zero, as $||\boldsymbol{h}|| \to \infty$, and, at the same time, the scale matrices $\boldsymbol{\Sigma}_n$ are constructed using a covariance function decaying to zero, as $||\boldsymbol{h}|| \to \infty$. Though processes having an autocorrelation function decreasing to zero might be thought of as a particular subclass of processes, they are nevertheless extremely important for real applications, where it is often necessary to recover the probabilistic structure of the process from a single (partial) realization of it.

Notice that this negative result is in agreement with a remark by Zhang and El-Shaarawi (2010). In Section 2 of their paper, they briefly discussed the spatial process

$$Z(\boldsymbol{x}) = \delta|X_0| + \sqrt{1-\delta^2}\ X(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{R}^2,$$

where $X(\boldsymbol{x})$ is a stationary Gaussian process with mean zero and variance one and $X_0$ is a scalar random variable with standard normal distribution, independent of the process $X(\boldsymbol{x})$, and $-1 \le \delta \le 1$. For the Gaussian process $X(\boldsymbol{x})$, let $\rho(\boldsymbol{h})$ be its (real valued) autocorrelation function for which $\rho(0) = 1$ and $\rho(\boldsymbol{h}) \to 0$, as $||\boldsymbol{h}|| \to \infty$. By construction, the process $Z(\boldsymbol{x})$ is weakly and also strongly stationary, and all finite-dimensional marginal distributions are multivariate skew-normal according to the definition in Azzalini and Dalla Valle (1996). In particular, the (stationary) univariate marginal distribution of the process $Z(\boldsymbol{x})$ is $\mathrm{SN}_1(0, 1, \delta/\sqrt{1-\delta^2})$. Whereas, in general, for any given set of $n$ spatial locations $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$, the $n$-dimensional random vector $\boldsymbol{Z} = (Z(\boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_n))^\top$ has distribution

$\mathrm{SN}_n(0, \boldsymbol{\Sigma}, \boldsymbol{\alpha})$, where $\boldsymbol{\Sigma} = (1-\delta^2)\boldsymbol{\Psi} + \delta^2$,

$$\boldsymbol{\alpha}^\top = \delta(1-\delta^2)^{-1/2}[(1-\delta^2) + \delta^2 \mathbf{1}_n^\top \boldsymbol{\Psi}^{-1}\mathbf{1}_n]^{-1/2}\mathbf{1}_n^\top \boldsymbol{\Psi}^{-1},$$

and $\boldsymbol{\Psi}$ is the $n \times n$ correlation matrix of the $n$-dimensional random vector $(X(\boldsymbol{x}_1), \ldots, X(\boldsymbol{x}_n))^\top$ built using the autocorrelation function $\rho(\boldsymbol{h})$.

For this process, the argument used in the counter-example in Section 2.1 does not lead to any inconsistency, since its finite-dimensional marginal distributions are self-coherent. However, as pointed out by Zhang and El-Shaarawi (2010), this process is useless for many practical purposes, since it is not ergodic and has an autocorrelation function

$$\mathrm{Corr}\big(Z(\boldsymbol{x}), Z(\boldsymbol{x}+\boldsymbol{h})\big) = \frac{\delta^2(1-2/\pi) + (1-\delta^2)\rho(\boldsymbol{h})}{\delta^2(1-2/\pi) + (1-\delta^2)},$$

that does not decay to zero. It is easy to see that, when $||\boldsymbol{h}|| \to \infty$, and $\rho(\boldsymbol{h}) \to 0$, the correlation converges to a (strictly) positive value that depends on the skewness parameter $\delta$. Finally, it might be the case to note that if we do not require any particular condition, we have a very simple way to define a stochastic process (made up of an infinite, but countable, number of random variables) having all its finite-dimensional marginal distributions belonging to some family of multivariate skew-normal distributions closed under marginalization (as the families defined by Equations (1) and (2)). For example, with respect to the family defined by Equation (2), consider the following infinite sequence of multivariate distributions of increasing dimension

