Article

A New Bivariate Family Based on Archimedean Copulas: Simulation, Regression Model and Application

Gabriela M. Rodrigues [1], Edwin M. M. Ortega [1,*], Roberto Vila [2] and Gauss M. Cordeiro [3]

1 Department of Exact Sciences, University of São Paulo, Piracicaba 13418-900, Brazil; gabrielar@usp.br
2 Department of Statistics, University of Brasilia, Brasilia 70910-900, Brazil; rovig161@unb.br
3 Department of Statistics, Federal University of Pernambuco, Recife 50670-901, Brazil; gauss@de.ufpe.br
* Correspondence: edwin@usp.br

Abstract: We use the Clayton and Frank copulas and the exponentiated odd log-logistic family to define a new flexible bivariate model to fit bimodal and asymmetry data. The copulas allow different distributions for the response variable, thus making analysis more suitable. We present some structural properties of the new model and describe a simulation study to show the consistency of the estimators. We construct a bivariate regression model based on the new family to fit oak lettuce plant data for different concentrations of silicon dioxide and organosilicon compounds. We check the response variables fresh weight and plant height together in order to verify the existing correlation between them. These variables exhibit a bimodal form, and the family used is able to model this behavior. Different marginal distributions are selected, which is an interesting point of the copula methodology. The variables have strong positive dependence, and the experiment is carried out comparing the control treatment with others leading to the following results: (i) the treatment 1-ethoxysilatrane (with concentrations 5×10^{-4} mL·L^{-1} and 10^{-3} mL·L^{-1}) is not significant for the response variables; (ii) the treatment amorphous silicon dioxide (with concentrations 50 mg·L^{-1} and 100 mg·L^{-1}) and the same treatment (with concentrations 5×10^{-3} mL·L^{-1} and 10^{-2} mL·L^{-1}) are significant and have positive effects on both responses; (iii) the treatment amorphous silicon dioxide (with concentrations 200 mg·L^{-1} and 300 mg·L^{-1}) are significant and have negative effects on the response variables. Overall, the proposed bivariate model is suitable for the current data and can be useful in other applications.

Keywords: bivariate model; copula function; maximum likelihood estimation; regression model; simulation study

1. Introduction

Various areas of knowledge can use modeling involving one or more response variables of interest, i.e., two or more attributes to be modeled. If these attributes are not independent and a practical explanation exists for this situation, multivariate statistical models should be used with the objective of explaining and capturing the correlation of these variables.

For this purpose, joint continuous distributions are commonly used. The bivariate normal and t-Student distributions are most often used, although they can be highly restrictive. Besides this, employing them implies a linear relationship and elliptical structure of the variables. These premises do not always hold.

The modeling of multivariate data based on copula functions is an interesting alternative to overcome these drawbacks. According to Nelsen (1986) [1], copulas provide a way to relate functions of multivariate distributions based on their marginal distribution functions. The main advantages from a statistical standpoint using this method are:

- For constructing joint distributions, copulas allow each of them to be modeled individually by a different marginal distribution, thus enabling more flexible associations by fitting different marginal distributions;
- Employment of copulas is a useful approach to model and understand the phenomenon of dependence, i.e., the dependence between these variables can assume diverse structures, including nonlinear ones, according to the type of copula utilized;
- The marginal distributions should not depend on the association of the variables under analysis; and
- A copula is invariant from continuous transformations of the marginals.

Various bivariate distributions have been proposed based on copula functions such as the Weibull bivariate derived from the Farlie–Gumbel–Morgenstern (FGM) copula functions, Ali–Mikhail–Haq (AMH), Gumbel–Hougaard, Gumbel–Barnett [2], generalized bivariate Rayleigh using the Clayton copula (El-Sherpieny and Almetwally [3]), bivariate Fréchet based on the FGM or AMH copulas [4], and generalized inverted Kumaraswamy using the Marshal–Olkin method [5]. Samanthi and Sepanski [6] defined families from four bivariate copulas using the Kumaraswamy distribution, the bivariate exponentiated half logistic based on the Marshall–Olkin class [7], and the bivariate Lindley distribution via the FGM copula [8]. Here, we focus on Archimedean copulas, which have closed-form expressions as an important characteristic, since they can be easily constructed starting from specific generator functions. Further, they are highly flexible, permitting the modeling of various forms of dependence, including asymmetry and extreme dependence. Due to the ease of their construction, these copulas have a large number of applications in various areas of knowledge, such as finance ([3,9,10]), health ([11,12]), hydrology ([13–15]), and survival analysis ([16–18]). Detailed studies of these families and their applications can be found in Nelsen (2007).

In many practical situations, the response variable presents an asymmetric and/or bimodal behavior. The motivation for this work comes from an experiment carried out at the Plekhanov Russian University of Economics that evaluated the growth of oak leaf lettuce (Lactuca sativa var. crispa). The histograms of the response variables fresh weight (grams) (y_1) and plant height (cm) (y_2) measured in the experiment are reported in Figure 1a,b. It is noted that they present bimodal behavior and positive asymmetries. In addition, the scatter plot between them (Figure 1c) indicates that they present a strong and positive correlation.
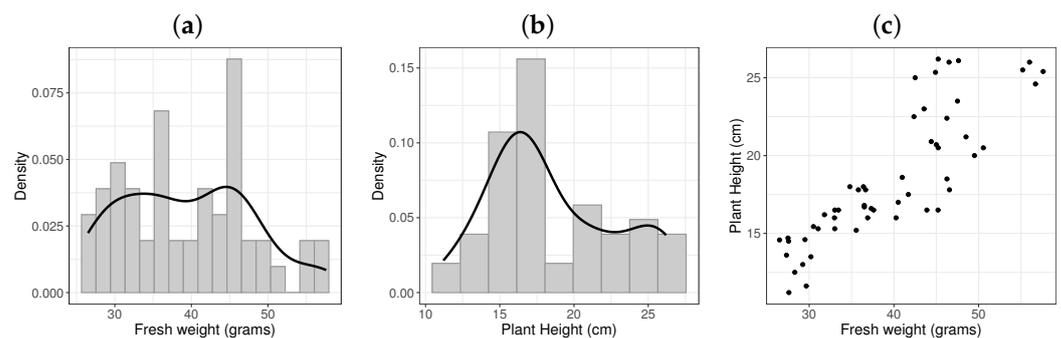

Figure 1. (**a**) Histogram of fresh weight, (**b**) Histogram of plant height, and (**c**) Scatter plot between fresh weight and plant height.

Although a wide range of flexible bivariate distributions can be found in the literature, we note a scarcity of bimodal bivariate distributions. For this purpose, we use the *exponentiated odd log-logistic-G* (EOLL-G) family (Alizadeh et al., 2020) [19]), whose densities have great flexibility in modeling data, such as bimodality and/or positive or negative asymmetry. The Clayton and Frank copulas are adopted, so the new models are called the BCEOLL-G and BFEOLL-G families, respectively. We consider these copulas because they are suitable to model data with positive correlation (Naifar, 2011) [9] according to the dataset in Figure 1c.

This paper is structured as follows: Section 2 provides a summary of copula functions. Section 3 proposes the bivariate Clayton and Frank copulas generated from the EOLL-G family. Section 4 introduces the Frank and Clayton copulas generated from the EOLL Normal distribution. Section 5 formulates the bivariate bimodal regression with copulas and presents inferential issues. Section 6 performs some simulations for different scenarios. Section 7 illustrates the new methodology for lettuce plants from an experimental trial. Section 8 concludes the article.

2. Archimedean Copulas

According to [20], copulas can be described as functions that link univariate marginal distributions, or alternatively, functions with a multivariate distribution whose marginals are uniform in the interval $[0, 1]$. The name and theory of copulas are based on the theorem of Sklar [21]. This important theorem pertaining to the copula method, which guarantees the mentioned relations, is described below. The proof can be found in [22].

Sklar's Theorem: Let $Y = (Y_1, \ldots, Y_d)$ be a random vector with marginal cumulative distribution functions (cdfs) $F_1(y_1), \ldots, F_d(y_d)$, and let $F(y_1 \ldots, y_d)$ be their joint cdfs. Define $u_i = F_i(y_i) = P(Y_i \leq y_i)$, $i = 1, \ldots, d$. Then, there is a copula function $C(\cdot)$ such that

$$C(u_1, \ldots, u_d) = P(Y_1 \leq y_1, \ldots, Y_d \leq y_d) = F(y_1, \ldots, y_d). \tag{1}$$

By differentiating (1), the joint probability density function (pdf) follows as

$$f(y_1, \ldots, y_d) = c(F_1(y_1), \ldots, F_d(y_d)) \prod_{i=1}^{d} f_i(y_i), \tag{2}$$

where $f_i(y_i)$ is the marginal pdf of Y_i and $c(F_1(y_1), \ldots, F_d(y_d))$ is the copula density by taking the derivative of the copula function. For independent marginals, $C(u_1, \ldots, u_d) = \prod_{i=1}^{d} u_i$ and $c(F_1(y_1), \ldots, F_d(y_d)) = 1$, and then $f(y_1, \ldots, y_d) = \prod_{i=1}^{d} f_i(y_i)$, which means the independence of the random variables.

Henceforth, we consider the bivariate case for the data introduced in Section 1.

Archimedean copulas are constructed by a strictly decreasing continuous generating function $\varphi : [0, 1] \rightarrow [0, \infty]$, where $\varphi(0) = \infty$ and $\varphi(1) = 0$. Thus, the distribution function of a two-dimensional Archimedean copula can be expressed as

$$C(\boldsymbol{u}, \lambda) = \varphi^{-1}(\varphi(u_1) + \varphi(u_2); \lambda), \quad \boldsymbol{u} = (u_1, u_2)^{\top},$$

where λ is the control parameter of the degree of dependency. The lower and upper tail dependence measures for a bivariate copula C are defined in Joe (1993) [23] (if the limits exist) as

$$\lambda_{\mathrm{L}} = \lim_{u \to 0} \frac{C(u, u)}{u}, \quad \lambda_{\mathrm{U}} = \lim_{u \to 1} \frac{1 - 2u + C(u, u)}{1 - u}.$$

For the Archimedean copula, these limits hold

$$\lambda_{\mathrm{L}} = \begin{cases} \lim_{u \to \infty} \dfrac{\varphi^{-1}(2u)}{\varphi^{-1}(u)}, & \text{if } \varphi(0) = \infty, \\ 0, & \text{otherwise,} \end{cases} \quad \lambda_{\mathrm{U}} = 2 - \lim_{u \to 0} \frac{1 - \varphi^{-1}(2u)}{1 - \varphi^{-1}(u)}. \tag{3}$$

Here, we employ the most cited bivariate Archimedean copulas (Clayton, 1978 [24]; Frank, 1979 [25]) for constructing new bivariate response regression models.

