Truncated skew-normal distributions: moments, estimation by weighted moments and application to climatic data

by C. Flecher[1,2], D. Allard[1] and P. Naveau[2]
[1] Biostatistics and Spatial Processes,
INRA, Agroparc 84914 Avignon, France
[2] Laboratoire des Sciences du Climat et de l'Environnement,
CNRS Gif-sur-Yvette, France

Running title: Truncated skew-normal distributions

Abstract In this paper we derive an expression of the m^{th} order moments and some weighted moments of truncated skew-normal distributions. We link these formulas to previous results for truncated distributions and non truncated skew-normal distributions. Methods to estimate skew-normal parameters using weighted moments are proposed and compared to other classical techniques. In a second step we propose to model the distribution of relative humidity with a truncated skew-normal distribution.

Key words: Truncated skew-normal distributions, inference methods, classical moments, weighted moments, relative humidity.

1 Introduction

In many applications, the probability distribution function (pdf) of some observed variables can be simultaneously skewed and restricted to a fixed interval. For example, variables such as pH, grades, and humidity in environmental studies, have upper and lower physical bounds and their pdfs are not necessarily symmetric within these bounds. To illustrate such behaviors, Figure 1 shows the histograms per season and the estimated pdfs of daily relative humidity measurements made in Toulouse (France) from 1972 to 1999.

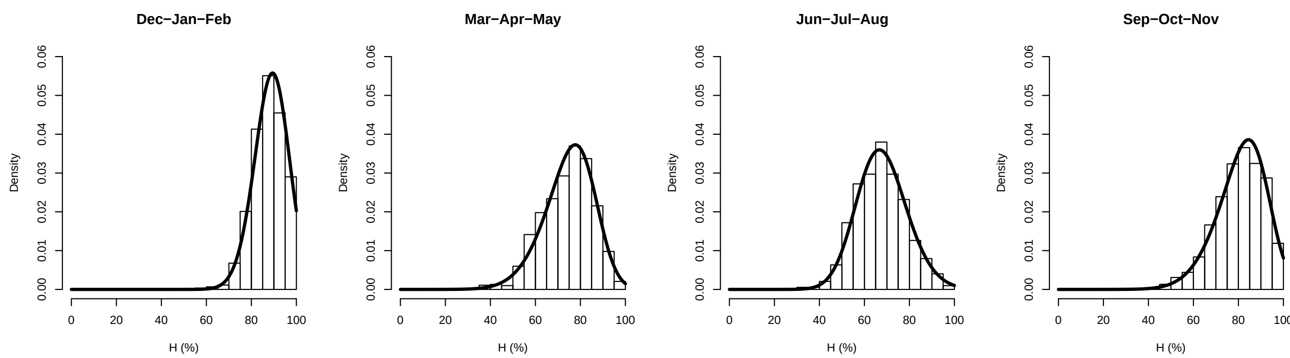

Figure 1: Histograms and estimated densities of daily relative humidity measurements made in Toulouse (France) from 1972 to 1999. Each panel corresponds to a season.

All observations belong to the interval $[0, 100]$ and skewness is apparent, especially during the Spring and Fall seasons. There are many strategies available to model such skewed and bounded data. In this paper our approach is to conceptually view such observations as truncated measurements originating from a flexible skewed distribution. More precisely we assume that the truncation bounds are known and we focus on the following skew-normal pdf defined by Azzalini (1985)

$$f_{\mu,\sigma,\lambda}(x) = \frac{2}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)\Phi\left(\lambda\frac{x-\mu}{\sigma}\right), \qquad (1)$$

where $\mu \in \mathbb{R}$, $\sigma > 0$ and $\lambda \in \mathbb{R}$ represent the location, scale and shape parameters, respectively. The notations ϕ and Φ correspond to the pdf and the cumulative distribution function (cdf) of the standard normal distribution, respectively. The skew-normal distribution (1) can be obtained from a bivariate Gaussian random vector, by conditioning one component on the second one being above (or below) a given threshold (Azzalini, 1985). The notation $X \sim SN(\mu, \sigma, \lambda)$ represents a random variable, X, following (1) and $F_{\mu,\sigma,\lambda}$ denotes its cdf. The particular case $\lambda = 0$ corresponds to the classical normal distribution with mean μ and variance σ^2. In the following, the pdf and the cdf of a $SN(0, 1, \lambda)$ are simply denoted f_λ and F_λ. In this context our main objective is to propose and study a novel method-of-moments approach for estimating the parameters of (1) in presence of truncation.

Among others, Martinez *et al.* (2008) recently studied the moments of the skew-normal defined by (1). Dhrymes (2005) provided a recursive relationship among the r-order moments of a standard normal distribution, say Z, truncated at a and b

$$m_r(a,b) = (r-1)m_{r-2}(a,b) - \frac{[z^{r-1}\phi(z)]_a^b}{[\Phi(z)]_a^b}, \quad \text{for } r = 1, 2, \ldots, \tag{2}$$

where $m_r(a,b) = \mathbb{E}[Z^r \mid a < Z \leq b]$ with $a < b$, $[\Phi(x)]_a^b$ denotes $\Phi(b) - \Phi(a)$, $m_0(a,b) = 1$ and $m_{-1}(a,b)$ can be any finite value. In Section 2 we combine the results of Martinez *et al.* (2008) with the approach of Dhrymes (2005) in order to derive the moments of a truncated skew-normal distribution.