$$\mathrm{SN}_1(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1, \boldsymbol{\alpha}_1), \quad \mathrm{SN}_2(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2, \boldsymbol{\alpha}_2), \quad \mathrm{SN}_3(\boldsymbol{\mu}_3, \boldsymbol{\Sigma}_3, \boldsymbol{\alpha}_3), \ldots$$

such that, for every $n$, $\mathrm{SN}_{n-1}(\boldsymbol{\mu}_{n-1}, \boldsymbol{\Sigma}_{n-1}, \boldsymbol{\alpha}_{n-1})$ can be obtained by marginalization from $\mathrm{SN}_n(\boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n, \boldsymbol{\alpha}_n)$. Though it might be somewhat impractical to work with such an infinite sequence, this sequence of distributions properly defines the (probabilistic) law of a stochastic process with indexing parameter in some countable set.

## 3. A Hierarchical Geostatistical Skew-Normal Stationary Process

In accordance with the remarks made in the Introduction, the negative results of the previous sections do not prevent the existence of stationary stochastic processes with a covariance function decreasing to zero, as the separation distance goes to infinity, having univariate marginal distributions that are skew-normal. In addition to the SETAR model recalled in Section 1, another example of a stationary stochastic process having a covariance function decreasing to zero, and univariate skew-normal marginal distributions, has been given, in the spatial domain, by Zhang and El-Shaarawi (2010), exploiting one of the stochastic characterizations of the skew-normal distribution. As in the SETAR model, even in this case, the finite-dimensional marginal distributions (with a number of dimensions greater than one) of the process do not belong to any of the commonly considered families of multivariate skew-normal distributions. In the continuous time domain, another interesting characterization of a skew-normal process has instead been advanced by Corns and Satchell (2007) to tackle the problem of pricing European options.

In the spatial domain, a strongly stationary geostatistical stochastic process, having (univariate) skew-normal conditional and marginal distributions, can be defined by building on a latent stationary Gaussian process, adopting the hierarchical approach of Diggle et al. (1998); see also Diggle and Ribeiro (2007).

Let $z(\boldsymbol{x}_i)$, for $i = 1, \dots, n$, be a set of geo-referenced data measurements relative to a regionalized variable, gathered at $n$ spatial locations $\boldsymbol{x}_i$. This regionalized variable is seen as a partial realization of a random function $\{Z(\boldsymbol{x})\colon \boldsymbol{x} \in \mathbb{R}^2\}$. Following Minozzo and Fruttini (2004) and Ferracuti (2005), let $\{Y(\boldsymbol{x})\colon \boldsymbol{x} \in \mathbb{R}^2\}$ be an unobservable mean zero stationary Gaussian process with $\mathrm{Var}\big[Y(\boldsymbol{x})\big] = \varsigma^2 > 0$ and $\mathrm{Cov}\big(Y(\boldsymbol{x}), Y(\boldsymbol{x}+\boldsymbol{h})\big) = \varsigma^2\rho(\boldsymbol{h})$, where $\rho(\boldsymbol{h})$ is a real valued spatial autocorrelation function for which $\rho(0) = 1$ and $\rho(\boldsymbol{h}) \to 0$, as $||\boldsymbol{h}|| \to \infty$. We assume that, given the latent process $\{Y(\boldsymbol{x})\colon \boldsymbol{x} \in \mathbb{R}^2\}$, the random variables $Z(\boldsymbol{x})$, for $\boldsymbol{x} \in \mathbb{R}^2$, are conditionally mutually independent, and that the conditional distribution of $Z(\boldsymbol{x})$, for each $\boldsymbol{x}$, depends only on the random variable $Y(\boldsymbol{x})$. In particular, we assume that, conditionally on the latent process $\{Y(\boldsymbol{x})\colon \boldsymbol{x} \in \mathbb{R}^2\}$, the random variables $Z(\boldsymbol{x})$, for any given $\boldsymbol{x}$, have conditional density $f_{Z|Y}(z; M(\boldsymbol{x}))$, specified by the value of the conditional expectation $M(\boldsymbol{x}) = \mathrm{E}\big[Z(\boldsymbol{x})\big|Y(\boldsymbol{x})\big]$, that is, $Z(\boldsymbol{x})\big|Y(\boldsymbol{x}) \sim f_{Z|Y}(z; M(\boldsymbol{x}))$, and that $h(M(\boldsymbol{x})) = \beta + Y(\boldsymbol{x})$, for some real parameter $\beta$ and some known link function $h(\cdot)$.