2.1. Clayton Copula

The Clayton copula [24] has generating function $\varphi(t) = (t^{-\lambda} - 1)/\lambda$ and distribution and density functions given by

$$C(u_1, u_2) = (u_1^{-\lambda} + u_2^{-\lambda} - 1)^{-\frac{1}{\lambda}}, \tag{4}$$

and

$$c(u_1, u_2) = (\lambda + 1)(u_1^{-\lambda} + u_2^{-\lambda} - 1)^{\frac{-(2\lambda+1)}{\lambda}} (u_1 u_2)^{-(\lambda+1)}, \tag{5}$$

respectively, where $\lambda > 0$. In this case, there is independence between the two variables when λ tends to zero. As $\varphi(0) = \infty$ and $\varphi^{-1}(t) = (1 + \lambda t)^{-1/\lambda}$, it follows from Equation (3)

$$\lambda_{\mathrm{L}} = \lim_{u \to \infty} \left(1 + \frac{\lambda u}{1 + \lambda u}\right)^{-1/\lambda} = 2^{-1/\lambda}$$

and

$$\lambda_{\mathrm{U}} = 2 - \lim_{u \to 0} \frac{1 - (1 + 2\lambda u)^{-1/\lambda}}{1 - (1 + \lambda u)^{-1/\lambda}} = 0.$$

2.2. Frank Copula

The Frank copula [25] has generating function $\varphi(t) = -\log[(e^{-\lambda t} - 1)/(e^{-\lambda} - 1)]$ and distribution and density functions given by

$$C(u_1, u_2) = -\frac{1}{\lambda} \log \left\{ 1 + \frac{(e^{-\lambda u_1} - 1)(e^{-\lambda u_2} - 1)}{e^{-\lambda} - 1} \right\} \tag{6}$$

and

$$c(u_1, u_2) = \frac{\lambda e^{-\lambda u_1 - \lambda u_2}(1 - e^{-\lambda})}{(e^{-\lambda} - e^{-\lambda u_1} - e^{-\lambda u_2} + e^{-\lambda u_1 - \lambda u_2})^2}, \tag{7}$$

respectively, where $\lambda \in \mathbb{R} \setminus \{0\}$. If λ tends to zero, the two variables can be independent, and if λ tends to infinity the two variables are correlated. As $\varphi(0) = -\infty$ and $\varphi^{-1}(t) = -\log\{1 + \exp(-t)[\exp(-\lambda) - 1]\}/\lambda$, it follows from Equation (3) using L'Hôpital rule $\lambda_{\mathrm{L}} = 0$ and

$$\lambda_{\mathrm{U}} = 2 - \lim_{u \to 0} \frac{\lambda + \log\{1 + e^{-2u}[e^{-\lambda} - 1]\}}{\lambda + \log\{1 + e^{-u}[e^{-\lambda} - 1]\}} = 0.$$

3. A Bivariate EOLL-G Family Based on Clayton and Frank Copulas

For any baseline cdf $G(y) = G(y; \boldsymbol{\eta})$ depending on a parameter vector $\boldsymbol{\eta}$, Alizadeh et al. (2020) [19] defined the cdf of the exponentiated odd log-logistic ("EOLL-G") family (for $y \in \mathbb{R}$) by

$$F(y) = F(y; \nu, \tau, \boldsymbol{\eta}) = \frac{G^{\nu\tau}(y)}{\{G^{\nu}(y) + [\bar{G}(y)]^{\nu}\}^{\tau}}, \tag{8}$$

where $\bar{G}(y) = 1 - G(y)$, $\nu > 0$ and $\tau > 0$ are two extra shape parameters.

The pdf associated to Equation (8) becomes

$$f(y) = f(y; \nu, \tau, \boldsymbol{\eta}) = \frac{\nu \tau G^{\nu\tau-1}(y)[\bar{G}(y)]^{\nu-1} g(y)}{\{G^{\nu}(y) + [\bar{G}(y)]^{\nu}\}^{\tau+1}}, \tag{9}$$

where $g(y) = dG(y)/dy$.

From now on, let $Y \sim$ EOLL-G$(\nu, \tau, \boldsymbol{\eta})$ be a random variable with pdf (9). Let X have cdf $G(\cdot; \boldsymbol{\eta})$ such that $\mathbb{E}(X^p) = \int x^p \, g(x)dx$ (for $p > 0$) are finite. Some properties of Y are reported below:

(a) The positive moments of Y, i.e., $\mathbb{E}(Y^p) = \int y^p f(y)dy$, $p > 0$, are finite when $\nu > 1$ (see Appendix A.1).

(b) The variable Y admits the stochastic representation: $Y = G^{-1}(S/(S+1))$, where S has the Dagum distribution with shape parameters ν and τ, and unit scale, and $G^{-1}(\cdot)$ denotes the inverse function of $G(\cdot)$ (see Appendix A.2).

(c) The positive moments of the standardization version of Y, say $[Y - \mathbb{E}(Y)]/\sqrt{\text{Var}(Y)}$, are finite if $\nu > 1$ (see Appendix A.3).

Definition 1. *Let $Y_1', \ldots, Y_d'$ be an independent copy of $Y_1, \ldots, Y_d$, i.e., $\boldsymbol{Y}' = (Y_1', \ldots, Y_d')$ is independent of $\boldsymbol{Y} = (Y_1, \ldots, Y_d)$ and both have the same joint cdf $F = F_{\boldsymbol{Y}}$. Following Koshevoy (1997) [26], the distance-Gini mean difference for F is*

$$M_D(F) \equiv \frac{1}{2d} \, \mathbb{E}(\|\boldsymbol{Y} - \boldsymbol{Y}'\|) = \frac{1}{2d} \int_{(0,\infty)^d} \int_{(0,\infty)^d} \|\boldsymbol{y} - \boldsymbol{y}'\| dF(\boldsymbol{y})dF(\boldsymbol{y}'), \qquad (10)$$

where $\| \cdot \|$ denotes the Euclidean distance in $\mathbb{R}^2$.

(d) The distance-Gini mean difference $M_D(F)$ corresponding to the joint cdf F in (1), where $Y_i \sim$ EOLL-G$(\nu_i, \tau_i, \boldsymbol{\eta}_i)$, $i = 1, \ldots, d$, is finite when $\nu > 1$ and X have moments of order greater than one (see Appendix A.4).

(e) If $U \sim U(0,1)$, then

$$Q_G \left\{ \frac{U^{1/(\nu\tau)}}{U^{1/(\nu\tau)} + (1 - U^{1/\tau})^{1/\nu}} \right\} \sim \text{EOLL-G}(\nu, \tau, \boldsymbol{\eta}), \qquad (11)$$

where $Q_G(u) = G^{-1}(u)$ is the quantile function (qf) of G.

The EOLL-G family includes as special cases: the OLL-G class for $\tau = 1$ (Gleaton and Lynch, 2006 [27]), and the exponentiated (Exp-G) class (Mudholkar et al., 1996 [28]) for $\nu = 1$. Clearly, Equation (9) becomes the baseline G when $\nu = \tau = 1$.

Thus, we consider the following marginal distributions

$$Y_1 \sim \text{EOLL-G}(\nu_1, \tau_1 \boldsymbol{\eta}_1) \qquad \text{and} \qquad Y_2 \sim \text{EOLL-G}(\nu_2, \tau_2, \boldsymbol{\eta}_2), \qquad (12)$$

where $\boldsymbol{\eta}_1$ and $\boldsymbol{\eta}_2$ are parameter vectors of the baseline G, and ν_1, τ_1, ν_2 and τ_2 are positive shape parameters.

3.1. BCEOLL-G Model

By inserting (8) and (9) in Equations (1), (2), (4) and (5), the bivariate joint BCEOLL-G cdf reduces to

$$F_{\text{BCEOLL-G}}(y_1, y_2) = \left[W_1^{-\lambda}(y_1; \boldsymbol{\eta}_1) + W_2^{-\lambda}(y_2; \boldsymbol{\eta}_2) - 1 \right]^{-\frac{1}{\lambda}}. \qquad (13)$$

The corresponding joint pdf has the form

$$
\begin{aligned}
f_{\text{BCEOLL-G}}(y_1, y_2) &= (\lambda + 1)\left[W_1^{-\lambda}(y_1; \boldsymbol{\eta}_1) + W_2^{-\lambda}(y_2; \boldsymbol{\eta}_2) - 1\right]^{\frac{-(2\lambda+1)}{\lambda}} \times \\
&\quad [W_1(y_1; \boldsymbol{\eta}_1)W_2(y_2; \boldsymbol{\eta}_2)]^{-(\lambda+1)} \times \\
&\quad \left[\frac{\nu_1\,\tau_1\,g_1(y_1; \boldsymbol{\eta}_1)\,G_1^{\nu_1\,\tau_1-1}(y_1; \boldsymbol{\eta}_1)\bar{G}_1^{\nu_1-1}(y_1; \boldsymbol{\eta}_1)}{H_1^{\tau_1+1}(y_1; \boldsymbol{\eta}_1)}\right] \times \\
&\quad \left[\frac{\nu_2\,\tau_2\,g_2(y_2; \boldsymbol{\eta}_2)\,G_2^{\nu_2\,\tau_2-1}(y_2; \boldsymbol{\eta}_2)\bar{G}_2^{\nu_2-1}(y_2; \boldsymbol{\eta}_2)}{H_2^{\tau_2+1}(y_2; \boldsymbol{\eta}_2)}\right],
\end{aligned} \tag{14}
$$

where (for $k = 1, 2$)

$$
H_k(y_k; \boldsymbol{\eta}_k) = G_k^{\nu_k}(y_k; \boldsymbol{\eta}_k) + [1 - G_k(y_k; \boldsymbol{\eta}_k)]^{\nu_k}, \qquad W_k(y_k; \boldsymbol{\eta}_k) = \frac{G_k^{\nu_k\,\tau_k}(y_k, \boldsymbol{\eta}_k)}{H_k^{\tau_k}(y_k; \boldsymbol{\eta}_k)}.
$$

3.2. BFEOLL-G Model

Similarly, the joint cdf of the bivariate BFEOLL-G model can be expressed as

$$
F_{\text{BFEOLL-G}}(y_1, y_2) = -\frac{1}{\lambda}\log\left\{1 + \frac{\{\exp[-\lambda\,W_1(y_1; \boldsymbol{\eta}_1)] - 1\}\{\exp[-\lambda\,W_2(y_2; \boldsymbol{\eta}_2)] - 1\}}{\exp(-\lambda) - 1}\right\}. \tag{15}
$$