Following the work of Flecher *et al.* (2009a), we also compute a special type of weighted moments of truncated skew-normal distributions. In Section 3, these theoretical expressions of moments allow us to propose a new method-of-moments approach to estimate the parameters of such truncated distributions. In Section 4, this estimation method is compared to the classical method-of-moments and maximum likelihood method (Mills, 1955) using simulated data. Daily air relative humidity measurements are also analyzed in this section. All proofs are relegated to the Appendix to improve readability of the paper.

2 Moments of truncated skew-normal variables

Lemma 1 presents an explicit, though cumbersome, solution to (2).

Lemma 1. *Let Z be a standard Gaussian random variable and $-\infty \leq a < b \leq +\infty$. Let us denote $m_r(a,b) = \mathbb{E}(Z^r \mid a < Z \leq b)$ the r^{th} moment of the truncated Z. Then,*

$$\begin{aligned} m_{2k}(a,b) &= (2k-1)!! \left(1 - \sum_{i=1}^{k} \frac{1}{(2i-1)!!} \frac{[z^{2i-1}\phi(z)]_a^b}{[\Phi(z)]_a^b}\right), \quad \textit{for } k = 1, 2, \ldots, \\ m_{2k+1}(a,b) &= -\sum_{i=0}^{k} \frac{(2k)!!}{(2i)!!} \frac{[z^{2i}\phi(z)]_a^b}{[\Phi(z)]_a^b}, \quad \textit{for } k = 0, 1, \ldots, \end{aligned} \tag{3}$$

where $n!!$ denotes the double factorial defined by Arfken (1985) as

$$n!! = \begin{cases} 1, & \textit{if } n = -1,\ n = 0 \textit{ or } n = 1, \\ n \times (n-2)!! & \textit{if } n \geq 2. \end{cases}$$

As $(a,b) \to (-\infty, \infty)$, the moments $m_{2k}(a,b)$ and $m_{2k+1}(a,b)$ tend to $(2k-1)!!$ and zero, respectively. The latter two values correspond to the classical Gaussian moments.

We now introduce the truncated skew-normal distribution as a truncation of the skew-normal distribution in 1. Its pdf is

$$f_{\mu,\sigma,\lambda}(x \mid a < X \leq b) = \begin{cases} \frac{1}{[F_{\mu,\sigma,\lambda}(x)]_a^b} f_{\mu,\sigma,\lambda}(x), & \text{if } a < x \leq b, \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

where $X \sim SN(\mu, \sigma, \lambda)$ and $-\infty \leq a < b \leq +\infty$ represents the range of the truncation. Let us consider the simple case $(\mu, \sigma) = (0, 1)$.

Proposition 1. *Let X be a $SN(0, 1, \lambda)$. Let us denote $s_{\lambda,r}(u, v) = \mathbb{E}[X^r \mid u < X \leq v]$ with $u < v$ the r^{th} moment of the truncated variable. The following recursive relationship holds,*

$$s_{\lambda,r}(u, v) = (r - 1)s_{\lambda,r-2}(u, v) + r_{\lambda,r}(u, v), \quad \textit{for } r = 1, 2, \ldots, \tag{5}$$

where $s_{\lambda,0}(u, v) = 1$ and $s_{\lambda,-1}$ can be any finite value,

$$r_{\lambda,r}(u, v) = -\frac{[x^{r-1} f_\lambda(x)]_u^v}{[F_\lambda(x)]_u^v} + \frac{2}{\sqrt{2\pi}} \frac{\lambda}{\lambda_*^r} \frac{[\Phi(\lambda_* x)]_u^v}{[F_\lambda(x)]_u^v} m_{r-1}(\lambda_* u, \lambda_* v),$$

where $\lambda_* = (1 + \lambda^2)^{1/2}$. From (5), we can derive

$$s_{\lambda,2p}(u, v) = (2p - 1)!! + \sum_{k=1}^{p} \frac{(2p - 1)!!}{(2k - 1)!!} r_{\lambda,2k}(u, v), \quad \textit{with } p = 1, 2, \ldots, \tag{6}$$

$$s_{\lambda,2p+1}(u, v) = \sum_{k=0}^{p} \frac{(2p)!!}{(2k)!!} r_{\lambda,2k+1}(u, v), \quad \textit{with } p = 0, 1, \ldots . \tag{7}$$

Martinez's *et al.* (2008) results can be viewed as limiting cases of (6) and (7)

$$\lim_{(u,v)\to(-\infty,+\infty)} s_{\lambda,2p}(u, v) = (2p - 1)!!, \text{ with } p = 0, 1, \ldots,$$

$$\lim_{(u,v)\to(-\infty,+\infty)} s_{\lambda,2p+1}(u, v) = \frac{2}{\sqrt{2\pi}} \sum_{k=0}^{p} \frac{(2p)!!}{(2k)!!} (2k - 1)!! \frac{\lambda}{(1 + \lambda^2)^{k+1/2}}, \text{ with } p = 0, 1, \ldots .$$

Equalities (6) and (7) tell us that odd (respectively even) moments of truncated skew-normal distributions can be interpreted as linear combinations of even (respectively odd) moments of the normal distribution truncated at $\lambda_* u$ and $\lambda_* v$. If $\lambda = 0$, Equation (5) and Proposition 1 are equivalent to the recursive equation provided in Dhrymes (2005) and Lemma 1, respectively. The restricting condition $(\mu, \sigma) = (0, 1)$ can be easily removed and leads to the following proposition.

Proposition 2. *Let $X \sim SN(\mu, \sigma, \lambda)$. Then, we have*

$$\mathbb{E}[X^m \mid a < X \leq b] = \sum_{r=0}^{m} C_m^r \mu^{m-r} \sigma^r s_{\lambda,r}(u, v). \tag{8}$$

where $u = (a - \mu)/\sigma$, $v = (b - \mu)/\sigma$, $C_m^r = \begin{pmatrix} m \\ r \end{pmatrix}$ is a binomial coefficient and $s_{\lambda,r}(u, v)$ is defined in Proposition 1.

