ENVIRONMETRICS
Environmetrics 2010; **21**: 318–340
Published online 13 July 2009 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/env.1004

Multivariate log-skew-elliptical distributions with applications to precipitation data

Yulia V. Marchenko[1,2] and Marc G. Genton[2*,†]

[1]*Stata Corp LP, College Station, TX 77845, U.S.A.*
[2]*Department of Statistics, Texas A&M University, College Station, TX 77843-3143, U.S.A.*

SUMMARY

We introduce a family of multivariate log-skew-elliptical distributions, extending the list of multivariate distributions with positive support. We investigate their probabilistic properties such as stochastic representations, marginal and conditional distributions, and existence of moments, as well as inferential properties. We demonstrate, for example, that as for the log-*t* distribution, the positive moments of the log-skew-*t* distribution do not exist. Our emphasis is on two special cases, the log-skew-normal and log-skew-*t* distributions, which we use to analyze US national (univariate) and regional (multivariate) monthly precipitation data. Copyright © 2009 John Wiley & Sons, Ltd.

KEY WORDS: climatic regions; extreme events; heavy tails; log-skew-normal; log-skew-*t*; moments; multivariate; precipitation; skewness

1. INTRODUCTION

In recent years, there has been a growing interest for more flexible parametric families of non-normal distributions with additional parameters allowing to regulate skewness and tails. This is especially important with environmental data which are often skewed and heavy tailed. The simplest representative of such families, as defined by Azzalini (1985), is the so-called skew-normal distribution. It extends the conventional normal model by introducing an additional parameter controlling the asymmetry of the distribution, the shape parameter. Azzalini and Dalla Valle (1996) proposed a multivariate analog of the univariate skew-normal distribution. Branco and Dey (2001) and Azzalini and Capitanio (2003) introduced the univariate and multivariate skew-*t* distributions, which extend the respective skew-normal distributions by allowing to control the tails of the distribution with the additional degrees of freedom parameter. A more detailed description of these and other skewed models may be found in the book edited by Genton (2004) and in the review by Azzalini (2005).

The support of the univariate skew-normal, skew-*t*, and, more generally, skew-elliptical distributions is the real line. For data that cannot be negative, such as income or precipitation, distributions with positive support, such as gamma, exponential, and log-normal, should be used for modeling purpose. The problem of modeling is exacerbated in the multivariate setting, where tractable distributions besides

*Correspondence to: M. G. Genton, Department of Statistics, Texas A&M University, College Station, TX 77843-3143, U.S.A.
†E-mail: genton@stat.tamu.edu

Received 20 February 2009
Accepted 7 April 2009

the multivariate log-normal are lacking. We expand the list of such distributions by introducing a family of multivariate log-skew-elliptical distributions.

Azzalini *et al*. (2003) introduced the univariate log-skew-normal and log-skew-*t* distributions, although without formal investigation of their properties and characteristics, and use them to model family income data. We extend their definition more generally to the class of multivariate log-skew-elliptical distributions. We also examine probabilistic properties of the multivariate log-skew-elliptical distributions, such as stochastic representations, marginal and conditional distributions, and existence of moments.

In climatology, various distributions have been used to model the distribution of precipitation data; among them there are the exponential, gamma (e.g., Wilks, 2006, p. 98), and log-normal (e.g., Crow and Shimizu, 1988). There is no definitive physical justification to what distribution has to be used to model precipitation. The distribution of precipitation data is often skewed and sometimes has heavy tails. For example, Wilson and Toumi (2005) study univariate heavy precipitation and point out that its distribution exhibits a heavy-tailed behavior. These characteristics motivate us to consider the log-skew-normal and log-skew-*t* distributions to model precipitation data. One appealing feature of these distributions is that their extension to the multivariate case is straightforward. Moreover, model parameters can be estimated using readily available estimation methods developed for the multivariate skew-normal and the multivariate skew-*t* distributions (Azzalini and Capitanio, 1999, 2003). We define these and, more generally, the multivariate skew-elliptical distributions next.

Let $\mathrm{EC}_d(\boldsymbol{\xi}, \Omega, g^{(d)})$ denote a family of d-dimensional elliptically contoured distributions (with existing probability density function, PDF) with a generator function $g^{(d)}(u)$, $u \geq 0$, defining a spherical d-dimensional density, a location column vector $\boldsymbol{\xi} \in \mathbb{R}^d$, and a $d \times d$ positive definite dispersion matrix Ω. If $\mathbf{X} \sim \mathrm{EC}_d(\boldsymbol{\xi}, \Omega, g^{(d)})$, then its PDF is $f_d(\mathbf{x}; \boldsymbol{\xi}, \Omega, g^{(d)}) = |\Omega|^{-1/2} g^{(d)}(Q_{\mathbf{x}}^{\boldsymbol{\xi},\Omega})$, where $Q_{\mathbf{x}}^{\boldsymbol{\xi},\Omega} = (\mathbf{x} - \boldsymbol{\xi})^\top \Omega^{-1}(\mathbf{x} - \boldsymbol{\xi})$ and $\mathbf{x} \in \mathbb{R}^d$ (Fang *et al*., 1990, p. 46). Genton (2004) summarizes various approaches for defining skew-elliptical distributions. We consider a class of skew-elliptical distributions with a PDF of the form

$$f_{\mathrm{SE}_d}(\mathbf{x}; \Theta) = 2 f_d(\mathbf{x}; \boldsymbol{\xi}, \Omega, g^{(d)}) F\{\boldsymbol{\alpha}^\top \omega^{-1}(\mathbf{x} - \boldsymbol{\xi}); g_{Q_{\mathbf{x}}^{\boldsymbol{\xi},\Omega}}\}, \quad \mathbf{x} \in \mathbb{R}^d \tag{1}$$

where $\Theta = (\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, g^{(d)})$, $\boldsymbol{\alpha} \in \mathbb{R}^d$ is a shape parameter, $\omega = \mathrm{diag}(\Omega)^{1/2}$ is a $d \times d$ scale matrix, $f_d(\mathbf{x}; \boldsymbol{\xi}, \Omega, g^{(d)})$ is the PDF of $\mathrm{EC}_d(\boldsymbol{\xi}, \Omega, g^{(d)})$ defined above, and $F(u; g_{Q_{\mathbf{x}}^{\boldsymbol{\xi},\Omega}})$ is the cumulative distribution function (CDF) of $\mathrm{EC}_1(0, 1, g_{Q_{\mathbf{x}}^{\boldsymbol{\xi},\Omega}})$ with the generator function $g_{Q_{\mathbf{x}}^{\boldsymbol{\xi},\Omega}}(u) = g^{(d+1)}(u + Q_{\mathbf{x}}^{\boldsymbol{\xi},\Omega})/g^{(d)}(Q_{\mathbf{x}}^{\boldsymbol{\xi},\Omega})$. Although $\boldsymbol{\alpha}$ is referred to as a shape parameter, the shape of the distribution (1) is regulated in a more complex way. In what follows, we use the notation $\mathrm{SE}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, g^{(d+1)})$ to refer to a family of skew-elliptical distributions with PDF (1). We also consider two special cases of multivariate skew-elliptical distributions: the skew-normal and the skew-*t*. The PDF of the multivariate skew-normal distribution is

$$f_{\mathrm{SN}_d}(\mathbf{x}; \Theta) = 2\phi_d(\mathbf{x}; \boldsymbol{\xi}, \Omega)\Phi\{\boldsymbol{\alpha}^\top \omega^{-1}(\mathbf{x} - \boldsymbol{\xi})\}, \quad \mathbf{x} \in \mathbb{R}^d \tag{2}$$

where $\Theta = (\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha})$, $\phi_d(\mathbf{x}; \boldsymbol{\xi}, \Omega)$ is the PDF of a d-variate normal distribution with location $\boldsymbol{\xi}$ and dispersion matrix Ω, and $\Phi(\cdot)$ is the CDF of the standard normal distribution. We denote a multivariate skew-normal random vector with PDF (2) as $\mathbf{X} \sim \mathrm{SN}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha})$. For more details and applications of the multivariate skew-normal distribution, see Azzalini and Dalla Valle (1996), Azzalini and Capitanio

DOI: 10.1002/env

(1999), Capitanio *et al*. (2003), and Azzalini (2005). The PDF of the multivariate skew-t distribution is

$$f_{\mathrm{ST}_d}(\mathbf{x};\Theta) = 2t_d(\mathbf{x};\boldsymbol{\xi},\Omega,\nu)T\left\{\boldsymbol{\alpha}^\top\omega^{-1}(\mathbf{x}-\boldsymbol{\xi})\left(\frac{\nu+d}{\nu+Q_{\mathbf{x}}^{\boldsymbol{\xi},\Omega}}\right)^{1/2};\nu+d\right\},\quad \mathbf{x}\in\mathbb{R}^d \tag{3}$$

where $\Theta = (\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, \nu)$, $t_d(\mathbf{x};\boldsymbol{\xi},\Omega,\nu) = \Gamma\{(\nu+d)/2\}(1+Q_{\mathbf{x}}^{\boldsymbol{\xi},\Omega}/\nu)^{-(\nu+d)/2}/\{|\Omega|^{1/2}(\nu\pi)^{d/2}\Gamma(\nu/2)\}$ is the PDF of a d-variate Student's t distribution with ν degrees of freedom, and $T(\cdot;\nu+d)$ is the CDF of a univariate Student's t distribution with $\nu+d$ degrees of freedom. We denote a multivariate skew-t random vector with PDF (3) as $\mathbf{X}\sim\mathrm{ST}_d(\boldsymbol{\xi},\Omega,\boldsymbol{\alpha},\nu)$. For more details and applications of the multivariate skew-t distribution see, for example, Azzalini and Capitanio (2003) and Azzalini and Genton (2008), among others.

The paper is organized as follows. The multivariate log-skew-elliptical distributions and their properties are defined and studied in Section 2, with emphasis on two special cases: the log-skew-normal and log-skew-t distributions. The relevant proofs of the results are given in the Appendix. Numerical applications of the log-skew-normal and log-skew-t distributions to US monthly precipitation data are presented in Section 3. The paper ends with a discussion in Section 4.

2. MULTIVARIATE LOG-SKEW-ELLIPTICAL DISTRIBUTIONS

In this section, we provide a formal definition of a family of multivariate log-skew-elliptical distributions and present their probabilistic properties including stochastic representations, conditional and marginal distributions, and moments, as well as inferential properties.

2.1. Definitions

Let $\mathbf{ln}(\mathbf{X}) = \{\ln(X_1),\ldots,\ln(X_d)\}^\top$, $X_i>0$, $i=1,\ldots,d$ be the component-wise logarithm of the positive random vector $\mathbf{X}=(X_1,\ldots,X_d)^\top$ and $\mathbf{exp}(\mathbf{Y}) = \{\exp(Y_1),\ldots,\exp(Y_d)\}^\top$ be the component-wise exponential of the random vector $\mathbf{Y}=(Y_1,\ldots,Y_d)^\top$.

Definition 1. *A positive random vector* $\mathbf{X}$ *has a multivariate log-skew-elliptical distribution, denoted as* $\mathbf{X}\sim\mathrm{LSE}_d(\boldsymbol{\xi},\Omega,\boldsymbol{\alpha},g^{(d+1)})$*, if* $\mathbf{ln}(\mathbf{X})$ *is a multivariate skew-elliptical random vector,* $\mathbf{ln}(\mathbf{X})\sim\mathrm{SE}_d(\boldsymbol{\xi},\Omega,\boldsymbol{\alpha},g^{(d+1)})$*, with PDF (1). Likewise, if* $\mathbf{X}\sim\mathrm{SE}_d(\boldsymbol{\xi},\Omega,\boldsymbol{\alpha},g^{(d+1)})$*, then* $\mathbf{exp}(\mathbf{X})\sim\mathrm{LSE}_d(\boldsymbol{\xi},\Omega,\boldsymbol{\alpha},g^{(d+1)})$.

If the multivariate log-skew-elliptical density exists, it is of the form

$$f_{\mathrm{LSE}_d}(\mathbf{x};\Theta) = 2\left(\prod_{i=1}^d x_i^{-1}\right) f_d\{\mathbf{ln}(\mathbf{x});\boldsymbol{\xi},\Omega,g^{(d)}\}F[\boldsymbol{\alpha}^\top\omega^{-1}\{\mathbf{ln}(\mathbf{x})-\boldsymbol{\xi}\};g_{Q^{\boldsymbol{\xi},\Omega}_{\mathbf{ln}(\mathbf{x})}}],\quad \mathbf{x}>\mathbf{0} \tag{4}$$

where $\Theta=(\boldsymbol{\xi},\Omega,\boldsymbol{\alpha},g^{(d)})$, and all other terms are as defined in Equation (1). Here, the term $\left(\prod_{i=1}^d x_i^{-1}\right)$ is the Jacobian associated with the transformation $\mathbf{ln}(\mathbf{x})\to\mathbf{x}$. Notice that when $\boldsymbol{\alpha}=\mathbf{0}$, the multivariate log-skew-elliptical PDF (4) reduces to a multivariate log-elliptical PDF as defined in Fang *et al.* (1990, p.56). Clearly, the interpretation of the parameters in Θ is not the same for $\mathbf{X}$ as for $\mathbf{ln}(\mathbf{X})$.