The corresponding joint BFEOLL-G pdf becomes

$$
\begin{aligned}
f_{\text{BFEOLL-G}}(y_1, y_2) &= \lambda[1 - \exp(-\lambda)]\exp\{-\lambda[W_1(y_1; \boldsymbol{\eta}_1) + W_2(y_2; \boldsymbol{\eta}_2)]\} \times \\
&\quad \{\exp(-\lambda) - \exp[-\lambda W_1(y_1; \boldsymbol{\eta}_1)] - \exp[-\lambda W_2(y_2; \boldsymbol{\eta}_2)] + \exp[-\lambda(W_1(y_1; \boldsymbol{\eta}_1) + W_2(y_2; \boldsymbol{\eta}_2))]\}^{-2} \times \\
&\quad \left[\frac{\nu_1\,\tau_1\,g_1(y_1; \boldsymbol{\eta}_1)\,G_1^{\nu_1\,\tau_1-1}(y_1; \boldsymbol{\eta}_1)\bar{G}_1^{\nu_1-1}(y_1; \boldsymbol{\eta}_1)}{H_1^{\tau_1+1}(y_1; \boldsymbol{\eta}_1)}\right] \times \\
&\quad \left[\frac{\nu_2\,\tau_2\,g_2(y_2; \boldsymbol{\eta}_2)\,G_2^{\nu_2\,\tau_2-1}(y_2; \boldsymbol{\eta}_2)\bar{G}_2^{\nu_2-1}(y_2; \boldsymbol{\eta}_2)}{H_2^{\tau_2+1}(y_2; \boldsymbol{\eta}_2)}\right].
\end{aligned} \tag{16}
$$

The BCEOLL-G and BFEOLL-G models include three special cases:

- The bivariate Clayton Exponentiated (Exp)-G and Frank Exp-G classes when $\nu = 1$;
- The bivariate Clayton odd log-logistic (OLL)-G and Frank OLL-G classes when $\tau = 1$.
- The bivariate Clayton and Frank baseline models when $\nu = \tau = 1$.

We can generate many bivariate models from Equations (14) and (16) by choosing different parent distributions.

3.3. Copula Dependence Measures

Pearson correlation is one of the most widely used measures, but it cannot be used in cases where the joint distributions are not normally distributed, and also cannot capture nonlinear relations between the variables.

We study the association of the variables by means of copulas using nonparametric concordance measures for the ranks of the variables, thus enabling coping with data not normally distributed and allowing nonlinear relations between the variables. A concordance measure can be defined as follows: let (y_{11}, y_{21}) and (y_{12}, y_{22}) be two observations of the bivariate random variable (Y_1, Y_2). Concordance exists when $(y_{11} - y_{21})(y_{12} - y_{22}) > 0$, while discordance exists when $(y_{11} - y_{21})(y_{12} - y_{22}) < 0$.

The Kendall (τ_k) and Spearman (ρ_s) correlations are concordance measures adopted in this methodology and their expressions derived from the copula dependence parameters are given below:

- Clayton copula: $\tau_k = \frac{\lambda}{\lambda+2}$. The expression for ρ_s is very complicated.

- Frank copula: $\tau_k = 1 + \frac{4}{\lambda}[D_1(\lambda) - 1]$ and $\rho_s = 1 + \frac{12}{\lambda}[D_2(\lambda) - D_1(\lambda)]$, where $D_k(\cdot)$ is the kth order Debye function $\frac{k}{\alpha^k}\int_0^\alpha \frac{t^k}{e^t - 1}dt$ (for $k = 1, 2$).

4. BCEOLL Normal and BFEOLL Normal Models

Here, we discuss special cases of the BCEOLL-G and BFEOLL-G models. The density functions (14) and (16) will be most tractable when the cdf $G_k(y_k; \boldsymbol{\eta}_k)$ has a simple analytic expression.

It is known that the data of many experiments follow a normal distribution. So, we present special models considering the normal baseline ($y_k \in \mathbb{R}$)

$$G_k(y_k; \boldsymbol{\eta}_k) = \Phi\left(\frac{y_k - \mu_k}{\sigma_k}\right) \qquad \text{and} \qquad g_k(y_k; \boldsymbol{\eta}_k) = \frac{1}{\sigma_k}\phi\left(\frac{y_k - \mu_k}{\sigma_k}\right), \tag{17}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the cdf and pdf of the standard normal, respectively, $\boldsymbol{\eta}_k = (\mu_k, \sigma_k)^\top$, $\mu_k \in \mathbb{R}$ is a location and $\sigma_k > 0$ is a scale (for $k = 1, 2$). Hereafter, let $Y_k \sim \text{EOLLN}(\nu_k, \tau_k, \mu_k, \sigma_k)$ be a random variable.

If the marginals follow the EOLLN distribution, the joint cdf is determined by inserting (17) in Equations (14) and (16). So, the joint pdfs of the BCEOLL Normal (BCEOLLN) and BFEOLL Normal (BFEOLLN) models are

$$
\begin{aligned}
f_{\text{BCEOLLN}}(y_1, y_2) &= (\lambda + 1)\left\{\left[\frac{\Phi^{\nu_1 \tau_1}(z_1)}{H^{\tau_1}(z_1)}\right]^{-\lambda} + \left[\frac{\Phi^{\nu_2 \tau_2}(z_2)}{H^{\tau_2}(z_2)}\right]^{-\lambda} - 1\right\}^{-\frac{(2\lambda+1)}{\lambda}} \left[\frac{\Phi^{\nu_1 \tau_1}(z_1)}{H^{\tau_1}(z_1)} \times \frac{\Phi^{\nu_2 \tau_2}(z_2)}{H^{\tau_2}(z_2)}\right]^{-(\lambda+1)} \times \\
&\quad \left\{\frac{\nu_1 \tau_1 \phi(z_1) \Phi^{\nu_1 \tau_1 - 1}(z_1)[1 - \Phi(z_1)]^{\nu_1 - 1}}{\sigma_1 H^{\tau_1}(z_1)}\right\}\left\{\frac{\nu_2 \tau_2 \phi(z_2) \Phi^{\nu_2 \tau_2 - 1}(z_2)[1 - \Phi(z_2)]^{\nu_2 - 1}}{\sigma_2 H^{\tau_2}(z_2)}\right\}
\end{aligned}
\tag{18}
$$

and

$$
\begin{aligned}
f_{\text{BFEOLLN}}(y_1, y_2) &= \lambda[1 - \exp(-\lambda)]\exp\left\{-\lambda\left[\frac{\Phi^{\nu_1 \tau_1}(z_1)}{H^{\tau_1}(z_1)} + \frac{\Phi^{\nu_2 \tau_2}(z_2)}{H^{\tau_2}(z_2)}\right]\right\} \times \\
&\quad \left\{\exp(-\lambda) - \exp\left[\frac{-\lambda \Phi^{\nu_1 \tau_1}(z_1)}{H^{\tau_1}(z_1)}\right] - \exp\left[\frac{-\lambda \Phi^{\nu_2 \tau_2}(z_2)}{H^{\tau_2}(z_2)}\right] + \right. \\
&\quad \left. \exp\left[-\lambda\left(\frac{\Phi^{\nu_1 \tau_1}(z_1)}{H^{\tau_1}(z_1)} + \frac{\Phi^{\nu_2 \tau_2}(z_2)}{H^{\tau_2}(z_2)}\right)\right]\right\}^{-2} \times \\
&\quad \left\{\frac{\nu_1 \tau_1 \phi(z_1) \Phi^{\nu_1 \tau_1 - 1}(z_1)[1 - \Phi(z_1)]^{\nu_1 - 1}}{\sigma_1 H^{\tau_1}(z_1)}\right\}\left\{\frac{\nu_2 \tau_2 \phi(z_2) \Phi^{\nu_2 \tau_2 - 1}(z_2)[1 - \Phi(z_2)]^{\nu_2 - 1}}{\sigma_2 H^{\tau_2}(z_2)}\right\},
\end{aligned}
$$

respectively, where (for $k = 1, 2$)

$$\Phi(z_k) = \Phi\left(\frac{y_k - \mu_k}{\sigma_k}\right), \quad \phi(z_k) = \phi\left(\frac{y_k - \mu_k}{\sigma_k}\right) \quad \text{and} \quad H(z_k) = \Phi^{\nu_k}(z_k) + [1 - \Phi(z_k)]^{\nu_k}. \tag{19}$$

Figures 2 and 3 report the joint densities, their contour plots, and bivariate cdfs for the Clayton and Frank copulas. The presence of bimodality is noted by the joint density. In the contour plots, we clearly find different association structures.
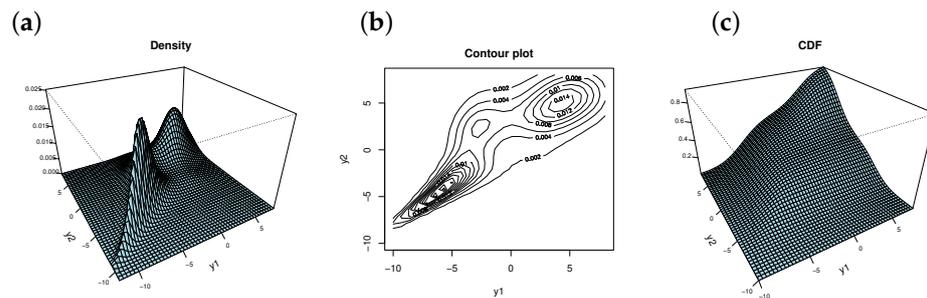
Figure 2. BCEOLLN copula for $Y_1 \sim \text{EOLLN}(\mu_1 = 0, \sigma_1 = 1, \nu_1 = 0.1, \tau_1 = 1)$, $Y_2 \sim \text{EOLLN}(\mu_2 = 0, \sigma_2 = 1, \nu_2 = 0.1, \tau_2 = 1.5)$ and $\lambda = 3$: (**a**) Bivariate pdf, (**b**) Contour plots of the pdf and (**c**) Bivariate cdf.
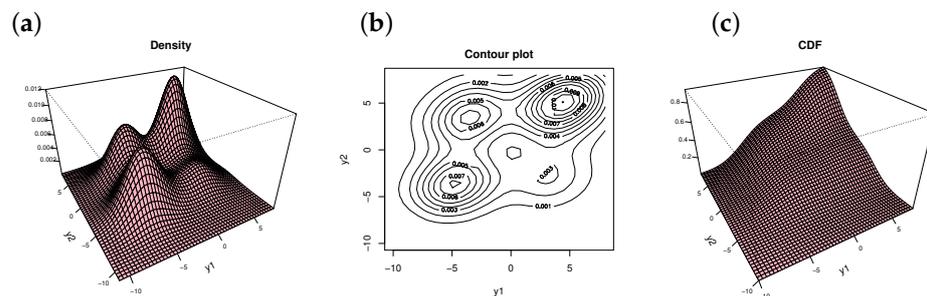

Figure 3. BFEOLLN copula for $Y_1 \sim \text{EOLLN}(\mu_1 = 0, \sigma_1 = 1, \nu_1 = 0.1, \tau_1 = 1)$, $Y_2 \sim \text{EOLLN}(\mu_2 = 0, \sigma_2 = 1, \nu_2 = 0.1, \tau_2 = 1.5)$ and $\lambda = 3$: (**a**) Bivariate pdf, (**b**) Contour plots of the pdf and (**c**) Bivariate cdf.