Proposition 2 provides moments of any order and consequently the first three moments can be used to implement a classical method-of-moment. Besides the complexity of deriving the explicit expression of the third moment, its estimation is usually tainted with a large variance because X^3 can take large values. An alternative route is to derive other types of moments. Following the work of Hosking *et al.* (1985) and Diebolt *et al.* (2008), Flecher *et al.* (2009a) introduced and studied probability weighted moments for skew-normal distributions. The basic idea is to compute moments of the type $\mathbb{E}[X^s \Phi^r(X)]$, where s and r are small integers. Here we concentrate on the case $s = 0$ and $r = 1$. We obtain the following proposition.

Proposition 3. *Let $X \sim SN(\mu, \sigma, \lambda)$ and $a < b$. Then,*

$$\mathbb{E}[\Phi(X) \mid a < X \leq b] = 2\Phi_2(\mathbf{0}; \boldsymbol{\nu}_+, \mathbf{I}_2 + \sigma^2 \mathbf{D}_+ \mathbf{D}_+^t) \frac{[F_+(x)]_a^b}{[F_{\mu,\sigma,\lambda}(x)]_a^b} = \varphi(\mu, \sigma, \lambda), \tag{9}$$

where F_+ is the cumulative distribution function of a closed skew-normal variable $CSN_{1,2}(\mu, \sigma^2, \mathbf{D}_+, \boldsymbol{\nu}_+, \mathbf{I}_2)$, $\mathbf{I}_2$ is the identity matrix of size two, $\mathbf{D}_+ = (1, \lambda/\sigma)^t$ and $\boldsymbol{\nu}_+ = (-\mu, 0)^t$.

The definition and some properties of closed skew-normal distributions can be found in the book edited by Genton (2004), see in particular Chapter 2 written by González-Farías *et al.*

3 A method-of-moments approach for truncated skew-normal distributions

To apply a method-of-moments approach to realizations generated from a truncated skew-normal pdf, three explicit equations are necessary to estimate the three skew-normal parameters. At this stage, two possible approaches can be implemented. Using Proposition 2, we can compute the three first moments $\mathbb{E}[X^m \mid a < X \leq b]$ with $m = 1, 2, 3$. This is the classical method-of-moments approach and we will call it MOM hereafter. A second alternative is to take advantage of Proposition 3. Flecher *et al.* (2009a) showed that this strategy applied to non-truncated skew-normal data leads to better estimates than the classical MOM approach. For truncated data, the following set of three equations is the core of our estimation scheme, hereafter called Method of Weighted Moments (MWM),

$$\begin{aligned} \mathbb{E}[X | a < X \leq b] &= \mu s_{\lambda,0}(u, v) + \sigma s_{\lambda,1}(u, v), \\ \mathbb{V}ar(X | a < X \leq b) &= \sigma^2 \left(s_{\lambda,2}(u, v) - s_{\lambda,0}^2(u, v) \right), \\ \mathbb{E}[\Phi(X) | a < X \leq b] &= \varphi(\mu, \sigma, \lambda), \end{aligned} \tag{10}$$

with $u = (a - \mu)/\sigma$, $v = (b - \mu)/\sigma$ and $s_{\lambda,r}(u, v)$ is defined in Proposition 1. The name "Method of Weighted Moments" can be justified by recalling that the system defined by (10) stems from the "weighted moments" $\mathbb{E}(X^r \Phi^s(X)|a < X \leq b)$ with $(r, s) = (1, 0)$, $(r, s) = (2, 0)$ and $(r, s) = (0, 1)$. To estimate the parameters, we used the following scheme:

- Compute the empirical moments corresponding to $\mathbb{E}[X \mid a < X \leq b]$,$\mathbb{V}ar(X \mid a < X \leq b)$ and $\mathbb{E}[\Phi(X) \mid a < X \leq b]$,
- solve (10) for (μ, σ, λ) using `nlminb` function in `R`.

The complexity of the equations, in particular the last equation related to the weighted moment, does not allow us to obtain distributional results for the estimators $(\hat{\mu}, \hat{\sigma}, \hat{\lambda})$. The performance of this estimation scheme must therefore be assessed by simulation.

Figure 2 shows how $\mathbb{E}[X^3 \mid a < X \leq b]$ (right panel) and $\mathbb{E}[\Phi(X) \mid a < X \leq b]$ (left panel) vary as functions of λ, when $(\mu, \sigma) = (0, 1)$ with 10% or 20% of left truncation. Boxplots have been obtained from 1000 replicates of the empirical moments, each one being computed on a sample of 500 truncated skew-normal random variables. Clearly, the estimates of $\mathbb{E}[\Phi(X) \mid a < X \leq b]$ are much less dispersed those of $\mathbb{E}[X^3 \mid a < X \leq b]$, specially for low to moderate values of λ. Both curves flatten out as λ reaches large values. It is thus very difficult to estimate large values of λ, say $\lambda \geq 4$. Similar results have also been obtained with different truncation schemes and/or truncation intensities, or without any truncation at all (see the next section for a precise definition of truncation intensity). These are good indications that the weighted moment should be preferred to the third moment, see also Tables 1 and 2 below.

