AIMS Mathematics

https://www.aimspress.com/journal/Math

AIMS Mathematics, 10(6): 12980–13005.
DOI: 10.3934/math.2025584
Received: 16 October 2024
Revised: 01 April 2025
Accepted: 03 April 2025
Published: 06 June 2025

Research article

Modeling of air pollution and Lossalae data by using new classes of skew bivariate family distribution and extreme distributions

Osama Mohareb[1], Oluwafemi Samson Balogun[2], Zeinab Youssef[3], Mohamed Yusuf[4], Mahmoud E. Bakr[5], Mohammad Abiad[6] and Yusra A. Tashkandy[5,*]

[1] Department of Mathematics and Computer Science, Faculty of Science, Port Said University, Port Said, Egypt

[2] Department of Computing, University of Eastern Finland, FI-70211, Finland

[3] Department of Mathematics, Faculty of Sciences, Damietta University, Damietta, Egypt

[4] Department of Mathematics, Faculty of Sciences, Helwan University, Egypt, mohammed

[5] Department of Statistics and Operations Research, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia

[6] College of Business Administration, American University of the Middle East, Kuwait

* **Correspondence:** Email: ytashkandi@ksu.edu.sa.

Abstract: Adding two more parameters for shaping provides a flexible way to make any basic bivariate distribution function (df) more adaptable. The extended bivariate dfs can better handle different kinds of bivariate data by including these parameters. They can handle a wide range of skewness and kurtosis indices well. The innovation lies in introducing a novel multiplicative bivariate stable-symmetric normal (MBSSN) distribution with two parameters, extending the conventional bivariate standard normal distribution. We conducted a detailed examination of the statistical properties of the MBSSN family and compared it to other significant competitors, such as generalized families of bivariate dfs, using real-world data like air pollution. The findings underscore the advantages and effectiveness of the MBSSN family in capturing the nuances of diverse datasets. We also applied the same methodology to develop a new two-parameter extension for two variations of bivariate extreme value distributions, which we then useed to analyze Lossalae datasets. This extension showcases the versatility and practicality of the proposed approach across different scenarios and distribution types.

Keywords: bivariate skew-normal distribution; mixtures of distributions; bivariate extremes; atmospheric contamination
Mathematics Subject Classification: 62E10, 62F10, 62H12, 62P12

List of abbreviations

df= distribution function
MBSSN= multiplicative bivariate stable-symmetric normal
SN= skew normal (SN)
BSN1= bivariate skew normal
BSN2= new version for bivariate Skew normal
BSSN= bivariate stable symmetric normal
BN= bivariate normal
FBSSN= Full bivariate stable symmetric normal
BMN= Bivariate Mixture Normal
BT-g-h= Bivariate g-h
BN= Bivariate Normal
LAQN= London Air Quality Network
AIC= The Akaike information criterion
BIC= Bayesian information criterion
BBLog= bivariate bi-logistic,
BAlog= bivariate asymmetric logistic
ALAE= allocated loss adjustment expense
loss= indemnity paymen.

1. Introduction

Studying multiple phenomena is important as it reveals how various factors interact and influence each other within the same environment. Hence, we introduce two types of bivariate distributions. The first type relies on diverse distributions to analyze data with significant skewness and kurtosis coefficient. The second type is founded on extreme data analysis. The data used in the second type was collected as block maxima (BM), representing the highest values observed over blocks of time, whether daily, monthly, or annually. The study of the second section is based on the concept of mixture distributions. The majority of these generation techniques seek to obtain an extended distribution (i.e., a family of distributions) by reshaping the baseline distribution. Although it keeps its baseline, its statistical properties are improved. It might have a broader range of kurtosis or skewness coefficient values, or its failure rate function (or risk function) has a specific mathematical form different from its baseline distribution. Azzalini [1] presented the following bivariate extension of the density of a random vector $\underline{Y} \in \mathbb{R}^2$, which follows a bivariate skew-normal (SN), denoted BSN_1, if the joint density of $\underline{Y}$ has the form $f_Y^{\langle 1\rangle}(\boldsymbol{Y};\boldsymbol{\lambda}) = c\phi(\boldsymbol{y};\Omega)\Phi(\lambda_1 y_1)\Phi(\lambda_2 y_2)$, where $\lambda_1, \lambda_2 \in \mathbb{R}$, c is the normalizing constant, $\Phi(.)$ is the cumulative distribution function of the standard normal distribution, and $\phi(\boldsymbol{y};\Omega)$ is the density of a bivariate normal distribution with correlation matrix $\Omega = (\rho_{ij}), i, j = 1, 2, \rho_{ij} = \rho$, if $i = j$, while $\rho_{ij} = \frac{\lambda_i}{\sqrt{1+\lambda_i}}\frac{\lambda_j}{\sqrt{1+\lambda_j}}\rho$, if $i \neq j$. In 1996, Azzalini and Dalla Valle [2] identified the drawback of the model BSN_1 and suggested a new version for bivariate SN distribution, denoted by BSN_2, with the following form:

$$f_V^{\langle 2\rangle}(\boldsymbol{v}; bm\lambda) = 2\phi(\boldsymbol{v};\Omega)\Phi(\lambda_1 v_1 + \lambda_2 v_2). \tag{1.1}$$

Azzalini [1] and Azzalini and Dalla Valle [2] showed that model (1.1) has some intriguing statistical characteristics.

Barakat [3] proposed a novel family of dfs that combines the baseline distribution and is more straightforward than the most well-known families of dfs. This family has a unique property; that, if it included any distribution, say F_X, it should have had its opposite F_{-X}, which made the author call it a stable symmetric family. Moreover, Barakat [3] used the baseline distribution of this family to be the conventional normal distribution and suggested that the resulting family is the stable symmetric-normal family. It can convey more statistical data than most other recognized families, as it encompasses eight distinct types of distributions, determined by kurtosis and skewness values of the nine possible variants. The stable symmetric-normal family also has many skewness and kurtosis indices. The family, which includes the nine potential forms of dfs, was referred to by Barakat and Khaled [4] as a whole family. To create a family of bivariate dfs that is more adaptable, Barakat et al. [5–7] suggested adding two shape parameters to any baseline bivariate df with comparison for this model using simulation and accurate data. Additionally, Barakat, Khaled [4], and Barakat et al. [5] applied the suggested models to air pollution. We can completely control the type of the resulting family thanks to the additional parameters. This technique gives the bivariate standard normal distribution a new two-parameter extension called the bivariate stable symmetric normal (BSSN) family. For any fitting data purpose, it is more convenient to consider the location-scale BSSN df.

$$\begin{aligned} F_{\boldsymbol{W}}(\boldsymbol{y};\alpha,l,\rho,\mu_1,\mu_2,\sigma_1,\sigma_2) &= \alpha\Phi_\rho(\frac{y_1-\mu_1}{\sigma_1}+l,\frac{y_2-\mu_2}{\sigma_2}+l)+\bar{\alpha}\Phi_\rho(\frac{y_1-\mu_1}{\sigma_1}-l,\frac{y_2-\mu_2}{\sigma_2}-l) \\ &= \alpha\Phi_\rho(\frac{y_1-l_1}{\sigma_1},\frac{y_2-l_2}{\sigma_2})+\bar{\alpha}\Phi_\rho(\frac{y_1-l_1^\star}{\sigma_1},\frac{y_2-l_2^\star}{\sigma_2}), \end{aligned} \tag{1.2}$$

where Φ_ρ is bivariate normal (BN) df, $l_1=\mu_1-l\sigma_1$, $l_1^\star=\mu_1+l\sigma_1$, $l_2=\mu_2-l\sigma_2$ and $l_2^\star=\mu_2+l\sigma_2$, and ρ is the correlation parameter. Barakat et al. [6, 8] extended and applied a new bivariate skew mixture normal distribution called location-scale full bivariate stable symmetric normal (FBSSN) df given by $F_{\boldsymbol{\eta}}(\boldsymbol{y};\alpha,l,\rho,\mu_1,\mu_2,\sigma_1,\sigma_2)$

$$\begin{aligned} &= \alpha\beta\Phi_\rho(\frac{y_1-\mu_1}{\sigma_1}+l,\frac{y_2-\mu_2}{\sigma_2}+l)+\bar{\alpha}\beta\Phi_\rho(\frac{y_1-\mu_1}{\sigma_1}-l,\frac{y_2-\mu_2}{\sigma_2}-l)+\bar{\beta}L_U(\frac{y_1-\mu_1}{\sigma_1},\frac{y_2-\mu_2}{\sigma_2}) \\ &= \alpha_1\Phi_\rho(\frac{y_1-l_1}{\sigma_1},\frac{y_2-l_2}{\sigma_2})+\alpha_2\Phi_\rho(\frac{y_1-l_1^\star}{\sigma_1},\frac{y_2-l_2^\star}{\sigma_2})+\alpha_3L_U(\frac{y_1-\mu_1}{\sigma_1},\frac{y_2-\mu_2}{\sigma_2}), \end{aligned} \tag{1.3}$$

where $l_2^\star=\mu_2+l\sigma_2$, $\alpha_1=\alpha\beta$, $\alpha_2=\bar{\alpha}\beta$ and $\alpha_3=\bar{\beta}$, where L_{bmU} is bivariate logistic df. Barakat et al. [9] compared several stable generalized families based on the mixture of normal dfs and skew-normal dfs, and applied these models to real data.

We present a novel bivariate distribution called "multiplicative bivariate stable-symmetric" (MBSSN). This distribution is derived from a standard mixture distribution by incorporating multiplication and division operations on the scale parameters. The significance of the MBSSN family lies in its enhanced flexibility and ability to fit various bivariate datasets effectively. Moreover, it extends the baseline two-form components of any bivariate distribution function, enabling it to encompass a wide range of skewness and kurtosis data. Our second objective focuses on extremes and rare data, as these models play a vital role in emergency and disaster management. Probability theory and statistics, known as the "extreme value theory," model the maximum or minimum values in a dataset. The last ten years have seen an increase in its use as a tool for risk management across several industries. This theory helps the government plan for storms, earthquakes, and other natural disasters. Banks can use it to forecast extreme investment losses and insurance companies to price their products. Barakat et al. [8] introduced many applications for powerful value modeling and related models in real data. Additionally, we present a newly proposed model and provides a list of potential future research directions. Aguirre-Salado et al. [10] developed an innovative spatial model utilizing generalized extreme values (GEVs) distribution and tree ensemble models to analyze the peak concentrations of particulate matter with a diameter of less than 2.5 microns (PM2.5) in the Mexico City metropolitan area from 2003 to 2021. The evolution of random networks and associated powerful value statistics are the subject of a survey by Markovich and Vaičiulis [11]. Bivariate extremes have received little research attention compared with other fields of statistical modeling. However, extensive probability theory is summarized in Galambos [12], Khaled et al. [13] and Barakat et al. [6]. Haj Ahmad [14] investigated the relationship between processor and memory reliability by Bivariate model approach. Qura et al. [15, 16] introduced a new bivariate distribution based on the copula function. Pikands submitted statistical applications [17] and Tawn [18]. All of them are affected by the conventional model for bivariate extremes, which use joint distributions for the extreme values in either margin of a bivariate sample. Bivariate radical value analysis in the environment included with the software system XTREMES. Mayer [19] compared the air pollutant characteristics. Yue [20] presented a comparison of two bivariate extreme value distributions. El-Sherpieny et al. [21] improved extreme bivariate distribution based on complete sample-to-sample in the presence of censored samples. Peng Shi [22] discussed the prediction of insurance liabilities by aggregating loss payment data from multiple insurers. The model is tested on a multivariate loss triangle dataset from the National Association of Insurance Commissioners.