5. Bivariate Regression Models

Let $(Y_{11}, Y_{12}), \cdots, (Y_{n1}, Y_{n2})$ be a bivariate random sample from the BCEOLLN and BEOLLN models, $Y_{ik} \sim \text{BCEOLLN}(\theta_{i1})$ and $Y_{ik} \sim \text{BEOLLN}(\theta_{i2})$, where $\theta_{i1} = (\mu_{i1}, \sigma_1, \nu_1, \tau_1)^\top$ and $\theta_{i2} = (\mu_{i2}, \sigma_2, \nu_2, \tau_2)^\top$ (for $i = 1, \ldots, n$ and $k = 1, 2$).

Considering a sample $(y_{1k}, \mathbf{x}_{1k}), \cdots, (y_{nk}, \mathbf{x}_{nk})$, a systematic component can be defined as

$$\mu_{ik} = \mathbf{x}_{ik}^\top \boldsymbol{\beta}_k, \tag{20}$$

where $\mathbf{x}_{ik}^\top = (1, x_{i1k}, \cdots, x_{ipk})$ is the explanatory variable vector of dimension $p + 1$ (for $k = 1, 2$ and $i = 1, \ldots, n$), and $\boldsymbol{\beta}_k = (\beta_{0k}, \beta_{1k}, \cdots, \beta_{pk})^\top$ is the vector of unknown parameters.

Considering n independent observations $(y_{1k}, \mathbf{x}_{1k}), \cdots, (y_{nk}, \mathbf{x}_{nk})$ (for $k = 1, 2$), the model defined by (20) and the joint pdf given in Equations (18) and (19). Further, let $z_{ik} = (y_{ik} - \mu_{ik})/\sigma_k$, $H(z_{ik}) = \Phi^{\nu_k}(z_{ik}) + [1 - \Phi(z_{ik})]^{\nu_k}$, and $W(z_{ik}) = \Phi^{\nu_k \tau_k}(z_{ik})/H^{\nu_k}(z_{ik})$.

If $\boldsymbol{\theta} = (\lambda, \boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)^\top$, $\boldsymbol{\theta}_k = (\boldsymbol{\beta}_k^\top, \sigma_k, \tau_k, \nu_k)^\top$, $\boldsymbol{\beta}_k^\top = (\beta_{0k}, \beta_{1k}, \cdots, \beta_{pk})$ (for $k = 1, 2$), the total log-likelihood functions for $\boldsymbol{\theta}$ have the forms below:

BCEOLLN regression model

$$
\begin{aligned}
l(\boldsymbol{\theta}) = {} & n \log\left[\frac{(\lambda + 1)\,\nu_1\,\tau_1\,\nu_2\,\tau_2}{\sigma_1\,\sigma_2}\right] - \frac{2\lambda + 1}{\lambda} \sum_{i=1}^{n} \log\left[W^{-\lambda}(z_{i1}) + W^{-\lambda}(z_{i2}) - 1\right] - \\
& (\lambda + 1) \sum_{i=1}^{n} \log[W(z_{i1})W(z_{i2})] + \sum_{i=1}^{n} \log\left\{\frac{\phi(z_{i1})\Phi^{\nu_1\,\tau_1 - 1}(z_{i1})[1 - \Phi(z_{i1})]^{\nu - 1}}{H^{\tau_1}(z_{i1})}\right\} + \\
& \sum_{i=1}^{n} \log\left\{\frac{\phi(z_{i2})\Phi^{\nu_1\,\tau_2 - 1}(z_{i2})[1 - \Phi(z_{i2})]^{\nu - 1}}{H^{\tau_2}(z_{i2})}\right\}.
\end{aligned} \tag{21}
$$

BFEOLLN regression model

$$
\begin{aligned}
l(\boldsymbol{\theta}) \;=\; & n\log\left\{\frac{\lambda[1-\exp(-\lambda)]\,\nu_1\,\tau_1\,\nu_2\,\tau_2}{\sigma_1\,\sigma_2}\right\} - \lambda\sum_{i=1}^{n}[W(z_{i1})+W(z_{i2})] - \\
& 2\sum_{i=1}^{n}\log\{\exp(-\lambda)-\exp[-\lambda\,W(z_{i1})]-\exp[-\lambda\,W(z_{i2})]+\exp[-\lambda(W(z_{i1})+W(z_{i2}))]\} + \\
& \sum_{i=1}^{n}\log\left\{\frac{\phi(z_{i1})\Phi^{\nu_1\,\tau_1-1}(z_{i1})[1-\Phi(z_{i1})]^{\nu-1}}{H^{\tau_1}(z_{i1})}\right\} + \\
& \sum_{i=1}^{n}\log\left\{\frac{\phi(z_{i2})\Phi^{\nu_1\,\tau_2-1}(z_{i2})[1-\Phi(z_{i2})]^{\nu-1}}{H^{\tau_2}(z_{i2})}\right\}.
\end{aligned}
\tag{22}
$$

For copulas, the estimation of the parameters is usually conducted in two stages to create bivariate response models. In the first stage, the events are considered independent and the parameters are estimated marginally. The estimates are then used in the second stage, where the association parameter α is estimated. This approach is coherent when the focus of the study is to estimate the association parameter λ. If the coefficients of the regression are the focus, this marginal two-stage approach does not add any additional information compared to the use of independent models. Then, it is best to conduct a joint estimation of the regression coefficients and λ.

The maximum likelihood estimate (MLE) $\widehat{\boldsymbol{\theta}}$ can be calculated by maximizing (21) and (22). We use the simplex method of Nelder and Mead (1965) [29] implemented in the `optim` function in the R software [30]. This method is a robust and direct search method, which uses only function values, i.e., it does not use gradient information. It compares the function values at the vertices of a general simplex, then replaces the vertex with the highest value by another point. The simplex adapts itself to the local landscape, and contracts on to the final minimum. Initial values for $\boldsymbol{\beta}$, σ_1 and σ_2 are taken from the fits of the sub-models with $\tau_1=\nu_1=0.5$ and $\tau_2=\nu_2=0.5$. It proves to be effective, computationally compact and provides the Hessian matrix, necessary for inference. See [29] for details. The asymptotic normal distribution of $\hat{\boldsymbol{\theta}}$ can be considered for inference, tests, and confidence intervals.

6. Simulation Study

A Monte Carlo simulation study examines the accuracy of the estimates in the bivariate EOLLN regression model. We use the Multivariate Copula Description (MVCD) from the COPULA package [31] in R. The MVCD function generates the required univariate marginals according to the supplied association parameter using the inverse transformation method.

We consider the sample sizes $n=50,100,200,500$, $r=1000$ replications and a covariate $x_1\sim\mathrm{Normal}(1,0.5)$ related to y_1 and y_2 by the identity link functions $\mu_{i1}=(\beta_{101}+\beta_{111}x_{i1})$ and $\mu_{i2}=(\beta_{102}+\beta_{112}x_{i1})$, respectively. Further, we take $\sigma_k=\exp(\beta_{20k})$, $\nu_k=\exp(\beta_{30k})$ and $\tau_k=\exp(\beta_{40k})$. Then, let $Y_k\sim\mathrm{EOLLN}(\mu_{ik},\sigma_k,\nu_k,\tau_k)$ (for $k=1,2$), the quantile function (qf) has the form

$$
y_k = Q(u) = Q_N\left\{\frac{u^{1/(\nu_k\,\tau_k)}}{u^{1/(\nu_k\,\tau_k)}+(1-u^{1/\tau_k})^{1/\nu_k}}\right\},\quad 0<u<1,
\tag{23}
$$

where $Q_N(p)=\Phi^{-1}(p;\mu_k,\sigma_k)$, $p\in(0,1)$, is the normal qf, namely

$$
\Phi^{-1}(p;\mu_k,\sigma_k)=\mu_k+\sqrt{2}\,\sigma_k\,\mathrm{erf}^{-1}(2p-1)
\tag{24}
$$

and $\mathrm{erf}^{-1}(\cdot)$ is the inverse error function.

The true parameter values are: $\mu_{i1}=(1.5+1.32x_{i1})$, $\sigma_1=\exp(1.5)$, $\nu_1=\exp(1.1)$, $\tau_1=\exp(1.4)$, $\mu_{i2}=(1.6+2.3x_{i1})$, $\sigma_2=\exp(0.6)$, $\nu_2=\exp(1.8)$, $\tau_2=\exp(1.9)$ and $\lambda=5$.

The average estimates (AEs), biases, and mean squared errors (MSEs) are given by

$$\text{AE}(\hat{\pmb{\eta}}) = \frac{1}{r}\sum_{i=1}^{r}\hat{\eta}_i, \quad \text{Bias}(\hat{\pmb{\eta}}) = \frac{1}{r}\sum_{i=1}^{r}(\hat{\eta}_i - \eta_i), \quad \text{MSE}(\hat{\pmb{\eta}}) = \frac{1}{r}\sum_{i=1}^{r}(\hat{\eta}_i - \eta_i)^2, \quad (25)$$

where $\hat{\pmb{\eta}}^{\top} = (\hat{\beta}_{101}, \hat{\beta}_{111}, \hat{\beta}_{201}, \hat{\beta}_{301}, \hat{\beta}_{401}, \hat{\beta}_{102}, \hat{\beta}_{112}, \hat{\beta}_{202}, \hat{\beta}_{302}, \hat{\beta}_{402}, \hat{\lambda})$.

For each of the Clayton and Frank copulas and sample size, the calculations follow the Algorithm 1.

Algorithm 1:

Input: n: sample size
$\beta_{10k}, \beta_{11k}, \beta_{20k}, \beta_{30k}, \beta_{40k}, \lambda$: parameter initial values
$n.par$: number of parameters
r: number of replicates
$\eta = \text{matrix}(0, r, n.par)$
$\text{se} = \text{matrix}(0, r, n.par)$
$i = 1$
while $i \leq r$ **do**
    $x_{i1} \sim$ Normal $(n, 1, 0.5)$
    $\mu_{ik} = (\beta_{10k} + \beta_{11k}x_{i1})$
    $\sigma_{ik} = \exp(\beta_{20k})$
    $\nu_{ik} = \exp(\beta_{30k})$
    $\tau_{ik} = \exp(\beta_{40k})$
    Get the copulas using the function MVDC
    Get $y_{ik}(n, \mu_{ik}, \sigma_{ik}, \nu_{ik}, \tau_{ik})$ using function RMVDC
    Fit the model using the function OPTIM
    if *Model converges* **then**
        $\eta[i,] = $ Estimates
        $\text{sd}[i,] = $ Standard errors
        $i = i + 1$
    else
        $i = i$
    end
end
Calculate the quantities in Equation (25).