4 Data analysis

4.1 Simulations

To assess our inference method we simulated 1000 vectors of size 500 of independent replicates, using the `rsn` function of the `R` package `sn` (Azzalini, 2009). We considered three values for the shape parameter $\lambda \in \{1, 2, 4\}$, corresponding to respectively low, medium and high levels of skewness. Because λ plays a symmetrical role, we restricted ourselves to positive values for λ; $\lambda = 0$ corresponds to absence of skewness. Note also that $\lambda = 4$ is located in a relatively flat area of the curve $\mathbb{E}[\Phi(X)|a < X \leq b](\lambda)$, i.e. different large values of λ give a similar weighted moment. Since

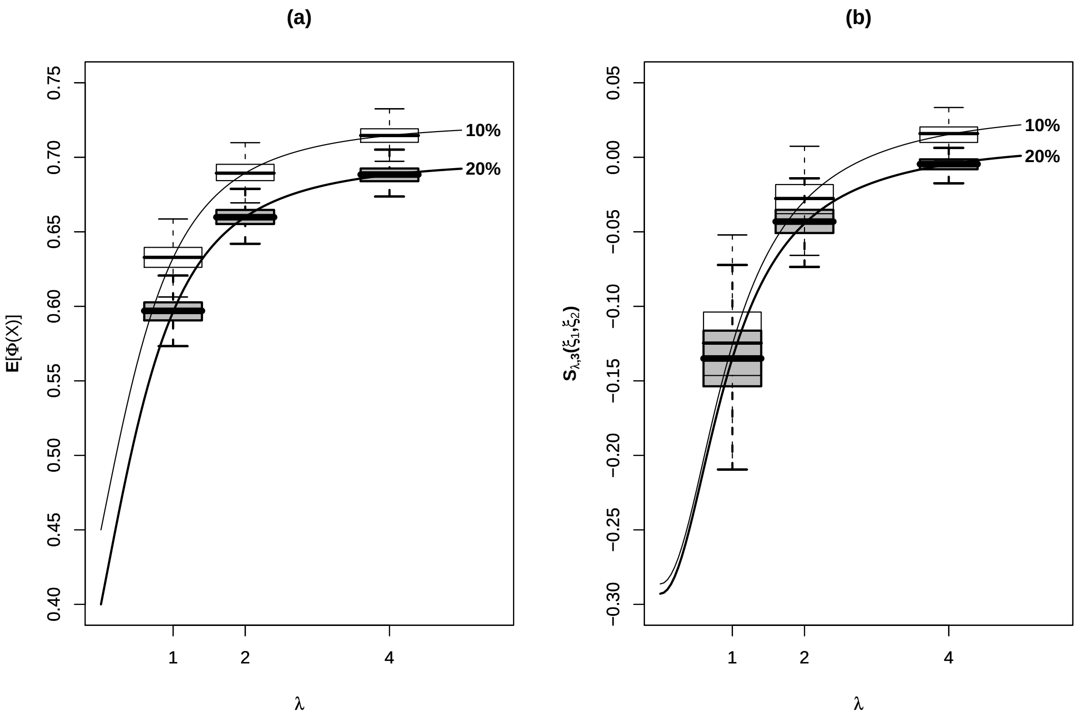

Figure 2: Theoretical and empirical moments of $\mathbb{E}[\Phi(X) \mid a < X \leq b]$ and $\mathbb{E}[X^3 \mid a < X \leq b]$ as a function of λ. Boxplots are computed on 1000 samples of size 500 of truncated $SN(0,1,\lambda)$ with $\lambda \in \{1,2,4\}$ and 10% and 20% left truncations.

Table 1: Bias and Mean-Square-Error (MSE) for the MWM, MOM and MLE (Maximum Likelihood Estimates) approaches obtained from 1000 replicates of size 500 for left or right truncation with intensity 20% and parameters $(\mu,\sigma,\lambda) = (0,1,2)$.

		MWM		MOM		MLE	
		Bias	MSE	Bias	MSE	Bias	MSE
Left truncation	$\hat{\mu}$	0.033	0.027	0.106	0.063	0.107	0.042
	$\hat{\sigma}$	-0.024	0.007	-0.033	0.010	-0.031	0.010
	$\hat{\lambda}$	-0.434	1.143	1994	$8.5\,10^7$	46.1	$5.7\,10^5$
Right truncation	$\hat{\mu}$	0.059	0.033	0.033	0.038	0.044	0.042
	$\hat{\sigma}$	0.004	0.159	0.107	0.543	0.063	0.163
	$\hat{\lambda}$	-0.003	1.806	0.372	5.546	0.213	2.214

left and right truncation do not play a symmetrical role, we considered left, right and bilateral truncation. The intensity of the truncation is defined as the probability that the non truncated

Table 2: Median and Inter-Quartile-Range (IQR) for the MWM, MOM and MLE approaches obtained from 1000 replicates of size 500 for left or right truncation with intensity 20% and parameters $(\mu, \sigma, \lambda) = (0, 1, 2)$.

		MWM		MOM		MLE	
		Median	IQR	Median	IQR	Median	IQR
Left truncation	$\hat{\mu}$	0.028	0.194	-0.077	0.232	0.089	0.139
	$\hat{\sigma}$	0.978	0.106	0.968	0.117	0.979	0.083
	$\hat{\lambda}$	1.717	1.352	1.649	2.522	2.287	2.851
Right truncation	$\hat{\mu}$	0.002	0.182	-0.012	0.198	-0.011	0.200
	$\hat{\sigma}$	1.020	0.318	1.017	0.418	1.009	0.414
	$\hat{\lambda}$	2.026	1.120	2.125	1.509	2.055	1.463

skew-normal random variable falls outside the truncation interval. We considered two intensities: 10% and 20%. In the case of a bilateral truncation, the same probability is applied to the left-hand side and right-hand side tails of the distribution.

