Iranian Journal of Fuzzy Systems

Volume 20, Number 3, (2023), pp. 159-175

University of Sistan and Baluchestan

Original paper

# A new copula-based bivariate Gompertz–Makeham model and its application to COVID-19 mortality data

M. Esfahani[1], M. Amini[2], G. R. Mohtashami-Borzadaran[3] and A. Dolati[4]

[1] *Department of Statistics, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad, Iran*

[2,3,4] *Department of Statistics, Ordered Data, Reliability and Dependency Center of Excellence, Ferdowsi University of Mashhad, P.O. Box 91775-1159, Mashhad, Iran*

[4] *Department of Statistics, Faculty of Mathematical Sciences, Yazd University, Yazd, Iran*

m.esfahani@mail.um.ac.ir, m-amini@um.ac.ir, grmohtashami@um.ac.ir, adolati@yazd.ac.ir

**Abstract**

One of the useful distributions in modeling mortality (or failure) data is the univariate Gompertz–Makeham distribution. To examine the relationship between the two variables, the extended bivariate Gompertz–Makeham distribution is introduced, and its properties are provided. Also, some reliability indices, including aging intensity and stress-strength reliability, are calculated for the proposed model. Here, a new copula function is constructed based on the extended bivariate Gompertz–Makeham distribution. Some of its features including dependency properties, such as dependence structure, some measures of dependence, and tail dependence, are studied. The estimation of the parameters of new copula is presented, and at the end, a simulation study and a performance analysis based on the real data are presented. So, by analyzing the mortality data due to COVID-19, the appropriateness of the proposed model is examined.

*Keywords:* Copula function, bivariate Gompertz–Makeham distribution, dependence measures, dependence structure, reliability.

## 1 Introduction

The univariate Gompertz distribution was used by [23] to analyze survival and mortality data. Sometimes we deal with data that existing models are not able to describe. For this reason, we are looking for generalized models to be able to describe these data. Makeham [15] used this model to provide a more accurate distribution for analyzing mortality data, known as the Gompertz–Makeham distribution. Next, to analyze the states of more than one variable, bivariate distributions based on the Gompertz–Makeham distribution are constructed. Some of these bivariate distributions are based on Gompertz marginal distributions. To read more about the most important distributions of bivariate Gompertz–Makeham, we refer to [12, 16, 17, 19]. Also, for this purpose, the extended bivariate Gompertz–Makeham ($EBGM$) distribution is introduced.

Copulas are functions used to link the univariate marginal distribution functions and their corresponding joint distribution function, first introduced in [21]. These functions have many applications in the field of probability and statistics and are one of the most useful tools for studying the structure of dependency between random variables. The most important and complete references for studying the copulas, their features, and applications are [9, 18]. Copula functions have also been considered in fuzzy probability studies, which can be referred to [6, 22, 24]. In recent years, based on the copulas function, new families of fuzzy implication functions have been introduced, which are called probabilistic implications. These new families make a link between probability theory and fuzzy logic and are also used to study the relationship between two or more random variables. Therefore, the proposed model can also be considered in this context. Grzegorzewski [5] introduced two new families of fuzzy implication functions based on copulas function, which

Corresponding Author: M. Amini

Received: August 2022; Revised: January 2023; Accepted: February 2023.

https://doi.org/10.22111/IJFS.2023.7645

make a link between probability theory and fuzzy logic. In follow, Dolati et al. [2] offered a conditional implication based on copulas. Also, Helbin et al. [7] proposed some properties of fuzzy implications based on copulas. Therefore, the model proposed in this article can be used in this direction, and its characteristics can be investigated. In [3], also proposed an extension of Archimedean copulas with concave multiplicative generators. In the following, we briefly explain the definition of the copula function and its basic features.

A function $C : [0,1]^2 \to [0,1]$ is a copula function if for all $u, v \in [0,1]$, the following properties hold:

i) $C(u,0) = C(0,v) = 0$.

ii) $C(u,1) = u, \quad C(1,v) = v$.

iii) For every $u_1, u_2, v_1, v_2 \in [0,1]$ such that $u_1 \leq u_2$ and $v_1 \leq v_2$, we have

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u1, v1) \geq 0.$$

In this paper, we present a new copula based on the $EBGM$ distribution and examine its features. The study of the dependence structure of this copula is done in Section 3. In Section 4, we obtain some dependency measures, such as Kendall's $\tau$, Spearman's $\rho$, Blomqvist's $\beta$, and Gini's $\gamma$ for the desired copula and provide numerical values for different values of the parameters. In Section 5, the tail dependence of the new copula is considered. In Section 6, some applications of the $EBGM$ distribution in reliability are presented. Aging intensity, stress-strength reliability, and conditional hazard function of the proposed model are calculated. In Section 7 by performing simulations, we estimate the copula parameters and study their performance by index MSE. An analysis of a real bivariate data set is performed to examine the applications of the proposed copula in Section 8. In the end, the data generation method for the proposed copula is disscused.

# 2 Proposed model

In this section, we introduce a new copula function by using the $EBGM$ distribution and examine its features.
It is worth noting that a large class of convex survival functions can be defined using the Laplace–Stiltjes transform. For this aim, if $Z$ is a nonnegative random variable with Laplace–Stiltjes transform $L_Z(t) = E(e^{-tZ})$, then it is clear that $\bar{F}(t) = E(e^{-tZ})$ is a survival function. In the following, we provide an example.

**Example 2.1.** *Let $Z$ be a random variable with the discrete probability density function $P(Z = k) = \frac{\theta^{k-1} e^{-\theta}}{(k-1)!}, \theta > 0, k = 1, 2, \ldots$. By performing simple calculations and obtaining the Laplace–Stiltjes transform of $Z$, we have*

$$S(t) = E(e^{-tZ}) = \exp\{-t - \theta(1 - e^{-t})\}, \quad t, \theta > 0. \tag{1}$$

*It is clear that $S(t)$ is a decreasing and convex function with $S(0) = 1$ and $S(\infty) = 0$. Therefore, it can be concluded that $S(t)$ is a survival function. Note that $S(t)$ is the survival function of the univariate Gompertz–Makeham distribution.*

Now, using the survival function $S(t)$, we define a bivariate distribution as follows. The Gompertz model is widely used in the study of mortality data. Therefore, it seems that the bivariate Gompertz model leads to useful results for comparing the mortality rates of two societies. For this purpose, using model (1), we introduce a bivariate survival model and study its properties.

**Proposition 2.2.** *Let $S(t)$ be a continuous survival function. Then, for $t = x + y + \alpha xy, \alpha \in [0,1]$, the function $R : [0,\infty]^2 \longrightarrow [0,1]$, defined by*

$$R(x,y) = S(x + y + \alpha xy) = \exp\{-x - y - \alpha xy - \theta[1 - \exp(-x - y - \alpha xy)]\}, \quad 0 \leq \alpha \leq 1, \quad \theta \geq 0, \tag{2}$$

*is a bivariate survival function.*

*Proof.* According to the conditions stated in [9, Section 2.1.1] for bivariate survival functions, $R(0,0) = 1$, $R(x,\infty) = R(\infty,y) = R(\infty,\infty) = 0$, and $R(x,y)$ applies to the rectangle inequality if $\frac{\partial^2 R(x,y)}{\partial x \partial y} \geq 0$. It can be easily written that $R(x,0) = P(X > x, Y > 0) = S(x)$ and $R(0,y) = P(X > 0, Y > y) = S(y)$. □

**Remark 2.3.** *Bivariate survival function (2) is a generalization of bivariate Gompertz–Makeham that we call it the extended version of bivariate Gompertz–Makeham ($EBGM$) distribution.*

Now, using the Sklar's theorem, we obtain the copula function of the proposed bivariable distribution and calculate the dependency coefficients for it. For this purpose, we have

$$R(x,y) = \hat{C}(S(x), S(y)), \tag{3}$$

where $\hat{C}(\cdot,\cdot)$ is the survival copula. By using the transformations $u = S(x)$ and $v = S(y)$, in view of Sklar's theorem, we have

$$\hat{C}(u,v) = R(S^{-1}(u), S^{-1}(v)), \qquad u,v \in (0,1).$$

Let $\psi(t;\theta) = \ln S(t) = \theta e^{-t} - t - \theta$, which is a decreasing function with $\psi(0) = 0$ and its derivative $\psi'(t) = -1 - \theta e^{-t}$ is a monotone function. Thus it can be written as $u = S(t) = \exp(\psi(t))$ and $S^{-1}(u) = \psi^{-1}(\ln(u))$. As a result

$$\hat{C}(u,v) = R(\psi^{-1}(\ln(u)), \psi^{-1}(\ln(v)), \qquad u,v \in (0,1), \tag{4}$$

and thus

$$\begin{aligned}\hat{C}(u,v) = &\exp\{-\psi^{-1}(\ln(u)) - \psi^{-1}(\ln(v)) - \alpha\psi^{-1}(\ln(u))\psi^{-1}(\ln(v)) - \theta[1 - \exp(-\psi^{-1}(\ln(u)) - \psi^{-1}(\ln(v)) \\ &- \alpha\psi^{-1}(\ln(u))\psi^{-1}(\ln(v)))]\}. \end{aligned} \tag{5}$$

According to the relationship between copula and survival copula functions $\hat{C}(u,v) = u + v - 1 + C(1-u, 1-v)$ presented in [18], the copula corresponding to (5) is given by

$$\begin{aligned}C(u,v) = u + v - 1 + &\exp\{-\psi^{-1}(\ln(1-u)) - \psi^{-1}(\ln(1-v)) \\ &-\alpha\psi^{-1}(\ln(1-u))\psi^{-1}(\ln(1-v)) - \theta[1 - \exp(-\psi^{-1}(\ln(1-u)) - \psi^{-1}(\ln(1-v)) \\ &-\alpha\psi^{-1}(\ln(1-u))\psi^{-1}(\ln(1-v)))]\}. \end{aligned} \tag{6}$$

We call the copula function (6) as the new Gompertz–Makeham copula and denoted by $NGMC$.