For both copulas, the biases and MSEs in Tables 1 and 2 decrease when n grows. So, the consistency of the estimators holds. The empirical coverage probabilities (CPs) in Table 3 show that their values converge to the 95% nominal level.

Table 1. Simulation findings from the fitted bivariate Clayton EOLLN regression model.

η	True Value	$n = 50$			$n = 100$		
		AEs	Biases	MSEs	AEs	Biases	MSEs
β_{101}	1.50	1.4840	-0.0160	0.1436	1.4949	-0.0051	0.0787
β_{111}	1.32	1.3206	0.0006	0.0523	1.3190	-0.0010	0.0244
β_{201}	1.50	1.5448	0.0448	0.0755	1.5383	0.0383	0.0405
β_{301}	1.10	1.1799	0.0799	0.0816	1.1577	0.0577	0.0450
β_{401}	1.40	1.4555	0.0555	0.0892	1.4233	0.0233	0.0453
β_{102}	1.60	1.5966	-0.0034	0.0056	1.5986	-0.0014	0.0024
β_{112}	2.30	2.3003	0.0003	0.0017	2.2991	-0.0009	0.0008
β_{202}	0.60	0.6135	0.0135	0.0484	0.6132	0.0132	0.0224
β_{302}	1.80	1.8481	0.0481	0.0487	1.8314	0.0314	0.0216
β_{402}	1.90	1.9731	0.0731	0.0956	1.9396	0.0396	0.0408
λ	5.00	5.2268	0.2268	0.1417	5.1446	0.1446	0.0601

Table 1. *Cont.*

η	True Value	$n = 200$			$n = 500$		
		AEs	Biases	MSEs	AEs	Biases	MSEs
β_{101}	1.50	1.4949	−0.0051	0.0437	1.5049	0.0049	0.0164
β_{111}	1.32	1.3172	−0.0028	0.0122	1.3219	0.0019	0.0049
β_{201}	1.50	1.5336	0.0336	0.0247	1.5185	0.0185	0.0108
β_{301}	1.10	1.1440	0.0440	0.0272	1.1253	0.0253	0.0112
β_{401}	1.40	1.4147	0.0147	0.0214	1.3992	−0.0008	0.0074
β_{102}	1.60	1.5985	−0.0015	0.0010	1.5993	−0.0007	0.0003
β_{112}	2.30	2.2996	−0.0004	0.0004	2.3003	0.0003	0.0002
β_{202}	0.60	0.6149	0.0149	0.0100	0.6089	0.0089	0.0048
β_{302}	1.80	1.8244	0.0244	0.0093	1.8135	0.0135	0.0045
β_{402}	1.90	1.9215	0.0215	0.0174	1.9078	0.0078	0.0057
λ	5.00	5.0736	0.0736	0.0223	5.0178	0.0178	0.0060

Table 2. Simulation findings from the fitted bivariate Frank EOLLN regression model.

η	True Value	$n = 50$			$n = 100$		
		AEs	Biases	MSEs	AEs	Biases	MSEs
β_{101}	1.50	1.3278	−0.1722	0.5729	1.3552	−0.1448	0.3244
β_{111}	1.32	1.3110	−0.0090	0.0905	1.3250	0.0050	0.0432
β_{201}	1.50	1.5802	0.0802	0.0891	1.5753	0.0753	0.0686
β_{301}	1.10	1.2140	0.1140	0.1023	1.1875	0.0875	0.0734
β_{401}	1.40	1.6102	0.2102	0.4140	1.5419	0.1419	0.2325
β_{102}	1.60	1.5604	−0.0396	0.0195	1.5730	−0.0270	0.0114
β_{112}	2.30	2.2977	−0.0023	0.0037	2.3015	0.0015	0.0017
β_{202}	0.60	0.6307	0.0307	0.0566	0.6419	0.0419	0.0446
β_{302}	1.80	1.8510	0.0510	0.0667	1.8546	0.0546	0.0483
β_{402}	1.90	2.1430	0.2430	0.3737	2.0495	0.1495	0.2159
λ	5.00	5.1239	0.1239	0.3300	5.0954	0.0954	0.2026
η	True Value	$n = 200$			$n = 500$		
		AEs	Biases	MSEs	AEs	Biases	MSEs
β_{101}	1.50	1.4714	−0.0286	0.1427	1.4853	−0.0147	0.0666
β_{111}	1.32	1.3145	−0.0055	0.0225	1.3264	0.0064	0.0086
β_{201}	1.50	1.5681	0.0681	0.0390	1.5407	0.0407	0.0189
β_{301}	1.10	1.1789	0.0789	0.0414	1.1461	0.0461	0.0210
β_{401}	1.40	1.4391	0.0391	0.0999	1.4102	0.0102	0.0421
β_{102}	1.60	1.5801	−0.0199	0.0052	1.5862	−0.0138	0.0024
β_{112}	2.30	2.2990	−0.0010	0.0008	2.3008	0.0008	0.0003
β_{202}	0.60	0.6271	0.0271	0.0217	0.6278	0.0278	0.0100
β_{302}	1.80	1.8334	0.0334	0.0225	1.8289	0.0289	0.0094
β_{402}	1.90	2.0108	0.1108	0.0961	1.9652	0.0652	0.0465
λ	5.00	5.0431	0.0431	0.0856	5.0227	0.0227	0.0428

Table 3. CPs for the fitted bivariate Clayton and Frank EOLLN regression models.

Clayton										
n	β_{101}	β_{111}	β_{201}	β_{301}	β_{401}	β_{102}	β_{112}	β_{202}	β_{302}	β_{402}	λ
50	0.995	0.971	0.999	0.996	0.997	0.999	0.972	1.000	1.000	0.999	0.961
100	0.996	0.972	0.998	0.997	0.998	1.000	0.974	1.000	1.000	1.000	0.962
200	0.999	0.958	0.999	0.999	1.000	0.999	0.965	1.000	0.999	1.000	0.968
500	0.999	0.962	0.999	0.999	1.000	1.000	0.964	1.000	1.000	1.000	0.948
Frank										
n	β_{101}	β_{111}	β_{201}	β_{301}	β_{401}	β_{102}	β_{112}	β_{202}	β_{302}	β_{402}	λ
50	0.983	0.956	1.000	0.998	0.986	0.988	0.937	1.000	1.000	0.988	0.950
100	0.993	0.961	0.999	1.000	0.992	0.989	0.944	1.000	0.999	0.990	0.951
200	0.990	0.956	0.999	0.998	0.991	0.996	0.956	1.000	1.000	0.997	0.962
500	0.995	0.955	0.999	0.999	0.995	0.998	0.946	1.000	1.000	0.997	0.953

7. Application to Lettuce Leaf Data

This dataset refers to the effect of the foliar application of different concentrations of silicon dioxide and organosilicon compounds on the growth and biochemical contents of Lactuca sativa var. crispa grown in phytotron conditions. The experiment was carried out on 12 March 2020, in the Department of Goods Commodity and Expertise of Goods Products in Moscow, Russian Federation. More details of the experiment are described in [32]. Here, we check the response variables fresh weight (grams) (y_1) and plant height (cm) (y_2) together in order to verify the existing correlation between them. We also present a regression model relating these covariates with the following treatments:

The levels of gradual concentrations of amorphous silicon dioxide (AS) and 1-ethoxysilatrane (ES) solutions are: AS1 (50 mg·L^{-1}), AS2 (100 mg·L^{-1}), AS3 (200 mg·L^{-1}), AS4 (300 mg·L^{-1}), and ES1 (5×10^{-4} mL·L^{-1}), ES2 (10^{-3} mL·L^{-1}), ES3 (5×10^{-3} mL·L^{-1}) and ES4 (10^{-2} mL·L^{-1}). These variables are defined by dummy variables as follows:

- y_{i1} fresh weight (grams);
- y_{i2} plant height (cm);
- x_{ijk} combination of AS and ES solutions (for $j = 1 \ldots 8$).

Hence, the systematic component has the form (for $k = 1, 2$ and $i = 1, \ldots, 54$):

$$\mu_{ik} = \beta_{0k} + \beta_{1k}x_{i1k} + \beta_{2k}x_{i2k} + \beta_{3k}x_{i3k} + \beta_{4k}x_{i4k} + \beta_{5k}x_{i5k} + \beta_{6k}x_{i6k} + \beta_{7k}x_{i7k} + \beta_{8k}x_{i8k}.$$

7.1. Descriptive Analysis

First, we present a descriptive analysis of the dataset. Table 4 indicates positive asymmetry and kurtosis for both response variables. Figure 1a,b displays the histograms of these variables, thus indicating bimodal behavior for both. Pearson's coefficient is 0.8497, which is a strong positive correlation, as can be noted in Figure 1. Figure 4 provides the boxplots for different experimental treatments.

Table 4. Descriptive analysis of fresh weight (y_1) and plant height (y_2) variables.

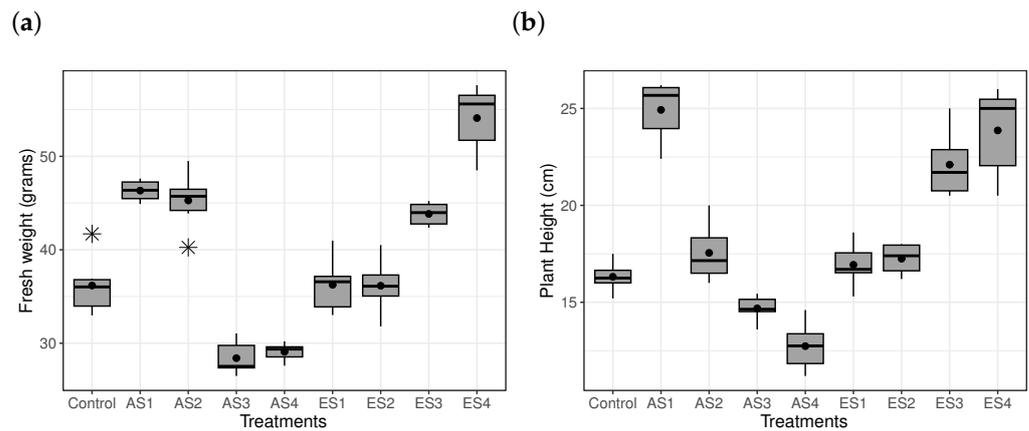| Variable	Mean	Median	s.d.	Min.	Max	Skewness	VC	Kurtosis
y_1	39.511	38.925	8.406	26.500	57.610	0.268	21.274	2.208
y_2	18.485	17.250	4.197	11.200	26.200	0.476	22.707	2.174

Figure 4. (**a**) Boxplots of fresh weight (y_1) and (**b**) Boxplots of plant height (y_2).