The bias and Mean-Square-Error (MSE) for the MWM, MOM and MLE approaches are reported in Table 1 for 20% left truncation and $\lambda = 2$. Similar results were obtained for other cases, but for sake of conciseness they are not reported here. Except in one case ($\hat{\mu}$ for right truncation) MWM yields better estimates than MOM and MLE, both in terms of bias and MSE. One can also observe that MOM and MLE have great difficulties in estimating λ in the case of a left truncation, a problem not encountered with MWM. This problem is related to the higher variability of the empirical third moment than that of the empirical weighted moment. In very few cases MOM and MLE provide some very large values of $\hat{\lambda}$ due to convergence problems. These rare but very large values blow up the bias and MSE of the estimators of λ. To remove this effect in the comparison, we added Table 2 which displays the median and inter-quantile range (IQR) values for each estimator. As expected, these statistics tell a better story for the MLE and MOM. Still MWM globally outperforms the other two methods. As a consequence of the results shown in Tables 1 and 2 and Figure 2, we will now only report simulation results for MWM.

Table 3 indicates the mean and the standard deviation (in brackets) of MWM estimates. Several comments can be made

- When the truncation intensity is 10%, estimates have generally smaller bias and standard deviation than when the truncation intensity is 10% is 20%.
- The most difficult parameter to estimate is λ. This fact has been consistently reported for

skew-normal distributions (Arellano–Valle and Azzalini, 2008); even with a truncation intensity of 10%, it is extremely difficult to estimate λ if the distribution is left truncated. MOM yields even worse result and cannot be considered as a viable alternative in this case.

- The scale parameter σ is in general very well estimated; with a tendency to i) underestimation for bilateral truncation and high values of λ, and ii) overestimation with a right truncation and low level of skewness.

- The quality of the estimation of the location parameter μ is closely related to the one of λ.

Table 3: Mean [standard deviation] of estimates for MWM; 1000 replicates of size 500 for left, right and bilateral truncation with varying amount of intensity; $(\mu, \sigma) = (0, 1)$.

		$\lambda = 1$	$\lambda = 2$	$\lambda = 4$
	Left 10%	$0.084_{[0.152]}$	$0.125_{[0.184]}$	$-0.074_{[0.167]}$
	Left 20%	$0.107_{[0.212]}$	$0.033_{[0.161]}$	$-0.201_{[0.250]}$
$\hat{\mu}$	Right 10%	$0.040_{[0.259]}$	$0.028_{[0.128]}$	$0.009_{[0.053]}$
	Right 20%	$0.037_{[0.324]}$	$0.059_{[0.173]}$	$0.018_{[0.062]}$
	Bilat. 10%	$0.031_{[0.209]}$	$0.098_{[0.144]}$	$0.118_{[0.092]}$
	Bilat. 20%	$0.064_{[0.153]}$	$0.172_{[0.077]}$	$0.034_{[0.166]}$
	Left 10%	$0.969_{[0.071]}$	$0.938_{[0.089]}$	$0.991_{[0.081]}$
	Left 20%	$0.979_{[0.092]}$	$0.976_{[0.077]}$	$1.032_{[0.094]}$
$\hat{\sigma}$	Right 10%	$1.041_{[0.249]}$	$0.987_{[0.162]}$	$0.990_{[0.107]}$
	Right 20%	$1.243_{[0.771]}$	$1.004_{[0.398]}$	$0.985_{[0.194]}$
	Bilat. 10%	$1.018_{[0.169]}$	$0.929_{[0.129]}$	$0.846_{[0.070]}$
	Bilat. 20%	$0.980_{[0.142]}$	$0.831_{[0.083]}$	$0.866_{[0.121]}$
	Left 10%	$0.819_{[0.421]}$	$1.255_{[0.863]}$	$1.301_{[0.729]}$
	Left 20%	$1.102_{[0.907]}$	$1.566_{[0.977]}$	$1.043_{[0.631]}$
$\hat{\lambda}$	Right 10%	$1.075_{[0.771]}$	$1.993_{[0.695]}$	$3.977_{[0.991]}$
	Right 20%	$1.455_{[1.788]}$	$1.997_{[1.344]}$	$3.931_{[1.436]}$
	Bilat. 10%	$1.105_{[0.738]}$	$1.646_{[0.966]}$	$0.911_{[0.367]}$
	Bilat. 20%	$0.920_{[0.578]}$	$0.891_{[0.205]}$	$0.834_{[0.055]}$

The most interesting feature of this simulation study is that truncation direction and intensity play an important role in the performance of the estimation procedure. Large truncation levels, especially for bilateral and left truncation when the shape parameter is large, diminish the overall estimation quality.

In some cases, identifiability issues arise. Consider for example a one-sided truncation at $a = 0$ (i.e. the density is equal to zero for negative values). Because a skew-normal distribution with infinite shape parameter λ is nothing but a standard normal distribution truncated at zero, the two sets of parameters $(\mu = 0, \sigma = 1, \lambda = 0, a = \infty)$ and $(\mu = 0, \sigma = 1, \lambda = \infty, a = 0)$ lead to the same density. In this case these two sets of parameter are undistinguishable. This is a special degenerate case, but obviously, some situations can lead to different sets of parameters with very similar distributions. Consider for example the estimation of a $SN(0, 1, 2)$ with 20% left truncation. The estimates are plotted in Figure 3. Values of $\hat{\lambda}$ ranges from 0.1 to 5.5, with a mean equal to 1.5. Figure 3a shows that the distribution of $\hat{\lambda}$ is bimodal, corresponding to two populations of estimates $(\hat{\mu}, \hat{\sigma}, \hat{\lambda})$ of comparable size (see Table 4).