DOI: 10.1002/env

For example, $\boldsymbol{\xi}$, Ω, and $\boldsymbol{\alpha}$ do not regulate strictly location, scale, and skewness, respectively, on the original scale compared to the log scale. From the definition, $\mathbf{X} = \mathbf{exp}\{\mathbf{ln}(\mathbf{X})\} = \mathbf{exp}\{\boldsymbol{\xi} + \omega\mathbf{ln}(\mathbf{Z})\} = \mathbf{exp}(\boldsymbol{\xi}) \odot \mathbf{exp}\{\omega\mathbf{ln}(\mathbf{Z})\}$, where $\mathbf{ln}(\mathbf{Z}) \sim \mathrm{SE}_d(\mathbf{0}, \bar{\Omega}, \boldsymbol{\alpha}, g^{(d+1)})$ and $\bar{\Omega} = \omega^{-1}\Omega\omega^{-1}$ is the correlation matrix. We can see that $\boldsymbol{\xi}$ affects the scale of the distribution of $\mathbf{X}$, and ω (and more generally Ω) together with the 'shape' parameter $\boldsymbol{\alpha}$ regulate the shape of the distribution. We investigate how these parameters affect the shape of the distribution in more detail using the log-skew-normal and log-skew-t distributions which we define next.

From Definition 1 and Equations (2) and (3), the PDF of a multivariate log-skew-normal distribution, denoted as $\mathrm{LSN}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha})$, is

$$f_{\mathrm{LSN}_d}(\mathbf{x}; \Theta) = 2\left(\prod_{i=1}^{d} x_i^{-1}\right)\phi_d\{\mathbf{ln}(\mathbf{x}); \boldsymbol{\xi}, \Omega\}\Phi[\boldsymbol{\alpha}^\top\omega^{-1}\{\mathbf{ln}(\mathbf{x}) - \boldsymbol{\xi}\}], \quad \mathbf{x} > \mathbf{0} \tag{5}$$

and the PDF of a multivariate log-skew-t distribution, denoted as $\mathrm{LST}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, \nu)$, is for $\mathbf{x} > \mathbf{0}$:

$$f_{\mathrm{LST}_d}(\mathbf{x}; \Theta) = 2\left(\prod_{i=1}^{d} x_i^{-1}\right)t_d\{\mathbf{ln}(\mathbf{x}); \boldsymbol{\xi}, \Omega, \nu\}T\left[\boldsymbol{\alpha}^\top\omega^{-1}\{\mathbf{ln}(\mathbf{x}) - \boldsymbol{\xi}\}\left(\frac{\nu + d}{\nu + Q_{\mathbf{ln}(\mathbf{x})}^{\boldsymbol{\xi},\Omega}}\right)^{1/2}; \nu + d\right]. \tag{6}$$

It is easily seen that the densities (5) and (6) reduce to the PDF of a multivariate log-normal distribution, when $\boldsymbol{\alpha} = \mathbf{0}$ and $\nu = \infty$. Also, the PDF (6) reduces to the PDF of a multivariate log-t distribution, when $\boldsymbol{\alpha} = \mathbf{0}$. For $d = 1$, the PDFs (5) and (6) correspond to the distributions in Azzalini *et al.* (2003), although they did not give the PDFs explicitly.

We illustrate the shapes of univariate log-skew-normal and log-skew-t distributions, and bivariate log-skew-normal distributions next. Figure 1 presents univariate log-skew-normal (left panel) and

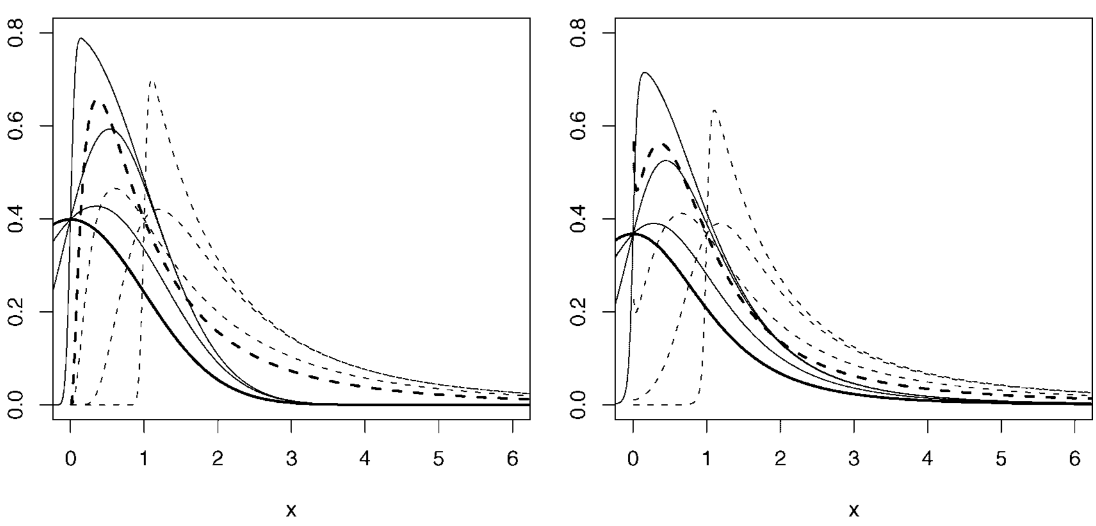

Figure 1. Left panel: skew-normal (solid curves) and log-skew-normal (dashed curves) densities with $\alpha = 0$ (thick curves) and $\alpha = 0.5, 2, 20$. Right panel: skew-t (solid curves) and log-skew-t (dashed curves) densities with $\alpha = 0$ (thick curves) and $\alpha = 0.5, 2, 20$ for fixed $\nu = 3$

DOI: 10.1002/env

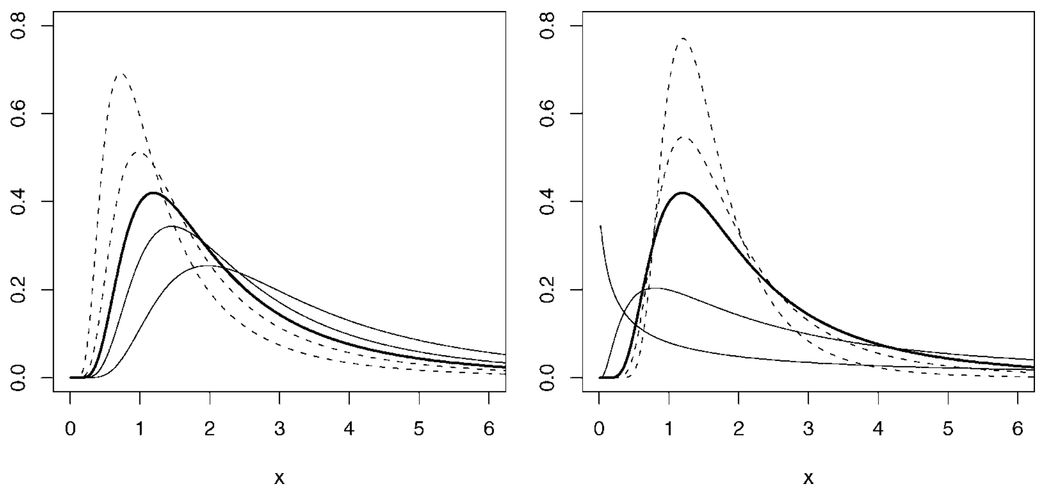

Figure 2. Log-skew-normal densities LSN(ξ, ω^2, 2) for varying ξ (left panel) and varying ω (right panel). Left panel: $\xi = 0$ (solid thick curve), $\xi = -0.2, -0.5$ (dashed curves), and $\xi = 0.2, 0.5$ (solid curves). Right panel: $\omega = 1$ (solid thick curve), $\omega = 0.6, 0.8$ (dashed curves), and $\omega = 2, 5$ (solid curves)

log-skew-*t* (right panel) densities for $\xi = 0$, $\omega = 1$, $\nu = 3$, and varying $\alpha = 0, 0.5, 2, 20$ (dashed curves). The respective skew-normal and skew-*t* densities are depicted for comparison as solid curves. The respective reference distributions are log-normal (log-*t*) and normal (Student's *t*) (thick curves). The additional parameter α allows the log-skew-normal and log-skew-*t* densities to have more flexible shapes than the reference log-normal and log-*t* distributions. The spike in the shape of the log-*t* density (right panel, thick dashed curve) is explained by the fact that this density has two stationary points for small values of ν. We observe a similar behavior for the log-skew-*t* density with $\alpha = 0.5$, but as in the case of the log-*t* distribution, it vanishes as α (or ν) increases.

Figure 2 depicts how the shape of a log-skew-normal density changes as a function of ξ and ω. The location ξ affects the shape in a multiplicative fashion (left panel). For positive values of ξ the density is stretched compared to the reference density with $\xi = 0$. For negative values of ξ, the density contracts toward the mode. Varying values of ω (right panel) change the look of the distribution, especially for large values of ω.

Figure 3 depicts bivariate log-skew-normal densities for various values of $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)^\top$. The shape $\boldsymbol{\alpha} = \mathbf{0}$ corresponds to the reference bivariate log-normal distribution. For negative values of α_1 and α_2, the density is skewed to the left in each direction, whereas for positive values of α_1 and α_2, it is more skewed to the right in each direction.

2.2. *Stochastic representations*

As in the case of the multivariate skew-elliptical distributions, there are other equivalent ways of defining the log-skew-elliptical distribution by using different stochastic representations. These representations may be used for simulation purposes. In this section, we formulate three stochastic representations for log-skew-elliptical random vectors.

DOI: 10.1002/env

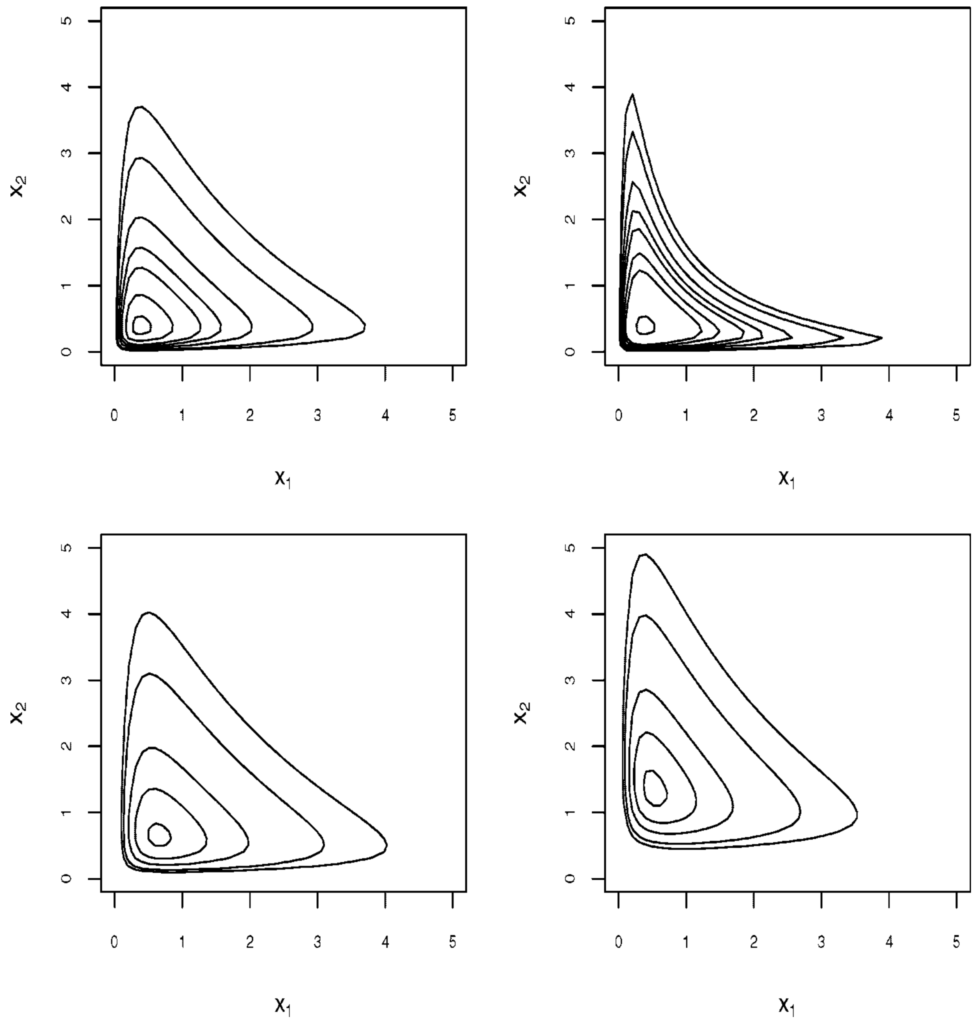

Figure 3. Contour plots of the standard bivariate log-skew-normal density plotted at levels 0.03, 0.05, 0.1, 0.15, 0.2, 0.3, 0.4, 0.8 for varying values of $\boldsymbol{\alpha}$: $\boldsymbol{\alpha} = (0, 0)^\top$ (top left panel), $\boldsymbol{\alpha} = (-2, -2)^\top$ (top right panel), $\boldsymbol{\alpha} = (0.5, 0.5)^\top$ (bottom left panel), and $\boldsymbol{\alpha} = (0.5, 2)^\top$ (bottom right panel)

