Working Paper Series
Department of Economics
University of Verona

On the existence of some skew normal stationary processes

Marco Minozzo

WP Number: 20 December 2011

ISSN: 2036-2919 (paper), 2036-4679 (online)

Chilean Journal of Statistics
Month 200x

On the existence of some skew-normal stationary processes

Marco Minozzo[1,*] and Laura Ferracuti

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

(Received: 1 October 2011 · Accepted in final form: 00 Month 200x)

Abstract

Recently some authors have introduced in the literature stationary stochastic processes, in the time and in the spatial domains, whose finite-dimensional marginal distributions are multivariate skew-normal. Here we show with a counter-example that the characterizations of these processes are not valid and so that these processes do not exist. Moreover, more generally, we also show that it is very unlikely that there might exist stationarity stochastic processes having all their finite-dimensional marginal distributions to be multivariate skew-normal. Besides, we point our attention to some valid constructions of stationary stochastic processes which can be used to model skewed data.

Keywords: multivariate skew-normal distribution · autocorrelation function · spatial process · stationary process · geostatistics · generalized linear mixed model.

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 have been the one that has received so far the widest attention by the statistical community. Following Azzalini and co-authors, we say that a random vector $\mathbf{Z} = (Z_1, \ldots, Z_n)^T$ has an *extended skew-normal distribution* with parameters μ, Σ, α and τ, and we write $\mathbf{Z} \sim \mathrm{SN}_n(\mu, \Sigma, \alpha, \tau)$, if it has probability density function of the form

$$f(\mathbf{z}) = \phi_n(\mathbf{z} - \mu; \Sigma) \cdot \Phi(\alpha_0 + \alpha^T D^{-1}(\mathbf{z} - \mu))/\Phi(\tau), \quad \text{for} \quad \mathbf{z} \in \mathbf{R}^n, \tag{1}$$

where $\mu \in \mathbf{R}^n$ is a vector of location parameters, $\phi_n(\,\cdot\,; \Sigma)$ is the n-dimensional normal density function with zero mean and (positive-definite) variance-covariance matrix Σ having elements σ_{ij}, $\Phi(\cdot)$ is the scalar N(0,1) distribution function, $D = \mathrm{diag}(\sigma_{11}, \ldots, \sigma_{nn})^{1/2}$ is the diagonal matrix formed with the standard deviations of the scale matrix Σ, $\alpha \in \mathbf{R}^n$

*Corresponding author. Department of Economics, University of Verona, Via dell'Artigliere 19, 37129 Verona, Italy. Phone: +39 045 802 8234. Fax: +39 045 802 8177. Email: marco.minozzo@univr.it

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

http://www.soche.cl/chjs

is a vector of skewness parameters, and $\tau \in \mathbf{R}$ is an additional parameter. Moreover, $\alpha_0 = \tau(1 + \alpha^T R\alpha)^{1/2}$ where R is the correlation matrix associated to Σ, that is, $R = D^{-1}\Sigma D^{-1}$. Clearly, this distribution extends the multivariate normal distribution through the parameter vector α, and for $\alpha = 0$ it reduces to the latter. When $\tau = 0$, also $\alpha_0 = 0$ and (1) reduces to

$$f(\mathbf{z}) = 2 \cdot \phi_n(\mathbf{z} - \mu; \Sigma) \cdot \Phi(\alpha^T D^{-1}(\mathbf{z} - \mu)), \quad \text{for} \quad \mathbf{z} \in \mathbf{R}^n. \tag{2}$$

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

The growing interest in these and other related multivariate families of distributions (see, for instance, Genton (2004); Azzalini (2005); Arellano-Valle and Azzalini (2006)) has led some authors to the specification of stochastic processes, in the time and in the spatial domains, that is, with indexing parameter in $\mathbf{R}$ or in $\mathbf{R}^2$ (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 the others, skew-normal processes in the time domain have been put forward by Gualtierotti (2004, 2005), by Pourahmadi (2007), and by Corns and Satchell (2007), whereas spatial skew-normal processes have been defined by Kim and Mallick (2002, 2004, 2005), by Kim et al. (2004), by Naveau and Allard (2004), by Allard and Naveau (2007), by Zhang and El-Shaarawi (2010), and by 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 adoption of a (stationary) autocorrelation function to characterize the covariance structure of the process.