In the case in which $f_{Z|Y}(z; M(\boldsymbol{x}))$ is skew-normal and $h(\cdot)$ is a translation by a constant, it is easy to verify that the process $\{Z(\boldsymbol{x})\colon \boldsymbol{x} \in \mathbb{R}^2\}$ is second-order, and also strongly stationary. In particular, assume that

$$Z(\boldsymbol{x}) = \beta + Y(\boldsymbol{x}) + \omega S(\boldsymbol{x}), \tag{10}$$

where $\omega \in \mathbb{R}^+$ and $S(\boldsymbol{x})$, for $\boldsymbol{x} \in \mathbb{R}^2$, are mutually independently distributed as skew-normals $\mathrm{SN}_1(0, 1, \alpha)$ such that, for every $\boldsymbol{x} \in \mathbb{R}^2$, the density of $S(\boldsymbol{x})$ is given by

$$f_S(s) = 2\phi_1(s; 1)\Phi(\alpha s), \qquad -\infty < s < \infty,$$

where $\alpha \in \mathbb{R}$. In assuming Equation (10), we have implicitly chosen the link function $h(M(\boldsymbol{x})) \doteq M(\boldsymbol{x}) - \omega(2/\pi)^{1/2}\alpha/(1+\alpha^2)^{1/2}$. Thus, we have that, for every $\boldsymbol{x} \in \mathbb{R}^2$, $Z(\boldsymbol{x})|Y(\boldsymbol{x}) \sim \mathrm{SN}_1(\beta + Y(\boldsymbol{x}), \omega^2, \alpha)$. Now, although the random variables $Z(\boldsymbol{x})$, for $\boldsymbol{x} \in \mathbb{R}^2$, are conditionally distributed as skew-normals, and also (see, e.g., Azzalini, 2005) marginally distributed as $\mathrm{SN}_1(\beta, \varsigma^2 + \omega^2, \alpha\omega/\sqrt{\varsigma^2(1+\alpha^2)+\omega^2})$, the other (multivariate) finite-dimensional marginal distributions of the process $\{Z(\boldsymbol{x})\colon \boldsymbol{x} \in \mathbb{R}^2\}$ are not skew-normal; see also the comments in Gupta and Chen (2004). Let us mention that the spatial process $Z(\boldsymbol{x})$ just defined can be obtained as a particular instance (that is, as a subclass) of the process given in Formula (4) of Zhang and El-Shaarawi (2010).

To derive the (stationary) autocorrelation structure of the process $\{Z(\boldsymbol{x})\colon \boldsymbol{x} \in \mathbb{R}^2\}$, consider first that, since, for every $\boldsymbol{x} \in \mathbb{R}^2$,

$$\mathrm{E}\big[S(\boldsymbol{x})\big] = \left(\frac{2}{\pi}\right)^{1/2} \frac{\alpha}{(1+\alpha^2)^{1/2}}, \qquad \mathrm{Var}\big[S(\boldsymbol{x})\big] = 1 - \frac{2}{\pi}\frac{\alpha^2}{(1+\alpha^2)},$$

it follows that, for every $\boldsymbol{x} \in \mathbb{R}^2$,

$$\begin{aligned}
\mathrm{E}\big[Z(\boldsymbol{x})\big|Y(\boldsymbol{x})\big] &= \beta + Y(\boldsymbol{x}) + \omega\left(\frac{2}{\pi}\right)^{1/2}\frac{\alpha}{(1+\alpha^2)^{1/2}},\\
\mathrm{Var}\big[Z(\boldsymbol{x})\big|Y(\boldsymbol{x})\big] &= \omega^2\left[1 - \frac{2}{\pi}\frac{\alpha^2}{(1+\alpha^2)}\right].
\end{aligned}$$

