Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 52 (3) (2023), 808 – 827
DOI : 10.15672/hujms.976348

Research Article

Modelling wind speed with a univariate probability distribution depending on two baseline functions

Fábio Veríssimo Jaques Silveira[1], Frank Gomes-Silva*[2], Cícero Ramos de Brito[3], Jader Silva Jale[2], Felipe Ricardo Santos de Gusmão[2], Sílvio Fernando Alves Xavier-Júnior[4], João Silva Rocha[2]

[1] *Federal Institute of Education, Science and Technology of Paraíba, João Pessoa/PB, Brazil*
[2] *Federal Rural University of Pernambuco, Recife/PE, Brazil*
[3] *Federal Institute of Education, Science and Technology of Pernambuco, Recife/PE, Brazil*
[4] *Paraíba State University, Campina Grande/PB, Brazil*

Abstract

Characterizing the wind speed distribution properly is essential for the satisfactory production of potential energy in wind farms, being the mixture models usually employed in the description of such data. However, some mixture models commonly have the undesirable property of non-identifiability. In this work, we present an alternative distribution which is able to fit the wind speed data decently. The new model, called Normal-Weibull-Weibull, is identifiable and its cumulative distribution function is written as a composition of two baseline functions. We discuss structural properties of the class that generates the proposed model, such as the linear representation of the probability density function, moments and moment generating function. We perform a Monte Carlo simulation study to investigate the behavior of the maximum likelihood estimates of the parameters. Finally, we present applications of the new distribution for modelling wind speed data measured in five different cities of the Northeastern Region of Brazil.

Mathematics Subject Classification (2020). 60E05, 62E10, 62F10

Keywords. Goodness-of-fit, identifiability, L-BFGS-B algorithm, maximum likelihood, Monte Carlo simulation

*Corresponding Author.
Email addresses: fabio.silveira@ifpb.edu.br (F.V.J. Silveira), frank.gsilva@ufrpe.br (F. Gomes-Silva), cicerocarlosbrito@yahoo.com.br (C.R. Brito), jsjale1983@gmail.com (J.S. Jale), felipe556@gmail.com (F.R.S. Gusmão), silvioxj@gmail.com (S.F.A. Xavier-Júnior), joaosilvarocha@hotmail.com (J.S. Rocha)
Received: 30.09.2021; Accepted: 08.11.2022

1. Introduction

The concern about the emission of greenhouse gases and environmental contamination from conventional energy generation procedures like coal and oil power plants encourages research on alternative resources. A smaller impact on the environment is an advantage of cleaner and sustainable energy production techniques, such as solar, geothermal and wind, over the combustion of fossil fuels.

The installed capacity of wind power in Brazil increased from 29 Megawatt in 2005 to roughly 16,000 Megawatt (9% of the total capacity of electricity generation) in 2020 [3]. The suitable choice of the wind turbine must match with the wind behavior at the site of installation. Perkin *et al.* [25] mention that inadequate turbine selection results in a financially sub-optimal investment. Thus, setting the probability distribution appropriately to model the wind speed is essential. Eltamaly [11] used the two-parameter Weibull distribution in a new computer program to perform the calculations required to precisely design the wind energy system and to seek the compatibility between sites and turbines.

Ilhan and Kantar [15] declare that despite the wide acceptance of Weibull distribution [18, 26, 29, 30], it may sometimes be poor to model all wind speed data available in nature. Hereupon, they propose using two possible models, namely, the skewed generalized error distribution [5] and the skewed t distribution [13]. Some other distributions used for wind speed and power modelling are Rayleigh [26], gamma [21], normal [29], generalized extreme value [18] and Birnbaum-Saunders [20]. Additionally, applications of nonparametric methods to wind speed modelling are also found in the literature [12, 14, 27].

Oftentimes, one requires more flexibility from the probability density function (pdf), as in case of bimodality [17] or calm winds regime [9]. In general, finite mixture models are more flexible than the typical single ones. Akdag *et al.* [1] compared the usual biparametric Weibull and the two-component mixture Weibull distribution in a study focused on wind regimes presenting nearly zero percentage of null speeds; they concluded that the mixture is more suitable to describe such wind conditions. Carta and Ramírez [8] used three different methods to estimate the parameters of the two-component mixture Weibull, namely, the method of moments, maximum likelihood and least squares; they verified that there is no significant difference among them.

The mixture density can be written as:

$$f(x;\boldsymbol{\psi}) = \sum_{i=1}^{d} w_i f_i(x;\boldsymbol{\theta}_i) \tag{1.1}$$

where the vector $\boldsymbol{\psi} = (w_1, \ldots, w_{d-1}, \boldsymbol{\eta}^\top)^\top$ contains the unknown parameters of the mixture model and the vector $\boldsymbol{\eta}$ contains all the distinct parameters in $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_d$. Since $w_1, \ldots, w_d$ are positive and sum up to one, the presence of w_d in $\boldsymbol{\psi}$ is unnecessary. The general definition of identifiability states that a family of densities $\{f(x;\boldsymbol{\psi}) : \boldsymbol{\psi} \in \boldsymbol{\Psi}\}$ is identifiable if:

$$f(x;\boldsymbol{\psi}) = f(x;\boldsymbol{\psi}^\star) \Leftrightarrow \boldsymbol{\psi} = \boldsymbol{\psi}^\star\,. \tag{1.2}$$

It is not seldom that (1.2) fails when two or more component densities in (1.1) belong to the same parametric family. Such is the case of the mixture of normal densities. Consider $d = 2$, f_1 and f_2 are normal densities, $w_1 = 0.5$ in (1.1) and let $\boldsymbol{\eta}_1 = (\mu_1, \sigma_1, \mu_2, \sigma_2)^\top$ and $\boldsymbol{\eta}_2 = (\mu_2, \sigma_2, \mu_1, \sigma_1)^\top$, where $\mu_1 \neq \mu_2$, $\sigma_1 \neq \sigma_2$. We have $\boldsymbol{\psi} = (w_1, \boldsymbol{\eta}_1^\top)^\top \neq (w_1, \boldsymbol{\eta}_2^\top)^\top = \boldsymbol{\psi}^\star \Rightarrow f(x;\boldsymbol{\psi}) = f(x;\boldsymbol{\psi}^\star)$. That is, (1.1) may be invariant under certain permutations of the elements in the parametric vector. McLachlan and Peel [19] mention an alternative definition of identifiability for mixture models, such that the mixture of d normal densities would be identifiable under specific conditions. Nonetheless, they remark that it does not overcome the complications due to the interchanging of component labels.

Models like the mixture of normal or Weibull densities are quite flexible tools, although the parametric estimation is only credible if the distribution under study is identifiable. We present in this paper a class, whose submodels may be feasible alternatives to mixtures of two components from the same parametric family. The class is derived using the method of generating classes of probability distributions of [6]. Its cumulative distribution function (cdf) is formulated as a composition of two baselines and under certain conditions, it satisfies (1.2), even if both baselines belong to the same parametric family.

The cdf of the general class is given by:

$$F(x) = \zeta(x) \sum_{j=1}^{n} \int_{L_j(x)}^{U_j(x)} \mathrm{d}H(t) - \nu(x) \sum_{j=1}^{n} \int_{M_j(x)}^{V_j(x)} \mathrm{d}H(t) \tag{1.3}$$

where H is a cdf, $n \in \mathbb{N}$, $\zeta, \nu : \mathbb{R} \mapsto \mathbb{R}$ and $L_j, U_j, M_j, V_j : \mathbb{R} \mapsto \mathbb{R} \cup \{\pm\infty\}$ are special functions that will be discussed in the next section.

The major contributions of this work are cited below:

- Few classes were created based on the normal distribution, like the notable one proposed by [4]. The class presented here is based on the normal distribution, which plays a significant role in statistical theory.
- The new class does not demand the inclusion of additional parameters. In this way, parsimony is taken for granted, since the only parameters in the generated submodels are brought by the baselines.
- Since the proposed class was created using the multibaseline method [6], it is possible to generate distributions using either continuous or discrete baselines.
- Parametric estimation of convex mixture models can be tricky and oftentimes requires a substantial computational cost, especially when the expectation maximization (EM) algorithm [24] is in action. In contrast, the proposed class can be a feasable alternative to describe data usually modelled by two-component mixture models. The estimation of the parameters of a submodel generated by the aforementioned class can be carried out using a method of optimization simpler than the EM algorithm, such as the limited memory Broyden-Fletcher-Goldfarb-Shanno with boundaries (L-BFGS-B) algorithm [7].
- A theorem establishing necessary conditions to assure the identifiability of generated distributions is formulated and demonstrated (2.1).