To clarify our point, let us underline that our negative result does not prevent 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 following particular *self exiting threshold autoregressive* (SETAR) model

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

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

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

that is, by $\tilde{f}_{Z_t}(z) = 2\phi_1(z; (1-\alpha^2)^{-1})\Phi(-\alpha z)$, where $\phi_1(\,\cdot\,; (1-\alpha^2)^{-1})$ is the 1-dimensional normal density function with zero mean and variance $(1-\alpha^2)^{-1}$. 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 $\mathrm{SN}_1(0, (1-\alpha^2)^{-1}, -\alpha)$. However, this fact does not imply that also the multivariate finite-dimensional marginal distributions of (3) are (multivariate) skew-normal. For instance, without loss of generality, consider the case in which $\alpha = 1/\sqrt{2}$. Then,

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

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 multivariate skew-normal family. Thus, Equation (3) specifies a stationary stochastic process having univariate, but not multivariate, marginal distributions that are skew-normal. We will argue in the following that, in general, the assumption of stationarity and the requirement that all finite-dimensional marginal distributions of a stochastic process had to be (multivariate) skew-normal are not compatible.

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 appeared in the literature and then discuss the implications of this counter-example in general. On the other hand, in Section 3, in a geostatistical setting and following a hierarchical approach, we show how to characterize a stationary stochastic processes having univariate skew-normal conditional distributions, for which we can derive some of its moments. Lastly, in Section 4 we conclude with some discussion.

2. Skew-normal stationary processes

In this section, to discuss the existence of stationary stochastic processes having all their finite-dimensional marginal distribution (multivariate) skew-normal, we 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

Let us consider the following characterization of a spatial skew-normal stationary stochastic process as appeared in Kim and Mallick (2004). Indicating with $\{Z(\mathbf{x}) : \mathbf{x} \in \mathbf{R}^2\}$ 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 $\mathbf{Z} = (Z(\mathbf{x}_1), \dots, Z(\mathbf{x}_n))^T$, where $\mathbf{x}_1,\dots,\mathbf{x}_n$ are n fixed spatial locations, has the following skew-normal density (Formula (3) of their paper)

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

where $\mathbf{F}$ is a known design matrix (of dimension $n \times q$) with full column rank, $\boldsymbol{\beta} \in \mathbf{R}^q$ are unknown regression parameters, $\sigma \in \mathbf{R}^+$ is a scale parameter, $\alpha \in \mathbf{R}$ is a skewness parameter, and $\mathbf{1}_n$ is the n-dimensional column vector of ones. Moreover, they also assume that $\mathbf{K}_\theta$ is a positive definite matrix (of dimension $n \times n$) with each entry given by $K_\theta(||\mathbf{x}_i - \mathbf{x}_j||)$, where $||\mathbf{x}_i - \mathbf{x}_j||$ denotes the Euclidean distance between $\mathbf{x}_i$ and $\mathbf{x}_j$, 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 following 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, let us note that for the stochastic process $\{Z(\mathbf{x}) : \mathbf{x} \in \mathbf{R}^2\}$ to be stationary we must have at least $\boldsymbol{\beta} = 0$; and that the scale matrix $\sigma^2\mathbf{K}_\theta$ is not, in general, the variance-covariance matrix of $\mathbf{Z} = (Z(\mathbf{x}_1), \dots, Z(\mathbf{x}_n))^T$. Indeed, Kim and Mallick (2004) use 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 $\{Z(\mathbf{x}) : \mathbf{x} \in \mathbf{R}^2\}$ only in the case in which $\alpha = 0$ (and, of course, $\boldsymbol{\beta} = 0$, otherwise the process is not stationary).

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 (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 its covariance structure.

To show our point, let us consider three distinct and fixed spatial locations $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \in \mathbf{R}^2$. Let also $K_\theta(\cdot)$ be a given (isotropic) spatial autocorrelation function as before, and let $k_{12} \doteq K_\theta(||\mathbf{x}_1 - \mathbf{x}_2||)$, $k_{13} \doteq K_\theta(||\mathbf{x}_1 - \mathbf{x}_3||)$ and $k_{23} \doteq K_\theta(||\mathbf{x}_2 - \mathbf{x}_3||)$. Then, being $\{Z(\mathbf{x}) : \mathbf{x} \in \mathbf{R}^2\}$ 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 $(Z(\mathbf{x}_1), Z(\mathbf{x}_2), Z(\mathbf{x}_3))^T$ should be multivariate skew-normal with density

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

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

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

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

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

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

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

and with a little algebra we see that a necessary condition for (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 $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$, it follows that 'admissible' isotropic spatial autocorrelation functions $K_\theta(d)$ are to be 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 (9), and so (8) hold true.