Then, with some algebra, we can derive both the autocovariance function and the variogram of the process $\{Z(\boldsymbol{x})\colon \boldsymbol{x} \in \mathbb{R}^2\}$ using standard techniques. For instance, for $\boldsymbol{h} \neq 0$, the variogram is given by

$$
\begin{aligned}
\gamma(\boldsymbol{h}) &= \frac{1}{2}\mathrm{Var}\big[Z(\boldsymbol{x}+\boldsymbol{h}) - Z(\boldsymbol{x})\big] \\
&= \frac{1}{2}\mathrm{E}\big[\mathrm{Var}\big[Z(\boldsymbol{x})|Y(\boldsymbol{x})\big]\big] + \frac{1}{2}\mathrm{E}\big[\mathrm{Var}\big[Z(\boldsymbol{x}+\boldsymbol{h})|Y(\boldsymbol{x}+\boldsymbol{h})\big]\big] \\
&\quad + \frac{1}{2}\mathrm{Var}\big[\mathrm{E}\big[Z(\boldsymbol{x})|Y(\boldsymbol{x})\big]\big] + \frac{1}{2}\mathrm{Var}\big[\mathrm{E}\big[Z(\boldsymbol{x}+\boldsymbol{h})|Y(\boldsymbol{x}+\boldsymbol{h})\big]\big] \\
&\quad - \mathrm{Cov}\big(\mathrm{E}\big[Z(\boldsymbol{x})|Y(\boldsymbol{x})\big],\ \mathrm{E}\big[Z(\boldsymbol{x}+\boldsymbol{h})|Y(\boldsymbol{x}+\boldsymbol{h})\big]\big) \\
&= \frac{1}{2}\omega^2\left[1 - \frac{2}{\pi}\frac{\alpha^2}{(1+\alpha^2)}\right] + \frac{1}{2}\omega^2\left[1 - \frac{2}{\pi}\frac{\alpha^2}{(1+\alpha^2)}\right] + \frac{1}{2}\varsigma^2 + \frac{1}{2}\varsigma^2 - \rho(\boldsymbol{h})\varsigma^2 \\
&= \omega^2\left[1 - \frac{2}{\pi}\frac{\alpha^2}{(1+\alpha^2)}\right] + \varsigma^2\big(1 - \rho(\boldsymbol{h})\big),
\end{aligned}
$$

which is discontinuous at zero, that is, $\gamma(0) \neq \gamma(0^+)$, and we have

$$
\gamma(0^+) = \omega^2\left[1 - \frac{2}{\pi}\frac{\alpha^2}{(1+\alpha^2)}\right], \qquad \lim_{||\boldsymbol{h}||\to\infty}\gamma(\boldsymbol{h}) = \omega^2\left[1 - \frac{2}{\pi}\frac{\alpha^2}{(1+\alpha^2)}\right] + \varsigma^2.
$$

For the process given in Equation (10), it is easy to verify that, for $\boldsymbol{h} \neq 0$, the autocovariance function is given by

$$
\mathrm{Cov}\big(Z(\boldsymbol{x}+\boldsymbol{h}), Z(\boldsymbol{x})\big) = \varsigma^2\rho(\boldsymbol{h}),
$$

and so the autocorrelation function converges to zero, as $||\boldsymbol{h}|| \to \infty$.