2. The Normal-(G_1, G_2) class and some structural properties

The method [6] states that if $H, \zeta, \nu : \mathbb{R} \mapsto \mathbb{R}$ and $L_j, U_j, M_j, V_j : \mathbb{R} \mapsto \mathbb{R} \cup \{\pm\infty\}$ for $j = 1, 2, 3, \ldots, n$ are monotonic and right continuous functions such that:

(c1) H is a cdf and ζ and ν are non-negative;

(c2) $\zeta(x)$, $U_j(x)$ and $M_j(x)$ are non-decreasing and $\nu(x)$, $V_j(x)$, $L_j(x)$ are non-increasing $\forall j = 1, 2, 3, \ldots, n$;

(c3) If $\lim_{x\to-\infty} \zeta(x) \neq \lim_{x\to-\infty} \nu(x)$, then $\lim_{x\to-\infty} \zeta(x) = 0$; **or**
$\lim_{x\to-\infty} U_j(x) = \lim_{x\to-\infty} L_j(x)\, \forall j = 1, 2, 3, \ldots, n$, and $\lim_{x\to-\infty} \nu(x) = 0$; **or**
$\lim_{x\to-\infty} M_j(x) = \lim_{x\to-\infty} V_j(x)\, \forall j = 1, 2, 3, \ldots, n$;

(c4) If $\lim_{x\to-\infty} \zeta(x) = \lim_{x\to-\infty} \nu(x) \neq 0$, then $\lim_{x\to-\infty} U_j(x) = \lim_{x\to-\infty} V_j(x)$ and
$\lim_{x\to-\infty} M_j(x) = \lim_{x\to-\infty} L_j(x)\, \forall j = 1, 2, 3, \ldots, n$;

(c5) $\lim_{x\to-\infty} L_j(x) \leq \lim_{x\to-\infty} U_j(x)$ and if $\lim_{x\to-\infty} \nu(x) \neq 0$, then
$\lim_{x\to+\infty} M_j(x) \leq \lim_{x\to+\infty} V_j(x)\, \forall j = 1, 2, 3, \ldots, n$;

(c6) $\lim_{x\to+\infty} U_n(x) \geq \sup\{x \in \mathbb{R} : H(x) < 1\}$ and $\lim_{x\to+\infty} L_1(x) \leq \inf\{x \in \mathbb{R} : H(x) > 0\}$;

(c7) $\lim_{x\to+\infty} \zeta(x) = 1$;

(c8) $\lim_{x\to+\infty} \nu(x) = 0$ or $\lim_{x\to+\infty} M_j(x) = \lim_{x\to+\infty} V_j(x)\, \forall j = 1, 2, 3, \ldots, n$ and $n \geq 1$;

(c9) $\lim_{x\to+\infty} U_j(x) = \lim_{x\to+\infty} L_{j+1}(x)\, \forall j = 1, 2, 3, \ldots, n-1$ and $n \geq 2$;

(c10) H is a cdf without points of discontinuity or all functions $L_j(x)$ and $V_j(x)$ are constant at the right of the vicinity of points whose image are points of discontinuity of H, being also continuous in that points. Moreover, H does not have any point of discontinuity in the set

$$\left\{ \lim_{x\to\pm\infty} L_j(x), \lim_{x\to\pm\infty} U_j(x), \lim_{x\to\pm\infty} M_j(x), \lim_{x\to\pm\infty} V_j(x) \right\} \text{ for some } j = 1, 2, 3, \ldots, n;$$

then Equation (1.3) is a cdf.

Let $n = 1$, $H(t) = \Phi(t)$, namely, the standard normal cdf, $\zeta(x) = 1$, $\nu(x) = 0$, $U_1(x) = G_1(x)/[1 - G_1(x)]$ and $L_1(x) = \log[1 - G_2(x)]$, where $G_1(x)$ and $G_2(x)$ are cdfs. The function in Equation (1.3) turns into:

$$F_{G_1,G_2}(x) = \int_{\log[1-G_2(x)]}^{\frac{G_1(x)}{1-G_1(x)}} \mathrm{d}\Phi(t). \tag{2.1}$$

We took U_1 and $-L_1$ from the table of differentiable and monotonically non-decreasing functions presented in the well-known paper [2], whose method was used to create generalized distributions of the T-X family. We have intentionally picked the two simplest functions from the cited table; alternative (and more complicated) choices for U_1 and L_1 would naturally give rise to different classes. Defining $M_1(x)$ and $V_1(x)$ is not relevant, since $\nu(x) = 0$. Also, for obvious reasons, there is no need to verify (c4), (c5) and (c9). The conditions (c1), (c7), (c8) and (c10) are straightforward. As $U_1(x)$ and $\zeta(x)$ are non-decreasing and $L_1(x)$ is non-increasing, (c2) is true. It is easy to verify that $\lim_{x\to-\infty} U_1(x) = 0 = \lim_{x\to-\infty} L_1(x)$; and since $\lim_{x\to-\infty} \nu(x) = 0$, (c3) is satisfied. The condition (c6) is also true because $\lim_{x\to+\infty} U_1(x) = +\infty = \sup\{x \in \mathbb{R} : \Phi(x) < 1\}$ and $\lim_{x\to+\infty} L_1(x) = -\infty = \inf\{x \in \mathbb{R} : \Phi(x) > 0\}$.

Thereby, in agreement with the method exposed above, Equation (2.1) is a cdf. As already mentioned, it can be viewed as a composite function of two baselines. Henceforth, let it be denoted by Normal-(G_1, G_2) class of probability distributions.

Since $\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$, and $\Phi(x) = \int_{-\infty}^{x} \phi(t)\mathrm{d}t$, one can write Equation (2.1) as follows:

$$F_{G_1,G_2}(x) = \Phi\left(\frac{G_1(x)}{1 - G_1(x)}\right) - \Phi\left(\log[1 - G_2(x)]\right). \tag{2.2}$$

In case of continuous $G_1(x)$ and $G_2(x)$, one can take the derivative of Equation (2.2) with respect to x to obtain the following pdf:

$$f_{G_1,G_2}(x) = \phi\left(\frac{G_1(x)}{1 - G_1(x)}\right) \frac{g_1(x)}{[1 - G_1(x)]^2} + \phi\left(\log[1 - G_2(x)]\right) \frac{g_2(x)}{1 - G_2(x)}, \tag{2.3}$$

where $g_i(x)$ is the pdf of the random variable whose cdf is $G_i(x)$, for $i \in \{1, 2\}$.

At this point, we need to define properly the support of the distributions that emerge from the new class. Submodels of classes that may be written as a composite function of one single baseline usually have the same support of the baseline. However, characterizing the support of a submodel from (2.1) is not so straightforward, especially if the two baselines have different supports. As previously mentioned, given that $U_1(G_1(x), G_2(x)) = G_1(x)/[1 - G_1(x)]$, $L_1(G_1(x), G_2(x)) = \log[1 - G_2(x)]$ and $S_H = \mathbb{R}$, namely, the support of $H(t)$ is $\mathbb{R}$, we have that:

(a) S_H is a convex set;

(b) $U_1(1, 1) = U_1(G_1(+\infty), G_2(+\infty)) = +\infty = \sup\{x \in \mathbb{R} : \Phi(x) < 1\}$, $L_1(1, 1) = L_1(G_1(+\infty), G_2(+\infty)) = -\infty = \inf\{x \in \mathbb{R} : \Phi(x) > 0\}$, $U_1(G_1(x), G_2(x))$ and $L_1(G_1(x), G_2(x))$ are monotonic functions.

According to the Theorem (T4) in [6], (a) and (b) entail that the support of a distribution from (2.1) is the union of the supports of G_1 and G_2.

In the following lines, we demonstrate that, under specific conditions, the distributions generated by (2.2) enjoy the attractive property of identifiability. It is important because it assures the uniqueness of the estimates of the parameters.

Theorem 2.1. *Let $G_1(x|\boldsymbol{\theta}_1)$ and $G_2(x|\boldsymbol{\theta}_2)$ be the baseline cdfs of the normal-(G_1, G_2) cdf $F_{G_1,G_2}(x|\boldsymbol{\theta})$ (2.2), $\boldsymbol{\theta}_1 = (\theta_1, \ldots, \theta_r) \in \boldsymbol{\Theta}_1$, $\boldsymbol{\theta}_2 = (\theta_{r+1}, \ldots, \theta_{r+m}) \in \boldsymbol{\Theta}_2$ and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_r, \theta_{r+1}, \ldots, \theta_{r+m}) \in \boldsymbol{\Theta}$, where $\boldsymbol{\Theta}_1$, $\boldsymbol{\Theta}_2$ and $\boldsymbol{\Theta}$ are the parametric spaces associated with G_1, G_2 and F_{G_1,G_2} respectively. If G_1 and G_2 are identifiable, then F_{G_1,G_2} is identifiable.*

Proof. Assume that $\Phi\left(\frac{G_1(x|\boldsymbol{\theta}_1)}{1-G_1(x|\boldsymbol{\theta}_1)}\right) = \Phi\left(\frac{G_1(x|\boldsymbol{\theta}_1^\star)}{1-G_1(x|\boldsymbol{\theta}_1^\star)}\right)$,

where $\{\boldsymbol{\theta}_1, \boldsymbol{\theta}_1^\star\} \subset \boldsymbol{\Theta}_1$ and $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_1^\star$. Since Φ is injective, $\frac{G_1(x|\boldsymbol{\theta}_1)}{1-G_1(x|\boldsymbol{\theta}_1)} = \frac{G_1(x|\boldsymbol{\theta}_1^\star)}{1-G_1(x|\boldsymbol{\theta}_1^\star)} \Rightarrow G_1(x|\boldsymbol{\theta}_1) = G_1(x|\boldsymbol{\theta}_1^\star)$; it is a contradiction, because it denies the identifiability of G_1. Therefore, if $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_1^\star$ then $\Phi\left(\frac{G_1(x|\boldsymbol{\theta}_1)}{1-G_1(x|\boldsymbol{\theta}_1)}\right) \neq \Phi\left(\frac{G_1(x|\boldsymbol{\theta}_1^\star)}{1-G_1(x|\boldsymbol{\theta}_1^\star)}\right)$. Analogously, it is easy to verify that for $\{\boldsymbol{\theta}_2, \boldsymbol{\theta}_2^\star\} \subset \boldsymbol{\Theta}_2$, if $\boldsymbol{\theta}_2 \neq \boldsymbol{\theta}_2^\star$ then $\Phi\left(\log[1 - G_2(x|\boldsymbol{\theta}_2)]\right) \neq \Phi\left(\log[1 - G_2(x|\boldsymbol{\theta}_2^\star)]\right)$.