This counter-example shows that the characterization of Kim and Mallick (2004) is improper and so that 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)). Basically, in this work, to characterize a spatial stochastic process $\{Z(\mathbf{x}) : \mathbf{x} \in \mathbf{R}^2\}$, the authors assume that for any given set of n spatial locations $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ in $\mathbf{R}^2$, the random variables $Z(\mathbf{x}_1), \ldots, Z(\mathbf{x}_n)$ are jointly distributed (see Formula (5) of Allard and Naveau (2007)) as a multivariate *closed skew-normal distribution* as defined by Gonzalez-Farias et al. (2004), with the elements of the scale matrix specified through a (stationary) covariance function. Hence, since the closed skew-normal distribution includes the skew-normal distribution of Azzalini and Capitanio (1999) as a special case, we can still apply to this characterization the counter-example just discussed to show that, in general, also the stochastic processes advocated by Allard and Naveau (2007) do not exist.

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 introduces 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 would try to build the scale matrices of its finite-dimensional marginal distributions using a (stationary) covariance function decaying to zero, as the separating 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 distribution 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: a negative result

Let us now consider the following general question. Are there strictly stationary stochastic processes having all their finite-dimensional marginal distributions to be (multivariate) skew-normal? Somehow, similar queries have been posed by Pourahmadi (2007) which in the context of *autoregressive and moving average* (ARMA) models tries to argue that there is a considerable trade-off between stationarity and skewness. As far as we are concerned, here we argue, by recalling and adapting to our case some of the arguments of Pourahmadi (2007), that the counter-example of Section 2.1 reveals that the answer is no.

To this aim, let us try to consider (by absurdum) a real-valued stochastic process Z, with indexing parameter $\mathbf{x}$ varying in some indexing set (which might be $\mathbf{R}$, $\mathbf{R}^2$, or some subset of it), for which, for every integer $n \geq 1$, and every set of indexing values $\mathbf{x}_1, \ldots, \mathbf{x}_n$, the n-dimensional random vector $\mathbf{Z} = (Z(\mathbf{x}_1), \ldots, Z(\mathbf{x}_n))$ has an extended skew-normal distribution $\mathrm{SN}_n(\mu_n, \Sigma_n, \alpha_n, \tau)$ as in (1), where μ_n, Σ_n and α_n depend on n, and for which it holds the hypothesis of strict stationarity. That is, for which, for every integer $n \geq 1$, and every set of indexing values $\mathbf{x}_1, \ldots, \mathbf{x}_n$ and $\mathbf{x}_1 + \mathbf{h}, \ldots, \mathbf{x}_n + \mathbf{h}$, the n-dimensional random vector $(Z(\mathbf{x}_1), \ldots, Z(\mathbf{x}_n))$ has the same distribution of $(Z(\mathbf{x}_1 + \mathbf{h}), \ldots, Z(\mathbf{x}_n + \mathbf{h}))$. In this case, it follows that Z is also second-order stationary, that is, that, for all couples of indexing values $\mathbf{x}$ and $\mathbf{x} + \mathbf{h}$:

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

Let us remember that for a second-order stationary stochastic process we have basically, that is, if its spectral distribution function is absolutely continuous, that the autocorrelation $\mathrm{Corr}[Z(\mathbf{x} + \mathbf{h}), Z(\mathbf{x})]$, which is a (real-valued) function of $\mathbf{h}$, goes to 0, as $||\mathbf{h}|| \to \infty$ (see, for instance, Cox and Miller (1965), Chapter 8).

On the other hand, for any given random vector $\mathbf{Z} = (Z(\mathbf{x}_1), \ldots, Z(\mathbf{x}_n))$ of dimension n, since it must have an extended skew-normal distribution, it is well known that we can express $\mathrm{E}[\mathbf{Z}]$ and $\mathrm{Cov}[\mathbf{Z}]$ in terms of μ_n, Σ_n, α_n and τ through the formulae

$$\mathrm{E}[\mathbf{Z}] = \mu_n + \zeta_1(\tau) D_n \delta_n \qquad \text{and} \qquad \mathrm{Cov}[\mathbf{Z}] = \Sigma_n + \zeta_2(\tau) D_n \delta_n \delta_n^T D_n,$$

where $\delta_n = (1 + \alpha_n^T R_n \alpha_n)^{-1/2} R_n \alpha_n$, and

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

(as usual, $\phi(\cdot)$ is the 1-dimensional standard normal probability density function of mean zero and variance one).