Though, as we have already noticed, the multivariate finite-dimensional marginal distributions of the process $\{Z(\boldsymbol{x})\colon \boldsymbol{x} \in \mathbb{R}^2\}$ are not skew-normal (in the sense of Equation (2)), it is easy to see that they are closed skew-normal, according to the definition of González-Farías et al. (2004a). To see this, consider $n$ spatial locations $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$, and the corresponding $n$-dimensional random vector $\boldsymbol{Z} = (Z(\boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_n))^\top$. Recalling that for any given $\boldsymbol{x} \in \mathbb{R}^2$ we can write $Z(\boldsymbol{x}) = \beta + Y(\boldsymbol{x}) + \omega S(\boldsymbol{x})$, the vector $\boldsymbol{Z}$ can be represented by

$$
\boldsymbol{Z} = \beta\mathbf{1}_n + \boldsymbol{Y} + \boldsymbol{D}_\omega\boldsymbol{S} = \boldsymbol{W} + \boldsymbol{V},
$$

where $\boldsymbol{W} = \beta\mathbf{1}_n + \boldsymbol{Y}$, $\boldsymbol{V} = \boldsymbol{D}_\omega\boldsymbol{S}$, $\boldsymbol{Y} = (Y(\boldsymbol{x}_1), \ldots, Y(\boldsymbol{x}_n))^\top$, and $\boldsymbol{S} = (S(\boldsymbol{x}_1), \ldots, S(\boldsymbol{x}_n))^\top$, with $\boldsymbol{D}_\omega$ being an $n \times n$ diagonal matrix with $\omega$ on the diagonal. Now, since $S(\boldsymbol{x})$, for $\boldsymbol{x} \in \mathbb{R}^2$, are independently and identically distributed as $\mathrm{CSN}_{1,1}(0, 1, \alpha, 0, 1)$, according to Theorem 3 of González-Farías et al. (2004b), we have that $\boldsymbol{S} \sim \mathrm{CSN}_{n,n}(0, \boldsymbol{I}_n, \boldsymbol{D}_\alpha, 0, \boldsymbol{I}_n)$, where $\boldsymbol{D}_\alpha$ is the $n \times n$ diagonal matrix with $\alpha$ on the diagonal.

Since $\boldsymbol{Y}$ follows a multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}_Y$ with entries given by $\mathrm{Cov}\big(Y(\boldsymbol{x}), Y(\boldsymbol{x}+\boldsymbol{h})\big) = \varsigma^2\rho(\boldsymbol{h})$, we also have that $\boldsymbol{Y} \sim \mathrm{CSN}_{n,1}(0, \boldsymbol{\Sigma}_Y, 0, 0, 1)$. Moreover, being $\boldsymbol{W}$ distributed as a multivariate normal with mean $\beta\mathbf{1}_n$ and covariance matrix $\boldsymbol{\Sigma}_Y$, we can write that $\boldsymbol{W} \sim \mathrm{CSN}_{n,1}(\beta\mathbf{1}_n, \boldsymbol{\Sigma}_Y, 0, 0, 1)$, and using Theorem 1 of González-Farías et al. (2004b) we can also write that $\boldsymbol{V} \sim \mathrm{CSN}_{n,n}(0, \boldsymbol{D}_{\omega^2}, \boldsymbol{D}_{\alpha/\omega}, 0, \boldsymbol{I}_n)$, where $\boldsymbol{D}_{\omega^2}$ is the $n \times n$ diagonal matrix with $\omega^2$ on the diagonal, and $\boldsymbol{D}_{\alpha/\omega}$ is the $n \times n$ diagonal matrix with $\alpha/\omega$ on the diagonal. Thus, considering that $\boldsymbol{Z} = \boldsymbol{W} + \boldsymbol{V}$, we can conclude, using Theorem 4 of González-Farías et al. (2004b), that $\boldsymbol{Z} \sim \mathrm{CSN}_{n,n+1}(\beta\mathbf{1}_n, \boldsymbol{\Sigma}_Y + \omega^2\boldsymbol{I}_n, \boldsymbol{D}^*, 0, \boldsymbol{\Delta}^*)$, for some appropriate matrices $\boldsymbol{D}^*$ and $\boldsymbol{\Delta}^*$.