Now consider $\{\boldsymbol{\theta}, \boldsymbol{\theta}^\star\} \subset \boldsymbol{\Theta}$ such that $\boldsymbol{\theta} \neq \boldsymbol{\theta}^\star$ and assume that $F_{G_1,G_2}(x|\boldsymbol{\theta}) = F_{G_1,G_2}(x|\boldsymbol{\theta}^\star)$. If $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_1^\star$ and $\boldsymbol{\theta}_2 \neq \boldsymbol{\theta}_2^\star$, then we can infer from (2.2) that $G_2(x|\boldsymbol{\theta}_2) = G_2(x|\boldsymbol{\theta}_2^\star)$, namely, an absurd. Likewise, if $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_1^\star$ and $\boldsymbol{\theta}_2 = \boldsymbol{\theta}_2^\star$, we get to similar contradiction. If $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_1^\star$ and $\boldsymbol{\theta}_2 \neq \boldsymbol{\theta}_2^\star$, then the assumption fails since $F_{G_1,G_2}(x|\boldsymbol{\theta}) \neq F_{G_1,G_2}(x|\boldsymbol{\theta}^\star)$ for almost all values of x within the support. Therefore, F_{G_1,G_2} is identifiable. □

2.1. Series representation

The normal cdf can be written in terms of the error function erf as follows:

$$\Phi(z) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{z}{\sqrt{2}}\right)\right], \tag{2.4}$$

where $\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}\mathrm{d}t$. Since $\operatorname{erf}(z/\sqrt{2})$ may be linearly represented by:

$$\begin{aligned}\operatorname{erf}\left(\frac{z}{\sqrt{2}}\right) &= \frac{2}{\sqrt{\pi}}\sum_{n=0}^{\infty}\frac{(-1)^n\cdot(z/\sqrt{2})^{2n+1}}{n!(2n+1)}\\ &= \sqrt{\frac{2}{\pi}}\cdot\sum_{n=0}^{\infty}\left(-\frac{1}{2}\right)^n\frac{z^{2n+1}}{n!(2n+1)},\end{aligned} \tag{2.5}$$

replacing Equation (2.5) in Equation (2.4), we get to:

$$\Phi(z) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}}\sum_{n=0}^{\infty}\left(-\frac{1}{2}\right)^n\frac{z^{2n+1}}{n!(2n+1)}. \tag{2.6}$$

Now using the result of Equation (2.6) in Equation (2.2), we have:

$$F_{G_1,G_2}(x) = \sum_{n=0}^{\infty}\frac{(-1/2)^n}{n!(2n+1)\sqrt{2\pi}}\left[\underbrace{\left(\frac{G_1(x)}{1-G_1(x)}\right)^{2n+1}}_{\text{A1}} - \underbrace{\left(\log[1-G_2(x)]\right)^{2n+1}}_{\text{A2}}\right]. \tag{2.7}$$

A well-known result on power series raised to a positive integer N states that:

$$\left(\sum_{k=0}^{\infty}a_k y^k\right)^N = \sum_{k=0}^{\infty}c_k y^k, \tag{2.8}$$

where $c_0 = a_0^N$, $c_k = \frac{1}{ka_0}\sum_{s=1}^{k}(sN-k+s)a_s c_{k-s}$ for $k \geq 1$ and $N \in \mathbb{N}$. Setting $N = 2n+1$, $y = G_1(x)$ and $a_k = 1$ for all $k \geq 0$, we can use the result in Equation (2.8) to rewrite A1 in Equation (2.7):

$$\begin{aligned} \text{A1} &= G_1(x)^{2n+1}\left(\frac{1}{1-G_1(x)}\right)^{2n+1} = G_1(x)^{2n+1}\left(\sum_{k=0}^{\infty} G_1(x)^k\right)^{2n+1} \\ &= G_1(x)^{2n+1}\sum_{k=0}^{\infty} c_{1,k}\cdot G_1(x)^k = \sum_{k=0}^{\infty} c_{1,k}\cdot G_1(x)^{k+2n+1}\,, \end{aligned} \tag{2.9}$$

such that $c_{1,0} = 1$ and $c_{1,k} = \frac{1}{k}\sum_{s=1}^{k}(2s[n+1]-k)c_{1,k-s}$ for $k \geq 1$. Equation (2.8) also allows us to rewrite A2 in Equation 2.7 as follows:

$$\begin{aligned} \text{A2} &= \left(-\sum_{m=1}^{\infty}\frac{G_2(x)^m}{m}\right)^{2n+1} = -\left(\sum_{k=0}^{\infty}\frac{G_2(x)^{k+1}}{k+1}\right)^{2n+1} \\ &= -G_2(x)^{2n+1}\left(\sum_{k=0}^{\infty}\frac{G_2(x)^k}{k+1}\right)^{2n+1} = -G_2(x)^{2n+1}\sum_{k=0}^{\infty} c_{2,k}\cdot G_2(x)^k \\ &= -\sum_{k=0}^{\infty} c_{2,k}\cdot G_2(x)^{k+2n+1} \end{aligned} \tag{2.10}$$

where $c_{2,0} = 1$ and $c_{2,k} = \frac{1}{k}\sum_{s=1}^{k}\frac{2s(n+1)-k}{s+1}c_{2,k-s}$ for $k \geq 1$. Now inserting (2.9) and (2.10) in (2.7), we have:

$$F_{G_1,G_2}(x) = \sum_{i=1}^{2}\sum_{n,k=0}^{\infty} c_{i,n,k}\cdot G_i(x)^{k+2n+1} \tag{2.11}$$

where $c_{1,n,k} = \frac{(-1/2)^n}{n!(2n+1)\sqrt{2\pi}}c_{1,k}$ and $c_{2,n,k} = \frac{(-1/2)^n}{n!(2n+1)\sqrt{2\pi}}c_{2,k}$. Using Fubini's theorem on differentiation we can write the derivative of (2.11) as follows:

$$f_{G_1,G_2}(x) = \sum_{i=1}^{2}\sum_{n,k=0}^{\infty} c_{i,n,k}\cdot g_{i,k+2n+1}(x) \tag{2.12}$$

where $g_{i,k+2n+1}(x) = (k+2n+1)g_i(x)G_i(x)^{k+2n}$ is the pdf of a random variable from the exponentiated family [22]. Thus, we can say that (2.12) is the Normal-(G_1, G_2) pdf (2.3) expressed as a linear combination of pdfs of exponentiated distributions.

2.2. Raw moments, incomplete moments and moment generating function

Given that X is a random variable following a distribution from the normal-(G_1, G_2) class, we can use (2.12) to write the r-th raw moment of X as follows:

$$E(X^r) = \sum_{i=1}^{2}\sum_{n,k=0}^{\infty} c_{i,n,k}\int_{-\infty}^{\infty} x^r g_{i,k+2n+1}(x)\mathrm{d}x \tag{2.13}$$

$$= \sum_{i=1}^{2}\sum_{n,k=0}^{\infty} c_{i,n,k}E(Y_{i,k+2n+1}^r) \tag{2.14}$$

where $Y_{i,k+2n+1}$ follows the exponentiated distribution whose pdf is $g_{i,k+2n+1}$.

Let Q_i be the quantile function of the baseline G_i. Replacing x in (2.13) by $Q_i\left(v^{1/k+2n+1}\right)$ we can also represent (2.14) as:

$$E(X^r) = \sum_{i=1}^{2} \sum_{n,k=0}^{\infty} c_{i,n,k} \int_0^1 \left[Q_i\left(v^{1/k+2n+1}\right)\right]^r \mathrm{d}v\,.$$

Similarly, one can write the r-th incomplete moment of X as follows:

$$\begin{aligned} m_r(z) &= \int_{-\infty}^{z} x^r f_{G_1,G_2}(x)\mathrm{d}x = \sum_{i=1}^{2} \sum_{n,k=0}^{\infty} c_{i,n,k} m_r^{\star}(z) \\ &= \sum_{i=1}^{2} \sum_{n,k=0}^{\infty} c_{i,n,k} \int_0^{[G_i(z)]^{k+2n+1}} \left[Q_i\left(v^{1/k+2n+1}\right)\right]^r \mathrm{d}v \end{aligned}$$

where $m_r^{\star}(z)$ is the r-th incomplete moment of $Y_{i,k+2n+1}$ mentioned above.

The moment generating function (mgf) of X is denoted by $M_X(t) = E\left(e^{tX}\right)$. It can be determined from (2.12) as:

$$\begin{aligned} M_X(t) &= \sum_{i=1}^{2} \sum_{n,k=0}^{\infty} c_{i,n,k} \int_{-\infty}^{\infty} e^{tx} g_{i,k+2n+1}(x)\mathrm{d}x \\ &= \sum_{i=1}^{2} \sum_{n,k=0}^{\infty} c_{i,n,k} M_{Y_{k+2n+1}}(t)\,, \end{aligned}$$

where $M_{Y_{k+2n+1}}(t)$ is the mgf of $Y_{i,k+2n+1}$.

2.3. Estimation and inference

Let $\boldsymbol{X} = (x_1, \ldots, x_n)$ be a complete random sample of size n from the random variable $X \sim$ normal-(G_1, G_2). Given that $\boldsymbol{\theta}_1 = (\theta_1, \ldots, \theta_r)^\top$ is the $r \times 1$ parametric vector associated with $G_1(x) = G_1(x|\boldsymbol{\theta}_1)$, $\boldsymbol{\theta}_2 = (\theta_{r+1}, \ldots, \theta_{r+m})^\top$ is the $m \times 1$ parametric vector associated with $G_2(x) = G_2(x|\boldsymbol{\theta}_2)$ and $f_{G_1,G_2}(x) = f_{G_1,G_2}(x|\boldsymbol{\theta})$ where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_r, \theta_{r+1}, \ldots, \theta_{r+m})^\top$, we can write the log-likelihood function of X as follows:

$$\ell(\boldsymbol{\theta}|\boldsymbol{X}) = \sum_{i=1}^{n} \log\left\{\phi\left(\frac{G_1(x_i)}{1-G_1(x_i)}\right)\frac{g_1(x_i)}{[1-G_1(x_i)]^2} + \phi\left(\log[1-G_2(x_i)]\right)\frac{g_2(x_i)}{1-G_2(x_i)}\right\}.$$