Proposition 1 (selection representation 1). *Consider a $d+1$-dimensional random vector $(\tilde{U}_0, \mathbf{U}^\top)^\top$ that follows a multivariate elliptical distribution $\mathrm{EC}_{d+1}(\mathbf{0}, \bar{\Omega}^\star, g^{(d+1)})$ with*

$$\bar{\Omega}^\star = \begin{pmatrix} 1 & \mathbf{0}^\top \\ \mathbf{0} & \bar{\Omega} \end{pmatrix}.$$

Let $\mathbf{X} = \mathbf{exp}\{\boldsymbol{\xi} + \omega \mathbf{ln}(\mathbf{Z})\}$, $\mathbf{V} = \mathbf{exp}(\mathbf{U})$ and $\mathbf{Z} \stackrel{d}{=} (\mathbf{V} | \tilde{U}_0 < \boldsymbol{\alpha}^\top \mathbf{U})$. Then $\mathbf{Z} \sim \mathrm{LSE}_d(\mathbf{0}, \bar{\Omega}, \boldsymbol{\alpha}, g^{(d+1)})$ and, so, $\mathbf{X} \sim \mathrm{LSE}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, g^{(d+1)})$.

DOI: 10.1002/env

Proposition 2 (selection representation 2). *Consider a* $d+1$*-dimensional random vector* $(U_0, \mathbf{U}^\top)^\top$ *that follows a multivariate elliptical distribution* $\mathrm{EC}_{d+1}(\mathbf{0}, \bar{\Omega}^\star, g^{(d+1)})$ *with*

$$\bar{\Omega}^\star = \begin{pmatrix} 1 & \boldsymbol{\delta}^\top \\ \boldsymbol{\delta} & \bar{\Omega} \end{pmatrix}$$

and skewness parameter $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_d)^\top \in (-1, 1)^d$. *Let* $\mathbf{X} = \mathbf{exp}\{\boldsymbol{\xi} + \omega \mathbf{ln}(\mathbf{Z})\}$, $\mathbf{V} = \mathbf{exp}(\mathbf{U})$ *and* $\mathbf{Z} \stackrel{d}{=} (\mathbf{V} | U_0 > 0)$. *Then* $\mathbf{Z} \sim \mathrm{LSE}_d(\mathbf{0}, \bar{\Omega}, \boldsymbol{\alpha}, g^{(d+1)})$ *and, so,* $\mathbf{X} \sim \mathrm{LSE}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, g^{(d+1)})$ *with*

$$\boldsymbol{\alpha} = \bar{\Omega}^{-1}\boldsymbol{\delta}/(1 - \boldsymbol{\delta}^\top \bar{\Omega}^{-1}\boldsymbol{\delta})^{1/2}. \tag{7}$$

Azzalini and Dalla Valle (1996) refer to the selection representation 2 as a *conditioning method* and apply it to define a multivariate skew-normal distribution. This stochastic representation corresponds to the so-called $\boldsymbol{\delta}$-parameterization of the shape parameter $\boldsymbol{\alpha}$ from $\mathbb{R}^d$ to $(-1, 1)^d$. The selection representation 1 may be obtained from the selection representation 2 by setting $U_0 = (1 + \boldsymbol{\alpha}^\top \bar{\Omega} \boldsymbol{\alpha})^{-1/2}(\boldsymbol{\alpha}^\top \mathbf{U} - \tilde{U}_0)$.

Branco and Dey (2001) present a special subclass of the skew-elliptical distributions, the scale mixture of a skew-normal distribution, denoted as $\mathrm{SMSN}_d\{\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, K(\eta), H(\eta)\}$, with a PDF of the form

$$f_{\mathrm{SMSN}_d}(\mathbf{x}; \Theta) = 2\int_0^\infty \phi_d\left\{\mathbf{x}; \boldsymbol{\xi}, K(\eta)\Omega\}\Phi\{\boldsymbol{\alpha}^\top \omega^{-1}(\mathbf{x} - \boldsymbol{\xi})K^{-1/2}(\eta)\right\} dH(\eta), \quad \mathbf{x} \in \mathbb{R}^d \tag{8}$$

where $\Theta = \{\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, K(\eta), H(\eta)\}$, η is a random variable (a so-called mixing variable) with CDF $H(\eta)$, and $K(\eta)$ is a weight function. We denote $\mathbf{X} \sim \mathrm{LSMSN}_d\{\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, K(\eta), H(\eta)\}$, if $\mathbf{ln}(\mathbf{X}) \sim \mathrm{SMSN}_d\{\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, K(\eta), H(\eta)\}$ with PDF (8).

Proposition 3 (log-scale-mixture-skew-normal). *Let* $\mathbf{Z} \sim \mathrm{LSN}_d(\mathbf{0}, \Omega, \boldsymbol{\alpha})$. *Suppose that* η *is a random variable with CDF* $H(\eta)$ *and* $K(\eta)$ *is a weight function. If* $\mathbf{X} = \mathbf{exp}(\boldsymbol{\xi}) \odot \mathbf{Z}^{K^{1/2}(\eta)}$, *then* $\mathbf{X} \sim \mathrm{LSMSN}_d\{\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, K(\eta), H(\eta)\}$.

For example, the multivariate log-skew-normal and log-skew-t distributions are special cases of the LSMSN distribution. The multivariate log-skew-normal distribution arises when $K(\eta) = 1$ and $H(\eta)$ is degenerate. The multivariate log-skew-t distribution arises when $K(\eta) = 1/\eta$ and η is distributed as $Gamma(\nu/2, \nu/2)$ with PDF $f(\eta) = (\nu/2)^{\nu/2}\exp(-\nu\eta/2)/\Gamma(\nu/2)$.

Other stochastic representations of skew-elliptical random vectors can be used to generate log-skew-elliptical random vectors. For example, similarly to Proposition 2, we can formulate a convolution-type stochastic representation (Arellano-Valle and Genton, 2009) for log-skew-elliptical random vectors. Also, various types of skew-elliptical random vectors can be viewed as linear combinations of order statistics of exchangeable elliptical random variables (Arellano-Valle and Genton, 2007). Coupled with Definition 1, this presents yet another stochastic representation of log-skew-elliptical distributions.

2.3. Marginal and conditional distributions

From Definition 1, it can be inferred that all marginal distributions of $\mathbf{X}$ are univariate log-skew-elliptical. Since $\mathbf{ln}(\mathbf{X})$ is multivariate skew-elliptical, then each component $\ln(X_i)$ is univariate skew-elliptical (e.g. Branco and Dey, 2001). Therefore, by definition X_i is univariate log-skew-elliptical.

Proposition 4 (marginal distribution). *Let* $\mathbf{X} \sim \mathrm{LSE}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, g^{(d+1)})$. *Consider the following partition of* $\mathbf{X}^\top = (\mathbf{X}_1^\top, \mathbf{X}_2^\top)$, $\boldsymbol{\xi}^\top = (\boldsymbol{\xi}_1^\top, \boldsymbol{\xi}_2^\top)$, *and* $\boldsymbol{\alpha}^\top = (\boldsymbol{\alpha}_1^\top, \boldsymbol{\alpha}_2^\top)$ *into* q *and* $d-q$ *components, respectively. Let* $\Omega = (\omega_{ij})_{i,j=1}^d$ *have the following partition:*

$$\begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}$$

with $\Omega_{21} = \Omega_{12}^\top$. *Define the square matrices* $\omega_1 = \mathrm{diag}\left(w_{11}^{1/2}, \ldots, w_{qq}^{1/2}\right)_{q\times q}$ *and* $\omega_2 = \mathrm{diag}\left(w_{q+1q+1}^{1/2}, \ldots, w_{dd}^{1/2}\right)_{(d-q)\times(d-q)}$. *Then the marginal distribution of* $\mathbf{X}_1$ *is* $\mathrm{LSE}_q(\boldsymbol{\xi}_1, \Omega_{11}, \boldsymbol{\alpha}_1^\star, g^{(d+1)})$ *and the marginal distribution of* $\mathbf{X}_2$ *is* $\mathrm{LSE}_{d-q}(\boldsymbol{\xi}_2, \Omega_{22}, \boldsymbol{\alpha}_2^\star, g^{(d+1)})$, *where*

$$\boldsymbol{\alpha}_1^\star = \frac{\boldsymbol{\alpha}_1 + \omega_1\Omega_{11}^{-1}\Omega_{12}\omega_2^{-1}\boldsymbol{\alpha}_2}{(1+\boldsymbol{\alpha}_2^\top\bar{\Omega}_{22\cdot 1}\boldsymbol{\alpha}_2)^{1/2}}, \qquad \boldsymbol{\alpha}_2^\star = \frac{\boldsymbol{\alpha}_2 + \omega_2\Omega_{22}^{-1}\Omega_{21}\omega_1^{-1}\boldsymbol{\alpha}_1}{(1+\boldsymbol{\alpha}_1^\top\bar{\Omega}_{11\cdot 2}\boldsymbol{\alpha}_1)^{1/2}},$$

$$\bar{\Omega}_{22\cdot 1} = \omega_2^{-1}(\Omega_{22} - \Omega_{21}\Omega_{11}^{-1}\Omega_{12})\omega_2^{-1}, \quad \bar{\Omega}_{11\cdot 2} = \omega_1^{-1}(\Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21})\omega_1^{-1}. \tag{9}$$

As in the case of the skew-elliptical family, the log-skew-elliptical family is closed under marginalization but not under conditioning. To present a conditional distribution of $\mathbf{X}_2|\mathbf{X}_1$ for a log-skew-elliptical random vector $\mathbf{X}$, we first need to define the so-called extended skew-elliptical family (Arellano-Valle and Genton, 2009; Arellano-Valle and Azzalini, 2006; Arellano-Valle *et al.*, 2006). A random vector $\mathbf{X}$ has a multivariate extended skew-elliptical distribution, denoted by $\mathbf{X} \sim \mathrm{ESE}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, \tau, g^{(d+1)})$, if its density is of the form

$$f_{\mathrm{ESE}_d}(\mathbf{x};\Theta) = \frac{1}{F\left(\tau/\sqrt{1+\boldsymbol{\alpha}^\top\bar{\Omega}\boldsymbol{\alpha}}; g^{(d)}\right)} f_d(\mathbf{x};\boldsymbol{\xi},\Omega,g^{(d)})F\{\boldsymbol{\alpha}^\top\omega^{-1}(\mathbf{x}-\boldsymbol{\xi})+\tau; g_{Q_{\mathbf{x}}^{\xi,\Omega}}\}, \quad \mathbf{x}\in\mathbb{R}^d \tag{10}$$

where $\Theta = (\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, g^{(d)}, \tau)$, $\tau \in \mathbb{R}$ is the extension parameter, and other parameters are as defined in Equation (1). Note that Equation (10) reduces to Equation (1) if $\tau = 0$.

Similar to Definition 1, we define $\mathbf{X}$ to be a random vector from a log-extended-skew-elliptical family, denoted as $\mathbf{X} \sim \mathrm{LESE}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, \tau, g^{(d+1)})$, if $\mathbf{ln}(\mathbf{X}) \sim \mathrm{ESE}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, \tau, g^{(d+1)})$ as defined in Equation (10). The density function of the LESE random vector may be obtained similarly to Equation (4) using rules for transformation of random vectors.

Proposition 5 (conditional distribution). *Let* $\mathbf{X} \sim \mathrm{LSE}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, g^{(d+1)})$. *Consider the partitions defined in Proposition 4. Then*

$$(\mathbf{X}_2|\mathbf{X}_1) \sim \mathrm{LESE}_{d-q}\left(\boldsymbol{\xi}_{2\cdot 1}, \Omega_{22\cdot 1}, \boldsymbol{\alpha}_2, \tau, g^{(d+1)}_{Q_{\mathbf{x}_1}^{\xi_1,\Omega_{11}}}\right)$$

where $\boldsymbol{\xi}_{2\cdot 1} = \boldsymbol{\xi}_2 + \Omega_{21}\Omega_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\xi}_1)$, $\Omega_{22\cdot 1} = \Omega_{22} - \Omega_{21}\Omega_{11}^{-1}\Omega_{12}$, *and the extension parameter is* $\tau = (\boldsymbol{\alpha}_1 + \omega_1\Omega_{11}^{-1}\Omega_{12}\omega_2^{-1}\boldsymbol{\alpha}_2)^\top\omega_1^{-1}(\mathbf{x}_1 - \boldsymbol{\xi}_1)$.