For this process, on the basis of the observations $Z(\boldsymbol{x}_1), \dots, Z(\boldsymbol{x}_n)$ at a set of spatial locations $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$, likelihood inference can be carried out by implementing some Monte Carlo EM algorithm, whereas a prediction at any given unobserved site $\boldsymbol{x}_0 \in \mathbb{R}^2$ can be obtained either by MCMC techniques (see also Zhang and El-Shaarawi, 2010), or by noticing that the joint distribution of $Z(\boldsymbol{x}_0), \dots, Z(\boldsymbol{x}_n)$ is closed skew-normal and so by exploiting the closure property with respect to conditioning of this distribution to find directly the conditional distribution of $Z(\boldsymbol{x}_0)$, given the observed data $Z(\boldsymbol{x}_1), \dots, Z(\boldsymbol{x}_n)$, which is again closed skew-normal; see, e.g., Proposition 2.3.2 of González-Farías et al. (2004a). Let us finally mention that following, for instance, Minozzo and Fruttini (2004), it would be possible to extend this (univariate) spatial process to a multivariate one having univariate marginal skew-normal distributions and multivariate (for $n \geq 2$) marginal closed skew-normal distributions, by building on the classical geostatistical proportional covariance model, or, more generally, on the linear model of coregionalization.

## 4. Conclusions

In this paper, we have raised the attention on some ill defined skew-normal processes that have recently appeared in the literature and showed with a counter-example that these processes do not exist. This counter-example is concerned with a particular, though important, family of skew-normal distributions, and with the poor characterization of their scale matrices. It does not prevent the existence of stationary stochastic processes with an autocorrelation function decreasing to zero, for which all the finite-dimensional marginal distributions are (multivariate) skew-normal, for some particular subclass or family of multivariate skew-normal distributions. Indeed, adopting a constructive approach to the definition of a stochastic process, we have built a strictly stationary spatial process having all its (multivariate) finite-dimensional marginal distributions belonging to the closed skew-normal family.

## Acknowledgements

The authors are grateful to the two guest editors and to the two referees for their constructive comments and suggestions, and would like to thank Antonio Gualtierotti and Philippe Naveau for very interesting exchanges of opinions, and also Hyoung-Moon Kim and Bani K. Mallick for useful discussions on their work. The first author would also like to thank Adelchi Azzalini for the many suggestions he received (not just on this occasion!). Finally, we gratefully acknowledge funding from the Italian Ministry of Education, University and Research (MIUR) through PRIN 2008 project 2008MRFM2H.

## References

Allard, D., Naveau, P., 2007. A new spatial skew-normal random field model. Communications in Statistics - Theory and Methods, 36, 1821–1834.

Allard, D., Soubeyrand, S., 2012. Skew-normality for climate data and dispersal models for plant epidemiology: when application fields drive spatial statistics. Spatial Statistics, 1, 50–64.

Arellano-Valle, R.B., Azzalini, A., 2006. On the unification of families of skew-normal distributions. Scandinavian Journal of Statistics, 33, 561–574.

Arellano-Valle, R.B., del Pino, G., San Martín, E., 2002. Definition and probabilistic properties of skew-distributions. Statistics and Probability Letters, 58, 111–121.

Azzalini, A., 1985. A class of distributions which includes the normal ones. Scandinavian Journal of Statistics, 12, 171–178.

Azzalini, A., 1986. Further results on a class of distributions which includes the normal ones. Statistica, 46, 199–208.

Azzalini, A., 2005. The skew-normal distribution and related multivariate families. Scandinavian Journal of Statistics, 32, 159–188.

Azzalini, A., Capitanio, A., 1999. Statistical applications of the multivariate skew-normal distribution. Journal of The Royal Statistical Society Series B - Statistical Methodology, 61, 579–602.