The solution of the system of equations $U(\boldsymbol{\theta}|\boldsymbol{X}) = \mathbf{0}_{r+m}$ provides the maximum likelihood estimates (MLEs) for $\boldsymbol{\theta}$, where $\mathbf{0}_{r+m}$ is an $(r+m) \times 1$ vector of zeros and $U(\boldsymbol{\theta}|\boldsymbol{X}) = \nabla_{\boldsymbol{\theta}}\ell(\boldsymbol{\theta}|\boldsymbol{X})$ is the score vector. The elements of $U(\boldsymbol{\theta}|\boldsymbol{X}) = (u_j)_{1\le j\le r+m}$ are:

$$\begin{aligned} u_j = \sum_{i=1}^{n} \frac{1}{f_{G_1,G_2}(x_i)}\phi\left(\frac{G_1(x_i)}{1-G_1(x_i)}\right)\frac{1}{(1-G_1(x_i))^2}\Bigg[&\frac{\partial}{\partial\theta_j}g_1(x_i) \\ &+\frac{g_1(x_i)}{1-G_1(x_i)}\left(2-\frac{G_1(x_i)}{[1-G_1(x_i)]^2}\right)\frac{\partial}{\partial\theta_j}G_1(x_i)\Bigg], \text{ for } 1 \le j \le r \end{aligned}$$

and

$$u_j = \sum_{i=1}^{n} \frac{1}{f_{G_1,G_2}(x_i)} \frac{\phi\left(\log[1-G_2(x_i)]\right)}{1-G_2(x_i)} \left[\frac{\partial}{\partial\theta_j} g_2(x_i) + \left(1+\log[1-G_2(x_i)]\right) \frac{g_2(x_i)}{1-G_2(x_i)} \frac{\partial}{\partial\theta_j} G_2(x_i) \right], \text{ for } r < j \leq r+m\,.$$

For testing hypotheses and constructing confidence intervals for $\boldsymbol{\theta}$, the information matrix $J(\boldsymbol{\theta}|\boldsymbol{X})$ is needed. The expectation of $J(\boldsymbol{\theta}|\boldsymbol{X})$, denoted by $\mathfrak{I}_{\boldsymbol{\theta}}$, is the expected Fisher information matrix. Given that certain conditions of regularity are fulfilled, the quantity $\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ follows approximately a multivariate normal distribution $N_{r+m}(\mathbf{0}_{r+m}, \mathfrak{I}_{\boldsymbol{\theta}}^{-1})$. The appendix A brings the expression for $J(\boldsymbol{\theta}|\boldsymbol{X})$.

3. The proposed model

The Weibull cdf is given by $G_W(x|k,\lambda) = 1 - \mathrm{e}^{-(x/\lambda)^k}$, for $x \geq 0$, $k > 0$ and $\lambda > 0$. Replacing G_1 and G_2 in (2.2) by $G_W(x|k_1,\lambda_1)$ and $G_W(x|k_2,\lambda_2)$ respectively, we get to the cdf of the Normal-Weibull-Weibull distribution (NWW, for short):

$$F_{NWW}(x|\boldsymbol{\theta}) = \Phi\left(\mathrm{e}^{(x/\lambda_1)^{k_1}} - 1\right) - \Phi\left(-\left(\frac{x}{\lambda_2}\right)^{k_2}\right),$$

where $\boldsymbol{\theta} = (k_1, \lambda_1, k_2, \lambda_2)^\top$. The corresponding pdf can be obtained using (2.3):

$$f_{NWW}(x|\boldsymbol{\theta}) = \phi\left(\mathrm{e}^{(x/\lambda_1)^{k_1}} - 1\right) \frac{k_1}{\lambda_1}\left(\frac{x}{\lambda_1}\right)^{k_1-1} \mathrm{e}^{(x/\lambda_1)^{k_1}} + \phi\left(-\left(\frac{x}{\lambda_2}\right)^{k_2}\right) \frac{k_2}{\lambda_2}\left(\frac{x}{\lambda_2}\right)^{k_2-1}.$$

The associated hazard rate function (hrf) becomes

$$h_{NWW}(x|\boldsymbol{\theta}) = \frac{\phi\left(\mathrm{e}^{(x/\lambda_1)^{k_1}} - 1\right) \frac{k_1}{\lambda_1}\left(\frac{x}{\lambda_1}\right)^{k_1-1} \mathrm{e}^{(x/\lambda_1)^{k_1}} + \phi\left(-\left(\frac{x}{\lambda_2}\right)^{k_2}\right) \frac{k_2}{\lambda_2}\left(\frac{x}{\lambda_2}\right)^{k_2-1}}{1 - \left\{\Phi\left(\mathrm{e}^{(x/\lambda_1)^{k_1}} - 1\right) - \Phi\left(-\left(\frac{x}{\lambda_2}\right)^{k_2}\right)\right\}}.$$

Figure 1 displays some plots of the NWW pdf for different values of the parameters. The distribution is able to fit unimodal right-skewed data (top-left chart) and also left-skewed data (top-right chart). Notice the different shapes of the bimodal curves in the remaining charts. For instance, in the bottom-left chart, the vertical distance between the modes and the local minimum in the purple curve is much greater than in the green one. We may also notice that λ_1 and λ_2 somehow behave like shape parameters, as in the original Weibull baselines, controlling the shape of the "bells" (compare purple and gray curves to see the effect of varying λ_1; same for blue and red curves concerning λ_2).

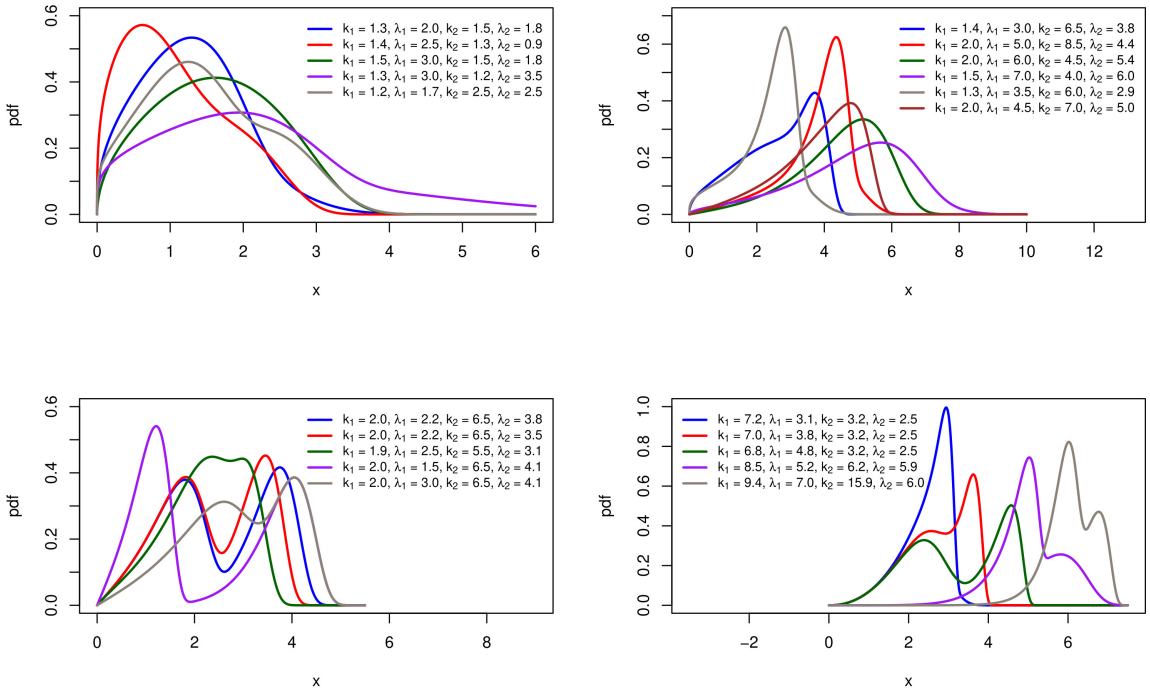

Figure 1. Plots of the NWW pdf

The hrf (Fig.2) can accommodate increasing, decreasing, S-shaped, upside-down bathtub, bathtub, and other shapes.

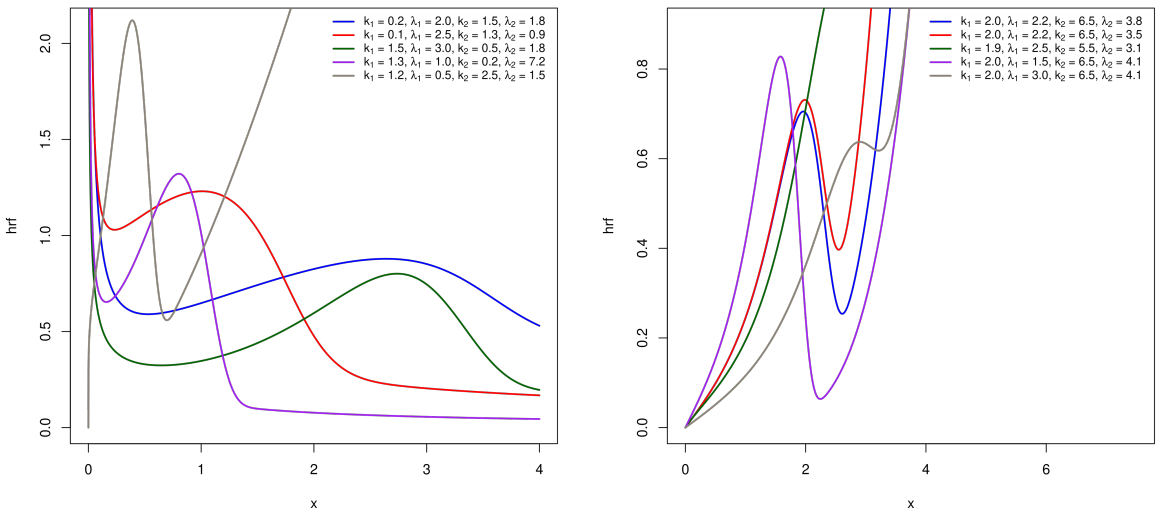

Figure 2. The hrf of the NWW distribution

Figures 3 and 4 show the behavior of skewness and kurtosis of the NWW distribution for some pairs of parameters.