DOI: 10.1002/env

2.4. Moments

Similar to the log-elliptical distribution (Fang *et al.*, 1990), if the mixed moments of a log-skew-elliptical distribution exist, they can be expressed conveniently using the characteristic function (or the moment generating function if it exists) of the skew-elliptical distribution.

Proposition 6 (mixed moments). *Let* $\mathbf{X} \sim \mathrm{LSE}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, g^{(d+1)})$ *and* $\mathbf{n} = (n_1, n_2, \ldots, n_d)^\top$, $n_i \in \mathbb{N}$, $i = 1, \ldots, d$. *If the mixed moments* $E\left(\prod_{i=1}^d X_i^{n_i}\right)$ *exist, then*

$$E\left(\prod_{i=1}^{d} X_i^{n_i}\right) = M_{\mathbf{ln(X)}}(\mathbf{n})$$

where $M_{\mathbf{ln(X)}}(\cdot)$ *is the moment generating function of the skew-elliptical random vector* $\mathbf{ln(X)} \sim \mathrm{SE}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, g^{(d+1)})$.

The mixed moment of the multivariate log-skew-normal random vector $\mathbf{X} \sim \mathrm{LSN}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha})$ can be expressed as

$$E\left(\prod_{i=1}^{d} X_i^{n_i}\right) = 2\exp\left(\boldsymbol{\xi}^\top \mathbf{n} + \mathbf{n}^\top \Omega \mathbf{n}/2\right) \Phi\left\{\boldsymbol{\alpha}^\top \bar{\Omega} \omega \mathbf{n}/(1 + \boldsymbol{\alpha}^\top \bar{\Omega} \boldsymbol{\alpha})^{1/2}\right\}$$

where the right-hand-side term in the above is the moment generating function of the skew-normal random vector (e.g., Genton, 2004, p. 17). For example, the first four moments of the univariate log-skew-normal random variable $X \sim \mathrm{LSN}(\xi, \omega^2, \alpha)$ are

$$E(X) = 2\exp(\xi + \omega^2/2)\Phi\{\alpha\omega/(1+\alpha^2)^{1/2}\} \tag{11}$$

$$\begin{aligned} E(X^2) &= 2\exp(2\xi + 2\omega^2)\Phi\{2\alpha\omega/(1+\alpha^2)^{1/2}\} \\ E(X^3) &= 2\exp(3\xi + 4.5\omega^2)\Phi\{3\alpha\omega/(1+\alpha^2)^{1/2}\} \\ E(X^4) &= 2\exp(4\xi + 8\omega^2)\Phi\{4\alpha\omega/(1+\alpha^2)^{1/2}\}. \end{aligned} \tag{12}$$

We can use these expressions to estimate the mean and skewness (and also kurtosis) of the fitted log-skew-normal distribution by substituting the parameters with the respective maximum likelihood estimates (MLEs). As for the log-*t* distribution, positive moments of the log-skew-*t* distribution do not exist.

Proposition 7 (log-skew-*t* moments). *The mixed moments* $E\left(\prod_{i=1}^d X_i^{n_i}\right)$ *of the log-skew-t random vector* $\mathbf{X} \sim \mathrm{LST}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, \nu)$ *are infinite for any* $n_i \geq 0$, $i = 1, \ldots, d$, *such that* $\sum_{i=1}^d n_i > 0$, *and any* $\boldsymbol{\xi} \in \mathbb{R}^d$, *positive definite matrix* Ω, $\boldsymbol{\alpha} \in \mathbb{R}^d$, *and* $\nu > 0$.

DOI: 10.1002/env

2.5. Inferential properties of log-skew-elliptical distributions

Let $l_{\mathrm{LSE}}(\Theta|\mathbf{x})$ be the log-likelihood function for the log-skew-elliptical model with PDF (4):

$$l_{\mathrm{LSE}}(\Theta|\mathbf{x}) = \ln\{f_{\mathrm{LSE}_d}(\mathbf{x};\Theta)\} = -\sum_{j=1}^{d}\ln(x_j) + \ln\{f_{\mathrm{SE}_d}(\mathbf{ln}(\mathbf{x});\Theta)\} = -\sum_{j=1}^{d}\ln(x_j) + l_{\mathrm{SE}}(\Theta|\mathbf{ln}(\mathbf{x})).\quad(13)$$

From relation (13), the LSE log-likelihood differs from the SE log-likelihood only by the term $-\sum_{j=1}^{d}\ln(x_j)$, which is free of any unknown parameters Θ. As such, inferences about Θ may be based on the SE log-likelihood. This allows us to use existing estimation methods developed for the skew-elliptical models to estimate parameters of the log-skew-elliptical model. To estimate the parameters, we simply fit the desired skew-elliptical model to the log-transformed original data. Since the LSE and the SE log-likelihoods are equivalent, we briefly summarize the inferential aspects for two important special cases, the log-skew-normal and the log-skew-*t*.

Inferential properties of skewed distributions received much attention in the literature, especially for the case of skew-normal and skew-*t* distributions (e.g., Azzalini and Capitanio, 1999, 2003; Sartori, 2006; Pewsey, 2000, 2006; Azzalini and Genton, 2008; Arellano-Valle and Azzalini, 2008). The two important inferential aspects are the existence of a stationary point at $\boldsymbol{\alpha} = \mathbf{0}$ of the profile log-likelihood function and the unboundedness of the log-likelihood function in some regions of the parameter space. In the case of the univariate skew-normal distribution, Pewsey (2006) proved the existence of a stationary point at $\alpha = 0$ of the profile log-likelihood function. Azzalini and Genton (2008) extended his argument to the multivariate case and showed that the Fisher information matrix of the profile log-likelihood of the skew-normal model is singular at $\boldsymbol{\alpha} = \mathbf{0}$. Azzalini (1985) proposed an alternative (centered) parameterization that alleviates the singularity of the resulting reparametrized information matrix for the univariate skew-normal distribution. Arellano-Valle and Azzalini (2008) extended this centered parametrization to the multivariate case. This unfortunate property seems to vanish in the case of the skew-*t* distribution. Azzalini and Capitanio (2003) noted that the behavior of the profile log-likelihood function of the skew-*t* distribution is more regular and demonstrate it numerically with several datasets. Although there is no rigorous proof of it, Azzalini and Genton (2008) presented a theoretical insight into why the Fisher information matrix is not singular at $\boldsymbol{\alpha} = \mathbf{0}$ in the case of the multivariate skew-*t* distribution.

The unboundedness of the MLEs of the shape and degrees of freedom parameters of the skew-*t* distributions was first discussed by Azzalini and Capitanio (2003). In the case of the univariate standard skew-*t* distribution with fixed degrees of freedom, the infinite estimate of the shape parameter is encountered when either all observations are positive or all observations are negative which can happen with positive probability. In other more general cases, such as unknown degrees of freedom and the multivariate case, the conditions under which the log-likelihood is unbounded are more complicated and, thus, more difficult to describe. Sartori (2006) and Azzalini and Genton (2008) presented ways of dealing with the unbounded estimates. Sartori (2006) proposed a bias correction to the MLEs. Azzalini and Genton (2008) suggested a deviance-based approach according to which the unbounded MLEs of $(\boldsymbol{\alpha}, \nu)$ are replaced by the smallest values $(\boldsymbol{\alpha}_0, \nu_0)$ such that the likelihood ratio test of $H_0 : (\boldsymbol{\alpha}, \nu) = (\boldsymbol{\alpha}_0, \nu_0)$ is not rejected at a fixed level, say 0.1.

DOI: 10.1002/env

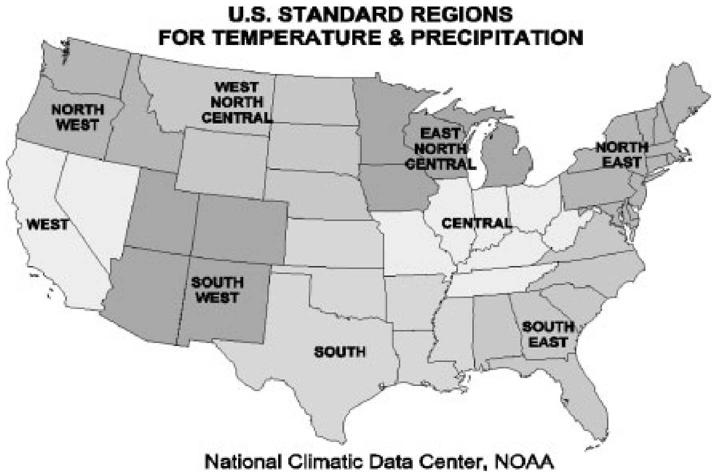

Figure 4. US climatic regions from NCDC

3. PRECIPITATION DATA ANALYSIS

US national and regional precipitation data are publicly available from the National Climatic Data Center (NCDC), the largest archive of weather data. We use monthly precipitation data measured in inches to hundreds for the period of 1895–2007 (113 observations per month). Monthly (divisional) precipitation data are obtained as monthly equally weighted averages of values reported by all stations within a climatic division. The regional values are computed from the statewide values (which are obtained from the divisional values weighted by area) weighted by area for each of the nine US climatic regions: Northeast, East North Central, Central, Southeast, West North Central, South, Southwest, Northwest, and West (see Figure 4). National values are obtained from the regional values weighted by area. Table 1 presents basic summary statistics of the data for each month and climatic region (including national values). The analyses were performed in R (R Development Core Team, 2008) and Stata (StataCorp, 2007) using, among other capabilities, the R package *sn* developed by Azzalini (2006) and a suite of Stata commands to be presented in Marchenko and Genton (2009).

3.1. US national scale

We analyze US national precipitation data by fitting univariate log-skew-normal and log-skew-*t* models and compare their fits to the conventional log-normal model (e.g., Crow and Shimizu, 1988). Separate analyses are carried out for each month. As discussed in Section 2.5, we estimate the parameters of the log-skew-normal and log-skew-*t* distributions by fitting the skew-normal and skew-*t* distributions, respectively, to the log-transformed data.

The attractiveness of the skew-normal and, more generally, the skew-elliptical distributions is that they preserve some pleasant properties of their respective elliptical counterparts. One of them is that the distribution of quadratic forms of skew-elliptical random vectors does not depend on the skewness parameter (and is chi-squared for the skew-normal model). This property is useful, for example, for

DOI: 10.1002/env

Table 1. Monthly precipitation minimum, mean, median, and maximum values for nine US climatic regions: Northeast (NE), East North Central (ENC), Central (C), Southeast (SE), West North Central (WNC), South (S), Southwest (SW), Northwest (NW), and West (W)