The distribution of extreme values can be studied using either the Block Maxima or Peaks Over Threshold approaches. In the univariate instance, these methods lead to the generalized Pareto distributions (GPD) with scale and shape parameters and the GEV distributions with location, scale, and shape parameters. In most practical applications, data often exists in multiple dimensions, necessitating the extension of the methods above to accommodate corresponding bivariate and multivariate distributions. To create a more adaptable family of bivariate dfs, we present a new technique for modifying any baseline bivariate extreme value model by adding two shape parameters.

A mixture distribution method has the ability to capture complex data characteristics more effectively than a single distribution. This approach enables us to combine components that individually model specific features of the data, such as skewness in the central distribution and the behavior of extreme values in the tails. Using a mixture, the model gains flexibility to handle heterogeneous data patterns and have a more accurate fit. Despite the two proposed distributions focusing on different aspects, both skewed distribution and extreme distribution describe non-symmetrical data in similar ways. An asymmetrical distribution occurs when data are not evenly distributed, with a tail to the left or right. Extreme value modeling, on the other hand, emphasizes the presence of rare events—outlier values that represent blocks of maximum or minimum values. While skewness can indicate the presence of extreme values, extreme distributions specifically highlight their occurrence, which can affect the overall skewness of the data. Additionally, both models are based on mixture methods. Finally, our main contribution of this article is to illustrate how to apply the multiplicative parameter technique to a mixture of bivariate normals, enabling us to assess the superiority of this model compared to others in the same family. It also illustrates the application of an extra parameter for a mixture of extreme value models to achieve the same objective.

This paper is arranged as follows: In section 2, we introduce the definition of the multiplicative bivariate stable-symmetric family. In section 3, we present the simulation results of the bivariate stable-symmetric family. In section 4, we focus on the applications of mixture distribution to model air pollution data. In Section 5, a new bivariate extreme value probability model is discussed. Finally, in in Section 6 conclude the paper, accompanied by remarks on the bivariate normal and extreme distribution.

2. The multiplicative bivariate stable-symmetric family

Below, we define a multiplicative bivariate stable-symmetric normal family (the MBSSN family) for any continuous baseline df $F_{\boldsymbol{Y}}(\boldsymbol{y})$, with the probability density function (Pdf) $f_{\boldsymbol{Y}}(\boldsymbol{y})$ and shape parameters $0 \le \bar{\alpha} = 1 - \alpha \le 1$, $l > 0$. The MBSSN family is defined via the mixture of the df $F_{\boldsymbol{Y}}(\boldsymbol{y})$ and its reverse $F_{-Y_1,-Y_2}(y_1, y_2) = \bar{F}_{Y_1,Y_2}(-y_1, -y_2)$ by

$$G(y_1, y_2; \alpha, l) = \alpha F\left(\frac{y_1}{l}, \frac{y_2}{l}\right) + \bar{\alpha}\bar{F}(ly_1, ly_2), \tag{2.1}$$

where $G(y_1, y_2; \alpha, l)$ is df, and

$$\bar{F}(y_1, y_2) = 1 - F_{Y_1}(y_1) - F_{Y_2}(y_2) + F_{Y_1,Y_2}(y_1, y_2).$$

The survival function of the bivariate df (2.1) is given by

$$\bar{G}(y_1, y_2; \alpha, l) = 1 - G(y_1; \alpha, l) - G(y_2; \alpha, l) + G(y_1, y_2; \alpha, l) = \alpha\bar{F}\left(\frac{y_1}{l}, \frac{y_2}{l}\right) + \bar{\alpha}F(-ly_1, -ly_2).$$

Moreover, the Pdf is given by

$$g_{Y_1,Y_2}(y_1, y_2; \alpha, l) = \frac{\alpha}{l^2} f_{Y_1,Y_2}\left(\frac{y_1}{l}, \frac{y_2}{l}\right) + l^2\bar{\alpha} f_{Y_1,Y_2}(ly_1, ly_2).$$

If the baseline df $F_{\boldsymbol{Y}}(\boldsymbol{y})$ is radial symmetric, i.e., $F_{\boldsymbol{Y}}(\boldsymbol{y}) = F_{-\boldsymbol{Y}}(\boldsymbol{y}) = \bar{F}_{\boldsymbol{Y}}(-\boldsymbol{y})$ the definition (2.1) becomes:

$$G(y_1, y_2; \alpha, l) = \alpha F\left(\frac{y_1}{l}, \frac{y_2}{l}\right) + \bar{\alpha}F(-ly_1, -ly_2). \tag{2.2}$$

It is evident that studying Eq 2.2 is futile unless we establish its marginal value for practical application. Therefore, we should familiarize ourselves with distributions that have broad applications, like the stable symmetric-normal family. This approach involves multiplicative combinations that further enhance the adaptability of the distribution families to fit empirical data more accurately; therefore, we focus on normal distribution. Let us now apply (2.2) on the bivariat normal distribution $\Phi_{\rho,\boldsymbol{\mu}}(y)$ with mean $\boldsymbol{\mu} = (\mu_1, \mu_2) \neq (0, 0)$, unit variance and correlation coefficient ρ. A novel skew bivariate normal distribution, also referred to as an MBSSN, is described as follows:

$$G_{\boldsymbol{Y},\boldsymbol{\mu}}(\boldsymbol{y}; \alpha, l, \rho) = \alpha\Phi_{\rho,\boldsymbol{\mu}}\left(\frac{y_1}{l}, \frac{y_2}{l}\right) + \bar{\alpha}\Phi_{\rho,\boldsymbol{\mu}}(-ly_1, -ly_2). \tag{2.3}$$

It is possible to derive the moment-generating function by some direct algebra by using (2.3) in the form:

$$M_{Y_1,Y_2,\boldsymbol{\mu}}(t_1,t_2) = L_{\alpha,l}(t_1,t_2) + L_{\bar{\alpha}.\frac{-1}{l}}(t_1,t_2), \tag{2.4}$$

where

$$L_{\alpha,l}(t_1,t_2,\boldsymbol{\mu},\rho) = \alpha\exp(l(t_1\mu_1 + t_2\mu_2) + \frac{l^2}{2}(t_1^2 + 2t_1t_2\rho + t_2^2)). \tag{2.5}$$

Therefore, the following lemma provides the means, variances, coefficients of skewness, and coefficients of kurtosis for the MBSSN family's marginals.

Lemma 2.1. About each of the marginals of MBSSN df, we have:

(1) The mean is $\mu_{Y_i} = (\alpha l - \frac{\bar{\alpha}}{l})\mu_i;\ i = 1,2.$

(2) The variance is $\sigma^2_{Y_i} = \frac{1}{l^2}((\alpha l^4 + \bar{\alpha}) + \alpha\bar{\alpha}(1+l^2)^2\mu_i^2),\ i = 1,2.$

(3) The coefficient of skewness is

$$\gamma^{[1]}_{Y_i} = \frac{(1+l^2)^2\mu_i\alpha\bar{\alpha}[3(l^2-1) - (2\alpha-1)(1+l^2)\mu_i^2]}{[(1+l^2)^2\mu_i\alpha\bar{\alpha} + (\alpha l^4 + \bar{\alpha})]^{\frac{3}{2}}}, i = 1,2.$$

Clearly,

$$\gamma^{[1]}_{Y_i} = \begin{cases} = 0, \text{if} \begin{cases} \mu_i = 0 \text{ , or } \alpha = 1 \text{ , or } \alpha = 0 \text{ , or} \\ \alpha = \frac{3(l^2-1)}{2(1+l^2)\mu_i^2} + \frac{1}{2}, \mu_i^2 \neq 0 \text{ , or} \\ \forall\mu_i, l = 1, \alpha = \frac{1}{2} \end{cases} \\ > 0,, \text{if} \begin{cases} \alpha < \frac{3(l^2-1)}{2(1+l^2)\mu_i^2} + \frac{1}{2}, \mu_i > 0, \text{or} \\ \alpha > \frac{3(l^2-1)}{2(1+l^2)\mu_i^2} + \frac{1}{2}, \mu_i < 0 \end{cases} \\ < 0, \text{if} \begin{cases} \alpha > \frac{3(l^2-1)}{2(1+l^2)\mu_i^2} + \frac{1}{2}, \mu_i < 0 \text{ , or} \\ \alpha < \frac{3(l^2-1)}{2(1+l^2)\mu_i^2} + \frac{1}{2}, \mu_i > 0. \end{cases} \end{cases}$$

(4) The coefficient of kurtosis is

$$\gamma^{[2]}_{Y_i} = \frac{3(\alpha l^8 + \bar{\alpha}) + \alpha\bar{\alpha}\mu_i^4(1+l^2)^4 + 6\alpha\bar{\alpha}\mu_i^2(\bar{\alpha}l^4 + \alpha)(1+l^2)^2}{(\alpha\bar{\alpha}\mu_i^2(1+l^2)^2 + (\alpha l^4 + \bar{\alpha}))^2}, i = 1,2.$$

Consequently,

$$\gamma^{[2]}_{Y_i} = \begin{cases} = 3, \text{if} \begin{cases} \alpha = A_\star \text{ , or } \alpha = A^\star \text{ , or } \mu_i \neq 0 \text{ , or} \\ l = 1 \text{ , } \mu_i = 0 \text{ , or} \\ \forall\mu_i, \alpha = 1, \alpha = 0, \end{cases} \\ < 3 \text{ , if}, \alpha \in (A_\star, A^\star), \mu_i \neq 0. \\ > 3, \text{if} \begin{cases} l \neq 1, \mu_i = 0, \alpha \neq 0, 1 \text{ , or} \\ \alpha \notin (A_\star, A^\star), \mu_i \neq 0. \end{cases} \end{cases}$$

where

$$A_\star = \frac{6(l^2-1) - \sqrt{18(l^2-1)^2 + 3(l^2+1)^2\mu_i^4)}}{6(l^2+1)\mu_i^2} + \frac{1}{2},$$

and

$$A^\star = \frac{6(l^2-1) + \sqrt{18(l^2-1)^2 + 3(l^2+1)^2\mu_i^4)}}{6(l^2+1)\mu_i^2} + \frac{1}{2}.$$

Its correlation coefficient is straightforward to control.
Theorem 2.1. The correlation coefficient, indicated by $\rho_{\underline{Y}}$, of the MBSSN family is given by

$$\rho_{\underline{Y}} = \frac{(\alpha l^4 + \bar{\alpha})\rho + \alpha\bar{\alpha}\mu_1\mu_2(l^2+1)^2}{\sqrt{[(1+l^2)^2\mu_1^2\alpha\bar{\alpha} + (\alpha l^4 + \bar{\alpha})][(1+l^2)^2\mu_2^2\alpha\bar{\alpha} + (\alpha l^4 + \bar{\alpha})}}.$$

for all $-1 \leq \rho \leq 1$, $l > 0$ and $0 \leq \alpha \leq 1$. Moreover,

(1) $\rho_{\underline{Y}} = \rho$, if and only if $\alpha = 1$, or $\mu_1 = \mu_2 = 1$.
(2) $\rho_{\underline{Y}} = 1$, if and only if $\rho = 1$ and $\alpha = 1$ or $\rho = 1$ and $\mu_1 = \mu_2 = 0$.
(3) $\rho_{\underline{Y}} = -1$, if and only if $\rho = -1$ and $\alpha = 1$ or $\rho = -1$ and $\mu_1 = \mu_2 = 0$.
(4) $\rho_{\underline{Y}} = 0$, if and only if $-1 \leq \rho = \frac{-\alpha\bar{\alpha}\mu_1\mu_2(l^2+1)^2}{\alpha l^4 + \bar{\alpha}} \leq 0$.
(5) $\rho_{\underline{Y}} = \rho_Y$, if and only if $\mu_1 = \mu_2 = \mu$ and satisfy

$$-1 \leq \rho = \frac{-\alpha\bar{\alpha}\mu^2(l^2+1)^2}{\alpha l^4 + \bar{\alpha}} \leq 0.$$