In Figure 1, we provide the contour plot of $NGMC$ due to the deep grasp of understanding the performance of mentioned copula function. In accordance with this plots, the depth of function is related to its side with some differences for given values. This figure shows the contour plot of $NGMC$ for $\theta = 2$ and different values of $\alpha$.

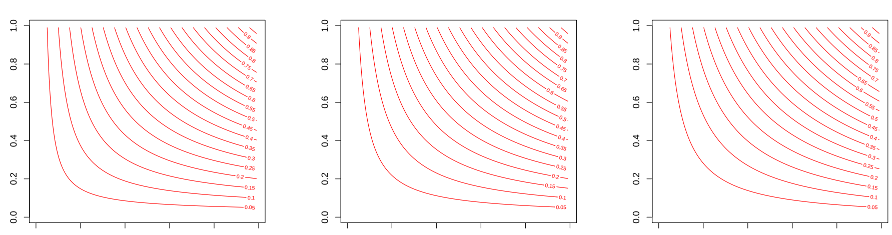

Figure 1: Contour plot of $NGMC$ for $\theta = 2$ and $\alpha = 0.01$ (left), $\alpha = 0.5$ (center), and $\alpha = 0.99$ (right).

# 3 Dependence structure

let $X$ and $Y$ be two random variables with the joint distribution and density functions $H$ and $h$ and the associated copula $C$. Nelsen [18] presented some dependence concepts for the two variables $X$ and $Y$ as follows:

- $X$ and $Y$ are said to be positively (negatively) quadrant dependent denoted by $PQD(NQD)$ if and only if $C(u,v) \geq (\leq) uv$ for all $u,v \in I = [0,1]$.

- $X$ and $Y$ are said to be left corner set decreasing, denoted by $LCSD(X,Y)$, if and only if
$$H(x,y)H(x',y') \geq H(x,y')H(x',y),$$
for all $x,y,x',y'$ in $\bar{R}$ such that $x \leq x'$ and $y \leq y'$.
- $X$ and $Y$ are said to be right corner set increasing, denoted by $RCSI(X,Y)$, if and only if
$$\bar{H}(x,y)\bar{H}(x',y') \geq \bar{H}(x,y')\bar{H}(x',y),$$
for all $x,y,x',y'$ in $\bar{R}$ such that $x \leq x'$ and $y \leq y'$.
- $Y$ is left tail decreasing in $X$, shown as $LTD[Y|X]$, if and only if $\frac{C(u,v)}{u}$ is nonincreasing in $u$, for all $v$ or $\frac{\partial C(u,v)}{\partial u} \leq \frac{C(u,v)}{u}$ for all $u$.
- $Y$ is said to be right tail increasing in $X$, briefly $RTI[Y|X]$, if and only if $\frac{1-u-v+C(u,v)}{1-u}$ is nondecreasing in $u$, for all $v$ or $\frac{\partial C(u,v)}{\partial u} \leq \frac{v-C(u,v)}{1-u}$ for all $u$.
- $Y$ is stochastically increasing(decreasing) in $X$, denoted by $SI(SD)[Y|X]$, if and only if $\frac{\partial C(u,v)}{\partial u}$ is nonincreasing(nondecreasing) in $u$ for all $v$.
- A function $f$ from $R^2$ to $R$ is totally positive of order two, abbreviated $TP2$, if $f(x,y) \geq 0$ on $R^2$ and whenever for $x \leq x'$ and $y \leq y'$, it satisfies the following inequality:
$$f(x,y)f(x',y') - f(x',y)f(x,y') \geq 0.$$

The following proposition presents some results for the proposed model (6).

**Proposition 3.1.** *The survival function $R(x,y)$ in (2) is $TP2$.*

*Proof.* Since the random variable $X$ and $Y$ are continuous, it is enough to show that $\frac{\partial^2 R(x,y)}{\partial x \partial y} \geq 0$; see tip (i) in section 2.1.1 of [9]. Let $g(x,y) = -x - y - \alpha xy$. Therefore
$$\frac{\partial R(x,y)}{\partial x} = (\frac{\partial g(x,y)}{\partial x} + \theta \frac{\partial g(x,y)}{\partial x} e^{g(x,y)}) \exp\{g(x,y) - \theta + \theta e^{g(x,y)}\},$$
and
$$\begin{aligned}
\frac{\partial^2 R(x,y)}{\partial x \partial y} =&[\frac{\partial^2 g(x,y)}{\partial x \partial y} + \theta \frac{\partial^2 g(x,y)}{\partial x \partial y} e^{g(x,y)} + \theta \frac{\partial g(x,y)}{\partial x} \frac{\partial g(x,y)}{\partial y} e^{g(x,y)} \\
&+ (\frac{\partial g(x,y)}{\partial y} + \theta \frac{\partial g(x,y)}{\partial y} e^{g(x,y)})(\frac{\partial g(x,y)}{\partial x} + \theta \frac{\partial g(x,y)}{\partial x} e^{g(x,y)})] \exp\{g(x,y) - \theta + \theta e^{g(x,y)}\} \\
=&[-\alpha - \alpha\theta e^{g(x,y)} + \theta(-1-\alpha x)(-1-\alpha y)e^{g(x,y)} + (-1-\alpha y + \theta(-1-\alpha y)e^{g(x,y)}) \\
&\times (-1-\alpha x + \theta(-1-\alpha x)e^{g(x,y)})] \exp\{g(x,y) - \theta + \theta e^{g(x,y)}\} \\
=&[(1-\alpha) + (1-\alpha)\theta e^{g(x,y)} + \theta(\alpha x + \alpha y + \alpha^2 xy)e^{g(x,y)} + \alpha x + \alpha y + \alpha^2 xy + \theta^2(1+\alpha x)(1+\alpha y)e^{2g(x,y)} \\
&+ 2\theta(1+\alpha x)(1+\alpha y)e^{g(x,y)}] \exp\{g(x,y) - \theta + \theta e^{g(x,y)}\},
\end{aligned}$$

which is nonnegative for $\alpha \in [0,1]$ and $\theta > 0$. □

Based on [18, Corollaries 5.2.16 and 5.2.17], since $R(x,y)$ is $TP2$, we have $RCSI(X,Y)$ and so the survival copula of $NGMC$ is $TP2$.

**Corollary 3.2.** *Let $(X,Y)$ be a random vector with the joint survival function (2) and corresponding $NGMC$, presented in (6). The following results hold:*
*(i) $X$ and $Y$ are $RTI(X,Y)$,*
*(ii) $X$ and $Y$ are $PQD$.*

*Proof.* In Proposition 3.1, we obtain that survival function $R(x,y)$ is $TP2$. Therefore according to [18, Corollary 5.2.16], we conclude that $RCSI(X,Y)$. So, since $RCSI(X,Y)$, by [18, diagram 5.8 p. 200 ], the model is $RTI(X,Y)$ and also concludes notions, such as $PQD$, which indicates positive dependence structure. □

## 4 Some measures of dependence

Now, to investigate the behavior of $NGMC$, we obtain some nonparametric measure of dependence for this model. The most famous of these measures are Kendall's $\tau$, Spearman's $\rho$, Blomqvist's $\beta$, and Gini's $\gamma$, which are defined as follows according to copula function, respectively [18]:

$$\tau_C = 4\int_0^1\int_0^1 C(u,v)dC(u,v) - 1 = 1 - 4\int_0^1\int_0^1 \frac{\partial C(u,v)}{\partial u}\frac{\partial C(u,v)}{\partial v}dudv, \tag{7}$$

$$\rho_C = 12\int_0^1\int_0^1 C(u,v)dudv - 3, \tag{8}$$

$$\beta_C = 4C(\frac{1}{2},\frac{1}{2}) - 1, \tag{9}$$

$$\gamma_C = 4[\int_0^1 C(u,1-u)du - \int_0^1 [u - C(u,u)]du]. \tag{10}$$

**Proposition 4.1.** *If $(X,Y)$ is a random vector with corresponding $NGMC$, then the measures of dependence for $(X,Y)$ with $NGMC$ are given below:*

$$\begin{aligned}\tau_{NGMC} =& 1 - 4\int_0^1\int_0^1 \{1 + [K(u,v) - \theta K(u,v)\exp\{-\psi^{-1}(\ln(1-u))) - \psi^{-1}(\ln(1-v))\\ &- \alpha\psi^{-1}(\ln(1-u))\psi^{-1}(\ln(1-v)\}]\hat{C}(u,v)\}\\ &\times\{1 + [J(u,v) - \theta J(u,v)\exp\{-\psi^{-1}(\ln(1-u))) - \psi^{-1}(\ln(1-v))\\ &- \alpha\psi^{-1}(\ln(1-u))\psi^{-1}(\ln(1-v))\}]\hat{C}(u,v)\}dudv,\end{aligned}$$