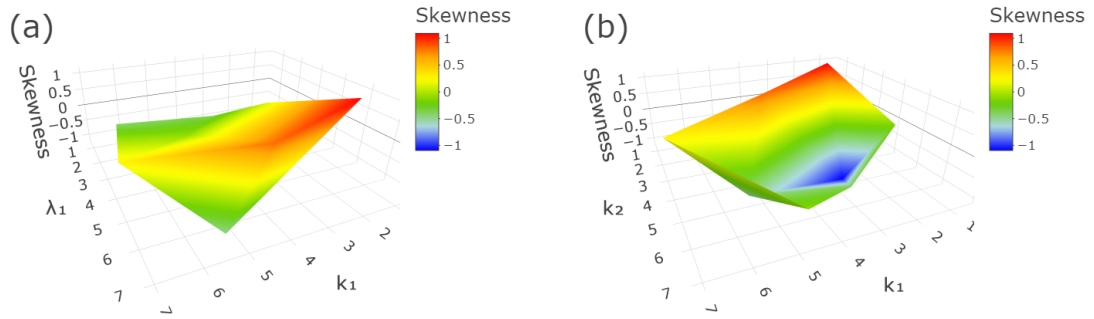

Figure 3. The skewness of the NWW distribution

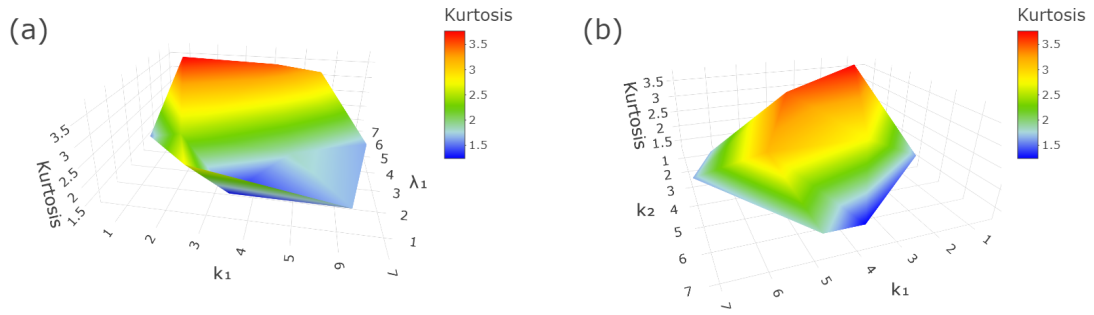

Figure 4. The kurtosis of the NWW distribution

4. Simulation

Performing Monte Carlo simulation studies is considerably relevant whenever one wants to test and confirm assumptions on new statistical methods. In this work, we want to investigate the behavior of the estimates of the parameters of the NWW distribution under the method of maximum likelihood. For this purpose, we used the software R version 3.4.4 [28].

Initially, we employed the Von Neumann's acceptance-rejection method [23] to generate pseudo-random samples from the NWW distribution; this simple method requires only the corresponding pdf $y = f(x)$, a minorant and a majorant for x and a majorant for y. The procedure was replicated 10,000 times and at each replication, six different sample sizes were considered, namely, $n = 15$, 30, 50, 100, 200 and 500. We examined scenarios with four different values of the parametric vector $\boldsymbol{\theta} = (k_1, \lambda_1, k_2, \lambda_2)^\top$, which are presented from the second to fifth columns of Tables 1, 2 and 3.

For each scenario, we calculated the bias and the mean squared error (MSE) as follows:

$$\text{Bias}_i = \frac{1}{10000}\sum_{j=1}^{10000}\left(\widehat{\theta}_{ij} - \theta_i\right), \qquad \text{MSE}_i = \frac{1}{10000}\sum_{j=1}^{10000}\left(\widehat{\theta}_{ij} - \theta_i\right)^2$$

where θ_i is the i-th element of $\boldsymbol{\theta}$ and $\widehat{\theta}_{ij}$ is the estimate for θ_i at the j-th replication.

The global maximum of the log-likelihood function was found by using the L-BFGS-B algorithm. It is based on the gradient projection and uses a limited memory BFGS matrix to approximate the Hessian of the objective function [7].

Besides presenting small values, the desired behavior for both bias and MSE is to decrease inasmuch as the sample size increases. According to Tables 2 and 3, the values of

Table 1. Parameter estimates under the maximum likelihood method for the NWW model

	Actual value				Estimates			
n	k_1	λ_1	k_2	λ_2	$\widehat{k}_1$	$\widehat{\lambda}_1$	$\widehat{k}_2$	$\widehat{\lambda}_2$
15	1.3	2	1.5	1.8	2.32486	1.76022	2.53552	1.76879
	3	1.5	2.8	2.5	3.04782	1.83043	3.86394	2.30280
	2	2.2	6.5	4.1	2.36468	2.30808	8.07782	3.92074
	1.4	1.6	4.8	5.1	1.48950	1.58997	6.01459	4.96175
30	1.3	2	1.5	1.8	1.92763	1.86712	1.92300	1.72723
	3	1.5	2.8	2.5	2.74367	1.73779	3.18720	2.38999
	2	2.2	6.5	4.1	2.17105	2.29667	7.36049	4.01863
	1.4	1.6	4.8	5.1	1.46719	1.59943	5.40679	5.06024
50	1.3	2	1.5	1.8	1.73308	1.93138	1.72891	1.70956
	3	1.5	2.8	2.5	2.77118	1.66219	3.00589	2.42949
	2	2.2	6.5	4.1	2.09837	2.26229	7.01746	4.05036
	1.4	1.6	4.8	5.1	1.44390	1.59971	5.14961	5.08213
100	1.3	2	1.5	1.8	1.55192	1.97899	1.60034	1.72567
	3	1.5	2.8	2.5	2.86621	1.55322	2.90938	2.48081
	2	2.2	6.5	4.1	2.04535	2.23250	6.76158	4.07480
	1.4	1.6	4.8	5.1	1.43147	1.59866	4.98942	5.09417
200	1.3	2	1.5	1.8	1.43420	1.98859	1.55542	1.75702
	3	1.5	2.8	2.5	2.93080	1.51561	2.85893	2.49575
	2	2.2	6.5	4.1	2.01272	2.21137	6.66010	4.09431
	1.4	1.6	4.8	5.1	1.42094	1.59690	4.88342	5.09672
500	1.3	2	1.5	1.8	1.36029	1.98180	1.52928	1.78791
	3	1.5	2.8	2.5	2.97256	1.50148	2.80837	2.49915
	2	2.2	6.5	4.1	1.99838	2.20363	6.57981	4.10146
	1.4	1.6	4.8	5.1	1.40897	1.59913	4.83454	5.09864

bias and MSE for all the estimated parameters are small and the greater the sample size, the smaller the values. Thus, the results presented in this section indicate that the MLEs of the parameters of the NWW distribution are well-behaved when calculated using the L-BFGS-B algorithm.

5. Application

The hourly wind speed data measured at 10 m above ground level were collected by the National Institute of Meteorology of Brazil (INMET). These data are not publicly available, but they can be requested from the INMET [16]. The anemometers used for measuring the wind speed (in m/s) are installed in stations located in five cities spread in four states of the Brazilian Northeastern Region, as illustrated in Figure 5 (blue dots indicate the geographical position of the stations).

Table 2. Bias of the estimates under the maximum likelihood method for the NWW model

	Actual value				Bias			
n	k_1	λ_1	k_2	λ_2	$\widehat{k}_1$	$\widehat{\lambda}_1$	$\widehat{k}_2$	$\widehat{\lambda}_2$
15	1.3	2	1.5	1.8	1.02486	-0.23978	1.03552	-0.03121
	3	1.5	2.8	2.5	0.04782	0.33043	1.06394	-0.19720
	2	2.2	6.5	4.1	0.36468	0.10808	1.57782	-0.17926
	1.4	1.6	4.8	5.1	0.08950	-0.01003	1.21459	-0.13825
30	1.3	2	1.5	1.8	0.62763	-0.13288	0.42300	-0.07277
	3	1.5	2.8	2.5	-0.25633	0.23779	0.38720	-0.11001
	2	2.2	6.5	4.1	0.17105	0.09667	0.86049	-0.08137
	1.4	1.6	4.8	5.1	0.06719	-0.00057	0.60679	-0.03976
50	1.3	2	1.5	1.8	0.43308	-0.06862	0.22891	-0.09044
	3	1.5	2.8	2.5	-0.22882	0.16219	0.20589	-0.07051
	2	2.2	6.5	4.1	0.09837	0.06229	0.51746	-0.04964
	1.4	1.6	4.8	5.1	0.04390	-0.00029	0.34961	-0.01787
100	1.3	2	1.5	1.8	0.25192	-0.02101	0.10034	-0.07433
	3	1.5	2.8	2.5	-0.13379	0.05322	0.10938	-0.01919
	2	2.2	6.5	4.1	0.04535	0.03250	0.26158	-0.02520
	1.4	1.6	4.8	5.1	0.03147	-0.00134	0.18942	-0.00583
200	1.3	2	1.5	1.8	0.13420	-0.01141	0.05542	-0.04298
	3	1.5	2.8	2.5	-0.06920	0.01561	0.05893	-0.00425
	2	2.2	6.5	4.1	0.01272	0.01137	0.16010	-0.00569
	1.4	1.6	4.8	5.1	0.02094	-0.00310	0.08342	-0.00328
500	1.3	2	1.5	1.8	0.06029	-0.01820	0.02928	-0.01209
	3	1.5	2.8	2.5	-0.02744	0.00148	0.00837	-0.00085
	2	2.2	6.5	4.1	-0.00162	0.00363	0.07981	0.00146
	1.4	1.6	4.8	5.1	0.00897	-0.00087	0.03454	-0.00136

Figure 5. Northeastern Region of Brazil and geographical position of the stations

Table 3. MSE of the estimates under the maximum likelihood method for the NWW model