Proof. See Appendix 1

Mardia [23] defined multivariate skewness and kurtosis for a population of p-variate, respectively, as

$$\beta_{1,p} = E[(\boldsymbol{Y} - \boldsymbol{\mu})\Sigma^{-1}(\boldsymbol{Y} - \boldsymbol{\mu})]^3,$$

and

$$\beta_{2,p} = E[(\boldsymbol{Y} - \boldsymbol{\mu})\Sigma^{-1}(\boldsymbol{Y} - \boldsymbol{\mu})]^2,$$

where Σ is the dispersion matrix, $\boldsymbol{Y} = (y_1, y_2)$ and $\boldsymbol{\mu} = (\mu_1, \mu_2)$. By applying these equations to the baseline, we find the skewness and the kurtosis (see Appendix 2), it is more feasible to consider the location scale for any data fitting purposes. MBSSN df

$$F_{\boldsymbol{Y}}(\boldsymbol{y}; \alpha, l, \rho, \mu_1, \mu_2, \sigma_1, \sigma_2, v) = \alpha\Phi_\rho(\frac{y_1 - l_1}{\sigma_1}, \frac{y_2 - l_2}{\sigma_2}) + \bar{\alpha}\Phi_\rho(\frac{y_1 - l_1^\star}{\sigma_1}, \frac{y_2 - l_2^\star}{\sigma_2}),$$

$$F_{\boldsymbol{Y}}(\boldsymbol{y}; \alpha, l, \rho, \mu_1, \mu_2, \sigma_1, \sigma_2, v) = \alpha\Phi_\rho(\frac{y_1 - l_1}{\sigma_1}, , \frac{y_2 - l_2^\star}{\sigma_2}),$$

where $l = \sigma_1 = \frac{1}{\sigma_2}$, $l_1 = \mu_1 + lv$, $l_1^\star = \mu_1 - \frac{v}{l}$, $l_2 = \mu_2 + \frac{v}{l}$, and $l_2^\star = \mu_2 - lv$. Moreover, we should obtain the estimates $\hat{\alpha}, \hat{\rho}, \hat{l}_1, \hat{l}_2, \hat{l}_1^\star, \hat{l}_2^\star, \hat{\sigma}_1$, and $\hat{\sigma}_2$ in any fitting problem. The estimation problem for these parameters is crucial in estimating mixtures of bivariate normal distributions. Several authors have proposed effective strategies to address this issue, including Quandt and Ramsey [24] and Lin et al. [25].

3. Simulation

The univariate g-h distribution was proposed by Tukey [26]. This distribution is defined by transforming the standard normal variable Z to

$$X = T_{g,h}(Z) = \frac{exp(gZ) - 1}{g} \exp\left(\frac{h}{2}Z^2\right), \tag{3.1}$$

where g is a real constant that regulates skewness, and h is a nonnegative real constant that governs kurtosis or elongation, pertaining to the definition of the g-h distribution when $h \in \mathbb{R}$. It can be demonstrated that $T_{0,0}(Z) = Z$ represents the normal distribution, whereas the distribution of $T_{0,h}(Z)$ has heavier tails than the normal distribution, with the heaviness increasing with h. Furthermore, $T_{g,0}(Z) = \frac{\exp(gZ)-1}{g}$ corresponds to the location-scale log-normal distribution. The sign of g determines the direction of skewness, but not its magnitude. Positive values of g induce right-tail skewness in the distribution, whereas negative values of g induce left-tail skewness. Thus, a remarkable range of distribution types are covered by this family of dfs. Utilizing (3.1) and using bivariate normal data and a predetermined correlation structure is a popular method for creating correlated g-and-h distributions (see Wilcox, [27]). Namely, a random vector $\boldsymbol{X} \in \mathbb{R}^2$ is said to have a standard bivariate g-h distribution, where $\boldsymbol{g} = (g_1, g_2)^T \in \mathbb{R}^2$ controls the skewness and $\boldsymbol{h} = (h_1, h_2)^T \in \mathbb{R}^2_+$ controls the kurtosis, if it can be represented as

$$\boldsymbol{X} = (T_{g_1,h_1}(Z_1), T_{g_2,h_2}(Z_2))^T, \tag{3.2}$$

where $(Z_1, Z_2)^T$ has a standard bivariate normal $\Phi_\rho(.,.)$ with correlation coefficient ρ and $T_{g_i,h_i}(Z_i), i = 1, 2$, is defined in (3.2). It is important to note that because the original (normal) correlation structure is not invariant to the transformation (3.1), it is not preserved. Wilcox [27] showed that $X_1 = T_{1,0}(Z_1)$ and $X_2 = T_{1,0}(Z_2)$ have the correlation coefficient of 0.378. In general, the correlation between $X_1 = T_{g_1,h_1}(Z_1)$ and $X_2 = T_{g_2,h_2}(Z_2)$ can be expressed as

$$\rho_X = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} T_{g_1,h_1}(z_1) T_{g_2,h_2}(z_2) \phi_\rho(z_1, z_2) dz_1 dz_2,$$

where $\phi_\rho(z_1, z_2)$ is the standard bivariate normal density, with correlation ρ, and, in general, we have $\rho_X \neq \rho$. A bivariate g-h distribution (BT-g-h) is used in our simulation investigation. The BT-g-h distribution is a statistical model used to model the joint distribution of two variables, where each variable has a skewed distribution. It is a generalization of the normal distribution that enables skewness and kurtosis to vary independently for each variable. The BT-g-h distribution is a family of distributions that includes the normal distribution as a particular case and is defined by two parameters: g and h. The g parameter controls the skewness of the distribution, while the h parameter controls the kurtosis (see [28, 29]).

Remark 1. The ghyp pakage has three types of BT-g-h distributions; we use two type of them. The first is the BT-g-h with "alpha.bar" parameter (BT-g-h1), from which we generate the simulation data with the parameters (alpha. bar =0.1 (shape parameter), $\mu_1 = \mu_2 = 0$ (location parameter), $g_1 = g_2 = 0$ (skewness parameter), $\sigma_1 = \sigma_2 = 1$ (dispersion parameter)) for symmetry data, and (alpha. bar =0.1 , $\mu_1 = 0, \mu_2 = 2, g_1 = 0, g_2 = 2, \sigma_1 = 1, \sigma_2 = 2$) for hight skew data. The second is BT-g-h with the h_1, h_2 parameter instead of the alpha.par parameter (BT-g-h2), which is used in estimation and fitting data in the case of simulation studies and real data.

The EM algorithm is an effective iterative method for calculating the maximum likelihood estimate when dealing with missing or hidden data. In Maximum Likelihood Estimation (MLE), the objective is to estimate the model parameter(s) that maximize the likelihood of the observed data. Each iteration of the EM algorithm has two processes: The E-step and the M-step. In the E-step, or expectation step, the missing data are inferred based on the observed data and the current estimation of the model parameters. This is accomplished by conditional expectation, elucidating the terminology selection.

During the M-step, the likelihood function is optimized based on the presumption that the missing data are available. The estimates of the missing data from the E-step are utilized instead of the real missing data. Convergence is guaranteed as the algorithm is designed to enhance the likelihood with each iteration. For instance, **Y** be a random vector that results from a parameterized family. We wish to find $\boldsymbol{\theta}$ such that $P(\boldsymbol{Y}|\boldsymbol{\theta})$ is a maximum. This is known as the MLE for θ. In order to estimate $\boldsymbol{\theta}$, it is typical to introduce the log likelihood function defined as $L(\boldsymbol{\theta}) = \log P(\boldsymbol{Y}|\boldsymbol{\theta})$. The likelihood function is considered to be a function of the parameter θ given the data $\boldsymbol{Y}$. Since $\log(y)$ is a strictly increasing function, the value of $\boldsymbol{\theta}$, which maximizes $P(\boldsymbol{Y}|\boldsymbol{\theta})$ also maximizes $L(\boldsymbol{\theta})$. The EM algorithm is an iterative procedure for maximizing $L(\boldsymbol{\theta})$. Assume that after the kth iteration the current estimate for $\boldsymbol{\theta}$ is given by $\boldsymbol{\theta}_k$. Since the objective is to maximize $L(\boldsymbol{\theta})$, we wish to compute an updated estimate $\boldsymbol{\theta}$, such that $L(\boldsymbol{\theta}) > L(\boldsymbol{\theta_k})$. The EM algorithm method is used here to derive the standard MLE of the MBSSN family's parameters. This algorithm is an iterative method for approaching a likelihood function's maximum.We generate a group of random samples of 1000 highly skewed parameters symmetric with parameters from the BT-g-h1 distribution. The parameters are determined using the following packages: ghyp, maxLik, and fMultivar.

As a result, we calculate the MLE for the model's parameters for each scenario, and we compare the location-scale MBSSN (see (2.1)) with the location-scale BSSN, FBSSN families, the location-scale BT-g-h2 family $F_{\boldsymbol{Y}}(\frac{y_1-\mu_1}{\sigma_1}, \frac{y_2-\mu_2}{\sigma_2}; \underline{h}, \underline{g})$ (where $\sigma > 0$ and the random vector (RV) $\boldsymbol{Y}$), the location-scale BSN_2 family $F_{\boldsymbol{V}}^{\langle 2\rangle}(\frac{y_1-\mu_1}{\sigma_1}, \frac{y_2-\mu_2}{\sigma_2}; \boldsymbol{\lambda})$,), the location-scale bivariate mixture normal df

$$\mathrm{BMN}(\boldsymbol{y}; \alpha, \rho, \boldsymbol{\mu}, \boldsymbol{\sigma}) = \alpha\Phi_\rho(\frac{y_1-\mu_1}{\sigma_1}, \frac{y_2-\mu_2}{\sigma_2}) + \bar{\alpha}\Phi_\rho(\frac{y_1-\mu_3}{\sigma_3}, \frac{y_2-\mu_4}{\sigma_4})$$

(the correlation coefficient of MBN is denoted by ρ_{BMN}) and the location-scale bivariate normal df $\mathrm{BN}(\boldsymbol{y}; \rho, \boldsymbol{\mu}, \boldsymbol{\sigma}) = \Phi_\rho(\frac{y_1-\mu_1}{\sigma_2}, \frac{y_1-\mu_1}{\sigma_2})$ (with correlation coefficient $\rho_{_{BN}} = \rho$). Using the R package, the following calculations are made to estimate the parameters of these families: We utilize the ghyp package for BT-g-h2, the Sn package for BSN_2, the stats package for BN families, the maxLik package for MBSSN, BSSN, and BMN families, and the maxLik and VGAM packages for FBSSN.

The Akaike information criterion (AIC) (Akaike, [30]) and the Bayesian information criterion (BIC) (Schwarz, [31]) are employed to compare the fits of these families. These criteria are based on the model's likelihood score, the number of observations, and the number of parameters. The estimation parameters of this study are shown in Tables 1 and 2. The comparison results are summarized in Tables 3, 4 and Figure 1. Based on the tables, it is evident that the MSSN family, closely followed by the FBSSN family, exhibits the lowest values of the AIC and BIC criteria in both cases when the data is symmetric or strongly skewed. The families BT-g-h2 and BSN_2, on the other hand, have the highest values for those criteria. The MBSSN family is the model that best fits these data sets. The BT-g-h2 and BSN_2 families are the worst two models that suit these data sets.

Remark 2. The identifiability problem is a key consideration while estimating the parameters of the mixture df's. However, Teicher [32] has shown that, generally speaking, the class of combination of normal df's is unidentifiable. This suggests directly that the family (2.3) is generally not distinguishable. In our study, however, the lack of identifiability is not disadvantage because our primary goal is to select an appropriate family (with estimated parameters) to represent the provided data, even if this family is not unique. For additional information regarding the identifiability issue and the estimation process's convergence issues presented in the mixture of normal df's, read Ho and Nguyen [33] and Ho [34].

Table 1. Simulation-estimation study for symmetric data.