*where* $K(u,v) = \frac{1+\alpha\psi^{-1}(\ln(1-v))}{(1-u)\psi'(\psi^{-1}(\ln(1-u))}$ *and* $J(u,v) = \frac{1+\alpha\psi^{-1}(\ln(1-u))}{(1-v)\psi'(\psi^{-1}(\ln(1-v))}$,

$$\begin{aligned}\rho_{NGMC} =& 12\int_0^1\int_0^1 \{u + v - 1 + \exp\{-\psi^{-1}(\ln(1-u)) - \psi^{-1}(\ln(1-v)) - \alpha\psi^{-1}(\ln(1-u))\psi^{-1}(\ln(1-v))\\ &- \theta[1 - \exp(-\psi^{-1}(\ln(1-u)) - \psi^{-1}(\ln(1-v))\}dudv - 3,\\ \gamma_{NGMC} =& 4\int_0^1 \exp\{-\psi^{-1}(\ln(1-u)) - \psi^{-1}(\ln(u)) - \alpha\psi^{-1}(\ln(1-u))\psi^{-1}(\ln(u))\\ &- \theta[1 - \exp(-\psi^{-1}(\ln(1-u)) - \psi^{-1}(\ln(u)) - \alpha\psi^{-1}(\ln(1-u))\psi^{-1}(\ln(u)))]\}du,\\ &- 4\int_0^1 (1 - u - \exp\{-2\psi^{-1}(\ln(1-u)) - \alpha(\psi^{-1}(\ln(1-u)))^2\\ &- \theta[1 - \exp(-2\psi^{-1}(\ln(1-u)) - \alpha(\psi^{-1}(\ln(1-u)))^2)]\})du,\\ \beta_{NGMC} =& 4\exp\{-\psi^{-1}(\ln(\frac{1}{2})) - \psi^{-1}(\ln(\frac{1}{2})) - \alpha\psi^{-1}(\ln(\frac{1}{2}))\psi^{-1}(\ln(\frac{1}{2}))\\ &- \theta[1 - \exp(-\psi^{-1}(\ln(\frac{1}{2})) - \psi^{-1}(\ln(\frac{1}{2})) - 1.\end{aligned}$$

The values of these measures of dependence for $NGMC$ model are provided in Table 1 for different values of the parameters $\alpha$ and $\theta$. From Table 1, it is obvious that by increasing each of the parameters $\alpha$ and $\theta$ results in increasing the situation for the coefficient values. In this regard, it is also clear that these situations are the same for all four coefficients. In addition, increasing the rate of these coefficients in parameter $\theta$ is less than in the parameter $\alpha$.

Table 1: Dependence coefficients for different values $\alpha$ and $\theta$.

| $\alpha$ | $\theta$ | $\tau$ | $\rho$ | $\beta$ | $\gamma$ |
|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 5 | 0.0077108 | 0.0091082 | 0.0020425 | 0.0071796 |
| 0 | 10 | 0.0079653 | 0.0093891 | 0.0032302 | 0.0082211 |
| 0 | 20 | 0.0085185 | 0.0095196 | 0.008971 | 0.0098259 |
| 0.01 | 0 | 0.0092329 | 0.0094661 | 0.0087945 | 0.0086769 |
| 0.01 | 5 | 0.0608539 | 0.0177004 | 0.0141939 | 0.0125076 |
| 0.01 | 10 | 0.0814848 | 0.0186435 | 0.0171494 | 0.0150217 |
| 0.01 | 20 | 0.0915941 | 0.0190114 | 0.0198565 | 0.0186842 |
| 0.25 | 0 | 0.1081035 | 0.1676177 | 0.1131810 | 0.1291653 |
| 0.25 | 5 | 0.1916431 | 0.1911486 | 0.1880398 | 0.1899215 |
| 0.25 | 10 | 0.2164509 | 0.3193787 | 0.1956151 | 0.1965621 |
| 0.25 | 20 | 0.2816711 | 0.2671732 | 0.2114139 | 0.2198954 |
| 0.5 | 0 | 0.4017420 | 0.4752950 | 0.4035514 | 0.4082505 |
| 0.5 | 5 | 0.4808417 | 0.5485596 | 0.4747397 | 0.4611259 |
| 0.5 | 10 | 0.4997715 | 0.5744469 | 0.5497783 | 0.4934377 |
| 0.5 | 20 | 0.5289899 | 0.6150295 | 0.6340040 | 0.5314749 |
| 0.75 | 0 | 0.6832735 | 0.7222555 | 0.6525617 | 0.6329195 |
| 0.75 | 5 | 0.7482679 | 0.7605206 | 0.7043329 | 0.7268439 |
| 0.75 | 10 | 0.7942453 | 0.7866408 | 0.7345665 | 0.7720722 |
| 0.75 | 20 | 0.8224745 | 0.7998752 | 0.7474535 | 0.8107038 |
| 0.99 | 0 | 0.8531718 | 0.8200414 | 0.8785191 | 0.8131689 |
| 0.99 | 5 | 0.9708391 | 0.9936896 | 0.9621588 | 0.9716251 |
| 0.99 | 10 | 0.9813484 | 0.9952591 | 0.9836202 | 0.9727606 |
| 0.99 | 20 | 0.9991689 | 0.9971459 | 0.9947255 | 0.9847479 |
| 1 | 0 | 0.9910357 | 0.9972385 | 0.9914979 | 0.9916316 |
| 1 | 5 | 0.9967582 | 0.9996211 | 0.9959702 | 0.9968233 |
| 1 | 10 | 0.9986077 | 0.9999452 | 0.9987492 | 0.9992055 |
| 1 | 20 | 0.9997132 | 0.9999922 | 0.9992689 | 0.9999969 |

## 5 Tail dependence

Let $X$ and $Y$ be two continuous random variables with distribution functions $F$ and $G$, respectively. The upper and lower tail dependence coefficient of $(X, Y)$ are, respectively, defined by

$$\begin{aligned}\lambda_U &= \lim_{t\to 1^-} \frac{P\big[Y > G^{-1}(t), X > F^{-1}(t)\big]}{P(X > F^{-1}(t))}\\ &= \lim_{t\to 1^-} \frac{R(S^{-1}(t), S^{-1}(t))}{1-t}\\ &= \lim_{t\to 1^-} \frac{\hat{C}(1-t, 1-t)}{1-t}.\end{aligned}$$

and

$$\lambda_L = \lim_{t\to 0^+} \frac{P\big[X \le F^{-1}(t), Y \le G^{-1}(t)\big]}{t},$$

provided that the above limits exist. For more details, we refer to [9, 18].

For random vector $(X, Y)$ with the copula $NGMC$, we have

$$\lambda_U = \lim_{t\to 1^-} \frac{\exp\{-2\psi^{-1}(\ln(1-t)) - \alpha(\psi^{-1}(\ln(1-t)))^2 - \theta[1-\exp(-2\psi^{-1}(\ln(1-t)) - \alpha(\psi^{-1}(\ln(1-t))^2)]\}}{1-t} = \frac{0}{0}.$$

Therefore, by using the Hopital's law, it results

$$\lambda_U = \lim_{t\to 1^-} -[\frac{2\alpha\psi^{-1}(\ln(1-t))+2}{(1-t)\psi'(\psi^{-1}(\ln(1-t)))} + \theta\frac{2\alpha\psi^{-1}(\ln(1-t))+2}{(1-t)\psi'(\psi^{-1}(\ln(1-t)))}\exp(-2\psi^{-1}(\ln(1-t)) - \alpha(\psi^{-1}(\ln(1-t))^2]$$
$$\times \exp\{-2\psi^{-1}(\ln(1-t)) - \alpha(\psi^{-1}(\ln(1-t)))^2 - \theta[1-\exp(-2\psi^{-1}(\ln(1-t)) - \alpha(\psi^{-1}(\ln(1-t))^2)]\}.$$

Now, using the point that if $F$ is a distribution function with the survival function $\bar{F}$ and quantile function $F^{-1}$, then $F^{-1}(t) = \bar{F}^{-1}(1-t)$, and for the $EBGM$ distribution, we have $\psi^{-1}(\ln(1-t)) = \bar{S}^{-1}(t)$, where $\bar{S}(t) = F(t)$ is a distribution function. Hence for $t = 1$, it results $\psi^{-1}(\ln(0)) = \bar{S}^{-1}(1) = \infty$, and by placing in the limit, the value of the exponential function multiplied by the bracket equals zero. As a result, $\lambda_U = 0$.

Similarly,

$$\begin{aligned}\lambda_L &= \lim_{t\to 0^+} \frac{\bar{R}(S^{-1}(t), S^{-1}(t))}{t} \\ &= \lim_{t\to 0^+} \frac{2\bar{S}(S^{-1}(t)) - 1 + R(S^{-1}(t), S^{-1}(t))}{t} \\ &= \lim_{t\to 0^+} \frac{1-2t+R(S^{-1}(t), S^{-1}(t))}{t}.\end{aligned}$$

Due to the complex form of the $NGMC$ function, the values of $\lambda_L(NGMC)$ can be obtained by performing numerical method.

# 6 Application in reliability

In this section, we calculate some reliability indices, including ageing intensity and stress-strength reliability for the extended bivariate Gompertz–Makeham distribution.

## 6.1 Hazard functions

Another main concept of reliability is the hazard rate function of a distribution. By examining its behavior, some properties of the studied distribution can be obtained. In this section, we present the results of the conditional hazard rate function for the $EBGM$ distribution. Let $X$ and $Y$ be variables with univariate distribution in (1). Then the hazard rate function of $X$ and $Y$ consists of $h(t) = \frac{f(t)}{\bar{F}(t)} = 1 + \theta\exp\{-t\}, t \geq 0$, is a decreasing function.

Now let $(X, Y)$ be a vector variable with the $EBGM$ distribution presented in (2). The conditional survival function of $(X|Y > y)$ and $(X|Y = y)$ based on the $EBGM$ model is given by

$$\bar{F}_{X|Y>y}(x|y) = \exp\{-x - \alpha xy - \theta[\exp(-y) - \exp(-x-y-\alpha xy)]\},$$

and

$$f_{X|Y>y}(x|y) = (1 + \alpha y + \theta[(1+\alpha y)\exp(-x-y-\alpha xy) - \exp(-y)])\bar{F}_{X|Y>y}(x|y).$$

Thus

$$h_{X|Y>y}(t) = \frac{f_{X|Y>y}(x|y)}{\bar{F}_{X|Y>y}(x|y)} = 1 + \alpha y + \theta[(1+\alpha y)\exp(-t-y-\alpha ty) - \exp(-y)]. \tag{11}$$

**Corollary 6.1.** *Let $(X, Y)$ be a random vector with the EBGM distribution. Then*

$$\frac{\partial h_{X|Y>y}(t)}{\partial t} = -\theta(1+\alpha y)^2\exp\{-t-y-\alpha ty\} \leq 0,$$

*and thus $h_{X|Y>y}(t)$ is a decreasing function of $t$.*

## 6.2 Aging intensity

One of the basic outlooks in reliability theory is “aging property”, which is an essential property of a unit that may be a system of components or even a living organism. Some significant measures of aging are hazard rate, mean residual life, inactivity time, and so on. These concepts are directly related to the important concept of the total time on a test. For more details, we refer to [14].