	Actual value				MSE			
n	k_1	λ_1	k_2	λ_2	$\widehat{k}_1$	$\widehat{\lambda}_1$	$\widehat{k}_2$	$\widehat{\lambda}_2$
15	1.3	2	1.5	1.8	4.02977	0.27152	5.66534	0.23375
	3	1.5	2.8	2.5	2.88454	0.34407	11.84968	0.2658
	2	2.2	6.5	4.1	2.06556	0.49107	23.48069	0.26423
	1.4	1.6	4.8	5.1	0.15695	0.17709	10.54719	0.23258
30	1.3	2	1.5	1.8	1.2032	0.18892	0.91974	0.16899
	3	1.5	2.8	2.5	0.55271	0.24052	2.68382	0.14055
	2	2.2	6.5	4.1	0.68471	0.26764	8.42731	0.14106
	1.4	1.6	4.8	5.1	0.10125	0.06775	2.92035	0.06796
50	1.3	2	1.5	1.8	0.56116	0.1402	0.40547	0.12994
	3	1.5	2.8	2.5	0.37297	0.17005	0.90875	0.09217
	2	2.2	6.5	4.1	0.36651	0.15461	3.74351	0.09739
	1.4	1.6	4.8	5.1	0.06836	0.02677	1.23391	0.03306
100	1.3	2	1.5	1.8	0.21934	0.09554	0.17321	0.08241
	3	1.5	2.8	2.5	0.20545	0.05313	0.29138	0.03336
	2	2.2	6.5	4.1	0.17503	0.07261	1.43608	0.05735
	1.4	1.6	4.8	5.1	0.03662	0.00872	0.49183	0.01518
200	1.3	2	1.5	1.8	0.09015	0.05808	0.09846	0.0448
	3	1.5	2.8	2.5	0.10626	0.01388	0.12841	0.01236
	2	2.2	6.5	4.1	0.06473	0.02238	0.58192	0.01879
	1.4	1.6	4.8	5.1	0.01867	0.00388	0.21766	0.0074
500	1.3	2	1.5	1.8	0.03165	0.02462	0.04952	0.01954
	3	1.5	2.8	2.5	0.04366	0.00053	0.04958	0.00315
	2	2.2	6.5	4.1	0.01686	0.0035	0.19132	0.00263
	1.4	1.6	4.8	5.1	0.00737	0.00146	0.08212	0.00295

Esperantina (denoted by Station 1) is located in the north part of the State of Piauí. Jaguaruana (denoted by Station 2) is located in the mesoregion of Jaguaribe in the State of Ceará. Cabaceiras (denoted by Station 3) and Monteiro (denoted by Station 4) are located in the mesoregion of Borborema, State of Paraíba. Arapiraca (denoted by Station 5) is located in the mesoregion of Agreste, State of Alagoas. Table 4 brings further details about the stations and the years of wind data available.

Table 4. Details of the regions where the wind speed was measured

	Station	Latitude	Longitude	Altitude (m)	Period
1	Esperantina	3°54′07″S	42°14′02″W	59	2007–2018
2	Jaguaruana	4°50′02″S	37°46′51″W	20	2007–2018
3	Cabaceiras	7°29′20″S	36°17′13″W	382	2008–2018
4	Monteiro	7°53′20″S	37°07′12″W	599	2007–2018
5	Arapiraca	9°45′07″S	36°39′39″W	264	2008–2018

As we can see in Table 5, Station 4 has the highest value of the mean among the stations in the study, whereas the highest value of variance belongs to Station 2.

Table 5. Descriptive statistics

St.	n	mean	median	min	max	variance	skewness	kurtosis
1	72637	1.56472	1.5	0.1	8.8	0.96082	0.74469	0.61585
2	71797	3.09401	3	0.1	9.4	2.94352	0.25522	−0.59481
3	83953	3.21937	3.3	0.1	9.9	2.49808	−0.08484	−0.70967
4	76738	3.28302	3.3	0.1	9.6	2.60513	0.16094	−0.50692
5	72675	2.87482	2.8	0.1	9.3	2.80844	0.17049	−0.92422

Except for Station 3, whose skewness is negative, the remaining stations have different degrees of positive skewness. On the other hand, Station 1 has the only positive value of kurtosis and Station 5 has the lowest one. Thus, the descriptive statistics indicate that the statistical characteristics of the wind speed differ from station to station.

We calculated the estimates of the parameters under the method of maximum likelihood for five distributions. Besides fitting the proposed model (3), we fitted the Normal-Normal mixture model (NN), the Weibull-Weibull mixture model (WW), the Normal distribution (N) and the Weibull distribution (W) to each one of the five datasets. Table 6 presents the MLEs along with the respective standard errors in parenthesis. The global maximum of the log-likelihood function was found using the L-BFGS-B algorithm [7] for the distribution NWW, whereas the optimization concerning the mixture models NN and WW was performed along the lines of the EM-algorithm presented in [24]. The standard errors are small in all scenarios, suggesting that the estimates in Table 6 are fairly accurate for the five distributions.

Four information criteria were used to perform comparisons among the fitted models. Generally, such criteria indicate that the best model is the one presenting the lowest value, since they are related to the amount of information lost by a given model. We used the well-known Akaike information criterion (AIC), consistent Akaike information criterion (CAIC), Bayesian information criterion (BIC) and Hannan-Quinn information criterion (HQIC). The statistics of Anderson-Darling (A^*), Cramér-von Mises (W^*) [10], and Kolmogrov-Smirnov (D) were also used to compare the fitted models. The largest p-value among the calculated D statistics is 2.2e-16. Since these statistics are measures of the difference between the empirical distribution function and the real underlying cdf, it is reasonable to say that the smaller their values, the better the fit.

Table 7 brings the aforementioned goodness-of-fit measures (except for the D statistic) for the five cited models fitted to each station. According to Table 7, the distributions NWW and WW present the better fits among the competing models for all stations. The four information criteria indicate that NWW presents a higher performance over WW concerning stations 1, 2, 4 and 5. Regarding station 3, the difference between the values of AIC of NWW and WW is not considerable, although the AIC for NWW is slightly smaller in such case; the same behavior states for CAIC, BIC and HQIC. Both goodness-of-fit statistics A^* and W^* (see last ten rows of Table 7) agree with the information criteria in pointing NWW and WW as the two better fits. However, they suggest that NWW outperforms WW in fitting the datasets for all the five stations. To get more insight into these results, plots of the fitted densities overlapping the histograms of the wind speed data for stations 1 to 5 and the corresponding cdfs are presented in Figure 6.

According to Table 8, despite the performance of the W distribution for the D statistic, the NWW model stands out as competitive for stations 2, 3 and 5.

Table 6. Estimates and standard errors in parenthesis

Distr.	Par.	St1	St2	St3	St4	St5
NWW	k_1	1.09455	0.97219	0.93287	1.16262	0.92095
		(0.0066)	(0.0047)	(0.0035)	(0.0051)	(0.0041)
	λ_1	2.44439	4.14273	5.11736	4.61159	3.26931
		(0.0116)	(0.0300)	(0.0151)	(0.0358)	(0.0157)
	k_2	1.24356	2.21499	2.74024	2.17717	2.87659
		(0.0090)	(0.0181)	(0.0141)	(0.0232)	(0.0135)
	λ_2	2.23302	4.82613	4.47508	4.74835	4.76046
		(0.0104)	(0.0129)	(0.0083)	(0.0187)	(0.0059)
NN	μ_1	0.99225	1.25346	1.05246	1.52464	1.09819
		(0.0088)	(0.0112)	(0.0135)	(0.0152)	(0.0091)
	σ_1	0.57673	0.73618	0.64260	0.80711	0.66916
		(0.0051)	(0.0067)	(0.0077)	(0.0082)	(0.0056)
	μ_2	2.15056	3.69950	3.71423	3.78985	3.69032
		(0.0121)	(0.0110)	(0.0094)	(0.0127)	(0.0118)
	σ_2	0.96181	1.50081	1.28725	1.42131	1.33344
		(0.0038)	(0.0051)	(0.0054)	(0.0056)	(0.0063)
	w	0.50577	0.24754	0.18591	0.22375	0.31460
		(0.0083)	(0.0041)	(0.0036)	(0.0054)	(0.0041)
WW	k_1	2.46232	2.61226	3.48661	2.71890	3.63802
		(0.0451)	(0.0162)	(0.0183)	(0.0146)	(0.0284)
	λ_1	2.10941	4.09404	4.26754	4.08741	4.41073
		(0.0089)	(0.0126)	(0.0070)	(0.0103)	(0.0110)
	k_2	1.27757	1.20035	1.30110	1.26447	1.36100
		(0.0106)	(0.0088)	(0.0069)	(0.0118)	(0.0056)
	λ_2	1.42208	1.54955	1.81537	1.65794	1.86015
		(0.0186)	(0.0392)	(0.0242)	(0.0526)	(0.0183)
	w	0.44618	0.75154	0.71389	0.83225	0.51531
		(0.0183)	(0.0075)	(0.0046)	(0.0069)	(0.0056)
N	μ	1.56471	3.09401	3.21937	3.28301	2.87482
		(0.0036)	(0.0064)	(0.0054)	(0.0058)	(0.0062)
	σ	0.98021	1.71565	1.58052	1.61403	1.67582
		(0.0025)	(0.0045)	(0.0038)	(0.0041)	(0.0043)
W	k	1.58959	1.77193	2.03559	2.07474	1.65091
		(0.0047)	(0.0054)	(0.0059)	(0.0061)	(0.0051)
	λ	1.73825	3.44821	3.59712	3.68213	3.18827
		(0.0042)	(0.0075)	(0.0063)	(0.0066)	(0.0074)

It is worth pointing out that mixture models are commonly used to fit non-unimodal datasets, such as those represented by the histograms of stations 2, 3 and 5. Nonetheless, the results attest that the NWW accommodates such data better than the two mixture models presented in this study. Furthermore, NWW has one parameter less than NN or WW do.