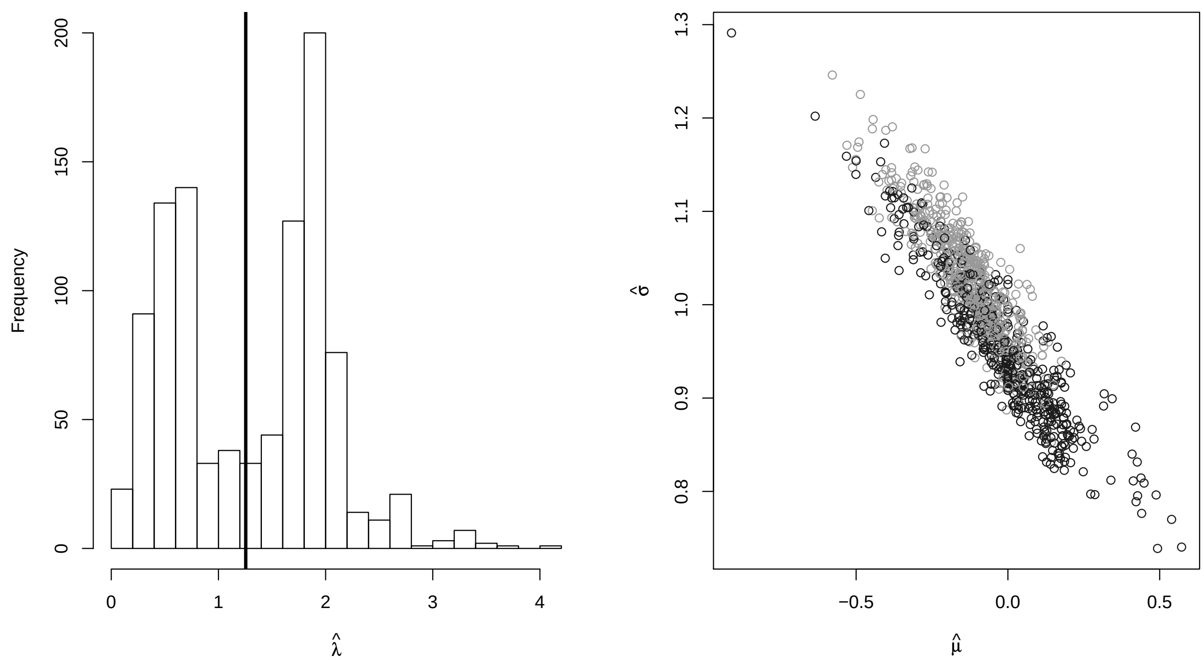

Figure 3: Estimates of (μ, σ, λ) computed from 1000 samples of size 500 of truncated SN (0, 1, 4) with a 10% left truncation.

Figure 4 displays the true density and two other densities corresponding to the two populations. This figure clearly indicates that the three densities are very close. They differ mainly near the truncation threshold.

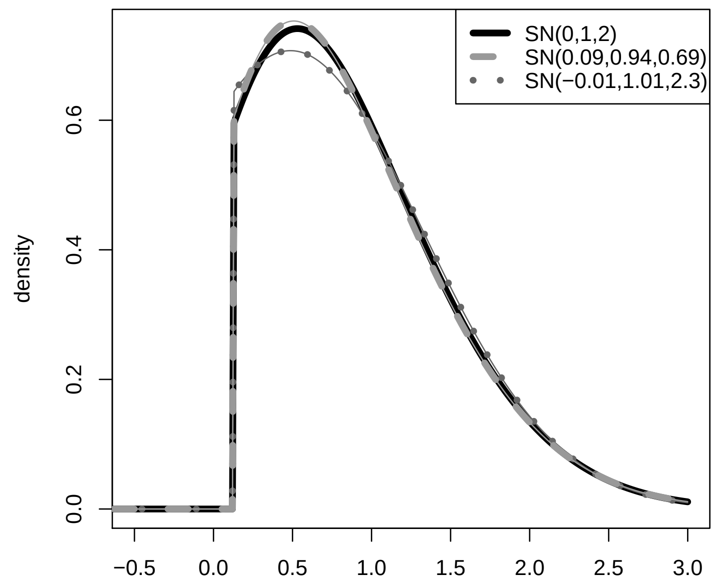

Figure 4: Black solid line: 20% left truncated SN(0,1,2); grey pointed line: typical density of group #1, SN(0.09,0.94,0.69); grey dashed line: typical density of group #2, SN(-0.01,1.01,2.3).

This example illustrates two important facts. Firstly MOM has greater difficulties than MWM for estimating the parameters since the third moment $\mathbb{E}[X^3 \mid a < X \leq b]$ is very sensitive to large values of X, see Table 1. Secondly, if λ is positive, the parameters are difficult to estimate in presence of a left truncation, but there are no particular problems for right truncations, see Table 3. It is the opposite when λ is negative.

4.2 Daily relative humidity measurements

The relative humidity of an air-water mixture is defined as the ratio of the partial pressure of the water vapor in the mixture to the saturated vapor pressure of water at a prescribed temperature (Perry, 2007). This quantity is normally expressed as a percentage and is thus restricted to the interval $[0, 100]$. The data set considered here consists of daily measurements of relative humidity recorded by the INRA weather station located in Toulouse, South of France, between 1972 and 1999. In temperate regions, relative humidity is always much larger than 0; truncation at 100 is the only truncation visible on the experimental histograms. We will therefore consider right

Table 4: Mean [standard deviation] of the estimates of a $SN(0,1,2)$ with 20% left truncation, when clustered in two populations.