7.2. Univarite Marginal Analysis

Univariate analysis of the response variables is conducted under the EOLLN distribution and its special cases: OLLN, Exp-N, and Normal. The Akaike information criterion (AIC) and Global Deviance (GD) in Table 5 reveal that the OLLN model is more adequate for y_1 (fresh weight) and the EOLLN model for y_2 (plant height).

Figures 5a and 6a report the histograms with estimated densities, and Figures 5b and 6b their empirical and estimated cumulative distributions, thus supporting the findings in Table 5.

Table 5. Adequacy measures of the univariate models for each response variable.

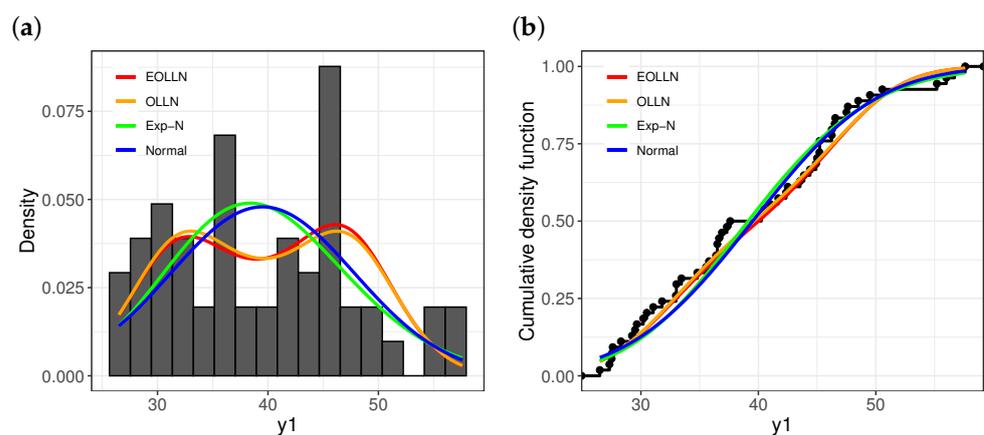| Model	Fresh Weight		Plant Height	
	AIC	**GD**	**AIC**	**GD**
EOLLN	385.98	377.98	**300.80**	**292.80**
OLLN	**384.05**	**378.05**	305.51	299.51
Exp-N	387.40	381.40	309.57	303.57
Normal	386.16	382.16	311.16	307.16

Figure 5. (**a**) Estimated densities and (**b**) Estimated cumulative functions and empirical cdf for the fresh weight.
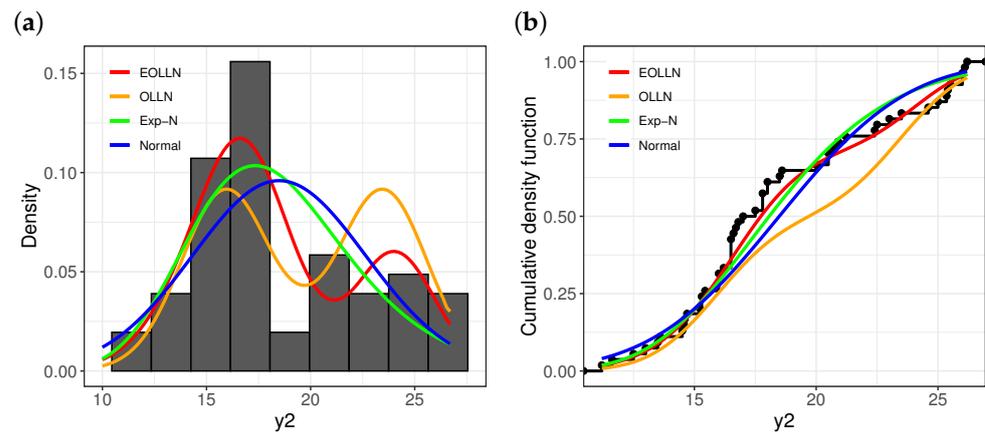
(a)

(b)

Figure 6. (**a**) Estimated densities and (**b**) Estimated cumulative functions and empirical cdf for the plant height.

7.3. Bivariate Analysis and Regression Model

Next, we carry out a joint analysis of the response variables from the current sixteen models for the Clayton and Frank copulas. Table 6 reveals that the bivariate (OLLN × EOLLN) regression model with Clayton and Frank copulas is the most suitable model to explain the response variables ($Y_1 \times Y_2$), thus agreeing with the univariate analysis.

Additionally, we provide the values of the copula dependence parameter (λ), Kendall's correlation (τ_k), and Spearman correlation (ρ_s). For both models and the two copulas, these values indicate that the variables have strong positive dependence, thus confirming the required dependence structure. Table 7 provides the estimated quantities for the best bivariate regression model.

We can conclude the following facts:

- **Interpretations for μ:**
  - Comparison of the treatments with the control indicates that treatments ES1 and ES2 are not significant for the metrics of fresh mass and plant height. The other treatments are significant for both variables.
  - Comparison of the treatments with the control indicates that treatments AS1, AS2, ES3, and ES4 have positive effects on the fresh mass. So, they increase with the fresh mass. On the other hand, the effects of treatments AS3 and AS4 are negative for the fresh mass.
  - For the plant height, the treatments AS1, AS2, ES3, and ES4 have positive effects, meaning greater height in relation to the control. In contrast, the treatments AS3 and AS4 have negative effects on plant height.
  - Table 8 compares all treatments with the corresponding control, from which other interpretations can be found.
  - All the results are clearly consistent with the descriptive analysis.

- **Interpretations for λ, τ_k and ρ_s:**
  - The values of λ, τ_k, and ρ_s in the best regression model reveal that the variables are correlated with moderate dependence. The dependence structure is necessary, as confirmed by the 95% CIs for the dependence parameter of the copula, which does not include zero.

Finally, Figure 7 displays the plots of the observed and predicted values and corresponding CIs, thus supporting that the best bivariate regression model is a good predictor for the current data.

Table 6. AIC, GD, copula dependence parameter (λ), Kendall correlation (τ_k) and Spearman correlation (ρ_s) for sixteen bivariate models fitted to lettuce data.

Model for ($Y_1 \times Y_2$)	Clayton					Frank				
	AIC	**GD**	λ	τ_k	ρ_s	**AIC**	**GD**	λ	τ_k	ρ_s
EOLLN × EOLLN	623.18	605.18	2.81	0.58	0.77	615.95	597.95	10.16	0.67	0.86
EOLLN × OLLN	634.02	618.02	2.27	0.53	0.72	627.92	611.92	9.67	0.66	0.85
EOLLN × Exp-N	629.52	613.52	3.38	0.63	0.81	628.37	612.37	8.60	0.62	0.82
EOLLN × Normal	625.02	611.02	3.58	0.64	0.83	630.43	616.43	10.23	0.67	0.87
OLLN × EOLLN	**620.25**	**604.25**	**2.93**	**0.59**	**0.78**	**612.78**	**596.78**	**10.79**	**0.69**	**0.88**
OLLN × OLLN	632.00	618.00	2.31	0.54	0.72	626.71	612.71	9.47	0.65	0.85
OLLN × Exp-N	628.19	614.19	3.36	0.63	0.81	623.88	609.88	10.25	0.67	0.87
OLLN × Normal	624.47	612.47	3.82	0.66	0.84	628.43	616.43	10.36	0.68	0.87
Exp-N × EOLLN	622.61	606.61	2.92	0.59	0.78	616.98	600.98	10.03	0.67	0.86
Exp-N × OLLN	631.90	617.90	2.87	0.59	0.78	632.15	618.15	9.11	0.64	0.84
Exp-N × Exp-N	632.43	618.43	3.47	0.63	0.82	636.22	622.22	10.85	0.69	0.88
Exp-N × Normal	626.75	614.75	3.33	0.62	0.81	632.04	620.04	9.80	0.66	0.86
Normal × EOLLN	620.62	606.62	3.00	0.60	0.79	615.77	601.77	10.30	0.67	0.87
Normal × OLLN	633.50	621.50	2.29	0.53	0.72	630.58	618.58	9.15	0.64	0.84
Normal × Exp-N	629.29	617.29	3.05	0.60	0.79	624.74	612.74	10.16	0.67	0.86
Normal × Normal	624.54	614.54	3.35	0.63	0.81	630.37	620.37	9.99	0.67	0.86

Table 7. Estimated quantities for the best bivariate regression model fitted to lettuce data.

θ	MLEs	CIs	SEs	*p*-Values
β_{01}	35.8036	[33.7814, 37.8257]	1.0086	<0.001
β_{11}	10.8022	[8.1460, 13.4583]	1.3248	<0.001
β_{21}	9.9009	[7.1183, 12.6833]	1.3879	<0.001
β_{31}	−7.6892	[−10.3286, −5.0497]	1.3165	<0.001
β_{41}	−6.7391	[−9.2985, −4.1795]	1.2766	<0.001
β_{51}	0.6059	[−2.1540, 3.3658]	1.3766	0.3308
β_{61}	−0.1776	[−3.0838, 2.7287]	1.4496	0.4515
β_{71}	8.4060	[5.3443, 11.4675]	1.5271	<0.001
β_{81}	18.6657	[15.5123, 21.8191]	1.5729	<0.001
$\log(\sigma_1)$	1.8421	[−4.0207, 7.7049]	2.9243	
$\log(\nu_1)$	1.0892	[−4.9764, 7.1548]	3.0255	
β_{02}	17.0122	[15.8189, 18.2056]	0.5952	<0.001
β_{12}	9.4337	[8.1786, 10.6887]	0.6259	<0.001
β_{22}	1.6790	[0.0778, 3.2802]	0.7986	0.0201
β_{32}	−1.6034	[−2.7452, −0.4617]	0.5694	0.0034
β_{42}	−3.2964	[−4.6713, −1.9214]	0.6857	<0.001
β_{52}	0.9021	[−0.4202, 2.2245]	0.6595	0.0885
β_{62}	1.0147	[−0.2440, 2.2736]	0.6279	0.0559
β_{72}	6.3580	[4.5401, 8.1759]	0.9067	<0.001
β_{82}	8.3849	[7.0828, 9.6870]	0.6494	<0.001
$\log(\sigma_2)$	0.7811	[−1.6658, 3.2282]	1.2205	
$\log(\nu_2)$	1.1064	[−1.0946, 3.3074]	1.0978	
$\log(\tau)$	−1.0721	[−2.6434, 0.4992]	0.7837	
λ	3.3776	[1.1464, 5.6086]	1.1129	
τ_k=0.339, ρ_s= 0.492			AIC: 456.7843 GD: 408.7843	