Finally, since the NWW distribution outperforms the competing models in fitting the wind speed data of the Northeastern Region of Brazil, according to different information criteria and formal goodness-of-fit statistics, we have good reasons to recommend its use to model similar data in future works. We also encourage practitioners of statistics to

Table 7. Goodness-of-fit measures

Crit.	Distr.	St1	St2	St3	St4	St5
AIC	NWW	189324.5	272976.1	306366.6	286607.1	268213.1
	NN	197221.2	277168.5	310138.0	289011.9	273287.3
	WW	189379.2	273075.2	306370.5	286647.3	268244.1
	N	203235.1	281266.5	315112.0	291251.2	281292.0
	W	190666.7	278737.7	319347.3	291232.9	277572.1
CAIC	NWW	189365.2	273016.8	306408.0	286648.1	268253.8
	NN	197272.2	277219.4	310189.7	289063.1	273338.3
	WW	189430.2	273126.1	306422.2	286698.5	268295.1
	N	203255.4	281286.9	315132.7	291271.7	281312.4
	W	190687.1	278758.0	319368.0	291253.4	277592.5
BIC	NWW	189361.2	273012.8	306404.0	286644.1	268249.8
	NN	197267.2	277214.4	310184.7	289058.1	273333.3
	WW	189425.2	273121.1	306417.2	286693.5	268290.1
	N	203253.4	281284.9	315130.7	291269.7	281310.4
	W	190685.1	278756.0	319366.0	291251.4	277590.5
HQIC	NWW	189335.8	272987.4	306378.0	286618.5	268224.4
	NN	197235.3	277182.7	310152.3	289026.1	273301.4
	WW	189393.4	273089.3	306384.8	286661.5	268258.3
	N	203240.7	281272.2	315117.7	291256.9	281297.7
	W	190672.4	278743.3	319353.0	291238.6	277577.7
A*	NWW	71.78	22.43	22.60	19.31	30.98
	NN	181.13	51.72	54.33	27.79	78.72
	WW	81.27	31.05	32.44	24.62	50.25
	N	531.70	230.18	303.82	118.73	494.43
	W	257.57	471.01	1117.26	337.07	784.70
W*	NWW	7.73	2.33	2.94	2.58	3.36
	NN	19.49	5.17	6.60	3.21	7.61
	WW	9.35	3.59	4.22	3.89	4.82
	N	72.08	31.04	45.14	16.72	72.83
	W	33.87	62.25	164.54	46.18	108.58

Table 8. Kolmogorov-Smirnov statistics

Crit.	Distr.	St1	St2	St3	St4	St5
D	NWW	0.7028925	0.8857891	0.9014568	0.9234669	0.8484073
	NN	0.7027823	0.8857891	0.9014687	0.9234669	0.8484348
	WW	0.7029200	0.8857891	0.9014925	0.9234669	0.8484348
	N	0.7029200	0.8857891	0.9014925	0.9234669	0.8484348
	W	0.6988862	0.8787972	0.8960609	0.9182413	0.8389405

investigate the modelling benefits of the NWW (and other submodels from the Normal-(G_1,G_2) class) with respect to data describing different phenomena usually modelled by mixtures.

6. Conclusions

An alternative distribution for modelling wind speed data is proposed and some mathematical properties of the class that generates it are discussed, like the series representation of the pdf, the moments and the moment generating function. The general cdf of the

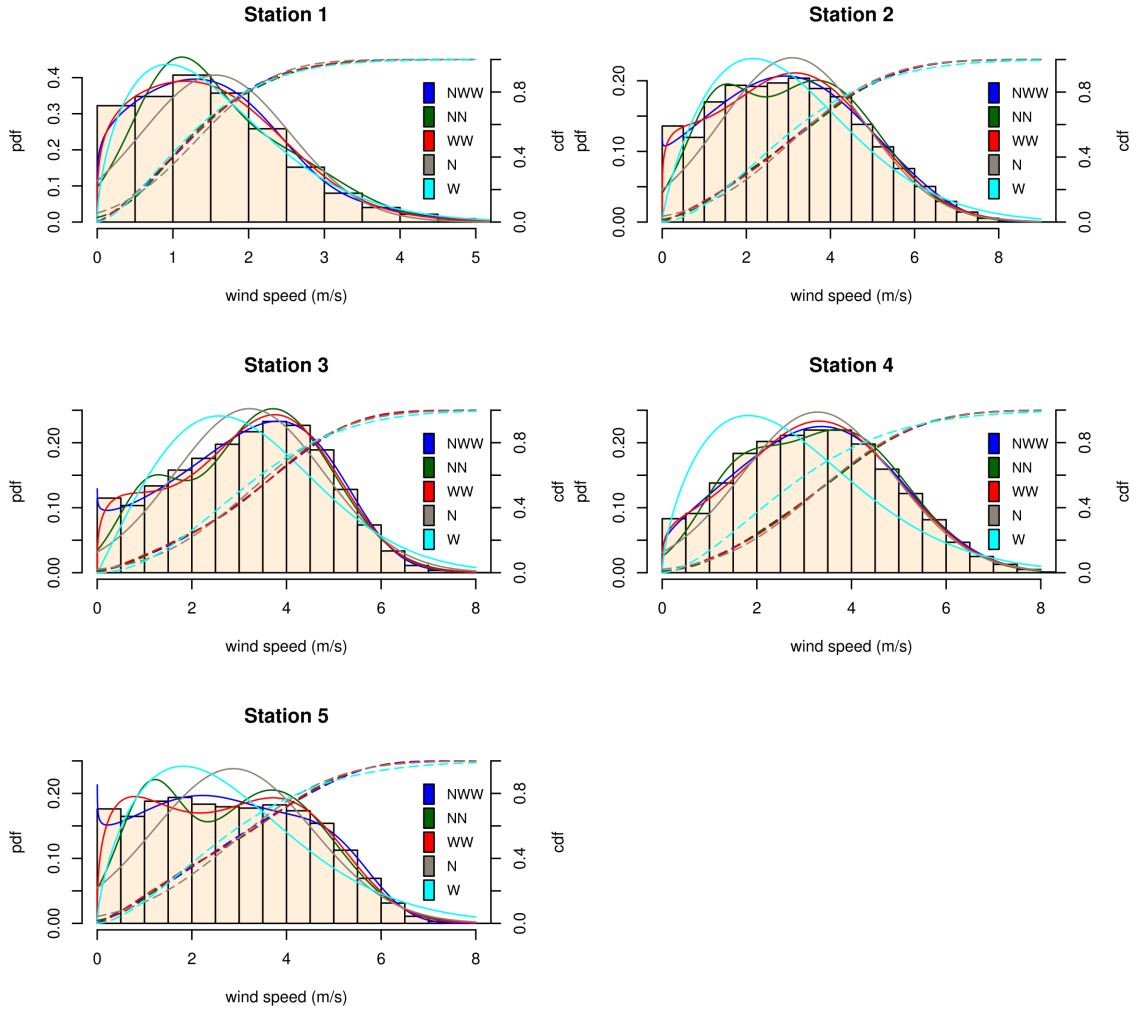

Figure 6. Histograms and fitted densities

Normal-(G_1, G_2) class is written as a composition of two baselines and its submodels are identifiable as long as both baseline cdfs are. Such is the case of the NWW distribution.

The novel model has four parameters and high flexibility. It is able to fit right-skewed and left-skewed data and its pdf presents unimodal and bimodal shapes.

The Monte Carlo simulation studies indicate that the MLEs of the NWW parameters behave appropriately when the optimization is performed via the L-BFGS-B algorithm.

The modelling gains of the NWW distribution are upheld by the satisfactory results concerning the application to the wind speed data collected in the Northeastern Region of Brazil. The considered information criteria and the formal goodness-of-fit statistics of Anderson-Darling and Cramér-von Mises suggest that the proposed model outperforms other competing distributions that are commonly employed in wind speed modelling, especially the highly competitive mixture model of two Weibull components.

We hope that this work may encourage the investigation of the modelling benefits of the NWW (and other identifiable submodels from the Normal-(G_1, G_2) class) with respect to data describing other natural phenomena usually modelled by mixtures.

References

[1] S. Akdag, H. Bagiorgas and G. Mihalakakou, *Use of two-component Weibull mixtures in the analysis of wind speed in Eastern Mediterranean*, Appl. Energy **87**, 2566-2573, 2010.

[2] A. Alzaatreh, C. Lee and F. Famoye, *A new method for generating families of continuous distributions*, Metron **71**, 63-79, 2013.

[3] J.C.H. Araújo, W.F. Souza, A.J.A. Meireles and C. Brannstrom, *Sustainability challenges of wind power deployment in coastal Ceará state, Brazil*, Sustainability **12** (14), 5562, 2020.

[4] A. Azzalini, *A class of distributions which includes the normal ones*, Scand. J. Stat. **12** (2), 171-178, 1985.

[5] T.G. Bali and P. Theodossiou, *Risk measurement performance of alternative distribution functions*, J. Risk Insur. **75** (2), 411-437, 2008.

[6] C.R. Brito, L.C. Rêgo, W.R. Oliveira and F. Gomes-Silva, *Method for generating distributions and classes of probability distributions: the univariate case*, Hacet. J. Math. Stat. **48** (3), 897-930, 2019.

[7] R.H. Byrd, P. Lu, J. Nocedal and C. Zhu, *A limited memory algorithm for bound constrained optimization*, SIAM J. Sci. Comput. **16** (5), 1190-1208, 1995.

[8] J. Carta and P. Ramírez, *Analysis of two-component mixture Weibull statistics for estimation of wind speed distributions*, Renew. Energy **32**, 518-531, 2007.

[9] T. Chang, *Estimation of wind energy potential using different probability density functions*, Appl. Energy **88**, 1848-1856, 2011.

[10] G. Chen and N. Balakrishnan, *A general purpose approximate goodness-of-fit test*, J. Qual. Technol. **27** (2), 154-161, 1995.

[11] A. Eltamaly, *Design and implementation of wind energy system in Saudi Arabia*, Renew. Energy **60**, 42-52, 2013.

[12] Q. Han, S. Ma, T. Wang and F. Chu, *Kernel density estimation model for wind speed probability distribution with applicability to wind energy assessment in China*, Renew. Sustain. Energy Rev. **115**, 109387, 2019.

[13] B.E. Hansen, *Autoregressive conditional density estimation*, Int. Econ. Rev. **35** (3), 705-730, 1994.

[14] B. Hu, Y. Li, H. Yang and H. Wang, *Wind speed model based on kernel density estimation and its application in reliability assessment of generating systems*, J. Mod. Power Syst. Clean Energy **5** (2), 220-227, 2017.

[15] U. Ilhan and Y.M. Kantar, *Analysis of some flexible families of distribution for estimation of wind speed distributions*, Appl. Energy **89**, 355-367, 2012.

[16] INMET, National Institute of Meteorology of Brazil. Official website, URL https://portal.inmet.gov.br/ accessed in 11/01/2021.