	US regions									
Month	NE	ENC	C	SE	WNC	S	SW	NW	W	National
January	0.87	0.32	0.72	0.92	0.16	0.53	0.23	0.43	0.28	0.92
	3.07	1.16	3.09	3.83	0.63	2.29	0.91	3.77	3.08	2.21
	2.93	1.04	2.87	3.61	0.61	2.18	0.79	3.83	2.59	2.17
	7.22	2.47	9.61	7.73	1.25	5.34	2.90	7.81	11.69	3.95
February	0.70	0.31	0.67	1.36	0.30	0.69	0.14	0.51	0.21	0.96
	2.71	1.06	2.71	3.95	0.60	2.30	0.81	2.94	2.42	2.01
	2.53	1.00	2.54	3.94	0.58	2.34	0.73	2.94	2.00	2.03
	5.43	2.40	5.47	7.45	1.07	5.63	2.07	5.75	7.57	3.20
March	0.71	0.23	0.55	1.45	0.39	0.89	0.20	0.58	0.25	0.97
	3.40	1.70	3.82	4.51	0.96	2.71	1.02	2.60	2.23	2.39
	3.33	1.66	3.70	4.55	0.93	2.79	0.93	2.47	1.97	2.35
	6.56	3.50	6.91	8.89	2.10	6.28	2.63	5.25	6.62	3.89
April	1.40	1.04	1.55	0.85	0.50	1.08	0.26	0.61	0.17	1.41
	3.39	2.56	3.90	3.62	1.57	3.22	0.94	1.90	1.27	2.42
	3.32	2.50	3.73	3.51	1.53	3.03	0.84	1.86	1.07	2.41
	6.81	4.85	6.82	7.06	2.83	6.92	2.50	3.81	3.33	3.56
May	0.98	1.15	1.65	0.97	0.65	1.53	0.19	0.36	0.08	1.78
	3.57	3.49	4.38	3.80	2.51	4.08	1.01	1.89	0.87	2.86
	3.43	3.32	4.25	3.62	2.43	4.09	1.01	1.79	0.69	2.87
	7.25	6.85	8.03	7.61	4.63	7.33	2.31	4.19	3.24	4.15
June	1.60	1.41	1.03	2.20	1.25	0.98	0.25	0.36	0.01	1.43
	3.76	4.01	4.24	4.98	2.86	3.66	0.90	1.46	0.41	2.90
	3.67	3.96	4.21	4.78	2.76	3.56	0.84	1.33	0.31	2.93
	8.53	6.68	9.11	8.37	5.27	7.05	1.94	3.02	1.14	4.21
July	2.02	0.85	1.47	2.94	0.84	1.34	0.90	0.16	0.01	1.81
	3.86	3.54	4.09	5.81	2.00	3.22	1.89	0.62	0.29	2.77
	3.74	3.41	4.02	5.74	1.89	3.17	1.81	0.54	0.25	2.79
	6.57	6.18	8.27	11.56	5.56	6.04	3.58	2.05	1.18	3.85
August	1.78	1.35	1.55	2.71	0.77	0.70	0.56	0.10	0.01	1.77
	3.84	3.54	3.61	5.26	1.72	2.93	1.92	0.74	0.34	2.59
	3.71	3.58	3.57	5.15	1.65	2.89	1.90	0.58	0.25	2.59
	8.01	6.27	6.30	9.77	3.03	6.06	3.25	2.98	2.01	3.55
September	1.25	0.95	0.71	1.91	0.47	0.88	0.09	0.12	0.03	1.45
	3.64	3.36	3.47	4.57	1.53	3.19	1.29	1.19	0.47	2.46
	3.42	3.28	3.34	4.42	1.46	3.15	1.34	1.11	0.30	2.47
	8.04	7.21	6.94	9.68	3.42	6.87	3.07	3.42	2.00	3.57
October	0.45	0.25	0.53	0.53	0.13	0.12	0.02	0.14	0.04	0.54
	3.36	2.33	2.90	3.18	1.14	2.81	1.10	2.22	0.97	2.14
	3.08	2.28	2.67	3.04	1.08	2.46	0.90	2.16	0.91	2.15
	9.43	4.66	7.15	7.33	2.95	7.07	3.67	5.20	3.24	3.72

Table 1. *Continued.*

November	0.88	0.24	0.72	0.83	0.06	0.20	0.09	0.30	0.04	0.88
	3.46	1.80	3.20	2.92	0.75	2.47	0.78	3.68	1.86	2.13
	3.36	1.73	3.13	2.73	0.71	2.36	0.70	3.64	1.57	2.09
	6.34	4.03	7.71	8.39	1.63	6.48	2.33	7.84	5.79	3.76
December	0.98	0.37	0.90	1.18	0.19	0.68	0.16	1.17	0.09	1.22
	3.26	1.27	3.13	3.71	0.62	2.55	0.88	3.91	2.50	2.23
	3.18	1.24	3.14	3.49	0.58	2.47	0.79	3.77	2.08	2.27
	6.74	2.62	7.58	7.05	1.20	5.51	2.29	8.42	7.05	3.60

The last column records summaries of monthly national values.

investigating how well the specified skew-elliptical model fits the data. Specifically, we can compare the probabilities of quantiles of the squared standardized residuals from a univariate skew-normal fit to the probabilities of quantiles of the chi-squared distribution with 1 degree of freedom. We present an example of such a probability plot (PP-plot) for the skew-normal fit for January in Figure 5 (top panel). From Figure 5, the skew-normal distribution fits the log-precipitation data slightly better than the normal distribution for January. In fact, PP-plots for the skew-normal and skew-*t* fits revealed that these distributions fit the log-precipitation data slightly better than the normal distribution for most months. For other months all distributions provided good fit.

Figure 6 presents estimated log-skew-normal means (left panel) and skew-normal skewness coefficients (right panel) for each month. Mean estimates are computed from the formula for the first moment of the log-skew-normal distribution given in Equation (11) with all parameters being replaced by the obtained MLEs. We present the skewness coefficients on the log scale. If desired, skewness coefficients on the original scale can be obtained similarly to means from Equation (12) using the recursive relationship between moments and central moments. The skew-normal model failed to converge when fitted to the June and October data due to unboundedness of the MLE for the shape parameter α. To alleviate this problem, we used the approaches of Sartori (2006) (labeled on the graph as MMLE for modified MLE) and of Azzalini and Genton (2008) (labeled on the graph as BMLE for bounded MLE), briefly mentioned in Section 2.5. The obtained two estimates (BMLE and MMLE) of the shape parameter are similar and their magnitude is comparable to the shape parameter estimates from other months.

The estimated mean precipitation ranges between 2.01 and 2.89 inch (after correcting for the unboundedness of the shape parameter). From the left panel of Figure 6, a particular trend is visible for the mean precipitation at the national level. The mean precipitation increases in Spring and Summer and decreases in Fall and Winter. The lowest mean precipitation of 2 inch is observed in February and the highest of 2.9 is observed in June. The estimated skewness coefficient (log scale, right panel) ranges between -0.7 and -0.1 and is negative for all months suggesting that the log-precipitation data are skewed to the right. The values of the skewness indexes are not too far from zero for some months which explains only a slight improvement in the fit of the skew-normal distribution. However, for other months based on the estimated skewness indexes the shape of the precipitation distribution deviates from symmetry and the fitted model allows to capture such deviations. More specifically, the normality assumption of the log-precipitation data was rejected by the likelihood ratio test of the null hypothesis $H_0 : \alpha = 0$ at the 1% significance level for June, September, October, November, and December.

Figure 7 (left panel) displays the estimated degrees of freedom parameter ν, controlling the heaviness of the distributional tails of the fitted log-skew-*t* models. According to this picture, the degrees of freedom ν is estimated to be 10 and higher for most months, including December, not depicted in the picture,

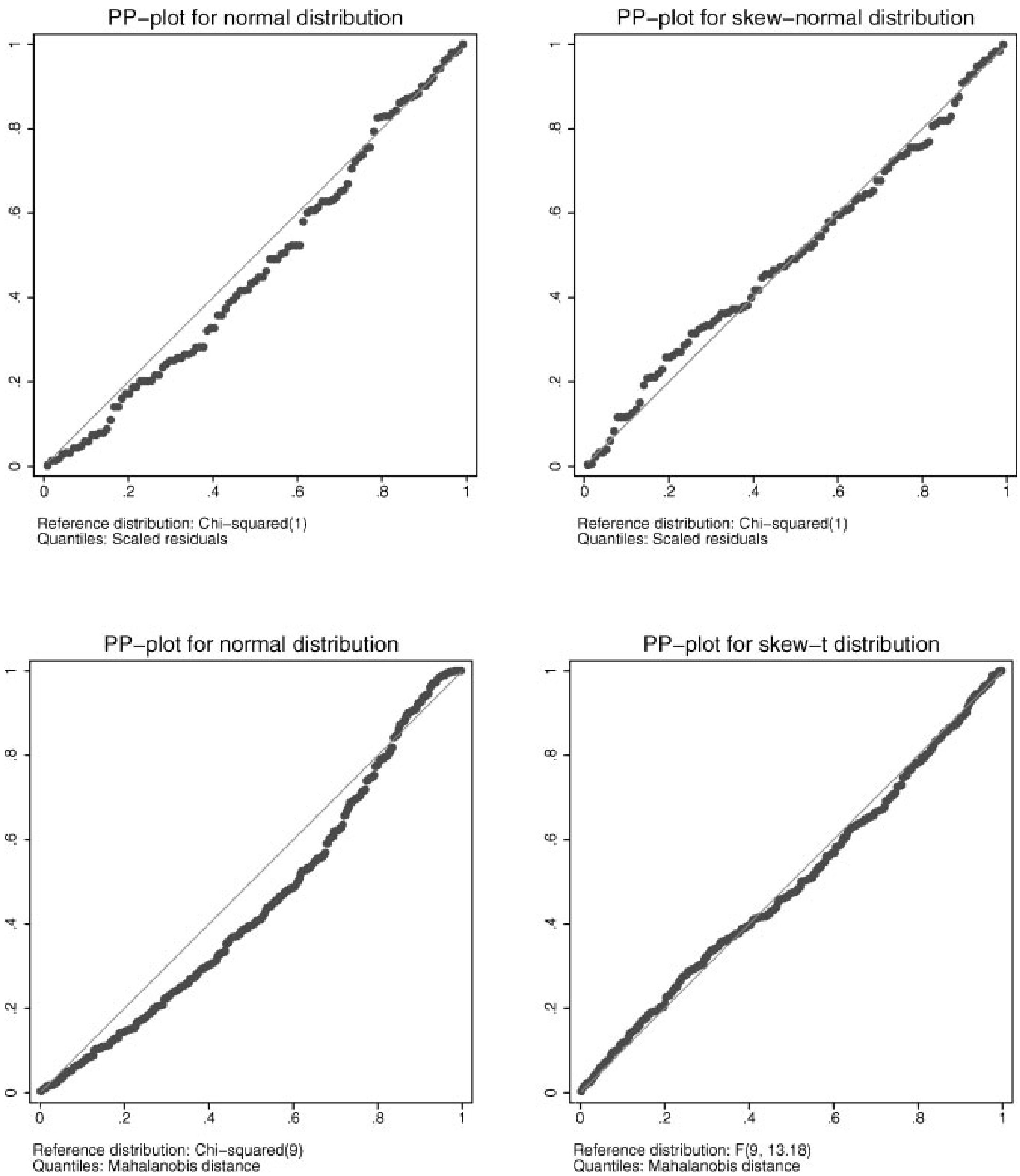

Figure 5. PP-plots for the univariate normal model (top left panel) and the univariate skew-normal model (top right panel) fitted to the log-precipitation for January. PP-plots for the multivariate normal model (bottom left panel) and the multivariate skew-*t* model (bottom right panel) fitted to the log-precipitation for November

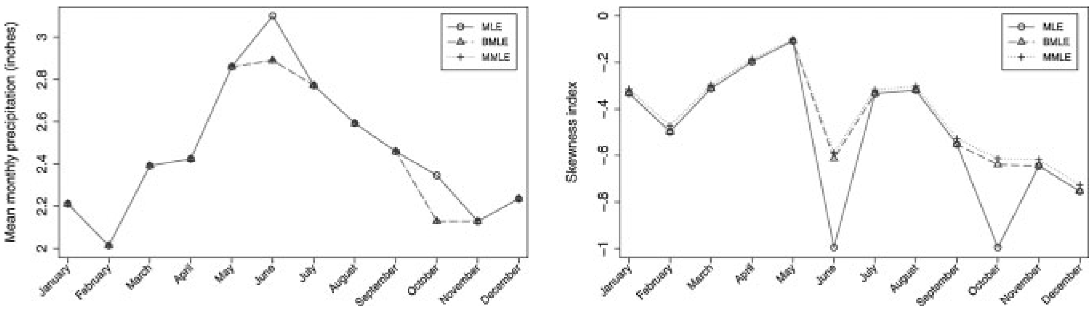

Figure 6. Estimated mean (left panel) and skewness indexes (log scale, right panel) of the log-skew-normal distribution fitted to precipitation for each month

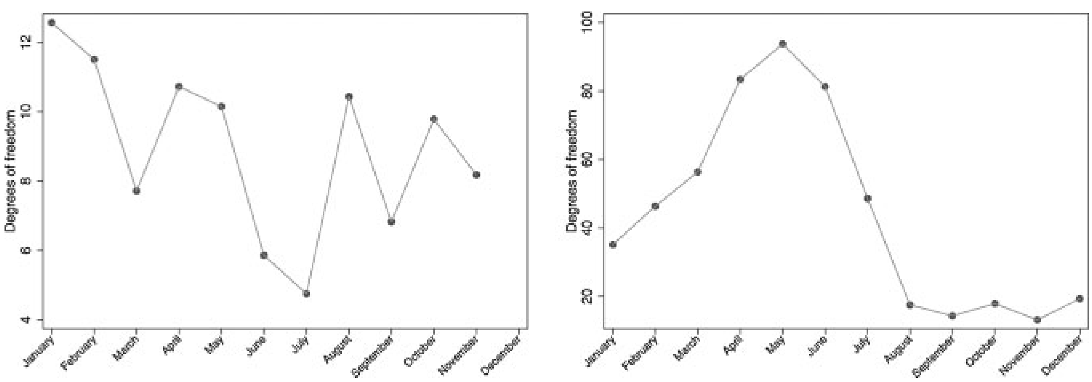

Figure 7. Estimated degrees of freedom of the log-skew-*t* distributions fitted to precipitation for each month. Left panel: univariate analysis of the national precipitation data (December is not included because of the very large estimate). Right panel: multivariate analysis of the regional precipitation data

for which the estimated degrees of freedom are very large. For these months the log-skew-*t* fit may be comparable to that of the log-skew-normal (or log-normal if the skewness index or the estimated shape parameter is not far from zero). For March and, especially, June and July, the estimated degrees of freedom are small suggesting heavier tails of the distribution of the precipitation compared to the log-skew-normal and the log-normal.