Jiang et al. [8] showed that the representation of aging of a system by failure rate is qualitative, and accordingly, they introduced a new notion called aging intensity ($AI$). Let $X$ be a absolutely continuous random variable with the distribution function $F$. The $AI$ for $X$ at time $t$, denoted by $L_X(t)$, is defined by

$$L_X(t) = \frac{h_X(t)}{H_X(t)}, \tag{12}$$

where $h_X(t)$ is the hazard rate function of $X$ and $H_X(t) = \frac{1}{t}\int_0^t h_X(x)dx$ is the hazard rate average. For the univariate Gompertz–Makeham distribution presented in (1), the univariate $AI$ is given by

$$L_X(t) = \frac{t + \theta t e^{-t}}{t - \theta e^{-t} + \theta}. \tag{13}$$

Figure 2 illustrates the behavior of $L_X(t)$ for the univariate Gompertz–Makeham distribution. Figure 2 shows that

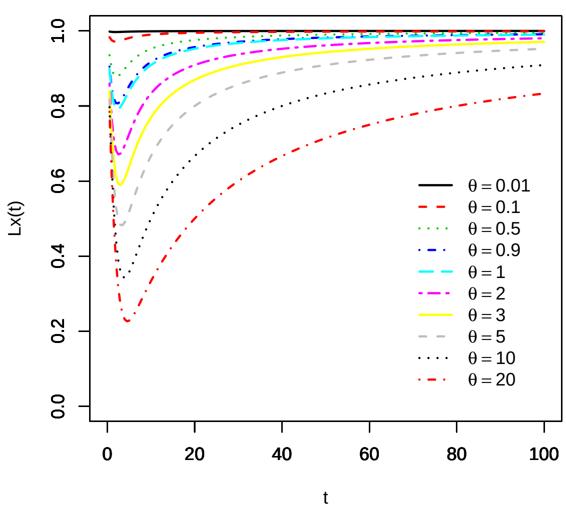


Figure 2: $L_X$ plot of the univariate Gompertz–Makeham distribution for different values of $\theta$.

$L_X(t)$ for the univariate Gompertz–Makeham distribution decreases when the parameter $\theta$ increases. Also, for a fixed $\theta$, the $L_X(t)$ graph first becomes strongly descending and then ascending. In the study of mortality models, this means that the rate of aging decreases rapidly at the beginning of birth and then increases with a gentler slope.

Now we will express the $AI$ concept in bivariate mode and present it for the $EBGM$ distribution. For this purpose, let $(X, Y)$ be a continuous random vector with the survival function $\bar{F}(x, y)$. The vector of $AIs$ for the pair $(X, Y)$ (denoted by $BAI(X, Y)$) is defined by $L_{X,Y}(x, y) = (L_1(x, y), L_2(x, y))$, where $L_i(x, y) = \frac{h_i(x,y)}{H_i(x,y)}$ for $i = 1, 2$, and $h_1(x, y) = -\frac{\partial}{\partial x}\ln \bar{F}(x, y)$, $h_2(x, y) = -\frac{\partial}{\partial y}\ln \bar{F}(x, y)$, $H_1(x, y) = \frac{1}{x}\int_0^x h_1(u, y)du$ and $H_2(x, y) = \frac{1}{y}\int_0^y h_2(x, u)du$.

Based on these notions, in the following result, we provide the $AI$ for the $EBGM$ distribution.

**Proposition 6.2.** *Let $(X, Y)$ be a vector variable with the $EBGM$ distribution presented in (2). Then, the bivariate $AI$ vector of $(X, Y)$ is given by $L_{X,Y}(x, y) = (L_1(x, y), L_2(x, y))$ where*

$$L_1(x, y) = \frac{(1 + \alpha y)(1 + \theta \exp\{-x - y - \alpha xy\})}{1 + \alpha y - \frac{\theta}{x}[\exp\{-x - y - \alpha xy\} - \exp\{-y\}]},$$

$$L_2(x,y) = \frac{(1+\alpha x)(1+\theta\exp\{-x-y-\alpha xy\})}{1+\alpha x - \frac{\theta}{y}[\exp\{-x-y-\alpha xy\} - \exp\{-x\}]}.$$

*Proof.* For $\alpha \in [0,1]$ and $\theta > 0$, we have

$$h_1(x,y) = (1+\alpha y)(1+\theta\exp\{-x-y-\alpha xy\}),$$
$$h_2(x,y) = (1+\alpha x)(1+\theta\exp\{-x-y-\alpha xy\}),$$

and thus

$$\begin{aligned} H_1(x,y) &= \frac{1}{x}\int_0^x h_1(u,y)du \\ &= \frac{1}{x}\int_0^x (1+\alpha y)(1+\theta\exp\{-u-y-\alpha uy\})du \\ &= 1+\alpha y - \frac{\theta}{x}[\exp\{-x-y-\alpha xy\} - \exp\{-y\}], \end{aligned}$$

and similarly $H_2(x,y) = 1+\alpha x - \frac{\theta}{y}[\exp\{-x-y-\alpha xy\} - \exp\{-x\}]$. Thus, we have $L_{X,Y}(x,y) = (L_1(x,y), L_2(x,y)) = (\frac{h_1(x,y)}{H_1(x,y)}, \frac{h_2(x,y)}{H_2(x,y)})$, which is the required result. □

Comparing random vectors using the $BAL$ concept has an interesting interpretation. For this purpose, let $X = (X_1, X_2)$ and $Y = (Y_1, Y_2)$ be the bivariate random vectors with $L_X(x,y) = (L_{1X}(x,y), L_{2X}(x,y))$ and $L_Y(x,y) = (L_{1Y}(x,y), L_{2Y}(x,y))$. Then $X$ is greater than $Y$ in the $AI$ order and is denoted by $X \geq_{AI} Y$ if and only if $L_{iX}(x,y) \leq L_{iY}(x,y)$ for $i = 1,2$. This means that if all the components of the aging intensity vector of random vector $X$ are smaller than the corresponding components of random vector $Y$, then random vector $X$ is more (better) in the $AI$ order with the bivariate aging intensity — a weaker (not stronger) tendency has old age.

## 6.3 Stress-strength reliability

The stress-strength model is a pivotal model in the field of reliability. The stress-strength model refers to intrinsic random strength $X$ of a component in a system that is subject to random stress $Y$ during its operation, such that it works only when $Y < X$. The probability of $Y < X$ is denoted as the reliability of the component, that is, $R = P(Y < X)$. The estimation of $R$ plays a main role in the reliability analysis. A general report on this subject is given by [13]. Calculating stress-strength based on competing risk is one of the things that have received attention in recent years and was presented in [20]. Shih and Emura [20] obtained the stress as follows. Let $(X,Y)$ be a continuous random vector with the survival function $\bar{F}(x,y)$. The stress-strength parameter $R = P(Y < X)$ is given by

$$R = \int_0^\infty -\frac{\partial}{\partial y}\bar{F}(x,y)|_{x=y=z}dz. \tag{14}$$

The stress-strength parameter based on the competing risk for the $EBGM$ distribution is given in the following proposition.

**Proposition 6.3.** *Let $(X,Y)$ be a random vector having $EBGM$ distribution presented in (2). Then $R = \frac{1}{2}$.*

*Proof.* Based on equation (14), we have

$$-\frac{\partial}{\partial y}\bar{F}(x,y) = (1+\alpha x)(1+\theta\exp\{-x-y-\alpha xy\})\exp\{-x-y-\alpha xy-\theta[1-\exp(-x-y-\alpha xy)]\},$$

and therefore

$$\begin{aligned} R &= \int_0^\infty (1+\alpha z)(1+\theta\exp\{-2z-\alpha z^2\})\exp\{-2z-\alpha z^2-\theta[1-\exp(-2z-\alpha z^2 y)]\}dz \\ &= -\frac{1}{2}\exp\{-2z-\alpha z^2-\theta[1-\exp(-2z-\alpha z^2 y)]\}|_0^\infty = \frac{1}{2}. \end{aligned}$$

□

This means that if $(X,Y)$ is a random vector with mortality model, then the effect of these variables on each other is equal and they also have the same effect on the mortality rate.

# 7 Parameter estimation

In this section, we consider the estimation of the dependence parameters $\theta$ and $\alpha$, by using the maximum likelihood and moment methods.

## 7.1 Maximum likelihood approach

Let $(u_i, v_i)$, $i = 1, \ldots, n$, be a random set with the uniform marginal distributions $u = F(x)$ and $v = G(y)$ and the dependence structure $NGMC$ in (6). Joe and Xu [10] proposed a new estimation technique consisting in two steps. We have used their method to estimate $NGMC$ parameters. The log likelihood function, denoted by Ł, is obtained as bellow:

$$
\begin{aligned}
Ł(\theta,\alpha) = \sum_{i=1}^{n} \hat{C}(u_i,v_i)[&\frac{\alpha}{(1-u_i)(1-v_i)\psi'(\psi^{-1}(1-u_i))\psi'(\psi^{-1}(1-v_i))} \\
&- \theta(\frac{\partial K(u_i,v_i)}{\partial v_i} + J(u_i,v_i))\exp\{-\psi^{-1}(\ln(1-u_i)) - \psi^{-1}(\ln(1-v_i)) - \alpha\psi^{-1}(\ln(1-u_i))\psi^{-1}(\ln(1-v_i))\}] \\
&+ K(u_i,v_i)(1-\theta\exp\{-\psi^{-1}(\ln(1-u_i)) - \psi^{-1}(\ln(1-v_i)) - \alpha\psi^{-1}(\ln(1-u_i))\psi^{-1}(\ln(1-v_i))\})\frac{\partial \hat{C}(u_i,v_i)}{\partial v_i}.
\end{aligned}
$$

Since there is no closed form to obtain analytically maximum for $Ł(\theta, \alpha)$, so the optim function in $R$ package has been applied.