Now, intuitively, and 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 $\mathbf{Z} = (Z(\mathbf{x}_1), \ldots, Z(\mathbf{x}_n))$, this would force δ_n and D_n to be of the form $\delta_n = \delta_0 \mathbf{1}_n$, $D_n = \gamma_0 I_n$, where I_n is the identity matrix of dimension n, and hence would force the scale matrix Σ_n and the covariance matrix $\mathrm{Cov}[\mathbf{Z}]$ of the random vector $\mathbf{Z}$ to be equal, up to an additive constant. For instance, if the indexing parameter of Z varies in the integers and for some n the random vector $\mathbf{Z} = (Z(\mathbf{x}_1), \ldots, Z(\mathbf{x}_n))$ is such that the indexing values $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are equally spaced, then, since the covariance matrix of $\mathbf{Z}$ should have a Toeplitz structure (with constant entries along diagonals), also the scale matrix Σ_n should have a Toeplitz structure. Thus, since the counter-example in Section 2.1 shows that there cannot be stochastic processes having all their finite-dimensional marginal distributions to be skew-normal and, at the same time, the scale matrices Σ_n to be constructed using a (stationary) covariance function going to 0, as $||\mathbf{h}|| \to \infty$, this counter-example also reveals that there cannot be strictly stationary stochastic processes having all their finite-dimensional marginal distributions to be skew-normal.

Let us note that if we do not require stationarity (or any other particular condition), with respect to whatever family of multivariate skew-normal distributions closed under marginalization (as the families defined by (1) and (2)) we can always define a stochastic process (made up of an infinite number of random variables) having all its finite-dimensional marginal distributions to belong to the same family, if, for instance, the indexing parameter of the process takes values in some countable set. For example, with respect to the family defined by (2), let us consider the following infinite sequence of multivariate distributions of increasing dimension

$$\mathrm{SN}_1(\mu_1, \Sigma_1, \alpha_1), \quad \mathrm{SN}_2(\mu_2, \Sigma_2, \alpha_2), \quad \mathrm{SN}_3(\mu_3, \Sigma_3, \alpha_3), \ \ldots$$

such that, for every n, $\mathrm{SN}_{n-1}(\mu_{n-1}, \Sigma_{n-1}, \alpha_{n-1})$ can be obtained by marginalization from $\mathrm{SN}_n(\mu_n, \Sigma_n, \alpha_n)$. Then, though it might result 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

As we stressed in the Introduction, the negative result of the previous section does not prevent the existence of stationary stochastic processes 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 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, still here the finite-dimensional marginal distributions of the process do not belong to any family 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.

Alternatively, always remaining in the spatial domain, a second-order, and also strongly, stationary geostatistical stochastic process, different in spirit from that of Zhang and El-Shaarawi (2010), having (univariate) skew-normal conditional and marginal distributions, can be defined building on a latent stationary Gaussian process, following the hierarchical approach of Diggle et al. (1998) (see also Diggle and Ribeiro (2007)). Following Minozzo and Fruttini (2004) and Ferracuti (2005), let $\{Y(\mathbf{x}) : \mathbf{x} \in \mathbf{R}^2\}$ be a mean zero stationary Gaussian process with $\mathrm{Var}[Y(\mathbf{x})] = \varsigma^2 > 0$ and $\mathrm{Cov}\big[Y(\mathbf{x}), Y(\mathbf{x}+\mathbf{h})\big] = \varsigma^2 \rho(\mathbf{h})$, where $\rho(\mathbf{h})$ is a real valued spatial autocorrelation function for which $\rho(0) = 1$ and $\rho(\mathbf{h}) \to 0$, as $||\mathbf{h}|| \to \infty$. Assume then that the random variables $Z(\mathbf{x})$, $\mathbf{x} \in \mathbf{R}^2$, satisfy:

$$Z(\mathbf{x}') \perp\!\!\!\perp Z(\mathbf{x}'') \;\big|\; Y(\mathbf{x}') \quad \text{and} \quad Z(\mathbf{x}') \perp\!\!\!\perp Y(\mathbf{x}'') \;\big|\; Y(\mathbf{x}'),$$