Table 8. Comparisons between treatments according to the best bivariate regression model.

| Hypotheses H_0	Fresh Weight			Plant Height		
	MLEs	**SEs**	*p*-**Values**	**MLEs**	**SEs**	*p*-**Values**
AS1 - Control	10.802	1.325	0.000	9.434	0.626	0.000
AS2 - Control	9.901	1.388	0.000	1.679	0.799	0.020
AS3 - Control	−7.689	1.317	0.000	−1.603	0.569	0.003
AS4 - Control	−6.739	1.277	0.000	−3.296	0.686	0.000
ES1 - Control	0.606	1.377	0.331	0.902	0.660	0.088
ES2 - Control	−0.178	1.450	0.452	1.015	0.628	0.056
ES3 - Control	8.406	1.527	0.000	6.358	0.907	0.000
ES4 - Control	18.666	1.573	0.000	8.385	0.649	0.000
AS1 - ES4	−8.160	2.453	0.001	0.959	0.745	0.102
AS2 - ES4	−9.017	2.097	0.000	−6.829	0.902	0.000
AS3 - ES4	−26.352	2.667	0.000	−9.975	0.754	0.000
AS4 - ES4	−25.410	2.495	0.000	−11.658	0.855	0.000
ES1 - ES4	−18.376	2.642	0.000	−7.505	0.843	0.000
ES2 - ES4	−18.741	2.638	0.000	−7.312	0.810	0.000
ES3 - ES4	−10.192	2.853	0.000	−1.956	1.061	0.035
AS1 - ES3	2.663	1.581	0.049	3.120	1.159	0.005
AS2 - ES3	1.935	1.436	0.092	−4.597	0.859	0.000
AS3 - ES3	−15.784	1.599	0.000	−7.802	1.088	0.000
AS4 - ES3	−14.686	1.417	0.000	−9.550	0.922	0.000
ES1 - ES3	−7.468	1.504	0.000	−5.323	0.892	0.000
ES2 - ES3	−8.035	1.667	0.000	−5.143	1.033	0.000
AS1 - ES2	10.396	1.446	0.000	8.377	0.678	0.000
AS2 - ES2	9.707	1.464	0.000	0.689	0.841	0.208
AS3 - ES2	−7.959	1.399	0.000	−2.592	0.636	0.000
AS4 - ES2	−6.875	1.367	0.000	−4.295	0.713	0.000
ES1 - ES2	0.255	1.492	0.432	−0.120	0.701	0.432
AS1 - ES1	10.414	1.358	0.000	8.562	0.788	0.000
AS2 - ES1	9.640	1.339	0.000	0.918	0.764	0.117
AS3 - ES1	−7.931	1.363	0.000	−2.409	0.718	0.001
AS4 - ES1	−7.025	1.243	0.000	−4.117	0.694	0.000
AS1 - AS4	17.472	1.218	0.000	12.768	0.794	0.000
AS2 - AS4	16.701	1.129	0.000	4.995	0.781	0.000
AS3 - AS4	−1.036	1.191	0.194	1.715	0.715	0.010
AS1 - AS3	18.338	1.252	0.000	10.951	0.591	0.000
AS2 - AS3	17.712	1.303	0.000	3.355	0.875	0.000
AS1 - AS2	1.090	1.315	0.205	7.782	0.977	0.000

(**a**)   (**b**)
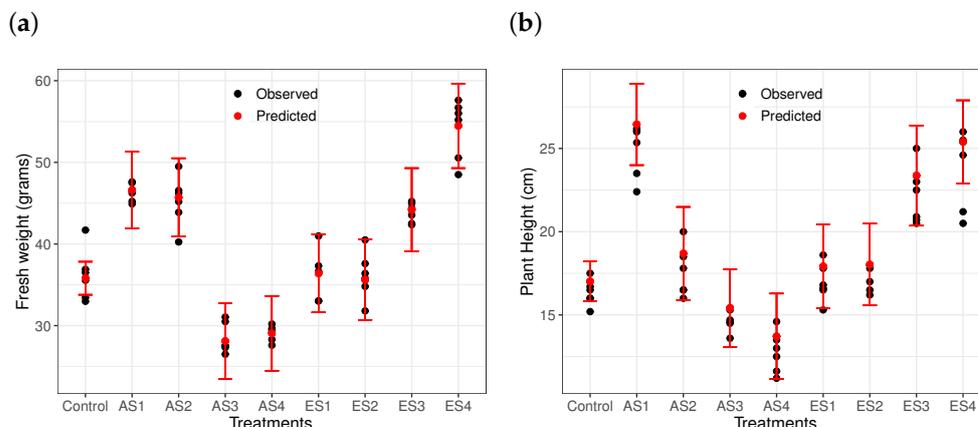
Figure 7. (**a**) Observed values (black) and predicted values (red) with CIs of the Clayton OLLN EOLLN regression model for fresh weight and (**b**) Observed values (black) and predicted values (red) with CIs of the Frank OLLN EOLLN regression for plant height.

8. Conclusions

This article proposed a new bivariate family based on Archimedean copulas. Motivated by an experiment that evaluated the growth of oak lettuce plants, which presented the variables fresh mass and height with bimodal behavior, we used the exponentiated odd log-logistic family, whose densities can model different types of data, including bimodality. Furthermore, these variables showed a strong positive correlation and the Clayton and Frank copulas were suitable for studying data with a positive correlation.

We presented some mathematical properties of the new family and a simulation study showed the consistency of the maximum likelihood estimators. We considered two categorical explanatory variables (each one with four treatments) to explain two response variables: fresh weight and plant height. Some important results were obtained from the application: (i) For both variables, the same treatments were significant or not; (ii) Two treatments had negative effects on the two variables and two other treatments had the greatest effects on the variables, i.e., with these treatments greater fresh mass and greater plant height were obtained. The bivariate regression model proved to be adequate for predicting fresh weight and plant height values of oak lettuce under the effect of different treatments. The choice of these copulas proved to be adequate due to the positive dependence structure between the variables. Other baseline distributions can be used as well as applications in other areas of knowledge.

Author Contributions: Conceptualization, G.M.R., E.M.M.O., G.M.C. and R.V.; methodology, G.M.R., E.M.M.O., G.M.C. and R.V.; software, G.M.R., E.M.M.O., G.M.C. and R.V.; validation, G.M.R., E.M.M.O., G.M.C. and R.V.; formal analysis, G.M.R., E.M.M.O., G.M.C. and R.V.; investigation, G.M.R., E.M.M.O., G.M.C. and R.V.; data curation, G.M.R., E.M.M.O., G.M.C. and R.V.; writing—original draft preparation, G.M.R., E.M.M.O., G.M.C. and R.V.; writing—review and editing, G.M.R., E.M.M.O., G.M.C. and R.V.; visualization, G.M.R., E.M.M.O., G.M.C. and R.V.; supervision, G.M.R., E.M.M.O., G.M.C. and R.V. All authors have read and agreed to the current version of the manuscript.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data Availability at https://www.sciencedirect.com/science/article/pii/S2352340921006120 (accessed on 10 July 2023).

Appendix A. Some Properties Related to the EOLL-G Model

Appendix A.1. Real Moments

Let $Y \sim \text{EOLL-G}(\nu, \tau, \boldsymbol{\eta})$. We obtain sufficient conditions that guarantee the existence of the real moments of Y. By using the well-known formula (for the moments of positive random variables)

$$\mathbb{E}(W^p) = p \int_0^\infty w^{p-1} \mathbb{P}(W > w)dw, \quad W > 0, \, p > 0, \tag{A1}$$

it is clear that $\mathbb{E}(Y^p) < \infty$ if and only if

$$I \equiv \int_c^\infty y^{p-1} \mathbb{P}(Y > y)dy < \infty,$$

for some $c \in (0, \infty)$ so that $0 < G(c; \boldsymbol{\eta}) < \infty$. We now verify the finiteness of I. In fact, let S be a Dagum random variable with shapes ν and τ, and unit scale, say $S \sim D(\nu, 1, \tau)$. Since

$$\mathbb{P}(Y \leqslant y) \stackrel{(8)}{=} \mathbb{P}\left(S \leqslant \frac{G(y; \boldsymbol{\eta})}{1 - G(y; \boldsymbol{\eta})} \right), \tag{A2}$$

the integral I becomes

$$I = \int_c^\infty y^{p-1} \, \mathbb{P}\left(S > \frac{G(y;\boldsymbol{\eta})}{1 - G(y;\boldsymbol{\eta})}\right) dy.$$

By applying Markov's inequality, the inequality holds

$$I \leqslant \mathbb{E}(S) \int_c^\infty y^{p-1} \frac{1 - G(y;\boldsymbol{\eta})}{G(y;\boldsymbol{\eta})} dy. \tag{A3}$$

If $y \longmapsto G(y;\boldsymbol{\eta})$ is a cdf, $G(y;\boldsymbol{\eta}) > G(c;\boldsymbol{\eta})$ for $y > c$. Let X have cdf $G(\cdot;\boldsymbol{\eta})$. So, the integral in (A3) can be expressed as

$$\frac{\mathbb{E}(S)}{G(c;\boldsymbol{\eta})} \int_c^\infty y^{p-1}[1 - G(y;\boldsymbol{\eta})]dy \leqslant \frac{\mathbb{E}(S)}{G(c;\boldsymbol{\eta})} \int_0^\infty y^{p-1}[1 - G(y;\boldsymbol{\eta})]dy = \frac{\mathbb{E}(S)\mathbb{E}(X^p)}{pG(c;\boldsymbol{\eta})},$$

where it is used Equation (A1). We obtain

$$I \leqslant \frac{\mathbb{E}(S)\mathbb{E}(X^p)}{pG(c;\boldsymbol{\eta})}.$$

As $\mathbb{E}(S) < \infty$ for $\nu > 1$ leads to the condition $\mathbb{E}(X^p) < \infty$, we have $I < \infty$.

Hence, under conditions $\nu > 1$ and $\mathbb{E}(X^p) < \infty$, we have $\mathbb{E}(Y^p) < \infty$.

Appendix A.2. Stochastic Representation

We can write from (A2),

$$\mathbb{P}(Y \leqslant y) = \mathbb{P}\left(G^{-1}\left(\frac{S}{S+1};\boldsymbol{\eta}\right) \leqslant y\right), \quad \forall y.$$

Finally, it is evident that

$$Y = G^{-1}\left(\frac{S}{S+1};\boldsymbol{\eta}\right)$$

provides a stochastic representation for $Y \sim \text{EOLL-G}(\nu, \tau, \boldsymbol{\eta})$.

Appendix A.3. Standardized Moments

Under the previous conditions, Appendix A.1 guarantees the existence of moments of positive order of Y, and then the existence of mean and variance. Let $\mu_Y = \mathbb{E}(Y)$ and $\sigma_Y^2 = \text{Var}(Y)$, and let $Z = (Y - \mu_Y)/\sigma_Y$ be the standardized version of Y.