## 7.2 Moments approach

Consider the observations of $(x_i, y_i)$, $i = 1, \ldots, n$, from a random vector $(X, Y)$, with the $NGMC$. We can estimate Kendall's $\tau$ and Spearman's $\rho$ using the corresponding sample version $\tau_n$ and $\rho_n$, respectively.
By solving the equations $\tau(\alpha, \theta) = \tau_n$ and $\rho(\alpha, \theta) = \rho_n$ simultaneously, the moment estimates (MM) of the parameters are obtained.

# 8 Numerical study

For illustrating the methodology, in the following, we consider a simulation study and real data analysis for the Gompertz–Makeham copula.

## 8.1 Monte Carlo simulation and comparison

A simulation study has been performed to compare the estimation performance of the parameters $\alpha$ and $\theta$, and we have obtained maximum likelihood and moment estimates (MLE, MME) for these. The mean square error (MSE), the mean absolute error (MAE), and bias are also calculated for each of the parameters. The coverage probability in estimating $\alpha$ and $\theta$ is also calculated and denoted by $C_p$.

For computing the estimates, we use an algorithm proposed in [11]. Using packages GoFKernel, numDeriv, and bbmle of R software, we simulate 100000 independent samples from the $NGMC$ using different values of dependence parameters $\theta \in \{1, 2, 5, 10\}$ and $\alpha \in \{0.01, 0.5, 0.75, 0.99\}$ and four sample sizes $n \in \{10, 20, 50, 100\}$.
We assumed that, the independence is achieved when $\alpha = 0.01$, weak for $\alpha = 0.2$, moderate for $\alpha = 0.5$, and strong for $\alpha = 0.99$. The results are given in Tables 2–5. As can be seen, by increasing the value of $n$, the MSE of the parameters $\alpha$ and $\theta$ decreases. Also, with increasing the value of the parameter $\theta$, a downward trend is seen in the values of MSE.

Table 2: The mean and empirical MSE of Simultaneous estimators of $\alpha = 0.01$ and $\theta$ with MM and ML methods.

| $(\alpha, \theta)$ | n | $MLE\alpha$ | $MME\alpha$ | $Bias\alpha$ | $MSE\alpha$ | $MAE\alpha$ | $MLE\theta$ | $MME\theta$ | $Bias\theta$ | $MSE\theta$ | $MAE\theta$ | $C_p$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (0.01,1) | 10 | 0.00970 | 0.00981 | -0.00024 | 0.02091 | 0.12501 | 0.99959 | 0.99932 | 0.00034 | 0.08350 | 0.25111 | 0.95026 |
| | 20 | 0.01017 | 0.01013 | -0.00019 | 0.02090 | 0.12458 | 1.00069 | 1.000143 | 0.00059 | 0.08348 | 0.24965 | 0.94946 |
| | 50 | 0.01017 | 0.01037 | -0.00014 | 0.02085 | 0.12514 | 0.99944 | 0.99988 | -0.00075 | 0.08348 | 0.24969 | 0.95047 |
| | 100 | 0.00985 | 0.01021 | 0.00011 | 0.02081 | 0.12474 | 0.99945 | 1.00062 | -0.00073 | 0.08345 | 0.24956 | 0.94964 |
| (0.01,2) | 10 | 0.00967 | 0.01019 | -0.00047 | 0.02085 | 0.12515 | 2.00003 | 1.99983 | 0.00009 | 0.08342 | 0.24982 | 0.95082 |
| | 20 | 0.01073 | 0.01038 | 0.00026 | 0.02085 | 0.12520 | 2.00109 | 1.99949 | 0.00105 | 0.08342 | 0.25018 | 0.95083 |
| | 50 | 0.00983 | 0.01015 | 0.00001 | 0.02081 | 0.12519 | 2.00054 | 1.99941 | 0.00049 | 0.08341 | 0.24973 | 0.95082 |
| | 100 | 0.00995 | 0.01040 | 0.00012 | 0.02080 | 0.12489 | 1.99944 | 1.99971 | 0.00121 | 0.08340 | 0.25027 | 0.95055 |
| (0.01,5) | 10 | 0.00976 | 0.01032 | 0.00029 | 0.02084 | 0.12509 | 5.00283 | 5.00022 | 0.00019 | 0.08339 | 0.25035 | 0.94951 |
| | 20 | 0.01061 | 0.00991 | -0.00038 | 0.02084 | 0.12475 | 5.00005 | 5.00057 | 0.00066 | 0.08338 | 0.25023 | 0.94984 |
| | 50 | 0.00985 | 0.00966 | 0.00005 | 0.02083 | 0.12473 | 4.99953 | 4.99921 | 0.00007 | 0.08335 | 0.25097 | 0.95098 |
| | 100 | 0.01015 | 0.01028 | 0.00066 | 0.02078 | 0.12530 | 5.00106 | 4.99976 | 0.00021 | 0.08330 | 0.25017 | 0.94964 |
| (0.01,10) | 10 | 0.00977 | 0.01007 | 0.00009 | 0.02083 | 0.12508 | 10.00076 | 10.00003 | 0.00048 | 0.08331 | 0.25059 | 0.95010 |
| | 20 | 0.00968 | 0.01002 | 0.00050 | 0.02079 | 0.12516 | 9.99881 | 10.00032 | -0.00018 | 0.08327 | 0.24990 | 0.94999 |
| | 50 | 0.01031 | 0.00993 | 0.00061 | 0.02079 | 0.12507 | 9.99991 | 9.99945 | -0.00067 | 0.08321 | 0.24968 | 0.95060 |
| | 100 | 0.01049 | 0.00977 | -0.00079 | **0.02075** | 0.12505 | 9.99969 | 9.99980 | -0.00015 | **0.08309** | 0.24995 | 0.94901 |

Table 3: The mean and empirical MSE of Simultaneous estimators of $\alpha = 0.2$ and $\theta$ with MM and ML methods.

| $(\alpha, \theta)$ | n | $MLE\alpha$ | $MME\alpha$ | $Bias\alpha$ | $MSE\alpha$ | $MAE\alpha$ | $MLE\theta$ | $MME\theta$ | $Bias\theta$ | $MSE\theta$ | $MAE\theta$ | $C_p$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (0.2,1) | 10 | 0.19961 | 0.20024 | 0.00046 | 0.02079 | 0.12471 | 1.00137 | 0.99912 | 0.00193 | 0.08359 | 0.25083 | 0.95017 |
| | 20 | 0.19879 | 0.19996 | 0.00019 | 0.02079 | 0.12514 | 0.99968 | 0.99986 | -0.00173 | 0.08356 | 0.25017 | 0.94955 |
| | 50 | 0.20074 | 0.19979 | 0.00027 | 0.02079 | 0.12458 | 1.00034 | 0.99979 | -0.00003 | 0.08352 | 0.25049 | 0.95043 |
| | 100 | 0.19983 | 0.19991 | 0.00035 | 0.02078 | 0.12537 | 1.00121 | 1.00015 | 0.00084 | 0.08351 | 0.24983 | 0.94971 |
| (0.2,2) | 10 | 0.20038 | 0.20023 | 0.00044 | 0.02079 | 0.12510 | 2.00084 | 2.00004 | -0.00056 | 0.08350 | 0.25053 | 0.95080 |
| | 20 | 0.20001 | 0.19969 | -0.00076 | 0.02076 | 0.12515 | 2.00054 | 2.00031 | 0.00209 | 0.08344 | 0.25057 | 0.95032 |
| | 50 | 0.20069 | 0.19988 | -0.00019 | 0.02076 | 0.12467 | 1.99889 | 2.00145 | -0.00054 | 0.08343 | 0.25059 | 0.94940 |
| | 100 | 0.20028 | 0.19997 | 0.00004 | 0.02075 | 0.12532 | 1.99982 | 1.99998 | -0.00173 | 0.08342 | 0.25058 | 0.94932 |
| (0.2,5) | 10 | 0.20005 | 0.20015 | -0.00019 | 0.02076 | 0.12509 | 5.00047 | 5.00001 | 0.00093 | 0.08338 | 0.24958 | 0.94962 |
| | 20 | 0.19959 | 0.20029 | -0.00002 | 0.02075 | 0.12509 | 4.99923 | 4.99939 | 0.00129 | 0.08335 | 0.24970 | 0.94985 |
| | 50 | 0.19949 | 0.19989 | 0.00008 | 0.02074 | 0.12488 | 4.99945 | 5.00027 | -0.00001 | 0.08333 | 0.24985 | 0.95001 |
| | 100 | 0.19953 | 0.20007 | 0.00001 | 0.02072 | 0.12523 | 4.99907 | 4.99946 | 0.00013 | 0.08333 | 0.24945 | 0.95091 |
| (0.2,10) | 10 | 0.20016 | 0.20010 | -0.00018 | 0.02072 | 0.12498 | 9.99872 | 10.00018 | -0.00029 | 0.08327 | 0.24879 | 0.95029 |
| | 20 | 0.20045 | 0.20006 | 0.00035 | 0.02071 | 0.12551 | 9.99940 | 10.00014 | 0.00098 | 0.08326 | 0.24943 | 0.94987 |
| | 50 | 0.20007 | 0.20005 | -0.00005 | 0.02070 | 0.12535 | 9.99909 | 10.00009 | 0.00078 | 0.08321 | 0.25001 | 0.94960 |
| | 100 | 0.20037 | 0.19981 | -0.00008 | **0.02063** | 0.12477 | 10.00006 | 9.99986 | 0.00112 | **0.08318** | 0.25058 | 0.94858 |