Since positive moments of the log-skew-*t* distribution are infinite (Proposition 7), we cannot estimate the mean and marginal skewness on the original scale as for the skew-normal distribution above. However, if estimation of the median (or any other quantile) is of interest, the log-skew-*t* model may be preferable, especially for estimating extreme events, than the skew-normal model due to its robustness properties (Azzalini and Genton, 2008).

If the significance or the precision of the considered quantities is of interest, one can perform a formal significance test (see, for example, Dalla Valle, 2007) or compute confidence intervals. Since all considered quantities are functions of MLEs, a classical Delta method can be used to make inferences about them. Another alternative is to use the bootstrap method to obtain confidence intervals. The latter approach may have better finite sample properties than the former, which relies on large sample sizes.

DOI: 10.1002/env

3.2. US regional scale

We apply multivariate log-skew-normal and log-skew-t models to fit monthly precipitation data over the nine US climatic regions. The multivariate aspect of the model accounts for possible dependence between the precipitation measurements from the regions. More data are usually required to reliably estimate parameters of the multivariate models, compared to their univariate analogs. For multivariate analysis of the regional precipitation data, we consider a 3-months moving window $(t-1, t, t+1)$, resulting in a total of 339 observations per region, to analyze precipitation data for month t.

Bivariate scatter and fitted contour plots visually confirmed a satisfactory fit of the multivariate log-skew-t model to the precipitation data. The multivariate log-skew-normal model exhibits some lack of fit for several months but still fits the data slightly better than the multivariate log-normal model. In this section, we concentrate on the multivariate log-skew-t model. We present the contour and PP-plots for the multivariate skew-t fit to the log-precipitation for November in Figures 8 and 5 (bottom panel), respectively. The bivariate scatter plot suggests that bivariate distributions for some pairs of regions, such as bivariate distributions involving the West region, may deviate from the bivariate normal distribution. The PP-plot confirms this by demonstrating some lack of fit of the multivariate normal distribution (left panel) and the improved fit by the multivariate skew-t distribution (right panel). The flexibility of the skew-t model in capturing both the skewness and heavier tails of the data results in a better fit. As with any richer model, however, such flexibility comes with the price of having to estimate more parameters and may also lead to the problem of overfitting.

The estimated degrees of freedom from the multivariate skew-t models fitted to each month (window) is presented in Figure 7 (right panel). From the graph, there is a noticeable separation in the tail behavior of the observed precipitation distribution over the seasons. The estimates of the degrees of freedom are around 40 and higher for the winter, spring, and summer months and they drop to under 20 for the fall and early winter months suggesting somewhat heavier tailed distributions of the precipitation in these months. We can also see that for November from the bivariate scatter plot depicted in Figure 8: a fair number of precipitation values are observed in the tails of the distribution for some regions. The estimates of the degrees of freedom are fairly large. This may be explained by the fact that only a single parameter, ν, controls the tails of the whole multivariate (nine-dimensional) distribution.

We cannot plot the skew-t skewness indexes for the data in the original scale. However, we can still infer the information about the changes in skewness from the skewness indexes obtained for the log-transformed data. Mardia (1970) defines the measures of multivariate skewness and kurtosis for multivariate data and Arellano-Valle and Genton (2009) present these measures for the multivariate extended skew-t distribution. For the purpose of this exposition, we consider marginal skewness indexes. We compute them as follows. First, we compute the parameters of the marginal distributions using the property of linear transformations of the skew-elliptical random vectors (e.g., Capitanio *et al*., 2003). Then, we use these parameters to compute the univariate skewness index (see Section 3.1), which we refer to as a marginal skewness index. Plots of these marginal skewness indexes computed for each region based on the log-transformed data are given in Figure 9.

As for the national-level precipitation data, the estimated marginal skewness indexes are close to zero for most months in all regions. This suggests that the marginal distributions of the log-precipitation corresponding to the regions are symmetric for these months. In some regions, a negative estimate of skewness is observed for some months. For example, the skewness values of −0.83, −0.89, −0.98, and −0.44 are observed for July, August, September, and October in the Southwest region. So, the use of the skew-normal or skew-t model may be justified for these months.

DOI: 10.1002/env

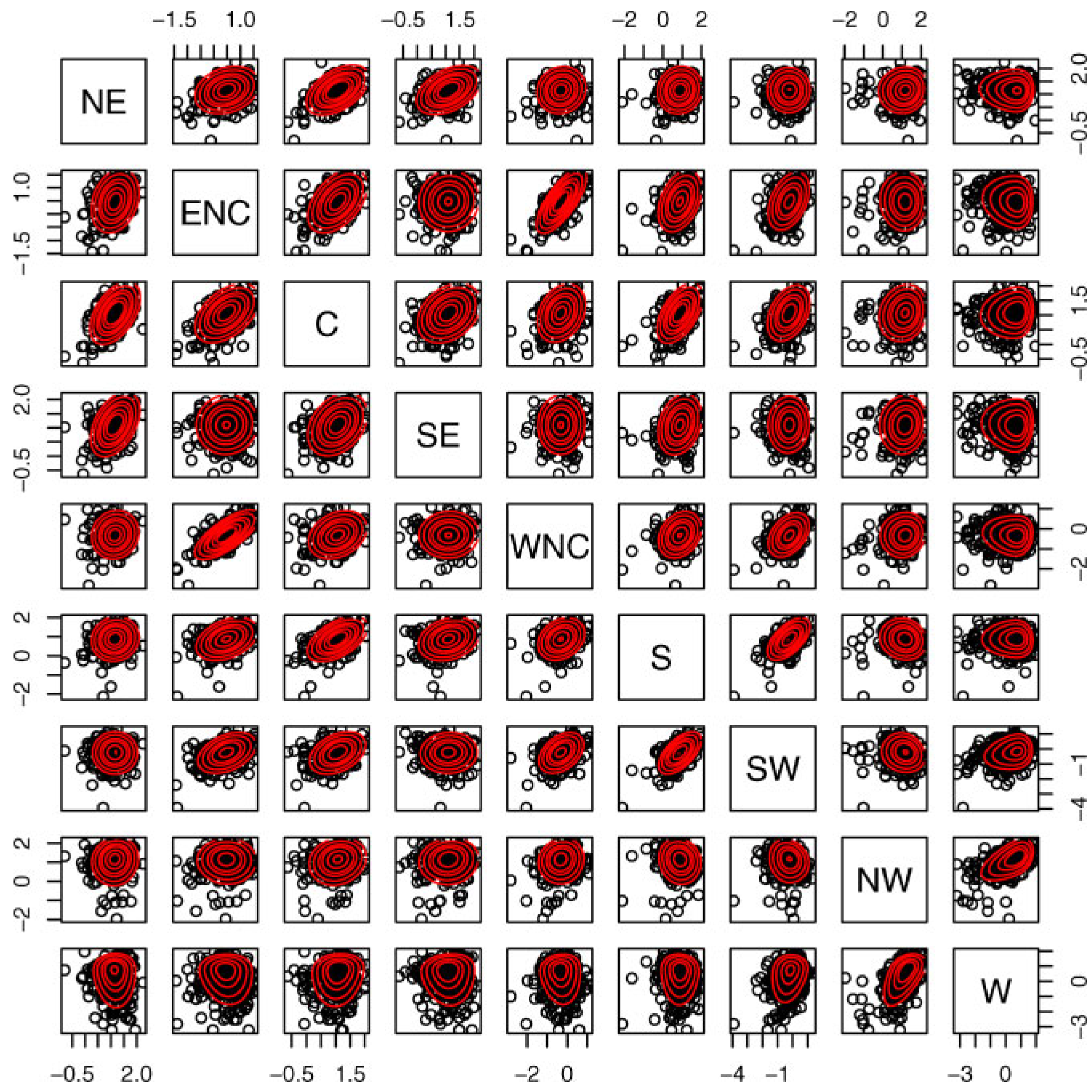

Figure 8. Bivariate scatter plots overlaid with contour plots for the multivariate skew-*t* model fitted to log-precipitation for November. This figure is available in colour online at www.interscience.wiley.com/journal/env

As we briefly mentioned in Section 3.1, using a skew-*t* distribution may be preferable to using a skew-normal distribution when modeling of the tails of the distribution is of interest. This is important, for example, to obtain accurate estimates of extreme quantiles (say $p > 0.95$). The estimated quantiles may then be used to make inferences about extreme events and their magnitude (Beirlant *et al*., 2004). The definition of a univariate quantile is straightforward but a concept of a multivariate quantile is more difficult. Various definitions of a multivariate quantile have been proposed and studied in the literature (e.g., Chaudhuri, 1996; Chakraborty, 2001). As for the skewness index above, we concentrate on the marginal quantiles, i.e. the quantiles obtained using the marginal distributions of a random vector.

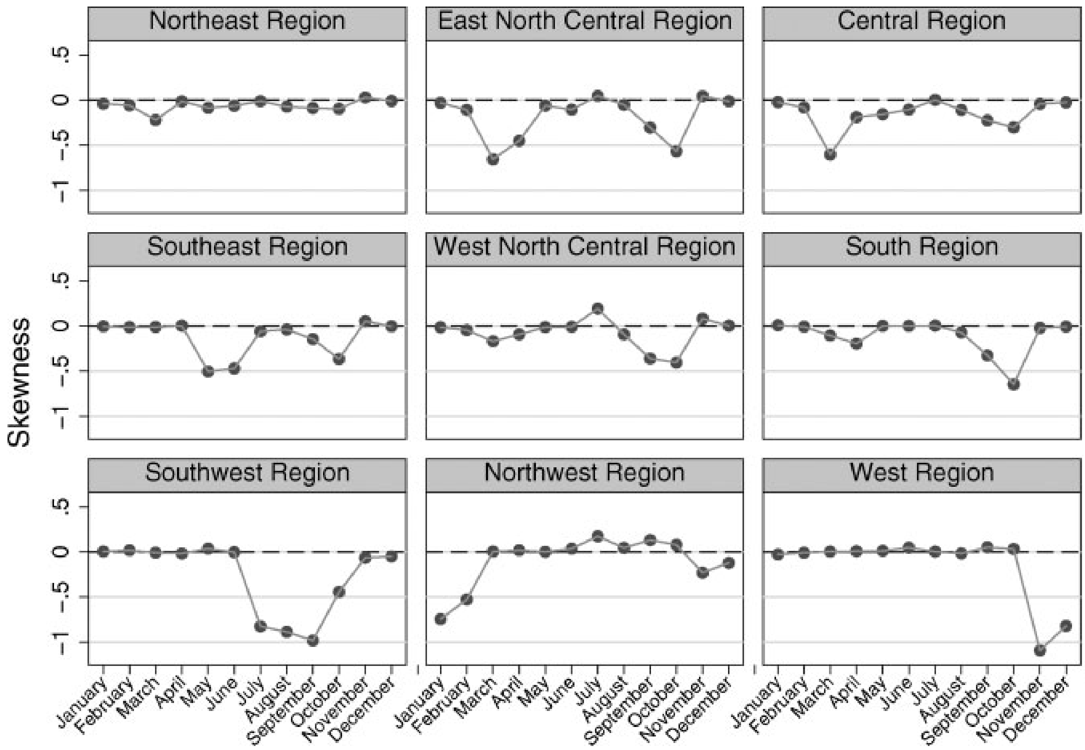

Figure 9. Estimated marginal skewness indexes of the multivariate skew-t distributions fitted to log-precipitation for each month

Figure 10 presents the 95th marginal percentiles corresponding to the multivariate log-skew-t distributions fitted to precipitation data over US climatic regions for all months. We compute the quantiles of the log-skew-t distribution by exponentiating the quantiles obtained from the respective skew-t distribution.

There are only slight variations in the estimated values of the 95th marginal percentile values over months in the Northeast, Central, Southeast, South, and Southwest regions. The variations are much more pronounced in the East North Central, West North Central, Northwest, and West regions. For example, in the West region the estimated values are declining rapidly from 6 inch in January and February to 1 inch in July and August and increase to 5 inch in December. This suggests that observing a monthly average precipitation of 6 inch is very unlikely for July or August whereas there is a 5% chance of observing it in January in the West region.