for any $\mathbf{x}' \neq \mathbf{x}''$; have conditional probability density $Z(\mathbf{x})\big|Y(\mathbf{x}) \sim f_{Z|Y}(z; M(\mathbf{x}))$, where $M(\mathbf{x}) = \mathrm{E}\big[Z(\mathbf{x})\big|Y(\mathbf{x})\big]$; and $h(M(\mathbf{x})) = \beta + Y(\mathbf{x})$, for some known link function $h(\cdot)$ and real parameter β. In the case in which $f_{Z|Y}(z; M(\mathbf{x}))$ is skew-normal and $h(\cdot)$ is a translation by a constant, it is easy to verify that the process $\{Z(\mathbf{x}) : \mathbf{x} \in \mathbf{R}^2\}$ is second-order, and also strongly, stationary. In particular, let us assume that $Z(\mathbf{x}) = \beta + Y(\mathbf{x}) + \omega S(\mathbf{x})$, where $\omega \in \mathbf{R}^+$ and $S(\mathbf{x})$, $\mathbf{x} \in \mathbf{R}^2$, are mutually independently distributed as skew-normals $\mathrm{SN}_1(0, 1, \alpha)$ such that, for every $\mathbf{x} \in \mathbf{R}^2$, the density of $S(\mathbf{x})$ is given by

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

where $\alpha \in \mathbf{R}$. In other words, we have that, for every $\mathbf{x} \in \mathbf{R}^2$, $Z(\mathbf{x})|Y(\mathbf{x}) \sim \mathrm{SN}_1(\beta +$

$Y(\mathbf{x}), \omega^2, \alpha)$. Now, although the random variables $Z(\mathbf{x})$, $\mathbf{x} \in \mathbf{R}^2$, are conditionally distributed as skew-normals, and also (see, for instance, 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 $\{Z(\mathbf{x}) : \mathbf{x} \in \mathbf{R}^2\}$ are not skew-normal (see also the comments in Gupta and Chen (2004)).

To derive the (stationary) autocorrelation structure of the process $\{Z(\mathbf{x}) : \mathbf{x} \in \mathbf{R}^2\}$, let us first consider that, since, for every $\mathbf{x} \in \mathbf{R}^2$,

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

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

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

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

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

it is discontinuos in zero, that is, $\gamma(0) \neq \gamma(0^+)$, and we have

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

Let us just 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, by building on the classical geostatistical proportional covariance model, or, more generally, on the linear model of coregionalization.

4. Discussion

In this paper we 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. Moreover, with the same counter-example we also showed, more generally, that there cannot exist strictly stationary stochastic processes, in the spatial or in the time domain, having all their finite-dimensional marginal distributions to be (multivariate) skew-normal. Though stationary processes 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.

Let us conclude noticing that, as it is, the above counter-example is relative to a particular, though important, family of skew-normal distributions and so that, ideally, it actually leaves open the possibility that there might exist strictly stationary processes having all their finite-dimensional marginal distributions to be (multivariate) skew-normal, for some particular subclass or family of multivariate skew-normal distributions. However, we think that this counter-example suggests that this possibility is very much unlikely, for whatever subclass or family of multivariate skew-normal distribution we might consider among the many proposals appeared in the literature.

Acknowledgements

We would like to thank Antonio Gualtierotti and Philippe Naveau for sending us a copy of their papers and 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!). 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.

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 & 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, XLVI, 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. J. Roy. Statist. Soc. B, 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.

Cox, D.R., Miller, H.D., 1965. The Theory of Stochastic Processes. Chapman & Hall, London.

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 & Hall/CRC, London.

González-Farías, G., Dominguez-Molina, J.A., Gupta, A.K., 2004. Additive properties of skew-normal random vectors. J. Statist. Plann. Infer., 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 Multiconference on Systemics, Cybernetics and Informatics (SCI 2004), Orlando, Florida, Vol. VI, 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. Ann. Inst. Statist. Math., 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., Mallick, B.K., 2002. Analyzing spatial data using skew-Gaussian processes, in: A. Lawson and D. Deninson, eds., Spatial Cluster Modelling, Chapman & 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

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

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 and Deutsch, eds., Proceedings of the Seventh International Geostatistics Congress, Kluwer Academic Publishers.

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, 31–47.