Table 4: The mean and empirical MSE of Simultaneous estimators of $\alpha = 0.5$ and $\theta$ with MM and ML methods.

| $(\alpha, \theta)$ | n | $MLE\alpha$ | $MME\alpha$ | $Bias\alpha$ | $MSE\alpha$ | $MAE\alpha$ | $MLE\theta$ | $MME\theta$ | $Bias\theta$ | $MSE\theta$ | $MAE\theta$ | $C_p$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (0.5,1) | 10 | 0.50002 | 0.49995 | 0.00017 | 0.02089 | 0.12490 | 1.00011 | 1.00028 | 0.00135 | 0.08352 | 0.24938 | 0.94993 |
| | 20 | 0.50014 | 0.49981 | 0.00015 | 0.02088 | 0.12518 | 1.00073 | 1.00060 | -0.00019 | 0.08352 | 0.24997 | 0.94991 |
| | 50 | 0.49947 | 0.49989 | 0.00019 | 0.02082 | 0.12508 | 1.00074 | 0.99994 | 0.00004 | 0.08350 | 0.24947 | 0.95061 |
| | 100 | 0.50025 | 0.50032 | -0.00010 | 0.02080 | 0.12493 | 1.00192 | 1.00031 | 0.00002 | 0.08351 | 0.24935 | 0.95114 |
| (0.5,2) | 10 | 0.50005 | 0.50028 | 0.00018 | 0.02076 | 0.12470 | 2.00195 | 1.99998 | 0.00066 | 0.08347 | 0.25008 | 0.94932 |
| | 20 | 0.50007 | 0.50022 | -0.00068 | 0.02075 | 0.12501 | 1.99971 | 2.00029 | 0.00039 | 0.08346 | 0.24996 | 0.95069 |
| | 50 | 0.49966 | 0.49988 | -0.00052 | 0.02070 | 0.12513 | 1.99935 | 2.00041 | 0.00052 | 0.08346 | 0.25077 | 0.95139 |
| | 100 | 0.50004 | 0.49996 | 0.00008 | 0.02066 | 0.12531 | 1.99908 | 1.99996 | -0.00037 | 0.08343 | 0.25017 | 0.94986 |
| (0.5,5) | 10 | 0.49961 | 0.50002 | 0.00007 | 0.020765 | 0.12449 | 4.99888 | 5.00034 | 0.00049 | 0.08345 | 0.25059 | 0.94954 |
| | 20 | 0.49967 | 0.50004 | -0.00077 | 0.02061 | 0.12472 | 4.99965 | 5.00001 | 0.00097 | 0.08336 | 0.25012 | 0.94955 |
| | 50 | 0.49987 | 0.49999 | 0.00032 | 0.02056 | 0.12454 | 4.99966 | 4.99965 | -0.00188 | 0.08341 | 0.25046 | 0.94965 |
| | 100 | 0.49951 | 0.49982 | 0.00063 | 0.02053 | 0.12492 | 4.99966 | 5.00032 | 0.00031 | 0.08340 | 0.25037 | 0.94945 |
| (0.5,10) | 10 | 0.50013 | 0.50005 | 0.00048 | 0.02052 | 0.12487 | 9.99965 | 9.99988 | -0.00078 | 0.08336 | 0.25097 | 0.94958 |
| | 20 | 0.50062 | 0.49998 | 0.00078 | 0.02051 | 0.12517 | 10.00014 | 9.99941 | 0.00071 | 0.08333 | 0.25028 | 0.95000 |
| | 50 | 0.49968 | 0.49985 | -0.00001 | 0.02047 | 0.12499 | 10.00008 | 9.99995 | 0.00059 | 0.08329 | 0.24976 | 0.95005 |
| | 100 | 0.50031 | 0.49951 | 0.00046 | **0.02043** | 0.12498 | 10.00067 | 10.00005 | 0.00006 | **0.08324** | 0.24994 | 0.95049 |

Table 5: The mean and empirical MSE of Simultaneous estimators of $\alpha = 0.99$ and $\theta$ with MM and ML methods.

| $(\alpha, \theta)$ | n | $MLE\alpha$ | $MME\alpha$ | $Bias\alpha$ | $MSE\alpha$ | $MAE\alpha$ | $MLE\theta$ | $MME\theta$ | $Bias\theta$ | $MSE\theta$ | $MAE\theta$ | $C_p$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (0.99,1) | 10 | 0.98886 | 0.98973 | 0.00054 | 0.02065 | 0.12546 | 0.99946 | 0.99935 | -0.00079 | 0.08354 | 0.25048 | 0.95089 |
| | 20 | 0.98945 | 0.99014 | 0.00081 | 0.02065 | 0.12513 | 0.99845 | 1.00065 | -0.00104 | 0.08350 | 0.25007 | 0.94998 |
| | 50 | 0.98981 | 0.99017 | 0.00066 | 0.02064 | 0.12518 | 1.00028 | 1.00091 | 0.00047 | 0.08347 | 0.25012 | 0.95032 |
| | 100 | 0.98960 | 0.99035 | 0.00017 | 0.02064 | 0.12504 | 0.99825 | 0.99956 | 0.00115 | 0.08345 | 0.25041 | 0.94924 |
| (0.99,2) | 10 | 0.99082 | 0.98963 | -0.00016 | 0.02066 | 0.12528 | 2.00221 | 1.99936 | 0.00147 | 0.08244 | 0.24959 | 0.94833 |
| | 20 | 0.99042 | 0.98995 | -0.00031 | 0.02061 | 0.12517 | 1.99951 | 1.99989 | 0.00087 | 0.08342 | 0.24976 | 0.94996 |
| | 50 | 0.98976 | 0.98956 | -0.00062 | 0.02061 | 0.12472 | 2.00081 | 2.00104 | 0.00024 | 0.08340 | 0.24963 | 0.95090 |
| | 100 | 0.98943 | 0.98989 | 0.00005 | 0.02060 | 0.12509 | 2.00069 | 2.00008 | 0.00055 | 0.08336 | 0.24978 | 0.95114 |
| (0.99,5) | 10 | 0.99026 | 0.98985 | -0.00012 | 0.02058 | 0.12473 | 5.00151 | 4.99982 | 0.00034 | 0.08333 | 0.24980 | 0.94926 |
| | 20 | 0.99037 | 0.99041 | -0.00023 | 0.02052 | 0.12443 | 4.99989 | 4.99988 | 0.00033 | 0.08319 | 0.24994 | 0.94997 |
| | 50 | 0.99022 | 0.98963 | 0.00024 | 0.02050 | 0.12494 | 5.00069 | 5.00025 | 0.00009 | 0.08332 | 0.24995 | 0.95048 |
| | 100 | 0.99026 | 0.99004 | -0.00038 | 0.02049 | 0.12476 | 4.99999 | 5.00001 | 0.00010 | 0.08331 | 0.24967 | 0.94915 |
| (0.99,10) | 10 | 0.99088 | 0.99002 | 0.00090 | 0.02048 | 0.12481 | 9.99958 | 1.00002 | 0.00031 | 0.08328 | 0.24972 | 9.49350 |
| | 20 | 0.99101 | 0.99015 | 0.00029 | 0.02044 | 0.12517 | 10.00021 | 10.00013 | -0.00041 | 0.08324 | 0.25004 | 0.95053 |
| | 50 | 0.98910 | 0.99028 | 0.00017 | 0.02040 | 0.12478 | 9.99909 | 1.00001 | 0.00091 | 0.08322 | 0.25025 | 0.94993 |
| | 100 | 0.98979 | 0.99008 | -0.00064 | **0.02037** | 0.12468 | 10.00077 | 9.99996 | -0.00036 | **0.08319** | 0.24985 | 0.95013 |

## 8.2 Real data analyze

In this section, we use a real data analysis to show the validity of the introduced copula. The study data set is daily mortality rate COVID-19 data from Italy, Belgium, and Canada. The data is available at https://github.com/CSSEGISandData/COVID-19/; see also [1]. It covers the interval from 1 April to 20 August 2020. We use $X$, $Y$, and $Z$, respectively, to represent the data sets of these three countries. Summary statistics and box plot for $X$, $Y$, and $Z$ are given in Table 6 and Figure 3.

Table 6: Summary statistics of mortality COVID-19 data of Italy, Belgium, and Canada

| | Min | 1st Qu | Median | Mean | 3rd Qu | Max |
|---|---|---|---|---|---|---|
| Italy ($X$) | 0.0043 | 0.0476 | 0.1390 | 0.1362 | 0.1987 | 0.4972 |
| Belgium ($Y$) | 0.0015 | 0.0221 | 0.0930 | 0.1365 | 0.1763 | 0.8585 |
| Canada ($Z$) | 0.0070 | 0.0242 | 0.0639 | 0.0694 | 0.1034 | 0.2551 |

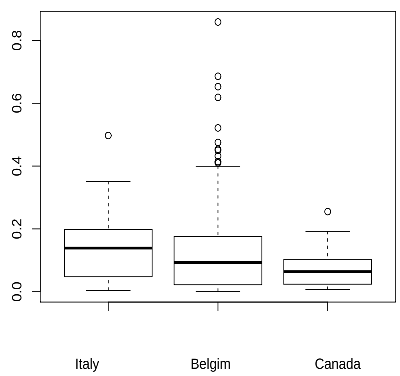


Figure 3: Box plot for mortality COVID-19 data.

If we fit distribution (1) to any of the desired data, then it is clear that the Gompertz–Makeham distribution is a suitable distribution for these data. Table 7 shows the estimation of parameters in fitting Gompertz–Makeham distribution to each of the data.