Azzalini, A., Dalla Valle, A., 1996. The multivariate skew-normal distribution. Biometrika, 83, 715–726.

Corns, T.R.A., Satchell, S.E., 2007. Skew Brownian motion and pricing European options. The European Journal of Finance, 13, 523–544.

Diggle, P.J., Moyeed, R.A., Tawn, J.A., 1998. Model-based geostatistics (with discussion). Applied Statistics, 47, 299–350.

Diggle, P.J., Ribeiro, Jr., P.J., 2007. Model-Based Geostatistics. Springer, New York.

Ferracuti, L., 2005. Geostatistical non-Gaussian factor models for multivariate spatial data. Ph.D. Thesis. University of Perugia, Perugia.

Genton, M.G., (ed.) 2004. Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality. Chapman and Hall/CRC, Boca Raton, FL.

González-Farías, G., Domínguez-Molina, J.A., Gupta, A.K., 2004a. The closed skew-normal distribution. In Genton, M.G., (ed.). Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality. Chapman and Hall/CRC, Boca Raton, FL, pp. 25–42.

González-Farías, G., Domínguez-Molina, J.A., Gupta, A.K., 2004b. Additive properties of skew-normal random vectors. Journal of Statistical Planning and Inference, 126, 521–534.

Gualtierotti, A.F., 2004. A family of (skew-normal) stochastic processes that can model some non-Gaussian random signals in dependent Gaussian noise. In 8th World Multi-conference on Systemics, Cybernetics and Informatics (SCI 2004). Volume VI. Orlando, FL, pp. 88–94.

Gualtierotti, A.F., 2005. Skew-normal processes as models for random signals corrupted by Gaussian noise. International Journal of Pure and Applied Mathematics, 20, 109–142.

Gupta, A.K., Chen, J.T., 2004. A class of multivariate skew-normal models. Annals of the Institute of Statistical Mathematics, 56, 305–315.

Hosseini, F., Eidsvik, J., Mohammadzadeh, M., 2011. Approximate Bayesian inference in spatial GLMM with skew normal latent variables. Computational Statistics and Data Analysis, 55, 1791–1806.

Kim, H.-M., Ha, E., Mallick, B.K., 2004. Spatial prediction of rainfall using skew-normal processes. In Genton, M.G., (ed.). Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality. Chapman and Hall/CRC, Boca Raton, FL, Chapter 16, pp. 279–289.

Kim, H.-M., Mallick, B.K., 2002. Analyzing spatial data using skew-Gaussian processes. In Lawson, A., Deninson, D., (eds.). Spatial Cluster Modelling. Chapman and Hall/CRC, London.

Kim, H.-M., Mallick, B.K., 2004. A Bayesian prediction using the skew Gaussian distribution. Journal of Statistical Planning and Inference, 120, 85–101.

Kim, H.-M., Mallick, B.K., 2005. A Bayesian prediction using the elliptical and the skew Gaussian processes. Technical report. `http://citeseer.ist.psu.edu/325606.html`.

Minozzo, M., Fruttini, D., 2004. Loglinear spatial factor analysis: an application to diabetes mellitus complications. Environmetrics, 15, 423–434.

Naveau, P., Allard, D., 2004. Modeling skewness in spatial data analysis without data transformation. In Leuangthong, O., Deutsch, C., (eds.). Proceedings of the Seventh International Geostatistics Congress. Springer, Dordrecht, pp. 929–938.

Pourahmadi, M., 2007. Skew-normal ARMA models with nonlinear heteroscedastic predictors. Communications in Statistics - Theory and Methods, 36, 1803–1819.

Tong, H., 1990. Non-Linear Time Series: A Dynamical System Approach. Oxford University Press, Oxford.

Zhang, H., El-Shaarawi, A., 2010. On spatial skew-Gaussian processes and applications. Environmetrics, 21, 33–47.