MLE's of parameters for different families based on symmetric data			
	h_1 h_2 g_1 g_2	μ_1 μ_2	σ_1 σ_2 $\rho_{\underline{x}}$
BT-g-h2	0.18 4.31 0.009 0.005	0.023 0.02	1.19 2.039 0.18
	λ_1 λ_2	μ_1 μ_2	σ_1 σ_2 $\rho_{\underline{v}}$
BSN_2	1.132 0.054	1.25 .028	2.13 1.392 0.723
	α_1 α_2	$l1$ $l_1^\star$ l_2 $l_2^\star$	σ_1 σ_2 $\rho_{\underline{w}}$
BSSN	0.79 0.21	15 8 16.012 10.016	19.986 11.98 .0255
	α_1 α_2	l_1 $l_1^\star$ l_2 $l_2^\star$	σ_1 σ_2 $\rho_{\underline{w}}$
MBSSN	0.98 0.02	5 5.99 14 0.0162	9.99 4 0.099
	α_1 α_2 α_3	l_1 $l_1^\star$ l_2 $l_2^\star$ μ_1 μ_2	σ_1 σ_2 $\rho_{\underline{\eta}}$
FBSSN	$5 * e^{-}10$ .073 0.93	0.48 19 9.99 14.94 20	5 12 0.002
	α_1 α_2	μ_1 μ_2 μ_3 μ_4	σ_1 σ_2 σ_3 σ_4 ρ_{BMN}
BMN	0.287 0.713	8 5 6 4	2.99 14.99 10.99 9.98 0.895
		μ_1 μ_2	σ_1 σ_2 ρ_{BN}
BN df		0.0094 0.049	1.423 1.3 0.07

Table 2. Simulation-estimation study for highly skewed data.

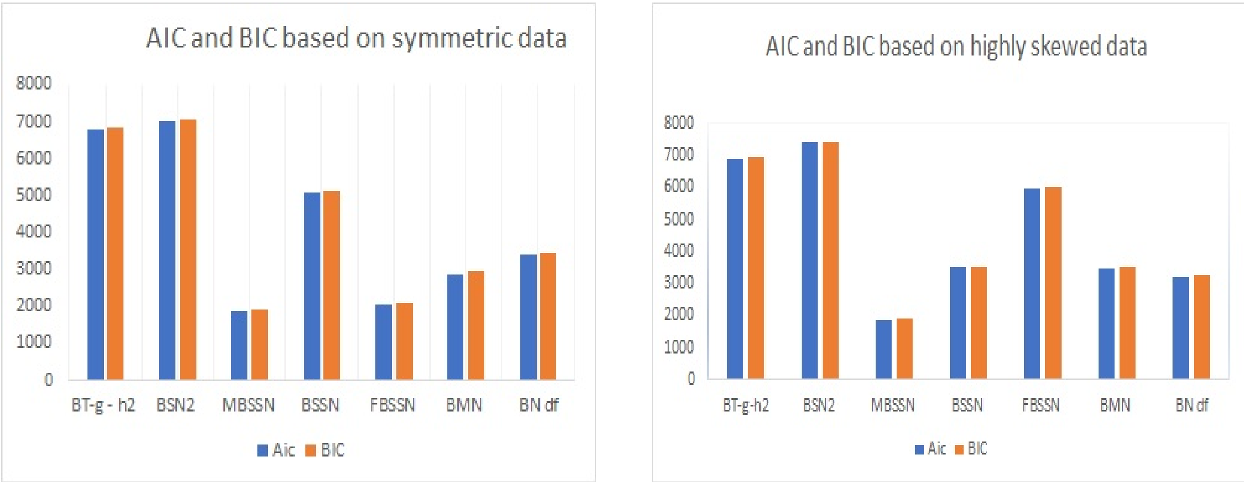
MLE's of parameters for different families based on highly skewed data															
	h_1	h_2	g_1	g_2	μ_1	μ_2					σ_1	σ_2	$\rho_{\underline{x}}$		
BT-g-h2	0.18	4.31	0.0014	0.121	.011	0.0737					1.306	1.458	0.027		
	λ_1	λ_2			μ_1	μ_2					σ_1	σ_2	$\rho_{\underline{v}}$		
BSN_2	.006	1.815			.019	1.383					1.317	2.105	.0137		
	α_1	α_2			l_1	$l_1^{\star}$	l_2	$l_2^{\star}$			σ_1	σ_2	$\rho_{\underline{w}}$		
BSSN	.8018	0.1982			2.158	9.756	6.336	1.846			2.078	9.35	1		
	α_1	α_2			l_1	$l_1^{\star}$	l_2	$l_2^{\star}$			σ_1	σ_2	$\rho_{\underline{w}}$		
MBSSN	0.9	0.1			5	5	14	6			14	0.02	0.1		
	α_1	$_2$	α_3		l_1	$l_1^{\star}$	l_2	$l_2^{\star}$	μ_1	μ_2	σ_1	σ_2	$\rho_{\underline{\eta}}$		
FBSSN	0.236	0.72	0.044		40	8	16	10	40	25.0	25	12	0.002		
	α_1	α_2			μ_1	μ_2	μ_3	μ_4			σ_1	σ_2	σ_3	σ_4	ρ_{BMN}
BMN	0.098	0.902			10.00	8.00	6.04	7.01			2.99	14.99	10.97	9.98	0.91
					μ_1	μ_2					σ_1	σ_2	ρ_{BN}		
BN df					0.0097	0.044					1.3175	1.546	.0120.		

Table 3. AIC and BIC for different distributions family based on symmetric data.

	Distributions family						
	BT-g-h2	BSN_2	MBSSN	BSSN	FBSSN	BMN	BN df
AIC	6820.014	7018.792	1893.703	5084.62	2059	2887	3433.332
BIC	6859.276	7054.084	1932.965	5123.882	2113.02	2945	3457
Rank	6	7	1	5	2	3	4

Table 4. AIC and BIC for different distributions family for highly skewed data.

	Distributions family						
	BT-g-h2	BSN_2	MBSSN	BSSN	FBSSN	BMN	BN df
AIC	6898.859	7393.558	1839.779	3502.8.	5957.16	3459.274	3229.908
BIC	6938.121	7423.005	1933.041	3542.062	6011.145	3508.352	3254.447
Rank	6	7	1	4	5	3	2



(a) AIC and BIC based on symmetric data

(b) AIC and BIC based on highly skewed data

Figure 1. AIC and BIC for different distributions family based on symmetric and highly skewed data.

4. Application of mixture distribution for air pollution

Bivariate mixture distribution can be used in air pollution modeling to estimate the joint distribution of two or more air pollutants, considering the presence of multiple sources or components of pollution. One example of how bivariate mixture distribution can be used in air pollution is in the modeling of sulfur dioxide (SO_2), nitrogen dioxide (NO_2), and nitric oxide (NO) concentrations in urban areas.

SO_2, NO_2, and NO are all harmful air pollutants that can negatively impact human health and the environment. These pollutants are often emitted from industrial activities, transportation, and other anthropogenic sources and can also be influenced by meteorological factors such as wind speed, temperature, and atmospheric stability.

By fitting a bivariate mixture distribution to historical data on SO_2, NO_2, and NO concentrations, it is possible to estimate the contributions of different pollution sources or components to each pollutant's overall concentration. For example, one component of SO_2 concentrations may be related to industrial emissions, while another may be related to natural sources such as volcanic activity or wildfires.

In summary, the bivariate mixture distribution is a valuable tool for comprehending the joint distribution of multiple air pollutants and their origins. This approach can significantly contribute to informed environmental and public health decision-making. Barakat et al. [5] used air pollution for their study. To apply the new distribution, between January 1 and December 31, 2019, the London Air Quality Network (LAQN) collected three daily data sets on air pollution. This network is a centralized source for measurements of air pollution, which are essential to support air quality management. To plan and improve air pollution monitoring in London, the LAQN was established in 1993. Through the website "www:londonair:org:uk", the residents of London can verify changes to the data on the amount of air pollution. The three contaminants, NO, NO_2, and SO_2 are used. Researchers can get these data sets from the website: "www.londonair.org.uk/london/asp/datadownload.asp". Table 5 provides the summary statistics for these data sets.

Table 5. Summary statistics.

Descriptive statistics for air pollutants								
pollution	n	minimum	maximum	median	mean	SD	skewness	kurtosis
Nitric oxide	2094	0.1	68.3	1.3	2.39	4.3	8.429	95.185
Nitrogen Dioxide	2094	0.4	57.1	12.2	14.74	9.97	1.345	1.887
Sulfur Dioxide	2094	0.7	8.5	2.8	2.53	1.022	0.463	0.637

In this study, we compare the location-scale MBSSN with the location-scale BSSN, FBSSN families, the location-scale BT-g-h2 family, the location-scale BSN_2, BMN DF, and BN df. Tables 6–8 summarize these estimates. Moreover, we compute AIC and BIC; as a consequence, Tables 9–11 and Figure 2 describe the comparative results. This demonstrates that, of the three tables, the MBSSN family has the lowest values for the AIC and BIC criteria, followed by the FBSSN family. The MBSSN family, followed by the FBSSN family, is the model that best fits these data sets (NO, NO_2), (NO, SO_2), and (NO_2, SO_2). The families BT-g-h2 and BSN_2, on the other hand, have the highest values for those criteria; thus, the BT-g-h2 and BSN_2 families are the terrible two models for these data sets.

Table 6. Parameter estimation for (NO, NO_2).

MLE's of parameters for different bivariate families of (NO, NO_2)			
	h_1 h_2 g_1 g_2	μ_1 μ_2	σ_1 σ_2 $\rho_{\underline{X}}$
BT-g-h2	0.05 4.09 1.182 3.685	0.83 10.59	2.549 9.49 0.342
	λ_1 λ_2	μ_1 μ_2	σ_1 σ_2 $\rho_{\underline{V}}$
BSN_2	24.879 -0.367	-.491 11.799	4.62 9.86 0.417
	α_1 α_2	l_1 $l_1^\star$ l_2 $l_2^\star$	σ_1 σ_2 $\rho_{\underline{W}}$
BSSN	0.813 0.187	8.37 11.76 3.36 20.00	10.58 2.099 0.30
	α_1 α_2	l_1 $l_1^\star$ l_2 $l_2^\star$	σ_1 σ_2 $\rho_{\underline{W}}$
MBSSN	0.78 0.22	10 14 14 0.015	12.00 3.00 0.24
	α_1 α_2 α_3	l_1 $l_1^\star$ l_2 $l_2^\star$ μ_1 μ_2	σ_1 σ_2 $\rho_{\underline{\eta}}$
FBSSN	0.392 .598 0.01	39.98 16 8 10 15 39.99	24.94 12 0.228
	α_1 α_2	μ_1 μ_2 μ_3 μ_4	σ_1 σ_2 σ_3 σ_4 ρ_{BMN}
BMN	0.004 0.996	9.98 8.049 6.06 6.92	2.44 14.96 11 10 0.02
		μ_1 μ_2	σ_1 σ_2 ρ_{BN}
BN df		1.9043 13.93	3.65 9.627 0.38

Table 7. Parameter estimation for (NO, SO_2).