4. DISCUSSION

The introduced family of multivariate log-skew-elliptical distributions enlarges the family of multivariate distributions with positive support. As for the classical log-elliptical families, its definition is based on the component-wise log transformation, traditionally used to map positive values onto a real line, applied to a skew-elliptical random vector. Although various approaches may be pursued to define multivariate

DOI: 10.1002/env

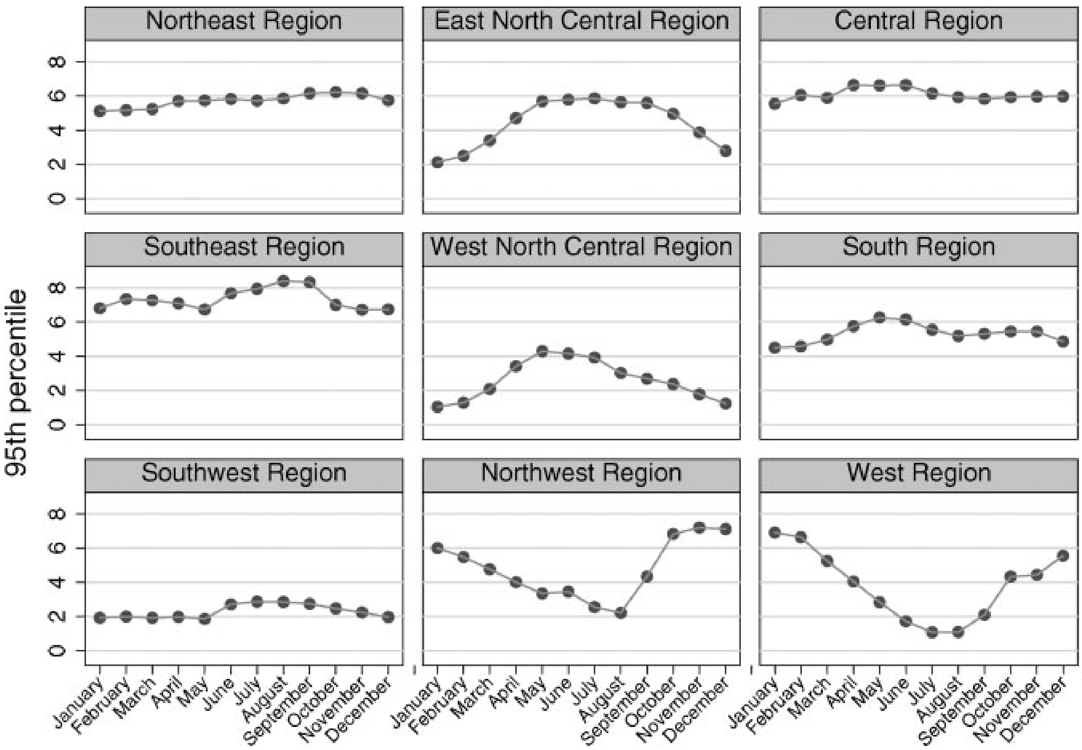

Figure 10. The estimated 95th marginal percentiles of the multivariate log-skew-t distributions fitted to the precipitation for each month

log-skew-elliptical distributions, the considered definition is the most natural one. Its attractiveness is that existing techniques, developed for the skew-elliptical family, can be used to estimate the parameters of and to generate from the defined log-skew-elliptical distribution.

Other extensions to the log-skew-elliptical family, similar to those suggested in the literature for the log-elliptical family, may be considered. One of them is, for example, the addition of the location parameter $\boldsymbol{\theta} \in \mathbb{R}^d$, by considering $\mathbf{X} = \boldsymbol{\theta} + \mathbf{Z}$, where $\mathbf{Z} \sim \text{LSE}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, g^{(d+1)})$. The corresponding PDF will be of the form (4), where $\mathbf{x}$ and $\mathbf{ln}(\mathbf{x})$ are replaced with $(\mathbf{x} - \boldsymbol{\theta})$ and $\mathbf{ln}(\mathbf{x} - \boldsymbol{\theta})$, respectively, and the constraint $\mathbf{x} > \boldsymbol{\theta}$ is placed on $\mathbf{x}$. However, this would require the development of specialized maximum-likelihood estimation techniques or at the very least incorporation of the profile-likelihood estimation approach to estimate the extra parameter, $\boldsymbol{\theta}$.

Two considered special cases, the log-skew-normal and log-skew-t distributions, provide more flexibility in capturing various shapes of the distribution compared to the conventional log-normal model, while introducing only d and $d + 1$ additional parameters, respectively. The d-dimensional shape parameter $\boldsymbol{\alpha}$ controls the skewness and the degrees of freedom parameter ν controls the heaviness of the tails of the underlying distribution on the log-transformed scale. Of course, such extra complexity may not be worthwhile in some applications due to limited availability of the data relative to the dimensionality of the model, in which case difficulties in estimating the model parameters and the problem of overfitting may occur. In the case of the multivariate log-skew-normal and log-skew-t

DOI: 10.1002/env

distributions, a sample size sufficient for estimation of the multivariate log-normal model should often suffice for estimation of these models as well.

We stated and proved some properties, known to hold for the skew-elliptical family and the log-elliptical family, in the case of the log-skew-elliptical family. As for the log-*t* distribution, the positive moments of the log-skew-*t* distribution do not exist. This introduces some limitations to its use in applications for which the estimation of mean (or other moments and their functions) is the goal. However, it may be preferred to the log-skew-normal and the log-normal distributions in applications for which the tails of the distribution are of interest (e.g., the analysis of extreme events).

We also presented the numerical application of both the univariate and the multivariate log-skew-normal and log-skew-*t* models to US monthly precipitation data. The application has some promising results demonstrating that, although for most months and US climatic regions the log-normal distribution provides a satisfactory fit (possibly due to data being averaged over time and space), there are months and regions for which the use of a more flexible parametric model, such as the log-skew-normal and the log-skew-*t*, is beneficial. Also, in the case of the multivariate log-skew-*t* (or log-*t*) models, it would be interesting to see if extending these models to allow for component-specific degrees of freedom results in an improved fit, although the construction of such a generalization is difficult, see the discussion in Azzalini and Genton (2008).

The log-skew-elliptical distributions may also be used to model daily precipitation data. In this case, we should see more improvement in the fit of the log-skew-elliptical models over the log-normal model since no averaging over time is done. Since daily precipitation often have zeros, a mixture model, which is a linear combination of continuous distributions and distributions with point mass at zero, can be considered. A log-skew-elliptical distribution can be used for the continuous component of the mixture model. For example, Chai and Baily (2008) investigated such mixture modeling using the univariate log-skew-normal distribution.

ACKNOWLEDGEMENTS

Genton's research was supported in part by NSF grants DMS-0504896, CMG ATM-0620624, and by Award No. KUS-C1-016-04, made by King Abdullah University of Science and Technology (KAUST). The authors thank Reinaldo B. Arellano-Valle, Kenneth P. Bowman, and a referee for their helpful comments and suggestions. The map of US climatic regions was made available by the National Oceanic and Atmospheric Administration/Department of Commerce.

REFERENCES

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

Arellano-Valle RB, Azzalini A. 2008. The centred parametrization for the multivariate skew-normal distribution. *Journal of Multivariate Analysis* **99**: 1362–1382.

Arellano-Valle RB, Branco MD, Genton MG. 2006. A unified view on skewed distributions arising from selections. *The Canadian Journal of Statistics* **34**: 581–601.

Arellano-Valle RB, Genton MG. 2007. On the exact distribution of linear combinations of order statistics from dependent random variables. *Journal of Multivariate Analysis* **98**: 1876–1894.

Arellano-Valle RB, Genton MG. 2009. Multivariate extended skew-*t* distributions and related families with applications. *Technical report # 080628*, Texas A&M University.

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

Azzalini A. 2005. The skew-normal distribution and related multivariate families (with discussion by Marc G. Genton and a rejoinder by the author). *Scandinavian Journal of Statistics* **32**: 159–200.

Azzalini A. 2006. R package sn: The skew-normal and skew-*t* distributions (version 0.4-6). URL http://azzalini.stat.unipd.it/SN

Azzalini A, Capitanio A. 1999. Statistical applications of the multivariate skew-normal distribution. *Journal of the Royal Statistical Society Series B* **61**: 579–602.
Azzalini A, Capitanio A. 2003. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew *t* distribution. *Journal of the Royal Statistical Society Series B* **65**: 367–389.
Azzalini A, dal Cappello T, Kotz S. 2003. Log-skew-normal and log-skew-t distributions as models for family income data. *Journal of Income Distribution* **11**: 12–20.
Azzalini A, Dalla Valle A. 1996. The multivariate skew-normal distribution. *Biometrika* **83**: 715–726.
Azzalini A, Genton MG. 2008. Robust likelihood methods based on the skew-*t* and related distributions. *International Statistical Review* **76**: 106–129.
Beirlant J, Goegebeur Y, Segers J, Teugels J. 2004. *Statistics of Extremes: Theory and Applications*. John Wiley & Sons: Chichester.
Branco MD, Dey DK. 2001. A general class of multivariate skew-elliptical distributions. *Journal of Multivariate Analysis* **79**: 99–113.
Capitanio A, Azzalini A, Stanghellini E. 2003. Graphical models for skew-normal variates. *Scandinavian Journal of Statistics* **30**: 129–144.
Chai HS, Baily KR. 2008. Use of log-skew-normal distribution in analysis of continuous data with a discrete component at zero. *Statistics in Medicine* **27**: 3643–3655.
Chakraborty B. 2001. On affine equivariant multivariate quantiles. *Annals of the Institute of Statistical Mathematics* **53**: 380–403.
Chaudhuri P. 1996. On a geometric notion of quantiles for multivariate data. *Journal of the American Statistical Association* **91**: 862–872.
Crow EL, Shimizu K. 1988. *Lognormal Distributions: Theory and Applications*. Marcel Dekker: New York.
Dalla Valle A. 2007. A test for the hypothesis of skew-normality in a population. *Journal of Statistical Computation and Simulation* **77**: 63–77.
Fang K-T, Kotz S, Ng KW. 1990. *Symmetric Multivariate and Related Distributions*. Chapman & Hall: New York.
Genton MG. 2004. *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*. Chapman & Hall/CRC: Boca Raton.
Marchenko YV, Genton MG. 2009. A suite of commands for fitting the skew-normal and skew-*t* distributions. *Stata Journal* (in preparation).
Mardia KV. 1970. Measures of multivariate skewness and kurtosis with applications. *Biometrika* **36**: 519–530.
Pewsey A. 2000. Problems of inference for Azzalini's skew-normal distribution. *Journal of Applied Statistics* **27**: 859–770.
Pewsey A. 2006. Some observations on a simple means of generating skew distributions. In *Advances in Distribution Theory, Order Statistics, and Inference*, Balakrishnan N, Castillo E, Sarabia JM (eds). Statistics for Industry and Technology, Birkhauser: Boston, MA; 75–84.
R Development Core Team. 2008. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria, ISBN 3-900051-07-0.
Sartori N. 2006. Bias prevention of maximum likelihood estimates for scalar skew normal and skew *t* distributions. *Journal of Statistical Planning and Inference* **136**: 4259–4275.
StataCorp. 2007. *Stata Statistical Software: Release 10*. StataCorp LP: College Station, TX.
Wilks D. S. 2006. *Statistical Methods in the Atmospheric Sciences* (2nd edn). Academic Press: New York.
Wilson PS, Toumi R. 2005. A fundamental probability distribution for heavy rainfall. *Geophysical Research Letters* **32**: L14812. doi:10.1029/2005GL022465

APPENDIX

Proof of Proposition 1 (selection representation 1). To show that $\mathbf{Z} \sim \mathrm{LSE}_d(\mathbf{0}, \bar{\Omega}, \boldsymbol{\alpha}, g^{(d+1)})$, we apply a property that if $\mathbf{Z} \overset{d}{=} (\mathbf{V}|\tilde{U}_0 < \boldsymbol{\alpha}^\top \mathbf{U})$, then $\psi(\mathbf{Z}) \overset{d}{=} (\psi(\mathbf{V})|\tilde{U}_0 < \boldsymbol{\alpha}^\top \mathbf{U})$ with $\psi(\cdot) = \mathbf{ln}(\cdot)$. Then, the result follows from Definition 1 and a selection representation of a skew-elliptical random vector as given by, for example, Arellano-Valle and Genton (2009). Then, $\mathbf{X} \sim \mathrm{LSE}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, g^{(d+1)})$, since $\mathbf{ln}(\mathbf{X}) = \boldsymbol{\xi} + \omega \mathbf{ln}(\mathbf{Z}) \sim \mathrm{SE}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, g^{(d+1)})$. ■

Proof of Proposition 2 (selection representation 2). Although this result can also be shown using similar arguments as in the proof for the selection representation 1, we choose a different approach here. We use properties of elliptical and log-elliptical distributions (Fang *et al*., 1990, pp. 45, 56) to show that the PDF of $\mathbf{Z}$ is of the form (4). It can be written as

$$f_{\mathbf{Z}}(\mathbf{z}) = f_{\mathbf{V}|U_0>0}(\mathbf{z}) = f_{\mathbf{V}}(\mathbf{z})\frac{P(U_0 > 0|\mathbf{V} = \mathbf{z})}{P(U_0 > 0)}. \tag{A.1}$$