Table 7: The MLEs values of Gompertz–Makeham distribution for Italy, Belgium, and Canada

| | $\theta$ | P-value | $D$-statistic |
|---|---|---|---|
| Italy ($X$) | 6.64 | 0.32 | 0.1145 |
| Belgium ($Y$) | 4.09 | 0.16 | 0.3164 |
| Canada ($Z$) | 7.65 | 0.53 | 0.0647 |

Table 8: Correlation coefficients of mortality COVID-19 data

| | Pearson's $r$ | Spearman's $\rho$ | Kendall's $\tau$(P-value) |
|---|---|---|---|
| Italy and Belgium ($X, Y$) | 0.443 | 0.658 | 0.450(2.484e-15) |
| Italy and Canada ($X, Z$) | 0.572 | 0.624 | 0.441(8.309e-15) |
| Belgium and Canada ($Y, Z$) | 0.330 | 0.539 | 0.351(6.557e-10) |

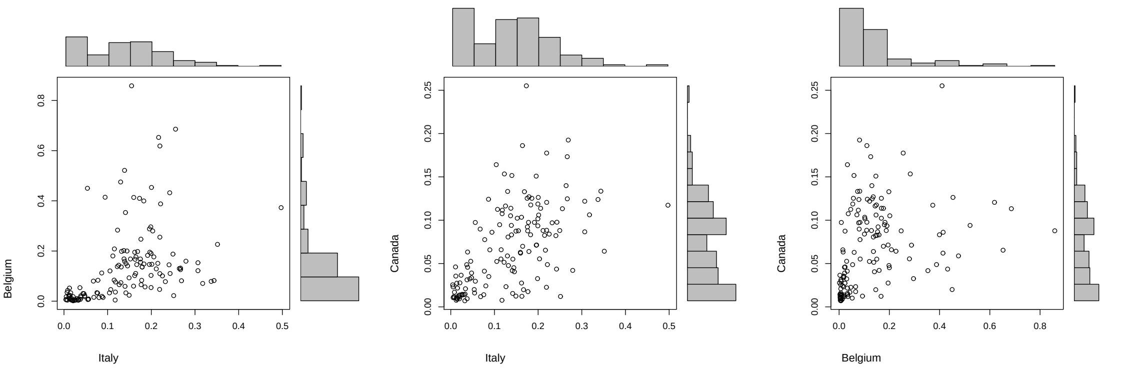


Figure 4: Scatter plot with histogram of margins for mortality COVID-19 data.

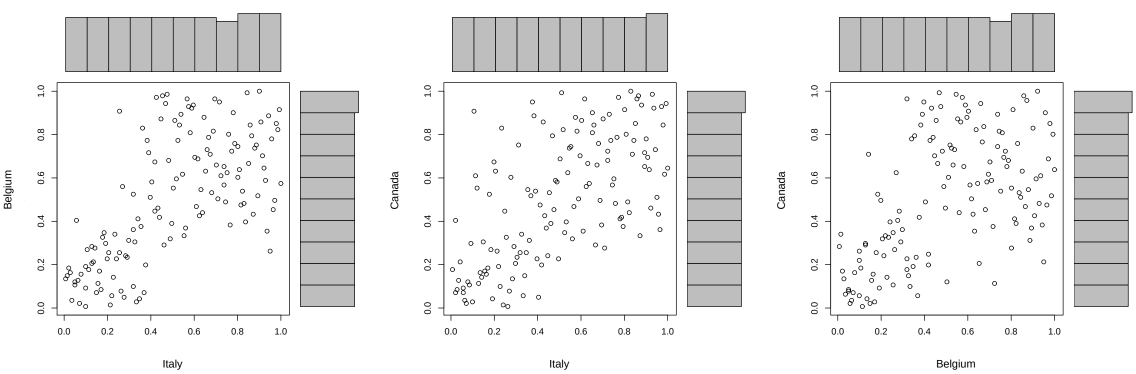


Figure 5: Scatter plot with histogram of uniform transformed data.

Figure 4 is a scatter plot of the data and shows the dependence between these variables.
Table 8 shows the values of Pearson, Spearman, and Kendall correlation coefficients between these three variables and the P-value of significance for Kendall dependence coefficient. According to the obtained values, it can be concluded that the mortality data in Italy and Belgium, as well as Italy and Canada, have a moderate dependence, and the dependence of mortality data in Belgium and Canada is weaker than that in the previous two cases. In all three cases, the dependence at the level of 0.05 is significant.
Figure 5 shows the scatter plot of uniformed data. Since the ranges of the time series are bounded in $[0, 1]$, the new scatter plot does not contain outliers like the original plot and is therefore more convenient to look at. The essential observation is that the relative orderings are maintained, that is, if two points in the original scatter plot were concordance, so are the two newly assigned points. Hence, for pure dependence-concordance measurement, it does not matter whether we apply the original plot or the transformed plot. It is now intuitively clear that measuring dependence with an eyeball check from the transformed plot might be more convenient than measuring dependence from the original scatter plot. What the performed transformation did mathematically was to standardize the marginal distributions to the uniform distribution on $[0, 1]$. The marginals were discarded, and only the pure dependence structure is maintained.

In the following, according to the goodness-of-fit procedures given in [4], we will perform the new Gompertz–Makeham copula's goodness-of-fit test for these three data sets. We will perform the famous copula's (Clayton, Gumbel, Normal, Joe, Frank, and t-student) goodness-of-fit test for these three data sets. Tables 9, 10, and 11 show the ML estimators of dependence parameters of copulas for each of the three cases studied. According to the obtained values, we conclude that $NGMC$ have the best fit over the rest of the copulas for each of the two datasets.

Table 9: The MLEs values of $NGMC$ and some copulas for Italy and Canada $(X, Z)$.

| Copula | Estimated parameter | P-value | Statistic of goodness of fit test |
|---|---|---|---|
| $NGMC$ | $\hat{\alpha} = 0.401, \quad \hat{\theta} = 4.148$ | **0.5907** | **0.0042** |
| Clayton | $\hat{\theta} = 0.76$ | 0.0435 | 0.0492 |
| Gumbel | $\hat{\theta} = 1.35$ | 0.0005 | 0.0949 |
| Normal | $\hat{\theta} = 0.44$ | 0.0045 | 0.0560 |
| Joe | $\hat{\theta} = 1.36$ | 0.0005 | 0.1928 |
| Frank | $\hat{\theta} = 3.15$ | 0.0065 | 0.0435 |
| t | $\hat{\theta} = 0.42$ | 0.0005 | 0.0653 |

Table 10: The MLEs values of $NGMC$ and some copulas for Italy and Belgium $(X, Y)$.

| Copula | Estimated parameter | P-value | Statistic of goodness of fit test |
|---|---|---|---|
| $NGMC$ | $\hat{\alpha} = 0.428, \quad \hat{\theta} = 6.714$ | **0.6871** | **0.0064** |
| Clayton | $\hat{\theta} = 0.77$ | 0.0545 | 0.0527 |
| Gumbel | $\hat{\theta} = 1.28$ | 0.0005 | 0.1294 |
| Normal | $\hat{\theta} = 0.43$ | 0.0005 | 0.0802 |
| Joe | $\hat{\theta} = 1.23$ | 0.0005 | 0.2334 |
| Frank | $\hat{\theta} = 2.86$ | 0.0005 | 0.0807 |
| t | $\hat{\theta} = 0.35$ | 0.0005 | 0.1019 |

Table 11: The MLEs values of $NGMC$ and some copulas for Belgium and Canada $(Y, Z)$.

| Copula | Estimated parameter | P-value | Statistic of goodness of fit test |
|---|---|---|---|
| $NGMC$ | $\hat{\alpha} = 0.377, \quad \hat{\theta} = 3.008$ | **0.4012** | **0.0012** |
| Clayton | $\hat{\theta} = 1.67$ | 0.0005 | 0.1080 |
| Gumbel | $\hat{\theta} = 1.84$ | 0.0005 | 0.1695 |
| Normal | $\hat{\theta} = 0.71$ | 0.0005 | 0.0917 |
| Joe | $\hat{\theta} = 1.92$ | 0.0005 | 0.3995 |
| Frank | $\hat{\theta} = 6.88$ | 0.0325 | 0.0405 |
| t | $\hat{\theta} = 0.68$ | 0.0005 | 0.1165 |

The goodness-of-fit test was performed for each of the three data sets, and the most suitable marginal distribution and its related results, are shown in Table 12.

Table 13 shows the values of these coefficients for $NGMC$ and the P-value of significance for the Kendall correlation coefficient. In all three cases, the dependence at the level of 0.05 is significant.

Remark 8.1. It is worth noting that the data of Covid-19 deaths are time series and have self-correlation. An alternative strategy to model these data is to use time series models.

Table 12: The goodness-of-fit test for mortality COVID-19 data of Italy, Belgium, and Canada

| | Distribution | P-value(Kolmogorov-Smirnov test) |
|---|---|---|
| Italy ($X$) | Dagum(0.13,7.28,0.26) | 0.344 |
| Belgium ($Y$) | Weibull(0.85,0.12) | 0.422 |
| Canada ($Z$) | Dagum(0.14,6.82,0.13) | 0.473 |

Table 13: Correlation coefficients of $NGMC$

| | Pearson's $r$ | Spearman's $\rho$ | Kendall's $\tau$(P-value) |
|---|---|---|---|
| $NGMC$ with $\hat{\alpha} = 0.428, \hat{\theta} = 6.714$ | 0.389 | 0.519 | 0.560(<0.01) |
| $NGMC$ with $\hat{\alpha} = 0.401, \hat{\theta} = 4.148$ | 0.601 | 0.598 | 0.541(0.0001) |
| $NGMC$ with $\hat{\alpha} = 0.377, \hat{\theta} = 3.008$ | 0.371 | 0.479 | 0.411(2.129e-6) |

## 8.3 Random variate generation

There are different methods to generate random observations from a pair of random variables $(X, Y)$ with a joint distribution function $H$. One of these methods is the use of the copula function and the concept of conditional probability, which is known as the conditional distribution method. In this section, we describe the method of generating a random sample of the copula function based on [18]. For this, we do the following steps, respectively:

1: Generate two independent uniform $(0, 1)$ variates $u$ and $t$;

2: Set $v = c_u^{(-1)}(t)$, where $c_u^{(-1)}$ denotes a quasi-inverse of $c_u$ and also $c_u(v) = \frac{\partial C(u,v)}{\partial u}$.