ML parameters MLE's of parameters for different bivariate families of (NO, SO_2)															
	h_1	h_2	g_1	g_2	μ_1	μ_2					σ_1	σ_2	$\rho_{\underline{x}}$		
BT-g-h2	0.088	4.15	1.098	0.016	.8867	2.624					6.58	1.36	-.049		
	λ_1	λ_2			μ_1	μ_2					σ_1	σ_2	$\rho_{\underline{v}}$		
BSN_2	27.923	.912			-0.459	2.67					21.24	1.166	-0.0128		
	α_1	α_2			l_1	$l_1^\star$	l_2	$l_2^\star$			σ_1	σ_2	$\rho_{\underline{w}}$		
BSSN	0.84	0.16			11.006	12.00	15.00	25.00			16.992	27.998	0.32		
	α_1	α_2			l_1	$l_1^\star$	l_2	$l_2^\star$			σ_1	σ_2	$\rho_{\underline{w}}$		
MBSSN	0.70	0.30			12.00	13.00	13.00	.019			13.00	3.00	0.297		
	α_1	α_2	α_3		l_1	$l_1^\star$	l_2	$l_2^\star$	μ_1	μ_2	σ_1	σ_2	$\rho_{\underline{\eta}}$		
FBSSN	$7.35e^{-05}$	$2.211e^{-02}$	0.978		10	10.1	30.	16.02	22	13.99	12.00	5.00	$2.49e^{-}02$		
	α_1	α_2			μ_1	μ_2	μ_3	μ_4			σ_1	σ_2	σ_3	σ_4	ρ_{BMN}
BMN	0.3124	0.6878			3.25	4.45	2.00	4.00			13	15	11	10	0.86
					μ_1	μ_2					σ_1	σ_2	ρ_{BN}		
BN df					1.9	2.63					3.95	1.08	-0.88		

Table 8. Parameter estimation for (NO_2, SO_2).

MLE's of parameters for different bivariate families of (NO_2, SO_2)															
	h_1	h_2	g_1	g_2	μ_1	μ_2					σ_1	σ_2	$\rho_{\underline{x}}$		
BT-g-h2	0.25	4.39	2.8	-.003	11.095	2.64					87.49	1.24	$9e^{-}03$		
	λ_1	λ_2			μ_1	μ_2					σ_1	σ_2	$\rho_{\underline{v}}$		
BSN_2	12.128	0.0945			2.48	2.76					223.92	1.18	.0095		
	α_1	α_2			l_1	$l_1^\star$	l_2	$l_2^\star$			σ_1	σ_2	$\rho_{\underline{w}}$		
BSSN	0.26	0.74			8.48	11.6	17.6	17.00			17.09	27.8	0.9963		
	α_1	α_2			l_1	$l_1^\star$	l_2	$l_2^\star$			σ_1	σ_2	$\rho_{\underline{w}}$		
MBSSN	0.165	0.835			1.998	1.99	2.997	.014			2.993	4.00	0.165		
	α_1	α_2	α_3		l_1	$l_1^\star$	l_2	$l_2^\star$	μ_1	μ_2	σ_1	σ_2	$\rho_{\underline{\eta}}$		
FBSSN	0.1302	0.84	0.025		43	13.82	10	10.29	16.38	19.89	27.68	10.017	0.70		
	α_1	α_2			μ_1	μ_2	μ_3	μ_4			σ_1	σ_2	σ_3	σ_4	ρ_{BMN}
BMN	$3.84e^{-}03$	0.99616			7.739	5.088	6.149	3.798			1.367	1.5	1.1	1	0.16
					μ_1	μ_2					σ_1	σ_2	ρ_{BN}		
BN df					13.93	2.64					9.63	1.08	0.108		

Table 9. AIC and BIC for (NO, NO_2).

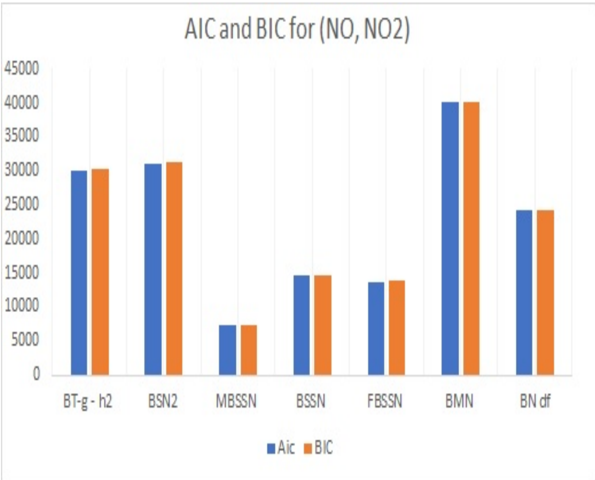
	Distributions family						
	BT-g-h2	BSN_2	MBSSN	BSSN	FBSSN	BMN	BN df
AIC	30127.92	31174.1	7359.71	14548.76	13760.1	40114.94	24176.1
BIC	30174.98	31209.39	7406.765	14607.58	13824.81	40162	24229.16
Rank	5	6	1	3	2	7	4

Table 10. AIC and BIC for (NO, SO_2).

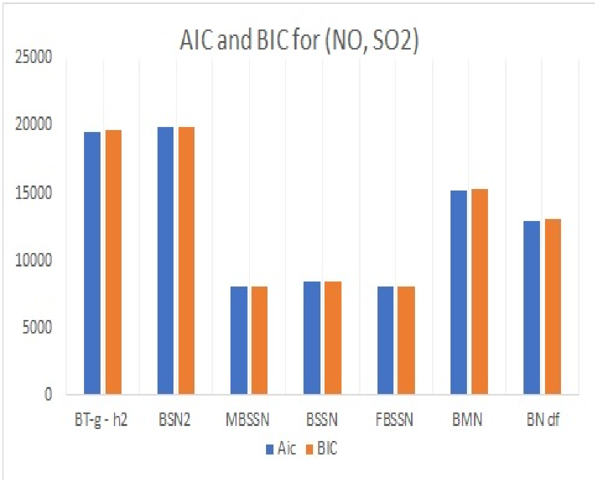
	Distributions family						
	BT-g-h2	BSN_2	MBSSN	BSSN	FBSSN	BMN	BN df
AIC	19604.01	19980.27	8026.984	8438.134	8027.522	15209.04	12979.81
BIC	19651.06	19946.26	8074.039	8485.189	8087.341	15273.74	13032.87
Rank	6	7	1	3	2	5	4

Table 11. AIC and BIC for (NO_2, SO_2).

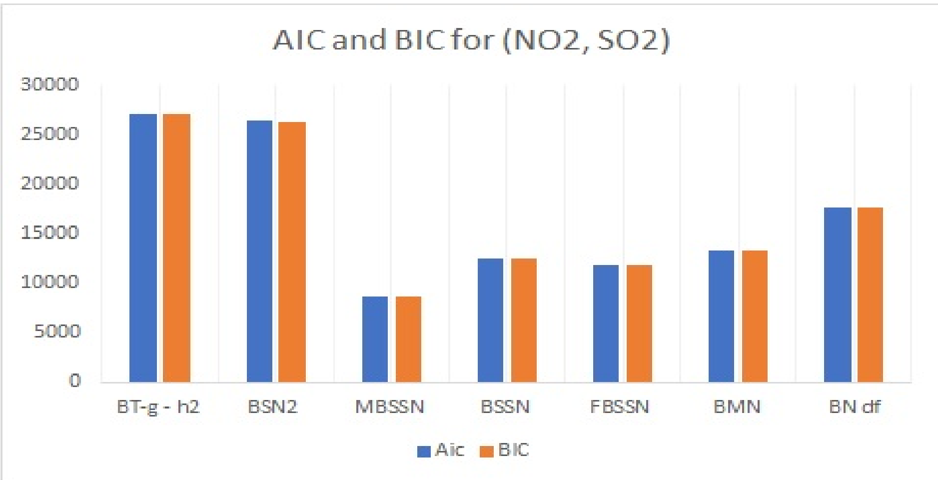
	Distributions family						
	BT-g-h2	BSN_2	MBSSN	BSSN	FBSSN	BMN	BN df
AIC	27132.00	26423.63	8643.022	12460.00	11846.62	13268.95	17680.43
BIC	27199.00	26388.34	8650.077	12525.59	11843.00	13327.77	17733.49
Rank	6	7	1	3	2	4	5



(a) AIC and BIC for (NO, NO_2)



(b) AIC and BIC for (NO, SO_2)



(c) AIC and BIC for (NO_2, SO_2)

Figure 2. AIC and BIC for different bivariate families.

5. A new bivariate extreme value probability model

The bivariate extreme value theory provides a framework for modeling the joint distribution of extreme values of two variables. A critical aspect of this theory is the consideration of the different types of asymptotic dependence that can occur between the two variables. Positive asymptotic support implies that the two variables are strongly and positively dependent on the tail of their distributions. In contrast, negative asymptotic dependence tends to be strongly negative at the bottom. Asymptotic independence, on the other hand, implies no dependency between the two variables in the tail. Some different models and approaches can be used to capture these types of asymptotic dependence, including copula-based models, max-stable processes, and conditional extreme value models. The choice of model depends on the specific application and the nature of the data being analyzed. For example, Coles and Tawn [35] used a max-stable process to model negative asymptotic dependence between two variables. In this section, we introduce a family of probability distributions by adding two shape parameters to any bivariate extreme value distribution baseline to get a more flexible family. We compare the two new two-parameter extension bivariate extreme value distributions with their bases.

5.1. Bivariate extreme value distribution

A bivariate extreme value probability model is a statistical model that describes the joint behavior of two random variables that have extreme values. Let set

$$z_i = z_i(y_i) = \{1 + \eta_i(y_i - \mu_i)/\sigma_i\}^{-1/\eta_i},$$

for $1 + \eta_i(y_i - \mu_i)/\sigma_i > 0$ and $i = 1, 2$, in which case the marginal parameters given by $(\mu_i, \sigma_i, \eta_i)$, $\sigma_i > 0$. η_i is the parameter for shape, μ_i is the parameter for location parameter, and σ_i is the parameter for scale . If $\eta_i = 0$ then z_i is defined by continuity.

The univariate margins in each of the functions for bivariate distributions $G(y_1, y_2)$ supplied below are the generalized extreme value, such that $G(y_i) = \exp(-z_i)$ for $i = 1, 2$. If $\frac{1+\eta_i(y_i-\mu_i)}{\sigma_i} \leq 0$ for some $i = 1, 2$, the value z_i is either more significant than the upper-end point (if $\eta_i < 0$) or less than the lower end point (if $\eta_i > 0$), of the i-th univariate marginal distribution.

The bivariate asymmetric logistic (BALog) df with the parameters ν, θ_1, and θ_2 was defined by Tawn [18] as

$$G_1(y_1, y_2) = \exp\{-(1 - \theta_1)z_1 - (1 - \theta_2)z_2 - [(\theta_1 z_1)^{1/\nu} + (\theta_2 z_2)^{1/\nu}]^{\nu}\},$$

where $0 < \nu \leq 1$ and $0 \leq \theta_1, \theta_2 \leq 1$. When $\theta_1 = \theta_2 = 1$, the BALog model is equivalent to the BLog model. If $\nu = 1$, $\theta_1 = 0$, or $\nu = 1$ $\theta_2 = 0$ the marginal of the BAlog becomes independent. Full dependency occurs in the limit when ν approaches 0 and $\theta_1 = \theta_2 = 1$. Different limits occur when θ_1 and θ_2 are fixed and r approaches zero.

The bivariate bi-logistic (BBLog) df with parameters θ_1 and θ_2, introduced by Smith [36], is defined as

$$G_2(y_1, y_2) = \exp\{-z_1 q^{1-\theta_1} - z_2(1 - q)^{1-\theta_2}\},$$

where $q = q(z_1, z_2; \theta_1, \theta_2)$ is the root of the equation

$$(1 - \theta_1)z_1(1 - q)_2^{\theta} - (1 - \theta_2)z_2 q_1^{\theta} = 0,$$

$0 < \theta_1, \theta_2 < 1$. The BBLog model is similar to the BLog model with the dependence parameter $\theta_1 = \theta_2$ when $\theta_1 = \theta_2$. As $\theta_1 = \theta_2$ gets closer to zero, complete dependency is realized in the limit. $\theta_1 = \theta_2$ approaches one, and when one of θ_1, θ_2 is fixed while the other approaches one, independence is obtained. When one of θ_1, θ_2 is fixed while the other is approaching zero, different limits are reached.