Using properties of elliptical distributions, $\mathbf{U} \sim \mathrm{EC}_d(\mathbf{0}, \bar{\Omega}, g^{(d)})$ and $U_0|\mathbf{U} \sim \mathrm{EC}(\boldsymbol{\delta}^\top \bar{\Omega}^{-1}\mathbf{u}, 1 - \boldsymbol{\delta}^\top \bar{\Omega}\boldsymbol{\delta}, g_{Q_{\mathbf{u}}^{\mathbf{0},\bar{\Omega}}})$, where $Q_{\mathbf{u}}^{\mathbf{0},\bar{\Omega}} = \mathbf{u}^\top \bar{\Omega}\mathbf{u}$. Then $\mathbf{V} = \mathbf{exp}(\mathbf{U})$ follows a d-dimensional multivariate log-elliptical distribution with PDF

$$f_{\mathbf{V}}(\mathbf{z}) = \left(\prod_{i=1}^{d} z_i^{-1}\right) f_{\mathbf{U}}\{\mathbf{ln}(\mathbf{z})\} = \left(\prod_{i=1}^{d} z_i^{-1}\right) f_d\{\mathbf{ln}(\mathbf{z}); \mathbf{0}, \bar{\Omega}, g^{(d)}\}.$$

Let $U_0^\star = \{U_0 - \boldsymbol{\delta}^\top \bar{\Omega}^{-1}\mathbf{ln}(\mathbf{z})\}/(1 - \boldsymbol{\delta}^\top \bar{\Omega}^{-1}\boldsymbol{\delta})^{1/2}$, then $U_0^\star \sim \mathrm{EC}(0, 1, g_{Q_{\mathbf{ln}(\mathbf{z})}^{\mathbf{0},\bar{\Omega}}})$, where $Q_{\mathbf{ln}(\mathbf{z})}^{\mathbf{0},\bar{\Omega}} = \mathbf{ln}(\mathbf{z})^\top \bar{\Omega}\mathbf{ln}(\mathbf{z})$. Hence

$$P(U_0 > 0|\mathbf{V} = \mathbf{z}) = P\{U_0 > 0|\mathbf{U} = \mathbf{ln}(\mathbf{z})\} = P\{U_0^\star < \boldsymbol{\alpha}^\top \mathbf{ln}(\mathbf{z})|\mathbf{U} = \mathbf{ln}(\mathbf{z})\} = F\{\boldsymbol{\alpha}^\top \mathbf{ln}(\mathbf{z}); g_{Q_{\mathbf{ln}(\mathbf{z})}^{\mathbf{0},\bar{\Omega}}}\}.$$

Since the distribution of U_0 is symmetric, $P(U_0 > 0) = 1/2$. Substituting the terms obtained above into Equation (A.1), we arrive at the PDF of $\mathbf{Z}$

$$f_{\mathbf{Z}}(\mathbf{z}) = 2\left(\prod_{i=1}^{d} z_i^{-1}\right) f_d\{\mathbf{ln}(\mathbf{z}); \mathbf{0}, \bar{\Omega}, g^{(d)}\}F\{\boldsymbol{\alpha}^\top \mathbf{ln}(\mathbf{z}); g_{Q_{\mathbf{ln}(\mathbf{z})}^{\mathbf{0},\bar{\Omega}}}\}$$

which is a special case of Equation (4) with $\boldsymbol{\xi} = \mathbf{0}$ and $\Omega = \bar{\Omega}$. Then, $\mathbf{X} \sim \mathrm{LSE}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, g^{(d+1)})$, since $\mathbf{ln}(\mathbf{X}) = \boldsymbol{\xi} + \omega\mathbf{ln}(\mathbf{Z}) \sim \mathrm{SE}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, g^{(d+1)})$. ■

Proof of Proposition 3 (log-scale-mixture-skew-normal)**.** This representation arises from the definition of a log-skew-elliptical random vector and skew-scale mixture representation of a skew-elliptical random vector: if $\mathbf{X} \sim \mathrm{SN}_d(\mathbf{0}, \Omega, \boldsymbol{\alpha})$, then $\mathbf{Y} \stackrel{d}{=} \boldsymbol{\xi} + K(\eta)^{1/2}\mathbf{X}$ has the PDF of the form (8), $\mathbf{Y} \sim \mathrm{SMSN}_d\{\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, K(\eta), H(\eta)\}$; see Azzalini and Capitanio (2003), among others. Since $\mathbf{X} = \mathbf{exp}(\boldsymbol{\xi}) \odot \mathbf{Z}^{K^{1/2}(\eta)} = \mathbf{exp}\{\boldsymbol{\xi} + K^{1/2}(\eta)\mathbf{ln}(\mathbf{Z})\} = \mathbf{exp}(\mathbf{Y})$, the result follows from Definition 1. ■

Proof of Proposition 4 (marginal distribution)**.** By definition, we have that the vector $\mathbf{ln}(\mathbf{X}) = \{\mathbf{ln}(\mathbf{X}_1)^\top, \mathbf{ln}(\mathbf{X}_2)^\top\}^\top$, $\mathbf{ln}(\mathbf{X}) \sim \mathrm{SE}_d(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, g^{(d)})$. Using the result about the marginal distribution of a skew-elliptical random vector, $\mathbf{ln}(\mathbf{X}_1) \sim \mathrm{SE}_q(\boldsymbol{\xi}_1, \Omega_{11}, \boldsymbol{\alpha}_1^\star, g^{(d)})$ and $\mathbf{ln}(\mathbf{X}_2) \sim \mathrm{SE}_{d-q}(\boldsymbol{\xi}_2, \Omega_{22}, \boldsymbol{\alpha}_2^\star, g^{(d)})$ with parameters as defined in Equation (9). Then, the result follows from Definition 1 of a log-skew-elliptical random vector. ■

Proof of Proposition 6 (mixed moments)**.** Provided the moment exists

$$E\left(\prod_{i=1}^{d} X_i^{n_i}\right) = E\left\{\prod_{i=1}^{d} e^{n_i \ln(X_i)}\right\} = E\left\{e^{\sum_{i=1}^{d} n_i \ln(X_i)}\right\} = E\left\{e^{\mathbf{n}^\top \mathbf{ln}(\mathbf{X})}\right\} = M_{\mathbf{ln}(\mathbf{X})}(\mathbf{n}). \quad ■$$

DOI: 10.1002/env

Proof of Proposition 7 (log-skew-t moments)**.** Without loss of generality, let $\boldsymbol{\xi} = \mathbf{0}$ and $\Omega = \bar{\Omega}$. We compute the moments directly from the log-skew-t PDF (6). Let $U_{\mathbf{ln}(\mathbf{x})} = \boldsymbol{\alpha}^\top \mathbf{ln}(\mathbf{x}) \left(\frac{\nu+d}{\nu+Q_{\mathbf{ln}(\mathbf{x})}}\right)^{1/2}$ and $\bar{\Omega}^{-1} = (\tilde{\rho}_{ij})_{i,j=1}^d$, where $\tilde{\rho}_{ii} > 0$ and $\tilde{\rho}_{ij} = \tilde{\rho}_{ji}$. Let $\mathbb{R}_a^d$ denote $(a, \infty)^d$ and $D^d\mathbf{x}$ denote $dx_1 dx_2 \ldots dx_d$. For any $n_i \geq 0$, $i = 1, \ldots, d$ the condition $\sum_{i=1}^d n_i > 0$ requires that at least one of the n_i's is nonzero. Suppose that $n_d > 0$.

$$
\begin{aligned}
E\left(\prod_{i=1}^d X_i^{n_i}\right) &= \int_{\mathbb{R}_0^d} \left(\prod_{i=1}^d x_i^{n_i}\right) f_{\mathrm{LST}_d}(\mathbf{x}; \mathbf{0}, \bar{\Omega}, \boldsymbol{\alpha}, \nu) D^d\mathbf{x}. \\
&= 2a_\nu b_\nu \int_{\mathbb{R}_0^d} \prod_{i=1}^d x_i^{n_i-1} \left(1 + \frac{Q_{\mathbf{ln}(\mathbf{x})}}{\nu}\right)^{-\frac{\nu+d}{2}} \int_{-\infty}^{U_{\mathbf{ln}(\mathbf{x})}} \left(1 + \frac{u^2}{\nu+d}\right)^{-\frac{\nu+d+1}{2}} \mathrm{d}u\, D^d\mathbf{x} \\
&= 2a_\nu b_\nu \int_{\mathbb{R}_0^{d-1}} \prod_{i=1}^{d-1} x_i^{n_i-1} \int_0^\infty x_d^{n_d-1} \left(1 + \frac{Q_{\mathbf{ln}(\mathbf{x})}}{\nu}\right)^{-\frac{\nu+d}{2}} \\
&\quad \times \int_{-\infty}^{U_{\mathbf{ln}(\mathbf{x})}} \left(1 + \frac{u^2}{\nu+d}\right)^{-\frac{\nu+d+1}{2}} \mathrm{d}u\, \mathrm{d}x_d D^{d-1}\mathbf{x} \qquad \text{(A.2)}
\end{aligned}
$$

where a_ν and b_ν are normalization constants.

The quadratic form $Q_{\mathbf{ln}(\mathbf{x})}$ can be rewritten as a function of x_d, $Q_{\mathbf{ln}(\mathbf{x})} = A_{\mathbf{ln}(\mathbf{x}_{-d})} + (\ln x_d + B_{\mathbf{ln}(\mathbf{x}_{-d})})^2$, where $A_{\mathbf{ln}(\mathbf{x}_{-d})}$ and $B_{\mathbf{ln}(\mathbf{x}_{-d})}$ do not depend on x_d (are functions of only $\ln x_1, \ldots, \ln x_{d-1}$ and $\tilde{\rho}_{ij}$, $i, j = 1, \ldots, d$). Then, $U_{\mathbf{ln}(\mathbf{x})} \longrightarrow \alpha_d \sqrt{\nu + d}$ as $x_d \to \infty$, and so $\int_{-\infty}^{U_{\mathbf{ln}(\mathbf{x})}} (1 + \frac{u^2}{\nu+d})^{-\frac{\nu+d+1}{2}}\, \mathrm{d}u \longrightarrow \int_{-\infty}^{\alpha_d\sqrt{\nu+d}} (1 + \frac{u^2}{\nu+d})^{-\frac{\nu+d+1}{2}}\, \mathrm{d}u > 0$ as $x_d \to \infty$. Thus, there is a $c_{\nu,\alpha_d} > 0$ (does not depend on $\mathbf{x}$), such that

$$
\int_{-\infty}^{U_{\mathbf{ln}(\mathbf{x})}} \left(1 + \frac{u^2}{\nu+d}\right)^{-\frac{\nu+d+1}{2}} \mathrm{d}u > c_{\nu,\alpha_d}, \text{ for large } x_d > x_d^\star.
$$

Then

$$
\begin{aligned}
\text{(A.2)} &> 2a_\nu b_\nu \int_{\mathbb{R}_0^{d-1}} \prod_{i=1}^{d-1} x_i^{n_i-1} \int_{x_d^\star}^\infty x_d^{n_d-1} \left(1 + \frac{Q_{\mathbf{ln}(\mathbf{x})}}{\nu}\right)^{-\frac{\nu+d}{2}} \int_{-\infty}^{U_{\mathbf{ln}(\mathbf{x})}} \left(1 + \frac{u^2}{\nu+d}\right)^{-\frac{\nu+d+1}{2}} \mathrm{d}u\, \mathrm{d}x_d D^{d-1}\mathbf{x} \\
&> 2a_\nu b_\nu c_{\nu,\alpha_d} \int_{\mathbb{R}_0^{d-1}} \prod_{i=1}^{d-1} x_i^{n_i-1} \left[\int_{x_d^\star}^\infty x_d^{n_d-1} \left\{1 + \frac{A_{\mathbf{ln}(\mathbf{x}_{-d})}}{\nu} + \frac{(\ln x_d + B_{\mathbf{ln}(\mathbf{x}_{-d})})^2}{\tilde{\rho}_{dd}\nu}\right\}^{-\frac{\nu+d}{2}} \mathrm{d}x_d\right] D^{d-1}\mathbf{x}.
\end{aligned}
$$

The innermost integral over x_d (in square brackets) diverges at infinity for any $n_d > 0$, and, therefore, the d-dimensional integral from the last step diverges for any $n_i \geq 0$, $i = 1, \ldots, d-1$. As such, the integral (A.2) also diverges and $E(\prod_{i=1}^d X_i^{n_i}) = \infty$ for any $n_i \geq 0$, $i = 1, \ldots, d-1$ and $n_d > 0$. More generally, $E(\prod_{i=1}^d X_i^{n_i}) = \infty$ for any $n_i \geq 0$, $i = 1, \ldots, d$, such that $\sum_{i=1}^d n_i > 0$. ■