3: The desired pair is $(u, v)$.

It can be easily shown that for the $NGMC$

$$\begin{aligned} c_u(v) =& 1 - [g^{'}(u)(1+\alpha g(v)) + (\theta g^{'}(u)(1+\alpha g(v)))\exp\{-g(u) - g(v) - \alpha g(u)g(v)\}] \\ & \times \exp\{-g(u) - g(v) - \alpha g(u)g(v) - \theta(1 - \exp\{-g(u) - g(v) - \alpha g(u)g(v)\})\}, \end{aligned}$$

where in $g(u) = \psi^{-1}[\ln(1-u)]$.

Now, for each of the three states, we generate random samples of $NGMC$ with the parameters obtained in Tables 9, 10, 11 and draw the scatter plot of the generated data with their histogram. Figure 6 shows the scatter plot of these data. Figure 7 also shows the graph of the real data (black dots) as well as the corresponding generated data (red dots). As can be seen in all three cases, the graphs of real data, and simulated data are similar. Figures 4, 6 and 7 show that there is no upper tail dependence and for the $NGMC$, $\lambda_U = 0$. Also, these figures clearly show the existence of lower tail dependence and for $NGMC$, $\lambda_L \geq 0$.

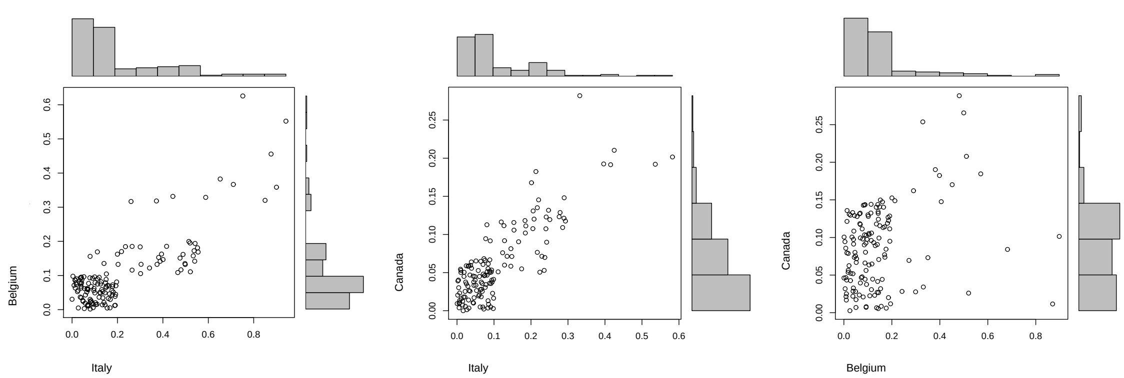


Figure 6: Scatter plot with histogram of simulated data.

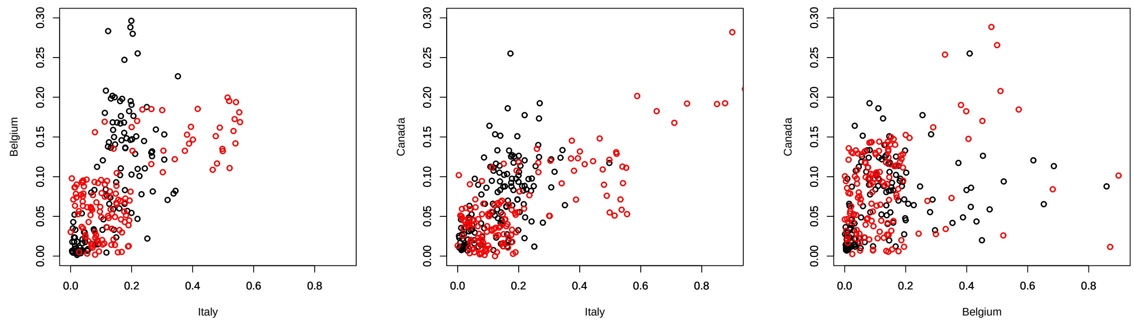


Figure 7: Scatter plot of real data (black dots) and simulated data (red dots).

# 9 Conclusions

In this article, we introduced the extended bivariate Gompertz–Makeham distribution and studied its reliability properties. We presented a new copula function and its properties based on the extended bivariate Gompertz–Makeham distribution. The dependence structure of this copula, as well as the tail dependence for it, has been studied. Some measures of dependency were also obtained for the introduced copula. We also studied the sensitivity analysis for the measures of dependency to the change of parameters. Then, by performing a simulation study, the maximum likelihood and moment estimators of the copula parameters were presented in different modes. Moreover, we used the mortality COVID-19 data to examine the copula function of the study and examined the usefulness of this copula function for modeling the studied data, and by comparing its performance with other copulas, the validity of the new copula has been shown.

# Acknowledgement

The authors are very grateful to the Associate Editor and the reviewers for their appropriate and relevant comments and suggestions which improved the presentation of the manuscript.

# References

[1] A. Algarni, A. Almarashi, I. Elbatal, A. Hassan, E. M. Almetwally, A. Daghistani, M. Elgarhy, *Type I half logistic Burr XG family: Properties, Bayesian, and non-Bayesian estimation under censored samples and applications to COVID-19 data*, Mathematical Problems in Engineering, 2021.

[2] A. Dolati, J. F. Sanchez, M. Ubeda-Flores, *A copula-based family of fuzzy implication operators*, Fuzzy Sets and Systems, **211** (2013), 55-61.

[3] N. Doodman, M. Amini, H. Jabbari, A. Dolati, *FGM generated Archimedean copulas with concave multiplicative generators*, Iranian Journal of Fuzzy Systems, **18**(2) (2021), 15-29.

[4] C. Genest, B. Remillard, D. Beaudoin, *Goodness-of-fit tests for copulas: A review and a power study*, Insurance: Mathematics and Economics, **42** (2008), 199-213.

[5] P. Grzegorzewski, *Probabilistic implications, in: EUSFLAT-2011 and LFA-2011*, Amsterdam, (2011), 254-258.

[6] P. Hájek, R. Mesiar, *On copulas, quasicopulas and fuzzy logic*, Soft Computing, **12**(12) (2008), 1239-1243.

[7] P. Helbin, M. Baczynski, P. Grzegorzewski, W. Niemyska, *Some properties of fuzzy implications based on copulas*, Information Sciences, **502** (2019), 1-17.

[8] R. Jiang, P. Ji, X. Xiao, *Aging property of unimodal failure rate models*, Reliability Engineering and System Safety, **79**(1) (2003), 113-116.

[9] H. Joe, *Dependence modeling with copulas*, CRC Press, 2014.

[10] H. Joe, J. J. Xu, *The estimation method of inference functions for margins for multivariate models*, Technical Report: Department of Statistics University of British Columbia, (1996), 166.

[11] M. E. Johnson, *Multivariate statistical simulation distributions*, New York, John Wiley and Sons, Statistics and Probability Letters, **46**(4) (1987), 359-364.

[12] N. Kolev, *Characterizations of the class of bivariate Gompertz distributions*, Journal of Multivariate Analysis, **148** (2016), 173-179.

[13] S. Kotz, M. Pensky, *The stress-strength model and its generalizations: Theory and applications*, World Scientific, 2003.

[14] C. D. Lai, M. Xie, *Stochastic ageing and dependence for reliability*, Springer Science and Business Media, New York, 2006.

[15] W. M. Makeham, *On the law of mortality and the construction of annuity tables*, Journal of the Institute of Actuaries and Assurance Magazine, **8** (1860), 301-310.

[16] A. W. Marshall, I. Olkin, *A bivariate Gompertz–Makeham life distribution*, Journal of Multivariate Analysis, **139** (2015), 219-226.

[17] H. A. Mohtashami Borzadaran, H. Jabbari, M. Amini, A. Dolati, *Stress-strength and ageing intensity analysis via a new bivariate negative Gompertz-Makeham model*, Journal of The Iranian Statistical Society, **20**(1) (2021), 219-246.

[18] R. B. Nelsen, *An introduction to copulas*, Springer-Verlag, New York, 2006.

[19] B. P. Rao, *On some analogues of lack of memory properties for the Gompertz distribution*, Communications in Statistics-Theory and Methods, **47**(18) (2018), 4415-4421.

[20] J. H. Shih, T. Emura, *Bivariate dependence measures and bivariate competing risks models under the generalized FGM copula*, Statistical Papers, **60**(4) (2019), 1101-1118.

[21] M. Sklar, *Fonctions de repartition an dimensions et leurs marges*, Publications de l'Institut de Statistique de l'Universit de Paris, **8** (1959), 229-231.

[22] C. Sun, Z. Bie, M. Xie, J. Jiang, *Fuzzy copula model for wind speed correlation and its application in wind curtailment evaluation*, Renewable Energy, **93** (2016), 68-76.

[23] W. S. B. Woolhouse, *Observations on Gompertzs law of mortality and dependence between it and Simpson's rule for finding the value of an annuity on three lives*, The Assurance Magazine, and Journal of the Institute of Actuaries, **10** (1863), 121-130.

[24] R. R. Yager, *Modeling holistic fuzzy implication using copulas*, Fuzzy Optimization and Decision Making, **5**(3) (2006), 207-226.