[17] O. Jaramillo and M. Borja, *Wind speed analysis in La Ventosa, Mexico: A bimodal probability distribution case*, Renew. Energy **29**, 1613-1630, 2004.

[18] R. Kollu, S. Rayapudi, S. Narasimham and K. Pakkurthi, *Mixture probability distribution functions to model wind speed distributions*, Int. J. Energy Environ. Eng. **3** (27), 2012.

[19] G. McLachlan and D. Peel, *Finite Mixture Models*, Wiley Interscience, 2000.

[20] K. Mohammadi, O. Alavi and J. McGowan, *Use of Birnbaum-Saunders distribution for estimating wind speed and wind power probability distributions: A review*, Energy Convers. Manag. **143**, 109-122, 2017.

[21] E.C. Morgan, M. Lackner, R.M. Vogel and L.G. Baise, *Probability distributions for offshore wind speeds*, Energy Convers. Manag. **52** (1), 15-26, 2011.

[22] G.S. Mudholkar and D.K. Srivastava, *Exponentiated weibull family for analyzing bathtub failure-rate data*, IEEE Trans. Reliab. **42** (2), 299-302, 1993.

[23] J. Von Neumann, *Various techniques used in connection with random digits*, Applied Mathematics Series **12**, National Bureau of Standards, Washington, DC, USA, 1951.
[24] H.D. Nguyen, D. Wang, G.J. McLachlan, *Randomized mixture models for probability density approximation and estimation*, Inf. Sci. **467**, 135-148, 2018.
[25] S. Perkin, D. Garrett and P. Jensson, *Optimal wind turbine selection methodology: A case-study for Búrfell, Iceland*, Renew. Energy **75**, 165-172, 2015.
[26] S. Pishgar-Komleh, A. Keyhani and P. Sefeedpari, *Wind speed and power density analysis based on Weibull and Rayleigh distributions (a case study: Firouzkooh county of Iran)*, Renew. Sustain. Energy Rev. **42**, 313-322, 2015.
[27] Z. Qin, W. Li and X. Xiong, *Estimating wind speed probability distribution using kernel density method*, Electr. Power Syst. Res. **81** (12), 2139-2146, 2011.
[28] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2018.
[29] B. Safari, *Modeling wind speed and wind power distributions in Rwanda*, Renew. Sustain. Energy Rev. **15**, 925-935, 2011.
[30] D. Weisser, *A wind energy analysis of Grenada: an estimation using the Weibull density function*, Renew. Energy **28** (11), 1803-1812, 2003.

Appendix A.

The information matrix cited in section 2.3 is given by $J(\boldsymbol{\theta}|\boldsymbol{X}) = -\nabla_{\boldsymbol{\theta}}\nabla_{\boldsymbol{\theta}}{}^{\top}\ell(\boldsymbol{\theta}|\boldsymbol{X}) = -(u_{jk})_{1\leq j\leq r+m,1\leq k\leq r+m}$ where:

$$
\begin{aligned}
u_{jk} =& \sum_{i=1}^{n} \frac{1}{f_{G_1,G_2}(x_i)} \phi\left(\frac{G_1(x_i)}{1-G_1(x_i)}\right) \frac{1}{(1-G_1(x_i))^2} \left\{ \left(\frac{2}{1-G_1(x_i)} - \frac{G_1(x_i)}{[1-G_1(x_i)]^3} \right) \right. \\
& \times \left(\frac{\partial}{\partial\theta_k} G_1(x_i) \frac{\partial}{\partial\theta_j} g_1(x_i) + \frac{\partial}{\partial\theta_j} G_1(x_i) \frac{\partial}{\partial\theta_k} g_1(x_i) \right) + \frac{\partial^2}{\partial\theta_j \partial\theta_k} g_1(x_i) - \frac{g_1(x_i)}{(1-G_1(x_i))^2} \\
& \times \left[\frac{\partial}{\partial\theta_j} G_1(x_i) \frac{\partial}{\partial\theta_k} G_1(x_i) \left(\frac{3{G_1}^2(x_i)}{(1-G_1(x_i))^2} + \frac{4G_1(x_i)}{1-G_1(x_i)} + 1 \right) + \frac{G_1(x_i)}{1-G_1(x_i)} \right. \\
& \left. \times \frac{\partial^2}{\partial\theta_j \partial\theta_k} G_1(x_i) \right] + \frac{g_1(x_i) G_1(x_i)}{(1-G_1(x_i))^5} \left(\frac{\partial}{\partial\theta_j} G_1(x_i) + \frac{\partial}{\partial\theta_k} G_1(x_i) \right) \\
& + \frac{2g_1(x_i)}{(1-G_1(x_i))^2} \left(3 - \frac{2G_1(x_i)}{(1-G_1(x_i))^2} \right) \frac{\partial}{\partial\theta_k} G_1(x_i) \frac{\partial}{\partial\theta_j} G_1(x_i) + \frac{\partial^2}{\partial\theta_j \partial\theta_k} G_1(x_i) \\
& \left. \times \frac{2g_1(x_i)}{1-G_1(x_i)} \right\} - \sum_{i=1}^{n} \frac{1}{f^2_{G_1,G_2}(x_i)} \phi^2 \left(\frac{G_1(x_i)}{1-G_1(x_i)} \right) \frac{1}{(1-G_1(x_i))^4} \\
& \times \left[\frac{\partial}{\partial\theta_j} g_1(x_i) + \left(\frac{2g_1(x_i)}{1-G_1(x_i)} - \frac{g_1(x_i) G_1(x_i)}{(1-G_1(x_i))^3} \right) \frac{\partial}{\partial\theta_j} G_1(x_i) \right] \\
& \times \left[\frac{\partial}{\partial\theta_k} g_1(x_i) + \left(\frac{2g_1(x_i)}{1-G_1(x_i)} - \frac{g_1(x_i) G_1(x_i)}{(1-G_1(x_i))^3} \right) \frac{\partial}{\partial\theta_k} G_1(x_i) \right], \text{for } 1 \leq j \leq r, 1 \leq k \leq r;
\end{aligned}
$$

$$
\begin{aligned}
u_{jk} =& \sum_{i=1}^{n} \frac{-1}{f^2_{G_1,G_2}(x_i)} \phi\left(\frac{G_1(x_i)}{1-G_1(x_i)}\right) \frac{1}{(1-G_1(x_i))^2} \frac{\phi(\log[1-G_2(x_i)])}{1-G_2(x_i)} \left(\frac{\partial}{\partial\theta_k} g_1(x_i) \right. \\
& + \left[\frac{2g_1(x_i)}{1-G_1(x_i)} - \frac{g_1(x_i) G_1(x_i)}{(1-G_1(x_i))^3} \right] \frac{\partial}{\partial\theta_k} G_1(x_i) \left(\frac{\partial}{\partial\theta_j} g_2(x_i) + (1 + \log[1-G_2(x_i)]) \right. \\
& \left. \left. \times \frac{g_2(x_i)}{1-G_2(x_i)} \frac{\partial}{\partial\theta_j} G_2(x_i) \right), \text{for } r < j \leq r+m, 1 \leq k \leq r;
\end{aligned}
$$

$$u_{jk} = \sum_{i=1}^{n} \frac{-1}{f^2_{G_1,G_2}(x_i)} \phi\left(\frac{G_1(x_i)}{1-G_1(x_i)}\right) \frac{1}{(1-G_1(x_i))^2} \frac{\phi(\log[1-G_2(x_i)])}{1-G_2(x_i)} \left(\frac{\partial}{\partial\theta_j} g_1(x_i) \right.$$
$$+ \left[\frac{2g_1(x_i)}{1-G_1(x_i)} - \frac{g_1(x_i)G_1(x_i)}{(1-G_1(x_i))^3}\right] \frac{\partial}{\partial\theta_j} G_1(x_i)\Big) \left(\frac{\partial}{\partial\theta_k} g_2(x_i) + (1+\log[1-G_2(x_i)])\right.$$
$$\left. \times \frac{g_2(x_i)}{1-G_2(x_i)} \frac{\partial}{\partial\theta_k} G_2(x_i)\right), \text{for } 1 \le j \le r, r < k \le r+m;$$

$$u_{jk} = \sum_{i=1}^{n} \frac{1}{f_{G_1,G_2}(x_i)} \frac{\phi(\log[1-G_2(x_i)])}{1-G_2(x_i)} \left[\frac{1+\log[1-G_2(x_i)]}{1-G_2(x_i)} \left(\frac{\partial}{\partial\theta_j} G_2(x_i) \frac{\partial}{\partial\theta_k} g_2(x_i)\right.\right.$$
$$\left. + \frac{\partial}{\partial\theta_k} G_2(x_i) \frac{\partial}{\partial\theta_j} g_2(x_i)\right) + \frac{\partial^2}{\partial\theta_j\partial\theta_k} g_2(x_i) + \frac{\partial}{\partial\theta_j} G_2(x_i) \frac{\partial}{\partial\theta_k} G_2(x_i) \frac{g_2(x_i)}{(1-G_2(x_2))^2}$$
$$\times \left(1 + \log^2[1-G_2(x_i)] + 3\log[1-G_2(x_i)]\right) + (1+\log[1-G_2(x_i)]) \frac{\partial^2}{\partial\theta_j\partial\theta_k} G_2(x_i)$$
$$\left. \times \frac{g_2(x_i)}{1-G_2(x_i)}\right] - \sum_{i=1}^{n} \frac{1}{f^2_{G_1,G_2}(x_i)} \frac{\phi^2(\log[1-G_2(x_i)])}{(1-G_2(x_i))^2} \left[\frac{(1+\log[1-G_2(x_i)])g_2(x_i)}{1-G_2(x_i)}\right.$$
$$\left. \times \frac{\partial}{\partial\theta_j} G_2(x_i) + \frac{\partial}{\partial\theta_j} g_2(x_i)\right] \left[\frac{(1+\log[1-G_2(x_i)])g_2(x_i)}{1-G_2(x_i)} \frac{\partial}{\partial\theta_k} G_2(x_i) + \frac{\partial}{\partial\theta_k} g_2(x_i)\right],$$
$$\text{for } r < j \le r+m, r < k \le r+m.$$