We suggest the following bivariate families

$$G_i(z_1, z_2; \alpha, l) = \alpha G_i(y_1 - l, y_2 - l) + \bar{\alpha} G_i(y_1 + l, y_2 + l), \qquad i = 1, 2.$$

By applying this equation to all perverse models, for fitting data purposes, we consider, in general, the location-scale bivariate extreme value distribution as

$$\begin{aligned} G_i(\boldsymbol{z}, \boldsymbol{\mu}, \boldsymbol{\sigma}; \alpha, l) =& \alpha G_i\left(\frac{y_1 - \mu_1}{\sigma_1} - l, \frac{y_2 - \mu_2}{\sigma_2} - l\right) + \bar{\alpha} G_i\left(\frac{y_1 - \mu_1}{\sigma_1} + l, \frac{y_2 - \mu_2}{\sigma_2} + l\right) \\ =& \alpha G_i\left(\frac{y_1 - (\mu_1 + l\sigma_1)}{\sigma_1}, \frac{y_2 - (\mu_2 + l\sigma_1)}{\sigma_2}\right) + \\ & \bar{\alpha} G_i\left(\frac{y_1 - (\mu_1 - l\sigma_2)}{\sigma_1}, \frac{y_2 - (\mu_2 - l\sigma_2)}{\sigma_2}\right) \\ =& \alpha G_i\left(\frac{y_1 - l_1}{\sigma_1}, \frac{y_2 - l_2}{\sigma_2}\right) + \bar{\alpha} G_i\left(\frac{y_1 - l_1^\star}{\sigma_1}, \frac{y_2 - l_2^\star}{\sigma_2}\right), \qquad i = 1, 2, \end{aligned}$$

where $l_1 = \mu_1 + l\sigma_1$, $l_2 = \mu_2 + l\sigma_2$, $l_1^\star = \mu_1 - l\sigma_1$ and $l_2^\star = \mu_2 - l\sigma_2$. Moreover, in any fitting problem we should get $\hat{\alpha}$, $\hat{l}_1$,$\hat{l}_2$,$\hat{l}_1^\star$, $\hat{l}_2^\star$, $\hat{\sigma}_1$, and $\hat{\sigma}_2$ clearly.

We suppose that the location-scale bivariate asymmetric logistic df function is

$$G_1(\boldsymbol{z}, \boldsymbol{\mu}, \boldsymbol{\sigma}; \alpha, l) = \alpha G_1\left(\frac{y_1 - l_1}{\sigma_1}, \frac{y_2 - l_2}{\sigma_2}\right) + \bar{\alpha} G_1\left(\frac{y_1 - l_1^\star}{\sigma_1}, \frac{y_2 - l_2^\star}{\sigma_2}\right),$$

where $l_1 = \mu_1 + l\sigma_1$, $l_2 = \mu_2 + l\sigma_2$, $l_1^\star = \mu_1 - l\sigma_1$, and $l_2^\star = \mu_2 - l\sigma_2$. Moreover, in any fitting problem we should get θ_1, θ_2, ν, $\hat{\alpha}$, $\hat{l}_1$, $\hat{l}_2$, $\hat{l}_1^\star$, $\hat{l}_2^\star$, $\hat{\sigma}_1$, and $\hat{\sigma}_2$ clearly.

We suppose that the location-scale bivariate Bi-Logistic df is

$$G_2(\boldsymbol{z}, \boldsymbol{\mu}, \boldsymbol{\sigma}; \alpha, l) = \alpha G_2\left(\frac{y_1 - l_1}{\sigma_1}, \frac{y_2 - l_2}{\sigma_2}\right) + \bar{\alpha} G_2\left(\frac{y_1 - l_1^\star}{\sigma_1}, \frac{y_2 - l_2^\star}{\sigma_2}\right),$$

where $l_1 = \mu_1 + l\sigma_1$, $l_2 = \mu_2 + l\sigma_2$, $l_1^\star = \mu_1 - l\sigma_1$, and $l_2^\star = \mu_2 - l\sigma_2$. Moreover, in any fitting problem we should get θ_1, θ_2, $\hat{\theta}_1$, $\hat{\theta}_2$ $\hat{\alpha}$, $\hat{l}_1$,$\hat{l}_2$, $\hat{l}_1^\star$, $\hat{l}_2^\star$, $\hat{\sigma}_1$, and $\hat{\sigma}_2$ clearly.

5.2. Modeling

The locale dataset is used, which includes observations for 1,500 liability claims. The allocated loss adjustment expense (ALAE) and the indemnity payout (loss) are reported in USD for each share. The ALAE are the additional costs incurred to resolve the claim (e.g., claims investigation expenses and legal fees). The indemnity payment row names capped at their policy maximum are listed in the dataset's additional attribute (see Beirlant et al. [37]). Maxima data have a bivariate extreme value distribution, so we generate a dataset of component-wise block maxima by randomly dividing the data into groups of size $m = 30$ and $k = 50$, where k component-wise maxima were taken, which are m observations. Location-scale asymmetric bivariate logistic families and location-scale bivariate Bi-Logistic families are all compared in this study. These bivariate extreme distributions on the locational scale are contrasted with the bivariate extreme distributions. The parameter estimation of two models is shown in Tables 12 and 13.

Table 12. Parameter estimation for bivariate extreme values distribution.

ML parameters estimation									
	ν	θ_1	θ_2	μ_1	μ_2	σ_1	σ_2	η_1	η_2
Balog	0.6823	0.999	0.5308	2.2320	0.5546	1.56	0.3128	0.3247	0.6538
	θ_1	θ_2	μ_1	μ_2	σ_1	σ_2	η_1	η_2	
BBlog	0.1006	.9681	2.332	0.5807	1.5476	0.3803	0.3837	0.5067	

Table 13. Parameter estimation for location-scale bivariate extreme values families.

MLE's												
	α	ν	θ_1	θ_2	l_1	l_2	$l_1^\star$	$l_2^\star$	σ_1	σ_2	η_1	η_2
BALog	.9995	.8812	.805	.691	1.4234	.9884	.7878	.8077	.8354	1.1288	.9619	1.0269
	α		θ_1	θ_2	l_1	l_2	$l_1^\star$	$l_2^\star$	σ_1	σ_2	η_1	η_2
BBlog	.9987		.0990	.9685	1.1309	.9496	1.0018	.4833	1.0104	1.0165	.96126	1.0260

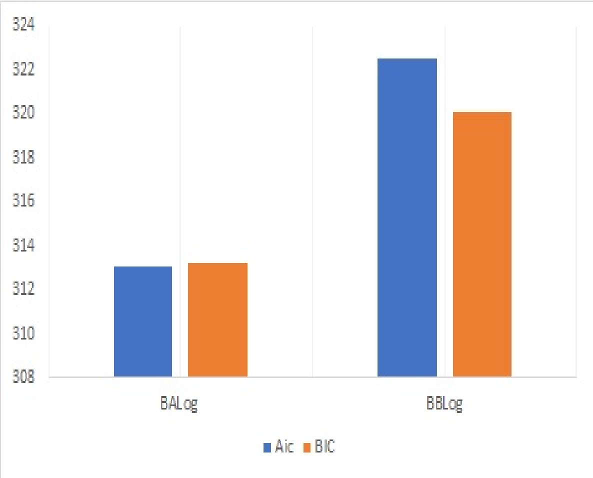
It is clear that when using this method and looking at Tables 14, 15 and Figure 3, we find improvements in the two models, the BALog and BBLog models, despite the convergence of the skewness coefficients of the indemnity payment (loss) and the allocated loss adjustment expense (ALAE), which are 1.848 and 1.791, and the kurtosis coefficients for them, which are 3.338 and 2.826, respectively.

Table 14. AIC and BIC for bivariate extreme distribution.

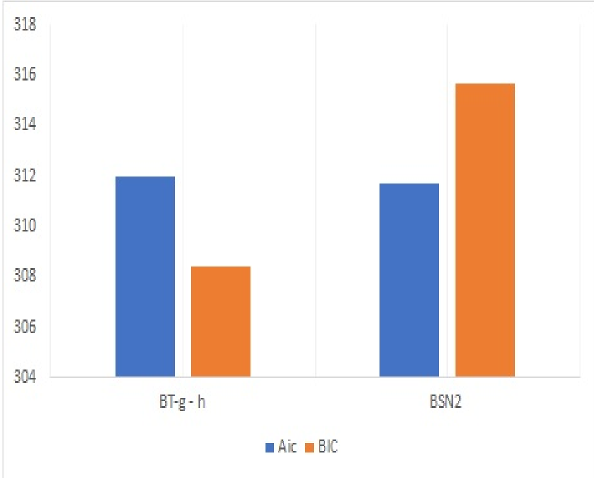
	BALog	BBLog
AIC	313.09	322.51
BIC	313.19	320.11

Table 15. AIC and BIC for location-scale bivariate extreme families.

	BALog	BBLog
AIC	312.01	311.738
BIC	308.39	315.69



(a) AIC and BIC for Bivariate Extreme distribution



(b) AIC and BIC for location-scale Bivariate Extreme families

Figure 3. AIC and BIC for different bivariate extreme families.

6. Conclusion and future work

We used the most widely accepted theoretical assumptions that enable any family of bivariate dfs to fit a wide range of bivariate data. Based on these theoretical preconditions, we created the MBSSN family, which can better characterize several bivariate data sets of various forms. We used three bivariate data sets of three air contaminants and compared the resulting families with other well-known families. We compared these families' fits using the AIC and BIC criteria. Whether in a simulation (see Tables 3 and 4 and Figure 1) or an application to real data (see Tables 9–11 and Figure 2), the MBSSN family won the comparisons, followed by the FBSSN family. Also, by applying the same method to the two bivariate extreme distributions, we obtained a better result in modeling than the original distributions for BALog and BBLog (see Tables 14 and 15 and Figure 3).

The models $G_i(z, \mu, \sigma; \alpha, \ell), i = 1, 2$, give us high accuracy compared with other models, which is clear in our study. Additionally, they adapt to a variety of data types, including extreme data with high skewness and kurtosis, making them adaptable to diverse fields such as air pollution, finance, and healthcare. The application of these models across domains, particularly in healthcare and environmental sciences, could lead to significant advancements because of their advantages. We believe that these directions will contribute to the ongoing development of our model and enhance its relevance in practical applications. We will develop inference procedures for the bivariate versions $G_i(z, \mu, \sigma; \alpha, \ell), i = 1, 2$, as highlighted in the paper by Chen and Xiao [38]. Also, developing multivariate versions of the proposed models (see Xu et al. [39]) could yield valuable insights. Another area of future work involves considering the method of probability weight moments for estimating parameters of the $G_i(z, \mu, \sigma; \alpha, \ell), i = 1, 2$, distributions. This method is known to perform well for small samples. In the case of the bivariate extreme model, the additive model is limited by the computational complexity of estimating the parameter and the use of different methods for distribution fitting. Although the suggested distributions provide better fits for various bivariate extreme data sets, a drawback with the proposed distributions is the lack of an exact theoretical motivation from a bivariate extreme value context. The bivariate extreme value distribution is the limiting distribution of a linearly normalized bivariate maximum as the sample size approaches infinity. Such a result does not appear to be possible for our distributions.

Author contributions

The first, second, and third authors have made mathematics; the fourth, fifth, sixth and seventh made simulation and application.

Use of Generative-AI tools declaration

The author declares she has not used Artificial Intelligence (AI) tools in the creation of this article

Acknowledgments

The study was funded by Ongoing Research Funding program, (ORF-2025-488), King Saud University, Riyadh, Saudi Arabia.