	Group # 1	Group # 2
N	463	537
$\hat{\mu}$	$0.085_{[0.187]}$	$-0.011_{[0.116]}$
$\hat{\sigma}$	$0.936_{[0.079]}$	$1.009_{[0.056]}$
$\hat{\lambda}$	$0.694_{[0.367]}$	$2.31_{[0.663]}$
Intensity	$33.2\%_{[7.6]}$	$19.4\%_{[7.3]}$

truncation only. To take into account the seasonality effect, the year is divided into four periods, (Semenov *et al.*, 1998; Flecher *et al.*, 2009b): December-January-February (DJF), March-April-May (MAM), June-July-August (JJA) and September-October-November (SON). Our goal is to fit a truncated skew-normal distribution for each one of the four periods assuming climatological stationarity throughout the whole period 1972-1999. The cooresponding histograms are represented in Figure 1.

For each season, the parameters were estimated using MOM and MWM. Results are presented in Table 5. In JJA, $\hat{\lambda}$ is positive, and the estimates obtained with the two methods are very close, as discussed in the previous section. For the other seasons, $\hat{\lambda}$ is negative, i.e. we encounter the difficult case described in the previous section. Hence, MOM and MWM estimates are different. The densities with the parameters estimated via MWM are depicted Figure 1. The agreement between histograms and estimated densities is clear.

Table 5: Estimated parameters by MWM, MOM and MLE for relative humidity in Toulouse for four seasons.

	MWM			MOM			MLE		
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\lambda}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\lambda}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\lambda}$
DJF	94.1	9.4	−0.98	92.0	9.0	0.00	96.1	11.5	−1.02
MAM	86.2	15.6	−1.98	86.0	16.0	−2.05	84.4	19.4	−1.32
JJA	58.8	14.6	1.29	59.0	14.0	1.17	68.1	13.3	0.40
SON	92.2	16.1	−2.12	88.0	14.0	0.01	89.6	14.7	0.02

5 Concluding remarks and discussion

In this paper we derived the m^{th} order moments of a truncated skew-normal distribution for any $m \geq 0$ as a linear combination of moments of truncated Gaussian distributions. We linked our results with classical results on non-truncated skew-normal distributions moments. We also derived expressions for some weighted moments by taking advantage of the formulas proposed by Flecher *et al.* (2009a).
We then presented two inference methods based on classical moments (MOM) and on weighted moments (MWM). Both methods were compared to the classical MLE approach for different values of the shape parameter λ and truncation intensities. Clearly the method based on weighted moments yields better estimates than the method of moments, especially for the scale parameter. We have not tried to implement a Bayesian approach (e.g., Liseo and Loperfido, 2006, 2003) that may offer an interesting alternative. Concerning the derivation of variance estimates, it would be of interest to explore a parametric bootstrap type (Davison and Hinkley, 1997).

We also noted that when the truncation intensity or the shape parameter are too large the parameters are nearly non identifiable.

Despite these limitations, the illustration on relative humidity indicates that reasonable estimates can be derived and the inferred density models fit the histograms adequately.

Acknowledgments. This work was supported by the ANR-CLIMATOR project, the NCAR Weather and Climate Impact Assessment Science Initiative and the ANR-MOPERA, GIS-PEPER and FP7-AQCWA projects. The authors would also like to credit the contributors of the R project.

References

Arfken, G., 1985. *Mathematical Methods for Physicists, 3rd ed.* Orlando, FL: Academic Press.

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

Azzalini, A., 1985. A class of distributions which includes the normal ones. *Scand. J. Statist*, **12**, 171–178.

Azzalini, A., 2009. R package `sn`: The skew-normal and skew-t distributions (version 0.4-12). URL http://azzalini.stat.unipd.it/SN.

Davison, A.C. and Hinkley, D. 1997. *Bootstrap methods and their applications.* Cambridge Series.

Diebolt, J., Guillou, A., Naveau, P. and Ribereau, P. Improving probability-weighted moment methods for the generalized extreme value distribution. *REVSTAT - Statistical Journal*, **6–1**, 33–50.

Dhrymes, P.J., 2005. Moments of truncated Normal distribution. Unpublished note. Avalaible at <http://www.columbia.edu/∼pjd1/papers.html> (28.05.2009).

Flecher, C. Naveau, P. and Allard, D., 2009a. Estimating the Closed Skew-Normal parameters using weighted moments. *Statistics & Probability Letters*, vol. 79, issue 19, pages 1977-1984.

Flecher, C. Naveau, P., Allard, D. and Brisson, N., 2009b. A Stochastic Daily Weather Generator for Skewed Data. *Water Ressources Research*, in press, doi:10.1029/2009WR008098.

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

Hosking, J.R.M., Wallis, J.R. and Wood,E.F., 1985. Estimation of the generalized extreme-value distribution by the method of probability-weighted moments, *Technometrics*, **27**, 251-261.

Liseo, B. and Loperfido, N. 2003. A bayesian interpretation of the multivariate skew-normal distribution. *Statistics and Probability Letters*, 61(4):395–401.

Liseo, B. and Loperfido, N. 2006. A note on reference priors for the scalar skew-normal distribution. *Journal of Statistical Planning and Inference*, 136(2):373–389.

Martinez, E.H., Varela, H., Gomez, H.W and Bolfarine H., 2008. A note on the likelihood and moments of the skew-normal distribution. *SORT*, **32**(1), 57–66.

Mills, F.C. (Ed.), 1955. *Satistical methods.* Hold.

Perry, R.H. and Green, D.W. (Ed.), 2007 *Perry's Chemical Engineers' Handbook, 7th Revised edition.* McGraw-Hill Publishing Co.