By using the triangle inequality and the C_p inequality $(x + y)^p \leqslant C_p\,(x^p + y^p), \forall x, y \geqslant 0$, where $p > 0$ and $C_p \equiv \max\{1, 2^{p-1}\}$, we have

$$\mathbb{E}(|Z|^p) \leqslant \frac{C_p}{\sigma_Y^p}\,[\,\mathbb{E}(|Y|^p) + |\mu_Y|^p\,] = \frac{C_p}{\sigma_Y^p}\,[\,\mathbb{E}(Y^p) + \mu_Y^p\,],$$

since $Y > 0$. As the function $x \longmapsto \varphi(x) = x^{1/p}$ is increasing for all $x > 0$ and $p > 0$, we have from the previous inequality

$$\|Z\|_p = \varphi(\mathbb{E}(|Z|^p)) \leqslant \varphi\left(\frac{C_p}{\sigma_Y^p}\,[\,\mathbb{E}(Y^p) + \mu_Y^p\,]\right) = \frac{C_p^{1/p}}{\sigma_Y}\,[\,\mathbb{E}(Y^p) + \mu_Y^p\,]^{1/p} \leqslant \frac{C_p^{1/p} C_{1/p}}{\sigma_Y}\,(\|Y\|_p + \mu_Y),$$

where in the last inequality again we use the C_p inequality, and $\|Z\|_p$ and $\|Y\|_p$ are the norms in L^p of Z and Y, respectively. Appendix A.1 leads to $\|Y\|_p < \infty$, $\mu_Y < \infty$ and $0 < \sigma_Y < \infty$, and then $\|Z\|_p < \infty$ from the above inequality.

Appendix A.4. The Multivariate Distance-Gini Mean Difference

Let $Y'_1, \ldots, Y'_d$ be an independent copy of $Y_1, \ldots, Y_d$, where $\boldsymbol{Y} = (Y_1, \ldots, Y_d)$ follows the joint cdf $F = F_{\boldsymbol{Y}}$ as given in (1), and $Y_i \sim \text{EOLL-G}(\nu_i, \tau_i, \boldsymbol{\eta}_i)$, $i = 1, \ldots, d$. The distance-Gini mean difference for F is defined as in (10) (see Koshevoy 1997 [26]).

Since $\|\boldsymbol{y}\| \leqslant \|\boldsymbol{y}\|_1 = \sum_{i=1}^d |y_i|$, where $\|\cdot\|_1$ is the Manhattan norm, it is clear that

$$M_D(F) \leqslant \frac{1}{2d} \sum_{i=1}^d \mathbb{E}(|Y_i - Y'_i|) = \frac{1}{2d} \sum_{i=1}^d \mathbb{E}(\max\{Y_i, Y'_i\} - \min\{Y_i, Y'_i\}). \qquad \text{(A4)}$$

As $Y'_1, \ldots, Y'_d$ is an independent copy of $Y_1, \ldots, Y_d$, we have that Y'_i is idenpendent of Y_i and both have the same univariate distribution EOLL-G$(\nu, \tau, \boldsymbol{\eta})$, with variance denoted by σ^2. By using the following inequality (see Item (5) of reference Vila et al., 2023 [33]), if Z is the standardized version of Y_1, then

$$\mathbb{E}(\max\{Y_i, Y'_i\} - \min\{Y_i, Y'_i\}) \leqslant 2\,\sigma\,\|Z\|_p \left(\frac{p-1}{2p-1}\right)^{(p-1)/p}, \quad p > 1, \forall i = 1, \ldots, d,$$

we have from (A4)

$$M_D(F) \leqslant \sigma \|Z\|_p \left(\frac{p-1}{2p-1}\right)^{(p-1)/p}, \quad p > 1.$$

By Appendices A.1 and A.3, we have $0 < \sigma < \infty$ and $\|Z\|_p < \infty$ if $\nu > 1$ and $\mathbb{E}(X^p) < \infty$.

Hence, under conditions $\nu > 1$, $p > 1$ and $\mathbb{E}(X^p) < \infty$, we obtain $M_D(F) < \infty$.

References

1. Nelsen, R.B. Properties of a one-parameter family of bivariate distributions with specified marginals. *Commun. Stat.-Theory Methods* **1986**, *15*, 3277–3285. [CrossRef]
2. Quiroz-Flores, A. Testing copula functions as a method to derive bivariate Weibull distributions. In Proceedings of the APSA Annual Meeting & Exhibition, Toronto, ON, Canada, 3–6 September 2009.
3. El-Sherpieny, E.S.; Almetwally, E.M. Bivariate generalized rayleigh distribution based on Clayton Copula. In Proceedings of the 54rd Annual Conference on Statistics, Computer Science and Operation Research, Cairo, Egypt, 9–11 December 2019; pp. 1–19.
4. Almetwally, E.M.; Muhammed, H.Z. On a Bivariate Frechet Distribution. *J. Stat. Appl. Probab.* **2020**, *9*, 1–21.
5. Muhammed, H.Z. On a bivariate generalized inverted Kumaraswamy distribution. *Phys. A Stat. Mech. Its Appl.* **2020**, *553*, 124281. [CrossRef]
6. Samanthi, R.G.M.; Sepanski, J. On bivariate Kumaraswamy-distorted copulas. *Commun. Stat.-Theory Methods* **2022**, *51*, 2477–2495. [CrossRef]
7. Alotaibi, R.M.; Rezk, H.R.; Ghosh, I.; Dey, S. Bivariate exponentiated half logistic distribution: Properties and application. *Commun. Stat.-Theory Methods* **2021**, *50*, 6099–6121. [CrossRef]
8. Vaidyanathan, V.S.; Sharon Varghese, A. Morgenstern type bivariate Lindley distribution. *Stat. Optim. Inf. Comput.* **2016**, *4*, 132–146. [CrossRef]
9. Naifar, N. Modelling dependence structure with Archimedean copulas and applications to the iTraxx CDS index. *J. Comput. Appl. Math.* **2011**, *235*, 2459–2466. [CrossRef]
10. Yang, L.; Cai, X.J.; Li, M.; Hamori, S. Modeling dependence structures among international stock markets: Evidence from hierarchical Archimedean copulas. *Econ. Model.* **2015**, *51*, 308–314. [CrossRef]
11. Novianti, P.; Kartiko, S.H.; Rosadi, D. Application of Clayton Copula to identify dependency structure of COVID-19 outbreak and average temperature in Jakarta Indonesia. *J. Phys. Conf. Ser.* **2021**, *1*, 012154. [CrossRef]
12. Li, H.; Lu, Y. Modeling cause-of-death mortality using hierarchical Archimedean copula. *Scand. Actuar. J.* **2019**, *3*, 247–272. [CrossRef]
13. Janga Reddy, M.; Ganguli, P. Risk assessment of hydroclimatic variability on groundwater levels in the Manjara basin aquifer in India using Archimedean copulas. *J. Hydrol. Eng.* **2012**, *17*, 1345–1357. [CrossRef]
14. Zhang, L.; Singh, V.P. Bivariate rainfall frequency distributions using Archimedean copulas. *J. Hydrol.* **2007**, *332*, 93–109. [CrossRef]
15. Tsakiris, G.; Kordalis, N.; Tigkas, D.; Tsakiris, V.; Vangelis, H. Analysing drought severity and areal extent by 2D Archimedean copulas. *Water Resour. Manag.* **2016**, *30*, 5723–5735. [CrossRef]

16. He, W.; Lawless, J.F. Bivariate location–scale models for regression analysis, with applications to lifetime data. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **2005**, *67*, 63–78. [CrossRef]
17. Wienke, A.; Locatelli, I.; Yashin, A.I. The modelling of a cure fraction in bivariate time-to-event data. *Austrian J. Stat.* **2006**, *35*, 67–76. [CrossRef]
18. Fachini, J.B.; Ortega, E.M.; Cordeiro, G.M. A bivariate regression model with cure fraction. *J. Stat. Comput. Simul.* **2014**, *84*, 1580–1595. [CrossRef]
19. Alizadeh, M.; Tahmasebi, S.; Haghbin, H. The exponentiated odd log-logistic family of distributions: Properties and applications. *J. Stat. Model. Theory Appl.* **2020**, *1*, 29–52.
20. Nelsen, R.B. *An Introduction to Copulas*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2007.
21. Sklar, M. Fonctions de répartition à n dimensions et leurs marges. *Ann. de L'ISUP* **1959**, *8*, 229–231.
22. Sklar, A. Random variables, joint distribution functions, and copulas. *Kybernetika* **1973**, *9*, 449–460.
23. Joe, H. Multivariate dependence measures and data analysis. *Comput. Statist. Data Anal.* **1993**, *16*, 279–297. [CrossRef]
24. Clayton D.G. A model for association in bivariate life tables and its applications in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* **1978**, *65*, 141–151. [CrossRef]
25. Frank M.J. On the simultaneous associativity of $F(x, y)$ and $x + y − F(x, y)$. *Aequationes Math.* **1979**, *19*, 194–226. [CrossRef]
26. Koshevoy, G.A. Multivariate Gini Indices. *J. Multivar. Anal.* **1997**, *60*, 252–276. [CrossRef]
27. Gleaton, J.U.; Lynch, J.D. Properties of generalized log-logistic families of lifetime distributions. *J. Probab. Stat. Sci.* **2006**, *4*, 51–64.
28. Mudholkar, G.S.; Srivastava, D.K.; Kollia, G.D. A generalization of the Weibull distribution with application to the analysis of survival data. *J. Am. Stat. Assoc.* **1996**, *91*, 1575–1583. [CrossRef]
29. Nelder, J.A.; Mead, R. A simplex method for function minimization. *Comput. J.* **1965**, *7*, 308–313. [CrossRef]
30. R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2022.
31. Kojadinovic, I.; Yan, J. Modeling multivariate distributions with continuous margins using the copula R package. *J. Stat. Softw.* **2010**, *34*, 1–20. [CrossRef]
32. Othman, A.J.; Eliseeva, L.G.; Ibragimova, N.A.; Zelenkov, V.N.; Latushkin, V.V.; Nicheva, D.V. Dataset on the effect of foliar application of different concentrations of silicon dioxide and organosilicon compounds on the growth and biochemical contents of oak leaf lettuce (Lactuca sativa var. crispa) grown in phytotron conditions. *Data Brief* **2021**, *38*, 107328. [CrossRef] [PubMed]
33. Vila, R.; Balakrishnan, N.; Saulo, H. An Upper Bound and a Characterization for Gini's Mean Difference Based on Correlated Random Variables. *Preprint* 2023. Available online: https://arxiv.org/pdf/2301.07229.pdf (accessed on 15 July 2023).