Conflict of interest

The authors declare no conflict of interest

References

1. A. Azzalini, A class of distributions which includes the normal ones, *Scand. J. Stat.*, **12** (1985), 171–178.
2. A. Azzalini, A. D. Valle, The multivariate skew-normal distribution, *Biometrika*, **83** (1996), 715–726. https://doi.org/10.1093/biomet/83.4.715
3. H. M. Barakat, A new method for adding two parameters to a family of distributions with application to the average and exponential families, *Stat. Methods Appl.*, **24** (2015), 359–372. https://doi.org/10.1007/s10260-014-0265-8
4. H. M. Barakat, O. M. Khaled, Toward the establishment of a family of distributions that may fit any dataset, *Commun. Stat. Simul. Comput.*, **46** (2017), 6129–6143. https://doi.org/10.1080/03610918.2016.1197245
5. H. M. Barakat, O. M. Khaled, N. Khalil, New families of distributions for modeling bivariate data, with applications, *J. Math. Stat.*, **14** (2018), 79–87. https://doi.org/10.3844/jmssp.2018.79.87
6. H. M. Barakat, A. W. Aboutahoun, N. El kadar, A new extended mixture skew normal distribution, with applications, *Rev. Colomb. Estad.*, **42** (2019), 167–183. https://doi.org/10.15446/rce.v42n2.70087
7. H. M. Barakat, O. M. Khaled, N. Khalil, Modern techniques in data analysis, with application to the water pollution, *Proc. Latvian Acad. Sci.*, **72** (2018), 184–192. https://doi.org/10.2478/prolas-2018-0005

8. H. M. Barakat, E. M. Nigm, O. M. Khaled, *Statistical techniques for modelling extreme value data and related applications*, Cambridge Scholars Publishing, 2019.

9. H. M. Barakat, A. W. Aboutahoun, N. N. El-kadar, On some generalized families arising from mixture normal distribution with application, *Commun. Stat. Simul. Comput.*, **50** (2021), 198–216. https://doi.org/10.1080/03610918.2018.1554110

10. A. I. Aguirre-Salado, S. Venancio-Guzmán, C. A. Aguirre-Salado, A. Santiago-Santos, A novel tree ensemble model to approximate the generalized extreme value distribution parameters of the pm 2.5 maxima in the mexico city metropolitan area, *Mathematics*, **10** (2022), 2056. https://doi.org/10.3390/math10122056

11. N. Markovich, M. Vaičiulis, Extreme value statistics for evolving random networks, Mathematics, **11** (2023), 2171. https://doi.org/10.3390/math11092171

12. J. Galambos, Order statistics of samples from multivariate distributions, *J. Am. Stat. Assoc.*, **70** (1975), 674–680. https://doi.org/10.1080/01621459.1975.10482493

13. O. M. Khaled, H. Mohamed Barakat, N. Khalil Rakha, Improved inference for the generalized Pareto distribution under linear, power, and exponential normalization, *Kybernetika*, **58** (2022), 883–902. https://doi.org/10.14736/kyb-2022-6-0883

14. H. Haj Ahmad, E. M. Almetwally, D. A. Ramadan, Investigating the relationship between processor and memory reliability in data science: A bivariate model approach, *Mathematics*, **11** (2023), 2142. https://doi.org/10.3390/math11092142

15. M. E. Qura, A. Fayomi, M. Kilai, E. M. Almetwally, Bivariate power Lomax distribution with medical applications, *PLOS One*, **18** (2023), e0282581. https://doi.org/10.1371/journal.pone.0282581

16. E. S. A. El-Sherpieny, H. Z. Muhammed, E. M. Almetwally, Bivariate Chen distribution based on copula function: Properties and application of diabetic nephropathy, *J. Stat. Theory Pract.*, **16** (2022), 54. https://doi.org/10.1007/s42519-022-00275-7

17. J. Pikands, Multivariate extreme value distributions, *Bull. Int. Statist. Inst.*, **1981** (1981), 859–878.

18. J. A. Tawn, Bivariate extreme value theory: Models and estimation, *Biometrika*, **75** (1988), 397–415. https://doi.org/10.2307/2336591

19. H. Mayer, Air pollution in cities, *Atmos. Environ.*, **33** (1999), 4029–4037. https://doi.org/10.1016/S1352-2310(99)00144-2

20. S. Yue, J. Wang, A comparison of two bivariate extreme value distributions, *Stoch. Env. Res. Risk A.*, **18** (2004), 61–66. https://doi.org/10.1007/s00477-003-0124-x

21. E. S. A. El-Sherpieny, H. Z. Muhammed, E. M. Almetwally, Data analysis by adaptive progressive hybrid censored under bivariate model, *Ann. Data Sci.*, **11** (2024), 507–548. https://doi.org/10.1007/s40745-022-00455-z

22. P. Shi, A multivariate analysis of intercompany loss triangles, *J. Risk Insur.*, **84** (2017), 717–737. https://doi.org/10.1111/jori.12102

23. K. V. Mardia, Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies, *Sankhya B*, **36** (1974), 115–128.

24. R. E. Quandt, J. B. Ramsey, Estimating mixtures of normal distributions and switching regressions, *J. Am. Stat. Assoc.*, **73** (1978), 730–738. https://doi.org/10.1080/01621459.1978.10480085

25. T. I. Lin, J. C. Lee, S. Y. Yen, Finite mixture modeling using the skew-normal distribution, *Stat. Comput.*, **17** (2007), 81–92. https://doi.org/10.1007/s11222-006-9005-8

26. J. W. Tukey, Modern techniques in data analysis, In: *Proceedings of the NSF-Sponsored Regional Research Conference*, 1977.

27. R. R. Wilcox, Simulation results on solutions to the multivariate Behrens-Fisher problem via trimmed means, *Statistician*, **44** (1995), 213–225. https://doi.org/10.2307/2348445

28. C. Field, M. G. Genton, The multivariate g-and-h distribution, *Technometrics*, **48** (2006), 104–111.

29. D. C. Hoaglin, Summarizing shape numerically: The g-and-h distributions, In: *Exploring data tables, trends, and shapes*, 2006, 461–513. https://doi.org/10.1002/9781118150702.ch11

30. H. Akaike, Information theory and the maximum likelihood principle, In: *Proceedings of the Second International Symposium on Information Theory*, 1973.

31. G. Schwarz, Estimating the dimension of a model, *Ann. Stat.*, **6** (1978), 461–464. https://doi.org/10.1214/aos/1176344136

32. H. Teicher, Identifiability of finite mixtures, *Ann. Math. Statist.*, **34** (1963), 1265–1269. https://doi.org/10.1214/aoms/1177703862

33. N. Ho, X. Nguyen, Singularity structures and impacts on parameter estimation in finite mixtures of distributions, *SIAM J. Math. Data Sci.*, **1** (2019), 4. https://doi.org/10.1137/18M122947X

34. N. Ho, Parameter estimation and multilevel clustering with mixture and hierarchical models, 2017.

35. S. G. Coles, J. A. Tawn, Modeling extreme multivariate events, *J. R. Stat. Soc. B*, **53** (1991), 377-392. https://doi.org/10.1111/j.2517-6161.1991.tb01830.x

36. R. L. Smith, Extreme value theory, In: *Handbook of applicable mathematics*, **7** (1990), 437–471.

37. J. Beirlant, Y. Goegebeur, J. Segers, J. L. Teugels, Statistics of extremes: Theory and applications, John Wiley & Sons, 2004. https://doi.org/10.1002/0470012382

38. P. Chen, X. Xiao, Novel closed-form point estimators for the beta distribution, *Stat. Theor. Relat. Fields*, **9** (2025), 12–33. https://doi.org/10.1080/24754269.2024.2419360

39. A. Xu, G. Fang, L. Zhuang, C. Gu, A multivariate t-student process model for dependent tail-weighted degradation data, *IISE Trans.*, 2024, 1–17 https://doi.org/10.1080/24725854.2024.2389538

Appendix

Appendix 1. Proofs of Theorem 2.1

By (2.4) and (2.5) we have

$$\frac{\partial L_{\alpha,l}(t_1,t_2,\underline{\mu})}{\partial t_1} = [l\mu_1 + l^2(t_1 + t_2\rho)]L_{\alpha,l}(t_1,t_2,\underline{\mu}).$$

$$\frac{\partial L_{\alpha,l}(t_1,t_2,\underline{\mu})}{\partial t_2} = [l\mu_2 + l^2(t_1 + t_2\rho)]L_{\alpha,l,\underline{\mu}}(t_1,t_2).$$

$$\frac{\partial^2 L_{\alpha,l}(t_1,t_2,\underline{\mu})}{\partial t_1\partial t_2} = [[l\mu_1 + l^2(t_1+t_2\rho)][l\mu_2 + l^2(t_1+t_2\rho)] + l^2\rho]L_{\alpha,l}(t_1,t_2,\underline{\mu}).$$

$$\frac{\partial^2 L_{\alpha,l}(0,0,\underline{\mu})}{\partial t_1\partial t_2} = (l^2\mu_1\mu_2 + l^2\rho)\alpha = \alpha l^2(\mu_1\mu_2+\rho)$$

$$E(Y_1Y_2) = \left.\frac{\partial^2 L_{\alpha,l}(t_1,t_2)}{\partial t_1\partial t_2}\right|_{t_1=0,t_2=0} + \left.\frac{\partial^2 L_{\bar{\alpha},\frac{-1}{l}}(t_1,t_2)}{\partial t_1\partial t_2}\right|_{t_1=0,t_2=0}$$

$$\begin{aligned}E(Y_1Y_2) &= (l^2\mu_1\mu_2 + l^2\rho)\alpha + (\bar{\alpha}\frac{\mu_1\mu_2}{l^2} + \frac{\rho}{cl^2})\\ &= \frac{1}{l^2}[(\alpha l^4+\bar{\alpha})\mu_1\mu_2 + (\alpha l^4+\bar{\alpha})\rho]\\ &= \frac{1}{l^2}[(\alpha l^4+\bar{\alpha})(\rho+\mu_1\mu_2)]\end{aligned}$$

and

$$E(Y_1)E(Y_2) = (\alpha l - \frac{\bar{\alpha}}{l})^2\mu_1\mu_2 = \frac{1}{l^2}(\alpha l^2-\bar{\alpha})^2\mu_1\mu_2.$$

Therefore, we have

$$\begin{aligned}COV(Y_1,Y_2) &= E(Y_1Y_2) - E(Y_1)E(Y_2)\\ &= \frac{1}{l^2}(\alpha l^4+\bar{\alpha})(\mu_1\mu_2+\rho) - (\alpha l^2-\bar{\alpha})^2)\mu_1\mu_2\\ &= \frac{1}{l^2}[(\alpha l^4+\bar{\alpha})\rho + \alpha\bar{\alpha}\mu_1\mu_2(l^2+1)^2].\end{aligned}$$

Since

$$\rho_{\underline{Y}} = \frac{COV(Y_1Y_2)}{\sigma_{Y_1}\sigma_{Y_2})}.$$

$$\rho_{\underline{Y}} = \frac{(\alpha l^4+\bar{\alpha})\rho + \alpha\bar{\alpha}\mu_1\mu_2(l^2+1)^2}{\sqrt{[(1+l^2)^2\mu_1\alpha\bar{\alpha} + (\alpha l^4+\bar{\alpha})][(1+l^2)^2\mu_2\alpha\bar{\alpha} + (\alpha l^4+\bar{\alpha})}}. \tag{.1}$$

Next, we can prove the statements

(1) By substituting with $\mu_1 = \mu_2 = 0$ in (.1) we get

$$\rho_{\underline{Y}} = \frac{(\alpha l^4+\bar{\alpha})\rho}{\sqrt{(\alpha l^4+\bar{\alpha})^2}} = \rho,$$

and by substituting with $\alpha = 1$ we get $\rho_{\underline{Y}} = \frac{l^4\rho}{\sqrt{l^4 l^4}} = \rho$.