Semenov, A.M., Brooks, R.J., Barrow, E.M. and Richardson, C.W., 1998. Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates. *Climate research*, **10**, 95–107.

6 Appendix

6.1 Proof of lemma 1

Dhrymes (2005) obtained the recursive representation (2). The proof is done by induction. It is only presented for odd moments; it is very similar for even moments. First note that the lemma is true for $k = 0$. Suppose now that (3) is true for the $(2k+1)^{th}$ moment. Then, using the recursive representation for the $(2k+3)^{th}$ moment:

$$\begin{aligned} m_{2k+3}(a,b) &= (2k+2)\left(-\sum_{i=0}^{k}\frac{(2k)!!}{(2i)!!}\frac{[z^{2i}\phi(z)]_a^b}{[\Phi(z)]_a^b}\right) - \frac{(2k+2)!!}{(2k+2)!!}\frac{[z^{2k+2}\phi(z)]_a^b}{[\Phi(y)]_a^b}, \\ &= -\sum_{i=0}^{k+1}\frac{(2k+2)!!}{(2i)!!}\frac{[z^{2i}\phi(z)]_a^b}{[\Phi(z)]_a^b}. \end{aligned}$$

Hence, (3) is true for any odd moment.

6.2 Proof of proposition 1

We first prove the first part of the proposition, i.e. equation (5). For $n > 1$, denoting $\lambda^* = \sqrt{1+\lambda^2}$, and integrating by part the quantity $s_{\lambda,r-2}$,

$$\begin{aligned} s_{\lambda,r-2}(u,v) &= \frac{2}{[F_\lambda(\xi)]_u^v}\int_u^v \xi^{r-2}\phi(\xi)\Phi(\lambda\xi)d\xi, \\ &= \frac{1}{r-1}\left(s_{\lambda,r}(u,v) + \frac{[\xi^{r-1}f_\lambda(\xi)]_u^v}{[F_\lambda(\xi)]_u^v} - \frac{2\lambda}{[F_\lambda(\xi)]_u^v}\int_u^v \xi^{r-1}\phi(\xi)\phi(\lambda\xi)d\xi\right), \\ &= \frac{1}{r-1}\left(s_{\lambda,r}(u,v) + \frac{[\xi^{r-1}f_\lambda(\xi)]_u^v}{[F_\lambda(\xi)]_u^v} - \frac{2\lambda}{\sqrt{2\pi}}\frac{1}{[F_\lambda(\xi)]_u^v}\int_u^v \xi^{r-1}\phi(\lambda_*\xi)d\xi\right), \\ &= \frac{1}{r-1}\left(s_{\lambda,r}(u,v) + \frac{[\xi^{r-1}f_\lambda(\xi)]_u^v}{[F_\lambda(\xi)]_u^v} - \frac{2}{\sqrt{2\pi}}\frac{\lambda}{\lambda_*^r}\frac{[\Phi(\lambda_*\xi)]_u^v}{[F_\lambda(\xi)]_u^v}m_{r-1}(\lambda_*u,\lambda_*v)\right), \end{aligned}$$

from which equation (5) follows directly by denoting

$$r_{\lambda,r}(u,v) = -\frac{[\xi^{r-1}f_\lambda(\xi)]_u^v}{[F_\lambda(\xi)]_u^v} + \frac{2}{\sqrt{2\pi}}\frac{\lambda}{\lambda_*^r}\frac{[\Phi(\lambda_*\xi)]_u^v}{[F_\lambda(\xi)]_u^v}m_{r-1}(\lambda_*u,\lambda_*v).$$

The second part of the proposition is proved by induction, similarly to the proof of proposition 1.

6.3 Proof of Proposition 2

Let $X \sim SN(\mu, \sigma, \lambda)$ with pdf $f_{\mu,\sigma,\lambda}$ and cdf $F_{\mu,\sigma,\lambda}$. Then,

$$\mathbb{E}[X^m \mid a < X \leq b] = \int_a^b x^m f_{\mu,\sigma,\lambda}(x)dx = \frac{2}{\sigma[F_{\mu,\sigma,\lambda}(x)]_a^b} \int_a^b x^m \phi(\frac{x-\mu}{\sigma})\Phi(\lambda\frac{x-\mu}{\sigma})dx.$$

Let us consider the change of variable $\xi = (x-\mu)/\sigma$. Then, with $\xi_i = (x_i - \mu)/\sigma$, $i = 1, 2$ we have

$$\begin{aligned}
\mathbb{E}[X^m \mid a < X \leq b] &= \frac{2}{[F_\lambda(\xi)]_u^v} \int_u^v (\sigma\xi + \mu)^m \phi(\xi)\Phi(\lambda\xi)d\xi, \\
&= \frac{1}{[F_\lambda(\xi)]_u^v} \int_u^v \left(\sum_{r=0}^{m} C_m^r \sigma^r \xi^r \mu^{m-r} \right) f_\lambda(\xi)d\xi, \\
&= \frac{1}{[F_\lambda(\xi)]_u^v} \sum_{r=0}^{m} C_m^r \sigma^r \mu^{m-r} \int_u^v \xi^r f_\lambda(\xi)d\xi, \\
&= \sum_{r=0}^{m} C_m^r \sigma^r \mu^{m-r} s_{\lambda,r}(u,v).
\end{aligned}$$

6.4 Proof of Proposition 3

This proposition is a direct application of a more general result regarding the weigthed moments of Closed Skew-Normal variables, established in Flecher *et al.* (2009a). Proposition 2 in Flecher *et al.* (2009a) is then applied for the function $h(x) = 1/[F_{\mu,\sigma,\lambda}(x)]_a^b$, for $a < x \leq b$ and $h(x) = 0$ elsewhere.