(2) For $\rho = 1$ and $\alpha = 1$ implies $\rho_{\underline{Y}} = 1$ and for $\rho = 1$ and $\mu_1 = \mu_2 = 0$ implies also $\rho_{\underline{Y}} = 1$.

(3) For $\rho = -1$ and $\alpha = 1$ implies $\rho_{\underline{Y}} = -1$ and for $\rho = -1$ and $\mu_1 = \mu_2 = 0$ implies also $\rho_{\underline{Y}} = -1$.

(4) If $\rho_{\underline{Y}} = 0$ then $(\alpha l^4+\bar{\alpha})\rho + \alpha\bar{\alpha}\mu_1\mu_2(l^2+1)^2 = 0$ which implies that

$$-1 \le \rho = \frac{-\alpha\bar{\alpha}\mu_1\mu_2(l^2+1)^2}{\alpha l^4+\bar{\alpha}} \le 0.$$

(5) By substituting with $\mu_1 = \mu_2 = \mu$ we get

$$\rho_{\underline{Y}} = \frac{(\alpha l^4 + \bar{\alpha})\rho + \alpha\bar{\alpha}\mu^2(l^2+1)^2}{(1+l^2)^2\mu\alpha\bar{\alpha} + (\alpha l\bar{\alpha})} = \rho_Y(\alpha, l, \mu). \tag{$*$}$$

By using (*), we get

$$\rho_Y(1, l, \mu) = \rho_Y(\alpha, l, 0) = \rho, \quad \rho_Y(\alpha, 1, \mu) = \frac{\rho + 4\alpha\bar{\alpha}\mu^2}{1 + \alpha\bar{\alpha}\mu^2}$$

and $\rho_Y(1, l, \mu) = 0$, if and only if $-1 \le \rho = \frac{-\alpha\bar{\alpha}\mu^2(l^2+1)^2}{\alpha l^4 + \bar{\alpha}} \le 0.$

This completes the proof.

Appendix 2. Proofs of Theorem 2.1 The skewness and the kurtosis of MBSSN df.

$$\begin{aligned}
\beta_{1,2} = {} & (-5(l^2\alpha - \bar{\alpha})^6\mu_1^3\mu_2^3 + 3(l^2\alpha - \bar{\alpha})^4(l^4\alpha + \bar{\alpha})\mu_1(1+\mu_1)\mu_2^3 \\
& - (l^2\alpha - \bar{\alpha})(l^6\alpha - \bar{\alpha})\mu_1(3 + \mu_1^2)\mu_2^3 \\
& + 9(l^2\alpha - \bar{\alpha})^4(l^4\alpha + \bar{\alpha})\mu_1^2\mu_2^2(\rho + \mu_1\mu_2) \\
& - 9(l^2\alpha - \bar{\alpha})^3(l^6\alpha - \bar{\alpha})\mu_1\mu_2^2(2\rho\mu_1 + \mu_2 + \mu_1^2\mu_2) \\
& + 3(l^2\alpha - \bar{\alpha})^2(l^8\alpha + \bar{\alpha})\mu_2^2(-\rho + \mu_1^2 + \mu_1\mu_2 + \mu_1^3\mu_2) \\
& + 3(l^2\alpha - \bar{\alpha})^2(l^4\alpha + \bar{\alpha})\mu_1^3\mu_2(1 + \mu_2^3) \\
& - (l^2\alpha - \bar{\alpha})^3(l^6\alpha - \bar{\alpha})\mu_1^3\mu_2(3 + \mu_2^2) \\
& - 9(l^2\alpha - \bar{\alpha})^3(l^6\alpha - \bar{\alpha})\mu_1^2\mu_2(3\rho\mu_2 + \mu_1(1 + \mu_2^2)) \\
& + 9(l^2\alpha - \bar{\alpha})^2(l^8\alpha + \bar{\alpha})\mu_1\mu_2(1 + 2\rho^2 + \mu_1\mu_2(4 + \mu_2) + \mu_1^2(1 + \mu_2^2)) \\
& - 3(l^2\alpha - \bar{\alpha})^2(l^{10}\alpha - \bar{\alpha})\mu_2(6\rho\mu_2 + 6\mu_1^2\mu_2 + \mu_1^3(1 + \mu_2^2) + \mu_1(3 + 6\rho^2 + \mu_2^2)) \\
& + (l^{12}\alpha + \bar{\alpha})(9\rho\mu_1^2 + \mu_1^3\mu_2(3 + \mu_2^2) + 3\mu_1\mu_2(3 + 3\rho + 6\rho^2 + \mu_2^2) + 3\rho(3 + \rho^2 + 3\mu_2^2)) \\
& - 3(l^2\alpha - \bar{\alpha})^2(l^8\alpha + \bar{\alpha})\mu_1^2(3\rho + \mu_2^2 + \mu_1(\mu_2 + \mu_2^3)) \\
& - (l^2\alpha - \bar{\alpha})^2(l^{10}\alpha - \bar{\alpha})\mu_1(6\mu_1(\rho + \mu_2^2) + \mu_2(3 + 6\rho^2 + \mu_2^2) + \mu_1^2(\mu_2 + \mu_2^3))) \\
& /(((l^4\alpha + \bar{\alpha})(\mu_1\mu_2 + \rho) - (l^2\alpha - \bar{\alpha})^2)\mu_1\mu_2)^3.
\end{aligned}$$

and the kurtosis as

$$\begin{aligned}
\beta_{2,2} = ((&l^2\alpha-\bar{\alpha})^4(l^8\alpha+\bar{\alpha})(3+6\mu_1^2+\mu_1^4)\mu_2^4-4(l^2\alpha-\bar{\alpha})^3(l^{10}\alpha-\bar{\alpha})\mu_2^3(12\rho\mu_1\\
&+4\rho\mu_1^3+3\mu_2+6\mu_1^2\mu_2+\mu_1^4\mu_2)-(l^2\alpha-\bar{\alpha})^4\mu_1^4(3-3(l^8-1)\alpha\\
&+6(1+l^2)^2(l^4\bar{\alpha}+\alpha)\alpha\bar{\alpha}\mu_2^2\\
&+(4-(17+20l^2+6l^4+6l^4+l^8)\alpha+4(1+l^2)^2(7+2l^2+l^4)\alpha^2\\
&-2(1+l^2)^3(11+3l2)\alpha^3+7(1+l^2)^4\alpha^4)\mu_2^4)\\
&+6(l^2\alpha-\bar{\alpha})^2(l^{12}\alpha+\bar{\alpha})\mu_2^2(24\rho\mu_1\mu_2+8\rho\mu_1^3\mu_2+\mu^4(1+\mu_2^2)\\
&+6\mu_1^2(1+2\rho^2+\mu_2^2)+3(1+4\rho^2+\mu_2^2))\\
&-4(l^2\alpha-\bar{\alpha})(l^{14}\alpha-\bar{\alpha})\mu_2(12\rho\mu_1^3(1+\mu_2^2)\mu_1^4\mu_2(3+\mu_2^2)\\
&+6\mu_1^2\mu_2(3+6\rho^2+\mu_2^2)+3\mu_2(3+12\rho^2+\mu_2^2)+12\rho\mu_1(3+2\rho^2+3\mu_2^2))\\
&+\bar{\alpha}(16\rho\mu_1^3\mu_2(3+\mu_2^2)+48\rho\mu_1\mu_2(3+2\rho^2+\mu_2^2)+12\rho\mu_1(3+2\rho^2+3\mu_2^2)\\
&+6\mu_1^2(3+12\rho^2+6(1+2\rho^2)\mu_2^2+\mu_2^4)+3(3+24\rho^2+8\rho^4+6(1+2\rho^2)\mu_2^2+\mu_2^4))\\
&+l^{16}\alpha(16\rho\mu_1^3\mu_2(3+\mu_2^2)+48\rho\mu_1\mu_2(3+2\rho^2+\mu_2^2)+\mu_1^4(3+6\mu_2^2+\mu_2^4)\\
&+6\mu_1^2(3+12\rho^2+6(1+2\rho^2)\mu_2^2+\mu_2^4)+3(3+24\rho^2+8\rho^4+6(1+4\rho^2)\mu_2^2+\mu_2^4))\\
&-4(l^2\alpha-\bar{\alpha})^3\mu_1^3(-4(l^2\alpha-\bar{\alpha})^3(l^4\alpha+\bar{\alpha})\mu_2^3(\rho+\mu_1\mu_2)\\
&+4(l^{10}-\bar{\alpha})\alpha\rho\mu_2-(l^{10}\alpha-\bar{\alpha})\mu_1(3+6\mu_2^2+\mu_2^4)\\
&-6(l^2\alpha-\bar{\alpha})^2(l^6\alpha-\bar{\alpha})\mu_2^2(2\rho\mu_2+\mu_1(1+\mu_2^2))\\
&-4(l^2\alpha-\bar{\alpha})(l^8\alpha+\bar{\alpha})\mu_2(3\rho(1+\mu_2^2)+\mu_1\mu_2(3+\mu_2^2)))\\
&+6(l^2\alpha-\bar{\alpha})^2\mu_1^2((l^2\alpha-\bar{\alpha})^4(l^4\alpha-\bar{\alpha})(1+\mu_1^2)\mu_2^4\\
&-4(l^2\alpha-\bar{\alpha})^3(l^6\alpha-\bar{\alpha})\mu_2^3(2\rho\mu_1+\mu_2+\mu_1^2\mu_2)\\
&+6(l^2\alpha-\bar{\alpha})^2(l^8\alpha+\bar{\alpha})\mu_2^2(1+2\rho^2+4\rho\mu_1\mu_2+\mu_2^2+\mu_1^2(1+\mu_2^2))\\
&-(l^2\alpha-\bar{\alpha})(l^{10}\alpha-\bar{\alpha})\mu_2(6\rho\mu_1(1+\mu_2^2)+\mu_1^2\mu_2(3+\mu_2^2)+\mu_2(3+6\rho^2+\mu_2^2))
\end{aligned}$$

$$\begin{aligned}
&+(l^{12}\alpha-\bar{\alpha})(3+12\rho^2+6(1+2\rho^2)\mu_2^2+\mu_2^4+8\rho\mu_1\mu_2(3+\mu_2^2)+\mu_1^2(3+6\mu_2^2+\mu_2^4)))\\
&+4(l^2\alpha-\bar{\alpha})\mu_1((l^2\alpha-\bar{\alpha})^4(l^6\alpha-\bar{\alpha})\mu_1(3+\mu_1^2)\mu_2^4\\
&+(l^2\alpha-\bar{\alpha})^3(l^8\alpha+\bar{\alpha})\mu_2^3(3\rho-3\rho\mu_1^2+3\mu_1\mu_2+\mu_1^3\mu_2)\\
&+(l^2\alpha-\bar{\alpha})^2(l^{16}\alpha-\bar{\alpha})\mu_2^2(6\rho\mu_2+6\rho\mu_1^2\mu_2)+\mu_1^3(1+\mu_1^2)+3\mu_1(1+2\rho^2+\mu_2^2))\\
&-4(l^2\alpha-\bar{\alpha})(l^{12}\alpha+\bar{\alpha})\mu_2(9\rho\mu_1^2(1+\mu_2^2)+\mu_1^3\mu_2(3+\mu_2^2)\\
&+3\mu_1\mu_2(3+6\rho^2+\mu_2^2)+3\rho(3+2\rho^2+3\mu_2^2))\\
&+(l^{14}\alpha-\bar{\alpha})(12\rho\mu_1^2\mu_2(3+\mu_2^2)+12\rho\mu_2(3+2\rho^2+\mu_2^2)+\mu_1^3(3+6\mu_2^2+\mu_2^4)\\
&+3\mu_1(3+12\rho^2+6(1+2\rho^2)\mu_2^2+\mu_2^4)))/(((l^4\alpha+\bar{\alpha})\rho+\alpha\bar{\alpha}\mu_1\mu_2(1+c^2)^2)^4.
\end{aligned}$$